Competition and Segmentation in Spoken-Word Recognition

Dennis Norris
Medical Research Council Applied Psychology Unit

James M. McQueen and Anne Cutler
Max Planck Institute for Psycholinguistics

Spoken utterances contain few reliable cues to word boundaries, but listeners nonetheless experience little difficulty identifying words in continuous speech. The authors present data and simulations that suggest that this ability is best accounted for by a model of spoken-word recognition combining competition between alternative lexical candidates and sensitivity to prosodic structure. In a word-spotting experiment, stress pattern effects emerged most clearly when there were many competing lexical candidates for part of the input. Thus, competition between simultaneously active word candidates can modulate the size of prosodic effects, which suggests that spoken-word recognition must be sensitive both to prosodic structure and to the effects of competition. A version of the Shortlist model (D. G. Norris, 1994b) incorporating the Metrical Segmentation Strategy (A. Cutler & D. Norris, 1988) accurately simulates the results using a lexicon of more than 25,000 words.

Subjectively, the task of recognizing words in continuous speech seems effortless. Yet the subjective ease of spoken-word recognition masks the fact that the speech stream itself is not broken up into units corresponding to words at all. The boundaries between words are not reliably marked in the input (Lehiste, 1972; Nakatani & Dukes, 1977). If we have the impression that speech comprises a discrete series of words, then this is a consequence of the output of the word-recognition process rather than a reflection of the nature of the input itself. However, how does the word-recognition process operate so effectively in the absence of clear cues to the location of word boundaries?

The literature contains three main suggestions as to how the speech-recognition system might proceed in the absence of reliable word boundary cues. The first suggestion is that recognition operates sequentially in a strictly left-to-right fashion (Cole & Jakimik, 1978, 1980; Marslen-Wilson & Welsh, 1978). As each word is identified, the boundary following that word can be located. Recognition of subsequent words follows in the same manner in a strictly left-to-right fashion. However, to work effectively, such an account requires that words can reliably be identified before their offset. Luce (1986) showed that, when frequency is taken into account, more than one third of words cannot be reliably identified until after their offset. Furthermore, Frauenfelder (1991; for Dutch) and McQueen and Cutler (1992; for English) showed that many polysyllabic words have shorter words embedded within

them and that these embedded words are most likely to appear at the onsets of the longer words. Shillcock (1990) pointed out that suffixes in particular pose a problem for sequential models. At their offset words like *run* are not yet unique because *running, runner, runs* are all possible continuations. The consequences of these theoretical concerns over the viability of strictly sequential processing have been demonstrated in experimental studies. Using the gating task, Grosjean (1985) and Bard, Shillcock, and Altmann (1988) confirmed that many words are not recognized until some time after their offset. However, the most serious failing of sequential models is that they claim that word recognition can proceed only in a strictly left-to-right manner. If recognition fails in the middle of an utterance, there is no way such a system can recover until it finds a clearly marked word boundary where the recognition process can be restarted.

In contrast to sequential procedures are theories that claim that recognition involves a separate procedure of segmentation of the input. This procedure allows lexical access to be guided by knowledge of where word boundaries are most likely to be located. Cutler and Norris (1988; see also Cutler & Carter, 1987) proposed a stress-based Metrical Segmentation Strategy (MSS), for stress-timed languages like English in which there is a contrast between strong and weak syllables. Strong syllables are those containing a full vowel; syllables with full vowels are strong regardless of whether they bear primary or secondary stress. Weak syllables, on the other hand, contain a reduced vowel, usually schwa. According to the MSS, the speech stream is segmented at the onset of strong syllables, and a new lexical access attempt is initiated at the onset of each strong syllable. The primary value of such a strategy lies with the fact that, in typical English speech, more than 90% of content words do begin with a strong syllable, and approximately 75% of all strong syllables are indeed the initial syllables of content words (Cutler & Carter, 1987).

In support of the MSS, Cutler and Norris (1988) presented data from a word-spotting task in which subjects listened to a list of bisyllabic nonsense words and were required to press a button as soon as they heard a nonsense word beginning with a real word. Detection of words with a CVCC consonant-vowel

structure was harder when they were embedded in strong–strong nonsense words (*mint* in /mɪnteɪf/) than in strong–weak nonsense words (*mint* in /mɪntəf/). According to the MSS, this result is expected because the second syllable (/teɪf/) in strong–strong strings triggers segmentation of the input, which disrupts processing of the embedded target word. When the second syllable is weak, the target is easier to detect because it is not interrupted by the segmentation process. Further support for the MSS came from a study of both natural and laboratory-induced misperceptions by Cutler and Butterfield (1992), which found that, as predicted by the MSS, listeners tended to assume that strong syllables were the onsets of content words.

This evidence supports the stress-based metrical segmentation procedure, which is appropriate for stress-timed languages such as English. Of course, this procedure could not operate for languages in which there is no alternation of strong and weak syllables. Nonetheless, a separate procedure of segmentation based on metrical structure can operate in any language; it is only different across languages insofar as the metrical structures of the languages themselves differ. Cross-linguistic studies have, in fact, demonstrated how segmentation procedures vary across languages with different rhythmic (metrical) structures. Mehler, Dommergues, Frauenfelder, and Segui (1981) and Cutler, Mehler, Norris, and Segui (1986, 1992) provided evidence that segmentation in French is syllable based, while Otake, Hatano, Cutler, and Mehler (1993) and Cutler and Otake (1994) proposed that segmentation in Japanese is mora based; in both cases, the phonological representations involved in segmentation are also those characterizing the metrical structure of each language. Metrically based segmentation can, therefore, be viewed as a language-universal procedure. However, in this article, we focus on English, and we use the term MSS to refer only to the stress-based metrical segmentation procedure.

A third approach to the segmentation problem is provided by models involving competition between candidate words. Competition theories contrast with those theories that postulate a separate segmentation process, because competition provides a way for lexical segmentation to emerge as a consequence of word recognition. In models like TRACE (McClelland & Elman, 1986) and Shortlist (Norris, 1991, 1994b), recognition is a consequence of competition between multiple lexical candidates beginning at many different points in the input. In TRACE, this process of competition is in principle open to any word in the vocabulary at any time, which is rendered possible by the alignment of a complete copy of the lexical network with each point in the input where a word might begin. All nodes corresponding to overlapping candidates have to be interconnected with inhibitory links. Obviously, there are a very large number of such links if the vocabulary is not trivially small, and this large number is repeated in each copy of the network, making TRACE a computationally expensive model.

In Shortlist, a two-stage model, the process of generating candidate words is separated from the process of competition, and this means that no problem of duplicated lexical networks arises. Shortlist deals with the temporal nature of speech *within the initial stage* of candidate word generation; in

principle, this stage may be implemented as a simple recurrent network (Norris, 1990), although in the present implementation of the model (Norris, 1994b) the initial stage is simulated with a lexical search procedure. The crucial feature of Shortlist is that it works in a completely bottom-up fashion. On the basis of an acoustic–phonetic analysis of the input, the initial stage generates a set of candidate items that are roughly consistent with the input. These candidates are then entered in the second stage into an interactive activation network in which overlapping candidates compete, in much the same way that overlapping lexical nodes compete in TRACE. However, in Shortlist, competition occurs only between that small set of candidates (the "shortlist") for which there is some bottom-up evidence. The competition network, therefore, remains small and contains only a small number of inhibitory connections. (For the 26,452-word lexicon used in the present work, Shortlist requires less than one billionth of the number of inhibitory connections required by TRACE!)

McQueen, Norris, and Cutler (1994) pointed out that sequential recognition, competition, and the existence of a separate segmentation procedure are not, in fact, incompatible. For example, competition models like TRACE and Shortlist operate in a mode that is predominantly sequential. When a word is recognized, other candidates that overlap that word will be inhibited. It will, therefore, be easier to recognize a word when the preceding word has already been recognized because the preceding word will help to inhibit any candidates that overlap with both the current and preceding words. Although competition models do not rely on strictly sequential recognition, they will be at their best when sequential recognition is possible. McQueen et al. (1994) also suggested that models using a separate segmentation procedure can be incorporated into competition models. Specifically, they suggested that the MSS can be implemented in Shortlist by giving an activation boost to candidates beginning at strong syllables.

A great merit of competition models is that their performance is quite independent of the availability or reliability of word boundary cues. Until recently, however, there was no direct evidence for lexical competition between alternative candidates. Considerable indirect support for competition is provided, admittedly, by studies using the cross-modal priming task (Swinney, 1979). Such studies support the conclusion that multiple lexical candidates are indeed activated during the early stages of auditory word recognition (Marslen-Wilson, 1987, 1990; Shillcock, 1990; Swinney, 1981; Zwitserlood, 1989). However, evidence for activation does not amount to direct evidence for competition because it is possible that multiple candidates may be activated without entering into competition with one another. More convincing evidence for activation combined with competition comes from priming studies by Goldinger, Luce, and Pisoni (1989) and Goldinger, Luce, Pisoni, and Marcario (1992) that suggested that recognition may be inhibited when words are preceded by similar sounding words. The inhibition is assumed to be caused by competition between the preceding word and the target, in accord with the predictions of the authors' Neighborhood Activation Model (Luce, Pisoni, & Goldinger, 1990).

By far the strongest and most direct evidence for lexical competition, though, comes from the study by McQueen et al.

(1994). They reported a series of three experiments that provide evidence supporting both the lexical competition mechanism proposed by Shortlist and the MSS of Cutler and Norris (1988). McQueen et al. (1994) used the same word-spotting task used by Cutler and Norris (1988). Subjects were asked to spot words in bisyllabic nonsense words such as /nɔmɛs/ and /dɔmɛs/. Each of this pair of nonsense words contains the embedded word *mess*. However, /dɔmɛs/ should also activate words like *domestic* and *domesticated*. These words should compete with *mess* and make it harder to detect *mess* in /dɔmɛs/ than in /nɔmɛs/ in which there are no such competitors. This is exactly what McQueen et al. (1994) found. They also looked at nonsense word pairs like /sækrɔf/ and /sækrɔk/, with a strong–weak stress pattern in which the embedded target word appeared in initial position. In this case, the target word is *sack* and the competitor, in the first member of the pair (/sækrɔf/), is *sacrifice*. Shortlist predicts that the presence of a competitor should make it harder to spot *sack* in /sækrɔf/ than in /sækrɔk/ but only if subjects respond slowly (i.e., if they have not already responded before the end of the following syllable). *Sacrifice* is a potential competitor for *sack* from word onset and continues to be so, in this example, up to and including the vowel of the second syllable; it is only with the arrival of the consonant at the end of the second syllable that a distinction between /sækrɔf/ and /sækrɔk/ emerges, whereby the former continues to offer bottom-up support to *sacrifice*, but the latter does not. For the same reason—that the competitor remains active for at least some time within both strings—the competition effect should also be smaller with these strong–weak items than with weak–strong items. If subjects respond very quickly, then /sækrɔf/ and /sækrɔk/ will not have diverged, and there will be no differential effect of competition. Again, these predictions conformed to what McQueen et al. (1994) found. The competition effect for strong–weak items was smaller than for weak–strong items and was only significant in the one experiment in which subjects responded most slowly.

In addition to providing support for the competition predictions of Shortlist, the McQueen et al. (1994) experiments also produced further support for the predictions of the MSS. According to the MSS, target words should be easier to spot in weak–strong items than in strong–weak items. In weak–strong items, the MSS will trigger at the onset of the target word and the target word ends in silence. The target is, therefore, effectively segmented out from the embedding nonsense word. In the case of strong–weak items, there is no segmentation at the offset of the target because the following syllable is weak. Exactly as predicted, and in all three of the experiments, words in strong–weak items took longer to spot than words in weak–strong items.

The McQueen et al. (1994) results are, therefore, important in three respects. First, they provide a direct demonstration of competition effects in an on-line task. Second, they show that competition occurs even between candidates with different onset locations (e.g., *domestic* and *mess*). Finally, they suggest that metrical segmentation effects (the difference between strong–weak and weak–strong strings) operate alongside and in addition to competition effects. A complete account of

spoken-word recognition must, therefore, incorporate such segmentation effects as well as lexical competition.

