Lexical access of resyllabified words: Evidence from phoneme monitoring

JEAN VROOMEN
Tilburg University, Tilburg, The Netherlands

and

BEATRICE DE GELDER
Tilburg University, Tilburg, The Netherlands,
and Université Libre de Bruxelles, Brussels, Belgium

Resyllabification is a phonological process in which consonants are attached to syllables other than those from which they originally came. In four experiments, we investigated whether resyllabified words, such as "my *bike* is" pronounced as "mai.*bai.k*is," are more difficult to recognize than nonresyllabified words. Using a phoneme-monitoring task, we found that phonemes in resyllabified words were detected more slowly than those in nonresyllabified words. This difference increased when recognition of the carrier word was made more difficult. Acoustic differences between the target words themselves could not account for the results, because cross-splicing the resyllabified and nonresyllabified carrier words did not change the pattern. However, when nonwords were used as carriers, the effect disappeared. It is concluded that resyllabification increases the lexical-processing demands, which then interfere with phoneme monitoring.

Spoken word recognition entails the matching of a sensory input with a stored lexical representation. Most research has focused on the details of the matching process. An issue that has received considerable attention during the last few years is concerned with the alignment or segmentation of the speech signal. The central question is how listeners know where words begin, in the absence of reliable acoustic cues for word boundaries. This issue has turned out to be an intricate one, because, in normal connected speech, there are no reliable acoustic correlates of word boundaries that function like the white space does in written language. Rather, it turned out that the "silence" that we hear between words is in the head of the speaker, and not in the signal.

One approach to the segmentation problem that has received empirical support is that the recognition system takes the beginning of a syllable as the beginning of a word. At first sight, this seems to be a plausible hypothesis, because words generally start at the beginning of a syllable. A syllabic segmentation procedure would be beneficial, when compared with phoneme-sized units, because the majority of lexical access attempts would be successful. In their seminal study, Mehler, Dommergues,

Frauenfelder, and Segui (1981) observed that French listeners were faster to detect a segment if it corresponded exactly to the first syllable of a word, rather than comprising either more or less than a syllable. Subjects were faster to detect *ba* in *ba.llon* than in *bal.con* (the dot indicates a syllable boundary), and they were faster to detect *bal* in *bal.con* than in *ba.llon*. The account given was that the syllable intermediates between the acoustic signal and the lexicon and that the recognition system classifies the speech stream into syllable-sized units.

More recent evidence for a role of syllable boundaries in speech segmentation was obtained in a study by Vroomen and de Gelder (1997a). They used cross-modal associative priming and observed that words embedded in other words were activated if their onset matched the beginning of a syllable. For example, hearing the Dutch word *fram.boos* (*raspberry*) facilitated lexical decision of the visual target *kwaad* (*angry*) via the activation of the embedded word *boos* (*angry*). Crucially, no effect of end-embedded words was observed when the onset did not match the beginning of a syllable: For example, *wijn* (*wine*) as embedded in *zwijn* (*swine*) did not prime its associate *rood* (*red*), whereas *wijn* presented in isolation did.

Thus, it seems likely that listeners take the beginning of a syllable as the onset of a word. However, this strategy fails on a number of occasions. First, not every syllable is a word, simply because there are many multisyllabic words. If the system takes each syllable as the onset of a word, too many candidates are active. This problem of multiple activation was, in fact, demonstrated in the previously mentioned study of Vroomen and de Gelder (1997a), where *fram.boos* activated *boos*. Ultimately, however, the system will have to discard the wrongly ac-

tivated candidate, because *framboos*, and not *boos*, should be recognized.

There are several solutions to the problem of multiple activations. First, selection among multiple activated candidates may be controlled by lexical competition. For instance, in TRACE (McClelland & Elman, 1986) and Shortlist (Norris, 1994), active words compete with each other by inhibiting the activation of other words (for empirical evidence concerning lexical competition, see McQueen, Norris, & Cutler, 1994; Norris, McQueen, & Cutler, 1995; Vroomen & de Gelder, 1995). Hence, embedded words such as *boos* in *framboos* may be activated but not recognized, because of lexical inhibition from *framboos* on *boos*. Second, not every syllable may be considered to be a possible word onset. For example, in the Metrical Segmentation Strategy (MSS), as proposed by Cutler and Norris (1988), only strong syllables (i.e., syllables with unreduced vowels) are considered to be word onsets, but weak syllables (i.e., syllables with schwa as a vowel) are not (see, also, Vroomen, van Zon, & de Gelder, 1996, for Dutch). Moreover, there may also be prosodic cues, such as word stress (Vroomen, Tuomainen, & de Gelder, 1998) or trochaic rhythm (Vroomen & de Gelder, 1997b), that signal where word boundaries are likely to occur. For example, Finnish is a language with fixed word-initial stress, and, consequently, listeners take stressed syllables to be a word onset (Vroomen et al., 1998).

