Journal of Experimental Psychology:
Learning, Memory, and Cognition
1993, Vol. 19, No. 2, 260–271

0278-7393/93/$3.00

The Role of Syllables in the Perception of Spoken Dutch

Pienie Zwitserlood, Herbert Schriefers, Aditi Lahiri, and Wilma van Donselaar

Three experiments are reported concerning the role of the syllable in the perception of spoken Dutch. Ss monitored spoken words for the presence of target strings that did or did not correspond to the words' first syllable. Effects of syllabic match were obtained for spoken words with unambiguous syllabic structure, as well as for words containing ambisyllabic consonants, which are shared by 2 syllables. For both types of words, monitoring latencies were shorter if the target matched the first syllable of the spoken word. Syllable effects were independent of the relation between targets and stem morphemes of the spoken words. Commonalities and differences between these results and those obtained in other languages such as English and French are discussed.

In spoken language, the input consists of a continuous signal. To achieve access to the mental lexicon, the listener has to match portions of this continuous input against lexical elements, which are represented in the mental lexicon as discrete entities. An important question in speech processing and word recognition is whether and how continuous speech is segmented and mapped onto discrete lexical representations. One possible solution to this problem is to assume that units smaller than the word mediate between the continuous speech input and the mental lexicon. The syllable ranks high among the candidates that have been proposed, and it has received a considerable amount of attention in empirical research.

Often, it is not easy to determine what researchers from different areas mean when they talk about syllables. We define the syllable with respect to its component parts. A syllable minimally consists of a nucleus peak (e.g., [aI], the first person pronoun, or [I] in *trip*) that depending on language-specific constraints, may or may not be preceded and followed by consonantal onsets ([tr] in *trip*, none in *egg*) and codas ([p] in *trip*, none in *tree*). In most instances, the nucleus is a vowel, although in some cases it may be a syllabic consonant, as in the final sound in the English word *sudden* or the German word *lesen*.

Cross-linguistic evidence for the syllable as an organizational unit in phonology is well established (Clements & Keyser, 1983; Hooper, 1972; Kahn, 1980; Selkirk, 1984). Empirical support for the relevance of the syllable can be found in research on speech production (see Levelt, 1989; Meyer, 1992) and on the early phases of first-language acquisition (Mehler, Dupoux, & Segui, 1991). In experimental research on the role of the syllable in speech perception, however, results are more controversial. Research in the Romance languages such as French, Spanish, and Portuguese show positive evidence for the listener's sensitivity to the syllable in speech perception, whereas the results for English are less clear. These results are intriguing because they suggest differences between the languages. Either the languages differ with respect to the availability of cues in the sensory input signaling syllable boundaries or, at a more abstract level, in the number and complexity of syllable types. Both levels could result in different syllabification strategies, different routines to access the lexicon, or both. The latter interpretation is in fact a more popular one. Unfortunately, the discussion in the literature has focused on English versus the Romance languages. The goals of this article are (a) to investigate empirically the role of the syllable in a language more closely related to English, namely Dutch (both belong to the family of Germanic languages) and (b) to determine whether Dutch listeners behave like English listeners, or like listeners in the Romance languages.

A caveat is necessary at this point. Contrary to a general trend in the literature (Mehler, Dommergues, Frauenfelder, & Segui, 1981; Segui, 1984; Segui, Dupoux, & Mehler, 1991), our research is not couched in terms of a quest for prelexical units of perception. Mehler and others interpret the syllable effects obtained in the Romance languages as evidence for the syllable as a fundamental unit of perception, with a prelexical level of stored syllables mediating between the speech input and the mental lexicon. Cutler and Norris (1988; Cutler, 1986, 1991; Norris & Cutler, 1985) have proposed an alternative to this unit of perception approach. They distinguished between segmentation, that is, the detection of relevant cues or boundaries in the speech

Pienie Zwitserlood, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; Herbert Schriefers, Free University, Berlin, Germany; Aditi Lahiri, University of Konstanz, Konstanz, Germany; Wilma van Donselaar, University of Utrecht, Utrecht, The Netherlands.

The results have been previously presented at the 31st Conference of Experimental Psychology, Bamberg, Germany, 1989. The research was supported by the Max Planck Institute for Psycholinguistics and by the Interfaculty Research Unit for Language and Speech, University of Nijmegen. We are grateful to Colin Brown, Uli Frauenfelder, Peter Hagoort, Allard Jongman, Antje Meyer, Thomas Pechmann, Dirk Vorberg, and three anonymous reviewers for their comments on earlier versions of this article. Wilma van Donselaar thanks Toni Rietveld for his activities as supervisor. We also thank the International School for Advanced Studies, Trieste, Italy, for providing the opportunity to discuss the topics in this article at the 1991 workshop on "The Psycholinguistic Consequences of Contrasting Language Phonologies."

Correspondence concerning this article should be addressed to Pienie Zwitserlood, Max Planck Institute for Psycholinguistics, Postbus 310, NL-6500 AH Nijmegen, The Netherlands.

signal, and classification, which is the actual identification of units existing at a prelexical level of representation. In their view, processing of incoming speech primarily involves the detection of points in the speech signal at which lexical access can be initiated. For this purpose, anything that is salient in a given language could be used: segments, feet, syllables, strong syllables only, or their full vowels. Segmentation does not imply classification. Yet another position states that there is no explicit segmentation involved in lexical access (cf. Bard, 1991; Marslen-Wilson, 1987; McClelland & Elman, 1986). In this view, elements of lexical form are directly activated by the information contained in the speech input, and differences in the levels of activation of such elements as well as the competition between them result in the surfacing of one strong candidate for recognition. What appears to be segmentation is in fact the side effect of lexical competition.

We do not a priori confine ourselves to any particular position; rather we leave this issue open and come back to it in the concluding section. At this point, we wish to investigate the sensitivity of Dutch listeners to the syllabic structure of their language during speech processing, contrasting performance in Dutch with results obtained in other languages. This article is organized as follows: We begin with a brief review of the experimental literature and present some facts about Dutch syllables. Three experiments in Dutch are then reported. The first is a replication of the English study by Cutler, Mehler, Norris, and Segui (1986). This is followed by two related experiments that make use of a greater variety of syllable types. The second experiment includes a replication in Dutch of the French study by Mehler et al. (1981) as well as an extension of our first experiment with more control conditions. The third experiment investigates the interplay between syllabic and morphological information in syllable monitoring.

Earlier Research on the Syllable

As we mentioned before, research in Romance languages showed a sensitivity of the listeners to the syllables of their language. In a study by Mehler et al. (1981), French subjects monitored for the presence of visually presented consonant-vowel-consonant (CVC) or consonant-vowel (CV) sequences in spoken carrier words such as *pal·mier* (*palm tree*; a centered dot is used throughout to indicate a syllable boundary) and *pa·lace* (*palace*). The subjects' reaction times in detecting the targets in the spoken words, measured from the acoustic onset, were faster when the target corresponded to the first syllable of the spoken carrier word (e.g., *PA* in *pa·lace* and *PAL* in *pal·mier*) than when it did not (e.g., *PA* in *pal·mier* and *PAL* in *pa·lace*). This resulted in an interaction between type of target (CV vs. CVC) and the first syllable of the spoken carrier words. These results have since been corroborated by data from other Romance languages such as Spanish (Sanchez-Casas, 1988; Sebastian-Galles, Dupoux, Segui, & Mehler, 1992), Portuguese (Morais, Content, Cary, Mehler, & Segui, 1989), and to some extent, Catalan (Sebastian-Galles et al., 1992). The commonly advanced interpretation of such results is that the syllable serves as a prelexical unit in speech perception (Segui et al., 1991).

