852

JOURNAL OF MEMORY AND LANGUAGE 32, 115–140 (1993)

The Structure of English Syllables and Polysyllables

CAROL A. FOWLER

Dartmouth College and Haskins Laboratories

REBECCA TREIMAN

Wayne State University

AND

JENNIFER GROSS

Wayne State University

A series of experiments was carried out to explore the structure of monosyllabic, disyllabic, and trisyllabic words and nonwords. In most of the experiments, we used a phoneme shift task with visually presented stimuli to compare the speed with which hypothetical constituents could be extracted from one item and substituted into another. When constituents of the syllable and of the stimulus as a whole were confounded in monosyllables, evidence of an onset/rime or onset/remainder structure was obtained. In addition, initial clusters beginning with /s/ were less cohesive than other initial clusters. When constituents of the syllable and of the entire stimulus were unconfounded in disyllables, no influence of syllable structure was evident. Finally, when edge effects were eliminated by focusing on the middle syllables of trisyllables, effects of syllable structure emerged. These syllable-structure effects appeared with the phoneme shift task and with an unspeeded task involving auditory stimulus presentation. The results suggest that *both* word structure and syllable structure characterize spoken words.

What is the internal structure of a syllable such as *drift?* Three general possibilities may be distinguished (Vennemann, 1988). The simplest idea is that the syllable is a linear string of phonemes, /d/ followed by /r/, /ɪ/, /f/, and /t/. In this *linear* view, the syllable has no internal structure. Each phoneme is linked both to the phoneme that precedes it and the phoneme that follows it. The second possibility is a *flat* structure. Here, the syllable has three parts—an *onset*, a *peak*, and a *coda*. The onset is the initial consonant or consonant cluster, /dr/ in the example. The peak (sometimes called the nucleus) is /ɪ/; the coda is /ft/. In this flat view, /r/ is more closely linked to /d/, the other member of the onset, than to /ɪ/. However, there is no difference in the affinity of /ɪ/ with /r/ and /f/ since /ɪ/ does not belong to the same unit as either the preceding or the following phoneme. Finally, the syllable may have a *hierarchical* structure. Two types of hierarchical structure have been distinguished depending on how the basic units of onset, peak, and coda are grouped together. The first is a *rime* structure. Here, the peak and the coda are joined at a higher level to form the rime. Thus, /ɪ/ is more closely linked to /f/ than /r/ because /ɪ/ and /f/ belong to the same constituent of the syllable, the rime. A second possible

---

This research was supported in part by NICHD Grant HD01994 to Haskins Laboratories, NICHD Grants HD20276 and HD00769 to R. Treiman, and NSF Grant BNS-9020956 to R. Treiman. We thank Thomas Berg, Catherine Browman, Stuart Davis, Jonathan Kaye, Andrea Levitt, and three anonymous reviewers for their helpful comments. Reprint requests may be addressed to Carol Fowler, Haskins Laboratories, 270 Crown St., New Haven, CT 06511.

0749-596X/93 $5.00

hierarchical structure is the *body* structure. In this view, the onset and the peak form a unit called the body. In this case, /ɪ/ is more closely linked to the preceding /r/, which also belongs to the body, than to the following /f/.

According to many linguists, English syllables have a hierarchical structure of the rime type (e.g., Cairns & Feinstein, 1982; Fudge, 1969, 1987, 1989; Kiparsky, 1979; Selkirk, 1982). However, some linguists favor a hierchical structure of the body type (Iverson & Wheeler, 1989) while others favor a flat structure (e.g., Clements & Keyser, 1983). A linear structure is not currently popular in linguistics.

Behavioral evidence has been thought to provide strong support for the idea that English syllables have a hierarchical onset/rime structure. Some of this evidence involves spontaneous errors in the production of speech (e.g., MacKay, 1972; Stemberger, 1983). For example, errors called blends often join the onset of one word (e.g., the /kl/ of *close*) with the rime of another word (e.g., the /ir/ of *near*) to produce an unintended utterance (*clear*). Blends that divide words at the onset/rime boundary are more common than blends that divide words at other points.

Other behavioral evidence in support of an onset/rime structure comes from laboratory studies. In a study by Treiman (1986, Experiment 7), college students heard consonant–vowel–consonant (CVC) items. The students attempted to learn a game in which the second and third phonemes of each CVC were replaced with a fixed sequence, for example /ʌl/. If the syllable has a hierarchical rime structure, this game should be relatively easy to learn because it replaces the entire rime as a unit. On another occasion, the students learned a game in which the first and second phonemes of each CVC were replaced with /lʌ/. This game should be harder to learn according to the rime view since what is replaced is part of two different constituents. The first game was easier to master than the second, leading Treiman (1986) to favor the rime theory.

In another set of laboratory studies, Fowler (1987) adapted a procedure developed by Carter and Bradshaw (1984). College students were to exchange particular phonemes in two visually presented words and pronounce the resulting words as quickly and accurately as possible. For example, subjects might see *ton* and *pick*. In the initial-consonant exchange condition of the experiment, they said /pʌn tɪk/. In the final-consonant exchange condition, they said /tʌk pɪn/. Performance was faster and more accurate for exchanges of initial consonants than for exchanges of final consonants. Fowler (1987) interpreted these results to support a hierarchical rime structure for the English syllable. On this view, the initial consonant is easily detached from the rest of the syllable because it forms a separate constituent. The final consonant, being part of the rime, is less easily separated.

The speech error results, together with experimental findings such as those of Fowler (1987) and Treiman (1986), are consistent with the idea that English syllables have a hierarchical onset/rime structure (see Treiman, 1989, for a detailed discussion of the evidence). However, Davis (1989) recently posed an important challenge to this interpretation. As Davis pointed out, many of the words that are involved in natural speech errors and most of the stimuli that have been used in experiments contain a single syllable. What appears to be a division between the syllable onset and the syllable rime may actually be a division between the word-initial consonant and the remainder of the word. Thus, the division between /t/ and /ʌn/ of /tʌn/, which looks like a division between the syllable's onset and the syllable's rime, may actually be a division between the onset of the word and the rest of the word. Because the word contains a single syllable, it is not possible to distinguish between the rime of the first syllable and that portion of the word that consists of everything except the word onset. Thus, the behavioral results may be explained by proposing that word onsets have a special status.

Consistent with Davis' (1989) proposal, there is evidence that the first phonemes of words are special. For example, spoonerisms typically involve consonants at the beginnings of words, as in "they cut their shair hort" for "they cut their hair short." This is true whether the words contain one syllable or more than one syllable (Shattuck-Hufnagel, 1987). Browman (1978) reported similar observations on tip-of-the-tongue errors, errors in which speakers cannot retrieve a word but can provide words that are phonologically similar to the one for which they are searching. Frequently, the guesses begin with the same phoneme or phonemes as the correct word (see Brown, 1991). The guesses are less likely to share a final phoneme with the correct word and less likely still to share a medial phoneme. The shared word-initial sequences are not necessarily onsets, however. In Browman's (1978) corpus, they were as likely to be consonant–vowel (CV) units (which are not a constituent according to the rime view) as consonant–consonant (CC) clusters (which are a constituent). These results suggest that suprasegmental representations may be structured in terms of the word (or morpheme) instead of the syllable, with the first few phonemes of the word having a special status.

To distinguish between word-based structure and syllable-based structure, it is necessary to go beyond previous work by examining words that contain more than one syllable. Only then is the "remainder of the word" unit not the same as the syllable rime. If Davis (1989) is correct, a word such as *breakfast* is divided into an initial onset unit, /br/, and a remainder unit, /ɛkfəst/. There is no separate /ɛk/ unit as the rime of the first syllable. On the other hand, if structure is a syllable-based matter, /br/ (the onset of the first syllable), /ɛk/ (the rime of the first syllable), /f/ (the onset of the second syllable), and /əst/ (the rime of the second syllable) should all behave as units. It is also possible that *both* words and syllables are important. In this view, word-initial onsets have a special status but the individual syllables of polysyllabic items are made up of onset and rime units (see Berg, 1989).

To address these issues, several of the experiments reported here use tasks similar to those of Fowler (1987) and Treiman (1986) with polysyllabic stimuli. We ask whether there is evidence for onsets and rimes within the syllables of these items.

EXPERIMENT 1

To date, Fowler's (1987) task has only been used with simple CVC stimuli. Before extending the task to more complex stimuli, including those that contain more than one syllable, we must verify that the task is sensitive to the cohesive nature of cluster onsets at the beginnings of words. This was the first goal of Experiment 1. We expected that the /br/ of *brim* (a word-initial consonant cluster) would detach from its syllable more easily than the /bɪ/ of *bilk* (a word-initial consonant plus part of a rime). Relatedly, the /b/ of *brim* should be harder to detach than the /b/ of *bilk*. These predictions follow from most syllable-based theories, according to which initial consonant clusters are cohesive units. They also follow from Davis' (1989) word-based view, which grants a special status to word-initial consonant clusters.

A second goal of Experiment 1 was to test a modification of Fowler's (1987) task. Fowler's subjects exchanged two specified consonants in visually presented stimuli. Thus, subjects in the initial-consonant exchange condition responded /pʌn tɪk/ upon seeing *ton pick*. In the present studies, subjects made only one substitution. On seeing *mat brim*, for example, subjects said /bæt/ in the one-phoneme shift condition (i.e., they moved /b/ only) and /bræt/ in the two-phoneme shift condition (i.e., they moved both /b/ and /r/). By having subjects make only one response, we could use a greater variety of initial clusters.

A third goal of Experiment 1 was to examine the effect of the sonority of the postvocalic consonant on performance in the phoneme shift task. Phonemes may be arrayed along a sonority scale according to their vowel-likeness or degree of loudness. Vowels are the most sonorous type of phoneme, followed in turn by liquids, nasals, glides, and obstruents (see Clements, 1990, for discussion of sonority). In a post hoc analysis, Fowler (1987) found that more sonorous consonants took longer to shift than less sonorous consonants. This held true for both prevocalic and postvocalic consonants. This result suggests that sonorous consonants form especially cohesive units with the vowel. Derwing, Nearey, and Dow (1987) found similar effects of sonority for pre- and postvocalic consonants in word game tasks; Treiman found these effects for postvocalic consonants (Treiman, 1984) but not for prevocalic consonants (Treiman, 1986). In Experiment 1, we manipulated the sonority of the postvocalic consonant. We asked whether subjects took longer to shift the initial CV when the consonant after the vowel was a liquid or nasal than when it was an obstruent. Thus, subjects received pairs like *mat bilk* (postvocalic liquid), *mat bins* (postvocalic nasal), and *mat bids* (postvocalic obstruent). We asked whether /bɪ/ was easier to shift in *mat bids* than in the other pairs.