Another problem for a syllabic segmentation strategy which is virtually absent from discussion in the literature concerns resyllabification, or its less extreme case, ambisyllabicity. Resyllabification is a phonological process in which a consonant is attached to another syllable than that from which it originally came. In the traditional view, the syllable has an obligatory nucleus, usually the vowel (V), preceded by an optional consonantal (C) onset and followed by an optional consonantal coda (Kahn, 1976). A primitive syllable inventory consists of {CV, VC, V, CVC}, but many languages have also more complex syllable structures, such as CVCC (such as *milk*), CCVCC (such as *priest*), or CCCVCC (such as *screamed*). In general, it is preferred that prevocalic consonants occupy the syllable onset, in order to avoid an onsetless syllable. The coda of a preceding syllable can therefore be attached to an onsetless nucleus of the following syllable. In many languages, the onsetting of a prevocalic consonant takes place even across a word boundary (Kenstowicz, 1994). Dutch is a well-known example, but one also finds resyllabification in French or English. For example, a phrase such as *my bike is* may be pronounced as *my.bi.k is*, so that the *k* of *bike* is resyllabified across a word boundary to the next vowel-initial word. Alternatively, one may argue that the /k/ is not exclusively joined to *is*, but that it is ambisyllabic in the sense that /k/ is shared by two syllables. Whatever the interpretation, the consequence is that words may not always start at the onset of a syllable (like *is* in *kis*), and words may contain partial syllables (such as the *k* in *bi.kis*).

If, indeed, the speech system takes syllable boundaries to be word boundaries, one may expect that resyllabified words will be more difficult to recognize than words in their canonic syllabic pattern. This is, in fact, demonstrated by the results of Cutler and Norris (1988), who found that English listeners had difficulty detecting words such as *mint* in *min.tayf*. In Dutch, a similar result was obtained by Vroomen et al. (1996), who found that listeners had difficulty detecting *melk* (*milk*) in *mel.koos*.

Another observation in accordance with the resyllabification prediction was made by Dejean de la Bâtie and Bradley (1995). They used a word-initial phoneme-monitoring task and found that French liaison phonemes were more difficult to detect than nonliaison phonemes. Liaison phonemes are latent word-final consonants that surface (i.e., they are pronounced) when followed by a vowel-initial word. The phoneme is then resyllabified to the next word. As an example, the /t/ in *petit* (*small*) is silent when the word is pronounced in isolation (*pe.ti*), but when followed by *éléphant* (*elephant*), it is pronounced as pe.ti.té.lé.phant. The authors observed that a word-initial phoneme such as the /t/ in *talent* was more difficult to detect if preceded by *petit* than by *vrai* (*real*). The crucial difference is that *petit*, but not *vrai*, has a latent /t/ that may surface when followed by a vowel. Hence, the /t/ in *petit talent* may belong to both *petit* and *talent*, but the /t/ in *vrai talent* can only belong to *talent*. The authors argued that the ambiguity of the /t/ in *petit talent* hampered recognition of the carrier word, which then interfered with phoneme-monitoring latencies.

However, a number of alternative explanations for this finding remain open. One possibility is that recognizing resyllabified words is indeed genuinely difficult. One can argue, then, that resyllabification requires more processing capacity or attentional demands, which then interferes with phoneme monitoring (see, e.g., Foss, 1969; Pitt & Samuel, 1990; Wurm & Samuel, 1997). The results of Dejean de la Bâtie and Bradley (1995) can also be accounted for by race models of phoneme monitoring (Cutler, Mehler, Norris, & Segui, 1987; Foss & Blank, 1980). For example, in the race model of Cutler et al., phoneme detection is based on a race between a phonetic and a phonological code. The phonetic code is derived from the acoustic signal, and the phonological code is derived from the lexical representation of the word. The decision about target presence is made on the basis of the first available code. If resyllabified words are more difficult to recognize, their phonological code becomes available relatively late, and this may slow down phoneme-monitoring times. Alternatively, at this stage it cannot be ruled out that there are phonetic differences between resyllabified and nonresyllabified phonemes that slow down the phonetic code. Either way, a race model may also predict that phoneme monitoring of resyllabified words is slower than that of nonresyllabified words.

In the present study, we tried to obtain a better understanding of the effects of resyllabification on word recognition by using a generalized phoneme-monitoring task. In the study of Dejean de la Bâtie and Bradley (1995), the authors argued that the difficulty in detecting liaison phonemes stemmed from a difficulty in recognizing the

carrier word. However, listeners were required to respond to word-initial phonemes only, and they were thus more or less forced to recognize the word before a response could be made. At present, though, it is unclear what happens if there is no requirement to respond to word-initial phonemes only. Aside from a genuine difficulty in recognizing resyllabified words, there may be a task-specific effect that only shows up with word-initial phoneme monitoring, but not in other tasks. For example, it may be that listeners were confused about whether the /t/ in *petit talent* was the final phoneme of *petit* (no response required) or the initial phoneme of *talent* (a response required). This confusion about the position of the phoneme in the word may affect word-initial phoneme monitoring but not lexical access in general, because the speech processing system may not need to compute the absolute phoneme position within a word. Hence, if there is indeed a genuine difficulty in recognizing resyllabified words, it is important to demonstrate it with a task in which phoneme position per se is not crucial. In the present study, we therefore used a generalized phoneme-monitoring task. In generalized phoneme monitoring, the position of the target phoneme is not relevant, because listeners respond to target presence irrespective of the position within a carrier word. If there is a genuine difficulty in recognizing resyllabified words, the effect of resyllabification should also be obtained with generalized phoneme monitoring. On the other hand, if the difficulty in detecting resyllabified phonemes is a task-specific effect that is only obtained with word-initial phoneme monitoring, the effect should disappear with generalized phoneme monitoring.

