Chapter 12

The Role of the Syllable in Speech Segmentation, Phoneme Identification, and Lexical Access

Juan Segui, Emmanuel Dupoux, and Jacques Mehler

In previous papers we proposed that the syllable is a natural unit in speech processing (Mehler, Segui, and Frauenfelder 1981; Mehler 1981; Segui 1984). According to this position, syllabic units are constructed from the continuous speech signal and constitute the input for lexical access and eventually for more detailed phonemic analysis. This hypothesis is independent of the phonological properties of different natural languages. In the three sections of this chapter we review some recent experimental findings concerning the role of the syllable in speech segmentation, phoneme identification, and lexical access.

The Role of the Syllable in Speech Segmentation

In an initial study Mehler, Dommergues, Frauenfelder, and Segui (1981) investigated the role of the syllable as a unit in speech segmentation. These authors demonstrated that subjects responded faster to a phonemic target sequence when it corresponded to the initial syllable of the target-bearing word than when it did not. For instance, the target *pa* was detected faster in the word *pa-lace* than in the word *pal-mier*, whereas the target *pal* was detected faster in *pal-mier* than in *pa-lace*. In this experiment, target-bearing items began with a stop consonant (/p/, /t/, /k/, /b/, or /g/) followed in all cases by the vowel /a/. The interaction between target type (CV or CVC) and word type (CV or CVC) was highly significant (C = consonant; V = vowel). According to Mehler, Dommergues, et al. (1981), these results indicate that the syllable constitutes a basic intermediary unit in speech perception.

These experimental findings were recently replicated in Spanish by Sanchez-Casas (1988) and Bradley, Sanchez-Casas, and Garcia-Albea (1988), who used a larger number of experimental words beginning with a step consonant, a fricative, a liquid, or a nasal followed by any single vowel permitted in Spanish. As in French, when the target syllabic structure (CV or CVC) and the syllabic structure of the experimental word (CV or CVC) matched, response times (RTs) were shorter than in the nonmatching case. The syllabic effect was larger when the target sequence was specified auditorily rather than visually.

Some caution is necessary concerning the interpretation of these results, since the global RTs obtained by Sanchez-Casas and Bradley et al. were much greater than those observed in French (600 msec versus 360 msec, respectively). Differences in RTs may result from the use of alternative processing strategies, namely prelexical versus lexical strategies. This is important for interpreting the obtained results. In particular, if detection responses are derived from a lexical code, they can reflect the syllabic format of this code and not necessarily the syllabic nature of a prelexical segmentation strategy.

Finally, Morais et al. (1989) observed a syllabic effect in Portuguese with illiterate and ex-illiterate subjects using a modified version of the experimental procedure of Mehler, Dommergues, et al. (1981). In this experiment the dependent variable was the number of correct detections rather than RTs, and the targets were, of course, specified auditorily. A significant interaction was found between target type and word type: detections were more numerous when the target coincided with the first syllable of the experimental word than when it did not. The syllabic effect was similar for the two populations even if ex-illiterate subjects performed better than illiterate ones. This result suggests that Portuguese subjects tend to represent spoken words as a sequence of syllabic units. Furthermore, the comparability of the syllabic effect on the two populations indicates that this form of representation depends essentially on informal experience with the language rather than on formal instruction.

The syllabic effect observed in the three previously examined experiments was obtained with romance languages having clear syllabic boundaries and rather restricted syllabic structures. But what about other families of languages, and in particular, languages having unclear syllabic boundaries? The results obtained on this point are ambiguous.

To test the generality of the syllabic hypothesis, Cutler et al. (1983, 1986) conducted their first English experiment. Contrary to French, many English words do not have a clear syllabic structure. For instance, the syllabic boundary for the word *balance* falls neither clearly before nor after the phoneme /l/. When English subjects are asked to segment this type of word, they do not agree about the location of the syllabic boundary

(Treiman and Danis 1988). To represent the syllabic structure of these words in which an intervocalic consonant precedes an unstressed vowel, linguists have proposed that, in fact, the intervocalic consonant belongs to both the first and the second syllable. In other words, this phonetic segment is ambisyllabic.

In the 1986 experiment by Cutler et al. subjects received words with clear initial syllabic structure like *bal-cony* (CVC words) and words with ambisyllabic second consonants like *ba(l)ance* (CV(C) words). As in the original French experiment, the target sequences used in this English experiment were CV and CVC. No trace of syllabifying segmentation found in this experiment for CVC and CV(C) words. Monitoring latencies for CV and CVC targets were practically identical in both types of words.

