

The Syllable's Differing Role in the Segmentation of French and English

ANNE CUTLER

MRC Applied Psychology Unit, Cambridge, United Kingdom

JACQUES MEHLER

Centre de Science Cognitive et Psycholinguistique, CNRS, Paris, France

DENNIS NORRIS

MRC Applied Psychology Unit, Cambridge, United Kingdom

AND

JUAN SEGUI

Laboratoire de Psychologie Expérimentale, Associé au CNRS, Paris, France

Speech segmentation procedures may differ in speakers of different languages. Earlier work based on French speakers listening to French words suggested that the syllable functions as a segmentation unit in speech processing. However, while French has relatively regular and clearly bounded syllables, other languages, such as English, do not. No trace of syllabifying segmentation was found in English listeners listening to English words, French words, or nonsense words. French listeners, however, showed evidence of syllabification even when they were listening to English words. We conclude that alternative segmentation routines are available to the human language processor. In some cases speech segmentation may involve the operation of more than one procedure. © 1986 Academic Press, Inc.

Speech is continuous; speakers provide few reliable cues to the boundaries of words, phrases, or other units of meaning. Thus a major part of the listener's task of

This research is a European Psycholinguistics Association collaborative project. The authors' names are listed in alphabetical order. Financial support for the collaboration was provided by a twinning grant from the European Training Programme in Brain and Behaviour Research of the European Science Foundation. J.M. and J.S. also acknowledge financial support from CNET (Convention 837 BD 2800790) and CNRS (ATP "Aspects Cognitifs et Neurobiologiques du Langage"). A brief report of Experiments 1, 2, and 4 appeared, along with the results of Mehler et al. (1981), as part of a letter to *Nature (London)* (Cutler et al., 1983). We thank Bert Jansen for technical assistance with Experiment 3. Requests for reprints should be addressed to the first author at MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, United Kingdom.

extracting meaning from speech is segmenting the continuous signal into portions which can be mapped onto such meaning units. Psycholinguists have therefore for years—almost for decades—concerned themselves with the question of what segmentation units are necessarily, or possibly, involved in speech understanding.

The two sublexical units which have received the greatest amount of psycholinguistic attention are the phoneme and the syllable. Considerable experimental evidence supports each of them. The phoneme, being the smallest linguistic unit into which an utterance can be sequentially decomposed, may be said to have a built-in advantage, since it provides the smallest set of perceptual units. The smaller the set of stored units against which sections of

the speech wave have to be tested, the faster the testing process will be; hence a phoneme-based recognizer would seem to be more efficient than a recognizer based on syllables, which, since they can be described as combinations of phonemes, constitute a rather larger candidate set.

Recognition efficiency may trade off with segmentation efficiency, however; it may be that a phoneme-based recognizer is saddled with a very costly front end, as segmenting speech into phonemes may be extremely difficult. The acoustic information signaling a particular phoneme can be distributed across or dependent on the characteristics of neighboring phonemes, such that it can be impossible to identify the phoneme without reference to its context. What syllables lose in terms of candidate set compactness, they may gain in terms of robustness with respect to the acoustic context.

Certainly there is impressive evidence for the reality of the syllable in speakers' and listeners' psycholinguistic representations of words. Slips of the tongue show numerous constraints of syllable structure on possible slip patterns (Fromkin, 1971; Mackay, 1972); linguistic games ("play languages") in many languages show similar syllabically defined rules (Sherzer, 1982) and language users can readily learn such rules (Treiman, 1983). Young children can identify the number of syllables in a word in a tapping task, where they cannot identify the number of phonemes (Lieberman, Shankweiler, Fisher, & Carter, 1974); Morais, Cary, Alegria, & Bertelson (1979) found that illiterate adults performed similarly to the preliterate children, suggesting that phonemic competence may depend on alphabetic literacy while syllabic competence does not. Number of syllables in a word is similarly highly likely to be preserved in word substitution slips (Fay & Cutler, 1977) and tip-of-the-tongue guesses (Brown & McNeill, 1966).

Studies of segmentation units in continuous speech perception have chiefly used

various monitoring tasks (see Cutler & Norris, 1979, for a review of these). In such tasks, listeners are consistently faster monitoring for syllable-sized targets than phoneme-sized targets (Savin & Bever, 1970; Segui, Frauenfelder, & Mehler, 1981). There has been considerable controversy over whether this robust finding is due to the processing sequence of the respective units of perception or to factors peculiar to monitoring tasks; for our present purposes what is important in these studies is that listeners are clearly capable of extracting syllabic units from continuous speech when required to do so.

Mehler, Dommergues, Frauenfelder, and Segui (1981) went beyond this to show that a potentially syllable-sized monitoring target is responded to faster when it actually corresponds to a syllable in the heard speech stream than when it comprises rather more or rather less than a syllable. This study, entitled "The Syllable's Role in Speech Segmentation," provided the stimulus for the present work and will be described in detail. Mehler et al. presented University of Paris students with lists of unrelated French words and required them to monitor within each list for a specified word-initial sequence of sounds. This target was either a consonant-vowel (CV) sequence such as *pa-* or a consonant-vowel-consonant (CVC) sequence such as *pal-*. The words which contained the target had one of two syllable structures: the initial syllable was either open (CV), as in *palace*, or closed (CVC), as in *palmier*. Note that both *palace* and *palmier* begin with the same three sounds /p//a//l//; thus a positive detection response would be appropriate to either word for either *pa-* or *pal-* targets. However, each target corresponds to a complete syllable only in one of the two words: *pa-* is exactly the first syllable of *palace*, but less than the first syllable of *palmier*, whereas *pal-* is exactly the first syllable of *palmier*, but more than the first syllable of *palace*. Response time was significantly faster when the target

matched a complete syllable of the word; thus *palace* produced faster RTs to *pa-* than to *pal-*, and *palmier* faster RTs to *pal-* than to *pa-*.