In Experiment 1, listeners monitored target phonemes that were resyllabified or not, depending on whether the word that followed the target started with a vowel or a consonant, respectively. In Experiment 2, we made recognition of the carrier word more difficult in order to examine whether our results reflected lexical processing demands of the carrier word. In Experiment 3, we cross-spliced the carrier word in order to test whether the carrier word itself or the context that followed the carrier word was responsible for the observed effects. Finally, in Experiment 4, we used nonwords as carriers, to test whether there were inherent phonetic differences between phonemes that were followed by vowels and those that were followed by consonants.

EXPERIMENT 1

A generalized phoneme-monitoring task was used to investigate the effect of resyllabification on spoken word recognition. The task of the subjects was to detect a prespecified target phoneme, irrespective of its position within a word. The target phoneme was either resyllabified or not, depending on the context that followed the target phoneme. In the resyllabified case, the phoneme occupied the onset position of a syllable; in the nonresyllabified case it was in the coda position. For example,

in the resyllabified condition, listeners had to detect the target phoneme /t/ as embedded in the Dutch sentence fragment *de boot is gezonken* (*the boat is sunk*) with a syllable structure as in *de. boo.t is. ge.zon.ken*. The vowel of *is* attracts the *t* so that it is not an onsetless syllable anymore. The critical *t* can, therefore, be considered as resyllabified across a word boundary. Alternatively, one can argue that the *t* is ambisyllabic, because it is shared by two words. In the nonresyllabified condition, resyllabification of the target phoneme was prevented, because there was a prosodic, a syntactic, and a phonotactic boundary that blocked resyllabification. As an example, the control condition for the previous sentence was *de boot, die gezonken is* (*the boat, which sunk is*) with a syllable structure as in *de. boot. die. ge.zon.ke.n is*. The syntactic boundary of the subordinate clause *die gezonken is* was signaled prosodically by a short silence that prevented resyllabification. Moreover, the voiced /d/ of *die* has to be a syllable onset, because /d/ cannot be in coda position in Dutch. In the nonresyllabified condition, there was, therefore, no ambiguity as to which word the critical target phoneme belonged to. A typical example of the waveforms is presented in Figures 1 and 2.

We also varied whether carrier words were unique at word offset or not (in terms of the uniqueness point of the cohort model; Marslen-Wilson, 1984). A word such as *boot* is not unique before the *t*, because there are other words such as *boon* (*bean*) or *boog* (*bow*) in the lexicon. In contrast, *kist* (*box*) is unique before *t*, because there is no other word that starts with *kis*. In dual-code models of phoneme monitoring, it is generally maintained that phoneme detection is based on a race between a phonetic and a phonological code (Cutler, Mehler, Norris, & Segui, 1987; Dell & Newman, 1980; Foss & Blank, 1980). Some authors have argued that the phonological code becomes available as soon as the word is unique (Marslen-Wilson, 1984). The phonological code may, therefore, be available sooner in early-unique words than in late-unique words. One expects, therefore, phoneme detection in early-unique words to be faster than in late-unique words. Moreover, early-unique words were unique before they were resyllabified, so phoneme detection in early-unique words may suffer less from resyllabification than does that in late-unique words.

Method

Subjects. A group of 20 students from Tilburg University was tested. They were equally divided across the two versions of the test.

Materials. The materials were constructed around 40 words—half of them unique before the final phoneme (early-unique words), the other half unique at the final phoneme or later (late-unique words). The Dutch lexical inventory is such that most monosyllabic early-unique words end with a consonant cluster and most late-unique words end with a single consonant. The distinction between early-unique and late-unique words is, therefore, almost completely confounded by word type ending: All but one late-unique word ended in a singleton consonant, and all but two early-unique words ended in a consonant cluster. Each word was embedded in two sentences: (1) a sentence in which the critical target phoneme was resyllabified to the vowel of the next word (e.g., *t* in *de boot is gezonken*, with a
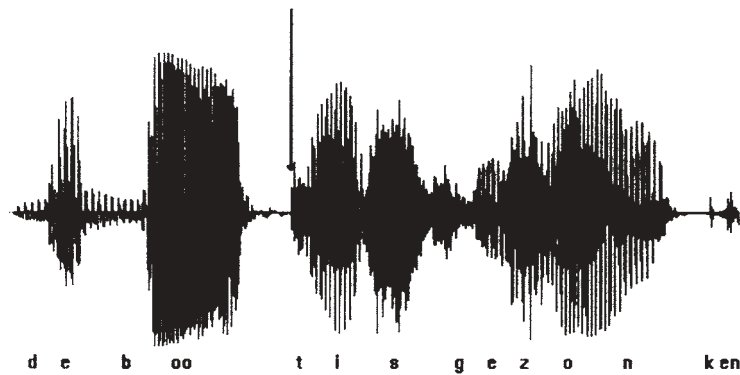
Figure 1. The waveform of the resyllabified sentence *de. boo.t is. ge.zon.ken*.

syllable structure as in *de. boo.t is. ge.zon.ken* and (2) a sentence in which the target phoneme was not resyllabified, because there was a prosodic, a syntactic, and a phonotactic boundary blocking resyllabification, as in *de. boot. die. ge.zon.ke.n is*. There were, therefore, four conditions: early-unique or late-unique words in resyllabified or nonresyllabified context.

