Activation of embedded words in spoken word recognition

Vroomen, J.; de Gelder, B.

Activation of Embedded Words in Spoken Word Recognition

Jean Vroomen
Tilburg University

Beatrice de Gelder
Tilburg University and Université Libre de Bruxelles

Three cross-modal associative priming experiments investigated whether speech input activates words that are embedded in other words. When the embedded word corresponded to the final syllable of a bisyllabic carrier (*boos*, meaning *angry* embedded in *framboos*, meaning *raspberry*), facilitatory priming effects were observed for related targets of the embedded word. No effects were found when the end-embedded word did not start at the onset of a syllable (*wijn* meaning *wine* in *zwijn* meaning *swine*). Beginning-embedded words were activated only if the carrier was a nonword (*vel* meaning *skin* in *velk*), but not when the carrier was a word (*vel* in *velg*, meaning *rim*). The results support the joint operation of metric segmentation and lexical competition: Words are activated if their onset matches the onset of a strong syllable; words are then excluded on the basis of interword competition.

Continuous speech contains very few reliable cues to the location of word boundaries. Nevertheless, listeners are rarely conscious of any difficulty resolving such ambiguities as words are recognized correctly and immediately without apparent mental effort. However, there is clear evidence that, before a word is recognized, numerous incorrect lexical hypotheses are entertained by the speech recognition system without the listener being aware of them. The way in which these lexical hypotheses are generated can provide insights into how continuous speech is segmented into discrete words.

In the literature, there are at least two general proposals about how speech is segmented. The first is a postlexical strategy by which words are recognized in the strict order in which they were spoken (Cole & Jakimik, 1980; Marslen-Wilson & Welsh, 1978). One of the best known examples is the early version of the Cohort model (Marslen-Wilson & Welsh, 1978). In early Cohort, word boundaries emerge when the current word is recognized. Words sharing the same onset are all accessed in parallel during recognition,

and a word is recognized as soon as the input distinguishes it from all other words in the lexicon. At this word's *uniqueness point*, the phonological code of the word is accessed, and the end is calculated so that the beginning of the next word can be predicted correctly. In early Cohort, thus, only words with aligned onsets are ever considered at the same time.

Continuous activation models contrast with this strict left-to-right sequential processing. Examples of continuous models are the more recent versions of Cohort (Marslen-Wilson, 1987, 1993), TRACE (McClelland & Elman, 1986), and Shortlist (Norris, 1994; Norris, McQueen, & Cutler, 1995). In continuous models, there is, at any moment in time, bottom-up input to lexical hypotheses, whatever part of the input they span. Thus, words not sharing the same onset also receive input and, eventually, these non-aligned words may also become active. Continuous models are thus less strict in requiring that all active words have aligned onsets. Therefore, the cohort of active words in continuous models can be different from the one of sequential models.

The commonality between sequential and continuous activation models is that both agree that multiple words are activated if they share the onset. One of the best known empirical demonstrations of this phenomenon comes from a study by Zwitserlood (1989) in which listeners heard *kapitein* or *kapitaal*, and in which participants concurrently saw a visual probe to which a lexical decision was made. In the critical conditions, the visual probe was presented at a point at which the input was still compatible with both possible words (e.g., at the *t* in *kapitein* or *kapitaal*). The probe was either associatively related to *kapitein* (i.e., BOOT, meaning *ship*) or to *kapitaal* (i.e., GELD, meaning *money;* it should be noted that auditory primes are written in italics and the visual targets in capitals). Reaction times for both probes were facilitated, relative to a control condition when presented at the *t* of either *kapitein* or *kapitaal*. By contrast, when the probe was presented at the end of the spoken

word, only the probe related to the actual word was facilitated. Thus, at the end of *kapitein,* BOOT but not GELD was facilitated and at the end of *kapitaal,* GELD but not BOOT was facilitated. These results strongly suggest that words matching from word onset are simultaneously activated until the ongoing speech input differentiates the competing candidates from each other.

However, sequential and continuous activation models differ in the way in which words are activated that do not match a word onset. As an example, one may take the word *bone* as being embedded in *trombone.* In strict sequential models, there is no way in which *bone* could be active, because only words starting with *trom* . . . could have been part of the word-initial cohort. Continuous models often make similar predictions, but for different reasons. In the TRACE model, there would be no facilitation of *bone* when *trombone* is heard because, when *bone* receives input, *trombone* is so active that it inhibits *bone* via lateral inhibition. Lexical inhibition from *trombone* thus prevents *bone* from becoming active (see Frauenfelder & Peeters, 1990 for a simulation). In Shortlist, there is also lexical competition, but one of the differences from TRACE is that Shortlist incorporates the metric segmentation strategy (MSS). The MSS, as put forward by Cutler and Norris (1988), is a prelexical strategy that locates where word boundaries are likely to occur in the ongoing speech signal. The original notion was that new lexical access attempts are initiated at the onset of strong syllables. (In English, strong syllables contain full vowels as opposed to weak vowels, which usually contain a schwa). The Shortlist model implements the MSS by giving words starting at the onset of strong syllables an extra boost over words that do not start at the onset of a strong syllable. Embedded words, such as *bone* in *trombone,* therefore, receive extra input that, depending on parameter setting, may give *bone* a short-lived activation.