McQueen et al. (1994) showed that Shortlist gave an accurate and detailed simulation of the competition effects observed in their experiments. However, as originally proposed, Shortlist cannot explain the advantage of weak–strong items over strong–weak items. Thus, the competition mechanism needs to be supplemented by the MSS to give a full account of the results that McQueen et al. obtained. Precisely how to provide an integrated account of the segmentation and competition effects raises interesting issues. As stated by Cutler and Norris (1988), part of the MSS involves initiating a new lexical access attempt at strong syllable onsets. However, the distinctive feature of competition models is that lexical access is effectively being performed at all possible locations, such that they are all capable of being word onsets. Therefore, McQueen et al. (1994) suggested that the MSS might be incorporated into Shortlist by giving a boost to all lexical candidates beginning at a strong syllable onset. This would introduce the inequality inherent in the MSS: that strong syllables are more likely to be word onsets.

Interestingly, however, this suggestion would implement only one aspect of the original MSS. As previously described, Cutler and Norris (1988) proposed that the MSS involves both segmenting the speech input at the onset of strong syllables and initiating new lexical access attempts from the points at which segmentation has occurred. Furthermore, it is the former process, segmentation, that appears to have the strongest direct support: The study by Cutler and Norris (1988) is most simply interpreted as showing evidence of segmentation at strong onsets. The misperception study of Cutler and Butterfield (1992) likewise showed systematic metrical effects in listeners' judgments of juncture locations.

A simple way to add the segmentation component of the MSS to Shortlist would be to impose a penalty on lexical candidates that straddle a strong syllable onset in the input. However, it would be more accurate to restrict the penalty only to words that straddle a strong onset when the lexical representation of the word does not explicitly mark the onset of a strong syllable, because this would avoid the problem of penalizing polysyllabic words with noninitial strong syllables. This procedure would penalize *mint* in /mɪntɜɪf/ because the /t/ in the input begins a strong syllable, but the lexical representation of *mint* does not mark the /t/ as being the onset of a strong syllable. Because the /t/ in /mɪntəf/ does not begin a strong syllable, there would be no penalty in this case and *mint* would be recognized more readily.

We can, therefore, think of the two components of the MSS proposal (the segmentation procedure and the consequent lexical access from segmentation sites) as being directly instantiated in the two components of the MSS-inspired additions to Shortlist (the Penalty and the Boost, respectively). Both can also be directly implemented in the procedure by which Shortlist computes the degree of fit between candidates and the input. In contrast to TRACE, Shortlist makes use of both match and mismatch information. Evidence in favor of a candidate increases its bottom-up score, and evidence against a candidate decreases the score. So a strong syllable onset in the input increases the score for candidates beginning with a

strong onset but decreases the score for candidates that are not marked as having a strong onset at that point. The two components of the MSS, therefore, have a direct parallel in the way Shortlist already uses match and mismatch information in computing the bottom-up score of each candidate.

In the present study, we report an implementation of the MSS in Shortlist, of the type just described. To ground this implementation in more than the arguments in principle just outlined, however, we also conducted a further empirical study of competition and metrical segmentation. To determine how best to give an integrated account of the evidence both for lexical competition and for the MSS, we need data that specifically address the issue of the interaction between the two. One way to achieve this is to repeat the Cutler and Norris (1988) experiment while explicitly manipulating the number of competitors beginning at the onset of the second syllable (e.g., at the /t/ in /mɪnteɪf/). Such a manipulation achieves two goals. First, by controlling the number of competitors, it establishes whether the original result might possibly have been an artifact of different numbers of competitors in strong–weak and strong–strong items. Second, it enables us to determine whether the MSS effect can be modulated by varying the number of competitors. If the MSS and competition processes are completely independent, then MSS effects should remain constant regardless of the number of competitors. Alternatively, if the MSS effect varies according to the number of competitors, then this suggests that it might indeed be plausible to instantiate the MSS in terms of the competition mechanism in Shortlist. In either case, the detailed pattern of results in this experiment will provide additional constraints in helping us decide how best to implement the MSS.

Experiment

In this experiment we pitted the MSS against the effect of lexical competition, using the Cutler and Norris (1988) word-spotting task. We varied the number of competitor words in the second syllable, beginning at the final consonant of CVCC target words (e.g., beginning at the /t/ in /mɪnteɪf/). This consonant will be referred to as the competitor-initial consonant. The competition factor had two levels: Either there were many possible words beginning with this consonant (as determined by a search of a machine-readable dictionary, see later discussion) or there were few such words. As in Cutler and Norris's (1988) Experiment 3, we included both CVCC and CVC target words. There were, therefore, three factors in the experiment, each with two levels: stress pattern (strong–strong vs. strong–weak); number of potential lexical competitors in the second syllable (many or few); and target word structure (CVCC or CVC).

The CVCC targets in the experiment were used to explore several hypotheses. First, they provide a direct test of competition effects irrespective of stress pattern. The simple competition prediction is that it should be more difficult to detect a target word when the competitor-initial consonant and the phonemes that follow it activate a large number of lexical hypotheses than when relatively few competitors are activated, because the larger the number of competitors the greater their combined inhibitory influence should be on the activation of the target word.

Second, it was possible to test for stress pattern effects with number of competitors controlled. All words beginning from the competitor-initial consonant should compete with target words that end with this consonant: The more there are, the harder it should be to detect the targets. It is the case in English that there are more words starting with strong syllables than with weak syllables (Cutler & Carter, 1987). It was, therefore, possible that the stress pattern effect in Cutler and Norris (1988) could have been a competitor effect. To test this hypothesis we counted the number of words in a machine-readable version of the *Longman Dictionary of Contemporary English* (*LDOCE;* Procter, 1975), which began with the target-final consonants and following vowels used in the Cutler and Norris (1988) materials (e.g., the words beginning /teɪ/ and /tə/ from /mɪnteɪf/ and /mɪntəf/). We found that there were, in fact, more words beginning from the weak syllables than from the strong syllables. In their Experiment 1 materials (excluding those that were discarded from the analysis), there were, on average, 36 words beginning from the target-final consonant in the strong–strong items and 151 words on average beginning from this consonant in the strong–weak items. This difference was even more marked in a similar count for their Experiment 3 materials: 44 words on average in strong–strong and 265 words on average in strong–weak. The reason for this imbalance is simple, even though the overall lexical statistics (Cutler & Carter, 1987) show the opposite pattern. The overall count for strong syllables sums across all instances of full vowels, whereas that for weak syllables is composed mainly of instances of the weak-vowel schwa (see Altmann & Carter, 1989). In the counts for these materials, therefore, most words with weak initial syllables were included for the strong–weaks, whereas only those with specific full vowels were included for the strong–strongs, thus reversing the overall trend.

Any competition process in the Cutler and Norris (1988) experiments should thus have been acting against the MSS. The finding that word spotting was more difficult in the strong–strong than the strong–weak items in spite of the imbalance in the number of competitors is, therefore, powerful evidence for the MSS. Nevertheless, it is possible to provide a rather different competition explanation for these findings. Bard (1990) argued that a large number of weakly activated lexical competitors (those beginning at the competitor-initial consonant in a strong–weak string), through inhibiting each other, should have a weaker combined inhibitory effect on the target word than a smaller number of more highly activated competitors (in a strong–strong string) that would inhibit each other less and thus have a greater combined inhibitory effect on the target word. This argument rests both on acoustic differences between strong and weak syllables (weak syllables are activated less because they tend to be phonetically more poorly specified) and on set-size differences. It is not possible to control for the acoustic differences because they, in part, define the strong–weak distinction, but it is possible to control the number of competitors.

In this experiment, therefore, we attempted to balance the number of second-syllable competitors in matched strong–

strong and strong–weak strings. Bard's (1990) argument predicts not only that there should be stress pattern effects (degraded weak syllables tending to activate more competitors than strong syllables even when the numbers have been matched on the basis of phonemic transcriptions) but also main effects of competition. Strong second syllables activating many competitors should inhibit target detection less than those activating few competitors (inhibition within the competitor set should be greater when it is larger so the combined effect of the set on the target is smaller), and weak second syllables with many competitors should also be less problematic than those with few competitors. Note that this claim is the opposite of the competition effect that emerges from Shortlist, which (like other interactive activation models) predicts that the greater the number of competitors, the greater their inhibitory effect on the target. Indeed, Shortlist simulations of Experiment 3 from Cutler and Norris (1988) show that competition effects do result in CVCC words in strong–strong strings being more highly activated at the vowel in the second syllable than CVCC targets in strong–weak strings. However, by the end of the second syllable, this effect was reversed. This ambiguity in the Shortlist simulations reinforced the need to examine the stress effect with number of competitors equated.

Third, and most important for our goal of implementing the MSS in Shortlist, the CVCC items allowed us to test for the interaction of metrical segmentation and lexical competition. The MSS is predicted to operate irrespective of the number of lexical hypotheses, but its influence on recognition might be modulated by the number of competitors. There could, therefore, be a larger stress pattern effect when there are many words activated by the second syllable than when there are few such words.

The CVC items were included primarily as a control condition, as in the earlier experiment. The MSS does not predict that CVC target detection will be influenced by the stress pattern of items such as /θɪnteɪf/ and /θɪntəf/. The target (thin) does not overlap with the second syllable in either case, so segmentation on the basis of whether this syllable is strong or weak should not influence detection of thin. Therefore, these items do not provide a test of the MSS. However,

they do give an additional test of lexical competition. If there are many candidates beginning from the competitor-initial consonant in CVC target-bearing items (e.g., the /t/ in /θɪnteɪf/ or /θɪntəf/), detection of the target word may be facilitated relative to when there are only a few such candidates. In strings with CVC targets, the more words there are that begin from the first phoneme after the end of the target, the more likely it is that there is indeed a word boundary at that point. Both the target and the words beginning from the competitor-initial consonant will inhibit words that straddle this boundary point, and the activation of the target will be boosted. This prediction is the opposite of the prediction for the CVCC targets, for which it is predicted that performance will be worse with many than with few competitors beginning from the competitor-initial consonant.

Method

Design and materials. The experiment had a 2 × 2 × 2 factorial design, with one within-item and two between-item factors. All test materials were bisyllabic nonsense strings with a C(C)VCCVC consonant-vowel structure. One hundred sixty were constructed in 80 pairs that had the same target word in initial position but that differed in stress pattern. Each of these pairs consisted of one item with two strong syllables (SS) and one item with a strong syllable followed by a weak syllable (SW). They were, in fact, identical apart from the vowel in the second syllable, which was either a full vowel (in the SS items) or schwa (in the SW items). Stress pattern was, therefore, the within-item factor.

One of the between-item factors was number of potential lexical competitors beginning in the second syllable. A search of the machine-readable dictionary (*LDOCE*) showed that, particularly for words beginning with weak syllables (containing schwa), the number of words found varied greatly depending on the initial consonant. Thus, for example, there were many words beginning with /kə/ but few words beginning with /tə/ (Table 1). Three consonants (/p, k, s/) were, therefore, chosen as competitor-initial consonants for the many-competitor condition and three (/t, d, tʃ/) were chosen for the few-competitor condition. The CVC syllables beginning with these

Table 1

Mean Number of Lexical Competitors Beginning in the Second Syllable for Each of the Six Competitor-Initial Consonants and Overall for the Many- and Few-Competitor Conditions

Condition	Competitor-initial consonant	CV				CVC			
		SS		SW		SS		SW	
		M	SD	M	SD	M	SD	M	SD
Many competitors	/p/	150	47	180	0	12	15	2	2
	/k/	279	129	521	0	30	35	24	40
	/s/	272	65	204	0	19	15	17	16
Overall		226	114	337	167	21	27	14	29
Few competitors	/t/	34	14	37	0	0	0	1	2
	/d/	31	13	39	0	0	0	0	0
	/tʃ/	7	5	2	0	0	0	0	0
Overall		29	16	32	13	0	0	1	2

Note. Shown are the mean number of competitors given the competitor-initial consonant and the following vowel (CV), and the mean number given the whole of the second syllable (CVC), as counted in the *Longman Dictionary of Contemporary English* (Procter, 1975). SS = strong–strong; SW = strong–weak.

consonants in the many-competitor condition were selected using the following constraints: They were not words, and as many words as possible began with the CV and with the CVC. The corresponding syllables for the few-competitor condition were again selected to be nonwords but were chosen to have as few words as possible beginning with the CV and the CVC. All words found that began with each CV (or CVC) were counted irrespective of length. The mean numbers of competitors in each condition are shown in Table 1.

The other between-item factor was target structure. Half of the target words had a C(C)VCC structure (referred to as CVCC), half had a C(C)VC structure (referred to as CVC). Every CVCC stress-pattern pair was yoked to a CVC stress-pattern pair, such that only their onsets differed (e.g., /stæmpɪdʒ-stæmpədʒ/ was paired with /præmpɪdʒ-præmpədʒ/).