TABLE 1
SAMPLE SET OF STIMULI FOR EXPERIMENT 1 WITH ITS ASSOCIATED RESPONSES

Trial type	Stimulus pair	Responses R1	Responses R2
IC (initial cluster)	*mat brim*	/bæt/	/bræt/
S1 (coda sonority degree 1)	*mat bilk*	/bæt/	/bɪt/
S2 (coda sonority degree 2)	*mat bins*	/bæt/	/bɪt/
S3 (coda sonority degree 3)	*mat bids*	/bæt/	/bɪt/

Method

Stimuli. The experimental trials were derived from 24 sets of four word pairs. A sample set of word pairs together with its associated responses appears in Table 1. Each set contained four pairs of monosyllabic real words. The phonological form of the first word was always CVC, as in *mat.* The form of the second word varied across the four members of the set, as described below. Within a set, all four pairs shared the same first word. The second words of each pair within a set shared their initial consonants and vowels, but differed in other ways that realized four distinct trial types. We will label the trial types IC (initial cluster), S1, S2, and S3 (sonority of postvocalic consonant degrees 1, 2, and 3). In trial type IC, the second word of each pair began with a consonant cluster, as in *brim.* In the S trials, the second word of each pair began with a single consonant, as in *bilk, bins,* and *bids.* Across trial types S1 to S3, the second members of the word pairs differed in the sonority of their postvocalic consonants. In trial type S1 the postvocalic consonant was a liquid, as in *bilk.* In trial type S2 the postvocalic consonant was a nasal, as in *bins;* in type S3 it was an obstruent, as in *bids.* Sonority of these consonant types decreases in the series from S1 to S3. If sonority affects a consonant's cohesiveness with the preceding vowel, then splitting a syllable between the vowel and the following consonant should be progressively easier in the series S1, S2, S3.

Across the four trial types, the words were matched in average length (in phonemes: 3.9, 3.6, 3.8, 3.6 for trial types IC and S1–S3, respectively; in letters: 3.9, 4.0, 4.0, 4.1) and in average frequency (25.5, 30.9, 25.5, and 30.4 words per million from Kučera & Francis, 1967).

Each of the four trial types was associated with two response conditions. In response condition R1, subjects substituted the initial consonant of the second word of a pair for the initial consonant of the first word. In the example of Table 1, stimulus pair *mat brim* is associated with the R1 response /bæt/. For these IC trials, response condition R1 required subjects to split the second word within the syllable onset, which was predicted to be difficult. For the other trial types (S1 to S3), response condition R1 required subjects to split the second word between its onset and its rime, which was predicted to be easier. Because the four word pairs of each stimulus set had the same first word and shared the initial consonant of the second word, responses were identical across the four trial types in the R1 response condition. Responses were real words with an average frequency of 68.9 (Kučera & Francis, 1967).

In response condition R2, subjects substituted the first two phonemes of the second word for the corresponding phonemes of the first word. In the example of Table 1, responses are /bræt/ for the IC trial and /bɪt/ for trial types S1 to S3. Response condition R2 involved moving just the syllable onset on IC trials. In trials of types S1 to S3, response condition R2 required moving both the onset and the vowel. Thus, we predicted that R2 responses would be faster for trial type IC than for trial types S1 to S3. In response condition R2, as in R1, results of the transformation were always real words. Now, however, responses were the same only for trial types S1–S3 (average frequency 46.1); they were different for IC trials (average frequency 21.5). This frequency difference was necessary given other constraints on stimulus selection; in any case, it favors conditions that we predict will have long response times.

The 24 sets of four trial types each participating in two response conditions produced 192 trials. These were partitioned into eight blocks of 24 trials. Each block contained one member of each stimulus set. Across the 24 trials, each trial type appeared three times in each of the two response conditions.

A feature of the design of the experiment about which we equivocated concerned whether each subject should experience all 192 trials or just one block of 24 trials. An advantage of the first option is that each subject provides a maximum of 24 responses in each of the eight cells of the design whereas each subject provides a maximum of just three responses in each block. A disadvantage is that, across blocks, subjects get asymmetrical practice on S as compared to IC trials. That is, subjects see three times as many trials in which they move the initial consonant in a singleton-initial stimulus than those in which they move the initial consonant in a cluster-initial stimulus, and they see three times as many trials in which they move a CV in a singleton-initial stimulus than those in which they move a CC in a cluster-initial stimulus. The first asymmetry might enhance a predicted difference while the other might reduce or eliminate another.

Our solution was to present all eight blocks to subjects, but to rotate the blocks across subjects using a Latin square design. In this way, across groups of eight subjects, every stimulus appeared once in each block in each stimulus and response condition. Using this procedure, we could compare the response pattern in the first block of trials with that across the eight blocks. In case the patterns were different, we decided to restrict the analysis to just the first block of trials for each subject. As it turned out, the results were generally quite similar in pattern in the first block as across the eight. However, there were some differences; accordingly, we present here the results on the first test block only.

Procedure. Subjects were run individually. They were instructed that, on each trial, they would see a pair of words on an exposed line on the computer-terminal screen in front of them. The words would be printed in lower case except for a subset of the letters at the beginning of the second word, which would be capitalized. The task was to replace the corresponding sounds of the first word with the sounds represented by the capitalized letters of the second word. Subjects were to make the substitution and say the resulting word as quickly as possible into a microphone before them. Spoken responses triggered a voice key and stopped a millisecond clock.

Examples of possible trials were presented on file cards to give subjects some initial practice. The examples were designed to clarify the idea of exchanging sounds rather than letters and the idea of replacing "corresponding" parts of the two words of a pair. When subjects indicated that they understood the instructions, the experiment proper began with a block of 24 practice trials. The practice trials used different words than the test trials. Each of the eight conditions of the test items was represented equally often in the practice trials.

In this and later blocks of trials, stimulus pairs were centered on the top line of the terminal screen. Lines below the top line were masked with opaque paper. The experimenter sat opposite the subject watching a different monitor on which the stimulus pair as well as the correct response was printed. If the subject made an error on a practice trial, including a mispronunciation of the stimulus, the experimenter corrected the subject. No such feedback was given on test trials. On all practice and test trials, the response time in milliseconds was printed on the subject's screen after the subject responded. In addition, the average response time was printed on the screen after each block. Following the practice trials, the eight blocks of test trials were presented. The order of blocks was varied across subjects as described earlier. In addition, the trials of a block were randomized differently for each subject. For this and the other experiments, subjects' responses were recorded on cassette tape to allow verification and transcription of errors.

Subjects. Sixteen Dartmouth College students participated in exchange for course credit. All subjects in this and subsequent experiments were native speakers of English who reported normal speech and hearing.

Results

The response-time analyses in this and the following experiments were based on response times to correct responses. We eliminated response times longer than 4000 ms (less than 1% of trials in this experiment) and trials on which either the voice key failed to trigger or was triggered by a stray sound (2% of trials). To reduce any disproportionate effect of long response times, we transformed response times to speed scores (that is, the reciprocals of response times). For ease of interpretation, the tables present averages transformed from speed scores back into response times.

Where appropriate, the data were analyzed both by subjects and by items. Only those results that reached the .05 level of significance in both types of analyses will be reported. The means presented in the tables are based on the subjects analyses.

Table 2 presents the response times and accuracies in Experiment 1 for the first block of trials. Accuracy was generally high and did not differ as a function of trial type or response condition. Response times were shorter overall in condition R1, where one phoneme was shifted, that in condition R2, where two phonemes were shifted ($F_1(1,15) = 22.56$; $F_2(1,23) = 25.95$; $p < .001$ for both). Response times did not differ across the four trial types. However, the

TABLE 2

MEAN RESPONSE TIMES (MILLISECONDS) AND ACCURACIES (PERCENTAGE CORRECT) IN FIRST BLOCK OF EXPERIMENT 1

	Response time		Accuracy	
Condition	R1	R2	R1	R2
IC (initial cluster)	1642	1578	94	92
S1 (coda sonority degree 1)	1327	1762	96	94
S2 (coda sonority degree 2)	1384	1620	88	92
S3 (coda sonority degree 3)	1494	1725	92	85

interaction of trial type and response condition was significant ($F_1(3,45) = 5.46, p = .003$; $F_2(3,69) = 3.74, p = .015$).

We used Scheffé tests to test predictions of the experiment. One prediction was that on R1 trials, subjects should be slower when the responses required breaking up a syllable onset (IC) than when it did not (S1–S3). This prediction was confirmed ($F_1(7,69) = 8.95$; $F_2(7,45) = 10.86$; $p < .001$ for both). On R2 trials, we predicted that the condition in which subjects split syllables between the initial cluster and the vowel (IC) would lead to faster response times than the conditions in which subjects split words between the vowel and the following consonant (S1–S3). This prediction was also confirmed ($F_1(7,69) = 2.71, p = .015$; $F_2(7,45) = 2.83, p = .016$).

Another way to examine the interaction of response condition and trial type is to ask whether subjects are faster in condition R2 than condition R1 for IC trials. Such a difference is expected because condition R2 involves splitting the syllable between the onset and the rime, whereas condition R1 involves splitting the onset itself. Similarly, we can ask whether subjects are faster in condition R1 than in R2 for trials of types of S1 to S3. Such a difference is expected because R1 involves splitting the syllable between the onset and the rime, whereas R2 involves splitting the syllable within the rime. Although numerical differences favored the first prediction, the results did not approach significance (nor were they significant when the results for all eight blocks of the experiment were pooled). However, the second prediction was confirmed both numerically and statistically ($F_1(7,45) = 5.01$; $F_2(7,69) = 4.63$; $p < .001$ for both).

Another aim of the experiment was to ask whether the sonority of the postvocalic consonant affects its cohesion with the vowel. If more sonorous consonants are more cohesive with a vowel, then we should find progressively slower response times in response condition R2 across trial types S3 (obstruents), S2 (nasals), and S1 (liquids). The ordering of response times for these trial types (1725, 1620, and 1762 ms, respectively) do not confirm the predictions. However, S3 trials produced unexpectedly long response times in response condition R1 as well as in R2. To eliminate the effects of those overall slow times, whatever their source may be, we subtracted response times in condition R1 from those in R2 in each of the three S trial types. These difference scores enable us to ask whether subjects are slower in condition R2 as compared to R1 when the vowel must be split off from a more, as compared to a less, sonorous consonant. Difference scores revealed the predicted ordering across the three conditions (S3: 231 ms; S2: 237 ms; S1: 435 ms). However, the difference scores did not differ significantly from one another.

