Journal of Experimental Psychology:
Human Perception and Performance
1995, Vol. 21, No. 1, 98-108

Metrical Segmentation and Lexical Inhibition in Spoken Word Recognition

Jean Vroomen
Tilburg University

Beatrice de Gelder
Tilburg University and Université Libre de Bruxelles

Two cross-modal priming experiments investigated whether speech segmentation is based on the occurrence of strong syllables (A. Cutler & D. Norris, 1988) or lexical competition (J. L. McClelland & J. L. Elman, 1986). Auditorily presented words with no, few, or many competitors served as prime for a visual target. Facilitatory effects were larger for primes with no or few competitors than for primes with many competitors. This difference disappeared when the interstimulus interval between the prime and target was shortened. The differential priming effects are interpreted as evidence for inhibition between competing lexical candidates.

Understanding spoken language involves the identification of the words that make up an utterance. Because the speech signal is continuous, listeners must segment a spoken utterance into words or into some smaller unit from which words can then be recognized. The most tangible dimension of this problem is that, at present, research has not been able to isolate discrete entities within running speech that correspond to segments. Admittedly, the crux of the difficulty lies in the fact that speakers do not provide reliable acoustic cues to indicate boundaries between words. Several decades of speech research have not yet led to a widely accepted solution. It therefore remains unclear how to start lexical access in the absence of reliable word boundary cues.

A possibility adopted by some psychological models of spoken word recognition is that the speech signal is segmented or classified into some intermediate prelexical linguistic unit, but among psycholinguists there is no consensus about whether this unit should be the phoneme, the syllable, the foot, or another relevant linguistic unit. For instance, several authors have claimed that speech is segmented into syllable-size units (for an overview, see Segui, Dupoux, & Mehler, 1990). The basic idea is that a lexical access attempt is initiated at the beginning of each syllable. The benefit of such a procedure would be that the majority of lexical access attempts would be successful (at least if one wants to maintain that syllable boundaries in continuous speech coincide with word boundaries). The seminal study by Mehler, Dommergues, Frauenfelder, and Segui (1981) provided empirical evidence for syllable-based segmentation. Listeners were faster to detect a segment if it corresponded exactly to the first syllable of a word, rather than if it comprised more or less than the syllable. Thus, subjects were faster to detect *ba* in *ba.llance* (the dot indicates the syllable boundary) than in *bal.con,* and they were faster to detect *bal* in *bal.con* than in *ba.llance.* However, one aspect complicating the prelexical segmentation issue is that linguistic variation plays an important role in the sense that perceptual procedures, and hence experimental results, may depend on the listener's native language. Thus, the above-mentioned experiments were run with native French speakers and French stimuli. Subsequent studies showed that this pattern of results did not hold up in English (Cutler, Mehler, Norris, & Segui, 1983, 1986). For English listeners, it did not matter whether the target did or did not correspond to the syllable of a word.

Cutler et al. (1986) ascribed the asymmetric results to phonological differences between French and English. A major phonological contrast between English and French that might be critical is the fact that English is a stress language with diverse syllable structures and speakers' intuitions about syllable boundaries are often vague. In contrast, French has less diverse syllable structures, and syllable boundaries are clearer. Cutler et al. argued that these factors made the syllable an appropriate segmentation unit for French but not for English. The proposal was that segmentation procedures of English listeners are tuned to the specifics of the English phonology. A major aspect of the English phonology is the metrical distinction between strong and weak syllables. Strong syllables have full, unreduced vowels, whereas weak syllables have reduced vowels that are usually realized as schwa. Cutler and Norris (1988) proposed the metrical segmentation strategy (MSS), which proposes that English listeners start a lexical access attempt

---

Jean Vroomen, Department of Psychology, Tilburg University, Tilburg, The Netherlands; Beatrice de Gelder, Department of Psychology, Tilburg University, The Netherlands, and Department of Psychology, Université Libre de Bruxelles, Brussels, Belgium.

The research was partly supported by a grant from the Human Frontier of Science Programme, Processing Consequences of Contrasting Language Phonologies. The research of Jean Vroomen has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences. We thank Dennis Norris for making available both the original SHORTLIST model (Norris, 1994) and the model adapted to include the MSS (McQueen, Norris, & Cutler, 1994). Antal van den Bosch was helpful in running the simulations. We wish to thank Theo Popelier and Monique van Zon for testing the study participants and Anne Cutler and James McQueen for valuable comments made on an earlier version of this article. We also thank the Institute for Perception Research for use of their studio facilities.

Correspondence concerning this article should be addressed to Jean Vroomen, Department of Psychology, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. Electronic mail may be sent via Internet to j.vroomen@kub.nl.

at the beginning of every strong syllable. Lexical words (i.e., content words, excluding functors) are expected to start just before the onset of such strong syllables.

The lexical statistics of the English vocabulary do indeed suggest that the success rate of the MSS will be quite high: Lexical words begin three times as often with strong syllables, and these words are twice as frequent as those beginning with weak syllables (Cutler & Carter, 1987). The empirical evidence for the MSS comes from two types of studies, "slips of the ears" and word spotting. With respect to the first, Cutler and Butterfield (1992) examined misperceptions of the locations of word boundaries in continuous speech. They presented unpredictable utterances to listeners at a level that was just above their threshold for speech perception. These barely audible sentences consisted of strings of alternating strong and weak syllables (e.g., *conduct ascents uphill*). The results of this study showed that listeners had a strong tendency to place erroneous word boundaries before strong syllables and to delete word boundaries before weak syllables (e.g., *conduct ascents uphill* → *the doctor sends her bill*). Thus, as predicted by the MSS, the study showed that listeners do rely on a strategy of assuming that strong syllables mark the beginning of lexical words.

The second line of evidence comes from word-spotting experiments (Cutler & Norris, 1988). In this task, listeners monitor for CVC (e.g., *thin*) or CVCC (e.g., *mint*) words (where C indicates a consonant and V indicates a vowel) that are embedded in bisyllabic pseudowords that end in either a strong (e.g., *thintayf* or *mintayf*) or a weak (e.g., *thintef* or *mintef*) syllable. In the case of a strong syllable (*thintayf* and *mintayf*), the MSS predicts that the word is segmented as *thin_tayf* and *min_tayf,* whereas in the case of a weak syllable, there is no segmentation at all (*thintef* and *mintef*). In line with these predictions, the results showed that CVCC words such as *mint* were harder to detect in *mintayf* than in *mintef,* whereas there was no difference for CVC words such as *thin* in *thintayf* and *thintef.* The CVCC target *mint* in *mintayf* was divided across the two segmentation units into *min_t,* with the impeding consequence that speech had to be assembled across a segmentation boundary. For CVC words (*thin*), there was no difference between *thintayf* and *thintef,* because the segmentation trigger in *thintayf* did not penetrate *thin* (*thin_tayf*).