The predictions of the later versions of Cohort are somewhat variable. The similarities with TRACE have been stressed in Marslen-Wilson (1987), but in later versions (e.g., Marslen-Wilson, 1993), there is not, as in TRACE, direct competition at the activation levels of lexical candidates. Rather, competition is now located at a decision stage in which the differential activation level of competing candidates determines the ease of recognition. The decision will take longer if two candidates have very similar activation levels. Crucial, for the present purpose, is that words in the most recent Cohort model do not influence each other directly. The activation level of a word is computed solely on the basis of the goodness-of-fit between the input and its lexical representation. Matching input facilitates and mismatching input directly inhibits the lexical activation of a word. This leads one to predict that embedded words, for which there is presumably no mismatching input, are as active as if presented in isolation. Thus, hearing *bone* or *trombone* would, ceteris paribus, make no difference for the activation level of *bone* per se. Rather, the difference between hearing *bone* or *trombone* is in the activation level of *trombone.*

At the empirical level, there is at present some mixed evidence that internally embedded words are active. First,

there is evidence from wordspotting by McQueen, Norris, and Cutler (1994) showing that listeners are able to detect words embedded at the end of nonword strings. The study by McQueen et al. showed that a listener can detect the word *mess* when embedded in *nemess.* However, spotting words at the end of nonsense strings appears to be very difficult (miss rates up to 46%), so one might argue that it requires a conscious strategy that is unrelated to normal listening in which words are recognized fast and reliably without much mental effort. However, less open to this kind of critique were papers by Vroomen and de Gelder (1995a; 1995b) and Norris, McQueen, and Cutler (1995), both showing that the number of words starting at the end of a bisyllabic string may have an impact on the recognition of a previously heard word.

In the study by Vroomen and de Gelder (1995b), the effect of competitor size was investigated in cross-modal identity priming. Auditorily presented words with no, few, or many competitors served as prime for a visual target. For example, the visual target MELK was preceded by the heard primes *melkem, melkeum,* or *melkaam.* The difference among the word endings is that there are no words in Dutch that start with *ke(m),* few with *keu(m),* and many with *kaa(m).* Thus, the *competitor size* of the words was defined as the number of words in the lexicon starting at the second syllable. The facilitatory effect of the prime was proportionate to the number of competitors: *melkem* produced the largest facilitatory effect; *melkeum,* the intermediate; and *melkaam,* the smallest. This finding strongly suggests that lexical candidates embedded at the end of a string can become active so that they can compete with previously heard words with which they overlap.

Similar effects from end-embedded words were obtained by Norris, McQueen, and Cutler (1995) using a wordspotting task. Listeners were asked to detect words at the onset of bisyllabic strings. For example, participants were asked to detect *pram* embedded in *prampidge* or *thin* in *thintaup.* The crucial difference between *taup* and *pidge* is that there are many words in English beginning with *pidge* but only few with *taup.* The results showed that the detection of consonant–vowel–consonant (CVC) targets followed by syllables with many competitors (*pram* in *prampidge*) was easier than the detection of targets followed by syllables with few competitors (*thin* in *thintaup*). The same pattern of results in wordspotting was also obtained in a correlational analysis by Vroomen, van Zon, and de Gelder (1996). It thus confirms that words embedded at the end of other strings may have an effect on the recognition of a previously heard target.

So far, the previously mentioned studies used words embedded in nonword strings, and the competition effects were attributed to the cohort of words activated by the second syllable. However, the evidence for lexical access of end-embedded words in real carrier words is far more mixed and less well-documented. Moreover, the evidence for the activation of a single end-embedded word instead of an entire cohort is sparse. There is a study by Shillcock (1990) showing that end-embedded words such as *bone* in *trombone* may be active when *trombone* is heard. Shillcock

observed in cross-modal associative priming that lexical decision times to targets were faster, compared with control targets, when a related word was embedded at the end of a carrier. For example, when *trombone* was heard, lexical decision times to the visual target RIB, which is related to *bone*, were faster than to the control target BUN.

However, at present, there are a number of questions concerning the generality of Shillcock's findings. First, there is a methodological problem with Shillcock's study in that the control condition for the related prime–target pairs (e.g., *trombone*–RIB) is a condition in which the target was changed (e.g., *trombone*–BUN), but not the prime. This may cause some problems for the interpretation of the data as one might attribute differences between related and unrelated pairs to differences between targets (i.e., RIB vs. BUN). Moreover, Marslen-Wilson, Tyler, Waksler, and Older (1994, Experiment 5) failed to observe a cross-modal repetition priming effect from *trombone* on BONE. Although this kind of form priming is different from associative priming, it surely poses a problem for a straightforward interpretation. There is also a study by Gow and Gordon (1995) using cross-modal associative priming. They failed to observe a facilitatory effect for the target KISS from *lips* when *tulips* was heard, whereas there was such an effect when listeners heard *two lips*. However, the data in this study were somewhat difficult to interpret because inspection of the table shows that the failure to obtain a priming effect was located in the different baselines of *tulips* and *two lips* but not in the absolute difference between the two experimental conditions. Moreover, the carrier words were rather mixed. Eight of them had the stress on the first syllable (e.g., *tulips, forwards*), the other 16 had the stress on the second (e.g., *attack, mistake*), and it is not clear whether this attenuated the priming effect.

Given the indeterminate status of these findings and their potential theoretical importance, we first wanted to establish whether words embedded at the end of real carrier words (e.g., *bone* in *trombone*) are temporarily activated. In Experiment 1, we conducted a study similar to Shillcock's (1990) using Dutch words and a somewhat different methodology (instead of varying the target, we varied the prime so as to keep target attributes constant). A participant thus heard a Dutch word such as *framboos* (meaning *raspberry*), which has the word *boos* (meaning *angry*) embedded in it. In the critical condition, the participant made a lexical decision concerning the associatively related visual target KWAAD (meaning *angry* as well). Strict sequential models would predict no facilitation for KWAAD because *boos* could not have been part of the word-initial cohort. The TRACE model would also predict no facilitation of KWAAD because lexical competition would prevent *boos* from becoming active. Particular versions of other continuous activation models might predict some activation of *boos*. In Shortlist, one might find activation of the embedded word if the MSS boost were stronger than the lexical competition effect. In the revised version of Cohort (Marslen-Wilson, 1993), there is no competition at the activation level so that *boos* receives only facilitatory bottom-up input, which should increase its activation.

In Experiment 2, we replicated Experiment 1, but this time with items that had a different stress pattern. In Experiment 3, we tested several predictions of the continuous activation model. This time, the embedded words used did not match a syllable boundary (e.g., *wijn*, meaning *wine* embedded in *zwijn*, meaning *swine*). This allowed us to test whether there is, as in Shortlist, a distinction between internally embedded words that match or do not match a syllable boundary. We also compared word versus pseudoword carriers because in the framework of lexical competition, pseudoword carriers should produce larger facilitatory effects than word carriers.

