Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Memory and Language 50 (2004) 231–244
Journal of Memory and Language
www.elsevier.com/locate/jml

Word onset patterns and lexical stress in English☆

Michael H. Kelly

National Analysts, Inc., 1835 Market Street, Philadelphia, PA 19103, USA

Received 18 November 2002; revision received 3 December 2003

Abstract

Theories of English phonology regard syllable onset patterns as irrelevant to the assignment of lexical stress. This paper describes three studies that challenge this position. Study 1 tested whether stress patterns on a large sample of disyllabic English words varied as a function of word onset. The incidence of trochaic stress increased significantly with the number of consonants in word onset position. This overall pattern was replicated when various subsets of the database were examined separately, thus demonstrating its robustness and generality. In Study 2, English speakers assigned trochaic stress to disyllabic pseudowords more often when they began with two consonants (e.g., "plonveen) rather than one (e.g., "ponveen"). Study 3 examined implications of these results for monosyllabic words and the rhythmic structure of English poetry. An analysis of Milton's *Paradise Lost* found that monosyllabic words beginning with two consonants (e.g., "stay") appeared in stressed positions in Milton's verse more often than matched words beginning with one consonant (e.g., "say"). These studies indicate that syllable weight, and consequently word stress is affected by onset as well as rime structure.

Keywords: Word onset; Lexical stress; Vocabulary structure; Phonological theory

Theories of English phonology generally propose that syllables are not simple concatenations of independent phonemes. Rather, the phonemes in words like /blɛnd/ form higher level units that act independently in operations on syllable structure. Most theories would agree in grouping the initial and final consonants into onset and coda units which surround a vowel nucleus (e.g., Clements & Keyser, 1983; Fudge, 1969, 1987; Halle & Vergnaud, 1980). Some phonologists go further and also join the vowel and coda into a "rime" unit (Fudge, 1969, 1987; Halle & Vergnaud, 1980).

There is substantial linguistic and psycholinguistic evidence supporting the onset-rime representation of syllable structure. For instance, English vowels are statistically more independent of onset than coda consonants. That is, English phonology contains many restrictions and probabilistic biases about which vowels and consonants can form rimes but few, if any, constraints on which onsets and vowels can appear in words together (Fudge, 1969, 1987; Kessler & Treiman, 1997). Various forms of lexical innovation also suggest that onsets and rimes are separate units. For example, English word blends like "brunch" and "smog" tend to be constructed so that the onset of one syllable (e.g., the /sm/ in smoke) is combined with the rime of another syllable (e.g., the American English /ag/ in "fog"). Potential blends that combine the onset and vowel of one syllable with the coda of another are statistically quite rare (Kelly, 1998). Thus, "breakfast" and "lunch" are merged into "brunch," not "brench." In addition, English speakers find it easier to learn word games that require manipulations of vowels and codas rather than manipulations of onsets and vowels (Treiman, 1983). For instance, they can learn relatively quickly to split the initial C and final VC in /joz/ and reassign them to a sequence like /jɛg soz/. But they find it more difficult to

☆ Thanks to Kay Bock, Shari Speer, Becky Treiman, and an anonymous reviewer for helpful comments on this research.

E-mail address: kellymh@verizon.net.

0749-596X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2003.12.002

split the initial CV and final C to create the new sequence /jog sεz/.

In sum, onsets and rimes seem to be distinct units in syllable representations and are manipulated separately in language processing. One might press the distinction further, and examine whether onsets and rimes differ in the types of linguistic processes they can affect. In this paper, I will examine this issue in the area of English word stress, which, according to many phonologists, is affected by syllable rime structure, but not by onset structure. For instance, in Chomsky and Halle's (1968) seminal work on the sound pattern of English, stress is attracted to strong syllables, which are defined using rime structure alone without any contribution from the onset. Halle and Vergnaud (1980) express this generalization explicitly when they state "in all languages known to us, stress assignment rules are sensitive to the structure of the syllable rime, but disregard completely the character of the onset" (p. 93; see Clements & Keyser, 1983; Selkirk, 1984; Hayes, 1995 for similar positions).

Since Halle and Vergnaud (1980), some examples of stress systems that consider syllable onsets have been presented (e.g., Everett & Everett, 1984; Davis, 1988). For instance, in the Australian language Western Aranda, words with three or more syllables receive initial stress when they begin with consonants but not vowels (Davis, 1988). In Piraha, a Brazilian language, the voicing status of onset consonants appears to affect stress placement (Everett & Everett, 1984). In English, stress variation on words suffixed with "-ative" may be explained by the sonority of syllable onsets (Davis, 1988; Nanni, 1977). Whereas the first syllable in the suffix normally receives secondary stress on words like "innovative" and "qualitative," this stress is absent on some words, such as "nominative" and "generative." The latter pattern occurs when the suffix opens with a sonorant consonant (i.e., the nasals /n, m/ or glides /r, l/).

One might acknowledge the data and analyses for such cases (though for a dispute of one such analysis, see Gahl, 1996), and still maintain that, in general, onsets do not affect lexical stress. Halle and Vergnaud (1987) present this interpretation in their discussion of Piraha when they state "What is novel about Piraha is that. . .syllable onset in particular plays a determining role in the location of stress" (p. 224). Relationships between onsets and stress may be "novel" because they appear in few languages or in only small, isolated regions of any given lexicon. The latter point may apply to the case of "-ative" in English. First, this case involves a relatively small number of English words. Second, it may not even be part of the synchronic description of English. In particular, there is no evidence that the pattern is generative in the sense that stress patterns on new words in this set would be affected by onset sonorance (though Davis does provide such evidence for an onset-stress relation in Italian). Hence, it may represent only a minor exception to the general principle of onset irrelevance that applies quite well across the English lexicon.

The research presented here will challenge this conclusion by documenting a relationship between onsets and stress that permeates the English lexicon. The investigation is motivated by the West Aranda example cited above, which appears to show a link between stress and the presence of consonants in word initial position. This pattern may reflect an association between syllable weight and stress. This association is well-recognized for rime structures, and could, in some languages, generalize to onset structures. If so, then the presence and number of consonants in onset position would increase a syllable's weight, and hence its attraction for stress. Some phonologists explicitly reject such an association. In particular, Clements and Keyser state that "The internal structure of the onset and coda are usually irrelevant [for stress assignment]. Thus, for example, the syllables /ro/ and /tro/ are equivalent in the operation of English stress rules" (p. 14).

In this paper, I will present three studies that directly contradict this position. In the first study, analyses of the English lexicon will reveal that the probability of first syllable, or trochaic, stress on disyllabic words is significantly related to the number of consonants that appear in word onset position. In Study 2, I will describe an experiment showing that English speakers weigh word onset information in determining stress, indicating that the psychological processes involved in generating stress are not blind to onset structure. Finally, Study 3 will show that word onset structure affects how words are aligned with stressed positions in formal English verse, thus suggesting that onset structure could have widespread effects on English prosody.