Error analysis. Subjects made few errors in the first block of trials. Some of the errors that they did make (stutters or substitutions of segments from outside the stimulus strings) do not provide information about the role of syllable structure in their performance. When these errors were eliminated, there were just 29 errors across the 16 subjects. We did not attempt to analyze the errors further.

Discussion

The results of Experiment 1 indicate that Fowler's (1987) phoneme exchange task, in which subjects exchange phonemes from two visually presented stimuli and produce two vocal responses, can be simplified. In the modified task, the phoneme shift task, subjects shift a phoneme or phonemes from one visually presented stimulus to another and produce one vocal response.

The phoneme shift task is sensitive to the cohesive nature of word-initial consonant clusters. In Experiment 1, subjects shifted a word-initial consonant more rapidly when it was an onset on its own (trial types S1 to

S3) than when it was part of a cluster onset (trial type IC). Conversely, subjects shifted a pair of phonemes more rapidly when the pair of phonemes was a cluster onset (trial type IC) than when the pair of phonemes was a consonant followed by a vowel (trial types S1 to S3). These results are compatible with the theory that monosyllabic English words have a hierarchical structure of the onset/rime type, or indeed with any theory in which word-initial consonant clusters behave as cohesive units.

The present results go beyond those of previous studies that have used non-speeded word games to examine the cohesive nature of word onsets (Treiman, 1983, 1986). In the word game studies, people more easily learn manipulations in which onsets behave as units than those in which they do not. In the present study, accuracy was over 90% and did not differ significantly across conditions. However, differences among the conditions were revealed in response times. Conditions in which an onset was shifted generally yielded faster responses than those in which the shift involved less than an onset or more than an onset. Together, Treiman's (1983, 1986) findings involving the learning of relatively difficult word games and the present findings involving the speed of performance of relatively easy tasks strongly support the psychological reality of the word onset.

In their numerical patterning, our findings on the sonority of the postvocalic consonant replicate those of Derwing et al. (1987), Fowler (1987), and Treiman (1984). That is, response times to shift a vowel before a liquid tended to be slower than times to shift a vowel before a nasal or an obstruent. However, variability was high in the present experiment and the differences were not significant.

EXPERIMENT 2

In Experiment 2, we looked further at initial consonant clusters. In particular, we compared those initial clusters that begin with /s/ to those initial clusters that do not. Linguists have noted some unusual features of clusters that start with /s/. First, the only three-consonant initial clusters that occur in English are composed of /s/ followed by a stop consonant, as in /spr/ and /skw/. Second, /s/-stop clusters deviate from a general pattern by which more sonorous consonants occur closer to the vowel while less sonorous consonants occur farther from the vowel. Initial clusters such as /pl/ respect this pattern; the liquid /l/ is more sonorous than the obstruent /p/. However, /s/-stop initial clusters such as /sp/ do not follow the typical sonority pattern; /p/ is not more sonorous than /s/.

The unusual features of /s/-stop clusters have led linguists to two quite different proposals. Some linguists have suggested that these apparent clusters are not really clusters at all. Instead, they are complex unitary segments that are represented orthographically as clusters (Ewen, 1982; Fudge, 1969; Selkirk, 1982). If /sp/, /st/, and /sk/ are single units, one avoids the need to postulate a third slot in the onset that only /s/ can fill. In this view, /s/-sonorant clusters such as /sl/ and /sw/ are similar to "normal" clusters such as /pl/ and /tw/. It is only /s/-stop clusters that are single units.

A second possibility is that the /s/ at the beginning of a cluster is not actually part of the onset, even when the cluster is at the beginning of a word (Kaye, Lowenstamm, & Vergnaud, 1990). Instead, /s/ is part of the rime of the preceding syllable. Kaye et al. (1990) cite several pieces of evidence from spoken Italian to support their claim, although they do not present any evidence from English. The proposal is that /s/ clusters differ structurally from clusters such as /bl/ and /tw/ whether the consonant that follows /s/ is a stop or a sonorant.

Treiman, Gross, and Cwikiel-Glavin (1992) provided behavioral evidence for the second perspective on /s/-consonant sequences in spoken English, at least when the sequence appears in the middle of a word. In several tasks, including multiple-choice selection of an appropriate syllabification for a spoken nonword and production of only the first or only the second syllable of a spoken nonword, people syllabified /s/-consonant sequences differently than sequences such as /pl/ and /tw/ in the middles of words. For example, in disyllables with final stress, people often syllabified between the consonants of /s/-consonant sequences. This occurred whether the consonant was a stop or a sonorant. In contrast, people typically syllabified before the first consonant in other clusters, thus maximizing the onset of the second syllable (see also Treiman & Zukowski, 1990).

Do /s/ clusters behave differently than non-/s/ clusters at the beginnings of words as well? Treiman (1986), studying adults' ability to learn novel word games, found that /s/ clusters as well as non-/s/ clusters were treated as units. There were no apparent differences in the cohesiveness of various initial clusters. However, Stemberger and Treiman (1986) reported that, in natural and experimentally elicited speech errors, the first consonant of a word-initial cluster was more likely to be lost when it was /s/ than when it was not /s/. This result is consistent with the idea that /s/ is less closely bound to the following consonant than is the /b/, say, of /blr/. Also, children often drop the /s/ of /s/ clusters, saying *spoon* as /bun/ or /pun/. In contrast, they typically drop the second phonemes of clusters such as /bl/, saying *blue* as /bu/ (Smith, 1973).

In Experiment 2, we looked further at clusters at the beginnings of words. Because Experiment 1 did not include enough /s/ clusters for separate analyses, Experiment 2 was designed to compare various types of /s/ clusters and non-/s/ clusters. We attempted to distinguish behaviorally the ideas that (1) /s/ clusters are like other clusters; (2) /s/-stop clusters are complex unitary segments, whereas /s/-sonorant clusters are true clusters (Ewen, 1982; Fudge, 1969; Selkirk, 1982); and (3) the /s/ of initial clusters does not belong to the syllable onset, regardless of the nature of the second element of the cluster (Kaye et al., 1990). If /s/ clusters are like other clusters, then people should find them no easier or harder to break up than other clusters. If /s/-stop clusters are complex unitary segments, they should be especially difficult to divide. Finally, if the /s/ does not form a constituent with the following consonant, it should be particularly easy to break apart the cluster.

Method

Stimuli. There were 240 test trials preceded by 30 practice trials. Each test trial consisted of two nonsense words.[1] The first nonword always had the phonological form CVC and the second always had the form consonant–consonant–vowel–consonant. Vowels were designed to be pronounced as lax. Word-initial singleton consonants were stops, fricatives, nasals, or /l/ balanced over the independent variables of the experiment, cluster type and response condition (R1 and R2). Four classes of cluster were each represented by 60 test trials: (1) /s/-stop clusters (/s/ followed by /p/, /t/, or /k/); (2) /s/-sonorant clusters (/s/ followed by /n/, /m/, /l/, or /w/); (3) fricative-sonorant clusters (/f/, /θ/, or /ʃ/ followed by /r/ and /f/ followed by /l/); (4) stop-sonorant clusters (/p/ followed by /r/ or /l/, /t/ followed by /r/ or /w/, /k/ followed by /r/ or /l/). Table 3 shows sample pairs in each category and their associated responses. Practice trials were designed in the same manner as the test items but used clusters that were not in the test list.

As in Experiment 1, either the first letter or letters corresponding to the first phoneme (condition R1) or the letters corresponding to the first pair of phonemes (condition R2) of the second member of each stimulus pair was capitalized on each trial.

---

[1] Although it would have been preferable to use real word stimuli, as in Experiment 1, this was not possible in Experiment 2 and the following experiments. There were not enough real words with the appropriate phonological structure that remained real words after the transformations required of our subjects.

TABLE 3
SAMPLE STIMULI AND RESPONSES FOR EXPERIMENT 2

Cluster type	Stimulus pair	Responses	
		R1	R2
/s/-stop	*gav spem*	/sæv/	/spæv/
/s/-sonorant	*hep snid*	/sɛp/	/snɛp/
Fricative-sonorant	*het fluk*	/fɛt/	/flɛt/
Stop-sonorant	*bup twiz*	/tʌp/	/twʌp/

TABLE 4
MEAN RESPONSE TIMES (MILLISECONDS) AND ACCURACIES (PERCENTAGE CORRECT) IN EXPERIMENT 2

	Response times		Accuracy	
Cluster type	R1	R2	R1	R2
/s/-stop	1083	1035	90	93
/s/-sonorant	1048	1016	94	93
Fricative-sonorant	1071	1000	86	92
Stop-sonorant	1089	978	94	94

Subjects' task was to shift the phonemes represented by the capitalized letter(s) from the second nonword to the corresponding position of the first nonword.

In this experiment, unlike in Experiment 1, response words differed in their initial consonants across the four types of clusters. Response times may therefore differ across cluster types because the voice key triggers later for some initial consonants than others. We did not correct for this, however, since our interest was in the within-cluster comparison of the response times for R1 and R2 shifts.

Two complementary test orders were devised so that, in one, a given nonword pair was in the R1 condition, whereas in the other it was in the R2 condition. Half of the test items in each list were in the R1 condition and half were in the R2 condition. Each subject received one test list and the lists were randomized differently across the subjects.

Procedure. The procedure was the same as in Experiment 1.

Subjects. Subjects were 24 students from the same population as in Experiment 1.

Results

The data were handled as in Experiment 1. Less than 1% of trials were eliminated either for response times exceeding 4000 ms or for failures of the voice key to trigger. Table 4 presents the descriptive statistics. As shown, responses were faster than those reported for comparable conditions in Experiment 1. One reason for this is that the data reported for Experiment 1 are for the first block of 24 trials only, whereas these data are for 240 trials. Subjects' responses in the phoneme shift task become faster with practice. Despite the overall difference in response times, the pattern of results for stimuli with initial consonant clusters is similar to that seen in Experiment 1. Specifically, responses were faster for whole-cluster shifts (R2) than for shifts involving the first consonant of the cluster (R1). The numerical differences do not support the idea that /s/-stops are particularly cohesive. Rather, they appear more compatible with the claim of Kaye et al. (1990) that /s/-consonant sequences do not form an onset constituent.