Experiment 1

Method

Participants. Participants were 24 students from Tilburg University. Equal numbers of participants received the two versions of the test.

Materials. The materials were constructed around 30 bisyllabic carrier words containing another word in their second syllable (Appendix A). Association norms were collected for the embedded words. Selection was done as follows. First, 68 carrier words were selected that contained other words in their final syllable. The carriers were chosen from a dictionary search (the CELEX dictionary; Baayen, Piepenbrock, & van Rijn, 1993) and met the following criteria: All carriers were nonderived bisyllabic words of frequent usage, carriers were lexical as opposed to grammatical words, both syllables of the carrier were strong (i.e., a strong–strong [SS] pattern), carriers had lexical stress on the second syllable, and the second syllable of the carrier was another word of frequent usage considered to have a close associate.

Because association norms for the embedded words were not available, it was necessary to conduct a pretest. The 68 embedded words were presented in random order in written form to 12 participants in a free-association test. The students were asked to write down the first associate that came to mind when they read each word. From these norms, the 30 highest associates were selected. The overall association strength was .52 (i.e., 52% of the subjects chose the probe as first associate, the minimum was .25). The 30 carriers that contained these embedded words plus the associated visual probe provided the basis for the experimental item set. For 80% of the carriers, the embedded word comprised the entire second syllable; in the remaining items, the initial consonant of the embedded word was ambisyllabic (e.g., the *l* in *balans*). Another 30 bisyllabic words were selected that could serve as unrelated primes for the control condition. The control primes were, like the carriers, bisyllabic lexical words with a SS pattern and with the stress on the second syllable. They were matched on frequency with the carrier words, and they did not have any semantic or phonological relation to the visual target.

An additional 90 filler items (spoken prime plus visual probe) that had no associative or phonological relation to prime or to its final syllable and the target were constructed. All auditory filler primes were bisyllabic, lexical SS words of frequent usage with the stress on the second syllable. Thirty fillers items were paired with a real word as visual probe (a "Yes" decision was required), and 60 fillers had a nonword as visual probe (a "No" decision was required). All nonword probes were legal nonwords matched for number of letters and syllables with the test targets.

An additional 20 unrelated prime/probe items were chosen as a

practice set. The practice set contained 10 real words and 10 nonwords as visual probes.

Design and procedure. Two sets of materials were constructed so that the experimental items were counterbalanced across the two conditions. Each critical target was observed under two priming conditions, but no participant saw the same probe more than once. The experimental trials in each set (15 related and 15 unrelated prime–target pairs) were pseudorandomly interspersed with the 90 fillers. Fillers and experimental trials appeared at exactly the same location across the two sets.

The auditory primes were spoken by a male native speaker of Dutch. They were recorded in a sound-attenuated studio on digital audiotape (Sony DAT-55). The items were then digitized at 22.05 kHz (16 bits precision), and the offset of the embedded word was determined under visual and auditory control. The offset point served as reference for the interstimulus interval (ISI), which was set at 0 ms. The target thus appeared at the offset of the prime on a cathode ray tube (CRT) screen. The presentation time of the visual probe was 50 ms (unmasked, white letters on black screen). A fixation point appeared 100 ms before the onset of the target. It was located 2 cm under the center of the visual probe, and it remained there for 50 ms. The intertrial interval was 3.5 s.

The participants were tested individually in a sound-attenuated booth. They were instructed to decide as quickly as possible whether the string of letters that was presented after each spoken stimulus was a real word or not. They responded by pressing a "Yes" or a "No" key in front of them. The "Yes" response was always made with the dominant hand. The spoken materials were presented via Sennheiser HD-410 headphones at a comfortable listening level. Each test session lasted about 10 min.

Results

In this and in all other experiments, error responses (i.e., answering "No" to a real word) were discarded from the analysis. Means per item and participant for each condition were computed from the remaining data points. Reaction times less than 200 ms and more than 1,000 ms were replaced by their cut-off values. Two items were discarded from the analysis: one because it appeared that not only the embedded word but also the carrier itself was associatively related to the target (i.e., *cacao* [meaning *chocolate*] with its embedded word *kou* [meaning *cold*] are both related to the target WARM [meaning *hot*]), the other item (DAK, meaning *roof*) was discarded because its error rate was 21%. The overall error rate after exclusion of these two items was 1% and was equally distributed across both conditions. In this and in all other experiments, no analyses of the error rates approached significance. The mean lexical decision latency for related targets was 537 ms; for the control condition, it was 564 ms. There was thus a 27-ms facilitatory effect of related primes, which was significant in the participant and item analysis, $F_1(1, 23) = 10.45, p < .004; F_2(1, 27) = 7.98, p < .009$. The effect is comparable in size to the effect obtained in similar experiments (e.g., Shillcock, 1990).

Several post hoc analyses were performed on the priming effects. If there is competition between the carrier and the embedded word, one would expect the embedded word to receive less activation when it is competing with a high-frequency carrier. There was, however, no significant correlation between the amount of priming and the (logarith-

mically transformed) frequency of the carrier word, $r(28) = .00$, suggesting that competition alone cannot account for the data. Second, the frequency of the embedded word might be important because embedded words might receive more activation if they are of high frequency. However, contrary to this prediction, the correlation between the frequency of the embedded word and the priming effect tended to be negative, $r(32) = -.32, p = .09$, indicating that low-frequency embedded words were in fact somewhat better primes. Third, a detailed prediction from TRACE is that, depending on parameter settings, end-embedded words might get a short-lived activation if the first syllable of the carrier word is short and, thus, only contains a few phonemes. Auditory primes such as *azuur* (meaning *azure*, and containing the critical word *zuur*, meaning *sour*) with only one phoneme in their first syllable might, ceteris paribus, be more effective primes than *framboos*, which has four phonemes in the initial syllable. However, when the item set was split into three categories (i.e., 5 items with one, 19 items with two, and 4 items with three and more phonemes in the first syllable), there was no difference in the priming effects of these categories. (Mean facilitatory effects were -6 ms, 36 ms, and 37 ms, $F_2(2, 25) = 1.29, p = .29$ for items with one, two, or three and more phonemes in the first syllable, respectively. It should be noted that the trend is in the wrong direction, but also that the post hoc comparison is restricted because of the unequal sizes of the categories.) However, there was a significant correlation between the amount of priming and the frequency of the *visual* target, $r(28) = -.42, p < .03$, indicating that high-frequency targets were less facilitated than low-frequency targets. This correlation probably suggests that one should be cautious in interpreting priming effects as a pure measure of the attained level of activation because the amount of priming may be modulated, in an intricate way, by the frequency of the visual target.