The picture became more complicated with a set of experiments reported by Cutler, Mehler, Norris, and Segui (1983, 1986, 1989). The pattern found for the Romance languages was not obtained for English. Half of the English materials had clear syllable boundaries, comparable to the French stimuli (e.g., *bal·cony*), but the other half had intervocalic consonants after short stressed vowels (e.g., *balance*). For English, there is agreement to treat these intervocalic consonants as ambisyllabic, that is, as being part of the first as well as of the second syllable (Anderson & Jones, 1974; Gussenhoven, 1986; Kahn, 1980; for empirical support, see Stemberger, 1983; Treiman, 1989; Treiman & Danis, 1988). Cutler et al. (1986) obtained no effects of a match between the targets (*BA* and *BAL*) and the first syllable of either type of spoken carrier word. Moreover, English subjects did not show any syllable effects when listening to French materials. When French subjects listened to English material, an effect of syllabic match between target and spoken word was obtained for the clear cases (e.g., *bal·cony*).

Cutler et al. (1986) interpreted these results in terms of language-specific processing strategies. Because syllable boundaries are clear in French, French speakers use a syllabic processing strategy, whereas speakers of English, a language with unclear syllable boundaries, do not use this strategy. Originally, Cutler et al. proposed a phoneme-based processing routine for English. In more recent research, Cutler and her colleagues have focused on stress and have argued that English listeners access the lexicon using the full vowels of stressed syllables (Cutler, 1991; Cutler & Norris, 1988). Under this view of a metrical segmentation strategy the concept of the syllable does not necessarily play a role. The implicit assumption remains that English listeners use this routine because the preponderance of ambisyllabicity renders syllabic segmentation inadequate.

Syllables in Dutch

To what extent can the lack of syllable effects in English be generalized to other related languages with unclear syllable boundaries? It has been suggested that listeners will not use a syllabic segmentation strategy (Cutler et al., 1986). To test this hypothesis we chose to study Dutch, which is a language with widespread ambisyllabicity and frequent instances of clear syllable boundaries;[1] examples are given in Table 1. For our present purposes, the most important Dutch syllable structure rule is that Dutch allows for syllables ending in a long vowel (Table 1: *kaa·de*), in a consonant or consonant cluster (Table 1: *maag, buk·sen,* or *bukt*), or in schwa when unstressed (the second syllable in *kaa·de*). But

[1] We queried the CELEX (Center for Lexical Information, Nijmegen) database for the first-syllable structure of disyllabic Dutch words, excluding compounds and prefixed words. Forty-five percent had a clear syllable boundary between the consonants of a medial cluster, 34% had open first syllables ending in long vowels, and no less than 21% were ambisyllabic.

Table 1
Examples of Dutch Syllable Types

Unambiguous syllable structure				Ambisyllabic
s s	s	s s	s	s s
C V V C V	C V V C	C V C C V C	C V C C	C V C V C
k a a d e	m a a g	b u k s e n	b u k t	b u [k] e n
(quay)	(stomach)	(rifles)	(stoops)	(stoop)

Note. Square brackets denote ambisyllabic consonants. C = consonant. V = vowel.

Dutch has no words, and no syllables, ending in short full vowels; syllables containing a short full vowel must be closed by a consonant (van der Hulst, 1984; see also De Haas, 1986; Trommelen, 1983). If this consonant is followed by another vowel, the principle of maximal onsets (Kahn, 1980; Selkirk, 1982) applies, ensuring that, wherever possible, consonants are assigned to the onset of a syllable. Thus, the consonant that closes the first syllable simultaneously forms the onset of the second syllable (Lahiri & Koreman, 1988; van der Hulst, 1984), resulting in ambisyllabicity, as in *bu[k]en* (see Table 1).[2]

Thus, Dutch provides an interesting testing ground in comparison with English and French. First, all three languages have words with clear syllable boundaries. Second, both Dutch and English have ambisyllabic consonants that are part of two syllables (cf. English *ca[n]on* and Dutch *ka[p]er* [*hairdresser*]). Third, Dutch, like English and unlike French, has varying stress assignment and vowel reduction (e.g., compare the first two vowels in *senile* and *senility*; the second two in *juweel, jewel,* and *juwelier, jeweller*).

English and Dutch, however, differ in the nature of ambisyllabicity. Ambisyllabicity differs from language to language, and language-specific phonological rules determine which consonants can become ambisyllabic. In English, stress largely determines when consonants are to be part of both syllables (Gussenhoven, 1986; Myers, 1987). Initial syllabification is insensitive to stress and follows general principles such as onset maximalization, and ambisyllabicity occurs crucially after stress assignment (Gussenhoven, 1986; Kahn, 1980). Indeed, only onset consonants followed by unstressed vowels are attached to preceding sonorant-final syllables, thus becoming ambisyllabic. Many phonological processes as aspiration and glottalization are dependent on whether a consonant is ambisyllabic. In Dutch, however, assigning the consonantal segment as part of two syllables depends on the quantity of the preceding vowel and not on stress. If the vowel is short, a single intervocalic consonantal segment closes the preceding syllable, thereby becoming part of two syllables and—by our definition—ambisyllabic. This syllable structure is crucial to the subsequent assignment of stress, which then treats such /cvc/ syllables as heavy (Kager, 1989; Lahiri & Koreman, 1988; van der Hulst, 1984).

It is entirely possible that clear syllable boundaries (e.g., *kaa·de, buk·sen*) are marked in some way in the speech signal and that listeners use these low-level cues in the syllable detection task or, more generally, for segmentation purposes (Zwitserlood, 1991). Ambisyllabic consonants in Dutch, however, do not provide any obvious cues that could aid segmentation. In English, cues are available: Ambisyllabic consonants, in contrast with those that are syllable-final only, can become aspirated in British English, and they may flap in American English (Gussenhoven, 1986; Kahn, 1980). Dutch does not have these features, and there are no other obvious qualitative cues such as a change in consonant duration or in voice onset time, as is the case with geminate consonants in Italian, Finnish, and Bengali. Comparing minimal pairs such as *ro[k]en* and *roo·ken,* Jongman and Sereno (1992) found that ambisyllabic consonants do not differ in duration from other intervocalic Dutch consonants. Thus, the fact that a single consonant follows a short vowel determines its ambisyllabicity, and the listener needs this abstract knowledge to correctly identify a syllable including an ambisyllabic consonant.

So, both English and Dutch have ambisyllabic cases, but the languages differ with respect to the availability of low-level cues and in the way ambisyllabicity is defined at a more abstract phonological level. Contrasting the two languages empirically may reveal whether these distinctions are crucial. To test our main question concerning the sensitivity of Dutch listeners to the syllabic structures of their language, we used all of the syllable types illustrated in Table 1. In a series of monitoring experiments, we used the same procedure and task as in the English and French studies. In a go–no-go task, listeners saw visually presented targets (e.g., *BUK, MAA*). The string of speech segments specified by such targets did or did not correspond to the first syllable of a subsequently presented spoken carrier word (e.g., *BUK* in *buk·sen* [*rifles*] and *MOL* in *mor·gen* [*morning*]). Subjects had to detect a match between the speech segments specified by the target and the first part of the spoken word.