Study 1: Analysis of the English lexicon

In this first study, I analyzed a large corpus of disyllabic English words to test the claim that word onset patterns are irrelevant to English stress assignment. An alternative possibility is that stress is drawn to syllables with high phonemic weight. This weight could be determined by a syllable's nucleus and coda structure, as traditional theories propose. If onset patterns also contribute to syllable weight, then the probability of trochaic stress on disyllabic words should be a direct function of the number of consonants in onset position.

Method

Word corpus

A corpus of disyllabic words was drawn from the MRC Psycholinguistic Database (Coltheart, 1981) at its website (www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm).The website contains a form that allows users to retrieve words from the database that match specified characteristics. I searched in particular for disyllabic words with trochaic or iambic stress. Stressed syllables are coded with a "2" in the database and unstressed syllables with a "1" or "0." Hence, words represented as "20" or "21" were considered to have trochaic stress. Words were considered to have iambic stress if they had "02" or "12" representations. 476 words were coded with a "22" stress pattern in the database, and so were not clearly identifiable as trochaic or iambic. Most of these seemed intuitively to have iambic stress, such as "sixteen" and "Chinese," but to be conservative they were excluded from the analyses. The results do not change substantially if these items are included in the iambic category.

Trochaic words were retrieved from the database if they occurred at least once in the Francis and Kucera (1982) frequency norms. The frequency constraint was not used in retrieving iambic words so that their representation in the corpus could be increased. The final corpus consisted of 6862 words, divided into 4126 trochaic and 2736 iambic. Each word in the corpus was placed into one cell of a 2×4 matrix formed by crossing the factors of word stress (iambic or trochaic) and number of consonants in word onset position (0, 1, 2, or 3). The tallies in these cells were then analyzed statistically to test the null hypothesis that lexical stress would be independent of word onset. As a stronger alternative hypothesis than mere lack of independence, I also tested whether the likelihood of trochaic stress increased with the number of consonants at word onset.

Results and discussion

The results presentation will be organized into six sections, with overall results described first. The subsequent five sections will examine the robustness and generality of the overall findings by seeing how consistently they appear in different word classes. These classes will be formed by separating words in the database according to grammatical, morphological, and phonological properties. To the extent that word onsets remain correlated with stress across these lexical variations, we have confidence that the patterns reflect general properties of the English lexicon.

Overall results

Table 1 shows the number of words that have trochaic or iambic stress as a function of word onset. The proportion of words with trochaic stress in each onset category is also displayed. Levels of trochaic stress clearly depend on word onset characteristics ($\chi^2(3) = 660.93$, $p < .0001$). Furthermore, trochaic proportions increase monotonically with the number of consonants beginning a word, and all of the critical pairwise comparisons are significant (Zero vs. One: $\chi^2(1) = 451.15$, $p < .0001$; One vs. Two: $\chi^2(1) = 76.96$, $p < .0001$; Two vs. Three: $\chi^2(1) = 4.95$, $p < .05$). This pattern is just what would be predicted if (a) onsets contribute to stress assignment and, more specifically, (b) increases in onset length attract stress.

Table 1
Distribution of disyllabic words with trochaic and iambic stress as a function of the number of consonants in word onset position

Number of onset consonants	Number trochaic	Number iambic	Proportion trochaic
0	441	806	.35
1	2862	295	.69
2	783	158	.83
3	40	1	.98

A potential confound exists in interpreting these results, however. As noted in Methods, different inclusion criteria were used for trochaic and iambic words, with the former included if they appeared at least once in the Francis and Kucera (1982) frequency counts and the latter permitted even if they did not appear. This difference served very practical purposes in the areas of sampling and statistical analysis. However, it raises the possibility that *frequency* rather than *stress* is correlated with onset patterns in English. More specifically, the analyses assume that the distribution of onset patterns among iambic words will not differ depending on presence or absence in Francis and Kucera. This assumption would be violated if, for example, iambic words with zero frequency were more likely than other iambic words to begin with zero or one consonant(s). This pattern would create the apparent association between stress and word onset shown in Table 1.

Table 2 tests this assumption directly by comparing the proportion of iambic words with different onset patterns as a function of word frequency (0 or greater than 0). The proportions are virtually identical across the two frequency classes, and provide no evidence that the overall differences between trochaic and iambic

Table 2
Proportion and raw counts of iambic words with various onset characteristics as a function of word frequency

Number of onset consonants	Word frequency			
	0 (Absent from Francis & Kucera)		Greater than 0 (Present in Francis & Kucera)	
	Prop.	Counts	Prop.	Counts
0	.350	328	.361	478
1	.580	543	.569	752
2	.070	66	.070	92
3	.001	1	.000	0

words are due to sampling characteristics. Rather, the results indicate that the patterns are quite reliable across different "slices" of the lexicon, a conclusion which will receive even more support in the following analyses.

Generality across grammatical class

The robustness of the onset effects can be evaluated further by seeing whether they appear despite variation in another factor known to be correlated with stress in English. Such tests can also eliminate the possibility that the onset effects are actually due to such other factors, with which they might be confounded.

It is well known that stress in English varies systematically with grammatical class. For instance, disyllabic nouns are more likely than disyllabic verbs to have trochaic stress (Chomsky & Halle, 1968; Kelly & Bock, 1988; Sherman, 1975). This difference is best illustrated in noun–verb homographs, such as "contest," "rebel," and "record," where the noun version has trochaic stress but the verb version has iambic stress. By examining the word onset effects across different grammatical classes, we can see whether they are a general feature of the English lexicon. We can also test whether the onset effects are an epiphenomenon, due to associations between onset structure and grammatical class. For instance, it may be the case that the correlation between word onset and stress vanishes when words are divided according to grammatical class. Perhaps the apparent correlation is due to a higher incidence of words with long onsets in the noun class. Such words would therefore tend to carry trochaic stress not because of their own phonological characteristics but rather because of their noun status.

To test these possibilities, onset effects were examined separately for nouns, verbs, and adjectives. The grammatical codes for the words were drawn from the MRC database. Table 3 presents the proportion of words that have trochaic stress as a function of grammatical class and word onset. One can see first that overall levels of trochaic stress are affected by grammatical class. Thus, trochaic stress is more common on nouns than on verbs. Nested within this grammatical class effect, however, is the persistent impact of word onset. In all three grammatical classes, the incidence of trochaic stress increases with the number of consonants that appear in word onset position. The probability that this pattern would occur by chance in all three grammatical classes is $1/24 \times 1/24 \times 1/24 = 1/13,824$.

Generality across morphemic structure

Cursory examination of the database suggests that prefixes are relatively common on words beginning with few or no consonants, such as "remind" and "inflame." If confirmed, this pattern raises the following alternative explanation for the apparent association between word onset and stress:

1. Prefix incidence in English is inversely related to the number of consonants in word onset position (i.e., prefixes are phonologically "light").
2. Disyllabic words with prefixes are more likely than words without prefixes to have iambic stress.
3. Once prefixation is taken into account, there will be no remaining association between word onset patterns and stress.