Although there are statistical as well as empirical facts that favor the MSS, one might also consider an alternative account that does not invoke a device specifically dedicated to perform prelexical speech segmentation. One might argue that the previously mentioned word-spotting results reflect lexical inhibition effects as, for instance, instantiated in the TRACE model of spoken word recognition (McClelland & Elman, 1986). In TRACE, inhibition among lexical candidates depends on the number of phonemes that lexical items share within the same time slices. Thus, a word like *mint* in *mintayf* will be inhibited by words starting with *tay* or *tayf,* because these words are competing for */t/*. There is thus competition at the lexical level for the proper assignment of the acoustic input. A critical observation, as noted earlier in this section, is that most lexical words in English start with strong syllables. Hence, it might be the case that lexical competition for a word like *mint* in *mintayf* is larger than that for *mint* in *mintef* because there are more words that start with *tay* or *tayf* than with *te* or *tef.* On the other hand, there are many different strong vowels (like the vowels in *hat, go, bee, hit, look,* etc.) but essentially only one weak vowel (schwa). Therefore, before predictions about the number of competitors can be made, a more fine-grained analysis of the lexical statistics would be required. Alternatively, however, it might be that strong vowels as in *tay* are acoustically more salient or transparent than weak vowels as in *te.* Hence, words starting with *tay* might be activated earlier or become more highly activated than words starting with *te.* Either way, *mint* in *mintayf* might be more strongly inhibited via lexical inhibition than *mint* in *mintef.* For CVC words such as *thin* in *thintayf* or *thintef,* there are no differential lexical inhibition effects, because words starting with *tay(f)* and *te(f)* do not overlap with *thin.* The amount of lexical inhibition *thin* receives from *thintayf* or *thintef* is thus the same. If this is so, then differential lexical inhibition effects for CVCC words, although not for CVC words, might be responsible for the word-spotting results.

Similarities Between English and Dutch

The present study was set up to investigate the separate or combined effects of the accounts given by the MSS and TRACE-like lexical competition. Before the actual study could be run, however, several issues had to be addressed. As the experiments to be reported were run in Dutch, the extent of the similarities and contrasts with English phonology had to be determined to assess whether reliance on the MSS might be a likely strategy for listeners of Dutch. Dutch is an interesting test case to compare with English, because both languages are related in terms of rhythmic structure. Dutch, like English, has various syllable structures (up to CCCVCCC syllables, as in *strengst, most strict*), and many syllables have opaque syllable boundaries (e.g., *ba[ll]et,* where the *[ll]* is an ambisyllabic consonant that belongs to both syllables). For phonological reasons similar to those given for English, Dutch does not seem to be a good candidate for a French-like syllabic segmentation procedure. Nevertheless, syllabic effects have been reported by Zwitserlood, Schriefers, Lahiri, and van Donselaar (1993), who found in a segment-monitoring task that segment-detection latencies were shorter if the target matched the first syllable of a spoken Dutch word.

Another important aspect of Dutch is whether its lexical statistics do indeed support the MSS such that betting on strong syllables would be a good strategy for lexical access. To answer this question, we computed word counts on the Dutch Center for Lexical Information (CELEX) lexicon and also, for comparison, on the English CELEX lexicon. There is a difference between English and Dutch in the definition of strong and weak syllables that might be relevant here. In English, most weak syllables contain a schwa, or they contain a short vowel followed by a syllable with primary stress. These criteria cover most instantiations of weak syllables in English. In Dutch, however, vowel reduction of unstressed short vowels is not as compulsory as in English. In some cases it is optional (e.g., *baNAAN, banana* [capital letters indicate primary stress], may be realized as *baNAAN* or *b@NAAN*), but in other, comparable cases it is rather uncommon to reduce the short unstressed vowel (*baNIER* is almost always realized as *baNIER* and not *b@NIER*; examples were taken from van Bergem, 1993). In the CELEX dictionary, all these examples are coded in their citation form, that is, with their full vowel (thus *banaan* and not *b@naan*). Because we have, at present, no databases of Dutch continuous speech from which the occurrence of these reductions can be estimated, we adopted for Dutch a strict criterion such that syllables are counted as weak only if the vowel in the CELEX dictionary is schwa. To facilitate the cross-linguistic comparison, however, we also used the English criterion for Dutch and vice versa. The word counts were computed on all lexical words including their derivations, but inflections, compounds, and derivational compounds were excluded. Words that have more than one pronunciation (particularly in English) or more than one spelling (particularly in Dutch) were counted only once in their citation form. Table 1 summarizes the results.

As can be seen in Table 1, in English, using the English criterion, 18.6% of lexical words start with a weak syllable. This is slightly less than the 27.0% and 22.4% that were reported by Cutler and Carter (1987). In Dutch, according to the Dutch criterion, this percentage is even lower: It is only 12.3%. Thus, at least according to these statistics, it seems that the MSS would be a good strategy for Dutch listeners. It should be mentioned, however, that there is a difference in the phonological structure of the weak initial syllables. In English, it is quite diverse: It can be a single V (*a-buse*), a CV (*sa-loon*), a VC (*ob-scure*), a CCV (*pro-tect*), or a CVC syllable (*com-bine*). In Dutch the situation is quite different: The overwhelming majority of weak initial items (i.e., 96%) are affixed words with the prefixes *be-*, *ge-*, or *ver-*. Whether this, or any other difference, plays a role in spoken word recognition remains to be determined empirically. The crucial aspect for the plausibility of the MSS-based approach, however, is that in both languages, strong syllables are more likely to be at the onset of lexical words than are weak syllables. Moreover, in English and Dutch, words with strong initial syllables are more than twice as frequent as words with weak initial syllables. Thus, it can be expected that the great majority of lexical words a Dutch listener encounters will have a strong syllable in the word-initial position.

The next question, then, is whether Dutch listeners do actually apply, at least to a first approximation, an MSS-like strategy. Thus, it is important to know whether the results of the English slips-of-the-ear (Cutler & Butterfield, 1992) and word-spotting studies (Cutler & Norris, 1988) can be replicated, this time using Dutch stimuli and listeners. For that purpose, we conducted two studies with a setup similar to the English experiments. The results from word spotting as well as from word boundary misperceptions provided positive evidence for the MSS (van Zon & de Gelder, 1993; Vroomen & de Gelder, 1993). For the investigation of word boundary misperceptions, we presented barely audible utterances made up of strong and weak syllables. As predicted by the MSS, insertions before strong syllables and deletions before weak syllables were far higher than could be expected from chance. A similar picture favoring the MSS emerged from the word-spotting studies. In this case, listeners monitored for CVCC (*melk,* milk) or CVC (*bel,* bell) words that were embedded in bisyllabic pseudowords that ended in strong (*melkoos* and *belkoos*) or weak (*melkes* and *belkes*) syllables. As in English, it was harder to detect *melk* in *melkoos* than in *melkes,* whereas, in fact, the opposite pattern was observed for CVC words. We can therefore conclude that Dutch resembles English in several aspects: There are statistical and empirical reasons to believe that strong syllables have a special status in both languages. The nature of the status, however, remains to be investigated.