In Experiment 1, we replicated the design of the original English study by Cutler et al. (1986) by using clear and ambisyllabic spoken Dutch carrier words. In Experiment 2A, we tested two different material sets. In the first material set, as in the French experiment by Mehler et al. (1981),

[2] Ambisyllabicity can have different theoretical interpretations in phonology. What is important here is that the intervocalic medial consonant is a single segment that is in the coda of the first syllable and in the onset of the second. This will be our definition of ambisyllabicity throughout the article.

we used spoken words with clear syllable boundaries, and we tested for interactions between the type of target and the structure of the first syllable of the carrier words. In the second set, unlike in the study by Cutler et al. (1986), we contrasted ambisyllabic words with spoken control words whose syllabic structure differs from the ambisyllabic words. In Experiment 2B we investigated syllabic matches and mismatches between one type of target and four different spoken words, covarying morphological overlap between the target and spoken words.

Experiment 1

Method

Subjects. Seventy-four students, aged 18–26 years, participated in the experiment; all were native speakers of Dutch. They were paid for their participation.

Materials. Twenty-four pairs of spoken words were used as critical spoken carrier words. All spoken words were disyllabic and had main stress on the first syllable. The two words of each pair had the same stem; one word contained an ambisyllabic consonant (e.g., *stre[m]ing, the curdling of milk*), and the other had a clear syllable boundary after the first consonant of a medial cluster (*strem·sel, rennet*). Half of the ambisyllabic spoken words were verbs in their infinitive form (*bla[f]en, to bark*), the other half were derived nouns (e.g., *stre[m]ing*), consisting of a stem and a derivational affix starting with a vowel (e.g., *-ing* and *-er*). The corresponding carrier words with clear boundaries were either regular third person past tense forms (*blaf·te, barked*) or derived words, created by adding an affix starting with a consonant (e.g., *-sel, -baar,* and *-ster*) to the stem shared with the ambisyllabic word (e.g., *strem·sel*). The resulting words were 5 nouns and 7 adjectives. The syllable boundary in the clear words was always located between the stem and the inflectional or derivational affix starting with a consonant.

Design and procedure. Each spoken word was presented in combination with two visually presented targets (e.g., *BLA* and *BLAF* with *bla[f]en* and *blaf·te*). We refer to these as CV and CVC, the first C being a single consonant or a cluster, the V being a short full vowel, and the second C a single consonant

The 2 × 2 combination of targets and spoken words replicates the design and materials used in the English study (Cutler et al., 1986). The only difference between the English and Dutch materials is that we used morphologically complex words, whereas the English stimuli were monomorphemic. In Dutch, monomorphemic polysyllabic words are rare, and it turned out to be impossible to find enough material corresponding to the English words. But as with words such as *ba[l]ance,* the Dutch ambisyllabic carrier words had an ambisyllabic intervocalic consonant following a full short vowel. The clear boundary words were like the English *bal·cony*: They had intervocalic consonant clusters, with the syllable boundary between the two consonants. The crossing of carrier words with targets yielded a total of four target–word combinations and thus four experimental conditions. Conditions were rotated across experimental versions, such that subjects were presented with only one out of four possible target–carrier word combinations for each word pair, resulting in 24 critical trials per subject.

A trial consisted of the visual presentation of a target string (of the CV or CVC type), followed by a short list of spoken words, including the critical carrier word. The spoken filler words, varying in number from one to four, preceded the critical carriers to ensure that the position of critical words in the list was not predictable. The carrier words of a particular pair were preceded by identical fillers in each version.

Two additional material sets of 12 trials each were constructed, which were identical in all versions. These were the no-go trials, on which subjects were expected not to give a response. In the first set, none of the words in a list contained the target that was visually presented. The second set consisted of catch trials, in which the onset of the last word in the list partially overlapped with the specified target. The overlap between the targets and the onset of the spoken words was minimally two segments (e.g., *MES* with *meppen* [*to hit*] and *ROL* with *rotsig* [*rocky*]). These catch trials were included to ensure that subjects responded only when there was complete overlap between the target and the onset of a spoken word. An additional 18 lists were used as practice, with equal numbers of go and no-go trials. In the total set of spoken words—including critical, filler, catch, and practice trials—form class and first-syllable types were distributed equally.

All spoken materials were recorded, in random order, by a female native speaker of Dutch. The words were digitized at a sampling frequency of 20 kHz, and timing pulses were set at the onsets of the words in list-final position. This was done under auditory and visual control. Separate audiotapes were created for each of the versions. The order of word lists was constant across tapes; only the critical carrier words varied. The timing sequence on each tape was as follows: Each trial started with a 200-ms warning tone, after which an inaudible 5-kHz pulse on the second channel of the tape triggered the visual presentation of the target. After 3.5 s, the first word of the auditory list was presented, followed by the next word after 2 s, and so on, until the list was complete. For critical, filler, and catch trials alike, an inaudible 1-kHz timing pulse at the onset of the last word of the list triggered the counter modules that registered the subjects' push-button responses. Time-out was set to 1,500 ms, at which point the visual target disappeared from the screen. The next trial started after 3 s, again with a warning tone.

Subjects were tested in groups of 2 to 4. Each subject was seated in a separate carrel, in front of a cathode ray tube (CRT) screen on which the targets were displayed in capital letters. The subjects did not have to memorize the targets because the targets remained on the screen until time-out. The spoken words were presented binaurally over closed-ear headphones. Subjects were instructed to monitor for the occurrence in the spoken words of the sound sequences specified by the visual targets. The task was of the go–no-go type: Subjects were to respond only when the sound sequence specified by the target was detected in one of the spoken words. A PDP 11/23 computer controlled the presentation of the targets and registered response latencies, measured from the pulses placed at word onset. Subjects indicated the detection of a target by pressing a single button on the response box in front of them. Each test session began with the same 18 practice trials and lasted approximately 20 min.

Results and Discussion

Two subjects were excluded from the analyses because of their high percentage of errors (>20%) on the catch trials. The analyses are based on the remaining 72 subjects. There were 0.8% cases in which subjects had not responded before timeout (1,500 ms). These missing data were equally distributed over conditions and were replaced by the subject mean over the remaining items in the relevant condition. A further 1.4% extreme values, defined as reaction times that

Table 2
Experiment 1: Monitoring Latencies for Two Targets in Two Different Carrier Words

	Carrier words	
Target type	Ambisyllabic	Clear boundary
CV	559	569
CVC	521	528

Note. C = consonant. V = vowel.

were more than two standard deviations outside the means for both subjects and items, were replaced using the procedure recommended by Winer (1971, p. 488).

Monitoring latencies to detect CV and CVC targets in the 24 pairs of carrier words were submitted to analyses of variance (ANOVAs) on subjects and items, with type of target (CV vs. CVC) and carrier word (ambisyllabic vs. clear boundary) as factors. The results are shown in Table 2. Both the subject and the item analysis showed a significant main effect of target type: $F_1(1, 71) = 31.63$, $p < .0001$, $MS_e = 3{,}480$; $F_2(1, 23) = 9.06$, $p < .007$, $MS_e = 4{,}047$. There was no main effect of carrier word, $F_1(1, 71) = 1.72$, $p = .19$, $MS_e = 3{,}179$; $F_2 < 1$, and the interaction was also not significant (both Fs < 1).[3]

There was an overall 39-ms advantage of CVC over CV targets, and this effect was independent of the type of spoken word. In the spoken words with clear boundaries, the syllable boundary is located after the first consonant of the medial cluster, and the CVC target corresponds to the first syllable of these words. The faster detection latencies for these targets showed that the syllabic match between target and spoken word affects the subjects' responses. But the same pattern of results was found for ambisyllabic words. If the CVC advantage is interpreted as an effect of syllabic match, we must conclude that the first syllable in ambisyllabic words includes the intervocalic consonant.