Response times and accuracies were subjected to analyses of variance with the factors of cluster type and response condition. In the analyses of response times, both main effects were significant (cluster type: $F_1(3,69) = 4.87$, $p = .004$; $F_2(3,472) = 6.81$, $p < .001$; response condition: $F_1(1,23) = 53.51$; $F_2(1,472) = 102.35$; $p < .001$ for both), as was the interaction between them ($F_1(3,69) = 18.56$; $F_2(3,472) = 11.63$; $p < .001$ for both). The interaction reflects the fact that the magnitude of the R1/R2 difference varied considerably across cluster types. Although R2 responses (shift of the whole cluster) tended to be faster than R1 responses (shift of just the first consonant of the cluster) for all cluster types, the superiority of R2 over R1 was smallest when the cluster began with /s/. In Scheffé tests, the R1/R2 difference was significant only for the fricative-sonorant and stop-sonorant clusters (fricative-sonorant clusters: $F_1(7,69) = 11.51$, $p < .001$; $F_2(7,472) = 2.39$, $p = .02$; stop-sonorant clusters: $F_1(7,69) = 25.51$; $F_2(7,472) = 14.64$, $p < .001$ for both). Tests contrasting the R2/R1 differences of the two /s/-initial groups with those of the other two groups were significant ($F_1(3,69) = 14.77$; $F_2(3,177) = 6.48$; $p < .001$ for both). Thus, there is no evidence that /s/-stop clusters are particularly cohesive, as Ewen (1982), Fudge (1969), and Selkirk (1982) proposed. The results seem more compatible with the proposal of Kaye et al. (1990) that the /s/ in a word-initial cluster is not a member of the syllable onset.

In the analyses of accuracy, only the two main effects were significant (cluster type: $F_1(3,69) = 4.49$, $p = .006$; $F_2(3,472) = 11.08$, $p < .001$; response condition: $F_1(1,23) = 4.16$, $p = .05$; $F_2(1,472) = 9.36$, $p = .0025$). The interaction was not significant in both by-subjects and by-items analyses. Overall, subjects were more accurate when moving both consonants of the cluster (R2 condition) than when moving just the first consonant (R1 condition).

Error analysis. The errors were classified into one of 32 cells according to response condition, cluster type, and error type. In the R1 condition, the most interesting errors for our purposes are those in which the whole cluster was moved (R1–CC). In the R2 condition, the complementary error type in which just C_1 was moved (R2–C_1) should be less common than R1–CC. (An error in which C_2 only was moved never occurred.) Occasionally, in both response conditions, subjects moved the designated segment or segments but also moved other segments as well. One common error, particularly in the R2 condition, was movement of CCV. This was the second error type that we scored. A third type of error included any other movement errors—that is, incorrect movements of segments from the second stimulus to the first or movements within the first stimulus. The final error category included all nonmovement errors; frequent errors in this category were substitutions of segments from outside the stimuli.

Table 5 presents the responses falling into each category as a percentage of all errors in the cell. For every cluster type but /s/-sonorant, R1–CC errors (shifts of the whole onset when subjects were supposed to shift just the first consonant) were more common than the complementary R2–C_1 errors (shift of just the first consonant when subjects were supposed to shift the whole cluster). Indeed, R2–C_1 errors, which break up syllable onsets, were hardly more frequent than the unlikely (on grounds of syllable structure) CCV movement error. An analysis of variance was performed to compare R1–CC errors and R2–C_1 errors. (The analysis was performed on numbers of errors rather than the percentages given in the table because it was impossible to compute percentages for several subjects who made no errors on a particular cluster

TABLE 5

ERRORS IN EXPERIMENT 2 EXPRESSED AS A PERCENTAGE OF ERRORS IN EACH CATEGORY AS A PERCENTAGE OF ALL ERRORS IN A CLUSTER TYPE BY R1 OR R2 CELL

	Response condition							
	R1				R2			
Cluster type	Move CC	Move CCV	Move other	Other error	Move C_1	Move CCV	Move other	Other error
/s/-stop	45	2	5	49	24	21	8	47
/s/-sonorant	29	0	23	49	34	25	2	39
Fricative-sonorant	69	0	4	27	28	13	4	54
Stop-sonorant	53	0	2	45	15	23	5	58

type.) In this analysis with the factors of response type (R1 or R2) and cluster type, both main effects and the interaction were significant (response type: $F_1(1,23) = 13.08$, $p = .0015$; cluster type: $F_1(3,69) = 6.81$, $p < .001$; interaction: $F_1(3,69) = 9.67$, $p < .001$). In general, R1–CC errors outnumbered R2–C_1 errors. However, the magnitude of the difference, and for /s/-sonorant clusters its sign, differed across cluster types. A planned comparison verified that clusters beginning with /s/ showed smaller differences than other clusters ($F_1(3,29) = 12.86$, $p < .001$). This finding suggests that initial clusters beginning with /s/ do not form a strong unit in the way that other initial clusters do.

Discussion

In Experiment 2, we compared shifts of an entire cluster to shifts of just the first consonant of the cluster. In general, whole-cluster shifts were faster and more accurate than first-phoneme-of-cluster shifts. The preference for whole-cluster shifts was confirmed in error analyses. A similar trend for faster responses to whole-cluster shifts was found for the initial cluster stimuli of Experiment 1.

The most important finding of Experiment 2 is that not all word-initial consonant clusters are equally cohesive. In particular, clusters that begin with /s/ seem to be less cohesive than clusters that do not, regardless of whether the phoneme that follows /s/ is a sonorant or an obstruent. Differences between /s/ clusters and non-/s/ clusters emerged in the analyses of reaction times. The superiority for whole-cluster shifts over first-phoneme-of-cluster shifts was significantly larger for clusters that did not begin with /s/ (91 ms) than for clusters that did begin with /s/ (40 ms). Differences between /s/ clusters and non-/s/ clusters also emerged in the error analysis. For non-/s/ clusters, subjects were much more likely to move the entire cluster when directed to move just the first phoneme, than to move just the first phoneme when directed to move the entire cluster. For /s/ clusters, the difference was smaller. However, analyses of accuracy did not show a significant interaction between cluster type and response condition.

The results of Experiment 2 do not support the idea that /s/-stop clusters are complex unitary segments whose elements are more closely bound than the elements of clusters such as /bl/ (Ewen, 1982; Fudge, 1969; Selkirk, 1982). To the contrary, the segments in /s/-stop and /s/-sonorant clusters seem to be *less* united than are the segments in clusters such as /bl/. Although /s/ clusters behave as units to some extent, it is easier to break up an /s/ cluster and shift just the first consonant than to do the same for a non-/s/ cluster. These results are compatible with findings of Stemberger and Treiman (1986) and Treiman et al. (1992), as well as with children's tendency to lose /s/ from /s/ clusters (Smith, 1973). The results may be interpreted to suggest that /s/ is not a part of the onset in the same way that the first phonemes of other syllable-initial clusters are (Kaye et al., 1990). However, /s/ clusters at the beginnings of words behave as units to some extent, consistent with the views of Kaye et al. (1990) on the interactions between constituents (Kaye, personal communication).

EXPERIMENTS 3A TO 3C

So far, we have found that whole-cluster shifts are favored when monosyllabic stimuli begin with a consonant cluster, especially a non-/s/ cluster, but that single-consonant shifts are favored when monosyllabic stimuli begin with a single consonant. These results may reflect the structure of the English syllable. However, because the stimuli have been composed only of monosyllables, we cannot know that syllable structure, as compared with word structure, is the important variable. As discussed in the introductory remarks, the literature suggests that word onsets have a special status.

Experiments 3A to 3C used disyllabic

stimuli to ask whether the evidence for cohesive onsets obtained in Experiments 1 and 2 reflects the structure of the syllable or the structure of the entire stimulus. The stimuli for Experiments 3A to 3C were pairs of nonsense words with the structure $C_1VC_2C_3VC_4$. Phonotactic constraints were used in an effort to promote syllabification between the two medial consonants. Thus, C_2 and C_3 were consonants such as /t/ and /k/ that cannot serve as clusters within a syllable. In one experiment (3A), subjects replaced either C_2, a coda consonant, or C_3, an onset consonant, in the first disyllable of a pair with the corresponding consonant from the second disyllable. If syllable structure determines the cohesiveness of a consonant with its neighbors, then C_3 should be easier to replace than C_2. In a second experiment (3B), subjects replaced one or the other onset consonant (C_1 or C_3). If *only* syllable structure determines the cohesiveness of a segment with its context, then these consonants should be equally easy to replace. However, if the onsets of words or nonwords have a special status, then C_1 should be easier than C_3. In the third experiment of the series (3C), subjects replaced a pair of segments: the initial or final rime of the disyllable or the initial or medial CV. If syllable structure determines cohesiveness, then rimes should be easier to replace than CVs. If syllable structure *alone* determines cohesiveness, with no effects of word structure, then initial and final rimes should be equally easy to replace.

The procedure in Experiments 3A to 3C was the same as that in Experiments 1 and 2. The data were handled as in Experiments 1 and 2. On average, less than 1% of responses were eliminated because they exceeded 4000 ms. However, 4.5% of responses were eliminated because the voice key did not trigger on the subject's response. This included failures of the voice key to trigger at all, triggering on an extraneous sound, such as "uh," and trials in which the subjects paused between the syllables of the disyllabic response.

EXPERIMENT 3A

Method

Stimuli. Stimuli consisted of two sets of 50 nonsense disyllable pairs composed from a sampling of five vowel spellings (*a, e, i, o, u* pronounced as lax vowels) and 18 single-letter consonant spellings. An example of a pair is *mupnav leFbok*, pronounced as /mʌpnæv lɛfbɑk/. In one transformation, subjects shifted C_2, as indicated by the capitalized letter in the example, to produce /mʌfnæv/. In another transformation, C_3 was capitalized (*mupnav lefBok*) and subjects shifted it to produce /mʌpbæv/. A second set of items, constructed from the first, switched the order of the two medial consonants of each disyllable. Across lists, then, the same consonants appeared as C_2 and C_3. All responses were nonwords. As far as possible, the individual syllables of the stimuli and responses were nonwords too.

Two lists of 100 items were created from the two base sets of 50. In each list, 50 items had C_2 of the second disyllable of a pair capitalized; the other 50 items had C_3 capitalized. In the second list, the assignment of capitalized C_2 and C_3 was reversed.

Test trials were preceded by 20 practice trials consisting of disyllables similar to those used in the test trials. Ten of the trials required that C_2 be shifted and 10 required that C_3 be shifted.