The Present Study

Having established that relevant similarities exist between English and Dutch, we return to the main objective of the present study, which is to investigate the separate or combined effects of the account given by the MSS and the one given by lexical competition. For that purpose, we needed to trace the lexical activation of words as they were spoken in time. The cross-modal priming paradigm (Swinney, 1979) appeared to be the most suited, as it can provide an on-line measure of levels of lexical activation. In cross-modal priming, study participants listen to a spoken word or sentence and simultaneously see a visual probe word that, in the critical condition, is related to the spoken word of interest.

Table 1
Metrical Characteristics of Word-Initial Syllables of Lexical Words and Their Frequency of Occurrence in the CELEX Database

	Dutch					English				
	Strong initial syllable		Weak initial syllable			Strong initial syllable		Weak initial syllable		
Criterion for "weak"	*N*	*f*	*N*	*f*	Proportion of weak-initial words	*N*	*f*	*N*	*f*	Proportion of weak-initial words
Dutch	26,888	12.8	3,802	6.4	**.123**	21,569	23.2	1,724	13.3	.074
English	24,956	13.5	5,735	5.7	.186	18,939	25.3	4,354	10.0	**.186**

Note. CELEX = Centre for Lexical Information, University of Nijmegen, The Netherlands; N = number.

The perception of the spoken word is intended to prime the visual probe. Listeners make a lexical decision in response to the visual probe, and the reaction time is compared with a control condition in which the spoken word has no relation to the probe. The difference in reaction time between the control and experimental conditions provides a measure for the lexical activation of the probe.

The major issue in the present study was to trace the lexical activation of words that were followed by either a strong or a weak syllable. The factors that were varied were whether a CVC (*bel*) or a CVCC (*melk*) word was followed by a strong or a weak syllable. In the case of a strong syllable, we were able to vary the cohort size of words that start with the strong syllable at its onset. Take as an example the words *bel* and *melk*: When these words were followed by a weak syllable (henceforth SW, i.e., strong–weak), these words were embedded in the bisyllabic nonsense string *belkem* and *melkem.* These nonsense words were heard while subjects made a lexical decision to the visual probe BEL or MELK. In the case of a strong second syllable (SS), the auditory primes were *belkeum* or *melkeum,* and *belkaam* or *melkaam.* The critical difference between the endings *keum* and *kaam* is the number of words that start with *keu(m)* or *kaa(m),* that is, the cohort size. According to the CELEX database, there are only 124 words in Dutch that start with *keu* (and zero words that start with *keum*), whereas there are 865 words that start with *kaa* (and 114 that start with *kaam*). Thus, there are many more words starting with *kaam* than with *keum* (we call the former condition SS-large; the latter, SS-small). If lexical competition plays a role, one should observe a difference between these two strong endings. More specifically, priming effects for CVCC words (*melk*) are expected to be largest for *melkem,* medium for *melkeum,* and smallest for *melkaam,* because there are no words that start with *ke(m),* few words that start with *keu(m),* and many words that start with *kaa(m).* Lexical competition is thus expected to increase in proportion with the cohort size of the competitors. In contrast, according to the MSS, there should be no difference between the SS-small and SS-large primes, because in both cases a segmentation trigger is set within the CVCC word that will be equally detrimental for the activation of the target. A different pattern of results is expected for CVC words such as *bel*: In this case, both hypotheses predict no differences among the SW, SS-small, and SS-large conditions. In the lexical inhibition framework, competition for /k/ in *belkem, belkeum,* and *belkaam* is irrelevant for the activation of *bel,* because there is no overlap. The MSS also predicts that there should be no differences among *belkem, belkeum,* and *belkaam,* because no segmentation trigger is set within *bel.* The CVC words therefore serve as a control condition for the general effects that the different weak and strong endings might have on the activations of the target words.

One further aspect that can be added to this picture is the possibility of frequency effects of the target. In the framework of lexical inhibition, it is possible that low-frequency targets suffer more from inhibitory influences than high-frequency targets, and this will be most pronounced in the SS-large condition. Such an effect may be observed because low-frequency targets will be less able to inhibit their competitors as compared with a high-frequency counterpart (see Bard, 1990, for simulations with TRACE on this issue). There might thus be an interaction between frequency of the target and cohort size of the competitors. Whether this interaction is actually observed, however, does depend on many parameters. Not only is there a function that relates activation to inhibition, but there is also a relation between frequency and activation, between prior activation and frequency, and among limits on activation and inhibition, decays, thresholds, biases, and so forth. One needs to set these parameters before quantitative predictions can be made. It would therefore be too strong a claim to say that the lexical inhibition hypothesis predicts an interaction between frequency of target and cohort size of competitors. Instead, it is better to say that it can account for a certain type of interaction. In contrast, the MSS is not explicit about frequency effects, because it is operational at a prelexical level.

Experiment 1

Method

Study participants. Forty native speakers of Dutch participated in the experiment, for which they were paid a small amount. Participants' vision was normal or corrected to normal.

Materials. We chose 96 words, 48 with CVCC and 48 with CVC structure. The words formed pairs like *melk* or *bel* such that (a) both words had the same short vowel, (b) both had the same consonant in the postvocalic position (in this example the *l*), and (c) neither word could be made into another word by adding or removing the final consonant of the CVCC word (*belk* and *mel* are both nonwords). All words were made into bisyllabic nonwords by the addition of an extra syllable. Three alternative endings were constructed: One had a weak vowel (schwa; SW), one had a strong vowel with a small cohort size (SS-small), and one had a strong vowel with a large cohort size (SS-large). The final consonant was constant within each triple. For the examples given above, the endings were *(k)em, (k)eum,* and *(k)aam,* yielding *melkem, melkeum,* and *melkaam,* and *belkem, belkeum,* and *belkaam* for the SW, SS-small, and SS-large conditions, respectively. All bisyllables had stress on the first syllable. The complete set of materials is presented in the Appendix.

The cohort sizes of the endings were determined with the CELEX database. Two cohort sizes for each of the two final strong syllables were computed: One was based on the initial CV and the other on the CVC. Thus, for the items *belkeum* and *melkeum* or *belkaam* and *melkaam,* cohort sizes were computed for *keu* and *keum* or *kaa* and *kaam*, respectively. The mean cohort sizes of the SS-small condition were 96.5 and 3.6 for CV and CVC onset, respectively, and for the SS-large they were 592.9 and 71.1 for CV and CVC onset, respectively. The cohort size of the SW condition is effectively zero, because there are no words in the CELEX lexicon that start with an unvoiced consonant followed by schwa.

The auditory words served as primes for a visual target. Pairing of primes and targets resulted in eight conditions, Four for CVCC visual targets and four for CVC targets (see Table 2). Condition 1 was a control condition for the CVCC targets. Listeners heard an unrelated CVCC word followed by a weak syllable (e.g., *last-em*). The control word was semantically and phonologically unrelated to the visual target (e.g., MELK), but the frequency of occurrence

Table 2
Sample of the Stimulus Set

Visual target/Condition	Spoken prime
MELK	
Control	*lastem*
SW	*melkem*
SS-small	*melkeum*
SS-large	*melkaam*
BEL	
Control	*wankem*
SW	*belkem*
SS-small	*belkeum*
SS-large	*belkaam*

Note. SW = strong-weak syllable; SS = strong second syllable.

of the control and target words was matched as closely as possible. In Condition 2 (SW), listeners heard a prime that was identical to the CVCC target. The prime was followed by a weak syllable (e.g., *melkem* for the visual target MELK). In Condition 3 (SS-small), listeners heard the target followed by a strong syllable with a small cohort size (*melkeum*), and in Condition 4 (SS-large), the target word was followed by a strong syllable with a large cohort size (*melkaam*). The same pairings were also constructed for the CVC targets.