I tested this scenario by first extracting all monosyllabic prefixes from the *American Heritage Electronic Dictionary*. Words in the database were then placed into "prefix" or "no prefix" categories depending on whether or not they began with any of these prefixes. Note that the matching process was strictly based on orthography. If the first n letters in a word matched the first *n* letters of any prefix, then the word was added to the "prefix" category. This process is conservative with regard to membership in the "no prefix" category though somewhat liberal toward membership in the "prefix" category. That is, the "no prefix" category has been stripped clean of any words with potential prefixes, though the "prefix" category could contain some monomorphemic words. For instance, the "im-" prefix not only matched "impress" but also "image," though the match in the latter case is merely accidental. However, no categorization system is perfect, and the advantages of this matching procedure are that it is automatic, easily replicable, and not subject to interpretive biases that human judgments might bring to the task.

The results strongly support hypotheses 1 and 2, but not the critical hypothesis 3. First, prefix likelihood *increases* as the number of onset consonants *decreases*. In particular, the proportion of prefixed words in the

Table 3
Distribution of disyllabic words with trochaic and iambic stress as a function of grammatical class and the number of consonants in word onset position

Number of onset consonants	Number trochaic	Number iambic	Proportion trochaic
Nouns	2411	646	.79
0	274	102	.73
1	1689	475	.78
2	429	68	.86
3	19	1	.95
Verbs	648	1228	.35
0	43	485	.08
1	468	667	.41
2	129	76	.63
3	8	0	1.00
Adjectives	966	183	.84
0	107	90	.54
1	632	81	.89
2	214	12	.95
3	13	0	1.00

Table 4
Proportion of disyllabic words with trochaic stress as a function of prefixation and number of onset consonants

Number of onset consonants	Prefixation pattern					
	Prefix			No prefix		
	Number trochaic	Number iambic	Proportion trochaic	Number trochaic	Number iambic	Proportion trochaic
0	289	703	.29	152	103	.60
1	496	799	.38	2366	496	.83
2	90	71	.56	693	87	.89
3	0	0	—	40	1	.98

sample rises from 0 to .17 to .31 to .80 as the number of onset consonants drops from three to zero. All adjacent comparisons are significant at $p < .01$ or better. Second, prefixed words are much more likely than non-prefixed words to have iambic stress (64–18%, $\chi^2(1) = 1444.65$, $p < .0001$).

Most critically, however, the relationship between word onset and stress is *not* eliminated when prefixation is factored out. Indeed, as Table 4 shows, trochaic stress increases with the number of onset consonants regardless of prefixation. In the "no prefix" category, all critical comparisons are significant except that between words with two and three onset consonants, though the results in the latter case are in the predicted direction (Zero vs. One: $\chi^2(1) = 78.73$, $p < .0001$; One vs. Two: $\chi^2(1) = 16.94$, $p < .01$; Two vs. Three: $\chi^2(1) = 2.25$, $p > .05$). Moreover, similar patterns can be seen in relevant comparisons within the "prefix" category (Zero vs. One: $\chi^2(1) = 20.54$, $p < .001$; One vs. Two: $\chi^2(1) = 17.72$, $p < .001$).

Based on these results, we can reject the hypothesis that the relationship between word onset and stress in English is an artifact of prefixation patterns. Not only is the relationship preserved among non-prefixed words, but it also overlays prefixation itself, modulating the generally greater incidence of iambic stress on prefixed words.

Generality across CV patterns

Another way to examine the robustness of the onset effects is to test whether they appear consistently across different phonemic structures. Consider, for example, the words "abyss," "canal," "blossom," and "stratum." Leaving aside the presence and number of consonants in onset position, these words share the same abstract sequence of vowels and consonants: VCVC. We can therefore test whether the relationship between onsets and stress appears when post-onset CV patterns are controlled. For instance, the proportion of words with trochaic stress is .28, .56, .71, and 1.00 for VCVC words beginning with 0, 1, 2, or 3 consonants respectively. By looking at a variety of CV patterns, we can evaluate the generality of these results.

To conduct this analysis, the words in the database were coded for their CV sequence, using the MRC phonemic representations of consonants and vowels. Words with the same CV pattern, leaving aside onsets, were grouped together. Within each of these groups, such as the VCVC set described above, words were separated into four categories depending on whether they had 0, 1, 2, or 3 onset consonants. The proportion of words with trochaic stress within each of these categories was calculated, and the results are shown in Table 5 for each of the CV patterns in the sample. Note that CV patterns were only included if words in each of the four onset categories appeared in the database.

Overall, the results of this analysis match the predictions, as trochaic stress generally increases with the number of consonants in onset position. Of course, there is variability in this general pattern. Some of the variability is due to the low numbers of words in certain cells, particularly those with three onset consonants. An analysis of variance (ANOVA) was therefore conducted to evaluate the statistical reliability of the results. The ANOVA fit a repeated measures design with the single factor of onset structure. This factor had four levels corresponding with the number of

Table 5
Proportion of disyllabic words with trochaic stress as a function of post-onset cv pattern and number of onset consonants

CV Pattern	Number of onset consonants			
	0	1	2	3
VCC	1.00	1.00	1.00	1.00
VCCC	1.00	1.00	1.00	1.00
VCCCV	.41	.85	.90	1.00
VCCV	.51	.85	.94	1.00
VCCVC	.38	.66	.81	1.00
VCCVCC	.28	.57	.86	1.00
VCCVVC	1.00	.88	1.00	1.00
VCV	.59	.82	.96	1.00
VCVC	.28	.56	.71	1.00
VCVCC	.26	.42	.61	1.00
VV	1.00	.89	.96	1.00

consonants in onset position. The CV patterns were the repeated measures in the ANOVA, with each contributing values to the four levels of the onset factor. These values were the proportion of words with trochaic stress in each onset condition.

The main effect of onset was statistically significant in the ANOVA ($F(3,30) = 4.86, p < .01$). Hence, across a range of CV patterns in English disyllabic words, there is a significant relationship between word onset and stress. As with the previous analyses, the means for the onset conditions are consistent with the claim that trochaic stress is correlated with the number of consonants at word onset. All the critical pairwise comparisons between these means were significant except for that between words with two and three onset consonants (Zero vs. One: $t(10) = 2.83, p < .02$; One vs. Two: $t(10) = 4.42, p < .002$; Two vs. Three: $t(10) = 0.23, p > .40$).

Generality across onset consonants

In the next analysis, I will focus on words beginning with one or two consonants and show that the onset effects appear consistently across a variety of initial phonemes. Table 6 presents data for eleven consonants that can occur at the beginning of English words either alone (C) or as part of a consonant cluster (CC). For each consonant, I calculated the proportion of C and CC words that had trochaic stress. As Table 6 shows, trochaic proportions are generally higher for CC words. A matched-pairs t test verified that the mean trochaic proportion of .87 in the CC category is significantly higher than the .73 mean in the C category ($t(10) = 2.25, p < .05$). Hence, the onset effects are not due to one or two phonemes in particular, but rather reflect a significant property of phonemes in general that can appear as onset clusters in English.