The absence of a target effect runs counter to the syllabic hypothesis. In particular, if CVC targets corresponded to the initial syllable of CVC words and at least in some cases to that of CV(C) words, a global CVC target advantage should be found. This was not the case, however. Nevertheless, a word-type effect was observed, CV(C) words being responded to faster than CVC words. The faster RTs for CV(C) words than for CVC words was attributed to the fact that alternating consonant-vowel patterns like CVC (*balance*) lend more readily to segmentation than nonalternating strings like CVCC (*balcony*). Thus the faster response to CVC words over CV words may be interpreted by pointing to the facility with which subsyllabic units were identified.

To interpret the total absence of a syllabic effect in English, Cutler et al. proposed that English subjects used a phonological, rather than a syllabic, segmentation device. The output of this device, which may be used to process languages with irregular syllabic structure, is a phonological representation of speech without any intermediate representation. The hypothesis of a phonological segmentation device may explain the differences observed in detecting target sequences as a function of the initial structural properties of the carrier word. Some sequences of phonemes (alternating consonant-vowel sequences) seem to be easier to deal with than others (nonalternating sequences). In any case, it is clear at this point that a purely phonetic segmentation device does not explain the absence of a difference between CV and CVC targets. Indeed, if words are computed phoneme by phoneme, one should predict shorter RTs for CV than for CVC targets.

More recently Sanchez-Casas (1988) and Bradley et al. (1988) confirmed the absence of an interaction between word type and target type for English. However, contrary to the 1986 results of Cutler et al., these authors observed that CVC targets are generally responded to faster than CV targets. This result was observed independently of the structure of the experimental words, CVC or CV(C). As noted above, these results may be considered to reflect the use of a syllabic strategy if we assume that the initial syllable of CV(C) words may be, at least in some cases, the initial CVC sequence. Nevertheless, Sanchez-Casas and Bradley et al. assumed that the advantage of CVC targets over CV targets is best explained in terms of the ease with which the mental representation of the target can be developed and maintained rather than as reflecting the use of a syllabic-segmentation routine. Independently of the different theoretical interpretations, it is clear that there is an empirical discrepancy between the results of Cutler et al. and those obtained by Sanchez-Casas and Bradley et al.

The hypothesis of Cutler et al., according to which French and English listeners employed two qualitatively different speech-perception routines, can be further evaluated on the basis of some recent results that Zwitserlood et al. (in preparation) observed in Dutch. Like English, Dutch is a language with widespread ambisyllabicity. In particular, in Dutch a syllable with a short vowel is *necessarily closed by a consonant*. Furthermore, if this consonant is followed by another vowel, this consonant is the onset of the second syllable, which results in ambisyllabicity (e.g., *bo*(*k*)*en*). However, if the second consonant is followed by another consonant, the second consonant belongs unequivocally to the first syllable (e.g., *bok-ser*). On the other hand, words having a long vowel following the initial consonant have clear syllable boundaries of either the CVV type (e.g., *maa-gen*) or the CVVC type (*maag-den*).

In the experiment by Zwitserlood et al. subjects monitored CV, CVV, CVC, and CVVC words with clear and unclear syllabic boundaries for the presence of target sequences consisting of two or three phonemes. The results obtained for words having a clear syllabic structure and a long vowel reproduce the interaction obtained in French between target type and word type (CVV targets are detected faster in CVV-C words than in CVVC words, whereas CVVC targets are detected faster in CVVC words than in CVV-C words). These results corroborate those previously observed in French.

The main finding in this experiment was the presence of a clear target-type effect for syllabically clear CVC words and for ambisyllabic CV(C) words. Reaction times for these two types of words were significantly shorter for CVC targets than for CV targets. On the other hand, monitoring times for CVC targets were identical in clear syllabic words like *win-de* and in ambisyllabic words like *wi(n)en*. According to the authors, these results indicate that Dutch listeners used a syllabification strategy for clear and ambisyllabic words, since they were just as successful in segmenting both word types.