Discussion

The results of Experiment 1 clearly show that the speech processing system generates lexical hypotheses that do not correspond to the strictly left-to-right parsing of the speech input: Words embedded at the end of other words are generated as lexical candidates. Therefore, these results are inconsistent with strong sequential models such as the early version of Cohort (Marslen-Wilson & Welsh, 1978). In early Cohort, end-embedded words could never have been part of the word-initial cohort, and so they should not have been active. The TRACE model is less deterministic in that lexical hypotheses are generated on the basis of a goodness-of-fit between a lexical entry and the speech input. Its sequentiality is less emphasized than in Cohort, but still, the details of the results do not favor TRACE either. In TRACE, end-embedded words will be effectively suppressed by the activation of the carrier (Frauenfelder & Peeters, 1990), although in some special cases there might be a residual activation of end-embedded words if (a) the carrier has a short first syllable, (b) the carrier is low in

frequency, and (c) the end-embedded word is high in frequency. None of the observed correlation patterns supported these predictions. On the other hand, the results might be accommodated by Shortlist whereby the MSS part would give an extra boost to words starting with a strong syllable. A word such as *boos* in *framboos*, therefore, receives an extra boost because the vowel of *boos* is strong. This might result in a short-lived activation of *boos*, which can be held responsible for the facilitatory effect on KWAAD. The later version of Cohort (Marslen-Wilson, 1993) also accounts for the present findings because there is a continuous match with the input without lateral inhibition at the lexical level. However, before elaborating on these issues, we tested the generality of the previous findings by using carriers with different stress patterns.

Experiment 2

Experiment 2 was similar to the previous one, except that all carriers were bisyllabic words with a weak–strong (WS) pattern. Examples are *beschuit* (meaning *biscuit*), *brevet* (meaning *certificate*), or *vervoer* (meaning *transport;* for a complete list see Appendix B). The first syllable of the carrier contained a schwa; the second syllable was strong and contained another word that was semantically unrelated to the carrier itself (i.e., *schuit* [meaning *boat*], *vet* [meaning *fat*], and *voer* [meaning *fodder*] do not have any semantic relation with their carriers). The morphological complexity of the carriers is somewhat debatable. In the CELEX lexicon, they are all coded as being monomorphemic, but most of the carriers (25 out of 36) began with a syllable that commonly occurs as a prefix in Dutch (i.e., *be-, ge-,* and *ver-*). If there is some form of prefix stripping, the processor must treat these word beginnings, which are homophonous with a prefix, as potential prefixes. On this account, one might expect end-embedded words to receive even more activation than in Experiment 1 because the prefix-stripping process segments the second syllable as a potential stem from a prefixed word. However, in contrast with this prediction, with comparable English WS carriers, Shillcock (1990) did not obtain priming of end-embedded WS carrier words (e.g., *report* did not activate *port*), although there was an effect with SS carriers (*trombone* activated *bone*). Thus, there is a potential conflict here, and we therefore conducted the next experiment using WS carriers.

Method

Participants. Twenty-four new participants from Tilburg University were tested. Equal numbers received the two versions of the test.

Materials. The materials were constructed around 36 bisyllabic carrier words with a WS pattern containing another word in their second syllable (Appendix B). For these embedded words, association norms were collected as in the previous experiments. First, 53 monomorphemic WS-carrier words of frequent usage were selected from the CELEX dictionary that contained other words in their final syllable. Association norms for these embedded words were collected from 18 new participants, after which the 36 highest

associated pairs were selected. The overall association strength was .48 (the minimum was .27). Another 36 bisyllabic WS words were selected that could serve as unrelated prime for the control condition. These control primes were also bisyllabic lexical WS words individually matched on frequency with their experimental counterpart.

An additional 108 filler items (spoken prime plus visual probe) were constructed that had no associative or phonological relation between prime or final syllable and target. All auditory filler primes were bisyllabic lexical words with a WS pattern. Thirty-six filler primes were paired with a real word as visual probe ("Yes" decision required), and the other 72 fillers had a nonword as visual probe ("No" decision required). All nonword probes were legal nonwords matched for number of letters and syllables with the test targets. An additional 20 unrelated items were chosen as practice set.

Design and procedure. Design and procedure were as in Experiment 1. Thus, two sets of materials were constructed so that the experimental items were counterbalanced across the two conditions. Each critical target was observed under two priming conditions, but no participant saw the same probe more than once. The experimental trials in each set were pseudorandomly interspersed with the fillers, and they appeared at exactly the same location across the two sets. The ISI was measured from word offset of the prime and was set at 0 ms.

Results and Discussion

The overall error rate was 1% and was equally distributed across both conditions. The mean lexical decision latency in the related condition was 542 ms; in the control condition, it was 570. The 28-ms facilitatory effect of related primes was significant by items and subject, $F_1(1, 23) = 10.15, p < .004; F_2(1, 35) = 10.70, p < .002$, and it was comparable in size with that obtained in Experiment 1.