The status of /s/

Some phonological and psycholinguistic evidence suggests that consonant clusters beginning with /s/ are atypical, and may even be illegal as syllable onsets (Kaye, Lowenstamm, & Vergnaud, 1990). For example, Treiman, Gross, and Cwikiel-Glavin (1992) asked English speakers to syllabify disyllabic pseudowords. At the boundary between the first and second syllables, the pseudowords contained consonant clusters beginning with /s/, such as the /sk/ in "fiskanch," or clusters beginning with obstruents, such as the /kl/ in "fiklanch." The subjects were much more likely to place the entire cluster into the onset position of the second syllable when it began with an obstruent other than /s/. The dominant response for /s/ clusters was to split them up, thus treating the /s/ as coda of the first syllable and the following phoneme as onset of the second. At minimum, such results indicate that /s/-initial clusters have less integrity than other clusters in English.

Treiman et al.'s experiments examined consonant sequences at syllable, though not word, boundaries. If such results extend to word onsets, then the /s/ in word initial clusters might not be part of the onset representation. As a consequence, relationships between word onset and stress might not generalize to clusters beginning with /s/. The latter phoneme would be invisible to stress assignment operations or "extrametrical," under this view. The data shown in Table 6 are not consistent with this prediction, however. Like many other word-initial consonants, /s/ is more strongly associated with trochaic stress when it begins a cluster than when it is the only consonant in word onset position. This difference, which was significant across the set of consonants in Table 6, is also significant when /s/ is considered on its own ($\chi^2(1) = 27.99$, $p < .001$).

We can look at the potential effects of /s/ extrametricality in another way. Consider a consonant /C/

Table 6
Distribution of disyllabic words with trochaic and iambic stress as a function of initial phoneme and onset structure (C or CC)

Initial phoneme	Onset structure					
	C			CC		
	Number trochaic	Number iambic	Proportion trochaic	Number trochaic	Number iambic	Proportion trochaic
/b/	231	131	.64	86	12	.88
/d/	129	241	.35	30	5	.86
/f/	189	63	.75	86	4	.96
/g/	78	23	.77	58	12	.83
/k/	255	209	.55	107	25	.81
/p/	243	79	.76	104	55	.65
/s/	298	90	.77	197	12	.94
/ʃ/	55	20	.73	4	0	1.00
/t/	156	23	.87	80	26	.76
/θ/	24	0	1.00	10	2	.83
/v/	94	21	.82	1	0	1.00

Table 7
Distribution of disyllabic words with trochaic and iambic stress as a function of phoneme in onset (C) or clusters with /s/ (sC)

Initial phoneme	Phoneme position					
	C			sC		
	Number trochaic	Number iambic	Proportion trochaic	Number trochaic	Number iambic	Proportion trochaic
/k/	255	209	.55	36	0	1.00
/l/	163	26	.86	12	0	1.00
/m/	240	68	.78	6	0	1.00
/n/	91	9	.91	8	0	1.00
/p/	243	79	.75	47	4	.92
/t/	156	23	.87	73	4	.95
/w/	129	28	.82	11	1	.92

that can appear in word onset position or follow /s/ in a two-phoneme cluster. If /s/ is not properly a part of a word's onset when it begins a cluster, then levels of trochaic stress should not differ between disyllabic words in these two classes. If, on the other hand, /s/ contributes to onset weight, and hence stress, then trochaic stress on disyllabic words should be more likely when they begin with /sC/ than /C/ alone. I tested this hypothesis in the database for all /C/s that could form word-initial clusters with a preceding /s/. The results of this analysis are shown in Table 7, and present a clear, consistent pattern. The incidence of trochaic stress is significantly higher for words beginning with /sC/ than /C/ alone ($t(6) = 3.57$, $p < .02$). Hence, /s/ does not appear to be treated extrametrically in this aspect of English stress assignment. The results also demonstrate further the generality of the onset effects since the higher levels of trochaic stress on /sC/ than /C/ onsets occurred for all seven C phonemes in the sample.

Conclusions

In sum, these statistical analyses of a large sample of disyllabic English words document a significant relationship between word stress and word onset characteristics. This relationship is both large in magnitude and broad in scope, as it permeates the English lexicon and cannot be localized to a small set of atypical words. As such, it conflicts with many theories of English phonology which do not consider onset structure as relevant to stress assignment.

It remains possible, however, that the psychological processes that assign word stress are blind to onset characteristics, as linguistic theories like Clements and Keyser's (1983) would predict. In this case, the patterns here identified in the English lexicon might reflect a variety of historical artifacts that are not properly a part of the synchronic structure of current English or the speaker's knowledge of that structure. This possibility will be examined in Study 2.

Study 2: Effects of word onsets on stress judgments

In this experiment, English speakers assigned trochaic or iambic stress to written disyllabic pseudowords. The pseudowords were created in pairs whose members were identical except for the word onset. One member of each pair began with one consonant whereas the other began with two consonants, as in "bontoon" versus "brontoon." The former pseudowords will be termed C items and the latter CC items. If English speakers associate trochaic stress more strongly with consonant cluster onsets than singleton onsets, then more trochaic stress judgments should be given to CC pseudowords like "brontoon" than to C pseudowords like "bontoon."

In addition to manipulating onset structure, I also varied the morphemic status of the first syllable in C pseudowords. For instance, the materials contained pairs like "formand–flormand," where "for" matches an English prefix. This factor was included to test further the possibility that the correlation between onsets and stress identified in Study 1 reflect morphemic influences on stress rather than onsets per se. Although analyses of the database cast doubt on this hypothesis, it is still important to test whether English speakers themselves associate stress with onset structure independently of prefixation. If so, then trochaic stress judgments in Study 2 should be more common on CC than C pseudowords regardless of the morpheme manipulation.

Method

Materials

The materials consisted of 48 pairs of pseudowords. The members of each pair were identical except that one began with two consonants (CC pseudowords) whereas the other began with one (C pseudowords). In some of the pseudoword pairs, the C member began with an English prefix like "per-" and "con-." Such syllables were considered as prefixes if they were listed as such in *The American Heritage Electronic Dictionary*. Hence, under this criterion the "bon" and "fes" at the beginnings of "bontoon" and "feslak" were not considered prefixes. The materials contained 11 pairs with a prefixed C item and 37 pairs without a prefixed C item. More of the latter pairs were included to increase the statistical power for testing the critical hypothesis that onset lengths per se can affect stress judgments. Appendix A lists the pseudoword pairs, with an asterisk marking the prefixed cases.

Two lists were constructed from these pairs. One member from each pair was randomly selected for inclusion in list 1, subject to the constraint that half of the items in the list begin with two consonants and the other half begin with one consonant. All pseudowords not selected for list 1 were included in list 2. One pseudoword pair, telpez–trelpez, was inadvertently repeated in the general list. To be conservative, this pair was deleted from the analyses, leaving 46 pseudoword pairs. Results change only marginally if telpez–trelpez is included.