Design and procedure. The auditory primes were recorded by a male speaker of Dutch in a professional sound-treated studio. The items were recorded on digital audiotape (Philips DAT850) and later digitized at 22.1 kHz (16 bits). The acoustic offset of each of the embedded words (for instance, at the /k/ in *melkem* or at the /l/ in *belkem*) was determined auditorily and visually from the time–amplitude waveform. This offset point served as reference for the interstimulus interval (ISI). In Experiment 1, the ISI was set at 250 ms, so that at approximately the end of the second syllable a visual target appeared on a cathod-ray tube screen. A counter module was started at the onset of the visual target to register the participant's reaction time and response key. The target was presented unmasked for 50 ms. A fixating point, located 2 cm under the center of the target, appeared 100 ms before the onset of the visual target and remained there for 50 ms. The auditory primes were presented at a comfortable listening level via a Sennheiser HD410-SL headphone. The intertrial interval was 3 s.

Four different versions of the stimulus set were made, such that each visual target appeared only once in the experiment. The prime type of an item pair was fixed within a version. Thus, *melkem* and *belkem* were presented in one version, *melkeum* and *belkeum* in another, and so forth. Each version contained twelve different items of each condition for a total of 96 experimental trials per version. There were an additional 192 bisyllabic filler items that were the same in all four versions. Half of the spoken filler primes started with a real Dutch CVC word and the other half started with a CVCC word. Over the complete set of word primes (fillers and experimental words), one third ended in a weak vowel (SW), one third ended in a strong vowel (SS) that was taken from syllables with small cohort sizes, and one third ended in strong vowels that were taken from syllables with large cohort sizes. Each member of an experimental quadruplet and all fillers appeared in exactly the same location in each of the four versions.

Half of the visual targets were nonwords ("no" decision), and half were real Dutch words ("yes" decision). The nonwords were phonotactically correct pronounceable CVC or CVCC pseudowords. Half of the nonwords were phonologically related to the prime in that they shared two or three phonemes. This was done to prevent participants from adopting a strategy by which any amount of phonological overlap between prime and target could result in a yes decision. Over the complete set of experimental and filler items, half of the primes were phonologically related to the visual target (for the experimental yes responses, the target was embedded in the prime; for the no responses the target shared two or three phonemes), and half of the prime–target pairings were phonologically unrelated. Testing lasted about 20 min, with a short pause halfway through the testing. A practice session of 24 trials preceded the experiment. Participants were tested individually in a soundproof booth and were asked to make a speeded lexical decision about the visual target by pressing a "yes" or a "no" key.

Results

Error rates for the items were inspected first. Four CVC items were missed (i.e., a response of no to a real word) by more than 30% of the experiment participants. These items, plus their matched CVCC pair, were discarded from the analyses. The overall error rate for the remaining items was 6%. An additional 4% of extreme values, defined as reaction times that were more than 2 standard deviations outside the relevant participant or item means, was discarded. Errors and extreme values were replaced by the mean of the item or participant for the relevant condition. Analyses of variance (ANOVAs) were performed with items and participants as random factors. The overall results are presented in Table 3. As can be seen, priming effects occurred in the exact direction predicted by the notion that lexical inhibition plays a role in spoken word recognition: Priming effects in CVCC words are largest for *melkeum,* medium for *melkeum,* and smallest for *melkaam.* For CVC words, there are no differences among the priming effects of *belkem, belkeum,* and *belkaam.*

A 2 (target type) × 4 (prime type) ANOVA on the reaction times showed that there was no overall difference between CVCC or CVC targets, $F_1(1, 39) = 3.05, p = .09$; $F_2 < 1$. The effect of the prime was significant, $F_1(3, 117) = 16.39, p < .001$; $F_2(3, 129) = 14.21, p < .001$, as was the interaction between target type and prime, $F_1(3, 117) = 3.42, p < .02$; $F_2(3, 129) = 3.65, p < .02$. To

Table 3
Mean Reaction Times (RTs) and Priming Effects for CVCC and CVC Targets With an Interstimulus Interval of 250 ms

Condition	Spoken prime	Visual target	RT (ms)	Priming
		CVCC words		
Control	*lastem*	MELK	621	
SW	*melkem*	MELK	578	43
SS-small	*melkeum*	MELK	589	33
SS-large	*melkaam*	MELK	602	19
		CVC words		
Control	*wankem*	BEL	621	
SW	*belkem*	BEL	599	23
SS-small	*belkeum*	BEL	600	21
SS-large	*belkaam*	BEL	595	27

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

investigate whether the priming effects were different for the three types of context (SW, SS-small, and SS-large), we performed ANOVAs on the priming effects, with target type and context as within-participants variables. The amount of priming was computed by subtracting the reaction time of the target-embedded prime words from the appropriate control condition. A 2 (target type) × 3 (context) ANOVA on the priming effects showed that the overall difference between CVCC and CVC targets was not significant by participants ($F_1 < 1$), but it was significant by items, $F_2(1, 43) = 5.53$, $p < .03$. There was no overall effect of the context, $F_1(2, 78) = 2.29$, $p = .11$; $F_2(2, 86) = 2.12$, $p = .13$, but the interaction between target type and context was significant by participants, $F_1(2, 78) = 7.42$, $p < .001$, and marginally significant by items, $F_2(2, 86) = 2.61$, $p = .08$. Separate ANOVAs on the priming effects of CVCC and CVC targets showed that the priming effects of CVCC targets were different for the three types of contexts, $F_1(2, 78) = 8.88$, $p < .001$; $F_2(2, 86) = 4.27$, $p < .02$. There was no difference in the priming effects of CVC targets (F_1 and $F_2 < 1$). Planned comparisons of the priming effects of CVCC targets indicated that SW primes (*melkem*) were faster than SS-large primes (*melkaam*) by 24 ms, $F_1(1, 39) = 17.47$, $p < .001$; $F_2(1, 43) = 6.63$, $p < .02$; that SW primes (*melkem*) were faster than SS-small primes (*melkeum*) by 10 ms, which was significant by participants, $F_1(1, 39) = 4.34$, $p < .05$, but not by items ($F_2 < 1$); and finally, that SS-small primes (*melkeum*) were faster than SS-large primes (*melkaam*) by 14 ms, $F_1(1, 39) = 4.34$, $p < .04$; $F_2(1, 43) = 5.31$, $p < .03$. Thus, priming effects of CVCC targets were proportionate with the cohort sizes of the competitors.

