Processing of Orthographic Structure by Adults of Different Reading Ability*

Marcus Taft
University of New South Wales

Abstract
The research presented here examines the proposal that orthographic processing in reading polysyllabic words takes place via an analysis of the word into an orthographic/morphological structure called the Basic Orthographic Syllabic Structure or BOSS. This structure includes the largest possible coda in the first component (e.g., the THUND of THUNDER) and, as such, it cuts across the phonological syllable boundary (e.g., THUN + DER). The existence of the BOSS has been previously supported by showing that words physically divided at their BOSS (e.g., THUND ER) are faster to recognize than those divided at their syllable (e.g., THUN DER). However, there has been little, if any, report of confirmatory evidence for this conclusion. Three experiments are reported here demonstrating that whether the BOSS division is faster than the syllable division depends crucially on the reading ability of the participants. Better adult reading, independently measured in terms of comprehension, is associated with a preference for the BOSS division while poorer adult reading is more associated with a syllable preference. Such a result potentially explains the conflicting findings that have been previously reported. A number of different explanations are offered for the pattern of results, with the suggestion that poorer adult reading is more oriented toward the pronunciation of the word than is better reading.

1 Introduction

In order to recognize that the letter-string THUNDER means ‘the noise that accompanies lightning’, a reader needs to gain access to the representation for THUNDER in lexical memory, which stores the word-form and provides the gateway to semantic information. A match has to be found between this stored lexical representation and the incoming sensory representation of THUNDER, and the issue that is addressed here is how such a match is made. In particular, is the letter-string treated merely as a linear concatenation of letters or is its internal structure utilized in some way. For example, is the fact that THUNDER is composed of two syllables relevant to the matching process?

* Acknowledgments: The research reported in this paper was funded by a grant awarded to the author by the Australian Research Council. The assistance of Nenagh Kemp, Jennifer Peña, Louisa Carroll, Jeesun Kim, Angelo Alonzo, and Eunice Wu is gratefully acknowledged.

Address for Correspondence: Marcus Taft, School of Psychology, University of NSW, Sydney NSW 2052, Australia; e-mail: <m.taft@unsw.edu.au>
Most of the previous research examining internal orthographic structure has either focused on morphemic units (e.g., the UN, HEAT and ED of UNHEATED, the BLACK and BOARD of BLACKBOARD; see e.g., Feldman, 1995; Sandra & Taft, 1994) or the subsyllabic units of onset and body or rime (e.g., the SHR and IEK of SHRIEK, the F and UND of FUND; see e.g., Bowey, 1996; Kay & Bishop, 1987; Treiman & Chafetz, 1987; Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995). Thus, most work has been concerned with the processing of polymorphemic or monosyllabic words, while the processing of monomorphemic polysyllabic words (like THUNDER) has been relatively neglected, despite the fact that such words are very common in the language.

Research by Taft and Forster (1976), however, suggested that the first syllable of such words plays an important role in their recognition. They showed a delay in lexical decision responses ("is this letter-string a word or not?") to nonwords that are actually first syllables of real words (e.g., the ATH of ATHLETE, the CHIM of CHIMNEY), and concluded that this arose from the fact that such syllables are represented in the lexical processing system. The result was replicated by Taft (1986).

1.1 Defining the syllable

If it is correct that the first syllable is an important unit in word recognition, it becomes necessary to define where the syllable boundary falls for any given word. However, if syllables are to be defined in terms of phonological considerations, it is not always clear-cut in English where this boundary is. For example, is the syllable structure of PANIC to be taken as PA + NIC, PAN + IC or even PAN + N + IC? The first analysis would be preferred on the basis of the “maximal onset principle” that is proposed on linguistic grounds (see, e.g., Spencer, 1996) because the N becomes an onset of the second syllable; though the idea of ambisyllabicity (e.g., Kahn, 1976) would favor the third analysis because it maintains N as an onset while avoiding having a syllable that ends in a short vowel (e.g., /æ/). Examples like CEMENT with an unstressed first vowel reduced to a schwa, SPIDER with a long first vowel, or THUNDER with a medial consonant cluster are clearer in their use of the maximal onset principle (i.e., CE + MENT, SPI + DER, and THUN + DER respectively). While linguists may or may not agree about the syllable structure of spoken English, there is a suggestion that native English speakers themselves are insensitive to the location of syllable boundaries in spoken words (e.g., Cutler, Mehler, Norris, & Segui, 1986). So the idea of a visual recognition unit being based on the spoken syllable is awkward, though such a proposal certainly exists in the literature (e.g., Adams, 1990; Spoehr & Smith, 1973).

Furthermore, if the recognition unit that represents the first syllable is phonologically defined, it would be hard to integrate such a unit with that representing a morpheme. For example, the phonological syllable structure of LEAKING and LEAKY has the boundary between the EA and the K, and that of LEAKS and LEAKED has no syllable boundary at all because they are monosyllabic. On the other hand, the morphemic structure for all of these words has a boundary immediately after the K. Thus, the phonological structure is entirely incompatible with the morphemic structure and it is therefore hard to see how both could be relevant at the same time in the recognition of a word. It would require the
assumption that words are represented in more than one way; in terms of syllable structure and in terms of morphemic structure, and that these structures are independently exploited.

1.2 Defining the BOSS

A suggestion has been made, however, that eliminates the above problems. Taft (1979) put forward the idea of an orthographic analysis of polysyllabic words that ignores the phonological syllable structure and that additionally incorporates morphological considerations. This is the Basic Orthographic Syllabic Structure or BOSS. The idea of the BOSS is that it maximizes the amount of information contained in the recognition unit by drawing a structural boundary after all of the consonants that follow the first vowel of the stem morpheme. For example, the BOSS of CEMENT is taken to be CEM, that of SPIDER is taken to be SPID and that of THUNDER is taken to be THUND. Other examples include the PAN of PANIC (giving a PAN + IC analysis), the CUST of UNACCUSTOMED (because CUSTOM is the stem morpheme) and the LEAK of LEAKY and LEAKS (because LEAK is the stem morpheme).

We see, then, that the BOSS totally violates the maximal onset principle that has been used to characterize spoken syllable structure. Instead of maximizing the onset of the second syllable, the orthographic syllabic structure actually maximizes the coda of the first syllable. A “coda” is the consonantal part of a syllable that follows the vowel (e.g., orthographically speaking, the K of SHRIEK and the ND of FUND).

With the BOSS following a “maximal coda principle” (cf. Taft & Radeau, 1995), the BOSS of CENTRAL, TABLE, ATHLETE and KIDNEY would be CENT, TAB, ATH, and KID respectively because /tr/, /bl/, /th/ and /dn/ cannot be codas. Taft (1979) originally proposed such an analysis for words of this type, but on the grounds that no monosyllabic words end in NTR, BL, THL, and DN. By redefining the BOSS in terms of the orthographic equivalent of the maximal coda, however, one avoids such problems as having to say, for example, that the BOSS of NAVY and NAVAL is NAV even though no (real) words actually end in V (see Taft, 1987). Because the V of NAV is an acceptable coda, NAV can be a BOSS by this definition. Although the maximal coda often cuts across the spoken syllable structure of a word, defining the BOSS in terms of a maximization of the coda implies that the basis for the BOSS is actually phonological in nature. Thus, while it can be seen as been an orthographic unit of processing, its characteristics are influenced by phonological considerations.