This result is strong evidence that listeners in this experiment were syllabifying the incoming words; that is, the syllable was functioning as an effective segmentation unit. Mehler et al. concluded that the "the syllable constitutes a unit of speech processing" and "syllable segments could well serve as accessing units" to the lexicon (Mehler et al., 1981; pp. 303–304).

If syllabification is to be considered a basic speech perception routine, then it should obviously be a routine available to speakers of all languages. Psycholinguistics aims to model the characteristics of the universal human language processing mechanisms rather than the processing of a specific language. Syllable structure, however, is notoriously an area in which languages differ widely. This is not to say that some languages have syllables while others don't; part of the attraction of the syllable as a universal processing unit is that it is a true linguistic universal. All languages have vowels and consonants, and all can be described in terms of syllables. But whereas some languages have very uniform syllable structure, for example, only CV sequences, other languages tolerate great variation in syllable weight, for instance from a short V to a CCCV:CCC sequence (where V: represents a long vowel). Clearly the task of syllabification is rather more complex in languages with a variety of syllable structures than in languages with simple structure.

Even more importantly for a segmentation unit hypothesis, languages differ in the degree to which syllable boundaries are clear and unambiguous. In general, stress languages tend to have greater variation in syllable weight than nonstress languages, and partly as a consequence of stress assignment, syllable boundaries in stress languages are frequently unclear. Compare, for instance, the English word *palace* with

French *palace*. In French the syllabification is clear—there is a syllable boundary between *pa-* and *-lace*. In English, however, the syllable boundary falls neither clearly before nor after the */l/*. Both [paɛl] and [læs] are permissible English syllables, [paɛ] and [æs] are less satisfactory. Phonologists represent the syllable structure of French *palace* as [pa][lace], but of English *palace* as [pa[l]ace]; that is, the */l/* properly belongs to both first and second syllables (Anderson & Jones, 1974; Kahn, 1976). Segments which belong to two syllables at once are said to be ambisyllabic. In stress languages, intervocalic consonants preceding an unstressed vowel are frequently ambisyllabic.

Ambisyllabicity obviously poses difficulties for syllabically based segmentation. If syllable boundaries are unclear, then segmenting speech into syllables runs up against the same kind of problems which we ascribed above to phonemic segmentation. Thus it is reasonable to ask whether syllabification would in fact be an efficient perceptual procedure in languages with ambisyllabic consonants, or whether it might not just make speech perception even more difficult. More specifically, it is reasonable to ask whether Mehler et al.'s clear evidence for syllabification using French materials will prove replicable in a language with widespread ambisyllabicity such as English.

Accordingly we decided to repeat Mehler et al.'s (1981) study in English. In the replication, we compared words with ambisyllabic consonants (e.g., *palace*) with words with unambiguous initial syllables. In *palpitate*, for example, the first syllable is clearly *pal-*. This comparison allows several opportunities for syllabification to be observed. Mehler et al., comparing RT to targets of the *pa-* and the *pal-* type, found RT to *pa-* faster in *palace* than in *palmier*, RT to *pal-* faster in *palmier* than in *palace*. If English listeners can ignore ambisyllabicity and impose differing initial syllable structures on *palace* and *palpitate*, it is just

possible that we might replicate the cross-over interaction found by Mehler et al. However, we consider this result to be unlikely. Rather, a syllabification procedure in English might be evinced with easily syllabified words, but not with hard-to-syllabify words. In this case we might expect *pal-* targets to be responded to faster than *pa-* targets in *palpitate*, but no difference in RTs to the two types of target to be found with *palace*. Yet a third possibility is that RTs to *pal-* targets might be faster than to *pa-* targets in both types of word, since *pal-* is the first syllable of *palace* even though the /l/ is part of the second syllable as well.

Thus three possible patterns of result from this replication could support the extension of the hypothesized syllabification procedure to the processing of English: (1) an interaction as found by Mehler et al. with CV targets responded to faster in CV[C]-syllabified words, and with CVC targets in CVC-syllabified words; (2) an advantage for CVC over CV targets with CVC-syllabified words, but no difference in the other word type; or (3) an overall RT advantage for CVC targets with both types of word.

EXPERIMENT 1

Method

Materials. The materials were chosen so as to mimic as closely as possible the stimuli used by Mehler et al. (1981). Seven pairs of unambiguous English content words (nouns and verbs) of similar frequency sharing the same initial three phonemes (CVC) were selected. In each pair one member had a syllable boundary after the initial CVC, while in the other member of the pair the third phoneme was ambisyllabic; that is, the second consonant of the initial CVC could be said to belong to both syllables. Thus in the pair *balcony/balance* the first three phonemes are identical, yet this CVC sequence forms a complete syllable only in *balcony*; *balance* has an ambisyllabic [l]. For each of the seven

pairs the vowel was [ae] and the second consonant [l] (Mehler et al. also used [r] as a second consonant, but in standard British English [r] pairs are not possible with constant vowel quality in the first syllable: compare *parish* with *party*). The seven pairs were: *balance–balcony*, *calorie–calculate*, *galaxy–galvanize*, *malady–malcontent*, *palace–palpitate*, *salad–salvage*, *talon–talcum*.

Fifty-six lists of unrelated words were compiled, varying in length from one to five words. Twenty-eight of the lists contained one of the 14 experimental words in final position (which could be second, third, fourth, or fifth position). Thus, as in the Mehler et al. experiment, each stimulus word occurred twice (although preceded by different filler items) in the experiment, so that it could be presented with both CV and CVC targets. The two members of any pair occurred in the same position in their respective sequences.

Of the 28 distractor sequences, 14 had no words matching the targets, in order to prevent subjects from responding to sequence-ending rather than target-occurrence, while the others had target matches anywhere from first to fifth position. The target-matching words in these sequences included some with third phonemes other than [l] for example, *tapestry* for target *ta-*. The complete set of 56 sequences and 10 practice sequences was recorded at a slow-normal rate by a male native speaker of British English; as in the Mehler et al. study, the words in each sequence were separated by 2-s intervals and the sequences themselves by 10-s intervals. An extra word added to the end of each sequence was removed from the recording during the creation of the final stimulus tape; this manipulation removed any prosodic cues to sequence ending which could have cued subjects' responses.