The experiments examined in this first section confirm the existence of a syllabic effect in syllable monitoring in languages having clear syllabic boundaries like French, Spanish, and Portuguese. However, as noted above, only the results obtained in French and eventually to be obtained in Spanish may be considered to reflect the use of an on-line syllabic segmentation strategy. Results obtained in experiments conducted with languages having less clear syllabic boundaries are open to alternative interpretations. In particular, although data obtained in Dutch are compatible with a reformulation of the syllabic hypothesis according to which phonological "knowledge" about the syllabic regularities of the language may be exploited on-line to segment the speech wave (even if no specific syllabic cues are present in the signal), results observed in English do not appear interpretable in the framework of this hypothesis and suggest on the contrary the use of a nonsyllabic segmentation strategy. Independent of the final issue concerning the nature of the segmentation strategy(ies), the experiments examined above indicate the necessity of taking into consideration the precise phonological structure of the languages employed in order to propose a realistic interpretation of the empirical data.

The Role of the Syllable in Phoneme Identification

The results obtained in previous experiments on phoneme and syllable monitoring suggest that phoneme detection is highly related to the internal structure of the syllable. In particular, Segui, Frauenfelder, and Mehler (1981) observed a strong correlation between phoneme and syllable RTs for words and nonwords. According to these authors, the syllable can be seen as the structural unit from which subsyllabic phonemic analysis originates. In other words, syllables are used to compute their corresponding underlying phonetic structures. In this section we revise some recent results that agree with this aspect of the syllabic hypothesis.

In a series of experiments conducted with French material, Cutler, Mehler, et al. (1987) report a correlation between the RT to detect the initial phonemes of a syllable and the structural complexity of this syllable. This structural complexity can be expressed in terms of the presence or absence of an initial consonant cluster and the presence or absence of a final consonant.

In their first experiment, conducted with French subjects, Cutler et al. observed shorter RTs to detect an initial phoneme in CV syllables than in CVC or CCV syllables. This effect is independent of the lexical status of the target-bearing item (word or nonword). However, RTs to words were shorter than RTs to nonwords.

In a second experiment English subjects were tested with the French experimental material. The effect of syllabic complexity is clearly replicated (CV RTs < CVC RTs < CCV RTs), whereas the lexical status of the target-bearing items does not have any significant effect. These results indicate that the syllabic-complexity effect seems to be independent of the subject's language. In fact, in an experiment conducted in English, Treiman et al. (1982) obtained results comparable to those reported by Cutler, Mehler, et al. (1987) using the same CV, CVC, and CCV types of syllables. RTs to detect the initial phoneme of CCV syllables differed significantly from those obtained with CVC and CV syllables. CV syllables were responded to faster than CVCs for seven out of the eight phoneme-bearing targets, but the difference was not statistically significant. Thus RTs were ranked in the same order as in the experiment by Cutler, Mehler, et al. (1987), CV < CVC < CCV.

Based on formal linguistic descriptions that ascribe an internal structure to the syllable in terms of onset and rhyme, Treiman et al. (1982) concluded that the onset component of the syllable constitutes a perceptual unit. It is the analysis of the syllable in elementary phonemic segments that accounts for the increase in reaction time to initial phonemes in CCV syllables.

In spite of the importance of the above results, which provide strong evidence that identification of the phoneme target depends on the internal structure of the syllable-bearing target, this effect must be reexamined in the light of Cutler, Butterfield, and William's (1987) critique of the experiments of Treiman et al. (1982). Cutler et al. maintain that the differences in RTs obtained by Treiman et al. for syllables bearing an initial consonant cluster (CCV) and those with a simple onset (CV or CVC) can be attributed to the fact that "phoneme monitoring targets are customarily specified for subjects in the form of a word mode. Morever, these models have usually had single-phoneme onset" (p. 407). Such a specification would lead the subject to represent the phoneme target as appearing in a simple CV context. This means that the "best correspondance" between the representation of the target induced by the experimental model and the properties of the stimulus accounts for the differences in RTs observed for CV and CCV stimuli.

To support this interpretation, Cutler, Butterfield, and Williams (1987) conducted an experiment in which subjects listened for word-initial target phonemes in continuous utterances. The target phoneme was a single onset in a CV syllable, or it corresponded to the first consonant of a consonant cluster onset in a CCV syllable. The model that subjects were given for the target matched or did not match the syllabic onset of the target bearing word. Faster RTs were observed when the target had been modeled with the same onset. Thus, when the instructions are to detect phoneme /b/ "as in *blue*," RTs are shorter for the stimulus *blend* (the same /bl/ consonantic cluster) than for the stimulus *besk* or *break*.

In view of these results it seems important to confirm the existence of a syllabic-complexity effect on phoneme detection with an experimental procedure capable of neutralizing the effect of the experimental model. The following two experiments were conducted to establish this point.