Whether a consonant should be processed as the coda of the first syllable or the onset of the second syllable may well become an important consideration for computational models of word recognition once they begin to incorporate polysyllabic words. Virtually all published computational models have been set up on the basis of a corpus of monosyllabic words (e.g., Coltheart, Curtis, Atkins, & Haller, 1993; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989) and connectionist models like that of Plaut et al. (1996) have an input system coded explicitly in terms of onsets, vowels, and codas. Once the recognition of polysyllabic words is required by such a model, as it must be for it to be generalizable to the reading of all words, a decision needs to be made about whether a consonant is to be coded as an onset or a coda (e.g., the D of THUNDER and SPIDER). It is the latter analysis that is being proposed here.
1.3 Evidence for and against the BOSS

What then is the evidence that has been presented in favor of the BOSS analysis? The main paradigm for eliciting such evidence has been one where the word is physically divided in terms of its BOSS or its phonologically defined syllable (PS) and recognition times compared. For example, Taft (1979) presented words split on the screen by means of a physical gap, and found that those divided at their BOSS (e.g., THUND ER) were associated with faster lexical decision responses than those divided according to their PS (e.g., THUN DER). Using several other splitting techniques, like temporal priming with an initial fragment (e.g., presenting THUN or THUND before THUNDER), Taft (1987) replicated this finding and, furthermore, showed that it did not arise from the fact that the BOSS was more informative than the PS simply because it had more letters. In particular, the items chosen were equally predictive of the final word on the basis of their PS and BOSS (e.g., THUNDER is already uniquely defined by THUN), and also, there was no further enhancement of reaction times when the initial fragment was made one letter longer than the BOSS (e.g., THUNDE).

There have been a few other studies that have obtained results that are at least consistent with the BOSS idea (e.g., Inhoff, 1987; Luszcz, Bungey, & Geffen, 1984; Prinzmetal, Treiman, & Rho, 1986), but there has been little success beyond Taft's studies in providing direct support for it. In fact, there have been several studies that have totally failed to find any evidence for the BOSS (e.g., Jordon, 1986; Katz & Baldasare, 1983; Knuijt & Assink, 1997; Lima & Pollatsk, 1983; Seidenberg, 1987). Some of these studies (e.g., Katz & Baldasare, 1983; Lima & Pollatsk, 1983) even employed a paradigm similar to that used by Taft (1979, 1987), namely, splitting the words physically on the screen, but still found no sign at all of a preference for a BOSS analysis over a PS one.

As a result of the failure of such research in laboratories other than Taft's, the idea of the BOSS has fallen into disrepute. However, there must be some reason for the discrepancy between the results from different laboratories. It is possible that Taft's findings arise from a Type I error, but this is unlikely given that he has presented at least seven experiments that show a preference for the BOSS over the PS (Taft, 1979, 1987, 1992). It could perhaps be argued that this preference arises from the way reading is (or was) taught in Australia compared to other English speaking countries. However, while it is true that most of Taft's research was carried out in Australia, at least one of the experiments, reported in 1979, was conducted in the U.S.A. (at the Massachusetts Institute of Technology). Is there any other factor that could have led to the discrepancy between findings?

There is evidence that adult readers vary in their use of phonological information when reading silently (e.g., Jared, Levy, & Rayner, 1999; Lewellen, Goldinger, Pisoni, & Greene, 1993). In particular, readers who perform relatively poorly in reading comprehension tests tend to be more phonologically oriented than are the better performers, as measured by homophone confusions. What these studies also show is that college students can show a wide range of abilities in their reading comprehension and that, therefore, an unselected group of students is going to include those who are phonologically oriented and those who are not. If a reader is phonologically oriented, it seems reasonable to suggest that a word that is split according to its BOSS structure will be disruptive to processing because it cuts across the phonological form of the word. For better readers, on the other hand, it can be argued
that the BOSS division should actually be helpful because the initial unit (e.g., THUND in THUND ER) provides more information about the word than does the initial unit of the PS division (e.g., THUN in THUN DER).

What this means is that an unselected group of university undergraduates might not show a preference for the BOSS division over the PS division simply because a BOSS preference for better readers is counteracted by a syllable preference for poorer readers. In order, then, to explain why Taft consistently found BOSS preferences, it must be assumed that he was biased in his selection of participants, and in particular, was using a preponderance of better readers. Such an idea is in fact quite possible because Taft included many 4th year honors students, graduate students, and even colleagues amongst his participants. So, it is possible that the reading ability of the participants used in Taft’s studies was biased toward the upper end whereas the participants used in the research of other laboratories drew upon the standard distribution of abilities found in the undergraduate population. The argument could therefore be made that evidence for the BOSS will only be observed when the participants are relatively good readers, and if so, not only would the discrepancy between research findings have a potential explanation, but the importance of the BOSS in efficient reading performance would be indicated.

In order to examine this possible relationship between reading ability and preference for the BOSS, a group of undergraduate students were tested on their recognition of words divided according to the BOSS or the PS, and also given an independent test of reading comprehension. A correlational analysis was then carried out between preference for the BOSS division (in terms of reaction time) and reading ability (in terms of reading comprehension scores). It was expected that, given the unbiased participant selection procedure, there would be no overall advantage for BOSS divided words over PS divided words, but that a positive correlation would be observed between BOSS preference and performance on the reading test.

2 Experiment 1

2.1 Method

2.1.1 Materials and procedure

In the first experiment, 30 polysyllabic words of a wide range of frequencies were selected (mean of 12.5 per million according to Carroll, Davies, & Richman, 1971) whose phonologically defined syllable could be differentiated from their BOSS. Thus the words either had a long first vowel (e.g., ROUTINE, MOVIE, ZERO) or a medial consonant cluster (e.g., DICTATE, TURKEY, TEXTILE). None of the words began with anything that could be construed as a prefix. Unlike the experiments of Taft (1987), the BOSS and the first phonological syllable were not equally predictive of the whole word. The reason for this was, first, that the generalizability of the BOSS preference was considered important and the 1987 study had already exhausted the small number of words whose PS and BOSS are equally predictive; and, second, that this experiment (as well as the one that follows) was actually designed with further experimental manipulations in mind (viz., interlanguage comparisons)
that would not allow such constraints to be placed on the items. The experimental items are presented in Appendix 1.

One group of participants saw half of the experimental items with a gap of three spaces at the phonological syllable boundary (the PS condition, e.g., ROU TINE, DIC TATE) and the other half with the gap at the BOSS boundary (the BOSS condition, e.g., MOV IE, TURK EY); while a second group saw the items in the opposite condition (e.g., ROUT INE, DICT ATE; MO VIE, TUR KEY).

The 30 experimental items were presented in random order along with 30 nonword distractors that had a similar structure to real words, but which were mostly more than one letter different from any real word (e.g., LEABARB, LOFAY, FELDOCK, NURBEE). These were divided either in terms of a phonological analysis (e.g., LEA BARB) or in terms of their BOSS (e.g., FELD OCK). Both groups received the same set of nonwords. Ten practice items preceded these words and nonwords.

Participants were told that they would see a letter-string split into two parts on the screen and were instructed to press a button as quickly but as accurately as possible in response to whether or not that letter-string formed a real word when the two parts were combined. Each item was presented under computer control in upper-case letters on a television monitor for 500 ms with an intertrial interval of 1 s after the response.

After completing the experiment, the participants were given a paper-and-pencil reading comprehension test. This test comprised a series of short passages each followed by three to seven multiple choice questions. The test was based upon the Cooperative Reading Comprehension Test developed by the Australian Council of Educational Research and resembled the comprehension component of the Nelson-Denny Reading Test (1973). In all, there were 57 multiple choice questions coming from a total of 12 passages. Each question had five alternative answers to choose from. A time limit of 15 mins was tightly adhered to and participants were told that they were not expected to complete the whole test, but that they should work as rapidly as possible without making careless mistakes.