Procedure

The study was conducted via the Internet. The study was advertised on University of Pennsylvania electronic bulletin boards. Interested participants could use an Internet browser to reach the author's home page, which had a link to the experiment. Once connected to the experiment, a JavaScript procedure assigned the participant randomly to list 1 or list 2.

The research was described as an investigation of English pronunciation, especially factors that determine how accent is determined on words like "zebra" and "gazelle." Participants were informed that they would read 48 invented words, each containing two syllables. For each item, they were asked to decide whether stress or accent would be placed on the first or second syllable. They could then indicate their decision by entering a "1" or a "2" in a text box. After judging a word, participants could use their computer's mouse to click on a button to receive the next word in the series. When stress judgments were made for all 48 pseudowords, participants clicked another button to send their responses to the author via electronic mail. The order of word presentation was determined randomly for each participant.

In addition to judging stress patterns on the pseudowords, participants were asked about their English background (native or non-native speaker) and their affiliation with the University of Pennsylvania (student/faculty/staff or unaffiliated). This information could be provided by clicking on appropriate radio buttons. Finally, participants were requested to enter their e-mail address in a text box. This information was used to send debriefings and to verify that each participant provided just one data set.

Participants

Twenty participants were in the experiment, with 10 assigned randomly to each list. All reported being native English speakers. Fourteen participants were from the University of Pennsylvania community, as determined from their self-report and electronic mail addresses.

Results and discussion

Each participant received two scores corresponding with the proportion of CC and C pseudowords that were assigned trochaic stress. Each pseudoword pair received two scores corresponding with the proportion of participants who assigned trochaic stress to the CC and C items in the pair. Statistical analyses were then run on these proportions, with the word onset factor varying within participants and items.

Results strongly supported the hypothesis that onset patterns affect English stress assignments. The mean proportion of trochaic assignments was .78 for CC pseudowords but only .61 for C pseudowords. This difference was highly significant across subjects ($t(19) = 5.32$, $p < .0001$) and items ($t(45) = 6.14$, $p < .0001$). Seventeen of the twenty subjects gave more trochaic judgments to CC as opposed to C pseudowords. Similarly, the CC member in 36 of the 46 pseudoword pairs received more trochaic assignments than the C member. Four pairs showed no difference, and only six exhibited the opposite pattern of results.

Table 8 displays the mean proportion trochaic judgments for pseudoword pairs with and without a prefix in the first syllable of the C items. The data show clear and significant effects of onset length on stress judgments in the non-prefixed cases (Subjects: $t(19) = 5.27$; Items: $t(36) = 6.92$; both $ps < .001$). Indeed, if anything, the onset effects here were larger than those observed in the prefixed items. Although trochaic judgments were slightly more common in the latter cases for CC than C pseudowords, the difference was not even statistically significant. (Of course, statistical power for detecting differences in the prefix condition was low given the small number of pseudoword pairs in that condition.) These results indicate that English speakers make direct links between word stress and the number of consonants in onset position.

Table 8
Mean Proportion of trochaic stress judgments in study 2 as a function of pseudoword onset (C or CC) and prefix on C pseudowords (present or absent)

Prefix	Onset structure	
	C	CC
Present	.67	.71
Absent	.60	.80

As in Study 1, we can look at results for particular subsets of items to test some further predictions of phonological theories. In particular, the potential extrametricality of /s/ at the beginnings of consonant clusters can be examined by comparing levels of trochaic stress on pseudoword pairs like "mernak" and "smernak." If /s/ is extrametrical in clusters like /sm/, then that phoneme should be invisible to stress assignment procedures. If so, "smernak" would be phonologically identical to "mernak," and levels of trochaic stress assignment should be equivalent for these and similar pseudoword pairs. However, results indicate strongly that clusters beginning with /s/ are treated like other word initial clusters. In particular, there were nine pseudoword pairs like "mernak–smernak" in the materials. The mean proportion of trochaic judgments was .81 for the sC cases but only .61 for the C cases, and this difference was statistically significant ($t(8) = 3.65,\ p < .01$). Furthermore, these means are very similar to those listed above for all the materials, indicating that clusters beginning with /s/ are no different from other clusters, at least in terms of word onset contributions to stress assignment.[1]

In sum, the results of this experiment indicate that English speakers do consider word onset characteristics when assigning lexical stress, contrary to the predictions of many phonological theories. In particular, they associate trochaic stress more strongly with disyllabic words that begin with consonant clusters rather than consonant singletons. Their behavior corresponds with statistical patterns in the English lexicon, but advances beyond Study 1 in three ways. First, the use of matched pairs like "fonjoob–flonjoob" allowed tests of onset effects with other aspects of word structure closely controlled. Second, the onset effects emerged strongly even when potential contributions from morphemic structure are removed. Third, Study 1's analysis of the lexical "environment" of English does not entail that word onset is considered by English speakers when they assign stress generatively. However, the results from this experiment indicate that the processes involved in stress assignment do weigh onset information, and hence this aspect of word structure will need to be incorporated into models of English stress assignment.

Study 3: Effects of word onset patterns on poetic rhythm

Principles of lexical phonology are often extended to phrasal phonology as a way of evaluating their scope and to situate lexical items into the larger domains of which they are typically a part. For example, both word and phrase level prosody in English are generally set to a rhythm of alternating strong and weak beats. At the lexical level, this basic pattern of rhythmic alternation can be heard in words like "constitution" and "Arizona." In phrasal contexts, lexical stress patterns are adjusted to prevent stress clashes. So, the iambic stress on "fifteen" is shifted toward trochaic stress in the phrase "fifteen books" to prevent the appearance of consecutive stressed syllables (Selkirk, 1984). These phrasal effects may feed back on the context-free pronunciation of words in that, over time, words may take on the stress patterns most suited to the phrasal environments in which they usually appear (Kelly, 1988, 1989).

In Study 3, I will take a similar approach to the relationship between word onset and stress by testing possible consequences for phrasal phonology. I will also examine the generality of onset effects by exploring possible associations with the prosodic salience of monosyllabic words. The context for the study will be English poetic rhythm, which has long proven to be a productive domain for testing the validity and scope of phonological theories, particularly in the domain of lexical and phrasal stress (e.g., Halle & Keysar, 1971; Kiparsky, 1975, 1977; various chapters in Kiparsky & Youmans, 1989).

The most important metrical structure in English poetry is iambic pentameter because of its prominence in the verse of Chaucer, Shakespeare, and Milton. The canonical form of this meter is illustrated in (1), which is quoted from the subject matter of Study 3, John Milton's *Paradise Lost*. Lines of iambic pentameter are constructed on an underlying plan of ten syllables that alternate between weak and strong beats. To achieve this meter, poets had to select and arrange appropriately a set of unstressed and stressed syllables. Hence, linguistic variables associated with stress have predictable relationships with poetic structure. For instance, closed-class words like articles and prepositions are typically unstressed in speech. Words from these classes accordingly appear most often in unstressed positions in iambic pentameter, as in "the" and "of" in (1). Monosyllabic words from open

(1) And justifie the ways of God to men (PL I.26).