Two target orders were prepared. Each of the 14 targets (*ba-*, *bal-*, *pa-*, etc.) occurred four times in the experiment, twice on experimental sequences, once on a dis-

tractor sequence without target match, and once on another distractor sequence. Each experimental word occurred once with one alternative target in the first half of the experiment and again with the other possible target in the second half. These target assignments in the first target order were exchanged in the second. No target occurred on two successive sequences, and target type (CV, CVC), sequence length, and position of distractor sequences with and without target matches were counterbalanced across first and second halves of the experiment.

Subjects. Subjects were 24 members of the University of Sussex community, who were each paid a small fee for participating. Twelve subjects were assigned to each target order condition.

Procedure. The experimental tape was presented binaurally over headphones. Targets were presented visually on a VDU screen and presentation of each new target was accompanied by the sounding of the VDU's bell which could be heard by the subjects despite the headphones. At the onset of each experimental item a signal on the other channel of the tape, inaudible to the subjects, started the clock of a Motorola 6809 microcomputer; timing was stopped by the subject pressing the response key.

Results

Mean response times for each item and each subject were computed, omitting responses shorter than 100 ms or longer than 1000 ms (data omitted for this reason, or because the subject failed to respond, comprised less than 5% of all responses).

The mean response times for each condition are displayed in Table 1. Separate analyses of variance were carried out on the subject and item means. No main effect or interaction reached the .05 level of significance in both analyses, and only two effects even reached this level in one analysis. As can be seen from the Table, words with ambisyllabic consonants elicited faster

TABLE 1
MEAN RT (ms) IN EXPERIMENT 1 (ENGLISH SUBJECTS,
ENGLISH WORDS)

	CV[C] words (e.g., <i>balance</i>)	CVC words (e.g., <i>balcony</i>)	
CV targets (e.g., <i>ba-</i>)	456	502	479
CVC targets (e.g., <i>bal-</i>)	448	514	481
	452	508	

responses than words with clear syllable boundaries, but although this effect was highly significant in the subjects analysis ($F_1(1,22) = 38.8, p < .001$), it failed to reach the set level of significance in the items analysis ($F_2(1,12) = 3.5, p < .09$). The same was true of the tendency for the second occurrence of a given word to elicit faster responses than the first ($F_2(1,22) = 9.9, p < .005$; $F_2(1,12) = 4.7, p < .06$).

What is particularly noticeable about these results is the absence of any pattern interpretable as an effect of syllabification. Recall that we suggested that any of three patterns of results could possibly be so interpretable: an overall advantage for CVC targets over CV with both kinds of word; an advantage for CVC over CV targets with CVC-syllabifiable words, but no difference with words containing ambisyllabic consonants; or an advantage for CVC targets with CVC-syllabifiable words, and for CV targets with words containing ambisyllabic consonants. Statistically, these patterns would have shown up as a main effect of target type or an interaction of target type with word type. Neither effect was found, nor did either interact with the additional variable of first versus second occurrence of an experimental word. The main effect of subject groups (i.e., target order condition), which could be examined in the subjects analysis only, was not significant, and although it did participate in some higher order interactions, none was such that the groups appeared to be differing specifically on the appearance of an effect of syllabification.

Discussion

As predicted, the interaction of CV versus CVC targets with initial syllable structure, which Mehler et al. found with French materials, failed to replicate with the phonologically different English materials. Moreover, no effect of any kind which could be interpreted as evidence for syllable-based segmentation was found in the present experiment. The subjects in our study were obviously not applying the same type of within-word segmentation in searching for the specified targets as were the French subjects in the earlier study.

There are several possible interpretations of the mismatch between the two findings. For instance, it might simply be the case that the materials used in Mehler et al.'s study incorporated some low-level asymmetry such that CV targets were physically better matched by the CV-syllabified words and CVC targets by the CVC-syllabified words. Alternatively, the materials in the present study might simply have inhibited our subjects from employing syllable-based segmentation, perhaps because they included a substantial subset of words with ambisyllabic consonants. In this latter case, one could still maintain the argument that syllabification is a universal basis for segmentation, but one which language users choose whether to apply as a function of whether the language materials they are using encourage its application, that is, whether they are easily syllabified.

On the other hand, it may be that our results are evidence of an interesting difference between English and French listeners—that syllabification is a routine available to some speakers and not to others. For instance, it may be used by native speakers of French (and perhaps other languages like French) but not by native speakers of English (or languages like it). In this case we would have to explain not experimental artifacts but cross-linguistic processing differences.

A simple route is open to us to shed light on these alternatives. If there is in fact no syllabification strategy at all, and the results from Mehler et al.'s study are an artifact of physical attributes of their materials, then English listeners should be just as susceptible to these low-level matching effects, which presumably do not depend on phonetic or linguistic structure; that is, they should show what would appear to be a syllabification effect with the materials from the earlier study.

Similarly, if English listeners can indeed employ syllable-based segmentation, but have merely been discouraged from doing so in an experiment containing words which were difficult to syllabify, we would expect them to be able to call on the procedure when presented with materials encouraging its use. Thus the easily syllabified French materials of the Mehler et al. study should provide an opportunity for English subjects to show evidence of syllabification.

If English listeners, however, simply do not have the syllabification routine available to them, they will be unable to impose it even upon easily syllabified French. Thus the simple replication of Mehler et al.'s experiment in French with English subjects will clarify interpretation of our English-language experiment: if we replicate the syllabification effect which Mehler et al. found, then the mismatch between the two experiments is probably due to inadequacy of the materials in either the first or the second study. However, if we fail to replicate the effect, that is, if English listeners fail to syllabify even when syllabification is easy, then syllabification may be a *language-specific* processing routine.

Experiment 2 was simple to initiate; the original tape from the Mehler et al. study was mailed across the English Channel.