It was considered that a 15 mins time limit was long enough to allow discrimination of reading ability, but short enough to minimize fatigue and boredom. The choice of the number of questions was driven by a desire to avoid a ceiling effect on reading scores. It was considered very unlikely that any reader would be able to complete all 57 questions in 15 mins and this expectation was upheld.

2.1.2 Participants

There were 53 participants (divided into a group of 26 and a group of 27) who primarily came from the first year Psychology class at the University of New South Wales, though 30% were taken from a class of third year students. All were native English speakers. Course credit was given to the first year students for participation.

2.2 Results and Discussion

Cutoff values for each subject were calculated as two standard deviations above or below the mean across all correct “yes” responses. In this and in all further experiments, any
response time falling outside of the cutoff values was replaced by the relevant cutoff value. As is standard in lexical decision experiments, two analyses were carried out for each contrast; one comparing the conditions within participants averaging across items (F₁) and the other comparing conditions within items averaging across participants (F₂).

The mean reaction times to the PS and BOSS conditions were 816 ms and 828 ms respectively, a difference favoring the PS that was by no means significant, F₁(1, 51) = 1.15, p > .1; F₂(1, 29) = 0.83, p > .1. Neither was there any difference in error rates, with 7.5% errors for the PS and 6.2% for the BOSS, F₁(1, 51) = 1.26, p > .1; F₂(1, 29) = 0.59, p > .1. The mean nonword response time was 1123 ms.

It is clear from these results that, unlike the experiments of Taft (1979, 1987), there is no sign of any advantage of the BOSS over the PS, a finding that is in keeping with the failure of other laboratories to find support for the BOSS. However, this does not preclude the possibility that the use of the BOSS in lexical processing is related to reading ability, which will now be examined.

2.2.1 The reading ability measure

There was considerable variability in performance amongst the participants on the reading comprehension test, with the number of questions answered ranging from 14 through to 43 (mean = 25.6) and the number correct ranging from 5 through to 39 (mean = 17.9). In determining whether the difference in RTs between the BOSS and PS conditions correlated with reading ability, the total number correct answers was taken as the primary measure of reading performance. This is in line with the reading comprehension score generated from the Nelson-Denny Reading Test (1973), where it is expected that most university students will complete all of the questions, and where percentage correct is calculated. In addition to the total number correct, however, the total number of questions completed will also be reported, in case this turns out to be a more informative measure.

2.2.2 The BOSS preference measure

The simplest way to measure BOSS preference for each participant would simply be to subtract the mean of the BOSS condition from that of the PS condition. Such an approach, however, ignores three things:

(a) Items were distributed differently to the two groups of participants such that the same word was seen under the BOSS condition in one group and under the PS condition in the other. This means that the items seen under the BOSS condition by one set of participants may not necessarily have been equated with the items seen under the PS condition by the same individuals in terms of word frequency, length, and so forth. This is immaterial for the analysis that compares the two conditions overall because of the use of the two groups, but the size of the BOSS preference for any one participant could be affected by an uneven distribution of words to conditions. For example, if the items used in the BOSS condition for Group 1 happen to be harder words than those used in the PS condition, then the reverse will be true for Group 2, and it might therefore be found that all participants show a BOSS preference in Group 1 and a
PS preference in Group 2 for that reason alone. If so, a correlation with reading ability measured across both groups will be contaminated by this factor.

In order to avoid this potential problem, the difference between the BOSS and PS means for each participant was simply standardized against the mean for all participants in the same group so that z-scores were used in the correlation.

(b) When comparing mean scores in a reaction time study, it is very possible that the larger the means are, the greater the absolute difference that is required between them for that difference to be meaningful. In other words, what is important might be the magnitude of the difference between means in proportion to the overall reaction time. Thus, the 20 ms difference between mean RTs of 1500 and 1520 is minuscule compared to the 20 ms difference between mean RTs of 500 and 520 ms. For this reason, the absolute difference between the mean RTs for the BOSS condition and the PS condition for each individual might be misleading because a 20 ms difference would indicate the same sized preference regardless of whether that subject was very fast or very slow.

Therefore, two types of z-scores were examined in the correlations: One based on the absolute RT difference between the BOSS and PS conditions for each participant, and one based on the ratio of that difference to the average RT for the participant. It can therefore be seen which of these two measures of “BOSS preference,” if either, is the more sensitive to reading ability. These BOSS preference measures take on a positive value when an individual is relatively faster on the BOSS items than the PS items and a negative value when the opposite is true.

(c) If a participant makes a large number of errors in one condition, the RT for that condition is going to be a less accurate reflection of the difficulty of that condition than when most of the items in that condition elicit a correct (i.e., useable) response.

The criterion adopted here was that if a participant made more than 30% errors in either of the word conditions, their RT data was not included in the correlation. In fact, only two participants were rejected from the experiment on these grounds, leaving 51 scores for use in the correlation.

2.2.3 The correlations

There was a strong indication in the data for a correlation between BOSS preference and reading ability. For the Total Correct score, the correlation was significant when using the ratio-difference z-scores, though only at the 1% level, $r = +0.252$ (and not significant at all when using absolute-difference z-scores, $r = +0.172$). When taking the Total Completed score as the reading measure, the correlation was clearly significant when using ratio-difference z-scores, $r = +0.334, p < .02$, (and, again, much reduced when using absolute-difference z-scores, $r = +0.267, p < .1$).

It can be seen, then, that the first experiment gives some suggestion that reading ability does modulate the preference in word recognition for a BOSS analysis relative to a phonologically consistent syllabic analysis. When reading performance is measured in terms of the number of items completed, it appears that the better the reader, the more likely a BOSS-divided word will coincide with the stored representation of that word.
relative to a syllabically divided word. There is similarly a strong trend when reading is measured in terms of the number of items correctly answered. It seems reasonable to suggest that the more efficiently a reader can decode the words of a passage, the more the resources that will be available for comprehending that passage. The argument can then be made that representing words in the decoding system in terms of their BOSS is more efficient than representing them in terms of their phonological syllable structure or, alternatively, that greater reliance on the phonological representation of a word is less efficient than reliance on a purely orthographic one.

If efficiency of decoding is the relevant factor, one might expect that mere speed of lexical decision would correlate with BOSS preference. Speed of lexical decision could be taken as a possible measure of efficiency of lexical processing and, indeed, faster average lexical decision times across word items were associated with better reading performance both in terms of number completed ($r = -0.407, p < .001$) and number correct ($r = -0.521, p < .001$). When examining the relationship between lexical decision time and BOSS preference, however, the correlation was not significant, though it was very close, at least when ratio-difference $z$-scores were used, $r = -0.268, p < .1$ (and $r = -0.156$ for absolute $z$-scores).

Note, though, that a lexical decision response reflects not only the ease of access to the representation of the word in lexical memory, but also the time taken to transform this process into a response. Adopting an activation account of lexical access (see e.g., Taft, 1991), the mere activation of a lexical representation cannot be taken to mean that the stimulus is necessarily a word, because a nonword will also partly activate lexical representations. Thus, words need to be discriminated from nonwords, and readers may well differ in their sensitivity to this discrimination. So there are likely to be readers who can efficiently access lexical information about a word presented in a passage, but who are very cautious in using that information to make the judgment that that item is a word in a lexical decision experiment. For this reason, speed of lexical decision might not necessarily be expected to correlate highly with BOSS preference if the latter is a reflection of efficiency of lexical access.