Similar correlational post hoc analyses were performed as in Experiment 1 on the priming effect and the frequency of the carrier or embedded word. There was only a small negative correlation approaching significance between the priming effect and the frequency of the carrier, $r(36) = -.29, p = .08$, indicating that there was less priming if words were embedded in high-frequency carriers. This pattern is congruent with a lexical competition account in which the activation of words is modulated by their frequency. To investigate whether (pseudo-)prefix stripping played a role, we divided the items into two sets depending on whether the first syllable of the carrier was a commonly occurring prefix (*be-, ge-,* or *ver-*) or not. There were 25 pseudoprefixed carriers and 11 carriers in which the first syllable was not a prefix. The analysis showed that there was no difference between these two categories, $F_2 < 1$. The mean priming effects of pseudoprefixed carriers was 23 ms; for the other ones, it was 32 ms. It should be noted that the trend for a pseudoprefix account is in the wrong direction.

The results of Experiment 2 thus clearly show that end-embedded words in WS carriers are activated. This finding thus replicates and extends the results of Experiment 1 in which only SS carriers were used.

Experiment 3

Experiment 3 was about the importance of a match between the onset of a word and a syllable. For that purpose, we included end-embedded words that were not aligned with a syllable boundary. For example, is *wijn* (meaning *wine*) considered to be a lexical candidate when the monosyllabic carrier *zwijn* (meaning *swine*) is heard? As before, sequential models would predict no facilitation of *wijn* in *zwijn* because the embedded word does not start at a word onset. In TRACE, competition will also be strong, although one might predict some priming effects because the embedded word now starts at the carrier's second phoneme. At the time the embedded word receives input, the carrier is not yet highly activated, so it is not yet a strong competitor for the embedded word. One might find even bigger priming effects if the carrier were a nonword like *twijn*, because nonwords are likely to be less effective competitors than real words (although it should be mentioned that the "lexical gang" of words starting with *twij* ... might inhibit the embedded word *wijn* as well).

The Shortlist model would predict no facilitation of *wijn* in *zwijn* because *wijn* does not receive the extra boost for beginning at a strong syllable onset. In the more recent version of the Cohort model (Marslen-Wilson, 1993), in which there is no direct competition at the activation level of the word, the lexical representation of *wijn* would receive equal amounts of activation from *wijn*, *zwijn*, and *twijn*, given that the bottom-up input to *wijn* is equal. In the recent Cohort, then, all three primes should produce a priming effect.

Another question was whether beginning-embedded words are activated as well. One may take as an example the word *vel* (meaning *skin*) as being embedded in *velg* (meaning *rim*). The embedded word *vel* matches the onset of *velg*, but it does not match the entire input. This situation is very similar to the ones Marslen-Wilson and collaborators have investigated. Their finding was that beginning-embedded words are part of a word-initial cohort until the input diverges from the lexical representation. However, in this account, it is not clear how short words that fully match the input are rejected for longer ones. It is thus left unexplained what the mechanism is which takes care that when *velg* is heard, only *velg*, and not *vel*, is recognized. Although Cohort is not very explicit about this situation, it seems to be in line with the theory arguing that words are excluded from the word-initial cohort mainly on the basis of mismatching acoustic information. So *vel* may have been active until the *g* in *velg* is heard. A crucial characteristic of Cohort is that there is no interword competition at the word level. So, presenting a beginning-embedded word in a word (*velg*) or pseudoword (e.g., *velk*) should make no difference in terms of inhibitory bottom-up input to the embedded word. The embedded word *vel* should thus be as active when presented in *velg* or in *velk*. On the other hand, in TRACE and Shortlist, there is lexical inhibition, so the lexical status of the carrier should matter because *velg*, though not *velk*, can be a competitor for *vel*. *Vel* should thus be active after presenting *velk*, but not after *velg*.

Method

Materials. The materials were constructed around 2 (beginning vs. end embedded) × 35 item pairs (such as *wijn/zwijn* or *vel/velg*; see Appendix C for a complete list). One member of the pair is referred to as the original begin (*vel*) or the original end word (*wijn*); the other is the carrier (*zwijn* and *velg*). The items were selected from a larger set of 240 words. Both words of a pair were nonderived, monosyllabic words of frequent usage; both were lexical words; the original word was embedded either at the beginning or at the end of the carrier; the orthography and phonology of the original word and the matching part of the carrier were identical; and the original word was considered to have a close associate.

Association norms were obtained by presenting the 240 original words in random order in written form to 12 students in a free-association test. From these norms, the 35 highest beginning-word associates and 35 highest end-word associates were selected. The overall association strengths for the beginning words was .31 (the minimum was .17); for the end words, it was .27 (with a minimum of .17).

For practical reasons, it was impossible to find adequate items that were embedded both at the beginning and at the end of a carrier word (e.g., *waar* in *zwaar* and *waard* would be one example). Therefore, we included a condition in which the original word served as prime so as to have a baseline against which priming effects of embedded words could be compared. In the original end-word condition, subjects heard *wijn*, and at the offset the associated visual probe ROOD was presented. In the embedded end-word condition, subjects heard *zwijn* followed by the same visual probe. The control for these two conditions was a word that did not have any phonological or semantic relation to the probe (e.g., *kwast*, meaning *brush*). The embedded end-nonword condition was made by changing the initial phoneme of *zwijn* so that the new prime became a legal nonword such as *twijn*. The baseline for *twijn* was a nonword control condition in which the original control word was changed by one phoneme so that it became a legal nonword as well (e.g., *kwast* became *kwest*).

The beginning-word conditions were made in a similar manner. Thus, in the original beginning-word condition, words like *vel* served as prime for the associated probe HUID. The embedded beginning-word condition had *velg* as prime followed by the same probe. An unrelated word (e.g., *reus*, meaning *giant*) served as baseline for these two conditions. Primes for the embedded beginning-nonword condition were made by changing the final phoneme so that *velg* became the nonword *velk*. The nonword-control prime for this condition was made by changing a phoneme of the control word so that it became a nonword as well (for example, *reus* became *beus*).

An additional 98 filler items (spoken prime plus visual probe) were constructed that had no associative or phonological relation between prime and target. Fourteen of the fillers items had a real word as visual probe preceded by a nonword prime. In this way it was accomplished that half of the real word probes were preceded by a real word prime, and the other half was preceded by nonword primes. The rest of the fillers were nonword probes, half of them preceded by a real word prime and half by a nonword prime. All nonword probes were legal nonwords matched for length with the test targets. An additional 20 prime–probe items were chosen as practice set. The practice set contained 10 real words and 10 nonwords as visual probes.