EXPERIMENT 2

Method

Materials. The materials were those

used (and fully described) by Mehler et al. (1981; Experiment 1).

Subjects. Subjects were 24 members of the University of Sussex community, who were each paid a small fee. Twelve subjects were assigned to each target order condition. No subject was fluent in French, and although most had studied some French, none had continued with it as far as or beyond the end of secondary schooling.

Procedure. The procedure was the same as in Experiment 1, except that as in Mehler et al.'s study targets were presented on a pack of cards, the subject turning to the next card on hearing the words "carte suivante" at the end of each sequence. Subjects were reminded of relevant French grapheme-phoneme correspondences where these differed from English (e.g., that the sequence "pal" is pronounced [pal] and not [paɛl]), and were given the opportunity to hear the initial set of practice items a second time if they found the French at first too confusing for the monitoring task to be easily performed. Only one or two subjects took this opportunity.

Results

Mean response times for each item and each subject were computed, omitting responses shorter than 100 ms or longer than 1000 ms. Data omitted for this reason or because the subject failed to respond comprised less than 5% of the total.

Mean response times for each condition are shown in Table 2. Separate analyses of

variance were carried out on the subject and item means. As in Experiment 1, no effect reached the .05 level of significance in both analyses; the RT advantage of words with CV first syllables over words with CVC first syllables was the only one to approach significance ($F_1(1,22) = 32.53, p < .001$; $F_2(1,8) = 4.63, p < .07$). The comparison of first with second occurrence of items did not reach significance in either analysis, nor did the crucial interaction of word type with target type. The effect of subject group (target order condition) in the subjects analysis was not significant, and interacted only with the word type by target type comparison (one group showed a weak crossover interaction, insignificant when analyzed separately, in the direction predicted by the syllabification hypothesis; the other group showed a strong interaction, significant when analyzed separately, in the opposite direction).

Discussion

Again English-speaking subjects showed no sign of the syllabification effect found by Mehler et al. with French, even though in this case the language materials presented to them could be easily syllabified and therefore provided a favorable environment for application of syllable-based segmentation. This experiment therefore suggests that the results from Experiment 1 are in no way artificial; English listeners appear not to be segmenting speech in the way French listeners do. Our results, therefore, suggest that segmentation strategies in continuous speech perception may be *language-specific*.

We will return to this question in the general discussion below. In the meantime, there is one puzzling aspect of the results so far which we would like to address. The prediction from the syllabification hypothesis for both experiments was for effects of target type; these were not found, but instead, there was evidence of effects of word type. In both experiments, there

TABLE 2
MEAN RT (ms) IN EXPERIMENT 2 (ENGLISH
SUBJECTS, FRENCH WORDS)

	CV words (e.g., <i>balance</i>)	CVC words (e.g., <i>balcon</i>)	
CV targets (e.g., <i>ba-</i>)	431	471	451
CVC targets (e.g., <i>bal-</i>)	419	500	459
	425	485	

was a tendency for CVC-syllabified items (*balcony*, *balcon*), irrespective of target type, to be responded to slower than other items (*balance*).

One possibility is that this result stems from a postlexical effect of some kind. Differences between the two types of word include that the CVC-syllabified set in the English experiment contained verbs as well as nouns, and that the same set in the French experiment contained some items without English cognates; neither was true of the other word type. A postlexical effect would be evidence that the monitoring task is tapping a rather later level of processing than is assumed in the prelexical explanation of the syllabification effect; it would leave us unclear whether the difference we had found between French and English listeners consisted in the segmentation procedures they used or the way they performed the monitoring task. Accordingly we decided to perform another analogous study in which no postlexical effects could reasonably be expected; that is, we presented subjects with the same monitoring task on lists of *nonwords*. If the pattern of results found in the previous experiments had been in any way determined by lexical factors, we should not expect to find that pattern, or that component of a pattern, replicated here.

EXPERIMENT 3

Method

Materials. Twenty-four nonwords were constructed, 12 with CVC initial syllable and 12 beginning with the same CVC sequences but with the second consonant occurring in an ambisyllabic position. They were *balgart*, *balvize*, *calfert*, *calcoze*, *dalmeece*, *daltade*, *galbime*, *galdape*, *palpone*, *palsuke*, *talvage*, *talzeen*; *balash*, *balic*, *calet*, *caluft*, *dalopp*, *dalun*, *galef*, *galince*, *palost*, *paluck*, *talatt*, *talect*. (The vowel and second consonant of the initial CVC sequence were always [ae] and [l] for the same reasons as in Experiment 1; the

vowel in the second syllable of the first 12 nonwords was always full, and that in the last 12 always reduced.)

Forty-eight lists of nonwords were compiled, 24 of which had one of the above nonwords in final position (which could be third, fourth, or fifth position). Of the other lists 10 had no occurrence of the specified target, and 14 had target-matching items in any position from first to fifth. Eight practice lists were also constructed.

The lists were recorded at a normal speaking rate by a male native speaker of British English. Again, an extra item at the end of each sequence was removed from the recording in the course of making the final tape.

Two target orders were constructed; each experimental item was assigned a CV target in one order and CVC in the other, with target type counterbalanced across orders. (Because it was easy to make up nonwords we could include more items in this experiment than in the preceding two, rendering it unnecessary to present each item twice; thus for any item, target type was a between-subjects comparison rather than a within-subjects comparison as in the previous studies.)

Subjects. Subjects were 24 members of the subject panel of the Applied Psychology Unit, Cambridge; each was paid a small fee for participating. Twelve subjects were assigned to each target order condition.

Procedure. The procedure was as in Experiment 1 except that the laboratory computer was a PDP 11/23.

Results

Item and subject means were computed and separate analyses of variance were conducted. Again, RTs over 1000 ms or shorter than 100 ms were omitted; these, with failures to respond, comprised less than 1% of the total data.