Two measures of BOSS preference were examined in this experiment; one where $z$-scores were based on the absolute difference between the BOSS and PS divisions and one where they were based on the ratio of this difference to average RT. It was suggested that the latter measure would be a more accurate reflection of preference and this expectation appeared to be upheld inasmuch as correlations using the ratio score were always higher than those using the absolute score. For this reason, in the further examination of BOSS preference to be reported here, the ratio measure will be exclusively used.

The first of these further studies simply repeats the same conditions as those examined in Experiment 1. The purpose of this experiment is twofold. First, the correlation of BOSS preference with Total Correct, perhaps the best a priori measure of reading ability, was only of marginal significance in Experiment 1 and needs to be replicated. Each participant saw only 15 items in each condition and therefore there was considerable opportunity for error in measuring BOSS preference. Experiment 2 doubles the number of items used. Second, by increasing the number of participants for whom there is a measure of BOSS preference (i.e., by combining the two experiments), it is possible to select a large group at each end of the reading ability scale in order to directly examine whether good readers actually process BOSS-divided words more quickly than PS-divided words and poor readers do the opposite.
3 Experiment 2

3.1 Method

Experiment 2 followed exactly the same procedure as Experiment 1 (except that the physical gap between units was reduced to two spaces because the gap of three spaces used in Experiment 1 seemed unnecessarily large).

An entirely new set of materials (see Appendix 2) was constructed for the second experiment, this time with 60 words (mean frequency of 22.8 per million) split equally between the BOSS and PS conditions for the two groups of participants. These words were of the same type as used in Experiment 1 (i.e., with a long first vowel, e.g., MUN ICIPAL vs. MUN ICIPAL, POI SON vs. POIS ON; or with a medial consonant cluster, e.g., CAR BON vs. CARB ON, CLUS TER vs. CLUST ER) though they also included cases where the first vowel was reduced to a schwa (e.g., FA TIGUE vs. FAT IGUE, SO CIETY vs. SOC IETY, BO TANICAL vs. BOT ANICAL\(^1\)). Sixty nonwords were also constructed in the same manner as in Experiment 1 (e.g., RAC TIOUS, JAND ER, KO BIN, NAB ULA TATE).

There were 63 participants (divided into groups of 31 and 32) taken from the same source as in Experiment 1, though fewer third year students participated (16%).

3.2 Results and Discussion

The mean response times for the PS and BOSS conditions were 806 ms and 798 ms respectively, a nonsignificant difference, both \(F's < 1\). The error rate difference was similarly nonsignificant, both \(F's < 1\), with percentage error rates of 7.7\% and 5.8\% for the PS and BOSS conditions respectively. Mean nonword response time was 1106 ms.

As in Experiment 1, then, there was no significant BOSS preference across the whole set of participants, so the focus now turns to the correlations with reading comprehension. There were seven participants who were rejected from the correlational analysis on the grounds that they made more than 30\% errors in at least one of the word conditions, thus leaving 56.

The mean number of comprehension questions completed was 25.4, with the mean number correct being 18.2. When it came to the correlations, a somewhat similar pattern was found in Experiment 2 as was observed in Experiment 1. The correlation between BOSS preference and Total Completed was highly significant, \(r = +0.390, p < .001\), but this time the correlation with Total Correct was even higher, \(r = +0.460, p < .001\).

4 Combining Experiments 1 and 2

Given that both experiments revealed a similar pattern of correlations, it is hardly surprising that correlations performed on the combined total of 107 participants showed the same thing:

---
\(^1\) Some of the items (like BOTANICAL, HABITUAL, CIRCULAR etc.) were not monomorphemic, but what is important for the experiment was simply that they had a polysyllabic stem.
\( r = 0.365, \ p < .001 \) for Total Completed, and \( r = 0.363, \ p < .001 \) for Total Correct. It is therefore apparent that reading comprehension performance (at least, under time pressure) is related to whether a BOSS or a PS analysis of words is favored. The scattergram mapping BOSS preference against Total Correct in the comprehension test is shown in Figure 1.

To examine directly whether good readers show a BOSS preference while bad readers show a PS preference, two groups were differentiated on the basis of their reading comprehension scores. In forming these groups of good and poor readers, two decisions had to be made. First, it needed to be decided on what measure the determination of “good” and “poor” should be based. Total Correct was chosen because, apart from revealing a significant correlation with BOSS preference, it was considered a priori to be the best measure of accurate reading performance.

The second decision that needed to be made was where the “good” and “poor” cutoffs should be drawn. It was found that 24 participants scored 23 or more correct while 26 scored 12 or less correct. These scores were therefore taken to be the cutoff values because they were the nearest approximation to providing the top and bottom quartiles.

The mean RTs for the BOSS and PS conditions were compared for the 24 good readers (mean reading score = 27.9 correct) and 26 poor readers (mean reading score = 10.2 correct) separately in a 2 × 2 factorial design. The mean scores are presented in Table 1. In the analysis of participant means that follows, there were four groups of good readers and four groups of poor readers, being taken from each of the two groups of participants in each of the two experiments. In the analysis of item means, the two experiments were treated as separate groups for each of the two levels of reading ability. A significant interaction was obtained between division type and reading ability, \( F_1(1, 42) = 12.79, \ p < .001; \ F_2(1, 88) = 8.84, \ p < .01 \). Analysis of the good readers alone revealed a significant

---

2 With the total number of participants being 107, each quartile should contain approximately 27 individuals. For good readers, the best approximation to this was 24 individuals because there were eight scores of 22. That is, if the top quartile were defined as including scores of 22, it would have included 32 individuals. Note that inclusion of these scores does not change the outcome of the ANOVA.
preference for the BOSS analysis, $F_1(1, 20) = 11.35, p < .01; F_2(1, 88) = 7.96, p < .01$, while analysis of the poor readers alone revealed a significant preference for the PS analysis, $F_1(1, 22) = 8.76, p < .01; F_2(1, 88) = 6.39, p < .05$.

Analysis of error rates revealed no significant main effect of division type ($F$'s < 1.6) nor an interaction with reading ability ($F$'s < 1). Neither was the main effect of reading ability significant, $F_1(1, 42) = 2.22, p > .1; F_2(1, 88) = 3.67, p > .05$. The 2.1% error difference for the good readers alone was not significant, $F_1(1, 20) = 2.32, p > .1; F_2(1, 88) = 3.58, p < .05$.

The results are clear. Faster responding when words are divided according to their BOSS than when divided according to their phonologically defined syllable is associated with proficient comprehension. Less proficient comprehenders appear to be more oriented to the phonological characteristics of the words and, therefore, find a BOSS analysis detrimental to their processing of a word because of its inconsistency with pronunciation.

The relationship between better reading comprehension and a BOSS preference can be interpreted in several ways. On the one hand, it may be the case that those who perform better on the reading comprehension test (even if it is because of a higher verbal IQ), are more proficient in making use of the fact that the BOSS is more informative about the whole word than the first syllable. How this greater informativeness might play its role is considered in the General Discussion. On the other hand, it is possible that a BOSS analysis is the optimal orthographic processing strategy and hence its use is most efficient for the decoding of words which, in turn, allows greater resources to be available for comprehension. According to this “shared limited capacity” view (Perfetti & Hogaboam, 1975), the more efficient the decoding process, the fewer the demands on the reader's higher comprehension processes.