All carrier words were monosyllabic nouns selected from the CELEX dictionary (Baayen, Piepenbrock, & van Rijn, 1993; for a complete list, see URL: http://cwis.kub.nl/~fsw_1/psychono/persons/jvroomen/resyl_appendix.htm). The critical phoneme was always a voiceless stop consonant (/p/, /t/, or /k/) in word-final position. The mean logarithmic frequency of occurrence of the words was 1.19 for early-unique words and 1.04 for late-unique words [$t(38) = 1.01$, n.s.]. Another 40 filler sentences, with *p*, *t*, or *k* as target, served as no-go trials. Before testing, 10 trials were given as practice.

Design and Procedure. Two counterbalanced versions of the test were made, so that a subject heard each carrier word only once, either in the resyllabified or in the nonresyllabified context. There were 40 experimental trials in each version (10 for each of the four conditions). The experimental trials were pseudorandomly interspersed with the 40 no-go filler trials. Fillers and each member of an experimental trial pair appeared in exactly the same location across the two sets.

The sentences were spoken by a male speaker of Dutch. They were recorded in a sound-treated studio on digital audiotape (Sony DAT-55). The sentences were digitized at 22.05 kHz (16-bit preci-

sion), and the onset of the critical target phoneme was determined under visual and auditory control. Reaction time (RT) was measured from the onset of the burst of the target phoneme.

The subjects were asked to press a button as rapidly as possible whenever they heard a previously specified target phoneme. The targets were shown for 1,500 msec on a computer screen. The sentence was played back from a soundcard 500 msec after presentation of the target. The sentences were presented via Sennheiser HD-410 headphones at a comfortable listening level. A session lasted about 12 min.

Results

In this and all the other experiments, RTs below 100 msec and above 1,000 msec were discarded: 3.6% of the data were left out this way. The subjects did not respond on 0.3% of the items (misses). Mean RTs are presented in Table 1. Analyses of variance (ANOVAs) were performed, with subjects and items as random factors. In the subject analyses, resyllabification (nonresyllabified vs. resyllabified) and word type (early-unique vs. late-unique words) were within-subjects variables; in the item analyses, resyllabification was a within-items variable, and word-type was a between-items variable. Phoneme detection latencies for nonresyllabified words were, on
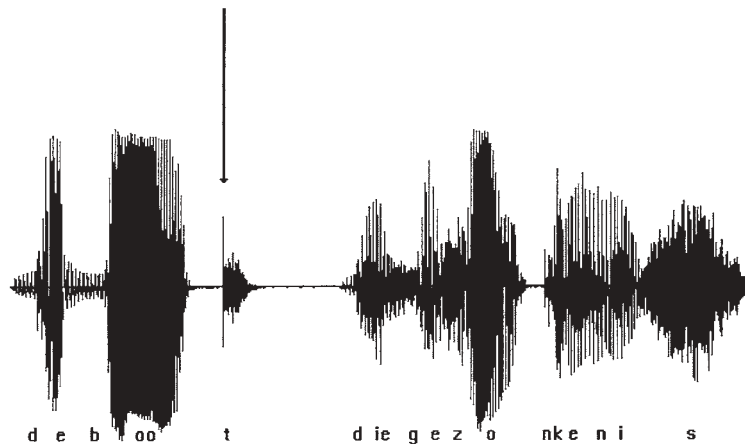

Figure 2. The waveform of the nonresyllabified phrase *de. boot. die. ge.zon.ke. nis*.

Table 1
Mean Phoneme Detection Latencies
(in Milliseconds) in Experiment 1

Word Type	Context		
	Nonresyllabified	Resyllabified	Difference
Early-unique	268	309	41
Late-unique	252	352	100

average, 70 msec faster than those for resyllabified words $[F_1(1,19) = 31.56, p < .001; F_2(1,38) = 68.22, p < .001]$. There was no main effect of early-unique versus late-unique words (all $ps > .10$), but there was a significant interaction between resyllabification and word type $[F_1(1,19) = 12.68, p < .002; F_2(1,38) = 10.54, p < .002]$. Separate t tests confirmed that both early-unique and late-unique words suffered from resyllabification [early-unique words, $t_1(19) = 2.67, p < .02; t_2(19) = 3.78, p < .001$; late-unique words, $t_1(19) = 6.58, p < .001; t_2(19) = 7.69, p < .001$], but the resyllabification effect was larger for late-unique than for early-unique words [100 vs. 41 msec; $t_1(19) = 3.56, p < .002; t_2(38) = 3.25, p < .002$].

A correlational analysis was conducted on the logarithmic frequency of occurrence of the carrier word and the phoneme monitoring latencies. A negative correlation is expected if there is a lexical involvement, because high-frequency carrier words should have faster latencies. The correlation was only significant in early-unique nonresyllabified words, indicating that high-frequency words had faster monitoring latencies than did low-frequency words [$r(20) = -.53, p < .02$]. For the other conditions, correlations were in the same direction but nonsignificant [nonresyllabified late-unique words, $r(20) = -.25$, $p = .27$; resyllabified early-unique words, $r(20) = -.16$, $p = .49$; nonresyllabified late-unique words, $r(20) = -.13, p = .56$].

Discussion

The main result of Experiment 1 is that nonresyllabified phonemes were detected faster than resyllabified phonemes. This result, as obtained with generalized phoneme monitoring, is in line with the results of Dejean de la Bâtie and Bradley (1995), who used word-initial phoneme monitoring. This convergence supports the idea that the difficulty of detecting resyllabified phonemes stems from a difficulty in recognizing resyllabified words.