Experiment 1

In this experiment we studied the effect of two structural parameters on the detection of an initial phoneme target: the presence or absence of an initial consonant cluster and the open or closed nature of the syllable (the presence or absence of a final consonant). The crossing of these two parameters produced the four following syllabic structures: CV, CCV, CVC, CCVC. The phoneme targets were the voiced and the unvoiced occlusive consonants /b/, /d/, /p/, and /t/. Each phoneme target was contained in a list made up of 40 monosyllabic items. Most of these items were nonwords. Eight items corresponding to two of each of the four types of syllable were added to two distinct vowels. The consonant clusters bore the liquids /r/ or /l/ for /p/ and /b/ and the liquid /r/ for /d/ and /t/. The final consonant was an occlusive. Thus, for example, for the phoneme target /p/ the following eight items were used: CV, /pa/ and /po/; CVC, /pad/ and /pob/; CCV, /pla/ and /pro/; CCVC, /plad/ and /prob/. The thirty-two filler items were made up of eight items out of each of the four types of syllabic structures. Thus each type of syllabic structure was represented by the same number of items. The lists were recorded by a native speaker of French at regular intervals of two seconds between items.

Subjects were instructed that their task consisted in reacting as fast as possible to a particular target phoneme. Each target phoneme was presented before the corresponding experimental list together with a random list of four nonexperimental CV, CVC, CCV, and CCVC items (e.g., for the phoneme target /p/, "the target is /p/, as in the syllables *pli, pe, pru, pred*"). Before presentation of the experimental lists, subjects were given a trial list

Table 1
Mean response time (msec) to detect the target phoneme as a function of onset type (the presence or absence of a consonant cluster) and syllable type (open or closed)

	Initial consonant cluster	
	− (CV-)	+ (CCV-)
Open syllables (-V)	444	495
Closed syllables (-VC)	483	532

constructed along the same lines as the experimental lists with the target phoneme /g/. Two groups of twenty subjects (students at the University of Paris V) were associated to voiced (/b/ and /d/) and voiceless (/p/ and /t/) targets. The observed RTs are similarly distributed for voiced and unvoiced targets.

The results are presented in table 1 and figures 1 to 4. An analysis of variance shows a significant effect of the type of syllable ($F(3, 114) = 14.9$, $p < .001$), whereas the nature of the phoneme target reveals no significant effect ($F < 1$) and does not interact with syllable type ($F < 1$). Specific comparisons indicated that the two structural parameters introduce highly significant effects: syllables without an initial consonant cluster were responded faster than syllables with an initial cluster (CV and CVC RTs < CCV and CCVC RTs; $F(1,38) = 35.1$, $p < .001$), and open syllables were responded to faster than closed syllables (CV and CCV RTs < CVC and CCVC RTs; $F(1,38) = 11.6$, $p < .005$).

Results obtained in this experiment clearly indicate the existence of a close relation between detection of a phoneme in the initial position in a syllable and the structural complexity of the syllable. This complexity was estimated by taking into account the presence or absence of a consonant cluster in initial position and the open or closed nature of the syllable. Our results agree with those obtained by Cutler, Butterfield, and Williams (1987) and Treiman et al. (1982). The observed syllabic effect cannot be coherently interpreted within the hypothesis of a correspondance between the model and the syllabic context.

Experiment 2

In a second experiment we decided to study more directly the roles of target model and syllabic complexity in an experimental situation similar to that employed in our previous experiments and in the experiment of Cutler, Mehler et al. (1987), i.e., one using very short experimental lists.

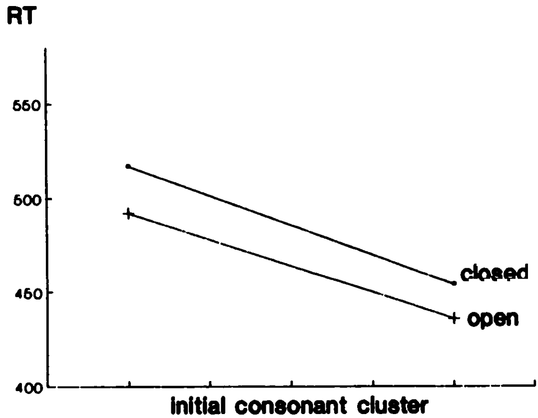

Figure 1
Mean response time (msec) to detect the target phoneme /b/ as a function of onset type (the presence or absence of a consonant cluster) and syllable type (open or closed)