Half of the subjects were assigned to each list. Of the subjects assigned to each list, half pronounced the disyllables with first syllable stress; the other half pronounced them with second-syllable stress. Examples were pronounced by the experimenter and then by the subject to illustrate the proper stress pattern.

Procedure. The procedure was identical to that of Experiment 1.

Subjects. The subjects were 32 students from the same population as Experiment 1.

Results

Table 6 presents response times and accuracies. Numerical differences do not sup-

TABLE 6
MEAN RESPONSE TIMES (MILLISECONDS) AND ACCURACIES (PERCENTAGE CORRECT) IN EXPERIMENT 3A

	Response times		Accuracy	
Stress pattern	Coda (C_2)	Onset (C_3)	Coda (C_2)	Onset (C_3)
First-syllable stress	1885	1918	85	78
Second-syllable stress	2003	2085	85	86

port the hypothesis that cohesion of word-internal consonants with their context reflects syllable structure. The onset consonant (C_3) was not shifted faster than the coda consonant (C_2).

In analyses of response times with the factors of consonant type (coda or onset) and word stress (first or second syllable), no effect was significant in both items and subjects analyses. In the analysis of accuracy, only the interaction was significant ($F_1(1,30) = 7.83$, $p = .009$; $F_2(1,49) = 9.10$, $p = .004$). The interaction reflects the crossover in accuracy depending on syllable stress. Subjects were more accurate when shifting a phoneme from a stressed syllable than one from an unstressed syllable.

Error analysis. We looked first at errors in which vowels were moved with consonants to ask whether whole rime errors were more common than CV movement errors. They were not: There were 11 VC errors and 11 CV errors. Thus, it does not appear that vowels cohere more with following than with preceding consonants, as the rime theory of syllable structure predicts. We also looked at errors in which a consonant was anticipated into an earlier slot or perseverated into a later slot in the response. In particular, we looked at anticipations into the C_1 and C_2 slots. For example, the response /fʌpnæv/ to *mupnav leFbok* would count as an anticipation into C_1. In these locations, the anticipated consonant can come from a compatible (C_3 for C_1 and C_4 for C_2) position or an incompatible (C_2, C_4 for C_1 and C_3 for C_2) position in the second stimulus. There were 3 anticipations into C_1 from C_3, the compatible location, and an average of 4.5 anticipations from C_2 or C_4. As for C_2, compatible and incompatible anticipations numbered 7 from C_4 and 18 from C_3. Analogously, we looked at perseveration errors into the C_3 and C_4 slots of the response. Perseverations into C_3 numbered 3 from the compatible location, C_1, and 32 from the incompatible location, C_2. The corresponding numbers for perseverations into C_4 were 12 and 11. In general, then, movements into compatible syllable locations were no more nor less common than movements into incompatible locations. The exception was movement into the C_3 slot by C_2 consonants and movement into the C_3 slot by C_2 consonants, both reflecting movement into an *incompatible* location. This result may reflect the fact that, in this experiment, C_2 and C_3 were the consonants being moved. Because these consonants are adjacent, subjects may have gotten confused on some trials.

EXPERIMENT 3B

Method

Stimuli. Stimuli were pairs of nonsense disyllables constructed from the same pool of consonants and vowels as in Experiment 3A. As in the earlier experiment, all stimuli and responses were nonwords whose component monosyllables were, insofar as possible, nonwords as well. A sample stimulus pair is *bepniz Kugfam,* pronounced as /bɛpnɪz kʌgfæm/. Here, the response was /kɛpnɪz/. In the other list, the second onset consonant of the word was capitalized (*bepniz kugFam*) and the response was /bɛpfɪz/.

In this set of test materials, unlike in Experiment 3A, response items in different conditions (C_1, C_3) had different initial consonants. To guard against spurious response-time differences across conditions due to systematic differences in voice-key triggering latencies, we matched response items approximately for initial consonant across C_1 and C_3 conditions. In the C_1 condition, there were 44 responses that began with stop consonants, 15 with nasals, 33 with fricatives, and 8 with nasals, liquids, or semivowels in the C_1 condition. The corresponding frequencies in the C_3 condition were 45, 16, 33, and 6. Test trials were preceded by 20 practice trials.

Half of the subjects were assigned to each stimulus list. Half of the subjects on each list pronounced the disyllables with first syllable stress; the other half pronounced them with second-syllable stress. Examples were pronounced by the experimenter and then by the subject to illustrate the proper stress pattern.

Procedure. The procedure was identical to that of Experiment 1.

Subjects. The subjects were 32 students from the same population as Experiment 1.

Results

Table 7 presents response times and accuracies. Response times were faster and accuracies higher for the first than for the second onset consonant. Response times to the second onset consonant were also faster than they were in Experiment 3A. This may reflect the fact that consonants to be moved in Experiment 3A were adjacent, making discrimination between transformation tasks more difficult than in the present experiment. Alternatively, the difference may reflect sampling differences among subjects or items.

TABLE 7
MEAN RESPONSE TIMES (MILLISECONDS) AND ACCURACIES (PERCENTAGE CORRECT) IN EXPERIMENT 3B

	Response times		Accuracy	
Stress pattern	Initial onset (C_1)	Medial onset (C_3)	Initial onset (C_1)	Medial onset (C_3)
First-syllable	1527	1803	91	83
Second-syllable	1613	1895	91	80

Analyses of variance on response times with the factors of onset consonant location (initial or medial) and stress (first or second syllable) showed only a main effect of onset location ($F_1(1,30) = 124.73$; $F_2(1,49) = 384.24$; $p < .001$ for both). Initial onsets moved more rapidly than medial onsets. In the analyses of accuracy, too, only the effect of onset location was significant ($F_1(1,30) = 33.51$; $F_2(1,49) = 45.89$; $p < .001$ for both). Performance was better on initial onsets than on medial onsets.

Because position in the syllable was held constant in this experiment, the large response time and accuracy differences do not reflect this variable. Rather, the findings suggest that the position of a consonant in the stimulus, initial or non-initial, affects its cohesion with the neighboring phonemes.

Error analysis. As in Experiment 3A, we looked at anticipation errors into C_1 and C_2 locations and perseverations into C_3 and C_4. An analysis of variance was performed with the factors of direction of movement (anticipation or perseveration), syllable position of the slot into which a consonant moved (onset or coda), preservation or not of the intended syllable position of the moved consonant, and word stress. Only the main effect of slot position and its interaction with direction of movement were significant (main effect: $F_1(1,30) = 8.11$, $p = .008$; interaction: $F_1(1,30) = 28.65$, $p < .001$). The interaction was significant because anticipation errors into an onset position were more frequent than into the coda position (26 errors versus 1 error, $F_1(3,30) = 21.17$, $p < .001$), whereas perseveration errors into the onset position (C_3) were marginally less frequent than into the coda position (4 errors vs. 13 errors, $F_1(3,30) = 2.49$, $p = .08$). In any case, the syllable position of the moved segment was not preserved. Onsets moved into coda positions as frequently as into onset positions and vice versa. At most, the pattern seems to suggest an attraction of consonants to word edges.

EXPERIMENT 3C

Method

Stimuli. In this experiment, sequences of two segments were moved. For half of the subjects, the sequences were the CV or VC (rime) in the first syllable of the second disyllable of a pair. For the other subjects, the sequences were the CV or rime of the second syllable. Disyllables were constructed from the same pool of vowels and consonants and using the same constraints as in Experiments 3A and 3B. As in Experiment 3B, we matched response items for initial consonant across the CV and VC conditions to avoid response-time differences due to voice-key latency differences. In the VC conditions, there were 38 initial stops, 28 fricatives, 13 nasals, and 21 nasals, liquids, or semivowels; in the CV conditions, the corresponding frequencies were 39, 28, 10, and 23.

Two matched lists of 50 stimulus pairs were constructed, one for presentation to subjects shifting phoneme sequences in the first syllables of the disyllables and one for subjects shifting phoneme sequences in the second syllables. The lists were matched so that a pair in the first list consisting of the syllables S_1S_2 S_3S_4 appeared in the second list with the ordering S_2S_1 S_4S_3. In this way, across conditions, subjects replaced the same CV and VC sequences with the same substituting CVs and VCs; only the position of the affected syllable in the disyllable differed across lists. A sample stimulus pair is *vapkem bufgon,* pronounced as /væpkɛm bʌfgɑn/. Either *bu* or *uf* was capitalized so that /bʌpkɛm/ or /vʌfkɛm/ was the response. In the other list, the disyllable pair appeared as *kemvap gonbuf,* again with either *bu* or *uf* capitalized so that /kɛmbʌp/ or /kɛmvʌf/ were the responses.

Each subject received each stimulus pair twice, once with the CV capitalized and once with the VC capitalized. Trials were randomized separately for each subject and were preceded by 20 practice trials. As in the earlier experiments, half of the subjects receiving each list pronounced the words with first-syllable stress and the other half pronounced them with second-syllable stress.

Procedure. The procedure was identical to that of Experiment 1.

Subjects. The subjects were 32 students from the same population as Experiment 1.

Results

Table 8 presents the response times and accuracies. Consistent with the findings of Experiment 3A, response times and accuracies do not pattern as expected if the cohesion among segments in the disyllables reflects an onset/rime syllable structure.

In analyses of response times, the main effect of syllable (first or second) was significant ($F_1(1,38) = 14.04$; $F_2(1,49) = 579.79$; $p < .001$ for both). Responses were faster when the phonemes to be shifted were in the first syllable than in the second syllable. There was also a main effect of sequence type (CV or rime) ($F_1(1,28) = 7.18$, $p = .01$; $F_2(1,49) = 51.77$, $p < .001$) and an interaction between syllable and sequence type ($F_1(1,28) = 13.55$, $p = .0011$; $F_2(1,49) = 48.30$, $p < .001$). The interaction was significant because, whereas CVs were moved more rapidly than syllable rimes in

TABLE 8

MEAN RESPONSE TIMES (MILLISECONDS) AND ACCURACIES (PERCENTAGE CORRECT) IN EXPERIMENT 3C

	Response times			
	First syllable		Second syllable	
Stress pattern	CV	VC	CV	VC
First-syllable	1535	1777	2140	2053
Second-syllable	1494	1660	2112	2135
Means	1515	1719	2126	2094
	Accuracy			
	First syllable		Second syllable	
	CV	VC	CV	VC
First-syllable	93	93	85	88
Second-syllable	93	91	83	88
Means	93	92	84	88

the first syllable of a stimulus (Scheffé: $F_1(3,28) = 13.49$; $F_2(3,49) = 40.71$; $p < .001$ for both), there was no significant difference in movement times between CVs and VCs in the second syllable (both *F*s < 1).