Classes like nouns and verbs on the other hand are usually aligned with stressed positions. However, these patterns are far from invariant. Open class words do appear in unstressed positions, such as "deep" and "say" in (2).

(2) Nor the deep Tract of Hell, say first what cause (PL I.28).

The goal of Study 3 is to try to account for some of this variability using word onset patterns. In particular, perhaps "say" was aligned with a weak beat in (2) because it is prosodically less salient than other open class words.

[1] Affiliation with the University of Pennsylvania was not correlated with the results. The mean proportion of trochaic assignments given to CC pseudowords was .78 for both the Penn and non-Penn subject groups. The corresponding values for C pseudowords were .63 and .59.

The results from studies 1 and 2 suggest that one component of prosodic salience is word onset structure. If those results reflect general properties of English prosody, then monosyllabic words beginning with consonant clusters should attract stress more strongly than otherwise identical words that begin with single consonants. If so, and if this difference affects poetic structure, then the former words should appear more often in stressed positions in iambic pentameter than the latter. Consider, for example, the words "stay" and "say." The former appears in stressed position 78% of the time in *Paradise Lost* whereas the latter appears there only 57% of the time. In Study 3, we will see whether this pattern holds generally across a large sample of similar word pairs.

Method

A set of word pairs like "say–stay" were collected using *The Random House Rhyming Dictionary*. The words in each pair were identical, as defined below, except that one began with a consonant cluster (a CC word, though one item, "strode," began with three consonants) and the other a consonant singleton (a C word). Using the rhyming dictionary, I tried to obtain an approximately exhaustive list of these word pairs, subject to some important constraints. In particular, precautions were taken when selecting word pairs to avoid potential confounds.

First, the rimes in each pair should ideally be identical. However, English phonology has changed considerably over the centuries, creating a significant obstacle in attaining this ideal. For example, "meat" and "meet" rhyme today, but they certainly did not do so in the past. A remnant of the earlier "meat" pronunciation can be found in "great" and "steak." Hence, there are clear dangers in using a contemporary rhyming dictionary to find minimal word pairs for searching a 17th century text. To address this problem, the words in each pair were required to have orthographically identical rimes. Thus, pairs like "speak" and "peek" were excluded from the analysis. This restriction is admittedly imperfect, since current English contains word pairs that do not rhyme even though they are spelled similarly, such as "cow" and "crow." However, there is a strong correlation between orthographic and phonological identity that extends back into English history. As noted above, for example, words containing an "ee" vowel spelling did not rhyme with words containing an "ea" spelling in earlier periods of English. Still, the words within each of these sets generally DID rhyme across English history (Wyld, 1965). By restricting the analysis to rhyming pairs agreeing in orthography, the chances are increased that the words rhymed for Milton as well as for us. Of course, for this criterion to be effective, the word pairs must be spelled identically by Milton, and this assumption might not be met given orthographic changes in the intervening three centuries. To deal with this issue, an edition of *Paradise Lost* was selected that did not modernize the spelling. In particular, an electronic version of the 1667 first edition was downloaded from the University of Oregon (web address: darkwing.uoregon.edu/~rbear/lost/lost.html). Milton's original spellings were retained in this version.

A second exclusion criterion concerned grammar. In a prior analysis of Milton's meter, Kelly and Bock (1988) reported significant variation in the alignment of words with strong beats depending on grammatical class. For instance, whereas monosyllabic nouns appeared in stressed positions 94% of the time in samples of Milton's and Shakespeare's verse, verbs appeared in such positions only 76% of the time. If uncontrolled, grammatical class could seriously confound the results. Consider, for instance, the word pair "pride" and "ride," which occur in stressed positions 100 and 60% of the time, respectively in *Paradise Lost*. This pattern conforms with predictions, but it might be due entirely to grammatical class since "pride" only acts as a noun in Milton's epic whereas "ride" only acts as a verb. To deal with this problem, the words in each pair had to appear in the same grammatical role in *Paradise Lost*.

Sixty-two word pairs satisfied these constraints and appeared in *Paradise Lost*. These pairs are listed in Appendix B. Other pairs met the orthographic and grammatical constraints, but one or both members of such pairs did not appear in the poem, and so could not be used in the analysis. For each word in the final sample, I calculated the proportion of citations in *Paradise Lost* that occurred in stressed positions. In iambic pentameter, these positions correspond to the even-numbered syllables in a line.

Results and discussion

The mean proportions of CC and C words in stressed position was .94 and .85, respectively. This difference was statistically significant in a matched-pairs *t* test ($t(61) = 2.95, p < .005$), and, like the results from Studies 1 and 2, supports the hypothesis that syllable onset patterns affect English stress. As in the prior studies, there was no evidence that /s/ was excluded from onset clusters. In particular, I compared Milton's treatment of word pairs like "slow" and "low," where one member began with a single consonant, C, and the other with sC. The sC and C items appeared in stressed positions 99 and 85% of the time, respectively. Though this difference was not statistically significant given the small number of pairs in the sample, it was, if anything, larger than the difference between CC and C words in the overall analysis.

In sum, the word pairs examined in this study differ in many idiosyncratic ways that likely influenced how Milton used them in *Paradise Lost*. Yet, despite large differences in semantics and phonology that should create corresponding variability in verse placement, certain words were significantly more likely than others to appear in stressed position. The only factor that consistently distinguished these two word types was onset structure, with CC onsets more often aligned with stressed position than C onsets. This pattern supports the view that words like "play" have more prosodic weight than words like "pay."

This study goes beyond Studies 1 and 2 in four ways. First, it generalizes the onset effects to monosyllabic words. Second, it demonstrates, for one speaker at least, that onset patterns were associated with stress in earlier periods of English history. Third, the results supplement those from Study 2 by showing again that word onsets can be correlated with stress even when separated from morphemic structure.

Finally, the study joins other research in linking lexical prosody with phrasal prosody. Verse structure has often provided a rich domain for exploring such relationships, and the initial effects of word onset on verse structure presented here could be extended in many directions. For example, it has been well-documented that adjective–noun phrases like "red rose" tend to be aligned in weak-strong positions in English poetry (Kiparsky, 1975, 1977). This alignment corresponds with principles of phrasal stress in English. However, adjective–noun phrases do not invariably appear in weak-strong position, such as "proud foot" in (5) drawn from Shakespeare's *King John*. The

(5) Lie at the proud foot of a conqueror (KJ, V.viii.113)

research presented here suggests one factor—word onset—that might account for some of this variability. Given the recent widespread accessibility of electronic corpora of poetry, this factor and others can be investigated systematically and efficiently, thus opening up new ways of testing relationships between English phonology and poetic structure.