Condition means are shown in Table 3. Once again no effect reached the .05 level of significance in both analyses, although

TABLE 3
MEAN RT (ms) IN EXPERIMENT 3 (ENGLISH SUBJECTS,
"ENGLISH" NONWORDS)

	CV[C] nonwords (e.g., <i>balic</i>)	CVC nonwords (e.g., <i>balgart</i>)	
CV targets (e.g., <i>ba-</i>)	391	399	395
CVC targets (e.g., <i>bal-</i>)	370	393	381
	380	396	

again the RT advantage of words with ambisyllabic consonants over words with CVC syllabification almost reached that level ($F_1(1,22) = 4.23, p < .06$; $F_2(1,18) = 4.85, p < .05$). The main effect of subject group (target order) in the subjects analysis was insignificant; it interacted only with the two-way comparison of item type and target type. Post hoc analyses showed that the source of this interaction was that RT to CVC targets with ambisyllabic words was significantly faster than RT in all other conditions for one group, but although this was also the fastest condition for the other group, it was not significantly faster.

Discussion

The pattern of results from this experiment is so similar to those from Experiments 1 and 2 that it seems highly probable that whatever subjects were doing, they were doing it in the same way in all three experiments. Thus the tendency to an RT effect of item type in the first two studies can hardly be dismissed as a postlexical artifact. Therefore this effect remains unexplained (and yet more in need of explanation due to the additional evidence in its favor from this experiment).

A further possibility is that the difference simply reflects a low-level acoustic difference between the two sets of words—for instance a length difference. It could be that CVC-syllabified items tend to be longer, and subjects wait till the end of the item before initiating a response, so that the longer type of item tends to produce longer RTs. If this were true, we would ex-

pect to find a positive correlation between measured item length and the average RT to that item. Accordingly we digitized the recordings of the experimental items from the preceding experiments, measured the length of each item, and correlated these measurements with RT. In Experiment 1, the CVC items indeed proved to be longer (mean = 667 ms) than the items with ambisyllabic consonants (540 ms), but length was not significantly correlated with RT ($r = .11$). The French items in Experiment 2 did not show the expected length difference: on the contrary, CV-syllabified items (747 ms) were longer than CVC (657 ms); the correlation of length with RT was in this case significant, but *negative* ($r = -.58, p < .01$)—the longer the item, the faster tended to be the response. The nonword items of Experiment 3 showed the same overall length asymmetry as the real English words of Experiment 1 (554 vs 716 ms), but again no significant correlation was found with RT ($r = .27$).

Another possible low-level acoustic factor is the difference in articulation of the postvocalic consonant in the different word types. Indeed, examination of spectrograms of the materials from Experiments 1 and 2 indicated that minor differences between the postvocalic consonants were visible in most pairs in both sets of materials, and were of the same nature in both sets of materials. These differences could serve as an acoustic basis for efficient syllabification. It is less clear that they would offer the basis for a simple word-type response difference, and it is still totally unexplained why the same articulatory effect in French and in English should be exploited in quite different ways by French and English listeners.

Although the results of Experiments 2 and 3 suggest that the difference between the segmentation behavior of French subjects (in Mehler et al.'s original experiment) and English subjects (in Experiment 1) is a real language-specific processing difference and not an artifact, it seemed to us

desirable to strengthen the argument by providing further evidence of French listeners' use of syllable-based segmentation. So far we have three demonstrations that English listeners fail to syllabify, but only one demonstration with this task that French listeners do syllabify.

The strongest test of our claim would be to show that French listeners use a syllabification routine irrespective of familiarity of the materials they are listening to. Such a test was, again, simple to undertake, since materials in English were available. Accordingly, the original tape from Experiment 1 was mailed across the English Channel and presented to French listeners in our next experiment.

Note that half the materials in this experiment—those with ambisyllabic consonants—are in fact not suitable for syllabification. Such linguistic materials are “un-French,” and hence provide a particular challenge to listeners attempting to segment by syllables. As with Experiment 1, we can make several alternative hypotheses about the results of this study. Our prediction is that the French listeners will show evidence of syllabification, as we claim that this is their regular segmentation routine for speech. Three patterns of results could be interpreted as such evidence:

(1) These subjects could show the cross-over interaction found by Mehler et al., with CV targets being responded to faster in CV[C] words, and CVC targets in CVC words. In this case we would have to assume that French listeners were ignoring the presence of ambisyllabicity in English phonology, and simply accepting un-English syllabifications such as [bae]. Such syllables also do not occur in French, but only because the English vowels in our materials do not occur in French; open syllables with short vowels *are* acceptable in French.

(2) The French subjects could show a syllabification effect with CVC words, but no effect with CV[C] words. This would be

evidence that the syllabification procedure simply failed when presented with ambisyllabic consonants.

(3) There could be an advantage for CVC targets over CV targets with both word types. This could provide evidence that French listeners are correctly attaching the ambisyllabic consonants to the preceding vowel and hence are treating both types of word as if they began with CVC syllables.

Any other pattern of results would argue against our claim that French listeners should employ syllable-based segmentation irrespective of materials.

EXPERIMENT 4

Method

Materials. The materials were those used in Experiment 1.

Subjects. Subjects were 20 members of the University of Paris V, who participated in the experiment on a voluntary basis. Ten subjects were assigned to each target order condition. No subject was fluent in English, and although some subjects had studied English, none had continued beyond secondary school.

Procedure. The procedure was essentially the same as for Experiment 2, except that the laboratory computer was a Telemechanique T1600. The subjects were reminded of English grapheme–phoneme correspondences (where these differ from French), and were given the opportunity of a repeated hearing of the initial set of practice items.

Results

Mean response times were computed across subjects and items, omitting responses shorter than 100 ms or longer than 1000 ms (these, with subjects' failures to respond, comprised less than 5% of the data). Means for each condition are shown in Table 4.