Another finding was that the resyllabification effect was larger for late-unique than for early-unique words. At this stage, there are several ways to account for this difference. In terms of a race model (Cutler et al., 1987; Dell & Newman, 1980; Foss & Blank, 1980), early-unique words may suffer less from resyllabification because their phonological code is already available before the word is resyllabified. In attentional models of phoneme monitoring (Foss, 1969; Pitt & Samuel, 1990; Wurm & Samuel, 1997), a similar prediction can be made by assuming that early-unique words require less processing

capacity than do late-unique words. Early-unique words may be recognized before they are resyllabified, which frees attentional capacity for the phoneme-monitoring task.

However, one should be cautious in explaining the difference between early-unique and late-unique words in terms of a lexical involvement. Most early-unique words ended in a consonant cluster, but late-unique words ended in a single consonant. It seems plausible that consonants in a cluster are much more predictable than singleton consonants, because there are heavy constraints on what constitutes a legal cluster. The final phoneme of early-unique words is, therefore, much more constrained in terms of transitional probability than that of late-unique words, and listeners may use this information to predict the upcoming phoneme, independent of the lexicon. To investigate whether the resyllabification effect reflects a lexical component, another experiment was conducted, in which recognition of the carrier word was made more difficult.

EXPERIMENT 2

Experiment 2 was similar to Experiment 1, except that noise was added at the onset of the carrier word. The noise was intended to increase the difficulty of recognizing the carrier word but to leave intact the phonetic realization of the target phoneme itself. It is well known that phoneme monitoring can be modulated by attentional factors (Foss, 1969; Pitt & Samuel, 1990; Wurm & Samuel, 1997), and it seems plausible that adding noise at the onset of a word increases the attentional demands in recognizing the word. If our resyllabification effect indeed reflects resource limitations in word recognition, one may expect the effect to increase when recognition of the word becomes more difficult.

Method

Subjects. Eighteen students from Tilburg University took part in the experiment. They were equally divided across the two versions of the test. None of them had participated in the previous experiment.

Materials. The materials were exactly the same as those in the previous experiment, except that a 50-msec burst of white noise was digitally added to the original waveform at the onset of the carrier word. The amplitude of the burst was approximately one fourth of the peak amplitude of the average vowel. The acoustic realization of the target phoneme and the preceding vowel was not changed by this manipulation, and all the words were still understandable.

Design and Procedure. The design and procedures were exactly the same as those in the previous experiment. All the items appeared in the same order.

Table 2
Mean Phoneme Detection Latencies (in Milliseconds)
in Experiment 2 with Noise Added

Word Type	Context		
	Nonresyllabified	Resyllabified	Difference
Early-unique	441	532	91
Late-unique	429	562	133

Results

In Experiment 2, 4.6% of the data was discarded because of time-outs. The subjects did not respond on 1.1% of the items (misses). The mean RTs are presented in Table 2. As can be seen, adding noise to the onset of the carrier word slowed responses by almost 200 msec, in comparison with Experiment 1, and the overall difference between resyllabified and nonresyllabified words was increased.

In the ANOVA on the RTs, there was a main effect of resyllabification because resyllabified phonemes were detected more slowly than nonresyllabified phonemes [$F_1(1,17) = 35.97, p < .001; F_2(1,38) = 51.24, p < .001$]. There was no main effect of early-unique versus late-unique words (all ps > .10), and the interaction between context and word type was this time significant by subjects only [$F_1(1,17) = 4.57, p < .05$], but not by items [$F_2(1,38) = 1.31, p > .10$].

In the analyses of the correlations, there were no negative trends anymore between the frequency of the carrier word and phoneme-monitoring latencies. If anything, the correlations tended to be positive for early-unique words and around zero for late-unique words [early-unique nonresyllabified words, $r(20) = .47, p < .04$; early-unique resyllabified words, $r(20) = .46, p < .04$; late-unique nonresyllabified words, $r(20) = -.04$, n.s.; late-unique resyllabified words, $r(20) = -.05$, n.s.].

In order to compare Experiments 1 and 2, an overall ANOVA was conducted, with experiment as a between-subjects and within-items factor. There was a main effect of experiment, because latencies were faster in Experiment 1 than in Experiment 2 [$F_1(1,36) = 33.53, p < .001; F_2(1,38) = 227.15, p < .001$]. There was again a main effect of resyllabification [$F_1(1,36) = 68.22, p = .001; F_2(1,38) = 95.26, p < .001$], because nonresyllabified phonemes were detected more quickly than resyllabified ones. The interaction between word type and resyllabification was significant, because early-unique words suffered less from resyllabification than did late-unique words [$F_1(1,36) = 15.57, p = .001; F_2(1,38) = 5.92, p < .02$]. Finally, there was an interaction in the analysis by items between experiment and resyllabification, which, however, failed to reach significance in the analysis by subjects [$F_1(1,36) = 3.49, p = .07; F_2(1,39) = 6.31, p < .02$]. It indicated, as expected, that the resyllabification effect tended to be smaller in Experiment 1 than in Experiment 2 (70 msec in Experiment 1 vs. 112 msec in Experiment 2). All other effects were nonsignificant (all ps > .10).