Target words were matched on mean frequency of occurrence both between the CVCC and CVC structure conditions and between the many and few competitor conditions. It was not possible to match each CVCC word with its CVC pair (e.g., *stamp* and *pram*) on frequency: Frequency matching was done over the complete item sets. A full list of the experimental materials is given in the Appendix, together with the mean word frequency for each condition. In addition to the target-bearing items, 120 filler items were constructed. These items were also all nonsense bisyllables, but none of them began with a real word.

Procedure. The materials were recorded in a sound-damped booth by a native male speaker of British English onto digital audiotape. They were then redigitized, sampling at 22.05 kHz with 16-bit resolution, and examined with a speech editor. Timing pulses were aligned with the onset of the competitor-initial consonant of each experimental item. At output, the items were upsampled to 44.1 kHz and then recorded onto the left channel of a digital audiotape at a rate of 1 every 3 s. The timing pulses were recorded onto the right channel of the tape (so they were inaudible to the subject during the experiment).

Two experimental tapes were constructed, each consisting of the same 120 filler items and 1 member of each of the 80 experimental item pairs. Stress pattern (SS vs. SW) was counterbalanced across tape, such that, for example, /mɪntaup/ appeared on one tape, whereas /mɪntəp/ appeared on the other tape. The nonword items and the target words appeared in the same positions on both tapes. Thus, the tapes were identical except that the slot for a given target word was filled with either the SS or SW item containing that target word. A short practice tape of 12 items was also prepared.

Subjects were tested individually in a quiet room. They were told that they would hear a list of nonsense words and that they should press a response key whenever they heard a nonsense word beginning with a real word, and then say aloud the word that they had spotted. Subjects' spoken responses were recorded onto audiotape. Subjects heard the practice tape, followed by one of the two experimental tapes. The items were presented binaurally over headphones. Timing and data collection were controlled by a Commodore microcomputer.

Results

The spoken responses of each subject were analyzed first. Occasionally, subjects made a manual response but then either failed to give a verbal response or responded with a word other than the target word; these responses were treated as missing data. Five subjects failed to detect at least two thirds of the 80 targets. The data from these subjects were not analyzed further. The responses from one other subject were lost because of equipment failure. The data from 30 subjects were, therefore, analyzed, 15 subjects per tape.

Analyses of variance (ANOVAs) were performed on reaction times (RTs) and miss rates, with both subjects (F_1) and items (F_2) as the repeated measure. Responses to CVCC and

CVC targets were analyzed separately. Some of the target words proved to be very difficult to detect. One CVCC item (fiend) and 11 CVC items (dim, gal, scum, truss, gull, flan, fen, ton, grin, goal, and brine) were removed from the analysis because they were missed by more than half of the subjects. Thus, the item ANOVAs had uneven numbers of targets per condition (20 targets in the many-CVCC condition, 19 in the few-CVCC condition, 15 in the many-CVC condition, and 14 in the few-CVC condition). Separately for CVCC and CVC targets, each subject's overall mean and standard deviation were computed. Responses more than two standard deviations from the mean were rejected as outliers and treated as missing data (this procedure rejected 6% of the data, distributed evenly across conditions). For the RT analyses, missing data points within each condition were replaced, separately for each subject or item, with the mean of the other responses of that subject or for that item in that condition. The mean RTs and miss rates are given in Table 2.

CVCC analyses. In the RT analysis there was a main effect of stress pattern. Targets were detected, on average, 65 ms faster in SW strings than in SS strings: $F_1(1, 28) = 17.62, p < .005, MSE = 7221; F_2(1, 37) = 5.62, p < .05, MSE = 16000$. Although targets were detected faster when there were few potential lexical competitors beginning in the second syllable than when there were many, this competition effect (22 ms on average) was not significant. The competition effect did not reliably interact with the stress pattern effect, but the stress effect was larger when there were many competitors (89 ms on average), $t_1(29) = 2.86, p < .01; t_2(19) = 2.07, p = .05$, than when there were few competitors (40 ms on average), $t_1(29) = 1.69, p = .10; t_2(18) = 1.25, p > .1$.

The size of the stress pattern effect was correlated with the frequency of the target word. The RT difference between the SS and SW responses for each item was positively correlated with the frequency of occurrence of the items: The greater the word frequency, the larger the difference (the greater the

Table 2

Mean Reaction Time (RT) and Miss Rate for the Detection of Target Words Embedded in the Onsets of Bisyllables With Potential Lexical Competitors Beginning in the Second Syllable

Number of competitors and stress pattern	Example	RT (ms)	Miss rate (%)
CVCC target structure			
Many			
SS	STAMPidge	696	22
SW	STAMP@dge	607	13
Few			
SS	MINTaup	650	19
SW	MINT@p	610	17
CVC target structure			
Many			
SS	PRAMpidge	586	19
SW	PRAMp@dge	625	27
Few			
SS	THINtaup	655	25
SW	THINt@p	696	19

Note. SS = strong–strong; SW = strong–weak; C = consonant; V = vowel.

advantage for SW over SS), $r(38) = 0.36, p < .05$. This effect was more robust when there were few second-syllable competitors, $r(18) = .46, p < .05$, than when there were many second-syllable competitors, $r(19) = .29, p > .10$.

There was also a correlation of target word frequency with mean speed of response to targets but only for target words followed by weak syllables, $r(38) = -.361, p < .05$. Examined in more detail, this effect was limited to targets followed by weak syllables with few competitors: many competitors, $r(19) = -.136, p > .5$; few competitors, $r(18) = -.508, p < .05$. There were no reliable correlations for targets followed by strong syllables. Note that the items producing a reliable frequency effect (those with weak second syllables, activating few competitors) are those that Shortlist predicts should show the largest frequency effects. These are the items for which frequency effects are least likely to be masked by the influence of strong second syllables and many activated competitors. These results are consistent with other frequency effects in word spotting (Freedman, 1992). They provide one explanation for the high between-item variance in the experiment (and hence for why the 40-ms stress effect when there were few second-syllable competitors failed to reach significance).

In the analysis of proportion of missed targets, there was again no competition effect (no difference between proportion of missed targets for many vs. few second-syllable competitors) and again a significant stress effect. Targets were detected, on average, 6% more accurately in SW strings than in SS strings: $F_1(1, 28) = 10.20, p < .005, MSE = 0.0090; F_2(1, 37) = 5.27, p < .05, MSE = 0.0121$. However, the stress effect was not equivalent across many–few competitors. The two-way interaction was significant by subjects, $F_1(1, 28) = 5.69, p < .05, MSE = 0.0090$, although not by items, $F_2(1, 37) = 2.53, p = .12, MSE = 0.0121$. As in the RT analysis, the stress effect was significant when there were many second-syllable competitors (9% on average): $t_1(29) = 2.47, p < .05; t_2(19) = 2.73, p < .05$, but the 2% effect for the items with few second-syllable competitors was not significant by subjects or items.

In the error analyses, the size of the overall stress pattern effect shown by each subject was found to correlate positively with the overall error rate of the subject, $r(29) = .47, p < .01$: The more errorful the subject, the larger the advantage for SW over SS strings. A post hoc split of the data into the 14 more accurate and the 16 more errorful subjects was, therefore, made. An ANOVA on error rates showed a significant interaction of stress pattern and overall accuracy across subjects, $F_1(1, 28) = 7.64, p = .01, MSE = 0.0121; F_2(1, 37) = 1.33, p > .2, MSE = 0.0248$: There was no difference in accuracy of target detection between SS and SW strings for the more accurate subjects (12% error in both conditions) but a mean difference of 10% (SS: 28% errors; SW: 18% errors) for the more errorful subjects. This ANOVA also produced a three-way interaction of subject accuracy with stress pattern and number of second-syllable competitors: $F_1(1, 28) = 4.45, p < .05, MSE = 0.0157$; $F_2 < 1$. An ANOVA on the accurate subjects alone yielded no significant effects, but one on the errorful subjects alone showed that the difference of 10% between SW and SS strings was significant, $F_1(1, 15) = 14.05, p < .005, MSE = 0.0131$; $F_2(1, 37) = 5.04, p < .05, MSE = 0.0323$. This effect was due almost entirely to items with many second-syllable competitors

(SS: 33% errors; SW: 13% errors), not to items with few competitors (SS: 24% errors; SW: 22% errors), as shown by the interaction of number of competitors with stress pattern, $F_1(1, 15) = 5.40, p < .05, MSE = 0.0222; F_2 < 1$.

The split into accurate and errorful subjects was also performed on the RT data. An ANOVA on RT showed a three-way interaction of subject accuracy, number of competitors, and stress pattern: $F_1(1, 28) = 5.89, p < .05, MSE = 13016; F_2(1, 37) = 3.17, p = .07, MSE = 6770$. Just as in the errors, an ANOVA on the accurate subjects alone gave no significant effects, although the stress pattern effect (37 ms on average; SS: 619 ms; SW: 582 ms) was marginal, $F_1(1, 13) = 3.03, p = .1, MSE = 6476; F_2(1, 37) = 3.04, p = .09, MSE = 20748$. For the errorful subjects, the advantage of SW over SS strings (89 ms on average; SS: 721 ms; SW: 632 ms) was significant, $F_1(1, 15) = 14.28, p < .005, MSE = 8942; F_2(1, 37) = 10.80, p < .005, MSE = 12091$, and, in contrast to the overall RT analysis, this effect interacted with number of competitors, $F_1(1, 15) = 5.26, p < .05, MSE = 15599; F_2(1, 37) = 2.47, p = .12, MSE = 12091$, with an average advantage of SW over SS of 161 ms (SS: 775 ms; SW: 614 ms) when there were many second-syllable competitors and an average advantage of only 17 ms (SS: 667 ms; SW: 650 ms) when there were few competitors.

The more errorful subjects tended to be those who responded more slowly (as shown by the positive correlation between subjects' overall mean speed and overall mean accuracy), $r(29) = .44, p < .05$. Although the size of the stress pattern effect in RT did not reliably correlate with subjects' speed, $r(29) = .22, p > .2$, the data were split into the fastest 15 and the slowest 15 subjects. (Nine of the 15 fast subjects also fell in the more accurate category, and 10 of the 15 slow subjects fell in the more errorful category. Although speed and accuracy should correlate quite highly, individual differences in performance will prevent a perfect correlation.) Just as in the accuracy split, the slower subjects produced larger stress pattern effects and larger interactions of stress pattern with number of second-syllable competitors. For these subjects' RTs, the advantage of SW over SS strings (85 ms on average; SS = 786 ms; SW = 701 ms) was significant, $F_1(1, 14) = 10.57, p < .01, MSE = 10080; F_2(1, 37) = 4.99, p < .05, MSE = 30307$, but the interaction with number of competitors failed to reach significance, $F_1(1, 14) = 3.05, p = .1, MSE = 16191; F_2(1, 37) = 1.41, p > .2, MSE = 30307$ (many competitors: SS = 819 ms, SW = 677 ms; few competitors: SS = 752 ms, SW = 726 ms). Likewise, in the slow subjects' errors, the advantage of SW over SS strings (8% on average; SS = 24% errors; SW = 16% errors) was also significant, $F_1(1, 14) = 6.69, p < .05, MSE = 0.0147; F_2(1, 37) = 4.36, p < .05, MSE = 0.0273$, but again the interaction with number of competitors failed to reach significance, $F_1(1, 14) = 2.35, p = .15, MSE = 0.0174; F_2(1, 37) = 1.34, p > .2, MSE = 0.0273$ (many competitors: SS = 27% errors; SW = 14% errors; few competitors: SS = 20% errors, SW = 17% errors). The fast subjects, on the other hand, showed a main effect of stress pattern in RT, $F_1 = 5.14, p < .05, MSE = 6165; F_2(1, 37) = 6.36, p < .05, MSE = 7564$ (SS = 561 ms on average; SW = 515 ms on average) but no interaction of stress pattern with number of competitors (F_1

and $F_2 < 1$), and there were no significant effects in their errors.

CVC analyses. The only significant effect in the RT analysis was a competition effect. Responses to target words in items in which the second syllable had many competitors were, on average, 69 ms faster than responses to target words in items in which the second syllable had few competitors. This effect was only significant by subjects, $F_1(1, 28) = 11.52, p < .005, MSE = 12748$; $F_2(1, 27) = 2.66, p > .1, MSE = 51545$. The stress pattern effect (responses to SW items 41 ms slower, on average, than responses to SS items) and the stress–competition interaction were not significant by subjects or items.