Separate analyses of variance were car-

TABLE 4
MEAN RT (ms) IN EXPERIMENT 4 (FRENCH SUBJECTS,
ENGLISH WORDS)

	CV[C] words (e.g., <i>balance</i>)	CVC words (e.g., <i>balcony</i>)	
CV targets (e.g., <i>ba-</i>)	448	467	458
CVC targets (e.g., <i>bal-</i>)	457	440	448
	453	454	

ried out on the subject and item means. The main effects of word type, target type, and first versus second presentation were not significant in either analysis. The interaction of word and target type, on the other hand, showed the effect found in the original Mehler et al. study: $F_1(1,18) = 5.49$, $p < .02$; $F_2(1,12) = 4.51$, $p = .055$. Once again, we have clear evidence that French speakers are using syllable-based segmentation.

However, although this crucial interaction of word type with target type was significant, in the predicted direction, we noted that the advantage of CVC targets with CVC words (27 ms) was very much larger than the advantage of CV targets with CV[C] words (9 ms). Accordingly, we conducted separate t tests on the target type effect for each word type. With CVC words, the difference was significant at the .05 level across subjects ($t(19) = 1.73$) and at the .07 level across items ($t(6) = 2.2$), whereas the difference with CV[C] words did not approach significance in either analysis.

As with the previous experiments, we correlated mean item RT with measured item length; once again, the correlation was not significant ($r = .002$).

Finally, note that the fact that the French subjects in this experiment showed no trace of the word type differences (RT advantage for CV-syllabified or ambisyllabic words) shown by the English subjects in the preceding three experiments provides conclusive evidence that this difference

was not due to some low-level acoustic artifact.

GENERAL DISCUSSION

Our four experiments, taken together, present a clear picture: the syllable's role in speech segmentation is different for speakers of different languages. Native speakers of French appear to use syllabification whether they are listening to familiar easy-to-syllabify French words or unfamiliar and hard-to-syllabify English words. Native speakers of English do not syllabify in the same way, whether they are listening to hard-to-syllabify English words, easily syllabified French words, or nonwords. Figure 1 compares the results of the present experiments with those of the Mehler et al. (1981) study.

The general picture is therefore a great deal more complicated than that suggested by the original work of Mehler et al. (1981). Their experiments, in which French listeners were presented with French words, suggested that the syllable functions as a segmentation unit in on-line speech processing. The present experiments were prompted, however, by consideration of phonological differences across languages. Not all languages have equally regular syllable structures, and not all languages have clear syllable boundaries. We reasoned that if the syllable were indeed a basic segmentation unit for all languages, such phonological differences would imply that segmentation should be far easier in some languages than in others: regular syllable structure and clear syllable boundaries should aid segmentation; irregular syllable structure and obscure syllable boundaries obstruct it. This would amount to a claim that some languages are intrinsically easier to perceive than others! However, we know of no evidence to date suggesting that this is so. Therefore we proposed an alternative hypothesis: that syllabification would not be used when the phonological structure of the language in question would

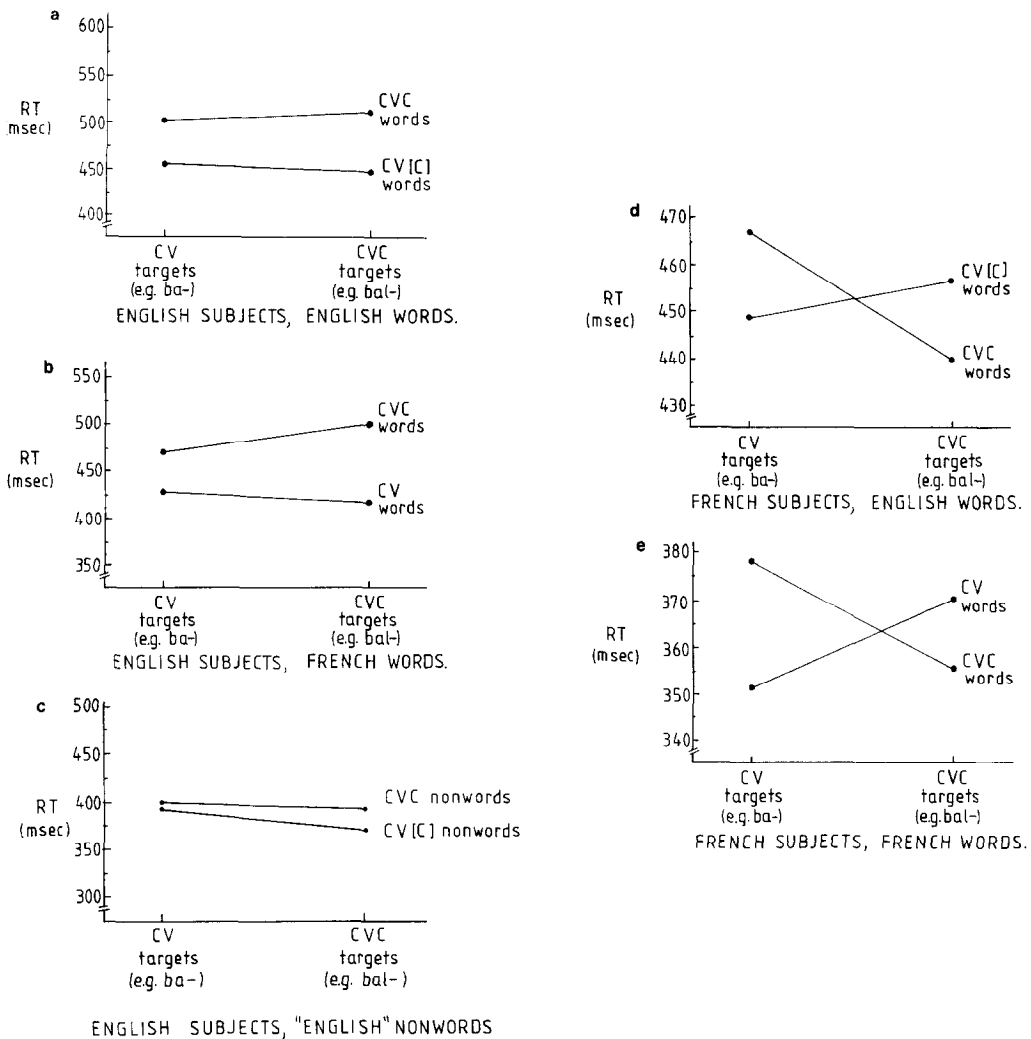


FIG. 1. Response time as a function of target size and word type for Experiments 1-4 (a-d) compared with Mehler *et al.* (1981), Experiment 1 (e).

render it inefficient. An experiment using a language with varied syllable structure and unclear boundaries should show no evidence of listeners using the syllable as a segmentation unit.