Discussion

Experiment 2 confirms and extends the findings of our Experiment 1. Phoneme monitoring of resyllabified words was again more difficult than that of nonresyllabified words. Moreover, the difference tended to increase when recognition of the carrier word was made more difficult. This suggests that resyllabified words require more attentional processing demands than do nonresyllabified words. However, as yet it is unclear where the difficulty stems from. It may be that resyllabified words themselves are more difficult to recognize, possibly because they differ phonetically from nonresyllabified words. Alternatively, it may also be that the context that follows the carrier words, a resyllabifying vowel or a nonresyllabifying consonant, is responsible for the effect. In the next experiment, we tried to examine these possibilities by cross-splicing the target words.

EXPERIMENT 3

In this experiment, we tried to determine whether the resyllabification effect should be attributed to the context that follows the carrier word (a resyllabifying vowel or not) or whether the carrier words themselves differ acoustically between conditions. We tried to control the acoustic differences between carrier words by cross-splicing them from one sentence to the other. If the resyllabification effect stems from acoustic differences between the carrier words, one expects the effect to reverse with cross-splicing. On the other hand, if the vowel or consonant that follows the target is critical, cross-splicing should have no effect, because the context is not changed. Moreover, to test the generality of our findings, we used new items, and instead of a phonotactic, syntactic, and prosodic boundary, we now used a phonotactic boundary only. An example is the Dutch sentence *de. poort. bleef.o.pen* (meaning *the gate remained open*), in which the critical target phoneme *t* cannot be assigned to the next syllable because there are no syllables that start with /tbl/. In contrast, in the sentence *de poort is open* (meaning *the gate is open*), the *t* is assigned to the next syllable (*de. poor.t is. o.pen*), because the vowel of *is* attracts the *t* into *t is*.

Method

Subjects. Twenty students from Tilburg University were tested. They were equally divided across the four versions of the test.

Materials. The materials were constructed around 28 monosyllabic words: Ten of them were unique before their final phoneme (early-unique words); the others were unique at or after the final phoneme (late-unique words). All the early-unique words ended in a consonant cluster; all the late-unique words ended in a single consonant. Each word was embedded in two sentences: (1) a sentence in which the critical target phoneme was resyllabified to the vowel of the next word (e.g., /t/ in *de poort is open*, with a syllable structure such as *de. poor.t is. o.pen*) and (2) a sentence in which the target phoneme was not resyllabified because there was a phonotactic boundary, such as *de. poort. bleef.o.pen*. The resyllabification was always blocked by a word starting with /b/, which, in Dutch, cannot be in a coda position. In order to control for acoustic differences between carrier words, we cross-spliced the critical carrier words from the original sentences, using a speech editor. For example, the word *poort* as excised from the resyllabified utterance replaced *poort* excised from the nonresyllabified utterance, and, vice versa, *poort* from the nonresyllabified utterance replaced *poort* from the resyllabified utterance. All the cuts were made at a zero crossing, and audible clicks were removed. All cross-spliced sentences sounded very natural, without any strange transition. There were thus four condi-

tions: resyllabified or nonresyllabified words that were or were not cross-spliced.

The critical phoneme of the carrier word was always a /t/ or /k/ in word-final position. The mean logarithmic frequency of occurrence of the words was 1.56 for early-unique words and 1.35 for late-unique words [$t(26) < 1$]. Another 28 filler sentences with /t/ or /k/ as target served as no-go trials. Before testing began, 10 trials were given as practice.

Design and Procedure. Four counterbalanced versions of the test were made, so that a subject heard each word only once, either in the resyllabified or in the nonresyllabified cross-spliced or non-cross-spliced context. There were thus 28 experimental trials in each version (7 for each of the four conditions). The experimental trials were pseudorandomly interspersed with the 28 no-go filler trials. Fillers and each member of an experimental quadruple appeared in exactly the same location across the four sets. All other procedures were the same as those in the previous experiments.

Results

The overall error rate was 3.4% and was equally distributed across the four conditions. Table 3 presents the mean RTs for the four conditions. As can be seen, the resyllabification effect was smaller than in the previous experiments, but resyllabified phonemes were still more difficult to detect than nonresyllabified phonemes. This difference was independent of whether the original or the cross-spliced version was heard.

In the 2 (syllabification) × 2 (splicing) ANOVA on the RTs, there was a main effect of resyllabification [$F_1(1,19) = 8.19, p < .01; F_2(1,27) = 5.50, p < .03$], indicating that nonresyllabified phonemes were detected, on average, 31 msec more quickly than resyllabified phonemes. There was neither an effect of splicing nor an interaction between resyllabification and splicing (all ps > .10). We therefore pooled the items over the splicing factor in order to investigate the effect of word type.

The RTs for the early-unique and late-unique words are presented separately in Table 4. The effect of resyllabification was again significant [$F_1(1,19) = 14.10, p < .001; F_2(1,26) = 6.25, p < .02$]. The effect of word type was significant by subjects only [$F_1(1,19) = 7.34, p < .02$], but not by items [$F_2(1,26) = 2.22, p > .10$]. The interaction between word type and resyllabification was not significant (all ps >.10).

We also computed the correlation between the logarithmic frequency of the carrier word and the monitoring latency. All correlations were negative but nonsignificant [early-unique nonresyllabified words, $r(18) = -.08$, n.s.; early-unique resyllabified words, $r(18) = -.09$, n.s.; late-unique nonresyllabified words, $r(10) = -.58, p = .07$; late-unique resyllabified words, $r(10) = -.35$, n.s.].