Design and procedure. Five sets of materials were constructed so that items were counterbalanced across conditions. Each probe word was observed under all five priming conditions, but no subject saw the same probe more than once. The 70 experimental trials in each set were pseudorandomly interspersed with the 98 fillers. The fillers and experimental trials appeared at the same location across the five sets.

The offset of the embedded word was determined under visual and auditory control. Note that the offset point for beginning-embedded words actually occurred within the carrier itself (e.g., at the *l* in *velg* or *velk*). This offset point served as reference for the ISI which was set at 250 ms. The target thus appeared 250 ms after offset of the embedded word on a CRT screen. The ISI was 250 ms as compared to 0 ms in the previous experiments because in this way it was assured that responses could not be initiated before the final phoneme of *velg* and *velk*, which was critical to distinguish words from pseudowords, was heard. In a control experiment not to be reported here, we investigated whether this particular ISI of 250 ms was important. In this control experiment, an ISI of 0 ms was used and 45 new participants were tested. The results were almost identical with those of the present experiment and exactly the same contrasts were significant. Therefore, we report data from only a single experiment.

Results and Discussion

Analyses on the items were done first. Two items (one beginning embedded and one end embedded) were discarded because of an experimenter error in the visual probes. After exclusion of these items, the overall error rate was 2% and equally distributed across all conditions. The mean lexical decision latencies are shown in Table 1. The data were analyzed in a 2 × 5 design (Location [beginning vs. end embedded] × Prime Type). In the participant analysis, all variables were repeated measures; in the item analyses, Location was a between-item measure.

In an overall analysis of variance, there was a main effect of Prime Type, $F_1(4, 176) = 6.75, p < .001; F_2(4, 264) = 7.79, p < .001$, and the interaction between Prime Type and Location was significant in the participant analyses $F_1(4,$

Table 1

Mean Reaction Times (RT, in ms) and Priming Effects in Experiment 3

Condition	Spoken prime	Visual target	RT	Priming
Begin words				
Original begin word	vel	HUID	480	32**
Embedded begin word	velg	HUID	504	8
Control begin word	reus	HUID	512	
Embedded begin nonword	velk	HUID	486	23**
Control begin nonword	beus	HUID	509	
End words				
Original end word	wijn	ROOD	484	18*
Embedded end word	zwijn	ROOD	498	4
Control end word	kwast	ROOD	502	
Embedded end nonword	twijn	ROOD	503	−5
Control end nonword	kwest	ROOD	498	

* $p < .05$ in F_1 and F_2. ** $p < .01$ in F_1 and F_2.

$176) = 3.88, p < .005$, but failed significance in the item analysis, $F_2(4, 264) = 2.33, p = .057$. Planned comparisons[1] showed the following results. When targets were preceded by an original beginning word (*vel*-HUID), latencies were 32 ms faster if compared to the appropriate control condition (*reus*-HUID), $F_1(1, 44) = 26.21, p < .001; F_2(1, 33) = 20.43, p < .001$. Similarly, when items were preceded by an original end word (*wijn*-ROOD), latencies were 18 ms faster as compared with the control condition for end words (*kwast*-ROOD), $F_1(1, 44) = 6.60, p < .02; F_2(1, 33) = 6.00, p < .02$. So, all original words produced significant priming effects.

For embedded words, only the 23-ms difference between beginning-embedded nonwords (*velk*-HUID) versus beginning-nonword controls (*beus*-HUID) reached significance, $F_1(1, 44) = 10.04, p < .003; F_2(1, 33) = 15.68, p < .001$. When beginning-embedded nonwords (*velk*-HUID) were compared with beginning-embedded words (*velg*-HUID), there was a significant lexical effect with the nonword carriers 18 ms faster than word carriers $F_1(1, 44) = 8.46, p < .006; F_2(1, 33) = 8.17, p < .007$. For end-embedded words and nonwords, no comparison reached significance (all $ps > .10$).

To summarize, for end-embedded words (*wijn*), no sign of lexical activation could be observed, either in words (*zwijn*) or in pseudowords (*twijn*). This failure cannot be attributed to poor associative relations because the original end words (*wijn*) did facilitate lexical decision times. For end-embedded words to be generated as lexical candidates, it thus seems crucial that their onset matches the onset of a syllable. For beginning-embedded words (*vel* in *velg* and *velk*), there is a match with the onset of a syllable, but in this case, only words embedded in pseudowords (*velk*) primed their related target. It thus shows that the lexical status of the carrier can have an impact on the activation of beginning-embedded words.

General Discussion

In the present study, we argued that listeners are frequently confronted with input that corresponds to more than one word, which, in turn, may cause the speech recognition system to produce false alarms. The focus of this study was whether the system does indeed generate multiple word candidates when words are embedded in other words, and if so, how among the array of possible candidates is the correct one selected. Taking the results of the three experiments together, it appears that they converge on the same conclusion: Embedded words are activated if their onset matches the onset of a syllable, and selection is driven by interword competition. This result most straightforwardly conforms to the predictions of the Shortlist model, but before we elaborate on that, we first summarize the experiments briefly.

Experiment 1 showed that lexical representations of embedded words are activated that do not start at a word onset.

---

[1] Planned comparisons do not need a significant F value.

In Experiment 1, using bisyllabic carriers with a metric SS pattern, we showed that the lexical entry of *boos* is accessed when *framboos* is heard. This result is consistent with the finding that the second syllable of a bisyllabic nonword string activates a cohort of word candidates (Norris, Mc-Queen, & Cutler, 1995; Vroomen & de Gelder, 1995a, 1995b). It also replicates the findings of Shillcock (1990) and extends them in an important way because, in this study, carrier words were presented in isolation and not, as in Shillcock, in a sentence context. One would expect that presenting words in isolation should decrease the chance of observing missegmentations because, obviously, a listener is given a reliable indication of where the word begins. Nevertheless, we observed missegmentations in isolated words. This shows that missegmentation in speech is a robust phenomenon because erroneous lexical hypotheses are generated even when the participant "knows" that only one word is presented for which there is no ambiguity about its onset. Therefore, the finding is strong evidence against sequential models, such as the early Cohort (Marslen-Wilson & Welsh, 1978). By contrast, continuous activation models may accommodate the results, but this is dependent on various other parameters, which we will discuss later.