General discussion

Theories of English phonology typically assume that syllable onset patterns do not affect stress assignment. The research presented here poses significant empirical challenges to this general assumption. First, the English lexicon in fact contains verifiable associations between onset patterns and stress. The incidence of trochaic stress on disyllabic words increases directly with the number of consonants in word onset position. This pattern is both large in magnitude and broad in scope, as the various analyses in Study 1 demonstrate. Second, this association seems to be part of the phonological representation that English speakers examine when determining stress, as shown by the stress judgment results of Study 2. Finally, the analysis of *Paradise Lost* in Study 3 extends the findings to monosyllabic words and indicates that onset patterns can have wide ranging effects on English stress. In the remainder of the paper, I will discuss programmatic implications of this research, generalizations to syllables other than word onsets, and a final alternative explanation for the results.

Programmatic implications

On a programmatic level, the research demonstrates the value of conducting large scale analyses of a lexicon when evaluating hypotheses about word structure. Of course, the main barrier to this methodology in the past has been technological. However, with the widespread and growing availability of lexical databases such as that used in Study 1, the prospects for statistically representative analyses of large lexicons are improving dramatically and quickly. Other recent studies testify to the potential of this approach (e.g., Kessler & Treiman, 1997). Furthermore, as Study 3 shows, such analyses are not restricted to the lexicon, but can be extended to texts of various sorts. Researchers in language acquisition have already been making extensive use of the CHILDES database of child speech (MacWhinney & Snow, 1990), and computational linguists have been making significant advances in automated parsing using large text databases like those developed by the Linguistic Data Consortium (Marcus, Santorini, & Marcinkiewicz, 1993).

Generality beyond word onsets

This research focused on potential associations between lexical stress and the structure of word onsets. The rationale for this focus was pragmatic rather than theoretical. In particular, complex relationships between stress and syllabification in English make word onsets most convenient for analysis. Consider, for example, the noun and verb versions of "record," which, like many noun–verb homographs in English, have trochaic and iambic stress respectively (Kelly & Bock, 1988). This stress difference is associated with a syllabification difference. Where the medial /k/ is ambisyllabic in the noun version, it is clearly the onset of the second syllable in the verb version. But is this syllabification difference a *consequence* of the stress difference (as traditional phonological theory would predict) or a partial *cause* of the stress difference (as the results presented here might lead one to conjecture)?

To determine whether syllable onsets affect stress beyond word-initial position, words must be examined (as in Study 1) or invented (as in Study 2) in which syllable boundaries are unambiguous and cannot be affected by stress placement. For example, voiced and voiceless obstruents cannot appear consecutively within a syllable. Hence, phonemes in sequences like the /tg/ in "outgrow" must belong to separate syllables. In such cases, iambic stress should be more common when the second syllable begins with a consonant cluster, such as the /gr/ in "outgrow," than a single consonant, such as the /b/ in "outburst."

Unfortunately, there were only 42 such words in the corpus (suggesting a bias against such phoneme sequences even *across* syllable boundaries). The second syllable began with one or two consonants in 28 and 14 of these words, respectively. The results were in the predicted direction, though not significant given the low sample size. In particular, 57% of the words beginning with a cluster in the second syllable had iambic stress, which is almost double the 29% rate for words in which the second syllable began with a single consonant ($\chi^2(1) = 2.13$, $p < .15$). Such results provide preliminary indications that the results reported here for initial syllables can be generalized to other syllables with properly controlled materials.

A link with etymology?

On a theoretical level, one must ask why word onsets are associated with stress in English. The most radical explanation would claim that lexical stress is directly related to syllable salience or weight, as many theories suppose. However, syllable weight may be affected not only by characteristics of the rime, as many have acknowledged, but also by characteristics of the onset. The fact that trochaic stress increases monotonically with number of consonants in word onset position is consistent with this view.

Before accepting this interpretation, however, we should continue to explore alternative explanations that do not include any *intrinsic* tendency to link stress with syllables containing long onsets. Perhaps for a variety of idiosyncratic historical reasons, the English lexicon encodes this relationship, and native speakers learn it over the course of their exposure to thousands of English words. They then generalize this knowledge to new cases, such as the pseudowords used in Study 2.

I will test one version of this hypothesis that attributes the onset effects historically to the substantial mixture of Germanic and Romance words in the English vocabulary (Baugh & Cable, 1963). This mixed vocabulary could be used to account for the association between word onset and stress in the following way:

1. Disyllabic words derived from Romance languages are more likely than Germanic words to have iambic stress.
2. Such Romance words are also more likely to begin with weak onsets (e.g., zero or one consonant).
3. The association between word onset and stress only appears when Romance and Germanic words are *combined* in analyses. There will be no such associations when Romance and Germanic words are examined *separately*.

To test these hypotheses, words were randomly drawn from the overall corpus used in Study 1 until at least 250 words in each of the Romance and Germanic categories were obtained. Etymological classifications were based on the online version of *Merriam-Webster's Collegiate Dictionary* (www.m-w.com). Words were classified as Romance if they were derived from Latin, French, Italian, Spanish, or Portuguese. Words were classified as Germanic if they derived from Old/Middle English, Danish, Dutch, German, Norwegian, or Swedish (Comrie, 1987). Words were placed in an "other" category if derived from other languages or if etymological information was not listed in their dictionary entries. Overall, the analysis included 703 Romance words, 250 Germanic words, and 388 "Others." The difference in sample sizes reflects variation in the probability that a random word drawn from the corpus belonged to each etymological category. This probability was substantially lower for Germanic words (presumably because of corresponding asymmetries in the population of English disyllabic words). So, by the time that the 250 quota on Germanic words had been reached, the sample contained many more words from the Romance and "Other" categories.

Hypothesis 1 was strongly supported as 43% of Romance words had iambic stress compared with only 12% of Germanic words ($\chi^2(1) = 44.37$, $p < .0001$). However, hypothesis 2 was only weakly supported. Although Romance words tended to begin with slightly fewer consonants than Germanic words (Romance mean: 0.94

Table 9
Distribution of disyllabic words with trochaic and iambic stress as a function of etymology and the number of consonants in word onset position

Number of onset consonants	Number trochaic	Number iambic	Proportion trochaic
Germanic	203	47	.81
0	18	22	.45
1	143	24	.86
2	39	1	.98
3	3	0	1.00
Romance	403	300	.57
0	41	107	.28
1	287	167	.63
2	71	26	.73
3	4	0	1.00
Other[a]	278	110	.72
0	27	42	.39
1	180	63	.74
2	70	5	.93
3	1	0	1.00

[a] Includes words for which etymology was not listed in source dictionary.

consonants vs. Germanic mean: 1.02 consonants), this difference was only marginally significant ($t(951) = 1.91$, $p < .06$). Most importantly, as Table 9 shows, the association between word onsets and stress holds in all three lexical categories: As number of onset consonants increases, so do levels of trochaic stress. The differences between C- and V-initial words are significant for all three lexical categories (Smallest $\chi^2(1) = 27.85$ for the "Other" category). The difference between CC- and C-initial words was significant for the "Other" category ($\chi^2(1) = 11.52$, $p < .001$) and marginally significant for the Romance and Germanic categories (Romance: $\chi^2(1) = 3.24$, Germanic: $\chi^2(1) = 3.07$, both ps $< .08$). There were too few CCC words for comparison, though as predicted these words had the highest levels of trochaic stress in all three lexical categories.