The following analyses were concerned with frequency effects of CVCC targets. As already mentioned, if lateral inhibition among lexical candidates plays a role, there might be an interaction between frequency of the target and cohort size of the competitors. Low-frequency targets might suffer more from the SS-large competitors than high-frequency targets, because low-frequency targets are less able to inhibit their competitors. To investigate that issue, we split the CVCC item set into low-frequency and high-frequency targets, with the criterion for frequency of occurrence set at an arbitrary 50 per million. There were 30 low-frequency targets (mean frequency = 11.67; cohort sizes for SS-small = 96.5 and 3.6, and for SS-large = 592.9 and 71.0 for CV and CVC cohorts, respectively) and 14 high-frequency targets (mean frequency = 176.71 and cohort sizes for SS-small = 73.4 and 2.0, and for SS-large = 601.1 and 104.5 for CV and CVC cohorts, respectively). Mean reaction times and priming effects for low- and high-frequency targets in the different competitor environments are presented in Table 4. As can be seen, the interaction was in the predicted direction: Priming effects of low-frequency targets suffered more than those of high-frequency targets from the SS-large environment.

A 2 (frequency) × 4 (prime type) ANOVA on the reaction times showed that there was a significant effect of prime type, $F_1(3, 117) = 18.01$, $p < .001$; $F_2(3, 126) = 13.03$, $p < .001$, and of frequency, $F_1(1, 39) = 34.68$, $p < .001$; $F_2(1, 42) = 6.70$, $p < .02$, but the interaction between prime type and frequency did not reach significance, $F_1(3, 117) = 1.67$, $p = .17$; $F_2(3, 126) = 1.26$, $p = .29$. However, although the overall interaction was not significant, planned comparisons on the priming effects did show the predicted interaction: For low-frequency targets, the 27-ms difference between SS-small and SS-large conditions was significant, $F_1(1, 39) = 9.24$, $p < .005$; $F_2(1, 29) = 8.76$, $p < .01$, but the 5-ms difference for high-frequency targets was not (both Fs < 1). The results thus point in the direction that priming effects of low-frequency targets suffer more from competitors than those of high-frequency targets.

Table 4
Mean Reaction Times (RTs) and Priming Effects for Low- and High-Frequency CVCC Targets

Condition	RT (ms)	Priming
	Low frequency	
Control	629	
SW	586	43
SS-small	594	35
SS-large	621	8
	High frequency	
Control	609	
SW	563	46
SS-small	569	40
SS-large	574	35

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

Discussion

The results indicate that lexical inhibition might be an important factor in the activation of lexical candidates. If there is phonetic overlap, as with the CVCC words, inhibition is proportionate to the cohort size of the competitors: Items with no or few lexical competitors are less inhibited than items with many competitors. In contrast, if there is no overlap, as with CVC words, lexical inhibition does not play a role, and consequently, priming effects are identical. Moreover, there is some evidence that lexical inhibition effects are proportionate to the frequency of the target: Low-frequency targets suffer more from competitors than high-frequency targets, because the former are less able to suppress their competitors.

There are, however, several concerns that should be considered. One potential problem is that there is an overall trend for CVC items to be less facilitated than CVCC items. From a theoretical point of view, one might expect just the opposite, because CVCC, though not CVC, items may suffer from competition effects.[1] One should consider, however, that the phonetic overlap between prime and target is larger with CVCC items than with CVC items. Thus CVCC

---

[1] We thank Arthur Samuel, one of the reviewers, for pointing this out.

words simply may receive more input and, hence, become more activated than CVC words.

Another possibility may be that for some unclear reason, the acoustic realizations of CVCC words in the SS-large and, to some extent in the SS-small context (*melkaam* and *melkeum*), were in a somewhat less canonical form than those of the control condition (*melkem*). In other words, unknown acoustic factors might have played a role. To investigate that, we conducted a second experiment in which the visual targets were presented earlier during presentation of the auditory prime. If indeed lexical inhibition was at stake, one would expect unequal priming effects of CVCC targets to disappear with short ISIs, because evidence for competing candidates is generally assumed to not arrive until later. Words inhibit each other to the extent that they are activated, and hence, lexical candidates must first be activated before they can start to inhibit each other. Inhibition thus lags behind activation, and if our results do reflect lexical competition effects, decreasing the ISI should have eliminated the difference between the priming effects of CVCC targets in the SW, SS-small, and SS-large conditions. On the other hand, if the acoustic realization of the target is responsible for the obtained results, decreasing the ISI should have had no effect on the observed pattern.

Experiment 2

Experiment 2 was conducted to test whether the unequal priming effects of CVCC words can be attributed to lexical inhibition effects or to an acoustic factor. Lateral inhibition lags behind activation, and decreasing the ISI should therefore eliminate the difference between the priming effects of CVCC targets in the SW, SS-small, and SS-large conditions.

Method

Study participants. Twenty participants were tested. None of them had taken part in the previous experiment.

Design and procedure. The design and procedures were the same as in the previous experiment except that the ISI was set at 0 ms.

Results

To maintain comparability with the previous experiment, we discarded the same four item pairs. A total of 6.5% of the remaining data were errors (a response of no to a real word) and another 4.5% were extreme values. These were replaced by the mean of the item or participant of the relevant condition. Mean reaction times are presented in Table 5. As can be seen, priming effects were large, but they were not different for the various conditions.

A 2 × 4 ANOVA was performed on the reaction times with target type (CVCC or CVC target) and prime type (control, SW, SS-small, and SS-large) as within-participants variables. There was no overall difference between CVCC or CVC targets (both F_1 and $F_2 < 1$), and there was a significant effect of the auditory prime, $F_1(3, 57) = 17.94$, $p < .001$; $F_2(3, 129) = 6.53, p < .001$), but the interaction between target type and prime was not significant, $F_1(3, 57) = 1.69$, $p = .18$; $F_2 < 1$. To investigate whether the priming effects were different for the three types of context (SW, SS-small, and SS-large), we performed a 2 (target type) × 3 (context) ANOVA on the priming effects. The ANOVA showed that there was no difference between the priming effects for CVCC or CVC targets, $F_1(1, 19) = 2.68$, $p = .12$; $F_2(1, 43) < 1$; there was no effect of the context (both F_1 and $F_2 < 1$); and the interaction between target type and context was also not significant (both F_1 and $F_2 < 1$). Separate ANOVAs on the priming effects of CVCC and CVC words were all nonsignificant (all *F*s < 1). Similarly, none of the planned comparisons that were conducted in Experiment 1 reached significance (all *F*s < 1). The results thus show that priming effects for CVCC words at a short ISI are not reliably different: neither is the distinction between weak and strong syllables relevant as it was in Experiment 1 nor is there an effect of the number of competitors.