Experiment 2 replicated and extended the results of Experiment 1 using carriers with a metric WS pattern. Using comparable English stimuli, Shillcock (1990) did not find any activation of embedded words in WS carriers. However, in the present study, we again observed that words were activated if embedded at the end of a bisyllabic carrier. The contrast with Shillcock's study is, at this stage, left unexplained. It might be that the sentence context as used by Shillcock induced a difference, but it is also possible that some other unknown cross-language difference plays a role. Whatever the reason, as concerns the activation of embedded words in Dutch, our conclusion can remain unchanged because there is no difference from Experiment 1.

In Experiment 3, we compared beginning- and end-embedded words having an onset that did or did not match a syllable boundary. The results for end-embedded words not matching a syllable boundary showed that these words were never generated as potential candidates: *zwijn* did not activate *wijn*. In principle, this failure may be accounted for by early Cohort, Shortlist, or TRACE without any modification. In terms of the early version of Cohort, the embedded word did not start at word onset; for Shortlist, the embedded word did not coincide with a strong syllable; and for TRACE, interword competition from the carrier or the lexical gang that was activated by the onset of the carrier might prevent the embedded word from becoming active. The results do not distinguish among these models. However, the results show that syllable boundaries—or maybe more critical, their acoustic-prosodic correlate—are important in lexical access. Taking the results for end-embedded words together, then, they underscore the relevance of a match *from a syllable boundary*.

The second finding in Experiment 3 was that beginning-embedded words survived mismatching acoustic information only if embedded in pseudowords (*vel* in *velk*) but not in words (*vel* in *velg*). This result is congruent with propos-

als in which there is competition at the lexical level, such as TRACE and Shortlist. In TRACE and Shortlist, the lexical status of the carrier should matter because words directly inhibit the activation of other words with which they overlap. So only words but not nonwords can compete with each other. In the revised version of Cohort (Marslen-Wilson, 1993), there is no such direct competition. In revised Cohort, words have an influence on each other during recognition such that the difference in the activation of the candidates determines the ease of recognition. However, the activation level of the word itself is only a function of its bottom-up input. Words receive facilitation from matching input and inhibition from mismatching input, but lateral effects between competitors do not play a role. These assumptions, however, cannot account for the present data. For example, in terms of bottom-up input both *g* and *k* in *velg* and *velk* (and because of coarticulatory influences, probably *l* as well) do not match the canonical lexical representation of *vel*. So, both *velg* and *velk* are expected to deliver equal amounts of facilitatory and inhibitory bottom-up input to *vel*. However, it was only the nonword *velk* that produced a priming effect, not *velg*. So, it rather seems that the activation of a word is also driven by inter-word competition because the lexical representation of *velg*, but not of *velk*, can inhibit *vel*.

However, as argued by Marslen-Wilson (1993), direct lateral inhibition at the lexical activation level may not be necessary to account for the difference between word and pseudoword carriers. Competition effects may also emerge at a recognition stage and not at the activation level per se (for the different predictions, see Vroomen & de Gelder, 1995a). One may account for the difference between *velk* and *velg* by maintaining that it reflects the output of the recognition level. Thus, when *velg* is heard, only *velg* is recognized, so *vel* will not be able to facilitate its target. By contrast, when *velk* is heard, no word is recognized, so *vel* may, in retrospect, be the best candidate that is left. This argument thus may account for the different priming effects of *velg* and *velk*. However, the problem is that there is only a single candidate at the output of the recognition level because, obviously, only one word is recognized at a time. So, if cross-modal facilitation reflects the output of the recognition stage, then only a single word can be held responsible. This monolithic nature of the recognition level conflicts with the parallel priming effects as found by Zwitserlood (1989), in which sequences such as *kapit . . .* primed both GELD and BOOT at the same time. Thus, the idea that cross-modal priming is concerned with decisions at the recognition level is, in this light, untenable.

Taken together, the results converge on the proposal made by Shortlist in the sense that (a) words are continuously generated, (b) words receive extra activation if their onset coincides with a strong syllable, and (c) there is direct competition at the lexical level. The results rule out strong sequential models, such as the early Cohort model. The early Cohort does not generate new words continuously because of its deterministic left-to-right parsing of the input. Therefore, sequential models would not predict activation of *boos* when *framboos* is heard. The later versions of

Cohort, TRACE, and Shortlist are all less deterministic because they generate lexical candidates continuously and, thus, independently of whether a word is recognized or not. However, the later version of Cohort misses competition at the lexical activation level, and it therefore cannot account for the difference between word and pseudoword carriers, as observed in Experiment 3. The TRACE model incorporates lateral inhibition, but lexical inhibition also prevents *boos* becoming active when hearing *framboos*. The interword competition in TRACE is thus responsible for a winner-take-all principle that conflicts with the data of the present study. Therefore, it seems necessary to combine interword competition with a boost for words that coincide with strong syllables. This is the key feature of the Shortlist model. In terms of parameter settings, the boost should be so powerful that it can overcome, at least very temporarily, the lexical competition effect.

At a more general level, the study provides evidence for the notion of lexical competition and metric information. As originally proposed in the MSS (Cutler & Norris, 1988), metric information serves to increase the efficacy of lexical segmentation. As in English, words in Dutch are most likely to start at the onset of a strong syllable (see Vroomen & de Gelder, 1995b). A listener may therefore take a strong syllable as the onset of a new word. Empirical evidence for the application of the MSS was found in English (for a review, see Cutler, Norris, & McQueen, in press) and more recently in Dutch (Vroomen, van Zon, & de Gelder, 1996). The joint application of the MSS and interword competition enriches this picture and can, we believe, overcome some of the limitations of earlier models of lexical access.