The French language, in which Mehler *et al.*'s original experiment had been conducted, has a comparatively regular syllable structure, and clear syllable boundaries. In particular, it has minimal consonantal ambisyllabicity, especially between the first and second syllable of polysyllabic words. English, on the other hand, has extremely irregular syllable structure (both *a*

and *screeched* are monosyllabic words), and frequently unclear syllable boundaries, particularly as a result of the widespread occurrence of ambisyllabicity in intervocalic consonants. Ambisyllabic consonants between the first and second syllable of polysyllabic words—as in *salad*—are especially common. English, therefore, seemed an appropriate language in which to test our hypothesis that syllabification would be a segmentation process specific to easily syllabified languages.

Experiment 1, in which English listeners heard English materials, showed, exactly

as predicted, no evidence that syllabification was being employed. Experiment 2, in which listeners from the same population heard French materials, added the further interesting finding that English listeners do not use syllabification even when the words they are listening to can be easily syllabified. Experiment 4, on the other hand, presented French listeners with English materials and showed that French listeners do employ syllabification even when some of the words they are listening to are hard to syllabify, and syllabification therefore fails on those words; with those words that were easy to syllabify (those with CVC initial syllable, e.g., *balcony*), the French showed a clear advantage for targets matching the initial syllable, an advantage which was not shown by the native English speakers.

Thus our results indicate that the segmentation processes characteristically employed by French speakers and English speakers differ: French speakers consistently make use of syllabification in segmentation; English speakers do not. We assume that this difference reflects the phonological differences between French and English which we outlined above. Furthermore, we assume that the effects are not specific to French and English, but that speakers of any language with clearly bounded regular syllables should show syllabification effects, while speakers of any language with irregular, hard-to-segment syllables should not.

Our results raise some very interesting issues for the study of segmentation processes in speech comprehension. As we noted earlier, the aim of psycholinguistics is the modeling of human language processing, not the processing of a particular language. Yet we have apparently established a language-specific speech segmentation strategy. This implies that the proper processing model for speakers of English (and languages like it) will differ from the proper model for speakers of French (and languages like it); a syllabification strategy

will be part of the latter but not the former. From the point of view of psycholinguistic theory, language-specific components in processing models are highly undesirable; if language specificity at this level is possible, why not dialect specificity or even speaker specificity? Or perhaps individual language users can switch processing modes according to the time of day, for instance? For this reason, we feel compelled to suggest a language-universal framework within which the present findings can be interpreted.

We speculate that syllabification is only one of a number of possible segmentation routines available to the human language processing device. During language acquisition, speakers adapt their perceptual routines so as to exploit with maximal efficiency the phonological properties of their native language. Effectively, they tend to favor some of the alternative segmentation routines over others, and incorporate them into their characteristic comprehension procedure. On this model, psycholinguistic theory would be called upon to enumerate the full range of possible segmentation routines and specify the conditions under which particular strategies are preferred to others; the general processing model would remain language-universal, with the language-specific variations being *predictable* from the phonological structure of each language.

What, then, are the alternatives to syllabification in speech segmentation? In particular, do we have any information about the segmentation routines employed by the nonsyllabifying native speakers of English? We suggest that certain aspects of our results will allow us to propose at least a tentative answer to this question.

Consider first the results of Experiment 4, in which French listeners heard English words. With CVC-syllabified words (*balcony*), they showed clear evidence of a processing advantage for targets which matched the initial syllable over targets which were shorter than the initial syllable.

But with the less easily syllabified English words like *balance* they showed no significant target effect. We argued that these subjects were applying syllabification where it was appropriate, that is, to those words which were easy to syllabify, and not where it would be inefficient, that is, to those words with ambisyllabic consonants. This argument alone, however, predicts that syllabification should be tried without success on the latter group of words, with the result that words like *balance* would be responded to slower by the French listeners. There was, however, no difference whatsoever in mean RT to the two types of word in this experiment. The absence of syllabic cues had no measurable effect on target detection time. This suggests that the listeners were quite efficiently segmenting the *balance* words, not with a syllabic routine since there was no significant difference between target types, but with some alternative routine employed in conjunction with syllabification. Whatever this alternative routine is, it is one which renders *balance* easier to segment than *balcony*. The syllabic routine works well with the *balcony* words, the alternative routine with the *balance* words (so that the net effect is no RT difference between the two word types).

Now consider the previously postponed question of the word-type effects found with English listeners. Recall that in Experiments 1 through 3 a small but highly consistent RT advantage was shown, irrespective of target type, for *balance*-type items over *balcony/balcon*-type items, whether English words, French words, or nonsense words. In other words, English listeners reliably found *balance* easier to segment than *balcony*. We tested and rejected several possible explanations of this effect. It could not be due to lexical factors such as frequency or word class since English listeners showed it even with French and with nonwords. It could not be an effect of item length since measured item length did not show a significant positive

correlation with RT in any of our four experiments, and additionally, the effect still held in Experiment 2 with French materials in which the *balance* words were consistently longer than the *balcon* words. Finally, a third explanation can also be ruled out on the basis of the Experiment 2 results. English is a stress language, and it could be that a preferred segmentation strategy for stress languages involves a preferred prosodic structure; words like *balance*, with an initial strong syllable followed by a weak syllable, might most closely match that preferred structure. However, this suggestion founders on the Experiment 2 results; French words have final accent, and all the CV-syllabified words in Experiment 2 had two full vowels, that is, from the point of view of English phonology, two strong syllables. The word-type effect for English listeners cannot be based on stress rhythm; otherwise it would not appear when these listeners were processing French.