Table 5
Mean Reaction Times (RTs) and Priming Effects for CVCC and CVC Targets With an Interstimulus Interval of 0 ms

Condition	Spoken prime	Visual target	RT (ms)	Priming
		CVCC words		
Control	*lastem*	MELK	574	
SW	*melkem*	MELK	523	52
SS-small	*melkeum*	MELK	528	47
SS-large	*melkaam*	MELK	517	57
		CVC words		
Control	*wankem*	BEL	563	
SW	*belkem*	BEL	531	32
SS-small	*belkeum*	BEL	527	36
SS-large	*belkaam*	BEL	531	32

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

General Discussion

Two cross-modal priming experiments were conducted to investigate whether speech segmentation is based on strong syllables, lexical competition, or both. The results clearly show that lexical competition does play an important role in the activation of word candidates. The study also allows us to specify the impact of competition in a more precise way than the literature has to date. Not only does competition play a role, but this role can also be quantified in the sense that the level of competition is a function of the number of available competitors. If target words overlap with competitors, lexical competition increases in proportion to the cohort size of the competitors. Frequency of the target also plays a role: High-frequency targets suffer less from competitors than do low-frequency targets. Finally, we have some idea about the time course of the competition effects: Initially embedded words are activated on-line irrespective of the right context, but the right context is responsible for differential inhibition effects if there is overlap with the embedded word. There is thus not only multiple activation of word candidates but also multiple competition. We argue that these competition effects are the signature of lateral inhibition. Before we discuss this claim, however, alternative explanations must be considered.

The first concern is that what we have called *cohort size* might, after all, not be relevant: Rather, it might be that frequency of the best competitor is the crucial aspect. Cohort size as such might be beside the point, because most items within a cohort are extremely rare (see Bard, 1990). As can be expected, cohort size is usually correlated with best-competitor frequency, as the chance of including high-frequency competitors in the cohort increases with the size of the cohort. In terms of processing consequences, the issue can be rephrased as the question of which competitors are relevant. In models such as TRACE, all active competitors contribute to recognition, but for instance, in cohort (Marslen-Wilson, 1987), decisions are made with respect to the strongest competitor only. A reexamination of the cohorts in our stimulus set showed that out of 44 SS-small and SS-large cohort pairs, 38 had the highest frequency competitor in the largest cohort. Therefore we do not, at present, take a firm stand on that issue; it is impossible to disentangle these alternatives on the basis of our results.

We now turn to the issue of whether the competition effects in our study are, in fact, manifestations of lateral inhibition. It is possible to obtain competition effects without lateral inhibition, such as, for instance, in the neighborhood activation model (NAM); (Goldinger, Luce, & Pisoni, 1989). In NAM, the perceptual choice criterion depends on the relative ratio of the activation of a target word to its competitors. Here, competition between lexical hypotheses is indirect, because there is no direct influence of the activation level of one candidate on another. Competition effects are thus manifested in higher level decision processes. The observed competition effects in our study might be attributed to such processes in that target words had to compete in the SS-small cohorts with different competitors than in the SS-large cohorts. On the other hand, competition between lexical candidates can also be more direct, as in TRACE. In this case, activation of one candidate is directly influenced by the activation of another. At the behavioral level, these two possibilities will be difficult to distinguish. There is, however, one specific aspect of our results that allows us to address this particular issue, that is, the interaction observed between frequency of the target and cohort size of the competitor. Such an interaction can be taken as evidence for direct competition through lateral inhibition (see Bard, 1990). Even though the manifestation of this interaction depends on a whole set of other parameters, we did observe such a trend. Thus, there are signs at the empirical level that lateral inhibition plays a role. Moreover, the interpretation of competition effects depends on one's view of cross-modal priming. We adopted the notion that priming effects reflect preactivation of lexical candidates. In this approach, activations are, at that stage in processing, not yet contaminated with decisional choice criteria. Thus, within this framework, choice rules do not yet apply. On this view, the present results are inconsistent with models in which lateral inhibition does not play a role. One of the most well-known examples of this type is the cohort model (Marslen-Wilson, 1987). In cohort, lexical representations are activated to the extent that they match the acoustic phonetic input. If there is a mismatch, words are strongly inhibited. A word is recognized as soon as it is the only active candidate that is left. There is only one way in which words can be deactivated, and that is through a mismatch with the acoustic input. (It should be mentioned that cohort also allows syntactic and semantic context to deactivate word candidates, but this is irrelevant in the present context, as there is no context of this type.) Thus, cohort has no mechanism by which active lexical candidates can directly inhibit other lexical candidates. It is therefore problematic for cohort to account for the present results, and it is probably also one of the main reasons why the contribution of cohort to the problem of speech segmentation is not obvious.

A model that has clearly devoted more attention to speech segmentation than the cohort model is TRACE (McClelland & Elman, 1986). The TRACE model is made up of a hierarchy of feature-, phoneme-, and word-level units that represent hypotheses about the sensory input. There are facilitatory connections between units from adjacent levels and inhibitory connections between units within a level. The excitatory and inhibitory effect of one unit on the other is proportional to the level of activation of the first. All units are aligned with each other in discrete time slices, and units facilitate and inhibit each other to the extent that they overlap within these time slices. This architecture ensures that the network's final interpretation of the input is unlikely to contain two words that receive input from the same phoneme. In this way, lexical embeddings of words within words (*cat* in *catalog*) can be handled, as these ambiguities are mainly solved through the interword inhibitory links.

It remains unclear, however, whether TRACE would be able to simulate the present results. There are some simulations by McClelland and Elman (1986) and more recently by Frauenfelder and Peeters (1990) that are of relevance here. A critical observation that has been made by these authors is that there is a distinction in TRACE between initially and finally embedded words. When the network receives as input a word like *carpenter,* the initially embedded word *car* is activated to some extent. Although the activation of *car* drops after *carpenter* is sufficiently activated, it does have an influence on the activation of the carrier word (see Frauenfelder & Peeters, 1990). In contrast, finally embedded words like *seed* in *precede* never become active, because *seed* is strongly inhibited by the highly active *precede* at the time that the /*s*/ is given as input to the model. There is thus a marked contrast between initially and finally embedded words: Initially embedded words can become active and, consequently, inhibit other words; finally embedded words do not become active at all.

Returning to the present experiments, we note that at first sight it seems that TRACE probably would not activate finally embedded cohorts such as *kaa* or *kaam* when receiving *melkaam,* because at the time /*k*/ is given as input, the *kaa(m)* cohort will be inhibited by the highly active *melk.* This means that, in fact, the cohort size of the competitors would be irrelevant in TRACE, because none of the members could become active anyway. There is, however, a difference between the *melkaam* and *precede* examples: *Melkaam* is a pseudoword, whereas *precede* is not. Moreover, *melk* and the *kaa* or *kaam* cohorts overlap in only one phoneme, whereas there is a complete overlap between *precede* and *seed.* As words inhibit each other in proportion with their overlap, *kaa(m)* words will receive less inhibition from *melk* than *seed* will from *precede.* In fact, a simulation that confirms this idea has previously been reported by McClelland and Elman (1986, p. 63). The TRACE model received two concatenated words, *par* and *key,* /*parki*/, as input. The word *park,* however, became more activated than *par,* because *park* has more phonemes and therefore receives more input. The important point is that there was also substantial activation for *key.* Hence, two lexical items (*park* and *key*) can become active in TRACE even if they have to share a phoneme. This bears directly on the *melkaam* example: Although *melk* might be highly active, it may nevertheless be possible for words starting with *kaa(m)* to become activated too. At present, however, no data have been reported showing whether the *kaa(m)* cohort in TRACE does inhibit the target word *melk* such that inhibition depends on the cohort size of the competitors. In principle, however, it is possible that the TRACE model does predict the observed pattern of results, but this will be difficult to evaluate. First, TRACE has a severely limited vocabulary of only 211 words, and if the size of the lexicon were increased, some of the parameter values would probably need to be reset. Moreover, in terms of hardware implementation, the model would become rather awkward, because the number of lexical inhibitory connections would have to increase more than exponentially with the number of lexical candidates. However, even if these restrictions could be overcome, one should be cautious about making assumptions with regard to the explanatory power of TRACE on this issue. For instance, Shillcock (1990) demonstrated that human listeners do activate words like *bone* in *trombone,* even though *bone* is completely embedded within the existing word trombone. This finding is compatible with the interpretation of the present results, but it is, of course, in contradiction with the TRACE simulation of *seed* in *precede.*