In the analyses of accuracy, only the effect of syllable was significant ($F_1(1,28) = 5.81$, $p = .02$; $F_2(1,49) = 5.21$, $p = .025$). Accuracy was higher for sequences in the first syllable than for sequences in the second syllable.

Overall, these results provide no evidence that the syllable constituents postulated by the onset/rime theory contribute to patterns of cohesion among consonants and vowels in disyllabic nonwords. Response times were not faster and accuracies were not higher to VCs, which are thought to be syllable constituents, than to CVs, which are not constituents.

Error analysis. We looked at four categories of error that should distinguish the CV and VC conditions if syllable structure is reflected in intersegment cohesion. Subjects sometimes made errors in which they moved only the capitalized C, leaving the V behind. Similarly, they sometimes moved only the capitalized V, stranding the C. We predicted that these errors should be more common in the CV shift condition because the C and V should be less cohesive in that condition than in the VC condition. Another error that occurred occasionally was movement of the whole CVC rather than just the capitalized CV or VC. Such errors should be more common in the CV condition, because the V is expected to be more cohesive with the following than the preceding consonant in a syllable. Accordingly, a to-be-moved CV should attract the following C, but a to-be-moved VC should be less likely to attract a preceding C. Finally, people sometimes moved CV when instructed to move VC or vice versa. If members of rimes are more cohesive than members of CVs, these errors should be more common in the CV condition.

Because there were relatively few errors in each category and because the prediction was the same for all error types, these errors were pooled within the CV and VC conditions. Overall, 24.2% of all VC errors as compared to 30.9% of all CV errors fell into the categories outlined above, a small difference in the predicted direction. Moreover, the predicted direction of difference held up numerically in all four error categories. However, the effect fell well short of significance in an analysis of variance with the factors of position of the moved CV or VC (first or second syllable), stress pattern (first- or second-syllable stress) and moved segments (CV or VC). In that analysis, the only significant effect was that of the position of the moved CV or VC ($F_1(1,28) = 9.66$, $p = .004$). There were more errors falling into the indicated categories in the second than in the first syllable. Thus, the error analysis is consistent with the analyses of response times and accuracies in providing no evidence favoring onsets and rimes as particularly cohesive units in the disyllables of our experiment.

Discussion of Experiments 3A–3C

The results of Experiments 3A to 3C fail to provide evidence that the cohesiveness of segments in disyllables reflects an onset/rime syllable structure. Rather, segments at the beginning of a stimulus, including onsets (Experiment 3B) and CVs (Experiment 3C), are more detachable from the remainder of the stimulus than are segments elsewhere. A major factor affecting performance on disyllables in the phoneme shift task seems to be whether the phoneme or phonemes being moved is at the beginning of the stimulus. If so, performance will be faster and better than if not. This advantage for earlier occurring phonemes may reflect linguistic factors, specifically the structure of the disyllable as a whole. Phonemes earlier in the stimulus may be more detachable from the remainder of the word than are phonemes later in the stimulus. This interpretation would be largely consistent with the proposal of Davis (1989), according to which the word-initial onset forms a separate unit.

Alternatively, or in addition, the advantage for earlier occurring phonemes may reflect the nature of our task. The left-to-right scanning that was necessitated by the visual presentation of the stimuli may have caused segments earlier in a stimulus to be shifted more rapidly than segments later in a stimulus. If this is the only determinant of performance, the results of Experiments 3A to 3C do not speak to the linguistic structure of the stimuli.

Supporting the idea that our results are at least partly due to linguistic factors, the findings of Experiments 3A to 3C are generally consistent with findings discussed in the Introduction on speech errors and tip-of-the-tongue errors. As in those studies, word-initial segments are least cohesive with the rest of the word while medial segments are most cohesive. One finding that is not consistent with findings in those domains is one obtained in Experiment 3B. In both speech errors (Shattuck-Hufnagel, 1987) and tip-of-the-tongue errors (Browman, 1978), the onsets of stressed noninitial syllables (e.g., the /l/ of *alone*) seem to participate in errors as if they were distinct units, behaving like word onsets in this respect. Onsets of unstressed noninitial syllables (e.g., the /l/ of *only*) show no particular tendency to be involved in speech errors or to be preserved in tip-of-the-tongue guesses. In Experiment 3B, however, response times to move the onset of the second syllable were no faster when that syllable was stressed than when it was not stressed.

Although the findings of Experiments 3A to 3C do not support syllable-based structure, it may be premature to conclude that the position of a phoneme in its syllable is unimportant. In Experiments 3B and 3C, phonemes in the first syllable of a disyllable were moved faster and more accurately than phonemes in the second syllable. This benefit for earlier occurring phonemes may have countered a disadvantage for codas, leading to the comparable performance for first-syllable codas and second-syllable onsets that was observed in Experiment 3A. To see clear effects of syllable-internal structure, it may be necessary to look at the middle syllables of stimuli that contain three or more syllables. Because these syllables are not at the beginning or the end of the stimulus, effects of syllable-internal structure may be visible.

In two final experiments, therefore, we examined the middle syllables of trisyllabic, medially stressed, nonsense words. One experiment of the pair (Experiment 4) used a phoneme shift task with visual presentation of the stimuli similar to that used in the foregoing experiments. The second experiment (Experiment 5) used a word game task with auditory presentation of the stimuli, similar to that used by Treiman (1986, Experiment 7). If different results are found in the two experiments, we could suggest that the left-to-right scanning that occurs with visual presentation of the stimuli has a major impact on the findings. In Experiment 4, we compared shifts of the onset of the middle syllable to shifts of the coda and shifts of the initial CV to shifts of the final VC. If the onsets of the middle syllables are detachable, then movements of the onset should be faster and easier than movements of the coda. In addition, the VCs of the middle syllables should be moved more readily than the CVs.

EXPERIMENT 4

Method

Stimuli. Stimuli were two matched lists of 30 trisyllabic nonsense words. All trisyllables had the structure $C_1VC_2VC_3C_4VC_5$ with medial stress. The medial stress attracts C_2 to the second syllable; C_3 and C_4 were selected so that phonotactic constraints promoted location of the syllable boundary between them. Therefore, the medial syllables were CVCs. A sample pair is *rupadkin yomefbug*. In different conditions, C_2, C_3, C_2V, or VC_3 of the second nonword was capitalized and subjects were instructed to shift those capitalized sounds from the second to the first nonword. Across the four conditions, the correct responses, then, were /rʌmædkɪn/, /rʌpæfkɪn/, /rʌmɛdkɪn/, and /rʌpɛfkɪn/. The trisyllable pairs in the two lists were identical except that a C_2VC_3 sequence in one list became a C_3VC_2 sequence in the second, so that the same consonants were shifted and replaced in all conditions.

Subjects received all 60 trisyllable pairs. Half of the subjects moved 30 CV and 30 rime sequences; the other subjects moved 30 C_2 and 30 C_3 consonants. Across subjects, each trisyllable pair participated in all four response conditions. Test trials were preceded by 15 practice trials designed after the test trials.

Procedure. The procedure was the same as that used in the previous experiments.

Subjects. Subjects were 24 students from the same population as Experiments 1–3.

Results

Response times were eliminated if they exceeded 6000 ms (<1% of all responses) or if the voice key failed to trigger to the response (2.6% of all responses). Response times and accuracies are presented in Table 9. The response times are longer than those observed in the preceding experiments, presumably because the stimuli are longer. More importantly, the pattern of response times is as expected if syllable constituents are reflected in patterns of segment cohesion. That is, response times to the two syllable constituents (C_2, VC) are faster than those to the two nonconstituents (CV, C_3).

TABLE 9

MEAN RESPONSE TIMES (MILLISECONDS) AND ACCURACIES (PERCENTAGE CORRECT) IN EXPERIMENT 4

Trial type	Response time	Accuracy
C_2	2541	83
C_2V	2764	84
C_3	2716	81
VC_3	2243	76

In analyses of variance with the factors of location of segment moved (beginning or end of the syllable) and number of phonemes moved (one or two), the effect of location was significant ($F_1(1,22) = 7.67$, $p = .01$; $F_2(1,59) = 5.80$, $p = .018$), as was its interaction with number of phonemes moved ($F_1(1,22) = 59.6$; $F_2(1,59) = 74.34$; $p < .001$ for both). Planned comparisons revealed that syllable constituents were moved faster than nonconstituents. That is, C_2 was moved faster than C_3 ($F_1(1,22) = 7.57$, $p = .011$; $F_2(1,59) = 31.75$, $p < .001$) and VC_3 was moved faster than C_2V ($F_1(1,22) = 66.68$; $F_2(1,59) = 43.07$, $p < .001$ for both). Analyses of errors yielded only a main effect of location, with segments at the beginning of the syllable moved move accurately than segments at the end of the syllable ($F_1(1,22) = 4.56$, $p = .04$; $F_2(1,59) = 4.82$, $p = .03$).

Error analyses. We asked whether, with one phoneme to be moved, subjects were more likely to move a vowel with C_3, the coda, than with C_2, the onset. Subjects rarely made errors in these conditions, and the errors that did occur did not differ between the conditions. We also asked whether, with two phonemes to be moved, the vowel was erroneously left behind more often in the C_2V condition (i.e., just the onset was moved) and the whole syllable was more likely to be moved as a unit in the C_2V condition (i.e., the integrity of the rime was preserved). Pooled errors of those sorts differed numerically as expected (27.3% of errors in the C_2V conditions were either V strandings or whole-syllable movements; 23.5% of VC_3 errors were of either sort); however, the difference did not approach significance.

Discussion

In the reaction times of Experiment 4, we found evidence for syllable-internal structure in the middle syllables of medially stressed trisyllables—syllables that are not at the edge of the stimulus. For these syllables, it was faster to shift an initial consonant than to shift a final consonant. This result is consistent with the idea that syllables are made up of onsets and rimes. The onset, being a separate constituent of the syllable, is relatively easy to detach from the rest of the syllable. The coda, being part of the rime, is harder to detach. In addition, it was faster to shift a VC than a CV. This result, too, is compatible with the idea that syllables are made up of onsets and rimes. The VC, or rime, is a constituent of the syllable; the CV is not. If the *only* structure within a polysyllable reflected the structure of the stimulus as a whole, with the polysyllable having an initial onset unit and a remainder unit (Davis, 1989), these findings would not be expected. Rather, the results suggest that, in addition to word-based structure, polysyllables also exhibit an onset/rime structure.