One thing that is clearly suggested by the results of Experiments 1 and 2 is that the mixed evidence arising from different laboratories in relation to the existence of the BOSS can potentially be put down to differing reading abilities of the participants used. In a nonselective group of participants, there are likely to be a wide range of reading abilities, even amongst university-level students. As is seen in the present study, no overall preference was observed, but this obscured the fact that those at the high end of the distribution of comprehension scores showed a BOSS preference, while those at the other end showed a PS preference. It is apparent that even university-based poorer readers can be discriminated from better readers in terms of their processing strategies.
5 Experiment 3

The next experiment looks again at the relationship between reading ability and BOSS preference. This time, however, the relationship is not examined within a correlational analysis, but rather, a factorial design. That is, a group of good readers were directly compared to a group of poor readers. This required an initial large-scale screening of first year Psychology students at the University of New South Wales using the reading comprehension test, in order to establish who might be suitable participants for the experiment.

The words used were all examples where the BOSS and PS did not coincide. However, the items were split into two types in order to look in more detail at the syllable preference of the poorer readers. If these readers are preferring the syllable because they are actually reading via a phonological conversion of the orthographic rendition of the word, they should prefer a syllabic analysis regardless of the type of word selected. There is another possibility however.

It may be the case that the poorer readers are using their orthographic recognition system (without on-line phonological conversion), but that the representations of words within that system are broken down in a way that parallels the phonological syllabic structure. That is, longer words are broken down within the orthographic system, and pronunciation is one obvious basis on which to guide this segmentation. However, it may not be the only basis for segmentation. For example, morphological factors may override the phonologically based segmentation such that, for example, ACTOR is broken down into ACT + OR rather than AC + TOR, and LEAKY as LEAK + Y rather than LEA + KY.

Morphological factors need not even have to be as obvious as these examples. It is possible for apparently monomorphemic words to have a BOSS that combines with other endings to make semantically related words. For example, although VIRUS is likely to be classified as constituting a single morpheme, it shares its BOSS with VIRAL, which is clearly related in meaning. Similarly, the putatively monomorphemic word CERTAIN shares its BOSS with the semantically related CERTIFY. It might be the case, then, that even poorer readers prefer VIR US (i.e., with a BOSS analysis) to VI RUS (i.e., with a PS analysis) because the existence of the related words alerts them to the importance of the BOSS in these cases. Experiment 3 tests this possibility by comparing such words to those for which there is no such “morphological” guidance to segmentation (e.g., BOGUS). Although this latter type of word might have a BOSS that occurs in another word (or indeed is a word in its own right, e.g., the BOG of BOGUS), this is totally unrelated to the presented word and therefore does not provide any morphological guidance to segmentation.

While there is a logical basis for the prediction that the existence of VIRAL might accentuate the preference of VIR US over VI RUS, a previous post-hoc examination of such items in the original Taft (1979) study actually suggests otherwise. Taft (1985) points out that those items whose BOSS was shared by another related word failed to show any BOSS preference at all. Indeed, a post-hoc analysis of the combination of Experiments 1 and 2 reported here is consistent with this. In 32 of the 90 words used in Experiments 1 and 2, the BOSS was shared with a related word (e.g., the DICT of DICTATE, cf. diction). For neither the good nor the poor readers was there a significant difference between type of division for these “morphemic cue” items (with 2 ms favoring the syllable for the former,
and 13 ms favoring the syllable for the latter, both $F_2's < 1$). The significant effects observed in the combined analysis of Experiments 1 and 2 arose overwhelmingly from those words whose BOSS was not shared with a related word (e.g., the POIS of POISON), with a 43 ms effect favoring the BOSS for the good readers, $F_2(1, 57)=9.34, p < .01$, and a 51 ms effect favoring the syllable for the poor readers, $F_2(1, 57)=3.98, p < .1$.

If it is true that words with a morpheme-like cue to their BOSS (e.g., VIRUS) fail to show a BOSS preference even for better readers, then further theoretical considerations will be required, since a more natural prediction is that such cues should exaggerate the BOSS preference. Experiment 3 tests this issue directly.

5.1 Method

5.1.1 Materials and procedure

Sixty words were selected whose BOSS and PS could be differentiated on the basis of their maximal coda or maximal onset analysis; that is, they had either a medial consonant cluster, a long first vowel, or an unstressed first vowel (e.g., VIR US vs. VI RUS; CUR TAIN vs. CUR TAIN; CRE MATE vs. CRE MATE). For half of these words, the BOSS could be found in another semantically related word with different letters immediately following it (e.g., VIRUS, cf. viral; CERTAIN, cf. certify; DONATE, cf. donor) and for the other half it could not (e.g., BOGUS, CURTAIN, CRE MATE). The former type of word was labeled the “Morphemic Cue” condition and the latter type the “No Morphemic Cue” condition. The words in the two conditions were designed so that they were roughly equivalent in orthographic structure (e.g., VIRUS vs. BOGUS; CERTAIN vs. CURTAIN; DONATE vs. CRE MATE). The items are presented in Appendix 3. The distractor nonwords were of the same type as in the previous experiments. In this experiment, the items were presented in lower case letters.

The procedure was identical to that used in Experiment 2, except that no reading comprehension test was administered at the time of performing the lexical decision task. The participants had already been classified as good or poor readers prior to the experiment on the basis of this test.

5.1.2 Participants

All first year Psychology students at the University of New South Wales were administered the reading comprehension test during their tutorial classes at the beginning of the academic year and indicated whether they were willing to be contacted at a later time for participation in a laboratory experiment. From the pool of willing students, 20 native English speakers who were classified as “good” readers and 20 who were classified as “poor” readers subsequently participated in the experiment. The criteria for good and poor reading were based on the scores determined from the combination of Experiments 1 and 2. Poor readers were taken to be those who got 12 or fewer items correct on the reading test, while good readers were those who got 22 or more items correct. The mean score for the good readers was 25.2 correct, while for the poor readers it was 9.5.
5.2 Results and Discussion

The mean reaction time and error rates for the word items of this experiment are presented in Table 2.

All of the following analyses are based on the RT data because no effects of error rate were observed. The only error rate contrast that approached significance (including interactions) was the good/poor difference for the Morphemic Cue words, $F_1(1, 36)=3.16$, $p<.1$; $F_2(1, 29)=4.01$, $p<.1$. Looking first at the analysis of the No Morpheme Cue condition (e.g., BOGUS), a similar pattern of results was found to that obtained in the combined analysis of Experiments 1 and 2, where the majority of words also provided no morphemic basis for a BOSS analysis. In particular, there was a main effect of reading ability, $F_1(1, 36)=14.59$, $p<.001$; $F_2(1, 29)=119.8$, $p<.001$, no main effect of division type, both $F_1$ and $F_2<1$, and importantly, an interaction between the two, $F_1(1, 36)=4.48$, $p<.05$; $F_2(1, 29)=6.24$, $p<.02$. Separate analyses of good and poor readers revealed a significant BOSS preference for the former, $F_1(1, 18)=4.42$, $p<.05$; $F_2(1, 29)=4.76$, $p<.05$, though only a trend in the direction of a PS preference for the latter, $F_1(1, 18)=1.30$, $p>.1$; $F_2(1, 29)=1.56$, $p>.1$.

3 The upper criterion is actually different to that used in the combined analysis of Experiments 1 and 2, where it was a score of 23 or more. The reason for lowering the criterion by one point in Experiment 4 was simply so that enough good readers could be included in the study.
These results again point to the association between better reading ability and BOSS analysis. The interaction indicates that better readers and poorer readers are responding differently to the division of the words, though the association of poorer reading with syllabic analysis is not given as strong statistical support here as it was in the combined analysis of Experiments 1 and 2.

Turning now to the Morphemic Cue items (e.g., VIRUS), it was found that there was no divisional preference for either the good or the poor readers, with the only significant difference being the main effect of reading ability, $F_1(1, 36) = 11.06, p < .001; F_2(1, 29) = 18.34, p < .001$.