A computational solution that is in this respect more promising is Norris's SHORTLIST model (Norris, 1994). SHORTLIST is a hybrid connectionist model in which a small number of lexical hypotheses (the shortlist) is generated by conventional programming techniques from a complete word lexicon. These lexical hypotheses are then wired into an interactive activation model in which competition takes place. The behavior of SHORTLIST resembles TRACE in many ways, but the model is more tractable and easier to extend. To investigate whether lexical inhibition in SHORTLIST depends on the cohort size of the competitors, we ran a small simulation. The input to the model was a CVCC target word (*melk*) embedded in three types of contexts—*melkem, melkeum,* and *melkaam.* The lexicon consisted of eight words. The words were chosen in such a way that the lexicon was roughly proportionate to the mean number of competitors that were found in the CELEX database. Thus, besides the target word *melk,* there were no words that started with *ke(m),* one word that started with *keu,* five words that started with *kaa,* and one word that started with *kaam.* The model was run with the default parameters (see Norris, 1994). The activation values of *melk* were traced in the three types of contexts, which are presented in Table 6. As can be expected, the competition effect of *melk* can be observed at the presentation of the critical vowel (V). At that time, activation of *melkaam* drops fastest, because the *kaa(m)* cohort becomes activated and its members mutually inhibit *melk.* In the *melkem* case, there is no lexical inhibition, and activation of *melk* remains just at the same level. Activation of *melk* in *melkeum* drops slightly, because the *keu* word is activated so that it inhibits, in its turn, *melk.* The decline, however, is less than in *melkaam,* because there is only a single word involved. This simulation thus displays, at least in a qualitative way, the behavior of our test participants: Inhibition increases if the cohort size of the competitors increases. A related aspect that deserves attention is that lexical inhibition effects are noticeable for only a short period of time. Lexical inhibition is only observed immediately after words have become sufficiently activated. The same effect was observed in the present study, and methodologically, this is an important point, because inhibitory effects might vanish within milliseconds.

In conclusion, we have argued that lateral inhibition at the lexical level is an important component of the word recognition system. If this factor is taken seriously, one can interpret the results of the word-spotting studies (Cutler & Norris, 1988; Vroomen & de Gelder, 1993) in a quite different light: Detection of a word such as *mint* in *mintayf* is more difficult than detection of *mint* in *mintef,* because the former receives more inhibition than the latter. The explanation might thus be that lexical inhibition is the driving force, instead of the MSS, which made a wrong decision about the beginning of a new word. The present results thus challenge the interpretation given by the MSS, but this is not to say that the MSS has been refuted. First, the MSS does not say anything about the nature of the competition that might follow the generation of a segmentation trigger. The MSS is not a model about word recognition but a model about the exploitation of rhythmic cues. So, the MSS can be easily accommodated within any kind of word recognition model and also within those that incorporate

Table 6
Activation of melk *at Each Phoneme Position in SHORTLIST*

	Phoneme position (av)					
Input	m	e	l	k	V	C
melkem	.00	.55	.63	.69	.69	.69
melkUm	.00	.55	.63	.55	.55	.55
melkAm	.00	.55	.63	.55	.45	.55

Note. av = activation values; V = vowel; C = consonant.

lateral inhibition. In fact, the SHORTLIST simulation that is reported here did instantiate the MSS hypothesis. That is, word activations were reset at the beginning of a strong syllable (i.e., at the occurrence of /k/ in *melkeum* and *melkaam*). This reset turned out to be a critical feature of the simulation that allowed lexical inhibition effects to take place (see Norris, 1994, for details). Moreover, empirical evidence for the simultaneous emergence of prelexical MSS-like effects and lexical competition comes from a recent study by McQueen, Norris, and Cutler (1994). McQueen et al. found in a word-spotting task that *mess* was easier to detect in *n@mes* than in *mest@m* (which is in accordance with the MSS), but at the same time they observed that *mess* was easier to detect in *n@mes* than in *d@mes*. The latter finding was explained as a lexical competition effect because *d@mes,* but not *n@mes,* is a possible word onset and, hence, *mess* in *d@mess* has to compete with *domestic*. Thus, at least in English, it seems likely that both metrical segmentation and lexical competition should be incorporated into a single model of spoken word recognition.

Moreover, returning to the present results, we note that there was a small difference between the SW and the SS-small context, and although the cohort size of the SS-small competitors was larger than that of the SW context, one nevertheless might argue that part of this difference lies in the distinction between the weak and strong syllables (the difference between the SS-small and SS-large cohorts would, of course, still need explanation). To investigate that issue, one would need to combine, in an orthogonal way, the cohort sizes of the competitors and the metrical weight of the second syllable. Unfortunately, at least in Dutch, language did not evolve in such a way that psycholinguists could run experiments of this type. However, there may be other languages in which these factors can be controlled systematically. Such cross-language studies are urgently needed. The outcome can have important consequences on the way spoken word recognition is envisaged: According to the lexical inhibition hypothesis, lexical candidates are activated and inhibited in a continuous way and not under the command of the MSS. The lexical inhibition hypothesis is a generic segmentation device that can be applied by any listener once he or she has acquired a lexicon. The MSS is more particular, because it is a language-specific strategy that has to be tuned to the rhythmic properties of the language (e.g., the stress pattern in English, the syllable in French, and the mora in Japanese). The number of lexical hypotheses that is simultaneously entertained is probably larger in the case of lexical inhibition, but lexical inhibition is more parsimonious in terms of processing mechanisms that are needed. Putting the issue in a broader context, we perceive the domain of applicability of the MSS as broader than that of lexical inhibition as it may work for children as a bootstrapping procedure for the acquisition of the lexicon (see Cutler, in press). Thus, future research may reveal that both mechanisms are required for a complete understanding of spoken word recognition.

References

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: Psycholinguistics and computational perspectives* (pp. 185–210). Cambridge, MA: MIT Press.

Cutler, A. (in press). Prosody and the word boundary problem. In J. Morgan & K. Denneth (Eds.), *Signal to syntax.* 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., Mehler, J., Norris, D., & Segui, J. (1983). A language specific comprehension strategy. *Nature, 304,* 159–160.