Our results clearly deviate from what was found for English in the original Cutler et al. (1986) study. With comparable materials and conditions, their only effect was that ambisyllabic words (*ba[l]ance*) elicited faster responses than words with clear syllable boundaries (*bal·cony*). Such an effect is not interpretable in terms of syllabic matches between targets and spoken words.

Therefore, we wondered whether we could take the faster responses to CVC targets in Dutch as unequivocal evidence for syllable effects and whether we could draw any conclusions regarding differences between English and Dutch. The problem we wish to address is that the syllable effect manifests itself as an advantage for one type of target. In the French study (Mehler et al., 1981) each target type, be it CV or CVC, acts as its own control in the situation in which it does not correspond to the first syllable of a spoken word. Monitoring times to *BA* in *bal·con* can be used as a baseline against which performance on *BA* in *ba·lance* can be compared, and the same applies for the CVC target. Of course, spurious properties of specific targets or spoken words can have an impact on reaction time, but the interaction between targets and spoken words can only be interpreted as an effect of syllabic match. But a main effect of target type such as we obtained could be caused by variables other than the manipulated syllabic factor.

For our Dutch stimuli, we can think of two confounding factors. First, unlike the English materials, we used morphologically complex spoken words. The CVC targets in our experiment not only corresponded to the first syllable of both words of a pair but also corresponded to their shared stem morpheme. Syllabic and morphological match were completely confounded, and obtaining a CVC advantage might have originated from either source. A second worry was that, for Dutch listeners, CV strings might be difficult targets because there are no syllables in the language ending in short full vowels.

We therefore conducted two additional experiments in which we either controlled or covaried these confounding factors. In Experiment 2A, we used two separate material sets. With the first material set, we tested syllabic influences on monitoring performance in the same way as was done in the original French study (Mehler et al., 1981). Keeping morphological match constant, we used pairs of spoken words with clear syllabic boundaries and targets specifying legal syllables. Dutch has open syllables ending in a long vowel (e.g., *maa·gen, stomachs*), as well as syllables with long vowels closed by a consonant (*maag, stomach*). With such materials we can cross targets and spoken words and predict an interaction between these factors if syllabic information is used in speech monitoring.

Although with spoken words that have clear syllable boundaries we can establish whether syllables are important in Dutch, for the comparison with English we have an interest in the ambisyllabic cases. We therefore included a second material set with pairs of ambisyllabic and control words sharing the same stem. In the control condition the CVC target does not match the syllabic structure of the spoken word. An English example would be the target *TAL* combined with *talc,* in comparison with the ambisyllabic carrier word *ta[l]on.* In Experiment 1, we obtained an advantage for CVC over CV targets in ambisyllabic spoken words, and we argued that three factors could have been responsible for such an effect: the manipulated syllabic factor, the fact that only the CVC targets correspond to the stem morpheme of the words, and finally, that CV targets do not specify legal syllables of Dutch. Including a control condition enables us to compare reaction times to the same legal CVC target in words that share the same stem but not the same syllabic structure.

Finally, in Experiment 2B, we varied the morphological relation between targets and word stems instead of keeping

---

[3] We measured the durations of the fragments of our spoken carrier words that correspond to the targets (e.g., /bla/ and /blaf/ in *bla[f]en* and *blaf·te*; see also Experiment 2B). There were no significant durational differences; the mean duration of the /CV/ fragments was 182 ms for the ambisyllabic words and 185 ms for the clear boundary words; this was 277 and 286 ms for the /CVC/ fragments. An analysis of covariance on items, with duration as a covariate, showed the same pattern of results as the original analysis: $F_2(1, 22) = 4.06$, $p = .05$, for the main effect of target type; $F_2 < 1$ for the effect of carrier word and for the interaction.

this factor constant. We investigated monitoring performance to CVC type targets in four different carrier words. We thus avoided the problematic CV targets, and we covaried morphological and syllabic match by using two spoken words whose stem matched the target and two others whose stem mismatched the target.

Experiment 2A

Method

Subjects. Experiments 2A and B were run concurrently. A total 55 subjects, aged 18–26 years, were tested, 48 of whom were included in the analyses.[4] They were all native speakers of Dutch and were paid for their participation.

Materials. For the combined Experiments 2A and 2B, two sets of spoken materials were constructed, each containing 18 quadruples of carrier words, all with stressed first syllables. In the first set (long vowel), all carrier words had first syllables with long vowels, whereas the vowels in the first syllables of the words in the second set (short vowel) were short. In Experiment 2A two spoken words of each quaduple were used; 18 pairs in the long-vowel set and 18 in the short-vowel set.

The spoken words in the long-vowel set all had clear syllable boundaries. Within each pair, one word had an open first syllable with a long vowel (*maa·gen, stomachs*); these words were disyllabic plural nouns. The corresponding closed-syllable words were all monosyllabic nouns (*maag, stomach*). The words of each pair shared the same stem morpheme. Two different targets were combined with each spoken word. For the long-vowel set, the targets were of the CVV type, consisting of a consonant (or consonant cluster) and a long vowel or CVVC, with a consonant (or cluster), a long vowel, and a single consonant (e.g., targets *MAA* and *MAAG* with *maa·gen* and *maag*).

The word pairs of the short-vowel set consisted of an ambisyllabic and a control word with the same stem. The syllabic structure of the two words differed: The ambisyllabic words had the single ambisyllabic consonant in the coda of the syllable; the control words had two consonants in the coda. The ambisyllabic spoken words were verbs in their infinitive form (*bu[k]en, to stoop*) or plural nouns (*ma[n]en, men*); Dutch inflectional morphology does not distinguish between infinitival and plural forms. When the ambisyllabic carrier word was an infinitive verb, the corresponding control word was a monosyllabic inflected form of this verb (*bukt, stoops*). For the plural nouns, the controls were compounds that required insertion of genitive *s* between the initial CVC sequence and the second noun of the compound (e.g., *mans·volk, men folk*). This was the case for 3 of 18 sets. Care was taken that the medial consonant cluster (e.g., /sv/) was an illegal syllable onset in Dutch. The two targets in the short-vowel set were CV, containing a full short vowel, or CVC, with a single consonant after the short vowel (e.g., *BU* and *BUK* with *bu[k]en* and *bukt*).

The crossing of targets and spoken words in the long-vowel set replicates the design used in the Romance studies in that each target matches the first syllable of only one carrier word. The short-vowel set contains spoken words with ambisyllabic consonants in intervocalic position and monosyllabic control words. The potential match between targets and spoken words is different in this set. The CV target does not correspond to the first syllable of either carrier word. The only matching case is the CVC target in the ambisyllabic word, because CV as well as CVC targets mismatch the monosyllabic words that have an additional consonant in the coda.

Design and procedure. The combined Experiments 2A and 2B had a total of 12 conditions, with the variable of vowel set as a between-materials factor. Of these 12 conditions, 8 are analyzed in Experiment 2A and 8 in Experiment 2B, partially overlapping with Experiment 2A. The 6 within-materials variables yielded a total of six experimental conditions within each vowel set. Six experimental versions were constructed, with conditions rotated across versions, such that subjects were presented with only one out of six possible target–carrier word combinations for each set of spoken words, resulting in 36 test trials per subject.

A trial consisted of a visually presented target and a critical spoken carrier word, preceded by a number of spoken filler words. The number of fillers varied from one to six, such that the critical carrier words occurred in Positions 2 to 7. The carrier words of a particular item set were preceded by identical filler words in each version.