Our explanation for the word-type effect is that it is again evidence of the application of an alternative segmentation routine. Moreover, this routine could even be, we suggest, the same alternative routine employed in conjunction with syllabification by the French listeners in Experiment 4. A possible characterization of this alternative routine is that the output of the segmenting device is simply a phonological representation of the incoming speech, without syllables or any other such intermediate representations. In any case, it appears that some sequences of phonemes may be intrinsically easier to deal with than others. Evidence from studies of speech perception in fact suggests that vowels are easier to identify if they are bounded by consonants (Strange, Verbrugge, Shankweiler, & Edman, 1976), and consonants are easier to identify if they are bounded by vowels (Liberman, Delattre, Cooper, & Gerstman, 1954). Thus alternating consonant-vowel patterns should lend themselves more readily to segmentation than nonalternating

strings. The speech materials in all of our experiments—English words, French words, nonsense words—contrasted item sets of which the first four phonemes were CVCV (*balance, balic*) with items sets of which the first four phonemes were CVCC (*balcony, balcon, balgart*). The former sets could have been facilitated, we claim, simply because their initial portions were easier to segment.

The postulated explanation would thus simultaneously account for both the consistent word-type effect in Experiments 1–3 and the failure of French listeners to experience difficulty with hard-to-syllabify words in Experiment 4. We conclude, therefore, that there is no single optimal segmentation strategy for the human language processor. Although the type of bottom-up information available to the language processor will be the same for speakers of all languages, there exist at least two, possibly many, segmentation routines which are applied to it, of which one is definitely a syllabic routine while another may be a phonemic routine. Moreover, it appears that speakers of some, possibly all, languages may have more than one routine available to them, and that speech segmentation may involve the application of more than one routine simultaneously.

Our final picture still has a few pieces missing. Our results suggest that English listeners are consistently segmenting in one way, whereas French listeners are employing more than one segmentation routine. To what extent are these procedures fixed? Could large amounts of pretraining with easily syllabifiable materials lead to application of a syllabic strategy by English listeners? On the other hand, could equivalent pretraining with hard-to-syllabify material force the abandonment of the procedure by French listeners? What segmentation processes are used by bilingual speakers who have equal facility in languages with regular and irregular syllable structure? Do such speakers vary their seg-

mentation routines as a function of the language they are listening to at a particular moment? We intend to explore these issues in further investigations. For the meantime, however, we believe the present results have shed considerable light on the question of segmentation processes in speech comprehension. At least some speakers appear to be able to use several segmentation procedures; there is no single optimal strategy of segmentation.

In conclusion, we would like to point to the usefulness of cross-language research of this kind to psycholinguistics. It is impossible to answer some psycholinguistic questions without considering different languages—particularly questions which involve the processing of linguistic features (such as syllables) which have widely varying structures across languages. A conclusive answer can only be obtained by comparing processing in languages with differing structures. But at simpler levels, too, cross-language research can be a valuable tool for the psycholinguist; for instance, to determine whether a particular effect is due to lexical factors or acoustic factors, it is often much easier to run the same experiment on a different language population (which should be impervious to lexical effects but susceptible to acoustic effects) than to design a further experiment. Psycholinguistics is a discipline with roots in two fields. Psycholinguists take for granted the resources of psychological expertise in modeling processes; they should equally take for granted knowledge about language structures.

REFERENCES

- ANDERSON, J., & JONES, C. (1974). Three theses concerning phonological representations. *Journal of Linguistics*, 10, 1–26.
- BROWN, R., & MCNEILL, D. (1966). The tip of the tongue phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 5, 325–337.
- CUTLER, A., MEHLER, J., NORRIS, D., & SEGUI, J. (1983). A language specific comprehension strategy. *Nature (London)*, 304, 159–160.
- CUTLER, A., & NORRIS, D. G. (1979). Monitoring sentence comprehension. In W. E. Cooper &

- E. C. T. Wakler (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*. Hillsdale, NJ: Erlbaum.
- FAY, D., & CUTLER, A. (1977) Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry*, 8, 505–520.
- FROMKIN, V. A. (1971) The non-anomalous nature of anomalous utterances. *Language*, 47, 27–52.
- KAHN, D. (1976) *Syllable-based Generalisations in English Phonology*. Doctoral dissertation, MIT, Cambridge, MA.
- LIBERMAN, A. M., DELATTRE, P. C., COOPER, F. S., & GERSTMAN, L. J. (1954) The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs*, 68, 1–13.
- LIBERMAN, I. Y., SHANKWEILER, D. P., FISHER, F. W., & CARTER, B. (1974) Reading and the awareness of linguistic segments. *Journal of Experimental Child Psychology*, 18, 201–212.
- MACKEY, D. G. (1972) The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology*, 3, 210–227.
- 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.
- MORAIS, J., CARY, L., ALEGRIA, J., & BERTELSON, P. (1979) Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, 7, 323–331.
- SAVIN, H. B., & BEVER, T. G. (1970) The non-perceptual reality of the phoneme. *Journal of Verbal Learning and Verbal Behavior*, 9, 295–302.
- SHERZER, J. (1982) Play languages: With a note on ritual languages. In L. K. Obler and L. Menn (Eds.), *Exceptional language and linguistics*. New York: Academic Press.
- SEGUI, J., FRAUENFELDER, U., & MEHLER, J. (1981) Phoneme monitoring, syllable monitoring and lexical access. *British Journal of Psychology*, 72, 471–477.
- TREIMAN, R. (1983) The structure of spoken syllables: Evidence from novel word games. *Cognition*, 15, 49–74.
- STRANGE, W., VERBRUGGE, R. R., SHANKWEILER, D. P., & EDMAN, T. R. (1976) Consonantal environment specifies vowel identity. *Journal of the Acoustical Society of America*, 60, 213–224.

(Received August 19, 1985)

(Revision received December 27, 1985)