In sum, this analysis disproves the hypothesis that the association between word onset and stress in English is an artifact of etymology. The association exists in the two great bands of the English vocabulary, and does not emerge only when the Romance and Germanic bands are combined. When these results are coupled with the diverse set of analyses in Study 1, we must conclude that the association between word onsets and stress pervades the English lexicon. Such findings support the hypothesis of a direct link between word onsets and lexical stress in English, contrary to what traditional phonological theory would predict.

Future research can test the cross-linguistic generality of these effects. For instance, if onset characteristics have an intrinsic tendency to affect syllable weight, and hence stress, then long onsets should be correlated with stress in many languages. Note that this view does not entail universality. Some, perhaps even most, languages might show no effects of syllable onsets in their stress systems, but few, if any, would show the opposite pattern, where stress would be inversely correlated with onset length. Such investigations will help to determine the cause(s) and generality of the onset effects found here, but their existence will need to be considered in future theories of the English stress system and the psychological mechanisms involved in implementing it.

Appendix A. Pseudoword pairs used in Study 2

beldop–breldop, bolay–brolay, botest–blotest, corlax–clorlax, covact–clovact*, dolmak–drolmak, feslak–freslak, fonjoob–flonjoob, fontrain–flontrain, formand–flormand*, garlag–glarlag, menlee–smenlee, mernak–smernak, pernew–spernew*, pinjub–plinjub, ransfoe–gransfoe, renell–drenell*, rignaz–grignaz, roncerp–troncerp, ronvoon–gronvoon, seldiz–sneldiz, telpez–trelpez*, torvoot–tworvoot, wispay–swispay, bendict–brendict, bontoon–brontoon, colvane–crolvane*, conzee–cronzee*, delpeen–drelpeen, delray–drelray, deltain–dreltain, delvoe–drelvoe, fornay–frornay*, lesbect–klesbect, pamdeen–plamdeen, peltact–pleltact, pomset–plomset, ponveen–plonveen, pelcrack–prelcrak, ponsect–pronsect, sestrow–slestrow, merset–smerset, pemit–spemit, pernor–spernor*, solray–spolray, torpez–strorpez, telmate–trelmate*.

Note. Pairs with a prefix on C member are asterisked.

Appendix B. Search pairs used in Study 3 analysis of "Paradise Lost"

band–brand, bought–brought, bow–blow, cope–scope, corn–scorn, dead–dread, dear–drear, fame–frame, fat–flat, feet–fleet, fight–fright, gain–grain, gay–gray, gaze–graze, keep–creep, lance–glance, last–blast, law–flaw, lay–play, lead–plead, led–fled, left–cleft, light–flight, lithe–blithe, loss–gloss, low–slow, pace–space, pain–plain, pay–pray, race–grace, rain–train, raise–praise, rape–grape, ray–fray, reach–breach, rest–crest, ridge–bridge, right–bright, rode–strode, rood–brood, roof–proof, rose–prose, rove–drove, rude–crude, rush–crush, sake–snake, say–stay, seed–speed, send–spend, sent–spent, serve–swerve, soil–spoil, sort–sport, sunk–slunk, talk–stalk, top–stop, tore–store, way–sway, west–quest, wine–twine, wore–swore, worn–sworn.

References

Baugh, A. C., & Cable, T. (1963). *A history of the english language*. Englewood Cliffs, NJ: Prentice-Hall.
Chomsky, N., & Halle, M. (1968). *The sound pattern of english*. New York: Harper and Row.
Clements, G. N., & Keyser, S. J. (1983). *CV phonology: A generative theory of the syllable*. Cambridge, MA: MIT Press.
Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, 33A*, 497–505.
Comrie, B. (1987). *The world's major languages*. New York: Oxford University Press.
Davis, S. (1988). Syllable onsets as a factor in stress rules. *Phonology, 5*, 1–19.
Everett, D., & Everett, K. (1984). On the relevance of syllable onsets to stress placement. *Linguistic Inquiry, 15*, 705–711.
Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton-Mifflin.
Fudge, E. (1969). Syllable. *Journal of Linguistics, 5*, 253–287.
Fudge, E. (1987). Branching structure within the syllable. *Journal of Linguistics, 23*, 359–377.
Gahl, S. (1996). Syllable onsets as a factor in stress rules: The case of Mathimathi revisited. *Phonology, 13*, 329–344.
Halle, M., & Keysar, S. J. (1971). *English stress: Its form, its growth, and its role in verse*. New York: Harper and Row.
Halle, M., & Vergnaud, J.-R. (1980). Three dimensional phonology. *Journal of Linguistic Research, 1*, 83–105.
Halle, M., & Vergnaud, J.-R. (1987). *An essay on stress*. Cambridge, MA: MIT Press.
Kaye, J., Lowenstamm, J., & Vergnaud, J.-R. (1990). Constituent structure and government in phonology. *Phonology, 7*, 193–231.
Kelly, M. H. (1988). Rhythmic alternation and lexical stress differences in English. *Cognition, 30*, 107–137.
Kelly, M. H. (1989). Rhythm and language change in English. *Journal of Memory and Language, 28*, 690–710.

Kelly, M. H. (1998). To "brunch" or to "brench": Some aspects of blend structure. *Linguistics, 36*, 579–590.

Kelly, M. H., & Bock, J. K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 389–403.

Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language, 37*, 295–311.

Kiparsky, P. (1975). Stress, syntax, and meter. *Linguistic Inquiry, 8*, 189–247.

Kiparsky, P. (1977). The rhythmic structure of English verse. *Linguistic Inquiry, 8*, 189–247.

Kiparsky, P., & Youmans, G. (1989). *Phonetics and phonology: Rhythm and meter*. New York: Academic Press.

MacWhinney, B., & Snow, C. (1990). The child language data exchange system: An update. *Journal of Child Language, 17*, 457–472.

Marcus, M. P., Santorini, B., & Marcinkiewicz, M. A. (1993). Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics, 19*, 313–330.

Nanni, D. (1977). Stressing words in -ative. *Linguistic Inquiry, 8*, 752–763.

Selkirk, E. O. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

Sherman, D. (1975). Noun-verb stress alternation: An example of lexical diffusion of sound change. *Linguistics, 159*, 43–71.

Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition, 15*, 49–74.

Treiman, R., Gross, J., & Cwikiel-Glavin, A. (1992). The syllabification of /s/ clusters in English. *Journal of Phonetics, 20*, 383–402.

Wyld, H. C. (1965). *Studies in english rhymes from surrey to pope*. New York: Russell and Russell.