EXPERIMENT 5

In the final experiment, we tested the generality of the findings of Experiment 4 by using a phoneme substitution game, as in Treiman (1986, Experiment 7). The word game procedure avoids the orthographic presentation of the previous experiments that may, particularly with the long stimuli of Experiment 4, introduce considerable reading and scanning times into the response latencies.

As in Experiment 4, the stimuli were trisyllabic nonwords with the structure $C_1VC_2VC_3C_4VC_5$ and medial stress. In this experiment, presentation was auditory. Subjects heard just one trisyllable, rather than the two they saw in Experiment 4. They attempted to learn a game that replaced one or two phoneme(s) in the middle syllable of the stimulus with one or two other phonemes that remained fixed throughout the experiment. Half of the subjects participated in the lax-vowel condition of the experiment. For these subjects, all stimuli had lax vowels in their middle syllables. Half of the subjects in the lax-vowel condition were assigned to the one-phoneme substitution condition. In different sessions, they were required to learn the rule C_2 goes to /g/ and the rule C_3 goes to /g/. For example, /ʃəpædnəð/ changed to /ʃəgædnəð/ under the first rule and to /ʃəpægnəð/ under the second. If the findings using this procedure replicate those of Experiment 4, then the first transformation should be easier than the second. The other subjects in the lax-vowel condition were assigned to the two-phoneme substitution condition. They learned the rules that C_2V goes to /gɛ/ and that VC_3 goes to /ɛg/. If the findings replicate those of Experiment 4, the second rule should be easier than the first.

The other subjects participated in the tense-vowel condition, receiving stimuli with tense vowels in their middle syllables. Half of these subjects learned the one-phoneme rules C_2 goes to /g/ and C_3 goes to /g/. The other half of the subjects learned the two-phoneme rules C_2V goes to /ge/ and VC_3 goes to /eg/. A comparison of the results for the lax-vowel and tense-vowel conditions should help to show whether the cohesiveness of phonemes within a syllable depends on the nature of the vowel. In some views (e.g., Clements & Keyser, 1983), tense vowels occupy two "slots" in the syllable structure. If so, they may be less closely bound to a following consonant. Results of word game studies by Derwing and Nearey (1990) provide some support for this claim.

Method

Stimuli. For the lax-vowel condition, 20 nonsense trisyllables with medial stress were devised using similar constraints as in Experiment 4. The nonwords all shared the first two phonemes (/ʃə/) and the last two phonemes (/əð/) because pilot work had indicated that this made the stimuli easier to remember. None of the nonwords contained English prefixes or suffixes. For the tense-vowel condition, the lax vowels in the middle syllables of these stimuli were replaced with tense vowels.

Procedure. Half of the subjects were assigned to the lax-vowel condition and half to the tense-vowel condition. Within each group, half of the subjects participated in the one-phoneme substitution condition and half in the two-phoneme substitution condition. Each subject participated in two sessions and learned one of the rules in each session. The subjects were informed that they would learn a word game involving nonsense words. They were told that their task was to transform each nonword into a new one by changing it according to a rule. They were instructed that all the nonwords were to be transformed according to a rule that they were to try to discover based on the examples that they would hear. On the first trial, the experimenter pronounced one of the trisyllables, randomly chosen from the list of 20, and gave the appropriate transformation as a response. For all subsequent trials, the experimenter pronounced the stimulus twice and the subject was required to repeat it twice correctly. After correct repetition, the subject was asked to respond by applying the rule of the game. The experimenter provided the correct answer if the subject responded incorrectly. We scored the first response that subjects gave; if the subject responded incorrectly and then self-corrected, their response was counted as incorrect.

At least one week elapsed between sessions. The order of rule learning was counterbalanced across subjects in a condition, and the order of stimuli was randomly chosen for each subject.

Subjects. Subjects were 80 students at Wayne State University who participated in exchange for course credit or pay.

Results

Table 10 shows the results in terms of several measures of performance. Number correct is the total number of correct responses on test trials. First correct trial is the earliest test trial on which the appropriate response was given. Longest run is the longest string of consecutive correct responses.

Analyses of variance with the factors of number of phonemes changed (one or two), location of change (beginning or end of the syllable), and type of vowel (lax or tense) were performed for each of the three nonorthogonal dependent measures. For the number of correct responses, the only significant effect was an interaction between the number of phonemes and the location of change ($F_1(1,76) = 35.91$; $F_2(1,19) = 88.38$; $p < .001$ for both). Planned comparisons revealed that replacement of the onset, C_2, was easier than replacement of part of the rime, C_3 ($F_1(1,38) = 9.78$, $p = .003$; $F_2(1,19) = 16.39$, $p = .001$). Compatibly, replacement of the entire rime, VC_3, was easier than replacement of C_2V ($F_1(1,38) = 27.87$; $F_2(1,19) = 81.47$; $p < .001$ for both). For location of first correct trial, the only significant effect was the interaction between the number of phonemes and the location of change ($F_1(1,76) = 4.52$, $p = .037$). Here, neither of the planned comparisons were significant. In the analysis of longest run of consecutive correct responses, there was a main effect of location of change ($F_1(1,76) = 6.50$, $p = .013$) and an interaction between the number of phonemes and the location of change ($F_1(1,76) = 31.85$, $p < .001$). Planned comparisons showed that replacement of the onset, C_2, produced longer runs of correct responses than replacement of part of the rime, C_3 ($F_1(1,38) = 4.27$, $p = .046$). The difference between VC_3 and C_2V was not significant. None of the subjects or items analyses showed a significant three-way interaction, as would have been expected if there were a different pattern of results for tense and lax vowels.

In sum, it was easier for subjects to learn games that substituted the onset of the medial syllable than the coda and easier to learn games that substituted the rime of the

TABLE 10

MEAN SCORES IN EXPERIMENT 5 AS A FUNCTION OF VOWEL TYPE, NUMBER OF PHONEMES MOVED, AND LOCATION OF CHANGE

Vowel type	Number correct		First correct trial		Longest run	
Lax			*One-phoneme condition*			
	C_2	C_3	C_2	C_3	C_2	C_3
	13.70	12.10	4.00	4.85	9.65	7.65
			Two-phoneme condition			
	C_2V	VC_3	C_2V	VC_3	C_2V	VC_3
	8.85	14.00	5.70	4.95	4.40	11.00
Tense			*One-phoneme condition*			
	C_2	C_3	C_2	C_3	C_2	C_3
	13.65	10.55	3.75	5.45	8.40	6.70
			Two-phoneme condition			
	C_2V	VC_3	C_2V	VC_3	C_2V	VC_3
	9.05	12.40	7.15	6.10	5.00	8.20
Lax and tense pooled			*One-phoneme condition*			
	C_2	C_3	C_2	C_3	C_2	C_3
	13.68	11.33	3.88	5.15	9.03	7.18
			Two-phoneme condition			
	C_2V	VC_3	C_2V	VC_3	C_2V	VC_3
	8.95	13.20	6.43	5.53	4.70	9.60

Note. Maximum possible score for number correct and longest run is 19.

medial syllable than the CV. This held true whether the vowel was tense or lax.

Error analyses. We asked whether, with one phoneme to be changed, subjects were more likely to replace the vowel along with the consonant in the C_3 condition (where the vowel and the consonant form a unit of the syllable) than in the C_2 condition (where the vowel and the consonant do not form a syllable constituent). In the C_3 condition, 18.2% of the errors involved a change of the vowel as well as a change of C_3 to /g/. In the C_2 condition, just 1.9% of the errors involved a change of the vowel as well as a change of C_2 to /g/. These data were analyzed using the factors of condition (C_2 or C_3) and vowel type (lax or tense); the results of four subjects who made no errors in one of the cells were omitted from the analysis. There was a main effect of condition ($F_1(1,34) = 18.79, p < .001$), a main effect of vowel type ($F_1(1,34) = 7.15, p = .011$), and an interaction between condition and vowel type ($F_1(1,34) = 6.29, p = .017$). The effect of condition was significant for both tense and lax vowels but the difference was larger for tense vowels (28.3% versus 1.9%) than for lax vowels (10.1% versus 1.9%).

We also asked whether, with two phonemes to be changed, the vowel erroneously remained unchanged more often in the C_2V condition than in the VC_3 condition. In 20.1% of errors in the C_2V condition, C_2 was changed to /g/ and the vowel remained the same. In 4.7% of errors in the VC_3 condition, C_3 was changed to /g/ and the vowel remained the same. These data were analyzed using the factors of condition and vowel type; the results of two subjects were omitted from the analysis. There was a main effect of condition ($F_1(1,36) = 10.31, p = .003$) and an interaction between condition and vowel type ($F_1(1,36) = 4.55, p = .040$). The interaction arose because the effect of condition was significant only for lax vowels (27.1% versus 6.0%), although it was in the same direction for tense vowels (13.1% versus 3.8%).

In the C_2V condition, 13.7% of the errors involved replacement of C_2 and V with the appropriate phonemes but also replacement of the final consonant; that is, they were whole-syllable errors. In the VC_3 condition, 7.8% of the errors involved replacement of V and C_3 with the appropriate phonemes but also replacement of the initial consonant. These data were analyzed using the factors of condition and vowel type; the results of two subjects were omitted. The only significant effect was that of condition ($F_1(1,36) = 28.25, p < .001$). Whole-syllable errors were more frequent in the C_2V condition than in the VC_3 condition, providing further evidence for the cohesiveness of the rime.

Discussion

The results of Experiment 5 provide clear evidence for syllable effects in the middle syllables of trisyllabic stimuli. In this experiment, a game involving a substitution of the onset was easier to learn than one involving a substitution of the coda. Also, a game involving a substitution of the rime was easier to learn than one involving a substitution of the syllable-initial CV. These results in a nonspeeded task are similar to those of the speeded task of Experiment 4, in which onset shifts were faster than coda shifts and rime shifts were faster than CV shifts. However, the results of Experiment 5 are even stronger than those of Experiment 4 in that Experiment 5 also found significant differences between conditions in the types of errors made.

A secondary issue investigated in Experiment 5 concerned possible differences between lax and tense vowels. We did not find significant evidence for such differences in the analyses which concerned subjects' ability to learn the various types of games. In the error analyses, however, we found some suggestions that vowels and coda consonants are more cohesive when the vowel is lax than when it is tense, consistent with the results of Derwing and Nearey (1990).