There were no significant main effects in the CVC error analyses. There was, however, a crossover interaction of stress pattern and number of competitors, significant only by subjects. There were more errors to SW than SS strings when there were many second-syllable competitors (8% on average) but more errors to SS than to SW strings when there were few competitors (6% on average), significant only by subjects, $F_1(1, 28) = 5.73, p < .05, MSE = 0.0265$; $F_2(1, 27) = 3.44, p > .05, MSE = 0.0212$. This interaction was explored with pairwise t tests; the stress pattern effects within each competitor condition were not significant by subjects or items.

Finally, the data for both the CVC and CVCC targets were submitted to combined ANOVAs. Several interactions were found, confirming that performance on the two types of target was quite different. In RT, target structure interacted with number of second-syllable competitors (because the competition effects were in opposite directions in the CVCCs and CVCs), $F_1(1, 28) = 16.38, p < .001, MSE = 7691$; $F_2(1, 64) = 4.11, p < .05, MSE = 31932$. Target structure also interacted with stress pattern (because there was a reliable advantage for SW over SS in the CVCCs but not in the CVCs), $F_1(1, 28) = 16.36, p < .001, MSE = 10164$; $F_2(1, 64) = 5.72, p < .05, MSE = 17554$. In errors, the only reliable interaction involving target structure was the three-way interaction of target structure, stress pattern, and number of competitors, $F_1(1, 28) = 9.87, p < .005, MSE = 0.0192$; $F_2(1, 64) = 6.38, p < .05, MSE = 0.0159$.

Discussion

Detecting CVCC targets is harder (responses are slower and less accurate) when they are followed by strong syllables than when they are followed by weak syllables. This effect is predicted by the MSS and replicates the Cutler and Norris (1988) result. When targets precede a strong syllable, they have to be assembled across a segmentation position. They do not when preceding a weak syllable. This result does more than replicate the previous finding, however, because the stress effect was measured here with the number of competitors controlled. The effect is, therefore, not due to the set size of competitors beginning from the second syllable. It seems clear that the difficulty of detecting CVCC targets in SS strings relative to SW strings is indeed due to stress pattern.

There was no main effect of competition in the CVCC responses. Overall, a subject's ability to detect CVCC targets was not influenced by the number of possible words beginning from the final consonant of the target. Although there were no

reliable differences, the trend in the RTs for the SS items was for responses to be slower when there were more second-syllable competitors, contrary to the prediction following from Bard (1990). There was no such effect in the SW items. This pattern of results can be interpreted as being due to the conjoint influences of the MSS and lexical competition. The number of second-syllable competitors has an effect only when the MSS indicates that the onset of those competitors is likely to be the beginning of a word, that is, in SS strings. In SW strings, however, no segmentation position is postulated, so the number of second-syllable word hypotheses should have a weaker influence on performance. Indeed, the lexical competition process should ultimately disfavor any parse of the input that leaves a CVC in the first syllable position in the CVCC target-bearing items, because this CVC will always be a nonword. Competition will tend to provide a parse of the input in which every segment is accounted for by a single lexical hypothesis. So it is, in fact, unlikely that a main effect of competition will be found in the CVCC items: There is some indication of an effect in the SS items only, because in these items the MSS acts to make it much more likely that there is a word boundary after the initial CVC.

Importantly, in the analysis of errors on CVCC targets, there was an interaction of stress pattern with number of second-syllable competitors. When there were many words activated by the second syllable, the stress effect was larger (the 9% effect was significant) than when there were few words activated by the second syllable (the 2% effect was not significant). Although the equivalent interaction in the overall RT analysis was not significant, the differences between the means were in line with the error rates: The advantage for SW over SS items was larger with many than with few second-syllable competitors.

Furthermore, the subjects who found the task more difficult (those who detected fewer targets or responded more slowly or both) showed larger interactions of stress pattern with number of competitors in both errors and RT. For the errorful subjects in the accurate–errorful split, these interactions were statistically reliable. For the slower subjects in the fast–slow split, these interactions, although larger than in the overall analysis, failed to reach significance. Shortlist predicts that subjects with higher response thresholds will tend to be slower and more errorful because threshold in these subjects will be reached less quickly and less often. If target activation remains below threshold, there is a greater opportunity for the following information to have an effect. Slower, more errorful subjects are thus more likely to be influenced both by the prosodic characteristics of the second syllable (whether it is strong or weak) and by the number of competitors activated by the second syllable. In keeping with Shortlist, the stress pattern effect and its interaction with number of competitors emerged more strongly in the responses of the more errorful and slower subjects.

Even though the Stress Pattern × Competitor interaction only appeared in RT in a subset of the subjects, the interaction in errors was present in the overall analysis. It therefore appears that lexical competition modulates the size of the metrical segmentation effect. The larger the number of second-syllable word candidates in SS strings, the greater is their pull

on the final consonant of the target; therefore, the harder it becomes to parse the input with this consonant as part of the first syllable.

For the CVC target words, there was no stress pattern effect, also as predicted by the MSS. The absence of this effect again replicates the results of Cutler and Norris (1988). There was, however, a marginal competition effect. Detecting CVC targets that were followed by syllables that activated only a few lexical hypotheses was more difficult (responses were slower but no less accurate) than detecting targets that were followed by syllables that activated many lexical hypotheses. This competition effect in the CVC items—in the reverse direction and larger than in the CVCC items—is consistent with the suggestion we advanced previously: namely, that a parse with a word boundary after the initial CVC should be favored in the case of a CVC target because in this case, in contrast to the strings containing CVCC words, the parse leaves a CVC word. The presence of many candidates in the second syllable should, therefore, boost the first-syllable hypothesis for both the SS and SW items. There is also some indication in the RTs that the competition effect is larger in the SS items than in the SW items. This is probably due to the added influence of the MSS in the SS items, increasing the likelihood of a word boundary after the CVC. In these items, then, the MSS appears to act to increase the size of the competition effect, even though no main effect of stress pattern is predicted, because for both the SS and the SW items the target does not span a segmentation position. Note, however, that the analyses in the CVC condition are based on fewer items. The pattern of CVC results should perhaps, therefore, be interpreted with caution.

Simulations

The experiment confirms the earlier finding (McQueen et al., 1994) that metrical segmentation, as instantiated in the MSS, and lexical competition, as instantiated in Shortlist, are both required to account for the process of continuous spoken-word recognition. It also constrains the way in which the MSS could be incorporated in Shortlist. The MSS appears to operate on the input, on the basis of strong syllables, irrespective of the number of words consistent with the input. However, lexical competition can then influence the effect on recognition that the MSS can have. When there are many lexical hypotheses beginning at a segmentation position (as determined by the MSS), recognition of a word that straddles this position is more difficult than when there are relatively few lexical hypotheses beginning at the segmentation position.

This finding represents an important extension of the results from McQueen et al. (1994). That study demonstrated that recognition was more difficult in the presence of a single highly active competitor than when no competitors were present. The current results demonstrate that such competition effects are sensitive to the number of competitors. At least in the case of SS strings, the more competitors there are, the harder recognition becomes. Can this pattern of results be simulated in Shortlist?

Recall that Shortlist is a two-stage model. In the first stage, a lexical search procedure generates a shortlist of lexical candidates on the basis of a completely bottom-up analysis of the input. The second stage is an interactive activation network in which all overlapping candidates are connected by inhibitory links with weights proportional to the number of phonemes by which they overlap. The distinctive feature of Shortlist is that the interactive activation network is generated dynamically so that it contains only those candidates produced by the first stage. After the presentation of each new phoneme as input, the set of lexical candidates is revised and the bottom-up evidence in favor of each candidate is updated. The network is then run for 15 cycles before the next phoneme is presented; this allows it to settle on a new interpretation of the current state of the input. The next section describes the steps we took to incorporate the MSS in Shortlist to enable it to simulate the interaction of competition and segmentation in human word recognition.

MSS Instantiation

We first tried to simulate the interaction between competition and segmentation using only the modification to Shortlist suggested by McQueen et al. (1994); that is, we gave a boost to all candidates with a strong initial syllable that began at a strong-syllable onset. The Boost took the form of adding 1.0 to the bottom-up score for the candidate. The Boost, therefore, had the effect of increasing the bottom-up score by the equivalent of one extra matching phoneme. Note that we never applied the Boost immediately after silence. The MSS is designed to suggest likely locations for word boundaries. After a substantial silence, the next phoneme must be the onset of a word, and it would, therefore, be inappropriate to favor words with strong onsets over words with weak onsets. There is, indeed, no experimental evidence that, after silence, words with weak initial syllables are any harder to process than words with strong initial syllables (Cutler & Clifton, 1984).

We made the assumption in all simulations that the fact that a syllable is strong would be detectable during the processing of the first phoneme of the syllable. The MSS should properly have its impact at the onset of the strong syllable; but given that the input to Shortlist is a discrete series of nonoverlapping phonemes with no co-articulation, the effective onset may not be identical to that in speech input to a listener. Accordingly, we decided to designate the first phoneme of each syllable as the syllable's onset. The impact of the Boost or Penalty caused by the strong-syllable onset can, therefore, be seen in the simulations at the end of the first phoneme of the syllable.

Recall that, as explained early in this article, the Boost procedure only allows the basic MSS effect (*mint* being harder in /mɪnteɪf/ than /mɪntəf/) to be produced indirectly via competition. The Boost has no direct effect on the activation of *mint*. In fact, despite experimenting with a wide range of model parameters, we found it impossible to reproduce anything more than a minute MSS effect using the Boost procedure. The reason for this is quite simple. By the time the competitors arrive, the target word (*mint*) has already reached a high level of activation. The target is so highly activated that the competitors are very strongly inhibited. The competitors never get off the ground, so they never get into a position where they can effectively inhibit the target. Note that this pattern of

behavior is not peculiar to Shortlist. Frauenfelder (1993) presented simulations of the effects of lexical competitors in TRACE. He also found that when competitors arise late in a word they have little or no impact on the activation level of that word.

Exactly the same factors also prevent the model from simulating the present data using the Penalty procedure. Although the Penalty procedure can reproduce the basic MSS effect by directly affecting the activation of words in strong–strong targets, the model is still insensitive to the effects of competitors beginning late in the word.

There is, however, a way in which we can modify Shortlist to allow late-arriving competitors to exercise the kind of effect they do, in fact, exercise in human word recognition (i.e., the effect observed in our experiment). This modification was discussed by Norris (1994b) as a way of allowing the model to produce an optimized interpretation of the input even under conditions of relatively high word-level inhibition. The modification involves regular recomputation of activation levels for all candidates, independently of earlier activation levels. The way in which this is implemented in practice is that, at the start of each new phoneme,[1] the current activation level of all candidates is annulled, the new state of the bottom-up evidence is evaluated, and the network then runs for 15 cycles (generally more than enough to reach asymptote) to recompute a new set of activation values. These values, which represent a new and optimized interpretation of the input, then become the new current output. (In all plots of the simulations presented here, only the final activation levels representing the recomputed interpretation after each phoneme are presented.) The bottom-up evidence to the candidate words is unaffected by this modification, which only affects computation of activation levels.

This mode of operation simulates a system in which new, optimized interpretations are continually being generated and updated. In practice, of course, Shortlist is implemented with discrete time-slices of input, corresponding to successive phonemes; consequently, the only point at which activation levels are annulled is when the input to the network changes with the presentation of each new phoneme. The important feature of this mode of operation is that it continually optimizes the output interpretation. All lexical candidates are treated equally, regardless of their position in time, and an existing interpretation of the input within the network will be abandoned when it ceases to be optimal. As each new phoneme arrives, the network can settle anew on the best current interpretation of the input, uninfluenced by decisions made earlier.

The effect can be appreciated by considering the problems that can arise when the standard version of Shortlist processes a string such as *bush allowance*. Assuming that this string is spoken naturally, there will be no interword pause, and the first two syllables will correspond to a rendition of the nonintended word *bushel*. At some point *bushel* will become more active than *bush*. However, even by the end of the input, when it should become apparent that the only interpretation that completely accounts for the input is "bush allowance," the inhibition generated by *bushel* can prevent *bush* and *allowance* from becoming active. If the lexical activations are recomputed

anew, however, all candidates will start from an equal footing and the correct candidates, which fully span the input, will win out over *bushel*. Essentially, the recomputation prevents the network from irreversibly latching on to the first analysis to achieve a slight advantage. The network is free to alter its interpretation as new evidence becomes available. As each new phoneme arrives, both existing candidates and new candidates are building their activation levels up from zero. New candidates can thus build up their activation without being suppressed by earlier candidates, which have previous bottom-up evidence in their favor. Competitors occurring late in a word can have an impact without inhibition from earlier words overwhelming them.