Three additional sets of 36 trials each were constructed, which were identical in all versions. With Set 1, the subjects had to provide a response, because the string of segments specified by the target matched the initial portion of the last spoken word of a trial. As with the critical trials, varying numbers (1 to 6) of fillers preceded the list-final word. These filler trials compensated for the unequal distribution of form class and word length in the test trials.

Sets 2 and 3, with the same list structure as the test and filler trials, provided the no-go trials. In Set 2, none of the words in a list contained the sound sequence specified by the target. Set 3 consisted of 36 catch trials, similar to those used in Experiment 1. In the total set of spoken words—including critical, filler, and catch trials—vowel length of the first syllable, form class, and first-syllable structure were distributed equally. All words had stressed first syllables. An additional 18 lists were used as practice, with equal numbers of go and no-go trials.

The manipulation of the spoken words, the production of the tapes, the timing sequence of the trials, the experimental procedure, and the task were identical to those of Experiment 1. Subjects were instructed to treat targets written with two vowel characters (e.g., *MAAG* or *MAA*) as specifying long vowels; this was made clear with some examples. Each test session started with the practice trials and lasted approximately 45 min. Subjects were given a short break after 25 min.

Results and Discussion

Seven subjects showed a high error rate on the catch trials (average 25%).[5] For the remaining 48 subjects, errors on the catch trials were minimal (3%). The analyses reported are on the data from these 48 subjects. There were 2.4% cases in which subjects had not responded before time-out (1,500

[4] Seven subjects were excluded from these analyses because of their high percentage of errors on the catch trials: 25% on average.

[5] The excluded subjects were apparently responding on the basis of some minimal overlap between target and spoken word, probably the first segment or segments only. Such a strategy results in errors on the catch trials, because in these trials the target and the spoken word only partially overlap (e.g., *BLON* with *blokken*). ANOVAs for syllabic factors for this group of subjects are quite revealing. Their data do not show any reliable syllabic effect (long-vowel set: $F < 1$; short-vowel set: $F_1(3, 18) = 1.12$, $p = .4$, $MS_e = 6{,}871$). These results emphasize the advantage, even the necessity, of catch trials to ensure that subjects perform the task properly.

ms). These missing data were equally distributed over conditions and were replaced by the subject mean over the remaining items in the relevant condition. Extreme values (2.4%), defined as reaction times that were more than two standard deviations outside the means for both subjects and items, were replaced using the procedure recommended by Winer (1971, p. 488). Because the relation between the targets and the first syllables of the carrier words was different in the two vowel sets, the data for each subset were analyzed separately.

For the materials with clear syllable boundaries (long-vowel set) an effect of syllabic structure should manifest itself as an interaction between targets and spoken words. The ANOVAs for the long-vowel set had two factors: target (CVV vs. CVVC) and spoken carrier word (open vs. closed syllable). The monitoring latencies for CVV and CVVC targets in the 18 pairs of carrier words were submitted to ANOVAs on subjects and on items. The results are shown in Table 3. No main effects of target or of carrier word were obtained (all $Fs < 1$). Both the subject and the item analysis revealed a significant interaction between the two factors: $F_1(1, 47) = 5.47, p < .03, MS_e = 4{,}210$; $F_2(1, 17) = 5.05, p < .04, MS_e = 1{,}712$. This outcome confirms our predictions for the materials with clear syllable boundaries, replicating the results obtained for Romance languages such as French and Spanish.

With respect to the short-vowel set, our main prediction is that mean monitoring times for CVC targets will be faster in the ambisyllabic words, in comparison with our baseline, which consists of the monosyllabic control words whose /*cvcc*/ structure does not match these targets. An additional comparison of interest is between detection latencies for CV and CVC targets in ambisyllabic words. Here, given our results from Experiment 1, we also expect a reaction time advantage for CVC targets.

Two one-factor ANOVAs, on subjects and on items, were carried out, with conditions as factor. A main effect was obtained in both: $F_1(3, 141) = 2.78\ p < .05\ MS_e = 5{,}283$; $F_2(3, 51) = 2.89, p < .05, MS_e = 1{,}906$. The results are shown in Table 4. To evaluate our predictions, we tested whether the mean latencies in the relevant conditions were statistically different. Comparisons with the Newman-Keuls procedure ($\alpha = .05$), with the error terms of both the subject and the item ANOVA, showed that detection times for CVC targets were significantly shorter in the ambisyllabic words than in the control words (34 ms). Also, in the ambisyllabic carrier words, the CVC targets were detected reliably faster (31 ms) than the CV targets. The two carrier words did not differ statistically when latencies to detect the CV target were compared (7 ms). There was no difference between latencies to detect the two targets in the control words (4 ms); it is important to keep in mind that both targets mismatched the syllable structure of these words.

Table 3
Experiment 2A: Monitoring Latencies in the Long-Vowel Set

	Carrier words	
Target type	Closed syllable	Open syllable
CVV	429	402
CVVC	412	428

Note. C = consonant. V = vowel.

Table 4
Experiment 2A: Monitoring Latencies in the Short-Vowel Set

	Word type	
Target type	Ambisyllabic	Control
CV	438	445
CVC	407	441

Note. C = consonant. V = vowel.

The data for both vowel sets provide clear evidence for syllabification effects in Dutch. When syllable boundaries are unambiguous, as is the case with the long-vowel words, an interaction is obtained between the factors target and spoken word. This effect is a sole function of the match between the target and the first syllable of the spoken carrier word. The pattern of results is strikingly similar to what was found for French and Spanish, with comparable materials (Mehler et al., 1981; Sanchez-Casas, 1988). Concerning the syllabic structure of ambisyllabic words, there is ample evidence for the match between the fragment specified by the CVC targets and the first syllable of ambisyllabic words. CVC targets are always detected faster in ambisyllabic words, whether compared with the CV targets in the same words or with CVC targets in the monosyllabic control words. The advantage for CVC over CV targets in ambisyllabic words from Experiment 1 is fully replicated in this experiment. However, to repudiate the hypothesis that this advantage is partially due to the morphological confound, the oddness of CV targets specifying nonexisting syllables in Dutch or both, the comparison of the same legal target, CVC, in the ambisyllabic and control words is crucial. The longer latencies for the mismatching control words (e.g., *BUK* in *bukt*) allow us to conclude that in Dutch, the first syllable of ambisyllabic words includes the intervocalic consonant.

Our worry that the results from Experiment 1 could have been due to the morphological match between the CVC target and the stem of the spoken words has been substantially reduced. As in Experiment 1, the word pairs in each set shared the same stem, but the effects we obtained allow us to separate the syllabic effects from morphological factors. First, the carrier words with long vowels show a syllable effect when the target mismatches the morphological make-up of the word (e.g., *MAA* in *maa·gen,* in which *maag* is the stem). Second, though the CVC target matches the stem of both ambisyllabic and control words (e.g., *BUK* in *bu[k]en* and *bukt*), reaction times are clearly shorter for ambisyllabic words. Although we can dismiss the idea that morphological factors are responsible for our effects, it is important to generalize effects of syllabic match to cases in which the target mismatches the morphological structure of the spoken word. Our only evidence for syllabic effects in the absence of a morphological match are the *MAA* in *maa·gen* cases. We therefore decided to investigate the contributions of morphological and syllabic match independently.