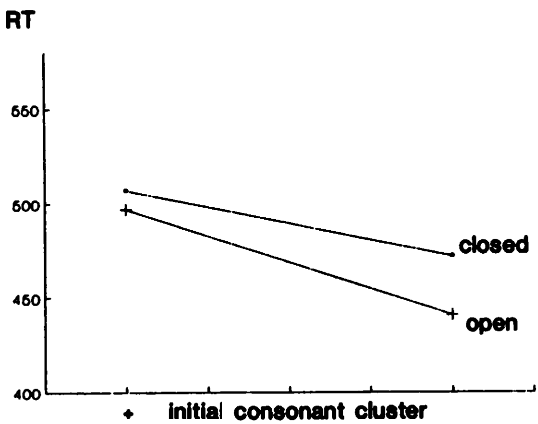

Figure 2
Mean response time (msec) to detect the target phoneme /d/ as a function of onset type (the presence or absence of a consonant cluster) and syllable type (open or closed)

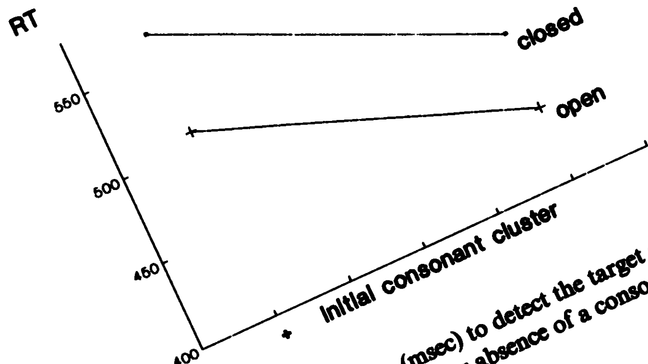

Figure 3
Mean response time (msec) to detect the target phoneme /p/ as a function of onset type (the presence or absence of a consonant cluster) and syllable type (open or closed)

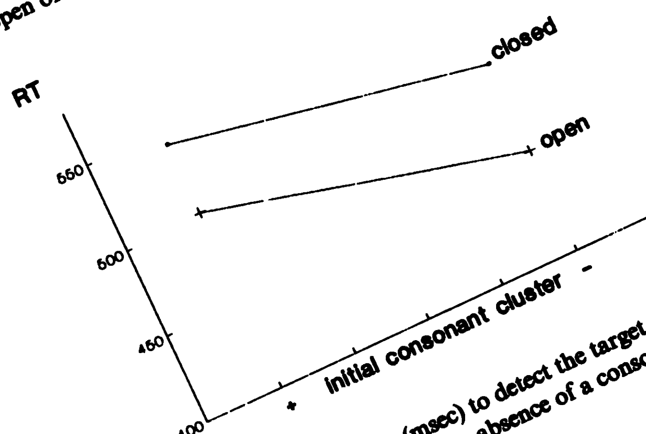

Figure 4
Mean response time (msec) to detect the target phoneme /t/ as a function of onset type (the presence or absence of a consonant cluster) and syllable type (open or closed)

Experimental stimuli consisted of six triplets made up of French bisyllabic words. Each triplet had three words beginning with the same initial phoneme. The initial syllable of these words (the structure of the target item) was CV, CVC, or CCV (for the phoneme /p/, for example, the triplet might be *palace*, *palmier*, and *plateau*). The experimental items were placed at the end of a short sequence (from one to six items). There were three experimental blocks, each including 18 experimental sequences and 12 filler sequences. In each block a particular experimental item was associated with a specific model. The model was a monosyllabic word corresponding to the initial syllable of a member of a triplet. Thus, for example, the experimental sequence bearing the item *palmier* was preceded in one block by the instruction "/p/ as in the word *pal*(*e*)," in another block by "/p/ as in the word *pa*(*s*)," and in the third block by the instruction "/p/ as in the word *pla*(*t*)." Thus, only in one list of a particular block (the first one in this example) did the model correspond to the initial syllable of the experimental item. The three blocks were recorded by a native speaker of French. The pauses between the words in a sequence were 1,500 msec long, and 10 seconds separated the presentation of each sequence. Each subject was given the three experimental blocks in a presentation order counterbalanced across subjects. Thirty subjects, psychology students at the University of Paris V, were tested.

The results obtained in this experiment are presented in table 2. An analysis of variance with subjects as a random variable reveals a very significant effect of type of target-item structure ($F(2, 58) = 16.5$, $p < .0005$), whereas the model structure does not introduce a significant effect ($F(2, 58) = 3.05, p > .10$). The interaction between these two factors is not significant ($F(4, 116) = 2.24, p > .10$).