By recomputing activation after each phoneme, we can simulate the basic SS–SW difference with both the Boost procedure (increasing the activation of words starting at strong onsets) and the Penalty procedure (penalizing words that do not contain strong onsets when there is a strong onset in the input). To achieve this, though, we found it necessary to increase the level of word–word inhibition relative to the bottom-up excitation. To keep the network stable, this also meant reducing the size of the changes in network activation made at each cycle by decreasing the absolute levels of word–word inhibition, bottom-up excitation, and decay. The new values for these parameters were 0.07, 0.03, and 0.06, respectively, as opposed to 0.12, 0.05, and 0.3 used in the earlier simulations with Shortlist reported in Norris (1994b) and McQueen et al. (1994).

Simulation of the Current Experimental Results

In presenting the results of the simulations, we plot the activation levels achieved at the end of each recalculation after the addition of each phoneme. A more complete account of spoken-word recognition could also include a decision-making mechanism to indicate exactly when each word is recognized. Such a mechanism might, for example, take the form used by Norris (1994a) to generate precise quantitative estimates of both speed and accuracy of responses in an interactive activation model (in that case, a model of speeded word naming). In that model, responses were generated when activation exceeded some prespecified response threshold and the dominant output was also by some winning margin more strongly activated than any competing responses. However, the general effect of this, or any other psychologically plausible decision mechanism, is to translate higher activation levels into faster and more accurate responses. So, because our current concern is with the operation of the competition and segmentation processes, and these are most directly revealed in the pattern of activation, we have not included a decision component in the current simulations.

Figure 1 shows the results of Shortlist simulations of CVCC items using a dictionary of 26,452 words, which is based on *LDOCE*. The first syllables of all the words in this lexicon were marked as being either strong or weak. The items from the

---

[1] In fact, resetting at intervals determined by the passage of time rather than at phoneme arrival would have the same effect.
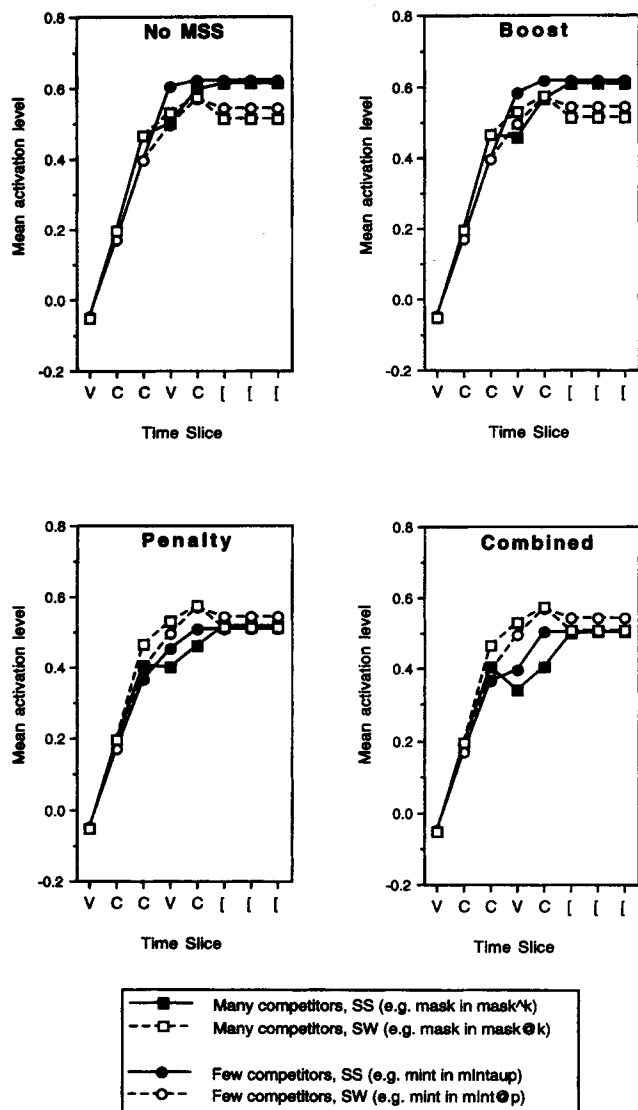
Figure 1. Mean activation levels for the CVCC materials from the experiment in the Shortlist model using a 26,452-word lexicon, with the No Metrical Segmentation Strategy (MSS) procedure, with the Boost procedure alone, with the Penalty procedure alone, and with the Combined Boost and Penalty procedures. The activation functions are shown for the items with many second-syllable competitors (squares) and for the items with few second-syllable competitors (circles). In each case, functions are given for the strong–strong (SS) strings (solid lines, solid symbols) and the strong–weak (SW) strings (dashed lines, open symbols). The activation functions are aligned with time slices for each of the last five phonemes of the input strings (consonants [C] and vowels [V]) and three following silent segments ("[").

experiment were transcribed, complete with strong-syllable markers (where appropriate) and given as input to the model. The graphs show the output of the model averaged over all experimental items. (Only the final states of activation reached after completion of each round of processing, after addition of each phoneme to the input, are shown.) Figure 1 shows the results with no MSS implemented at all, followed by the Boost

procedure alone, the Penalty procedure alone, and the Boost and Penalty combined.

The first thing to note about these simulations is that, with no MSS, there is a difference between SS and SW items in the opposite direction from that observed in the data. After the end of the target, targets are more highly activated in SS than SW items. With the Penalty procedure, however, this difference is eliminated. In the Penalty simulation, words in SW strings are somewhat more highly activated than those in SS strings throughout the final syllable of the target. Moreover, in accord with the data, the advantage of SW over SS items is greater for words with many competitors than for words with few competitors. The results of the Penalty simulation are, therefore, more in line with the data. The Boost simulation also shows a small interaction between stress and number of competitors, but this effect is restricted to the region of the final vowel. The Boost fails to remove the reversed stress effect after target offset seen in the No-MSS simulation.

The most impressive simulation of the data, however, comes from the Combined Boost and Penalty simulation. The Combined simulation is similar in general form to the Penalty simulation but shows, at the final vowel and consonant of the input, a larger effect of stress and a large interaction between stress and number of competitors. The Combined simulation thus shows most fully the pattern observed in the experiment.

The qualitative effects of the different implementations of the MSS remain constant across a wide range of model parameters. In general, Boost alone causes a transient decrease in activation in the final syllable (particularly on the final vowel) for targets embedded in SS strings with many competitors. SS strings with few competitors are almost unaffected by the Boost. Penalty alone also causes a decrease in the activation of targets in the second syllable of SS strings with many competitors but also results in a reduction in the asymptotic activation level for SS targets, relative to the No-MSS simulation, in both competitor conditions. As the size of the Penalty is increased, its effect on the asymptotic activation level becomes more pronounced until, ultimately, the target fails to reach a higher level of activation than its competitors. High levels of Penalty also lead to a reduction in the transient decrease in the activation of targets in SS strings with many competitors.

The effect of the Boost disappears completely at the end of the target string because of the way the matching procedure computes the bottom-up score in Shortlist. The competitors here are almost all more than two phonemes long, so they overlap with the following silence. The silence counts as a mismatch (score $= -3.0$), which, combined with their bottom-up score of two phonemes plus the Boost of one phoneme, leaves them with a total bottom-up score of zero. Therefore, once the silence arrives, the competitors have no further impact on the target. The Penalty, on the other hand, has an effect that lasts indefinitely because it permanently lowers the bottom-up score of penalized candidates.

The largest effect of stress and the largest interaction between stress and number of competitors appeared in the Combined simulation. We thought it advisable to investigate whether increasing the Penalty or Boost alone might make them behave a little more like the Combined. In fact, even

doubling the size of the Penalty mainly influenced the asymptotic difference between SS and SW. Stress pattern and number of competitors became almost completely additive. Doubling the size of the Boost proved somewhat more helpful in that it slightly increased the differences between the SS and SW items in the many-competitors condition. However, even when the Boost was doubled in size, there remained an asymptotic difference in the wrong direction between the SS and SW activations. The use of both Boost and Penalty together gives a better account of the data than any level of Boost or Penalty alone.

These simulation results were also analyzed quantitatively. ANOVAs were carried out on the activation values for each item, over the last three phonemes of the input string (i.e., the second syllable) and the three following silent segments (i.e., in all, the final six positions plotted in the Figure 1 graphs). These analyses were performed separately for each of the four Shortlist simulations: no MSS, Boost alone, Penalty alone, and Boost and Penalty combined. In the No-MSS simulation, there was a reliable stress pattern effect, in the wrong direction, with the mean activation of targets in SS strings (0.574) higher than that of targets in SW strings (0.516), $F(1, 38) = 6.79, p < .05$, $MSE = 0.0591$. The Boost MSS only slightly reduced this difference, but it remained in the wrong direction (SS mean = 0.564; SW mean = 0.516) and reliable, $F(1, 38) = 4.42, p < .05, MSE = 0.0614$. As can be seen in Figure 1, the Penalty MSS produces a stress pattern effect in the correct direction (SS mean = 0.472, SW mean = 0.516). This effect was only marginally significant, $F(1, 38) = 3.56, p = .07$, $MSE = 0.0658$, but interacted reliably with position, $F(5, 190) = 3.62, p < .005, MSE = 0.0071$. Separate analyses were performed by position. The stress pattern effect was reliable at the vowel of the second syllable (SS mean = 0.426, SW mean = 0.514), $F(1, 38) = 5.51, p < .05, MSE = 0.0280$, and at the final consonant (SS mean = 0.484, SW mean = 0.572), $F(1, 38) = 18.64, p < .001, MSE = 0.0084$.

For the Penalty-alone simulations, there was no main effect of number of second-syllable competitors in either the overall analysis or the analyses by position, nor was there an interaction of stress pattern and competitors. It was only in the Combined Boost and Penalty simulation that this interaction was obtained. In this analysis, there was a main effect of stress pattern (SS mean = 0.454, SW mean = 0.516), $F(1, 38) = 6.60, p < .05, MSE = 0.0711$, which again interacted with position, $F(5, 190) = 8.05, p < .001, MSE = 0.0073$. In the analyses by position, the stress pattern effect was again reliable both at the vowel and final consonant of the second syllable (see Figure 1); at the vowel, SS mean is 0.369 and SW mean is 0.514, $F(1, 38) = 13.25, p < .001, MSE = 0.0316$; at the consonant, SS mean is 0.455, SW mean is 0.572, $F(1, 38) = 27.08, p < .001, MSE = 0.0101$. However, on the consonant, there was also a reliable interaction of stress pattern with number of second-syllable competitors, $F(1, 38) = 5.12, p < .05, MSE = 0.0101$. These analyses make it clear that the Combined simulation provides a better account of the data than the Penalty-alone simulation. The activation functions shown in Figure 1 for the Combined simulation are not simply further apart than those shown for the Penalty alone: it is only in the Combined simulation that there is the statistically

significant interaction of stress pattern and competition effects that is present in the human data.

The CVC target-bearing items from the experiment were also presented as input to Shortlist, with the large dictionary and the same parameters. The mean activations of these items are shown in Figure 2. Here the overall pattern accurately captures the experimental results: There were no significant
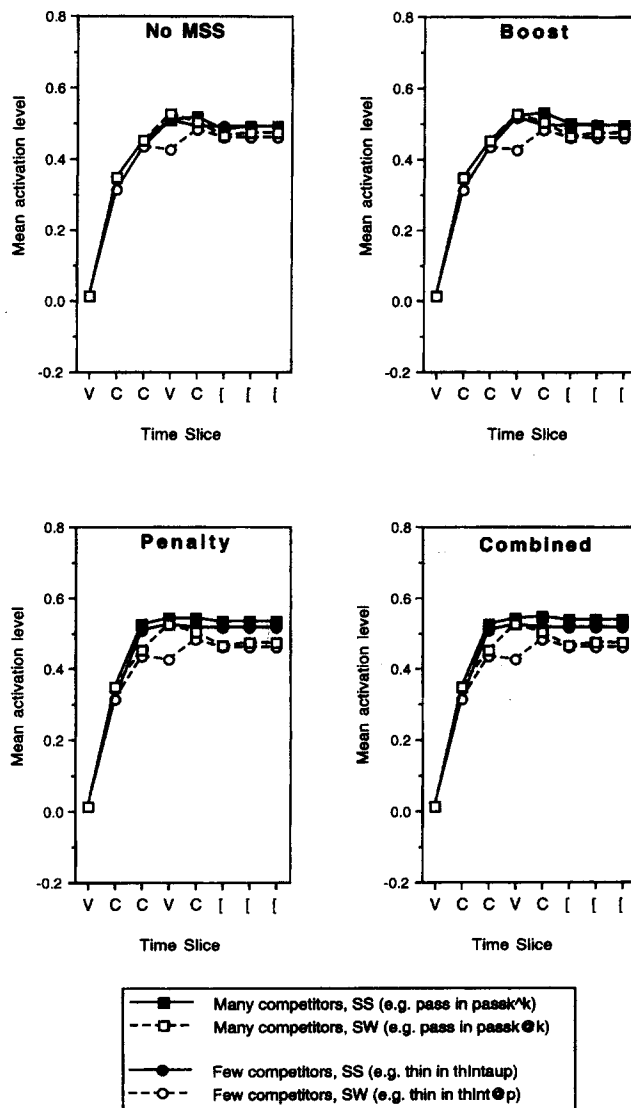

Figure 2. Mean activation levels for the CVC materials from the experiment in the Shortlist model using a 26,452-word lexicon, with the No Metrical Segmentation Strategy (MSS) procedure, with the Boost procedure alone, with the Penalty procedure alone, and with the Combined Boost and Penalty procedures. The activation functions are shown for the items with many second-syllable competitors (squares) and for the items with few second-syllable competitors (circles). In each case, functions are given for the strong–strong (SS) strings (solid lines, solid symbols) and the strong–weak (SW) strings (dashed lines, open symbols). The activation functions are aligned with time slices for each of the last five phonemes of the input strings (consonants [C] and vowels [V]) and three following silent segments ("[").

differences between conditions for CVC items, and the activation functions in the simulations are very similar for all conditions. By comparing the three MSS simulations with the No-MSS simulation, we can see that, as predicted by Cutler and Norris (1988), the MSS has almost no effect with CVC target-bearing items.