Table 3
Mean Phoneme Detection Latencies (in Milliseconds)
of Experiment 3 Averaged Across Early/Late Unique Words

Splicing	Context		
	Nonresyllabified	Resyllabified	Difference
Original	295	316	21
Cross-spliced	277	318	41

Table 4
Mean Phoneme Detection Latencies (in Milliseconds)
of Experiment 3 Averaged Across Spliced/Nonspliced Words

Word Type	Context		
	Nonresyllabified	Resyllabified	Difference
Early-unique	257	308	51
Late-unique	302	321	19

Discussion

The main result of Experiment 3 is that, even after cross-splicing, phonemes in nonresyllabified words were detected more quickly than those in resyllabified words. An acoustic explanation in terms of a better realization of the nonresyllabified carrier word is, thus, effectively ruled out. Rather, it seems that the context that follows the target phoneme is critical. The interaction between word type and resyllabification was not significant anymore, suggesting that the difference between early-unique and late-unique words is not very reliable, even though it was still the case that phonemes of resyllabified late-unique words were detected slowest of all.

EXPERIMENT 4

In Experiment 3, we tried to disentangle lexical from phonetic factors. As yet another test, we used nonwords as carriers in the present experiment. If the resyllabification effect stems from a difference in the lexical processing demands of resyllabified and nonresyllabified words, it should disappear when nonword carriers are used, because it seems likely that no lexical access attempt is made when all the items are nonwords. The alternative possibility is that there is a phonetic confound between syllable-final and syllable-initial phonemes. So far, in all our previous experiments, resyllabified phonemes were in syllable-initial position, and nonresyllabified phonemes were in syllable-final position. For phonetic reasons, it may be that syllable-final phonemes are always easier to detect than syllable-initial phonemes. This contrasts with Redford and Diehl (1996), who found that syllable-initial consonants are perceptually clearer than syllable-final consonants. However, they did not use phoneme monitoring, and their results may therefore not apply to our case. Thus, if phonetic differences are to account for the resyllabification effect, one should obtain the same result when nonwords are used as carriers.

Method

Subjects. Twelve students from Tilburg University were tested. They were equally divided across the two versions of the test.

Materials. All the carriers were bisyllabic nonsense strings derived from those of Experiment 3. As an example, the string *oort.blif* replaced the sentence *de. poort. bleef.o.pen*, and *oor.tif* replaced *de. poort. is o. pen*. The same target phoneme was used in the same syllabic position (syllable-initial or syllable-final), and the phonemes that immediately preceded or followed the target were also unchanged. The filler items were changed in a similar way. All other experimental details were exactly the same as those in Experiment 3.

Results

Preliminary analyses showed that one item was missed by all the subjects. Later inspection showed that the target was acoustically not realized. The item was therefore skipped from further analyses. The overall error rate was 3.6% and was equally distributed across the two conditions. The mean RT for syllable-initial targets was 499 msec, and for syllable-final targets, it was 562 msec. Thus, in contrast with previous experiments, syllable-initial phonemes were detected *more quickly* than syllable-final phonemes. In the ANOVA on the RTs, the effect of syllable position was highly significant [$F_1(1,11) = 44.14$, $p < .001$; $F_2(1,26) = 13.52$, $p < .001$].

Discussion

The results of Experiment 4 show that syllable position per se cannot account for the resyllabification effect. With nonwords as carriers, syllable-initial phonemes were detected more quickly than syllable-final phonemes, whereas with words, the opposite was the case. This result, therefore, strengthens the idea that the resyllabification effect should be attributed to a difficulty in recognizing the carrier word, and not to a phonetic difference between syllable-initial and syllable-final phonemes.

GENERAL DISCUSSION

We investigated the effects of resyllabification on word recognition with a phoneme-monitoring task. The results showed that the final phoneme of a resyllabified carrier word was more difficult to detect than that of a nonresyllabified word. When word recognition was made more difficult by adding noise to the onset of the carrier word, the difference increased, suggesting that the effect was modulated by the attentional requirements of recognizing the carrier word. Acoustic differences between resyllabified and nonresyllabified carrier words could be ruled out as a factor, because cross-splicing the carrier words had no effect. Phonetic differences between syllable-initial and syllable-final phonemes could also be ruled out, because the effect was reversed when nonsense strings were used as carriers. Taken together, the results suggest that the difference between resyllabified and nonresyllabified words emerges from a difference in their lexical-processing demands. Thus, resyllabified words are more difficult to recognize, which slows down phoneme monitoring.

An important question, then, is why resyllabified words are more difficult to recognize than nonresyllabified words? One possibility is that words are segmented at the onset of a syllable (Cutler & Norris, 1988; Vroomen et al., 1996). One can argue, then, that there will be two lexical access attempts in resyllabified words (for example, in *boo.tis* at *boo* and at *tis*), but only one in the nonresyllabified case. Moreover, in resyllabified words, the phonetic information may need to be reassembled across a syllable boundary (the *t* belongs to the previous syllable *boo*), but not in nonresyllabified words. From this per-

spective, our results are similar to those that have been obtained with the word-spotting task (Cutler & Norris, 1988), in which English listeners had difficulty detecting words such as *mint* in *min.tayf*, or in which Dutch listeners had difficulty detecting *melk* (*milk*) in *mel.koos* (Vroomen et al., 1996). The present study extends these findings in an important way by showing that this effect also emerges in natural utterances when words are resyllabified across word boundaries.