References

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX Lexical database (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

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., & 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., Norris, D., & McQueen, J. (in press). Lexical access in continuous speech: Language-specific realisations of a universal model. In T. Otake & A. Cutler (Eds.), *Phonological structure and language processing: Cross-linguistic studies*. Berlin: Mouton de Gruyter.

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: Psycholinguistic and computational perspectives* (pp. 50–86). Cambridge, MA: MIT Press.

Gow, D. W., Jr., & Gordon, P. C. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 344–359.

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

Marslen-Wilson, W. D. (1993). Issues of process and representation in lexical access. In G. T. Altmann & R. Shillcock (Eds.), *Cognitive models of speech processing: The second Sperlonga meeting* (pp. 187–210). Hillsdale, NJ: Erlbaum.

Marslen-Wilson, W. D., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review, 101*, 3–33.

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., 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.

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

Norris, D. G., McQueen, J. M., & Cutler, A. (1995). Competition and segmentation in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1209–1228.

Shillcock, R. (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.

Vroomen, J., & de Gelder, B. (1995a). Lexical inhibition in spoken word recognition. *Proceedings of the Fourth European Conference on Speech Communication and Technology*, Madrid, Spain, 1711–1714.

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

Vroomen, J., van Zon, M., & de Gelder, B. (1996). Cues to speech segmentation: Evidence from juncture misperceptions and word spotting. *Memory and Cognition, 24*, 744–755.

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

Appendix A

Materials of Experiment 1

Item	Spoken prime	Embedded word	Target word	Item	Spoken prime	Embedded word	Target word
1	accent	cent	geld	16	klavier	vier	vijf
2	advies	vies	vuil	17	kolos	los	vast
3	akkoord	koord	touw	18	lakei	kei	steen
4	atleet	leed	verdriet	19	libel	bel	deur
5	azuur	zuur	zoet	20	nihil	hiel	hak
6	balans	lans	ridder	21	olijf	lijf	lichaam
7	cacao	kou	warm	22	opaal	paal	water
8	figuur	guur	koud	23	papil	pil	ziek
9	framboos	boos	kwaad	24	prive	vee	koe
10	galant	land	zee	25	roman	man	vrouw
11	galei	lei	dak	26	safier	fier	trots
12	gazon	zon	maan	27	sate	thee	koffie
13	ivoor	voor	achter	28	seizoen	zoen	kus
14	kalief	lief	stout	29	terrein	rein	schoon
15	kameel	meel	brood	30	vampier	pier	worm

Appendix B

Materials of Experiment 2

Item	Spoken prime	Embedded word	Target word	Item	Spoken prime	Embedded word	Target word
1	belang	lang	kort	19	percent	cent	geld
2	beschuit	schuit	boot	20	bezwaar	zwaar	licht
3	gelijk	lijk	dood	21	bezoek	zoek	kwijt
4	brevet	vet	dik	22	gezin	zin	woord
5	sezoen	zoen	kus	23	beroep	roep	schreeuw
6	gering	ring	vinger	24	verraad	raad	advies
7	terrein	rein	schoon	25	gelei	lei	krijt
8	meneer	neer	op	26	secreet	kreet	gil
9	prelaat	laat	vroeg	27	vergift	gift	cadeau
10	vervoer	voer	eten	28	gewei	wei	koe
11	bestuur	stuur	auto	29	vertoon	toon	muziek
12	gebed	bed	slaap	30	verschil	schil	appel
13	verweer	weer	regen	31	benul	nul	niks
14	bevel	vel	huid	32	bretel	tel	seconde
15	vertrek	trek	honger	33	verstaan	staan	zit
16	gerecht	recht	krom	34	fregat	gat	zwart
17	verkeer	keer	maal	35	bedrijf	drijf	nat
18	bezit	zit	stoel	36	meteen	teen	voet

(Appendixes continue)

Appendix C

Materials of Experiment 3

Item	Original	Word embedded	Visual target	Item	Original	Word embedded	Visual target
		Begin words				Begin words (cont'd)	
1	bon	bont	politie	19	mees	meest	vogel
2	brie	brief	kaas	20	mes	mest	snijden
3	bron	brons	water	21	moe	moed	slaap
4	bui	buil	regen	22	nee	neef	ja
5	faa	faal	douche	23	paar	paars	twee
6	haas	haast	konijn	24	pen	pens	schrijven
7	hal	halt	gang	25	pin	pint	geld
8	hel	help	duivel	26	plan	plant	idee
9	hek	heks	tuin	27	ree	reep	hert
10	keu	keus	biljart	28	rij	rijk	wachten
11	koor	koord	zingen	29	tol	tolk	draai
12	kou	kous	winter	30	vel	velg	huid
13	kus	kust	zoen	31	ver	vers	weg
14	la	laaft	kast	32	zee	zeep	strand
15	maag	maagd	eten	33	por	port	duw
16	maan	maand	zon	34	bas	bast	gitaar
17	mal	mals	gek	35	gal	galg	vies
18	man	mand	vrouw				
		End words				End words (cont'd)	
1	laag	blaag	hoog	19	laan	slaan	boom
2	roos	broos	bloem	20	reuk	kreuk	neus
3	ring	kring	trouwen	21	ruit	kruit	glas
4	rat	krat	muis	22	roet	groet	zwart
5	wijn	zwijn	rood	23	lap	klap	stop
6	lomp	klomp	dom	24	wik	kwik	weeg
7	les	fles	school	25	lijk	slijk	dood
8	los	klos	vast	26	ros	gros	paard
9	rol	trol	toneel	27	lok	slok	haar
10	lans	glans	ridder	28	laat	plaat	vroeg
11	wal	kwal	schip	29	nul	knul	niets
12	roei	groei	boot	30	riem	priem	broek
13	raak	kraak	mis	31	lip	slip	mond
14	rond	grond	bal	32	lof	slof	eer
15	lucht	klucht	blauw	33	lang	slang	kort
16	lens	flens	oog	34	ras	gras	hond
17	week	kweek	dag	35	rang	drang	orde
18	rem	brem	auto				