Simulation of the Previous Experimental Results

In our efforts to simulate our data, we were forced to make a change to Shortlist that might possibly have important implications for the model's ability to simulate other data. Not only did we alter the model to recompute lexical activation after each phoneme, but we also increased the level of word–word inhibition relative to the amount of bottom-up excitation. To check that these changes did not alter the character of the model, we further simulated, with the current model parameters, the earlier data of McQueen et al. (1994) showing evidence in favor of lexical competition and segmentation. The results of these simulations are shown in Figures 3 and 4. As before, all the materials from that experiment were given as input to the model, and mean activation functions are shown. For these simulations, we again used a large dictionary. However, we could not use the LDOCE-based lexicon because, to assess the activation of polysyllabic words like domestic and sacrifice, we required a lexicon in which all syllables, not just the initial syllables, were marked as strong or weak. Because the strong–weak marking had to be performed by hand, a smaller dictionary was selected: one of 8,925 words, which were the most frequently occurring words in the CELEX lexical database (Burnage, 1990).

The main competition effect, in the WS strings (see Figure 3), is that activation of targets (e.g., mess) is lower in word onsets (e.g., /dəmɛs/) because of the activation of the longer words (e.g., domestic) than in the matched nonword onsets (e.g., /nəmɛs/). This effect is present in the No-MSS simulation, replicating the simulations reported in McQueen et al. (1994), and in each of the MSS implementations, Boost alone, Penalty alone, and Combined Boost and Penalty. The competition effect is also replicated in the SW strings (see Figure 4). Note that Figure 4 is not divided into separate figures because the results for the SW items were identical across all four simulations; the MSS, however it might be implemented, does not trigger on these items. Here, competition is shown by the difference in activation of, for example, sack in /sækrəf/ versus /sækrək/. Target activation is lower because of the competition from the longer words (e.g., sacrifice). In line with the human data, competition effects are small until relatively late, because the SW strings did not diverge until after the end of the target (i.e., sacrifice is a viable competitor in both /sækrəf/ and /sækrək/ until the final consonant). Clearly, none of the changes we have made to the model have undermined Shortlist's ability to simulate the competition effects observed by McQueen et al. (1994). Indeed, our experience with the model shows that it continues to demonstrate these competition effects despite wide-ranging variations in the model parameters.

In addition to demonstrating competition effects, the McQueen et al. (1994) experiments also showed an effect of the

MSS, in that targets in WS strings were identified more easily than targets in SW strings. The MSS-augmented version of Shortlist now also allows us to simulate this feature of the earlier results. Not surprisingly, the Penalty procedure makes little contribution to the WS–SW difference; this difference is mainly explained by the Boost procedure, which triggers for strong onsets in running speech but not for onsets immediately after silence. Figure 5 shows a subset of the data in Figures 3 and 4, giving a comparison of mean activations for the same targets embedded in WS and SW strings, respectively (e.g., mess in /nəmɛs/ vs. /mɛstəm/) for the four different simulations. A main effect of stress pattern, in the absence of any competition effects, can be seen only in the Combined and in the Boost-alone simulations: Activation of targets such as mess is higher in WS strings than in SW strings. Exactly this effect, namely an advantage for targets in WS over SS strings, was also observed in McQueen et al.'s (1994) experiments, but this main effect could not be captured in the simulations they reported, in which there was no MSS implementation.

As with the simulation results for the current experiment, the activation values from the simulations of the McQueen et al. (1994) data were analyzed statistically. The activation values for each item, over the six segmental positions beginning at the segment before the final phoneme of the target (i.e., from −1 to +4 in Figures 3, 4, and 5), were compared separately for the four different simulations. In an ANOVA of the nonword onsets alone, there was no main effect of stress pattern in the No-MSS simulation. In a similar analysis for the Penalty-alone simulation, there was a main effect of stress pattern: Mean activation of targets in WS strings (0.433) was higher than that of targets in SW strings (0.392), $F(1, 31) = 5.10, p < .05, MSE = 0.0053$. However, this effect interacted with position, $F(5, 155) = 12.69, p < .001, MSE = 0.0034$, and in pairwise t tests by position for the WS-matched nonword onsets and the SW-unmatched nonword onsets (e.g., mess in /nəmɛs/ vs. /mɛstəm/; i.e., the materials used to plot Figure 5), the advantage of WS strings over SW was limited to the phoneme immediately after the final phoneme of the target (WS mean = 0.538, SW mean = 0.498), $t(15) = 2.17, p < .05$.

In the Boost-alone and Combined simulations, on the other hand, there was a large advantage for WS over SW strings, which extended over most positions (see Figure 5). In the overall analyses, there was a main effect of stress pattern, Boost alone: $F(1, 31) = 157.96, p < .001, MSE = 0.0060$; Combined: $F(1, 31) = 211.68, p < .001, MSE = 0.0049$, which interacted with position, Boost alone: $F(5, 155) = 19.32, p < .001, MSE = 0.0030$; Combined: $F(1, 31) = 19.34, p < .001, MSE = 0.0030$. In pairwise t tests by position for the WS-matched nonword onsets and the SW-unmatched nonword onsets (e.g., mess in /nəmɛs/ vs. /mɛstəm/), the advantage of WS strings over SW appeared at all positions except the position before the final phoneme of the target. For Boost alone, the reliable differences were at the final phoneme of target (C in Figure 5), $t(15) = 2.38, p < .05$; at Position +1, $t(15) = 6.25, p < .001$; at Position +2, $t(15) = 7.10, p < .001$; at Position +3, $t(15) = 6.09, p < .001$; at Position +4, $t(15) = 6.19, p < .001$. The results for the Combined Boost and Penalty simulation were very similar: At the final phoneme of target C, $t(15) = 2.55, p < .05$; at Position +1, $t(15) = 7.31$,
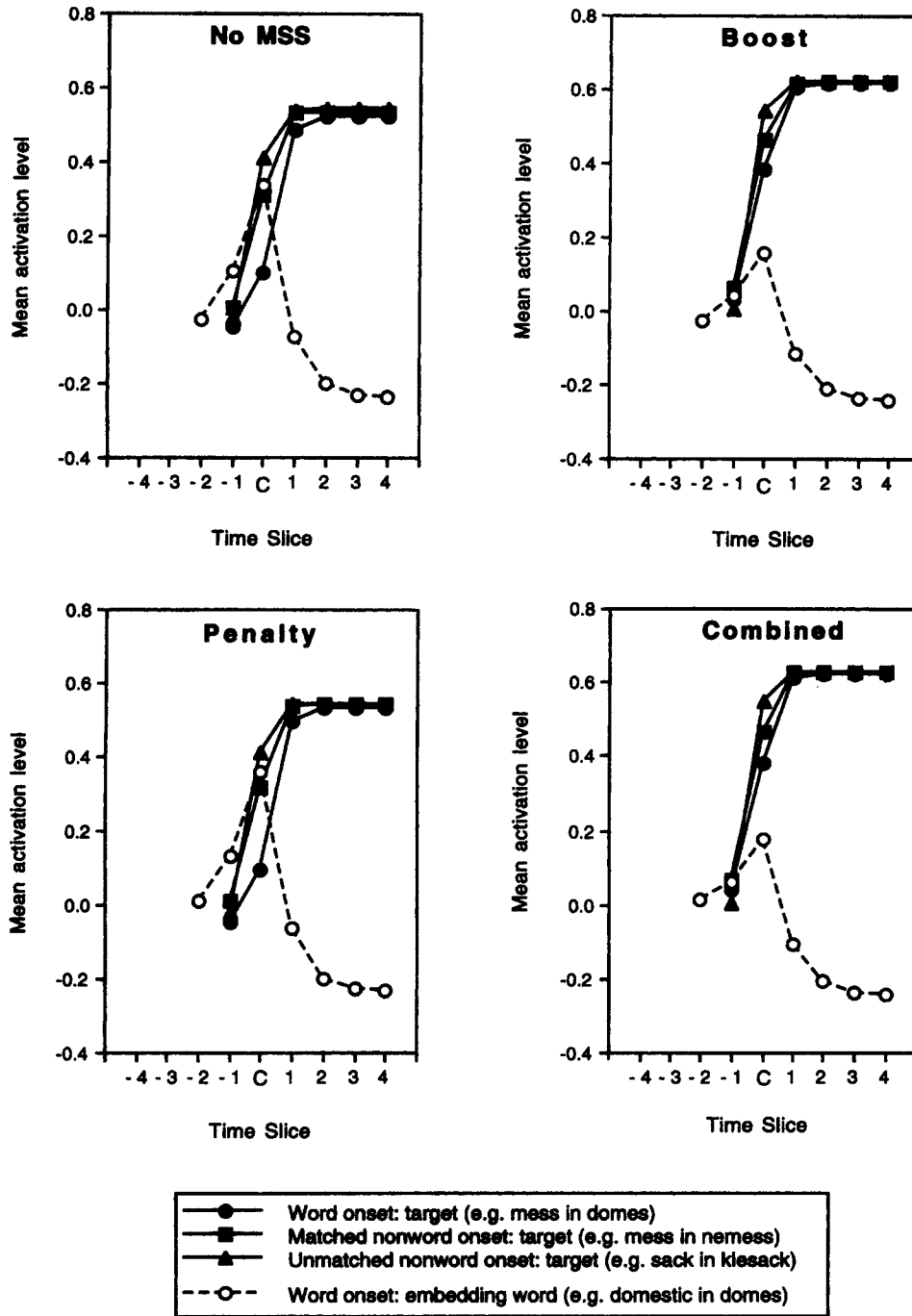
Figure 3. Mean activation levels of target words and embedding words for the weak–strong items from McQueen, Norris, and Cutler (1994), over time slices, in Shortlist. Filled symbols show activation of targets embedded in word onsets (circles; e.g., *mess* in /dɔmɛs/, the onset of domestic), in nonword onsets matched to word onsets (squares; e.g., *mess* in /nɔmɛs/), and in unmatched nonword onsets (triangles; e.g., *sack* in /klɔsæk/). Open circles show the activation of the embedding words in the word-onset items (e.g., *domestic* in /dɔmɛs/). The time slices are marked to indicate the alignment of the activation functions relative to the last consonant of the target word (C). Slices before C are for each phoneme working back through each item; slices after C contained silence markers. Results are shown for simulations with the No Metrical Segmentation Strategy (MSS) procedure, with the Boost procedure alone, with the Penalty procedure alone, and with the Combined boost and penalty procedures.
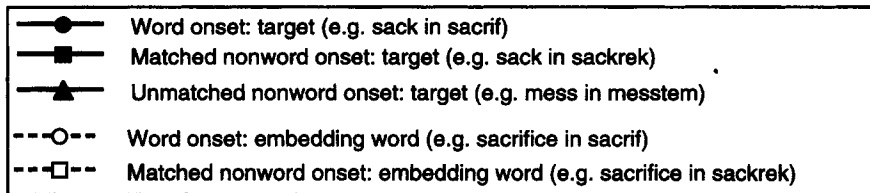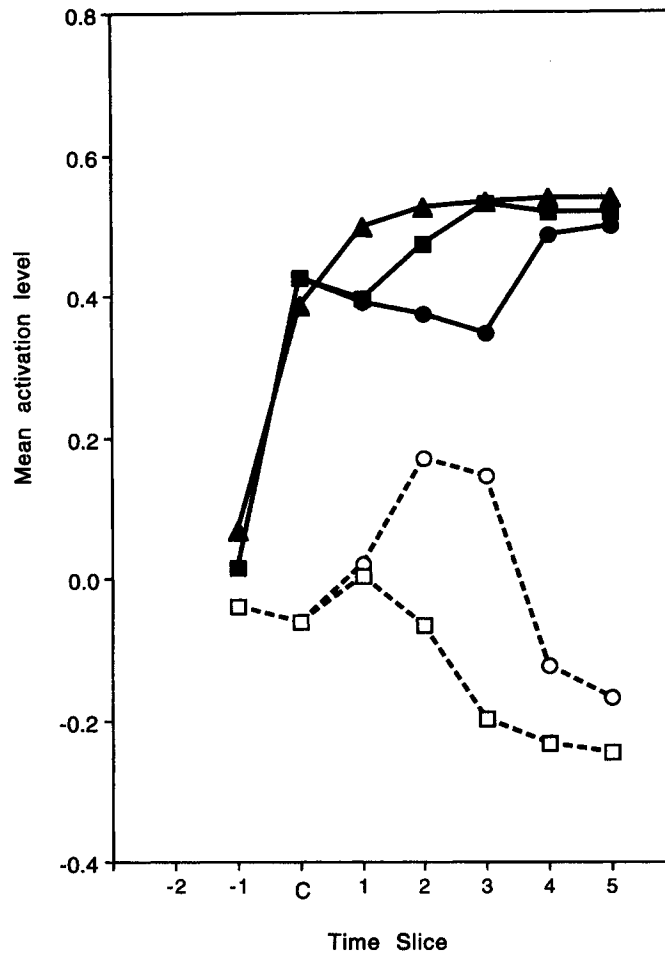
Mean activation level