Our study also puts the results of Dejean de la Bâtie and Bradley (1995) in a somewhat different perspective. Using word-initial phoneme monitoring, they found that potential liaison phonemes (the /t/ in *petit talent*) were more difficult to detect than nonliaison phonemes (the /t/ in *vrai talent*). We argued that there are two alternatives to account for the effect. The difference may be caused by intrusions of phoneme position within the word (i.e., the /t/ may be word final or word initial), or, as was argued by the authors, it may be caused by difficulties in recognizing the carrier word *talent*. Our study allows us to rule out the former explanation, because, with generalized phoneme monitoring, we still obtained a reliable difference between resyllabified and nonresyllabified words.

There may, however, be other possibilities that can account for our data. One may speculate that nonresyllabified words are easier to recognize than resyllabified words because, in the former case, the canonic syllabic pattern of the intended word is in the signal, but in the latter case it needs to be recovered. Thus, the notion may be that the syllabic structure of a spoken word needs to be recovered before the word itself can be recognized. In resyllabified words, this may be more difficult than in nonresyllabified words, and it may interfere with phoneme monitoring. However, this interpretation seems unlikely. There is, at present, no evidence that the syllabic structure of a word needs to be recovered. In fact, there are several reasons why this is unlikely. First, syllabic structure per se does not reduce the number of lexical candidates beyond what is already given in the phoneme sequence. So it is not clear what the functional role of syllabic structure in word recognition should be. Second, many affixes change the syllabic structure of the root. For example, the plural of the Dutch CVC word *boek* (*book*) is the CV.CVC sequence *boe.ken*. So the syllabic structure of the root in affixed words quite often does not match the syllabic structure of the root in its canonic form, and if *boek* and *boe.ken* share a common lexical entry (see, e.g., Marslen-Wilson, Tyler, Waksler, & Older, 1994), it is better that their syllabic structure not be taken into account.

However, this kind of reasoning does point toward an important distinction about the role that a syllable might play in spoken word recognition. We distinguish between a prelexical role, in which a syllable boundary guides speech segmentation, and a postlexical role, in which the syllabic structure of a word may or may not be recovered. As we argued, syllable boundaries may guide prelexical speech segmentation, but this does not imply that the syllabic structure of a word is computed as it is in

models of word production (Levelt, 1989). We conjecture that a syllable boundary serves as a cue for a word boundary, just as a long pause or word-final vowel-lengthening signals word boundaries. The drawback to this procedure, as we have shown, is that, because of resyllabification, the system is sometimes misled.

REFERENCES

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

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1987). Phoneme identification and the lexicon. *Cognitive Psychology*, **19**, 141-177.

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

Dejean de la Bâtie, B., & Bradley, D. C. (1995). Resolving word boundaries in spoken French: Native and nonnative strategies. *Applied Psycholinguistics*, **16**, 59-81.

Dell, G. S., & Newman, J. E. (1980). Detecting phonemes in fluent speech. *Journal of Verbal Learning & Verbal Behavior*, **19**, 608-623.

Foss, D. J. (1969). Decision processes during sentence comprehension: Effects of lexical item difficulty and position upon decision times. *Journal of Verbal Learning & Verbal Behavior*, **8**, 457-462.

Foss, D. J., & Blank, M. A. (1980). Identifying the speech codes. *Cognitive Psychology*, **22**, 609-632.

Kahn, D. (1976). *Syllable-based generalizations in English phonology*. Unpublished doctoral dissertation, MIT, Cambridge.

Kenstowicz, M. (1994). *Phonology in generative grammar*. Cambridge, MA: Blackwell.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Marslen-Wilson, W. D. (1984). Function and processes in spoken word recognition. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance: Control of language processes* (pp. 125-149). Hillsdale, NJ: Erlbaum.

Marslen-Wilson, W. D., Tyler, L., Waksler, R., & Older, L. (1994). Morphology and meaning in the English lexicon. *Psychological Review*, **101**, 3-32.

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, & 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 & Verbal Behavior*, **20**, 298-305.

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, & Cognition*, **21**, 1209-1228.

Pitt, M. A., & Samuel, A. G. (1990). Attentional allocation during speech perception: How fine is the focus? *Journal of Memory & Language*, **29**, 611-632.

Redford, M. A., & Diehl, R. L. (1996, December). *A study on the perceptibility of syllable-initial and syllable-final consonants*. Poster session presented at a meeting of the Acoustical Society of America, Honolulu.

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

Vroomen, J., & de Gelder, B. (1997a). The activation of embedded words in spoken word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 710-720.

Vroomen, J., & de Gelder, B. (1997b, November). *Trochaic rhythm in speech segmentation*. Paper presented at the 38th Annual Meeting of the Psychonomic Society, Philadelphia.

Vroomen, J., Tuomainen, J., & de Gelder, B. (1998). The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory & Language*, **38**, 133-149.

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

Wurm, L. H., & Samuel. A. G. (1997). Lexical inhibition and attentional allocation during speech perception: Evidence from phoneme monitoring. *Journal of Memory & Language*, **36**, 165-187.