The manipulation of morphemic cues was carried out to see whether poorer readers lose their syllable preference when the BOSS of the word is morphemic in nature, while better readers maintain their BOSS preference regardless of this manipulation. The morphemic cues used here were quite subtle in that the word in question would be considered to be monomorphemic even though a semantically related word sharing the same BOSS does exist (e.g., VIRUS is related to VIRAL, CERTAIN is related to CERTIFY). Despite their subtlety, however, it was thought that even poorer readers might be sensitive to these relationships and therefore give the word an internal structure compatible with the BOSS. In the event, it was certainly the case that poorer readers showed no syllable preference when the BOSS was morphemically cued, though the interaction between type of item (i.e., existence of a morphemic cue) and type of division (i.e., BOSS or syllable) was only marginally significant on the participant analysis, $F_1(1, 18) = 3.92, p < .1; F_2(1, 29) = 1.82, p > .05$. However, the presence of a morphemic cue also eliminated the BOSS preference for the better readers, for whom the interaction between type of item and type of division was significant, $F_1(1, 18) = 5.13, p < .05; F_2(1, 29) = 4.65, p < .05$. Thus, the post-hoc analyses of the Morphemic Cue items from Taft (1979), and from Experiments 1 and 2 are supported by the direct manipulation of this factor here, namely, there was no BOSS preference.

How is such a pattern of results to be interpreted? Taft (1985) suggested that splitting a word according to its BOSS (e.g., VIR US) might facilitate access to that word, but also access to the word that shares its BOSS (i.e., VIRAL). Therefore, the facilitated access is counteracted by an increase in competition (see also Leong, 1989). This suggestion needs expansion, though, in order to determine whether it really provides a potential explanation for the data. After all, in about a third of the items in the No Morphemic Cue condition, the BOSS was also shared with another word (e.g., the BOG of BOGUS), where this other word was not semantically related to the item. Therefore, there is the potential for competition here as well. In fact, this type of item shows just as strong a BOSS preference for the good readers (a 33 ms difference) as for those items whose BOSS is not a word (a 24 ms difference). So, it needs to be explained why words like BOGUS do not also experience an increase in competition when split according to the BOSS.

Figure 2 presents one way of conceptualizing how this might come about. It illustrates one possible framework for describing the structure of lexical representation for better readers. The orthographic system comprises onsets and bodies (see Taft, 1991). Not only are monosyllabic words broken down into onsets and bodies (e.g., BOG broken down into onset B and body OG), but so are polysyllabic words, such that a BOSS comprises an onset and body (Taft, 1992).
There are also abstract units of representation that intervene between form and meaning, and which essentially bring together a set of semantic features to link them to the relevant onset and body combinations. These are labeled as “lemmas” (Taft, in press), and may be seen as units that capture correlations between form and meaning. Lemmas themselves can be interlinked such that when two words are related in both meaning and form (e.g., VIRUS and VIRAL), a shared unit might develop with a link to their mutual semantic features. This unit is essentially a morpheme, capturing the fact that a particular form is associated with a particular meaning, and its strength of representation will be related to how constant its associated meaning is in the face of varying contexts (see Taft, in press). Thus, bound morphemes will be represented, but more weakly than will free morphemes (i.e., free-standing words) because the latter show more consistency in meaning as the context changes. A meaningless BOSS (like the BOG of BOGUS) will have no lemma representation because the link to the semantic level is captured entirely through the lemma for the whole word in which it occurs.

It can be seen from the figure that the relationship between VIRUS and VIRAL is quite different to that between BOG and BOGUS. The former two words are both activated via the same lemma unit (‘vir’), with the competition resolved primarily through top-down activation, that is, by determining which of the two activated lemmas is more compatible with the presented stimulus. Syntactic constraints could also contribute to the resolution when the word is presented within a sentential context. On the other hand, two orthographically related but semantically unrelated words (e.g., BOGUS and BOG) activate different lemma units via overlapping orthographic units, with the competition resolved through the activation arising from the nonoverlapping orthographic units. Thus, there is a qualitative difference in the nature of the competition engendered by the two types of item. In order to explain the obtained pattern of results, however, it needs to be argued that any influence of BOG on the recognition of BOGUS will arise under both presentation conditions (i.e., both BOG US and BO GUS), whereas the influence of VIRAL on the recognition
of VIRUS will have more of an impact under the BOSS condition (i.e., VIR US) than under the PS condition (i.e., VI RUS), thus counteracting the BOSS preference.

One possible way of incorporating this requirement is to speculate that there are actually two orthographic routes to the lemma. One is where the lemmas are activated via onset and body units as described so far, while the other is where the lemmas are activated directly from lower level letter units (e.g., from units for V, I, R, U, and S to the “virus” lemma). It can then be argued that the onset/body system develops as a more efficient route to the lemma (at least for better readers), such that the direct letter route simply exists as a slower back-up mechanism.

Perhaps the onset/body system is more efficient because it provides more information about the position and ordering of the letters within a word. That is, if I and R are incorporated into a body IR, then their possible position and ordering in the word is more constrained than if they were merely treated as individual letters (see also Plaut et al., 1996). Furthermore, classification into an onset/body structure provides greater constraints on the number of activated candidates. For example, if a letter is classified as an onset, then activation will occur in only those lemmas where that letter corresponds to an onset, as opposed to a coda. Finally, by grouping letters together, commonly used subunits can be activated more rapidly than rarely used ones, which is something that could not be captured through activation via single letters.

The argument is then that the presentation of VIR US allows rapid activation of the “vir” lemma which leads to competition between the lemmas for “virus” and “viral”. When VI RUS is presented, on the other hand, the delay in activating the lemma “vir” gives time for the letter route to come into play which directly activates the “virus” lemma thus generating less competition from “viral”. In the case of BOG US, the physical structuring of the stimulus allows rapid activation of the “bogus” lemma, though there might also be some competition from “bog”. However, such competition also exists when BO GUS is presented because the direct letter activation also activates “bog” along with “bogus”.

The above is obviously highly speculative, but offers a possible scenario whereby the BOSS preference may or may not be affected by competition between words that share the BOSS depending on the semantic relationship between the competing words. The suggestion is that the lack of a BOSS preference in the recognition of words that share their BOSS with another related word should not be taken to imply that such words lack a BOSS-like internal structure. On the contrary, it may even be that poorer readers represent them with such a structure, overriding any structure based on phonological considerations. It is the competition between the related words sharing the BOSS that counteracts any advantage of having the word presented with its BOSS highlighted. Obviously, such an argument would be unconvincing if made purely on the basis of the failure to find a BOSS preference for the Morphemic Cue items per se; but it is from the contrast with the No Morphemic Cue items, where there is freedom from such competition, that the argument gains its force. Under the No Morphemic Cue condition, the preference for a BOSS analysis is revealed, though for the better readers only. What poorer readers might be doing when there are no morphemic cues is something that will be taken up in the General Discussion.
The three experiments presented in this paper provide evidence that better adult reading is associated with a preference for reading words divided according to the principle of maximal coda rather than the principle of maximal onset. Poorer readers, on the other hand, prefer a division that coincides with the spoken syllable structure, namely, a maximal onset analysis, though this preference is apparently overridden when morphological considerations suggest a nonphonological analysis. These results indicate that evidence for the use of the BOSS, at least within the split-word lexical decision task, will only be observed if the selection of participants is biased toward the upper end of the range of reading abilities. This approach therefore has the potential to say something important about optimal reading strategies and how poorer readers differ from better readers. However, the current experiments do not explicitly identify what this might be because there are various ways of interpreting the results. What follows is a consideration of some of the possible accounts of the difference between good and poor readers.