There is considerable debate in the literature as to whether morphological information, represented in the mental lexicon,[6] influences speech processing and, more specifically, syllable monitoring (see Mehler & Segui, 1987). Segui, Frauenfelder, and Mehler (1981) provided evidence against lexical involvement in syllable monitoring; results by Dommergues, Segui, and Mehler (cited in Segui, 1984) showed a somewhat different picture. In their monitoring study, targets such as *CRI* were detected faster in monosyllabic words (*cri, shout*) than in polysyllabic words (*critere, criterion*) or in pseudowords. Their interpretation is that lexical influences in syllable monitoring are restricted to monosyllabic spoken words. But unlike the polysyllabic cases, *CRI* and *cri* overlap at the lexical level of morphological representation: They share the same stem morpheme. In other words, monosyllabicity and morphological match appear to be confounded.

In Experiment 2B, we covaried syllabic and morphological overlap between targets and spoken words to investigate whether syllabic effects are obtained in situations in which the target mismatches the morphological structure of the spoken word. We can also assess the influence of morphological information on syllable monitoring performance. To each word of the pairs used in Experiment 2A, we added a carrier word with the same syllabic structure, but with a different stem, resulting in four words per set. These quadruples of carrier words were presented in combination with CVC (or CVVC) targets only, avoiding the problematic CV targets. Keeping morphological match (or mismatch) constant, we compare responses to, for example, the target *BUK* (*stoop*) in the carriers *bu[k]en* (*to stoop*) and *buk·sen* (*rifles*).

Experiment 2B

Method

Subjects. Experiment 2B was run concurrently with Experiment 2A. The same 55 subjects, aged 18–26 years, participated, 48 of whom were included in the analyses.

Material. A total of 36 quadruples of spoken words were used, 18 with short vowels, 18 with long vowels. All words had stressed first syllables. The quadruples consisted of the word pairs from Experiment 2A plus two additional spoken words. The words in each pair in Experiment 2A shared the same stem morpheme, and this morpheme corresponded to the CVC (or CVVC) target (e.g., *BUK* in *bu[k]en* [*stoop*] and *bukt* [*stoops*]). The two additional words also shared the same stem (e.g., *buk·sen* [*rifles*] and *buks* [*rifle*]), but this stem (*buks*) did not correspond to the CVC target. As with the spoken words from Experiment 2A, the target matched the first syllable of one of the new carrier words (e.g., *BUK* in *buk·sen*); with the other word (*buks*) this was not the case. The additional spoken words whose first syllable matched the target were disyllabic nouns (*maag·den* [*virgins*] and *buk·sen* [*rifles*]). The other new carriers were mainly monosyllabic, mismatching the target because of an extra consonant in the coda position of the syllable (*maagd* [*virgin*] and *buks* [*rifle*]). In four cases a compound was used. Each spoken word in a quadruple was combined with a CVC or CVVC target type. Table 5 illustrates the materials and design of Experiment 2B; for further details on design and procedure see Experiment 2A.

Table 5
Conditions in Experiment 2B for CVC/CVVC Targets

	Short vowel words			
CVC (BUK)	bu[k]en	bukt	buk·sen	buks
Morpheme	+	+	−	−
Syllable	+	−	+	−
	Long vowel words			
CVVC (MAAG)	maag	maa·gen	maag·den	maagd
Morpheme	+	+	−	−
Syllable	+	−	+	−

Note. A plus sign indicates a match, and a minus sign indicates a mismatch. C = consonant. V = vowel.

Results and Discussion

Missing data (2.6%) and extreme values (2.4%) were treated as in Experiment 2A. Monitoring latencies for CVC and CVVC targets in four carrier words were analyzed in two ANOVAs on subjects and on items. Variables in the ANOVAs were syllabic match, morphological match, and vowel type, with the two sets of 18 items (with short and long vowels) nested under the latter factor in the analysis on items. In both subject- and item-based analyses, there was a strong 35-ms main effect of the syllabic factor: $F_1(1, 47) = 24.27$, $MS_e = 4{,}964$, $p < .0001$; $F_2(1, 34) = 11.75$, $MS_e = 3{,}844$, $p < .002$. The effect of morphological match (14 ms) just failed significance. Neither the factor vowel type nor any of the interactions reached significance. The results are shown in Table 6. As in the earlier experiments, a clear 35-ms syllable effect was obtained. Although the syllable effect seems smaller when there is morphological overlap (26 ms vs. 46 ms), the interaction between the syllabic and the morphological factors was not significant. Because we were interested in the generalization of the syllable effect to situations in which there is no morphological match between stems and targets, it is important that the syllable effect was quite large in those conditions.

In these analyses of syllabic and morphological factors, we compared detection times for the same target in four different spoken carrier words. One might argue that the pattern of results could be confounded by differences in either the frequency of the spoken carrier words or in the durations of the fragment of speech in each spoken word that has to be monitored to detect the target. With respect to the frequency of the tokens used, we can be brief. The spoken words, whose first syllable matched the target, had

[6] Given the evidence in the domain of spoken-word recognition against the position originally held by Taft and Forster (1975) that there is a prelexical level of morphological processing, we take morphological information to be lexically represented (cf. Schriefers, Zwitserlood, & Roelofs, 1991; Tyler, Marslen-Wilson, Rentoul, & Hanney, 1988).

a frequency per million of 24; this was 23 for mismatching words (Celex database; corpus size is 42,003,800). The frequency of the words that shared their stem morpheme with the target was 21 per million; this was 25 for the mismatching words.

As to the second point, it is obvious that response times in speech monitoring are a function of how much of a spoken word has to be monitored to detect a match with the target. A string of segments, for example /buk/, can be acoustically realized in different ways, depending on—among other things—the length of the total utterance and speaking speed. Of particular interest here are differences in duration: If /buk/ in *buksen* is of a shorter duration than in *buks*, then the subject can detect the match with the target sooner when hearing *buksen*. Because reaction times are measured from word onset, this will have an effect on the monitoring latencies.

We therefore assessed whether the acoustic realization of the same string of segments was of the same duration in the spoken words used (e.g., /buk/ in *bu[k]en, bukt, buk·sen,* and *buks*). For each word of a quadruple, we measured the duration of the fragment (see Table 7) corresponding to the target with speech editing software on a Vax 750. Onsets and ends of the fragments were determined under visual and auditory control. These measurements included the first consonant and the (short or long) vowel, plus some critical amount of the following consonant.[7] The durations in Table 7 showed clear differences between the word fragments as large as 34 ms. In particular, the mean duration in the conditions in which the spoken words morphologically match the target was 281 ms; this was only 261 ms for the mismatching cases. Because reaction times were measured from word onset, subjects could have detected a match with the target earlier in the conditions of a mismatch between the target and the stem morpheme of the spoken words. Such a 20-ms difference could have diminished the impact of our morphological factor.[8]

We introduced the morphological variable to investigate whether syllable effects are obtained in situations in which the targets do not correspond to the stem of the spoken word, and we showed that the effects of the syllabic factor are as strong as when there is also morphological overlap. A second issue concerned lexical effects in speech monitoring, and given the trend in the latency data and the observed differences in word-fragment duration, we tentatively conclude that the monitoring task is not free from lexical influences.

Table 6
Experiment 2B: Mean Monitoring Latencies (in Milliseconds) for CVC/CVVC Targets for Short- and Long-Vowel Carrier Words

	+Syllable		−Syllable	
Target	+ Morpheme	− Morpheme	+ Morpheme	− Morpheme
Short vowel (CVC)	407	403	441	458
Long vowel (CVVC)	412	424	429	460

Note. A plus sign indicates a match, and a minus sign indicates a mismatch. C = consonant. V = vowel.