To confirm the role of target-item structure without interference from the model, we compared the mean values corresponding to the three cases

Table 2
Mean response time (msec) to detect the target phoneme as a function of target-item structure and model structure

Target-item structure	Model structure CV	CVC	CCV	Mean RT
CV	341	378	389	369
CVC	373	390	397	387
CCV	420	414	409	414
Mean RT	378	394	398	

Table 3
Mean response time (msec) to detect the target phoneme as a function of matching and nonmatching target-item and model structures

	CV	CVC	CCV
Matching	341	390	409
Nonmatching	384	385	417
Mean RT	362	387	413

in which there existed a strict matching between the model and the target structure, i.e., the three diagonal cells in table 2. The observed difference is highly significant ($F(2, 58) = 11.3$, $p < .0005$). The order of response times as a function of target structure is the same as in experiment 1: CV < CVC < CCV. Contrary to the suggestion proposed by Cutler, Butterfield, and Williams (1987), this result shows that the differences in RTs introduced by the syllable structure of the target bearing item cannot be attributed exclusively to the degree of correspondence between the model and the target item.

Finally, we compared responses for matching and nonmatching target-item and model structures to estimate the role of model-structure correspondance. The global difference between matching and nonmatching responses was significant ($F(1, 29) = 4.19$, $p < .05$). This effect agrees with Cutler, Butterfield, and Williams's model hypothesis, though this effect interacts with item structure ($F(2, 58) = 4.52$, $p < .02$). As can be seen in table 3, the matching effect is due essentially to CV items.

The results observed in this experiment confirm the role of syllable structure in determining phoneme monitoring latencies. This structural effect is still found when there is a total correspondance between the model used to characterize the target phoneme and the initial syllable of the target-bearing item. In our experimental situation the effect of the correspondance between the model and the target-bearing item is strongly dependent on this latter factor. In fact, only the RTs to CV items are very sensitive to the model used to characterize the target phoneme.

Results obtained in the two previous experiments and in other recent experiments conducted by Dupoux indicate that, *at least for French*, the effect of syllabic complexity on phoneme-monitoring reaction time is robust and well established. This effect cannot be interpreted as being due to the degree of correspondance between the model given to the subject and the structural organization of the stimulus item. The original hypothesis according to which the segmental information necessary for phoneme detection is derived from a syllabic code does account for the data obtained.

The Role of the Syllable in Lexical Access

The syllabic hypothesis assumes that words are accessed by their initial syllable. In other words, initial syllables constitute the access code to the internal lexicon in spoken-word recognition. In this section we present some recent experimental results concerning this particular aspect of the syllabic hypothesis.

If this hypothesis is correct, we can predict that in phoneme monitoring, the role of lexical parameters should be qualitatively different for monosyllabic and polysyllabic items. Monosyllabic words automatically have access to their corresponding lexical representation without mediation by prelexical syllabic routines. Lexical information becomes immediately available and may be used to detect the initial phoneme of these items. On the other hand, polysyllabic words are accessed through their initial syllable, but the extraction of this syllable is generally not sufficient to isolate a particular word. In this case, phoneme detection may be based exclusively on prelexical information without the intervention of the lexical level.

This hypothesis about the differential nature of processing according to the syllable length of the target-bearing item has been supported in many experiments. In particular, whereas with monosyllabic items, phoneme targets are detected faster in words than in nonwords (Cutler, Mehler, et al. 1987; Rubin, Turvey, and Van Gelder 1976), this effect of lexical superiority is never found with polysyllabic items (Foss and Blank 1981; Segui, Frauenfelder, and Mehler 1981).

Our interpretation of the lexical-superiority effect assumes that the important parameter in these experiments was the *syllabic* length of the target-bearing items rather than the phonetic length or an even more crude, linguistically neutral length parameter, such as the total duration of the target-bearing word (Cutler and Norris 1979). It is clear, however, that these two interpretations, syllabic and durational, are able to explain the observed lexical-superiority effect. Consequently, it is important to obtain new empirical data that will enable us to choose between these two alternative interpretations.