The first possibility is that poorer adult readers are using an orthographic input system where polysyllabic words are broken down into manageable sized units on the basis of pronunciation. Better adult readers, in contrast, use a system that is sensitive to the greater efficiency in coding a polysyllabic word in terms of its maximal first unit (i.e., the BOSS) than in terms of its spoken syllable structure. That is, while better readers store THUNDER as THUND and ER (or TH, UND, and ER, if onsets are separated from bodies), poorer readers store it as THUN and DER. The results of Experiment 3, however, indicate that morphological considerations, even weak ones, can override the phonological cues used by poorer readers inasmuch as there was no sign of a syllable preference for poorer readers when morphological cues cut across the syllable structure (as with VIR US). The fact that such items did not actually generate a BOSS preference can either be explained in the same way as was proposed for the better readers (i.e., in terms of counterbalancing competition from the related word), or else by suggesting that only some of the poorer readers were influenced by morphological cues. So, this first account says that all adult readers rely on their orthographic input system when processing words on-line, and that the difference in reading ability is related to the nature of the orthographic representations.

A second possibility, however, is that poorer adult readers typically make greater use of their phonological input system when recognizing words than do better adult readers. This idea has been supported by the work of Lewellen et al. (1993) and Jared et al. (1999) looking at homophone confusions in silent reading, but is complicated by the results of Experiment 3 where morphological considerations seemed to take precedence over a syllabically based analysis.

In order to account for the possible involvement of morphological structure within the phonological system, it can be suggested that the phonological representations used in reading do not always coincide with the spoken form of the word. That is, perhaps phonological representations are more abstract than the spoken surface form, at least when there are morphological considerations, and that the surface syllabic structure is generated from this abstract representation only when overt output is required.
If the results for the poorer readers are indeed interpreted in terms of their greater reliance on the phonological representations of words, then the nature of their orthographic representations remains an open question. The syllabically coded phonological representations that the poorer adult reader uses could be activated via orthographic representations that have no systematic structure at all. It is possible, though, that their orthographic representations are structured according to BOSS principles, just like the better readers, but their lexical decision performance does not reflect this because of their reliance on phonological representations for word recognition.

The idea that all readers have similarly structured orthographic representations is compatible with the idea that units representing the internal structure of words emerge automatically from the statistical characteristics of the language (cf. Seidenberg, 1987, 1989). That is, as the orthographic processing system develops in a way that will optimize the discriminability of one word from another, the BOSS (or an approximation to the BOSS) emerges as a unit that will allow this. Hence, readers will develop an orthographic input system with similar representational units if they have been exposed to much the same distribution of words in the language regardless of their reading ability. Developmentally, the BOSS is only likely to become relevant once there is a large enough vocabulary of polysyllabic words for it to be advantageous for the system to make use of it. With exposure to more and more words, the BOSS-like nature of these units might become increasingly refined. This in itself might be used to explain the difference between good and poor readers in their use of the BOSS because the latter are likely to have been exposed to fewer words than the former. However, it is hard to see why poorer readers would then be more biased toward a phonological syllable than would better readers rather than simply being biased toward the BOSS but to a lesser extent.

Instead, it can be suggested that poorer adult readers make less use of their (BOSS-based) orthographic processing system when they read, and place more reliance on their phonological processing system than do better readers. In this way, while the better readers demonstrate the existence of a BOSS-oriented orthographic system, poorer readers find that the PS structure is more compatible with their predominantly phonological processing.

It may well be that phonological processing is important in the initial stages of learning to read (see e.g., Adams, 1990), but that readers become more advanced in their skills as they increasingly make use of an orthographic system that by-passes phonological mediation. Perhaps the poorer adult readers are those who have not made this progression. Such a conclusion appears to be at odds with recent claims that proficient adult readers use a predominantly phonological processing system in silent reading (e.g., Frost, 1998; Lukatela, Lukatela, & Turvey, 1993; Van Orden, 1987), but is consistent with the conclusions drawn by Taft and van Graan (1998) that a direct orthographic route to meaning dominates over a phonologically mediated route. In fact, according to the account presented here, evidence for the importance of phonological mediation will depend on the reading ability of the individuals being tested (cf. Jared et al., 1999; Lewellen et al., 1993).

The idea that poorer reading is influenced by spoken syllable structure, however, would be hard to accept if it were shown that English speakers are not actually sensitive to the locus of syllable boundaries in spoken words. The research of Cutler et al. (1986) suggests that they are not. In their study, a target unit was to be detected at the beginning of a spoken word (the target being presented visually), and it was found that syllable targets were no easier to judge than nonsyllable targets (e.g., detecting BAL in “balance” was no
easier than detecting BA in “balance”). As it happens, however, other research into spoken syllable detection has failed to support these results, and has found a significant detection advantage for syllable targets over nonsyllable targets for English speakers (Bradley, Sánchez-Casas, & García-Albea, 1993). Therefore, the research on spoken syllabic processing is not necessarily incompatible with the results obtained here.

It can therefore be proposed that poorer adult reading is characterized by a greater sensitivity to phonological considerations, either in terms of the way orthographic representations are structured (e.g., according to spoken syllabification) or in terms of a reliance on phonological representations during silent reading. Better adult reading, on the other hand, is characterized by a greater sensitivity to orthographic structure. Such an idea has been previously suggested in relation to other aspects of orthographic structure. For example, Mason (1978) showed that better adult readers were more sensitive to orthographic regularity than poorer readers in a naming task, though Massaro and Taylor (1980) failed to find effects of orthographic structure in a letter searching task.

Butler, Jared and Hains (1984) also concluded that better readers make greater use of orthographic structure than poorer readers because the former made more errors in a tachistoscopic word identification task when a nonsyllabic temporal presentation was imposed (e.g., co-ni-tue) relative to a syllabic one (e.g., con-tin-ue). Such a result may at first seem to be in conflict with the present study in that the poorer readers in the Butler et al. experiment showed no difference in error rates between syllabically and nonsyllabically presented words. This might suggest that the syllable has no special status for poorer readers, which would imply that no preference for the syllable over the BOSS should have been found for such readers in the present study. In response to this, however, while acknowledging that the studies used different tasks, it can be pointed out that the poorer readers in the present study also showed no syllable preference in relation to error rates (see Tables 1 and 2). It was only on reaction times that any sign of a syllable preference emerged.

If poorer readers are basing their responses on phonological representations, it was suggested above that these representations would be not always be veridical with pronunciation, notably when morphological structure cuts across the spoken syllable boundary. This idea that phonological representations can be abstract raises another interpretation for the performance of the better readers. In particular, rather than the BOSS preference arising from their orthographically oriented reading processes, it may be that they also make use of phonological representations when reading, but that these are structured in terms of the maximal coda principle such that they do not coincide with the spoken form. For example, ROUTINE might be stored as /ru:t + i:n/, with the use of readjustment rules leading to maximization of the onset (i.e., the spoken syllable structure) only when pronunciation is overtly generated. Such an idea is a direct way of capturing the suggestion that the BOSS is influenced by phonological principles (i.e., because the legality of a coda is defined phonologically⁴). This possible interpretation of the experiments therefore suggests that reading ability relates to the degree of abstractness of phonological representations, with better readers using a structure that optimizes the informativeness of the first unit regardless of the pronunciation of the word.

⁴ In fact, the idea that a BOSS cannot contain an illegal coda (like the DN of KIDNEY or the BL of TABLE) has only been assumed on the grounds of logic, but has never been explicitly tested.
Although plausible, there are two comments to make about this possibility. First, it seems incompatible with findings that better readers show little evidence of phonological processing in other tasks (e.g., Jared et al., 1999) and, second, that it becomes virtually impossible to distinguish between phonological and orthographic processing. Indeed, at least in relation to internal word structure, the two become functionally equivalent.