GENERAL DISCUSSION

In our first two experiments, we used a speeded task to explore the internal structure of monosyllables. Our findings in these experiments converged with evidence obtained in other ways to suggest that monosyllables have a hierarchical onset/rime structure. In Experiment 1, subjects who were required to shift the first two phonemes of a syllable shifted CC clusters faster than CV sequences. When required to shift just the first phoneme of a syllable, subjects in Experiments 1 shifted singleton consonants more rapidly than the first consonants of clusters. These findings are compatible with the idea that consonants before a syllable's vowel are one constituent, the onset, whereas the vowel and postvocalic consonants are another constituent, the rime. The results of Experiment 2 qualified that conclusion slightly in a direction consistent with recent theorizing in linguistics (Kaye et al., 1990) and with other behavioral evidence (Treiman et al., 1992). Specifically, the results of Experiment 2 suggested that the /s/ of initial consonant sequences such as /sp/ and /sl/ is less cohesive with the following consonant than are the first consonants of other initial consonant–consonant sequences.

Our next experiments were designed to test whether the constituents that we and others had observed with monosyllables were in fact constituents of a syllable or, as Davis (1989) had suggested, constituents of a word. The stimuli in Experiments 3A to 3C were disyllabic nonsense words. Response times to shift segments in these disyllables provided evidence against an interpretation based on syllabic constituents.

Whereas subjects were rapid and accurate at shifting the onsets of the first syllables of disyllables, they were slower and less accurate at shifting the onsets of the second syllables of disyllables (Experiment 3B). Indeed, the onset of the second syllable of a disyllable was moved as slowly as an internal coda (Experiment 3A). Finally, comparing CV and VC shifts for the two syllables of a disyllable, there was a gradual slowing of response times for sequences later in the stimulus, with some reduction in this slowing for the final VC sequence. In contrast to findings in the literature on tip-of-the-tongue errors, there was no special advantage for the onsets of internal stressed as compared to unstressed syllables. Had we stopped with disyllables, then, we might have concluded that there was no syllable structure in the individual syllables of polysyllables.

Pursuing our investigations, however, we obtained a strikingly different outcome on the stressed, medial syllables of trisyllables than we had obtained with disyllables. A major reason for examining trisyllables is that words may have both a word structure (that is, a word onset constituent and a rest-of-word constituent) and a syllable structure. These two structures coincide in monosyllables, making constituent effects easy to see. In disyllables, the word and syllable constituents only coincide at the word onset. Perhaps word-edge effects predominated with our disyllabic stimuli. To reduce these effects, we turned to trisyllables, requiring subjects to shift segments of the stressed medial syllable. A further motivation behind our trisyllable experiments was that the orthographic presentation used in Experiments 1 to 3 requires left-to-right scanning of the stimulus pair. Thus, task-specific characteristics may have been wholly or partly responsible for the findings that response times became longer as the to-be-shifted segments occurred later in the stimuli. To test for these orthographic effects, and to examine performance with the effects eliminated, we ran two trisyllable experiments. One trisyllable experiment, Experiment 4, used the speeded phoneme shift task of the preceding experiments with orthographic presentation. The other trisyllable experiment, Experiment 5, used the unspeeded word game task of Treiman (1986, Experiment 7) with auditory presentation. Both procedures, but especially the auditory one, yielded clear evidence of syllable constituency for the stressed medial syllables of trisyllabic nonwords.

Why were effects of syllable structure seen in Experiments 4 and 5 but not in Experiments 3A to 3C? Task-specific factors cannot be the *only* explanation, because Experiment 4, which used the same type of orthographic presentation as Experiments 3A to 3C, did find syllable effects. However, orthographic presentation may be a contributing factor. Specifically, orthographic presentation of stimuli in the speeded task may create strong early-to-late gradients in segment shifting times that, under some conditions, override syllable structure effects. Consistent with this idea, effects of syllable structure were stronger in Experiment 5, which used auditory presentation, than in Experiment 4, which used visual presentation. To further explore the role of presentation modality, we are now running some conditions of Experiment 3 using auditory presentation of the stimuli.

Another possible reason for the different findings in Experiments 3A to 3C as compared to Experiments 4 and 5 is that words or nonwords are not *necessarily* chunked into syllabic constituents. Rather, syllabic constituents emerge as the stimuli participate in certain tasks. For simple stimuli and tasks, people may only chunk into word-initial consonant and word remainder. Finer-grained chunks may become necessary with long stimuli, with tasks that stress working memory and with subjects of lower verbal abilities. When a need to chunk arises, constituents respect the natural inhomogeneities in cohesion among consonants and vowels that linguistic theories of syllable structure describe. According to this hypothesis, the disyllables and the visually presented phoneme shift task did not require chunking by our subjects beyond chunking into word-onset and word-remainder. The trisyllable tasks did require such chunking. According to this hypothesis, if correct, we should find evidence of syllable structure with our disyllables by making the task more difficult or by using subjects with lower verbal abilities. Conversely, we should eliminate evidence of syllable structure with trisyllables if we can make the task easier. We are currently pursuing these possibilities.

The present disyllable and trisyllable experiments were motivated by Davis' (1989) criticisms of the evidence for an onset/rime structure for English syllables. Davis pointed out that most of the previous evidence for onsets and rimes involved monosyllabic stimuli. With monosyllables, what appears to be a division between the syllable onset and the syllable rime may instead be a division between the word onset and the word remainder. Word onsets do seem to have a special status (Berg, 1989; Browman, 1978; Shattuck-Hufnagel, 1987). However, our findings do not support the idea that polysyllables are composed of *only* two constituents—an onset and a remainder. At least under some conditions, there is clear evidence for onsets and rimes within the syllables of polysyllabic items.

REFERENCES

BERG, T. (1989). On the internal structure of polysyllabic monomorphemic words: The case for superrimes. *Studia Linguistica*, **43**, 5–32.

BROWMAN, C. P. (1978). Tip of the tongue and slip of the ear: Implications for language processing. *University of California Working Papers in Phonetics*, **42**.

BROWN, A. (1991). A review of the tip of the tongue experience. *Psychological Bulletin*, **109**, 204–223.

CAIRNS, C. E., & FEINSTEIN, M. H. (1982). Markedness and the theory of syllable structure. *Linguistic Inquiry*, **13**, 193–226.

CARTER, R., & BRADSHAW, J. (1984). Producing "Spoonerisms" on demand: Lexical, phonological and orthographic factors in a new experimental paradigm. *Speech Communication*, **3**, 347–360.

CLEMENTS, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. E. Beckman (Eds.), *Between the grammar and physics of speech* (pp. 283–333). New York: Cambridge University Press.

CLEMENTS, G. N., & KEYSER, S. J. (1983). *CV phonology: A generative theory of the syllable*. Cambridge, MA: MIT Press.

DAVIS, S. (1989). On a non-argument for the rhyme. *Journal of Linguistics*, **25**, 211–217.

DERWING, B. L., & NEAREY, T. M. (1990). *Real-time effects of some intrasyllabic collocational constraints in English.* Paper presented at the International Conference on Spoken Language Processing. Kobe, Japan.

DERWING, B. L., NEAREY, T. M., & DOW, M. L. (1987, December). *On the structure of the vowel nucleus: Experimental evidence.* Paper presented at the Linguistic Society of America, San Francisco.

EWEN, C. J. (1982). The internal structure of complex segments. In H. Van der Hulst & N. Smith (Eds.), *The structure of phonological representations* (*Part II*) (pp. 27–67). Dordrecht: Foris.

FOWLER, C. A. (1987). Consonant–vowel cohesiveness in speech production as revealed by initial and final consonant exchanges. *Speech Communication*, **6**, 231–244.

FUDGE, E. C. (1969). Syllables. *Journal of Linguistics*, **5**, 253–286.

FUDGE, E. (1987). Branching structure within the syllable. *Journal of Linguistics*, **23**, 359–377.

FUDGE, E. (1989). Syllable structure: A reply to Davis. *Journal of Linguistics*, **25**, 219–220.

IVERSON, G., & WHEELER, D. (1989). Phonological categories and constituents. In R. Corrigan, F. Eckman, and M. Noonan (Eds.), *Linguistic categorization* (pp. 93–114). Amsterdam/Philadelphia: John Benjamins.

KAYE, J., LOWENSTAMM, J., & VERGNAUD, J.-R. (1990). Constituent structure and government in phonology. *Phonology*, **7**, 193–231.

KIPARSKY, P. (1979). Metrical structure assignment is cyclic. *Linguistic Inquiry*, **10**, 421–441.

KUČERA, H., & FRANCIS, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

MACKAY, D. G. (1972). The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology*, **3**, 210–227.

SELKIRK, E. O. (1982). The syllable. In H. Van der Hulst & N. Smith (Eds.), *The structure of phonological representations* (*Part II*) (pp. 337–383). Dordrecht: Foris.

SHATTUCK-HUFNAGEL, S. (1987). The role of word-onset consonants in speech production planning:

New evidence from speech error patterns. In E. Keller & M. Gopnik (Eds.), *Motor and sensory processes of language* (pp. 17–51). Hillsdale, NJ: Erlbaum.

SMITH, N. V. (1973). *The acquisition of phonology: A case study*. Cambridge, England: Cambridge University Press.

STEMBERGER, J. P. (1983). *Speech errors and theoretical phonology*. Bloomington, IN: Indiana University Linguistics Club.

STEMBERGER, J. P., & TREIMAN, R. (1986). The internal structure of word-initial consonant clusters. *Journal of Memory and Language*, 25, 163–180.

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

TREIMAN, R. (1984). On the status of final consonant clusters in English syllables. *Journal of Verbal Learning and Verbal Behavior*, 23, 343–356.

TREIMAN, R. (1986). The division between onsets and rimes in English syllables. *Journal of Memory and Language*, 25, 476–491.

TREIMAN, R. (1989). The internal structure of the syllable. In G. Carlson & M. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 27–52). Dordrecht: Kluwer.

TREIMAN, R., GROSS, J., & CWIKIEL-GLAVIN, A. (1992). The syllabification of /s/ clusters in English. *Journal of Phonetics*.

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

VENNEMANN, T. (1988). The rule dependence of syllable structure. In C. Duncan-Rose & T. Vennemann. (Eds.), *On language: Rhetorica, phonologica, syntactica: A festschrift for Robert P. Stockwell from his friends and colleagues* (pp. 257–283). London: Routledge.

(Received February 7, 1992)
(Revision received May 12, 1992)