In a recent series of experiments Dupoux and Mehler (in press) attempted to test the validity of the syllabic and durational interpretations. In a preliminary experiment these authors attempted to confirm the existence of a differential lexical effect for mono- and polysyllabic words using as a diagnostic variable the word-frequency effect. Subjects had to monitor initial phonemes of monosyllabic or bisyllabic words. Words were pre-

Table 4
Mean response time (msec) to detect the target phoneme, and percentage of errors, as a function of target syllable length and frequency (normal speech)

	Monosyllabic words	Bisyllabic words
Low frequency	414 (3.7%)	435 (1.7%)
High frequency	377 (2.0%)	422 (2.3%)
Difference	37*	13

*This difference is significant to $p < .001$.

sented in continuous lists, and target phonemes were specified auditorily (e.g., /p/ as in *Paris, Perpignan, Poitier*). Target words were of either high frequency (a mean of 148 for monosyllabic words and 102 for bisyllabic words) or low frequency (a mean of 3 for monosyllabic and bisyllabic words). They were paired according to their initial phoneme and syllabic structure (for example, the high-frequency /poule/ was paired with the low-frequency /pouf/). Identification points (in terms of phonemes) were also matched. Thirty students aged between 18 and 20 served as subjects. RTs were measured from the onset of the burst of the initial phoneme.

The results are presented in table 4. These results agree with those obtained for lexical superiority: in both cases lexical parameters affected the response time only to monosyllabic items ($F(1, 29) = 15.9, p < .001$, for monosyllabic words, and $F < 1$ for bisyllabic words). Monosyllabic words seem to be responded to on the basis of a lexical code, whereas bisyllabic items seem to be responded to on the basis of a prelexical code.

However, as noted before, the syllabic and durational interpretations both predict that with normal speech the lexical effect should be found only or essentially with monosyllabic items. In the syllabic interpretation, it is the number of syllables of the target-bearing item that determines the prelexical or lexical locus of the response, while in nonsyllabic models like Cutler and Norris's (1979) race model, it is the length (or duration) of the target-bearing item that determines the locus of the response. As for the interactive TRACE model (McClelland and Elman 1986), the possibility of obtaining a lexical effect in phoneme monitoring is a function of the level of lexical activation capable of affecting the phoneme unit corresponding to the target-bearing word. For phoneme-initial targets, detection responses cannot be affected by the lexical status of the target-bearing word, since in this case, when the phoneme target occurs, the level of activation of the target-bearing word is very low.

To test these interpretations in a second experiment Dupoux and Mehler (in press) presented subjects the same list of items but at a rate that was

Table 5
Mean response time (msec) to detect the target phoneme, and percentage of errors, as a function of target syllable length and frequency (compressed speech)

	Monosyllabic words	Bisyllabic words
Low frequency	439 (3.3%)	436 (3.8%)
High frequency	410 (1.3%)	434 (2.9%)
Difference	29*	2

*This difference is significant to $p < .002$.

twice as fast. Such a compression was obtained using an algorithm that averages adjacent periods of a 16-kHz digitized signal. This transformation leaves intact the spectral characteristics of speech, such as pitch, timbre, and formant structures. The important point is that in compressed format, bisyllabic words were shorter than uncompressed monosyllabic words but were still very intelligible. Since compression does not modify the number of syllables, the syllabic interpretation predicts the same result as that obtained with normal speech rate. In contrast, a durational interpretation predicts that under compressed presentation, frequency effects should be found with monosyllabic and bisyllabic items. In fact, the identification points of the compressed bisyllabic words occur temporally earlier than those of the normal monosyllabic word.

Thirty subjects participated in this experiment, and the main results are presented in table 5. As can be seen in table 5, only monosyllabic words showed a frequency effect ($F(1, 129) = 11.7$, $p < .002$, for monosyllabic words, and $F < 1$ for bisyllabic words). These results mirror perfectly those obtained in the previous experiment (RTs between these two experiments are significantly correlated; $r = .73$) and agree with the syllabic interpretation.

To reject an alternative interpretation according to which the absence of a frequency effect for bisyllabic items was artifactual and related to a bias in the choice of these items, a lexical-decision experiment was conducted with the same experimental lists. Results showed a robust and comparable frequency effect for monosyllabic and bisyllabic items.

Dupoux and Mehler's results are incompatible with a durational interpretation of the observed frequency effect and agree with the predictions derived from the syllabic hypothesis. The important experimental factor in this research was the syllabic length of the target-bearing words and not their global duration.

In recent papers Cutler proposed an initial-lexical-segmentation strategy well adapted to the structure of English vocabulary (Cutler and Norris

1988; Cutler, this volume). According to Cutler, the speech stream is segmented at strong syllables because these syllables generally occur in the initial position of polysyllabic words. At this level of formulation Cutler's hypothesis can be considered a particular version of the syllabic hypothesis. However, Cutler assumed that the distinction between strong and weak syllables is made not on the basis of the properties of a syllable per se but on the basis of the strong or weak nature of the vowel.