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

Frauenfelder, U. H., & Peeters, G. (1990). Lexical segmentation in TRACE: An exercise in simulation. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistics and computational perspectives* (pp. 50–86). Cambridge, MA: MIT Press.

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.

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

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

McQueen, J. M., Norris, D., & 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., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior, 20,* 298–305.

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

Segui, J., Dupoux, E., & Mehler, J. (1990). The role of the syllable in speech segmentation, phoneme identification, and lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistics and computational perspectives* (pp. 263–280). Cambridge, MA: MIT Press.

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

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

van Bergem, D. R. (1993). On the perception of acoustic and lexical vowel reduction. *Proceedings of the 3rd European Conference on Speech Communication and Technology: Vol. 1* (pp. 677–681). Berlin: European Speech Communication Association.

van Zon, M., & de Gelder, B. (1993). Perception of word boundaries by Dutch listeners. *Proceedings of the 3rd European*

Conference on Speech Communication and Technology: Vol. 1 (pp. 689–692). Berlin: European Speech Communication Association.

Vroomen, J., & de Gelder, B. (1993). The role of strong syllables for spoken word recognition in Dutch. In P. Smith (Ed.), *Proceedings of the Sixth Conference of the European Society for Cognitive Psychology* (pp. 37–38). Reading, United Kingdom: University of Reading.

Zwitserlood, P., Schriefers, H., Lahiri, A., & van Donselaar, W. (1993). The role of syllables in the perception of spoken Dutch. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 260–271.

Appendix

Experimental Materials

CVCC example					CVC example					Cohort size of second syllable			
Target	*f*	SW prime	SS-small prime	SS-large prime	Target	*f*	SW prime	SS-small prime	SS-large prime	SS-small CV	SS-large CVC	CV	CVC
nacht	266	nachtel	nachtuul	nachtaal	dag	935	dagtel	dagtuul	dagtaal	35	3	397	127
kalf	7	kalfel	kalfeul	kalfiel	val	35	valfel	valfeul	valfiel	8	0	250	65
kalk	11	kalkes	kalkuus	kalkaas	gal	4	galkes	galkuus	galkaas	68	0	865	37
kalm	48	kalmes	kalmuus	kalmees	bal	36	balmes	balmuus	balmees	172	9	808	35
dank	79	dankel	dankuul	dankool	lang	1003	lankel	lankuul	lankool	68	5	915	210
mank	4	mankep	mankeep	mankaap	gang	187	gankep	gankeep	gankaap	136	12	865	59
gans	19	gansek	gansoek	gansiek	man	1190	mansek	mansoek	mansiek	80	1	408	22
lans	5	lansel	lansuul	lansaal	pan	35	pansel	pansuul	pansaal	118	0	365	78
larf	1	larfel	larfeul	larfiel	bar	30	barfel	barfeul	barfiel	8	0	250	65
park	38	parkes	parkuus	parkaas	kar	16	karkes	karkuus	karkaas	68	0	865	37
hart	190	hartel	harteul	harteel	nar	1	nartel	narteul	narteel	18	0	672	203
mast	5	masteg	mastieg	masteeg	das	7	dasteg	dasteig	dasteeg	186	0	673	173
vast	332	vastem	vasteum	vastoom	ras	25	rastem	rasteum	rastoom	18	0	303	27
recht	232	rechtel	rechtiel	rechteel	pech	7	pechtel	pechtiel	pechteel	186	0	672	203
geld	280	gelteg	gelteug	gelteeg	cel	46	celteg	celteug	celteeg	18	6	672	173
melk	51	melkem	melkeum	melkaam	bel	34	belkem	belkeum	belkaam	124	0	865	114
kelk	4	kelker	kelkeer	kelkaar	fel	61	felker	felkeer	felkaar	136	34	865	160
cent	26	centes	centies	centees	den	7	dentes	denties	dentees	186	0	672	13
kerk	205	kerkem	kerkeum	kerkoom	ver	546	verkem	verkeum	verkoom	124	0	915	35
wesp	4	wespel	wespuul	wespool	bes	7	bespel	bespuul	bespool	92	0	299	246
pest	14	pestel	pestiel	pestool	zes	127	zestel	zestiel	zestool	186	0	303	5
nicht	11	nichtel	nichtuul	nichtaal	big	3	bigtel	bigtuul	bigtaal	35	3	397	127
gids	22	gidset	gidsoet	gidsiet	lid	230	lidset	lidsoet	lidsiet	80	8	408	58
lift	28	lifteg	lifteug	lifteeg	rif	1	rifteg	rifteug	rifteeg	18	6	672	173
niks	89	nikset	niksoot	niksiet	dik	158	dikset	diksoot	diksiet	203	5	408	58
film	106	filmes	filmuus	filmees	gil	8	gilmes	gilmuus	gilmees	172	9	808	35
mild	21	miltef	miltuuf	milteef	pil	27	piltef	piltuuf	pilteef	35	0	672	3
pink	6	pinkes	pinkees	pinkaas	ding	371	dinkes	dinkees	dinkaas	136	7	865	37
bink	1	binket	binkuut	binkoot	ring	34	rinket	rinkuut	rinkoot	68	0	915	22
gips	5	gipsel	gipsoel	gipsiel	lip	108	lipsel	lipsoel	lipsiel	80	1	408	27
golf	61	golfes	golfeus	golfees	lol	8	lolfes	lolfeus	lolfees	8	0	142	45
pomp	5	pompes	pompuus	pompaas	bom	23	bompes	bompuus	bompaas	92	0	667	25
mond	228	montel	monteul	monteel	non	19	nontel	nonteul	nonteel	18	0	672	203
honk	0	honkel	honkuul	honkaal	long	20	lonkel	lonkuul	lonkaal	68	5	865	76
vonk	11	vonkem	vonkeem	vonkoom	tong	54	tonkem	tonkeem	tonkoom	136	5	915	35
lont	3	lontef	lontief	lonteef	ton	30	tontef	tontief	tonteef	186	6	672	3
korf	3	korfem	korfuum	korfaam	dor	10	dorfem	dorfuum	dorfaam	39	3	168	84
worm	10	wormel	wormeul	wormeel	tor	4	tormel	tormeul	tormeel	24	0	808	84
post	13	posteg	postieg	posteeg	vos	7	vosteg	vostieg	vosteeg	186	0	397	173
lucht	183	luchtef	luchteuf	luchteef	mug	6	mugtef	mugteuf	mugteef	18	0	672	3
zucht	15	zuchtel	zuchtuul	zuchtaal	rug	180	rugtel	rugtuul	rugtaal	35	3	397	127
tulp	3	tulpes	tulpeus	tulpaas	nul	9	nulpes	nulpeus	nulpaas	31	0	667	25
punt	172	punt	punteum	puntoom	dun	42	duntem	dunteum	duntoom	18	0	303	27
muts	7	mutset	mutsoot	mutsiet	put	10	putset	putsoot	putsiet	203	5	408	58
M	62.8					127.4				96.5	3.6	592.9	71.0
SD	91.5					269.6				62.8	6.6	240.4	61.8

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

Received November 22, 1993
Revision received March 21, 1994
Accepted March 21, 1994 ■