The above interpretations of the data adopt the notion that BOSSes are explicitly represented in the lexical processing system (as exemplified in Figure 2). It is necessary, however, to consider possible interpretations of the BOSS preference that do not make such an assumption, but instead, explain it on the basis of other structural factors that are confounded with the BOSS/PS contrast. In all experiments reported here, the BOSS was usually more predictive of the whole word than was the PS. For example, the MU of MUNICIPAL could orthographically be the beginning of MUTATE, MUSIC, MUSEUM, MUSTARD, MUNDANE, and so forth, while the only other completions of MUN are MUNDANE and MUNITIONS. By presenting the first fragment of the word quite separately from the second fragment, it may be that the participants were activating a set of candidates on the basis of the first fragment, hence priming the recognition of the whole word if it turned out to be amongst this candidate set. The more predictive the initial fragment is of the word, the greater the likelihood that the candidate set will include that word and, therefore, the BOSS divided items would be recognized more quickly than the PS divided items. Perhaps good readers are better than poor readers at making use of such a strategy; a strategy that must be orthographically based because the BOSS division is not phonologically predictive of the whole word given that it cuts across the syllable boundary. Poor readers, on the other hand, may rely more on phonological factors in generating the word from the fragment.

What follows from this idea is that adding one letter to the BOSS should make it even more informative than the BOSS and therefore a BOSS + 1 division (e.g., THUNDE R) should show faster response times than a BOSS division. While Lima and Pollatsek (1983) did report such a BOSS + 1 advantage in one of their fragment priming experiments, no evidence for this was observed by Taft (1987) using three different tasks (including fragment priming) and, if anything, the trend was for slower response times to BOSS + 1 items than BOSS items.

Nonetheless, there still remains the possibility that the BOSS effects observed in the present set of experiments for better readers might really reflect their efficiency in the use of larger orthographic units and this awaits further testing. However, even if true, the results of the present study point to an important difference between good and poor readers, namely that poorer readers are more reliant on phonology than are better readers whose processing of words is more sensitive to orthographic information.

Another potential structural confound that could explain the results is found at the level of the bigram. In particular, it is conceivable that the frequency of the two letters straddling the boundary is different for the BOSS division and the PS division. For example, the bigram IN occurs more often in English words than does the bigram NA, which means that FIN AL (i.e., BOSS division) provides a more natural break than does FI NAL (i.e., PS division). If it were systematically the case that the BOSS division optimized the clustering of letters in terms of bigram frequency, then it could be argued that better readers are simply more sensitive to letter co-occurrence information than are poorer readers and
that their reading system is more responsive to the statistical characteristics of the language. As it happens, however, the example of FINAL was an exception in Experiments 1 and 2, and the bigram frequencies straddling the PS boundary ($V = 1196$, using the versatility measure of Solso, Barbuto, & Juel, 1979) were actually less rather than more frequent than those straddling the BOSS boundary ($V = 1714$), $t(89) = 3.42, p < .001$. In fact, the bigrams straddling the PS boundary formed a “bigram trough” (see Seidenberg, 1987), because they were also less frequent than the bigrams preceding the PS boundary ($V = 1788$), $t(89) = 3.11, p < .01$. For example, OU is more common than UT which is less common than TI, which means that the word ROUTINE naturally splits at the PS boundary (giving ROUTINE). Therefore, to explain the results of the first two experiments, it would need to be said that it is the poorer readers who are more sensitive to bigram frequencies than are the better readers, hence preferring the PS division. Such an argument cannot be maintained, however, because a nonsignificant trend in the opposite direction between the frequency of PS boundary bigrams and BOSS boundary bigrams was found for the materials used in Experiment 3 for both the No Morphemic Cue (1624 vs. 1461), $t(29) = 0.12, p > .1$, and Morphemic Cue conditions (1837 vs. 1249), $t(29) = 1.29, p > .1$. So, even though the bigram frequency distribution of the words used in Experiments 1 and 2 was quite different from that of the No Morphemic Cue items of Experiment 3, a similar pattern of response times was observed. Such an outcome therefore suggests that bigram frequencies do not offer an explanation for the results of the present research.

Finally, it must be emphasized that only one experimental paradigm was used in the studies reported here, namely, a word-splitting technique that required a lexical decision response. The same task was maintained in Experiments 2 and 3 to confirm, under the same conditions, the suggestive result of Experiment 1 that reading ability has an impact on BOSS preference. Having achieved this end (with a further elaboration in terms of the impact of morphemic relatedness), the conditions under which the effects emerge can now be explored. Taft (1987) used a number of different paradigms to support the BOSS notion (e.g., fragment priming), and it is possible that these and other tasks will uncover greater complexities than does the simple word-splitting paradigm, leading to a more definitive interpretation of what a BOSS or PS preference means.

7 Conclusions

The three experiments reported here show that there is a relationship between reading ability and preference for syllabification of polysyllabic words that maximizes the coda of the first syllable. As such, the preference of good readers is for an analysis that ignores the phonological syllabic structure of the word, and this implies the use of orthographic representations, though it may be interpreted in terms of abstract phonological representations. Poorer readers are more influenced by the pronunciation of words when processing them silently, though it is unclear whether this arises from the use of phonologically based orthographic units or from the use of their phonological mechanisms when reading.

The mixed evidence arising from different laboratories in relation to the existence of the BOSS can potentially be put down to differing reading abilities of the participants used. When an experiment uses participants from a general pool of undergraduate college students, there is likely to be sufficient variation in reading ability for those with a PS
preference to counterbalance those with a BOSS preference. It therefore remains to be seen whether other laboratories are now able to observe a BOSS preference when reading ability is taken into account.

Received: March 01, 2001; revised manuscript received: August 03, 2001; accepted: September 11, 2001

References


---

**Appendices**

**Appendix 1**

The following are the words used in Experiment 1.

routine, ornate, movie, fantastic, quartet, window, turkey, trombone, monkey, antics, stampede, furnish, sturdy, verbatim, radar, journey, dictate, cortex, garlic, banjo, zero, garbage, laundry, umbrella, torpedo, antique, gastric, tarmac, rhubarb, textile

**Appendix 2**

The following are the words used in Experiment 2.

bacon, plaza, bacteria, iguana, carbon, final, circular, sentimental, drama, radio, factor, pension, familiar, original, fatal, municipal, formal, habitual, memorial, ideal, balcony, botanical, carpentry, culture, dictation, dignity, doctorate, during, factory, fatigue, favorable, future, humanity, humor, imagine, manipulate, potential, silence, society, vacant, thunder, radiate, danger, journal, harmony, dandelion, flavor, burden, thimble, cluster, staple, piracy, poison, vibrate, drastic, spider, beagle, capable, copious, fumble

**Appendix 3**

The following are the words used in Experiment 3.

*No Morphemic Cue:*

ankle, crisis, biscuit, feature, lenient, dubious, rendezvous, cable, marsupial, pirate, nectar, barnacle, nurture, bargain, tyrant, fatal, curtain, pony, cremate, termite, jubilant, glucose, blister, petroleum, serpent, zero, cedar, loiter, cactus, bogus

*Morphemic Cue:*

angle, basis, circuit, rupture, gradient, arduous, horrendous, noble, pictorial, private, sonar, spectacle, denture, terrain, vacant, legal, certain, navy, donate, finite, turbulent, verbose, plaster, pendulum, student, solo, radar, neuter, minus, virus