Conclusion

The experimental findings briefly reviewed in the three sections of this chapter generally agree with the predictions derived from a syllabic hypothesis for languages having clear syllabic boundaries, like French. According to this hypothesis, syllables correspond to natural units in speech segmentation and constitute the input to lexical access and phonemic analysis.

In the first section we examined data obtained in syllable-detection experiments indicating that French words and probably Spanish words too are segmented in syllabic units. Results obtained in Portuguese indicate that words are represented in the internal lexicon in a syllabic format. Finally, cross-language research shows that we must take into consideration the phonological structure of the languages studied to interpret the observed results. Segmentation strategies seem to be strongly related to the phonological and prosodic properties of natural languages.

Experimental results considered in the second section show that phoneme-detection RTs vary as a function of the structural complexity of the target-bearing syllables. The time to identify a particular phoneme depends on the manner in which this phoneme is encoded at the syllabic level. In particular, two structural parameters affected RTs: the presence or absence of a consonant cluster at the onset of the syllable and the open or closed nature of the syllable.

In the third section we presented results suggesting that French words are accessed by their initial syllable. The presence of a frequency effect depends on the syllabic length of the words and not on their global duration. According to the syllabic hypothesis, monosyllabic words are responded to from the lexical code, while polysyllabic words are responded to from a prelexical syllabic code.

As noted in Mehler, Dupoux, and Segui, this volume, our syllabic hypothesis may be related in a natural way to the development of language capacities and in particular to the construction of an internal lexicon.

Acknowledgments

The authors would like to thank Marie Eliane Binaghi, Cristophe Bonneuil, and Toby Mintz for assistance in stimulus selection and setting up experiments 1 and 2.

References

Bradley, D. C., Sanchez-Casas, R. M., and Garcia-Albea, J. E. 1988. Language specific segmentation strategies: I, Performance over native language materials. In preparation.

Cutler, A., and Norris, D. 1979. Monitoring sentence comprehension. In W. E. Cooper and E. C. T. Walker (eds.), *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*. Hillsdale, N.J.: L. Erlbaum.

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

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

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

Cutler, A., Mehler, J., Norris, D., and Segui, J. 1987. Phoneme identification and the lexicon. *Cognitive Psychology* 19: 141–177.

Cutler, A., Butterfield, S., and Williams, J. N. 1987. The perceptual integrity of syllabic onsets. *Journal of Memory and Language* 26: 406–418.

Dupoux, E., and Mehler, J. In press. Monitoring the lexicon with normal and compressed speech: Frequency effect and the prelexical code. *Journal of Memory and Language* 29, no. 3: 316–335.

Foss, D. J., and Blank, M. 1980. Identifying the speech codes. *Cognitive Psychology* 12: 1–31.

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

Mehler, J. 1981. The role of syllables in speech processing: Infant and adult data. *Philosophical Transactions of the Royal Society*, series B, 295: 333–352.

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

Mehler, J., Segui, J., and Frauenfelder, U. 1981. The role of the syllable in language acquisition and perception. In. T. F. Myers, J. Laver, and J. Anderson (eds.), *The cognitive representation of speech*. Amsterdam: North-Holland.

Morais, J., Content, A., Cary, L., Mehler, J., and Segui, J. 1989. Syllabic segmentation and literacy. *Language and Cognitive Processes* 4: 57–67.

Rubin, P., Turvey, M. T., and Van Gelder, P. 1976. Initial phonemes are detected faster in spoken words than in spoken nonwords. *Perception and Psychophysics* 19:394–398.

Sanchez-Casas, R. M. 1988. Access representations in visual word recognition. Ph.D., *Monash University*, Australia.

Segui, J. 1984. The syllable: A basic unit in speech perception? In H. Bouma and D. G. Bouwhuis (eds.), *Attention and Performance*, vol. 10, *Language control mechanisms*. Hillsdale, N.J.: L. Erlbaum.

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

Treiman, R., Salasoo, A., Slowiaczek, L. M., and Pisoni, D. 1982. Effects of syllable structure on adults' phoneme monitoring performance. Progress report no. 8, Speech Research Laboratory, Indiana University.

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

Zwitserlood, P., Schiefers, H., Lahiri, A., and van Donselaar, W. In preparation. The role of the syllable in the perception of Dutch.