Table 7
Mean Durations (in Milliseconds) of Word Fragments Corresponding to CVC/CVVC Targets for Short- and Long-Vowel Words

Short-Vowel (CVC)	bu[k]en	buk·sen	bukt	tuks
	246	224	243	239
Long-Vowel (CVVC)	maag	maag·den	maa·gen	maagd
	326	282	308	299

Note. C = consonant. V = vowel.

Conclusion

In three monitoring experiments, we investigated syllable effects in Dutch, and the data from all three provide compelling evidence in favor of the listeners' sensitivity to the syllabic structure of spoken words. In Experiment 1, we replicated the design of the English experiment by Cutler et al. (1986). We obtained shorter monitoring latencies when the string of segments specified by the target matched the first syllable of the spoken word. This advantage was obtained for spoken words with clear syllable boundaries as well as for words with ambisyllabic consonants. In the second experiment, we corroborated the effects for ambisyllabic words by introducing a control condition, in which the target and spoken word mismatched syllabically. Latencies were consistently shorter for the matching ambisyllabic words. The data from Experiments 1 and 2A demonstrate that the intervocalic consonant forms the coda of the first syllable in ambisyllabic words in Dutch.

In Experiment 2A, we also replicated the design of the studies on Romance languages, by using materials with clear syllabic boundaries. As in the original study on French by Mehler et al. (1981), an interaction was obtained between targets and spoken words showing that monitoring times are a function of the syllabic match between the two. In Experiment 2B we generalized the syllable effect to situations in which there is a morphological mismatch between target

[7] Because the carrier words had either no segments or different segments following the second consonant (e.g., /t/ vs. /s/ in *bukt* and *buksen*), we decided not to include the full second consonant in our measurements. The critical amount of consonantal information was determined on the basis of available literature: for the stops (12) the measurement included 20 ms of plosive information after burst onset; for nasals (11) 50 ms from the moment that the high formants disappear; and for liquids (12) and fricatives (1), 70 ms of consonantal information (Ohde & Sharf, 1981; Warren & Marslen-Wilson, 1987).

[8] In an ANOVA on corrected mean item scores (mean reaction time − fragment duration), with the same factors as the ANOVA for the mean reaction times the effect of morphological match (34 ms) was highly significant, $F_2(1, 34) = 24.66$, $p < .0001$, $MS_e = 1{,}649$.

and spoken word. A strong effect of syllabic match was obtained with latencies to detect identical targets in four different spoken words. The matching cases included words with clear syllable boundaries (e.g., *maag·den* and *buk·sen*) as well as words with ambisyllabic consonants (e.g., *bu*[*k*]-*en*). We showed that although morphological overlap between target and spoken word seems to play a role in speech monitoring, positive effects of syllabic structure are independent of the match between the target and the stem morpheme of the spoken word.

So, Dutch listeners are clearly sensitive to the syllabic structure of spoken words, and this holds for the unambiguous as well as for the ambisyllabic cases. One point of departure for our study was the finding by Cutler et al. (1986) that speakers of English do not show any signs of a sensitivity to syllabic structure, whereas French speakers do. The argument put forward by Cutler et al. was that French has unambiguous syllable boundaries as opposed to English, which shows a greater diversity of syllable types and a preponderance of ambisyllabicity. How do these results relate to our own data on Dutch?

Although French and Dutch differ with respect to stress—French has fixed word-final stress, and Dutch has varying stress assignment—we demonstrated that Dutch listeners behave as the French do when confronted with spoken words with clear syllable boundaries. English and Dutch both have varying stress and words with clear syllable boundaries as well as with ambisyllabic consonants. Our first experiment, like the study by Cutler et al. (1986), included ambisyllabic and clear cases. Whereas Cutler et al. found no effects for English, we obtained a syllable effect with comparable materials and design. As argued earlier, ambisyllabic consonants in the two languages differ with respect to the availability of low-level cues for segmentation. English has some (aspiration or flapping); Dutch does not have such cues. Surprisingly, our Dutch listeners show clear syllable matching effects with ambisyllabic materials, whereas the English listeners do not. We also argued that ambisyllabicity in English and in Dutch differ with respect to the role of stress assignment and to the rules involved, and apparently such facts are important. Because there are no obvious low-level cues in Dutch, the diverging pattern of data for English and Dutch most probably results from differences between English and Dutch at a more abstract level of phonological structure. As such, we would agree with Cutler et al. (1986) that different processing routines are used for different languages, depending on their phonological structure and on the availability of low-level cues in the signal.

Of course, it is possible that clear syllable boundaries in Dutch are marked in the speech signal. In fact, we have evidence that this is so (Zwitserlood, 1991). But clearly, the fact that our subjects were successful in monitoring for the first syllable in ambisyllabic words shows that our syllable effects cannot be exclusively due to local cues available in the acoustic input. What could induce the kind of syllabification behavior that includes the intervocalic consonant into the first syllable of our ambisyllabic words? A successful strategy would be "Always include a consonant after a short vowel, unless the vowel is a schwa." This would result in correct syllabification for the ambisyllabic cases, but what about medial consonant clusters after short vowels? With consonant clusters, the strategy has to know about maximizing onsets (to correctly segment *mon·ster* [*monster*] and not *mons·ter*) and about permissable onsets and codas (to segment *pant·ser* [*armour*] and not *pan·tser*). Because the speech signal does not seem to contain consistent low-level segmentation cues for ambisyllabic consonants, such knowledge is better characterized as information about how speech segments may be combined into syllables.

We now turn to some consequences of our findings for the locus of syllabic information in speech perception. The question here is how our data relate to the different approaches summarized earlier: the prelexical storage of syllables, as proposed by Mehler and his colleagues (1981), the metrical segmentation approach (Cutler & Norris, 1988), and the view that segmentation is a side-effect of lexical processing (see Bard, 1991).

In our view, Dutch listeners combine information from acoustic cues with knowledge about the structural regularities of the syllables in their language during lexical access. A crucial question is from where this knowledge might derive. Is it instantiated as a matching process between the sensory input and the units of a prelexical syllabic level? Such a process would profit from a limited set of unambiguous syllables or syllable types and preferably operate on low-level acoustic cues. Because this situation clearly is not always obtained in Dutch and English, we feel that, although it might be appropriate for Romance languages, prelexical classification in terms of syllables cannot be the universal strategy.

A metrical segmentation routine, defined as the initiation of lexical access at the full vowel of stressed syllables, was suggested for English because English listeners showed no sensitivity to syllables, as such. Our Dutch stimuli always had stressed first syllables and, except for the few compounds, unstressed second syllables. The metrical strategy predicts lexical access to be initiated at the strong syllable and no segmentation to occur before a second weak syllable. Thus, no differences should be found between our spoken words, because lexical access would be attempted with the full vowel of all words, without further segmentation within words. In short, no syllable effects are predicted, but we do obtain them. Of course, a more general segmentation-without-classification approach, in which syllables are relevant, is compatible with our data.

The third option is that the syllable does not play a role before or during lexical access, because there is no segmentation or classification. If the incoming sensory input is continuously mapped onto lexical elements and if segmentation is an epiphenomenal side effect of lexical processing, then one possibility is to have syllabic information represented in the lexicon. It is possible that the information about a word's syllabic structure is stored at the level of lexical form representations and retrieved during the process of lexical access. Given the trend for lexical effects in the morphological conditions, a lexical locus of syllabic effects cannot be completely excluded. However, although representing syllabic information in the mental lexicon could account for our results with isolated words, it would not explain how connected speech is syllabified. Syllabification of connected speech goes across word boundaries; slowly spoken, a string such as *date it* would be */det·It/*, with a syllable boundary between the two words. In normal connected speech, however, the flapping of the second consonant in American English, or the aspiration in British English, tells us that resyllabification has occurred. The second consonant is now the onset of the second syllable: */de·dIt/* for the flapped version and */de·thIt/* for the aspirated version.