Time Slice

●———	Word onset: target (e.g. sack in sacrif)
■———	Matched nonword onset: target (e.g. sack in sackrek)
▲———	Unmatched nonword onset: target (e.g. mess in messtem)
---○---	Word onset: embedding word (e.g. sacrifice in sacrif)
---□---	Matched nonword onset: embedding word (e.g. sacrifice in sackrek)

Figure 4. Mean activation levels of target words and embedding words given the strong–weak items from McQueen, Norris, and Cutler (1994), over time slices, in Shortlist. Filled symbols show activation of targets embedded in word onsets (circles; e.g., *sack* in /sækrəf/, the onset of sacrifice), in nonword onsets matched to word onsets (squares; e.g., *sack* in /sækrək/), and in unmatched nonword onsets (triangles; e.g., *mess* in /mestəm/). Open symbols show the activation of the embedding words in word onsets (circles; e.g., *sacrifice* in /sækrəf/) and in matched nonword onsets (squares; e.g., *sacrifice* in /sækrək/). The time slices are marked to indicate the alignment of the activation functions relative to the last consonant of the target word (C). Slices before C are for each phoneme working back through each item; slices after C contained the following phonemes in the bisyllable and then silence markers. Only one graph is shown because the results were identical across all four simulations: with no Metrical Segmentation Strategy procedure, with Boost alone, with Penalty alone, and with Combined Boost and Penalty.

$p < .001$; at Position +2, $t(15) = 8.25$, $p < .001$; at Position +3, $t(15) = 7.54$, $p < .001$; at Position +4, $t(15) = 7.72$, $p < .001$.

It should be noted that the competition effects in these ANOVAs were also reliable. They had the same pattern of significance over positions in the data from all four simulations. In the WS strings (e.g., *mess* in /dəmes/ vs. /nəmes/), activation of targets in word onsets was reliably lower than that of targets in matched nonword onsets at the final phoneme of the target word and at the phoneme positions immediately
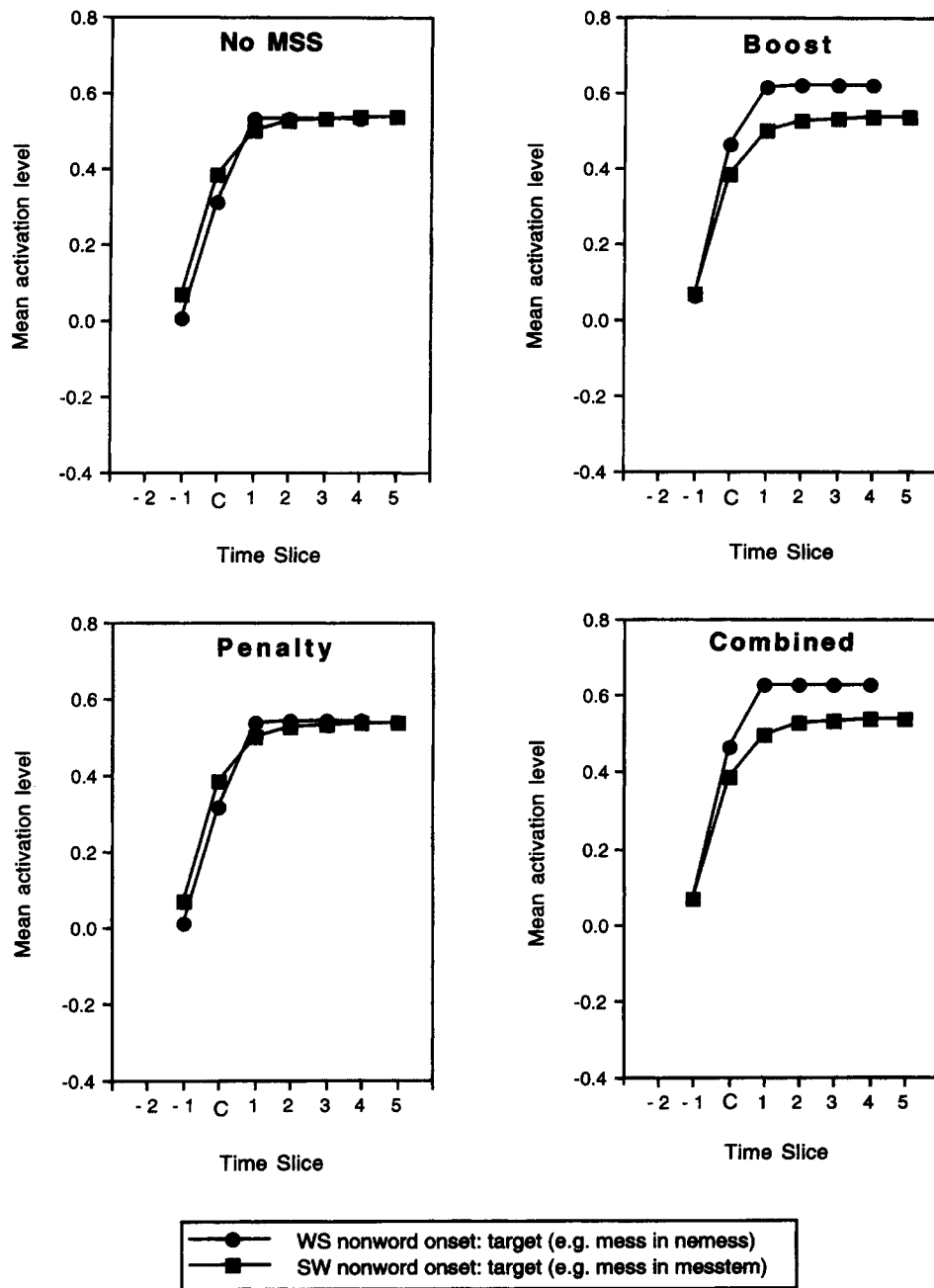
Figure 5. Mean activation levels of target words from weak–strong (WS) and strong–weak (SW) strings used in McQueen, Norris, and Cutler (1994). Squares show the activation of targets embedded in WS nonword onsets (e.g., *mess* in /nəmɛs/); triangles show the activation of the same targets embedded in SW nonword onsets (e.g., *mess* in /mɛstəm/). These activations were obtained when the model was running with the No Metrical Segmentation Strategy (MSS) procedure, with the Boost procedure alone, with the Penalty procedure alone, and with the Combined Boost and Penalty procedure. The time slices are marked to indicate the alignment of the activation functions relative to the last consonant of the target word (C). Slices before C are for each phoneme working back through each item; slices after C contained either silence (the WS items) or following phonemes and then silence markers (the SW items).

before and after this final phoneme (see Figure 3). In the SW strings (e.g., *sack* in /sækrəf/ vs. /sækrək/), activation of targets in word onsets was reliably lower than that of targets in matched nonword onsets at Phoneme Position 2 after the final phoneme of the target and the two positions after that (see Figure 4).

Thus, the Boost MSS and the Combined Boost and Penalty MSS, in Shortlist, both simulate the data from McQueen et al.

(1994) very well, producing statistically reliable stress pattern and competition effects. Although the Penalty MSS simulates the competition effects in these data, it fails to produce a reliable stress pattern effect.

This final simulation reinforces our earlier conclusion that neither the Boost nor the Penalty alone suffices to capture an MSS simulation. Although Penalty alone gives a good account of the data from the present experiment, we need to use the Boost procedure to obtain an acceptable simulation of the complete data from McQueen et al. (1994). Even when considering just the current data, only the Combined procedure produces the interaction between stress and number of competitors evident in the human data. When all of the simulations are considered together, the Combined Boost and Penalty simulations consistently provide a closer fit to the data than simulations using either Boost or Penalty alone. The two-component view of the MSS is also in accord with the way in which Shortlist uses both match and mismatch information at the segmental level. Therefore, neither the data nor the simulations give us any reason to abandon our original view of the MSS as a two-component system with a Boost at strong onsets combined with a Penalty for candidates mismatching with strong onsets.

General Discussion

The present study builds on and extends the work of both Cutler and Norris (1988) and McQueen et al. (1994). We have extended the earlier results in three respects. First, we have shown that competition effects can arise from competitors beginning later than the target word. In the McQueen et al. (1994) study, competitors began before the target (*mess* in /dɔmes/) or coincident with the target (*sack* in /sækrɔf/). In the present study, the competitors began at the final phoneme of the target word. Second, we have shown that competition effects are sensitive to the number of active competitors. The more competitors, the greater the effect of competition, as long as those competitors begin at a segmentation point. Finally, we have shown that competition effects interact with the effects of metrical segmentation. Effects of metrical segmentation of the kind first observed by Cutler and Norris (1988) become significant only when a large number of competitors begin at the onset of the strong syllable. This is an important finding because it demonstrates that the results of Cutler and Norris (1988) are not in any way artifacts of lexical competition; they can be explained only by a theory that takes explicit account of the metrical information proposed by the MSS. Indeed, simulating Experiment 3 from Cutler and Norris (1988) with the new version of Shortlist incorporating the MSS now shows an appropriate effect of stress instead of the ambiguous effect that was present in simulations using the original version of Shortlist with no MSS.

Thus, the present study provides further support for the notion of lexical competition embodied in Shortlist and for the use of metrical information as proposed by the MSS. In addition to these experimental findings, we have also shown that Shortlist, a model of spoken-word recognition incorporating interword competition, allows the MSS to be incorporated within it in quite a straightforward way and, moreover, in a way

that intuitively captures the nature of the original MSS proposal. The extension of Shortlist to give an account of metrical segmentation effects does not sacrifice any of its ability to account for competition effects. Our preferred method of incorporating the MSS in Shortlist implements the two components of the MSS in terms of a Boost to candidates with strong onsets that begin at strong onsets combined with a Penalty for candidates that do not contain strong onsets when there is a strong onset in the input. Not only can this implementation of the MSS give a convincing simulation of the present data, but it can also simulate the complete pattern of data from McQueen et al. (1994).