So, prelexical or lexical levels of representation might not be appropriate loci for syllabic information (see also Lahiri & Marslen-Wilson, 1991). From where then, might information about syllabic structures derive? One hypothesis about how information about syllables might be implemented is in terms of a parsing routine. In analogy to the syntactic parsing of sentences, the speech input could be mapped onto phonological representations of lexical form. Assuming that phonological representations specify what their segments are (Lahiri & Marslen-Wilson, 1991), these could be used as constituents, as building blocks in the construction or computation of syllables. Acoustic cues can aid this process, and knowledge about possible syllables of the language could be implemented in terms of constraints on computations. In such an approach, syllables are not represented units but rather derived or computed entities. A similar position is advanced by Frazier (1987). In her view, knowledge of permissible syllables aids the structuring of the acoustic input. Syllabification, then, does not exclusively depend on the availability of acoustic cues signaling syllable boundaries, but rather, it depends on the interplay between properties of the acoustic input and the listeners' knowledge about the phonological structure of their language. Such an approach is compatible with a continuous propagation of sensory information onto the representations in the mental lexicon.

Although at this point, more research is needed to establish where and how information about syllables might be implemented, we feel that approaches that emphasize the interplay between salient acoustic information and language-specific phonological information during lexical access are promising. They can explain our findings for Dutch, results for Romance languages, and the results for English. In this perspective, it is not surprising that segmentation behavior varies across languages, because languages vary with respect to their phonological structure and, as a consequence, with respect to the knowledge listeners may use in structuring the speech input.

References

Anderson, J., & Jones, C. (1974). Three theses concerning phonological representation. *Journal of Linguistics, 10,* 1–26.

Bard, E. G. (1991). Competition, lateral inhibition, and frequency: Comments on the chapters of Frauenfelder and Peeters, Marslen-Wilson, and others. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives.* Cambridge, MA: MIT Press.

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

Cutler, A. (1986). Phonological structure in speech recognition. *Phonology Yearbook, 3,* 161–178.

Cutler, A. (1991). Exploiting prosodic probabilities in speech segmentation. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives.* Cambridge, MA: MIT Press.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1983). A language specific comprehension strategy. *Nature, 304,* 159–160.

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

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1989). Limits on bilingualism. *Nature, 340,* 229–230.

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.

De Haas, W. (1986). Partial syllabification and schwa epenthesis in Dutch. *Gramma, 10,* 143–161.

Frazier, L. (1987). Structure in auditory word recognition. *Cognition, 25,* 157–187.

Gussenhoven, C. (1986). English plosive allophones and ambisyllabicity. *Gramma, 10,* 119–141.

Hooper, J. B. (1972). The syllable in phonological theory. *Language, 48,* 525–540.

Hulst, van der H. (1984). *Syllable structure and stress in Dutch.* Dordrecht, The Netherlands: Foris.

Jongman, A., & Sereno, J. A. (1992). Compensatory effects due to vowel length. *Proceedings of the 12th International Congress of Phonetics, 2,* 294–297.

Kager, R. (1989). *A metrical theory of stress and destressing in English and Dutch.* Dordrecht, The Netherlands: Foris.

Kahn, D. (1980). *Syllable-based generalisations in English phonology.* New York, NY: Garland.

Lahiri, A., & Koreman, J. (1988). Syllable weight and quantity in Dutch. In *Proceedings of the 7th West Coast Conference on Formal Linguistics.* Stanford, CA: Stanford University Linguistics Association.

Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition, 38,* 245–294.

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

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

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

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

Mehler, J., Dupoux, E., & Segui, J. (1991). Constraining models of lexical access: The onset of word recognition. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives.* Cambridge, MA: MIT Press.

Mehler, J., & Segui, J. (1987). English and French speech processing: Some psycholinguistic investigations. In M. E. H. Schouten (Ed.), *The psychophysics of speech perception.* Dordrecht, The Netherlands: Nijhoff.

Meyer, A. S. (1992). Investigation of phonological encoding through speech error analysis: Achievements, limitations, and alternatives. *Cognition, 42,* 181–211.

Morais, J., Content, A., Cary, L., Mehler, J., & Segui, J. (1989).

Syllabic segmentation and literacy. *Language and Cognitive Processes, 4,* 57–67.

Myers, S. (1987). Vowel shortening in English. *Natural Language and Linguistic Theory, 5,* 485–518.

Norris, D., & Cutler, A. (1985). Juncture detection. *Linguistics, 23,* 689–705.

Ohde, R. N., & Sharf, D. J. (1981). Stop identification from vocalic transition plus vowel segments of CV and VC syllables: A follow up study. *Journal of the Acoustical Society of America, 69,* 297–300.

Sanchez-Casas, R. M. (1988). *Access representations in visual word recognition.* Unpublished doctoral disseration, Monash University, Clayton, Victoria, Australia.

Schriefers, H., Zwitserlood, P., & Roelofs, A. (1991). The identification of morphologically complex spoken words; Continuous processing or decomposition? *Journal of Memory and Language, 30,* 26–47.

Sebastian-Galles, N., Dupoux, E., Segui, J., & Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language, 31,* 18–32.

Segui, J. (1984). The syllable: A basic perceptual unit in speech processing? In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Language control mechanisms.* Hillsdale, NJ: Erlbaum.

Segui, J., Dupoux, E., & Mehler, J. (1991). The role of the syllable in speech segmentation, phoneme identification, and lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational prespectives.* Cambridge, MA: MIT Press.

Segui, J., Frauenfelder, U., & Mehler, J. (1981). Phoneme monitoring, syllable monitoring and lexical access. *British Journal of Psychology, 72,* 471–477.

Selkirk, E. O. (1982). The syllable. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representation* (Part 2). Dordrecht, The Netherlands: Foris.

Selkirk, E. O. (1984). On the major class features and syllable theory. In M. Aronoff & R. T. Oehrle (Eds.), *Language sound structure: Studies in phonology presented to Morris Halle by his teacher and students.* Cambridge, MA: MIT Press.

Stemberger, J. P. (1983). The nature of /r/ and /l/ in English: Evidence from speech errors. *Journal of Phonetics, 11,* 139–147.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14,* 638–647.

Treiman, R. (1989). The internal structure of the syllable. In G. N. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing.* Dordrecht, The Netherlands: Kluwer.

Treiman, R., & Danis, C. (1988). Syllabification of intervocalic consonants. *Journal of Memory and Language, 27,* 87–104.

Trommelen, M. (1983). *The syllable in Dutch: With special reference to diminutive formation.* Dordrecht, The Netherlands: Foris.

Tyler, L. K., Marslen-Wilson, W., Rentoul, J., & Hanney, P. (1988). Continuous and discontinuous access in spoken word-recognition: The role of derivational prefixes. *Journal of Memory and Language, 27,* 368–381.

Warren, P., & Marslen-Wilson, W. (1987). Continuous uptake of acoustic cues in spoken word recognition. *Perception and Psychophysics, 41,* 262–275.

Winer, B. J. (1971). *Statistical principles in experimental design.* New York: McGraw-Hill.

Zwitserlood, P. (1991). *Monitoring for syllables in morphologically complex words.* Paper presented at the 32nd Annual Meeting of the Psychonomic Society, San Francisco, CA.

Received October 3, 1991
Revision received April 17, 1992
Accepted May 26, 1992 ■