To be sure, our implementation of the MSS in Shortlist may appear to have a rather different flavor to the original specification of the MSS by Cutler and Norris (1988). The original MSS consisted of a segmentation procedure and a procedure to initiate lexical access attempts at strong onsets. The MSS was entirely prelexical. In the Shortlist implementation, the effects of the MSS arise at the level of the lexical competition process. However, our current implementation still depends on the operation of a prelexical process that identifies the onsets of strong syllables. The main change in the MSS has been its evolution from a framework in which lexical access attempts were discrete events that might possibly be under the control of some external strategy to a view of lexical access as a continuous and autonomous process. The MSS cannot determine whether or not lexical access takes place, but it can modulate the strength of the bottom-up evidence in favor of candidates beginning with strong onsets that are aligned with strong onsets in the input. This is exactly the function that the MSS was originally designed to achieve. The character of the MSS may have changed slightly, but its role has not. The MSS takes advantage of knowledge of the relation between metrical information and the likely locus of word onsets to help facilitate the identification of words in continuous speech.

We believe that the present results confirm the claim put forward early in this article and in McQueen et al. (1994) that competition and the exploitation of a separate process of segmentation are not at all mutually exclusive; their compatibility can easily be demonstrated both experimentally and computationally. The experimental evidence for competition effects now seems to us conclusive. The many demonstrations of multiple activation of candidates for word recognition (e.g., Cluff & Luce, 1990; Marslen-Wilson, 1990; Shillcock, 1990; Zwitserlood, 1989) have been further augmented by a variety of different findings showing that, as predicted by competition models, candidates that are simultaneously active mutually inhibit one another (Goldinger et al., 1989, 1992; McQueen et al., 1994). The present finding shows, moreover, that the more such simultaneously active words there are, the more inhibition they generate. An effect of inhibition dependent on the number of competitors, similar to the present finding, has also been reported by Vroomen and de Gelder (1995). Interword competition is a robust finding that models of word recognition must now necessarily take into consideration. Models that involve competition between candidate words, such as the Neighborhood Activation Model (Luce et al., 1990) or the interactive activation models such as TRACE (McClelland &

Elman, 1986) and Shortlist (Norris, 1994b) are all favored by the present findings.

Likewise, however, metrically based segmentation is also supported by a rich and robust variety of evidence. Thus, we also believe that models that incorporate a separate procedure of segmentation are favored by our findings. Shortlist is, as we have seen, such a model. As described early in this article, the evidence for segmentation procedures comes from a number of phonologically quite diverse languages. We believe that the principles that guided our implementation of the MSS in Shortlist should also be directly extensible to languages other than English. Shortlist is a model that can be given input in any language and can draw on a dictionary of any language. Metrically based segmentation is a language-universal procedure, which happens to have language-specific realizations. Implementation of metrically based segmentation in Shortlist could, therefore, be accomplished for French, for example, by invoking the segmentation procedure at syllable boundaries; for Japanese it could be accomplished by invoking the segmentation procedure at mora boundaries. In either case, the model could be run with no significant alteration on the appropriate language-specific dictionary and input. Most important, all of the present simulations have used a realistically sized dictionary. Whatever the language, the present model is one that captures word-recognition effects, both of competition and of segmentation, on a scale appropriate to those actually experienced by listeners.

References

Altmann, G. T. M., & Carter, D. M. (1989). Lexical stress and lexical discriminability: Stressed syllables are more informative, but why? *Computer Speech and Language, 3,* 265–275.

Bard, E. G. (1990). Competition, lateral inhibition, and frequency: Comments on the chapters of Frauenfelder and Peeters, Marslen-Wilson, and others. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 185–210). Cambridge, MA: MIT Press.

Bard, E. G., Shillcock, R. C., & Altmann, G. T. M. (1988). The recognition of words after their acoustic offsets in spontaneous speech: Effects of subsequent context. *Perception & Psychophysics, 44,* 395–408.

Burnage, G. (1990). *CELEX: A guide for users.* Nijmegen, The Netherlands: CELEX.

Cluff, M. S., & Luce, P. A. (1990). Similarity neighborhoods of spoken two-syllable words: Retroactive effects on multiple activation. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 551–563.

Cole, R. A., & Jakimik, J. (1978). Understanding speech: How words are heard. In G. Underwood (Ed.), *Strategies of information processing* (pp. 67–116). London: Academic Press.

Cole, R. A., & Jakimik, J. (1980). A model of speech perception. In R. A. Cole (Ed.), *Perception and production of fluent speech* (pp. 133–163). Hillsdale, NJ: Erlbaum.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language, 31,* 218–236.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language, 2,* 133–142.

Cutler, A., & Clifton, C. E. (1984). The use of prosodic information in word recognition. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 183–196). Hillsdale, NJ: Erlbaum.

Cutler, A., Mehler, J., Norris, D. G., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language, 25,* 385–400.

Cutler, A., Mehler, J., Norris, D. G., & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology, 24,* 381–410.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 113–121.

Cutler, A., & Otake, T. (1994). Mora or phoneme? Further evidence for language-specific listening. *Journal of Memory and Language, 33,* 824–844.

Frauenfelder, U. H. (1991). Lexical alignment and activation in spoken word recognition. In J. Sundberg, L. Nord, & R. Carlson (Eds.), *Music, language, speech and brain: Wenner-Gren International Symposium Series* (pp. 294–303). London: Macmillan.

Frauenfelder, U. H. (1993, August–September). *Modelling the time-course of spoken word recognition.* Paper presented at the International Conference on the Psychology of Language and Communication, Glasgow.

Freedman, D. (1992). *Frequency effects in auditory word recognition using the word-spotting task.* Unpublished undergraduate project, Department of Experimental Psychology, University of Cambridge, Cambridge, United Kingdom.

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language, 28,* 501–518.

Goldinger, S. D., Luce, P. A., Pisoni, D. B., & Marcario, J. K. (1992). Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1211–1238.

Grosjean, F. (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception & Psychophysics, 38,* 299–310.

Johansson, S., & Hoflund, K. (1989). *Frequency analysis of English vocabulary and grammar* (Vol. 1). Oxford, England: Clarendon.

Lehiste, I. (1972). The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America, 51,* 2018–2024.

Luce, P. A. (1986). A computational analysis of uniqueness points in auditory word recognition. *Perception & Psychophysics, 39,* 155–158.

Luce, P. A., Pisoni, D. B., & Goldinger, S. D. (1990). Similarity neighborhoods of spoken words. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 122–147). Cambridge, MA: MIT Press.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25,* 71–102.

Marslen-Wilson, W. D. (1990). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148–172). Cambridge, MA: MIT Press.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10,* 29–63.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1–86.

McQueen, J. M., & Cutler, A. (1992). Words within words: Lexical statistics and lexical access. In *Proceedings of the International Conference on Spoken Language Processing* (Vol. 1, pp. 221–224). Banff, Alberta, Canada: University of Alberta.

McQueen, J. M., Norris, D. G., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 621–638.

Mehler, J., Dommergues, J.-Y., Frauenfelder, U. H., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior, 20,* 298–305.

Nakatani, L. H., & Dukes, K. D. (1977). Locus of segmental cues for word juncture. *Journal of the Acoustical Society of America, 62,* 714–719.

Norris, D. G. (1990). A dynamic-net model of human speech recognition. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 87–104). Cambridge, MA: MIT Press.

Norris, D. G. (1991). Rewiring lexical networks on the fly. In *Proceedings of Eurospeech '91* (Vol. 1, pp. 117–120). Genova, Italy: European Speech Communication Association with Istituto Internazionale delle Comunicazioni.

Norris, D. G. (1994a). A quantitative model of reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1212–1232.

Norris, D. G. (1994b). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52,* 189–234.

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language, 32,* 258–278.

Procter, P. (Ed.). (1975). *Longman dictionary of contemporary English.* London: Longman.

Shillcock, R. C. (1990). Lexical hypotheses in continuous speech. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 24–49). Cambridge, MA: MIT Press.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18,* 645–659.

Swinney, D. A. (1981). Lexical processing during sentence comprehension: Effects of higher order constraints and implications for representation. In T. Myers, J. Laver, & J. Anderson (Eds.), *The cognitive representation of speech* (pp. 201–209). Amsterdam: North-Holland.

Vroomen, J., & de Gelder, B. (1995). Metrical segmentation and lexical inhibition in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 98–108.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition, 32,* 25–64.

Appendix
Experimental Materials for Each of the Six Competitor-Initial Consonants

Number of words beginning in the second syllable						
Many				Few		
Competitor-initial consonant	Stress pattern	Target structure C(C)VCC	Target structure C(C)VC	Competitor-initial consonant	Stress pattern	Target structure C(C)VCC	Target structure C(C)VC

Competitor-initial consonant	Stress pattern	C(C)VCC	C(C)VC	Competitor-initial consonant	Stress pattern	C(C)VCC	C(C)VC
/p/				/t/			
	SS	STAMPidge	PRAMpidge		SS	MINTaup	THINtaup
	SW	STAMP@dge	PRAMp@dge		SW	MINT@p	THINt@p
	SS	CHIMPidge	DIMpidge		SS	MELTaudge	GELtaudge
	SW	CHIMP@dge	DIMp@dge		SW	MELT@dge	GELt@dge
	SS	STUMPol	DRUMpol		SS	LILTaug	PILLtaug
	SW	STUMP@l	DRUMp@l		SW	LILT@g	PILLt@g
	SS	SHRIMPol	SLIMpol		SS	HINTeig	SINteig
	SW	SHRIMP@l	SLIMp@l		SW	HINT@g	SINt@g
	SS	CLUMPaedge	GLUMPaedge		SS	PELTeish	YELLteish
	SW	CLUMP@dge	GLUMP@dge		SS	PELT@sh	YELLt@sh
	SS	SCALPaeb	GALPaeb		SS	LIFTeig	STIFFteig
	SW	SCALP@b	GALP@b		SW	LIFT@g	STIFFt@g
	SS	PULPoash	DULLpoash		SS	FRONTeish	GUNteish
	SW	PULP@sh	DULLp@sh		SW	FRONT@sh	GUNt@sh
	SS	PUMPul	SCUMpul		SS	GRUNTeithe	NUNteithe
	SW	PUMP@l	SCUMp@l		SW	GRUNT@the	NUNt@the
/k/					SS	FLINTuth	GRINtuth
	SS	RISKom	KISSkom		SW	FLINT@th	GRINt@th
	SW	RISK@m	KISSk@m		SS	GIFTudge	CLIFFtudge
	SS	TASKodge	GRASSkodge		SW	GIFT@dge	CLIFFt@dge
	SW	TASK@dge	GRASSk@dge		SS	JOLTul	GOALtul
	SS	TUSKom	TRUSSkom		SW	JOLT@l	GOALt@l
	SW	TUSK@m	TRUSSk@m	/d/			
	SS	DISKael	MISSkael		SS	CHILDap	VILEdap
	SW	DISK@l	MISSk@l		SW	CHILD@p	VILEd@p
	SS	DESKael	MESSkael		SS	BLINDaethe	BRINEdaethe
	SW	DESK@l	MESSk@l		SW	BLIND@the	BRINEd@the
	SS	HUSKaethe	PLUSKaethe		SS	BLANDarthe	FLANdarthe
	SW	HUSK@the	PLUSk@the		SW	BLAND@the	FLANd@the
	SS	DUSKoadge	FUSSkoadge		SS	FIENDardge	QUEENdardge
	SW	DUSK@dge	FUSSk@dge		SW	FIEND@dge	QUEENd@dge
	SS	FLASKoash	GLASSkoash		SS	RINDarsh	VINEdarsh

(table continues)

Appendix (continued)

		Many				Few	
		Number of words beginning in the second syllable					
Competitor-initial consonant	Stress pattern	Target structure		Competitor-initial consonant	Stress pattern	Target structure	
		C(C)VCC	C(C)VC			C(C)VCC	C(C)VC
	SW	FLASK@sh	GLASSk@sh		SW	RIND@sh	VINEd@sh
	SS	MASKuk	PASSkuk		SS	GLANDauk	CLANdauk
	SW	MASK@k	PASSk@k		SW	GLAND@k	CLANd@k
/s/				/tʃ/			
	SS	SENSEul	GLENsul		SS	TRENCHoib	FENchoib
	SW	SENSE@l	GLENs@l		SW	TRENCH@b	FENch@b
	SS	PULSEim	GULLsim		SS	MUNCHoab	FUNchoab
	SW	PULSE@m	GULLs@m		SW	MUNCH@b	FUNch@b
	SS	RINSEib	BINsib		SS	LUNCHaithe	TONchaithe
	SW	RINSE@b	BINs@b		SW	LUNCH@the	TONch@the
Mean frequency		28.7	32.5			24.4	21.6

Note. The target words are shown in uppercase, and the second syllables have orthographic transcriptions, with the vowel schwa being represented by @. Also shown are the mean frequencies of occurrence of the target words as given in Johansson and Hoflund (1989). C = consonant; V = vowel; SS = strong–strong; SW = strong–weak.