JOURNAL OF MEMORY AND LANGUAGE 31, 18–32 (1992)

Contrasting Syllabic Effects in Catalan and Spanish

NÚRIA SEBASTIÁN-GALLÉS

Universitat de Barcelona, Barcelona, Spain

EMMANUEL DUPOUX

University of Arizona, Laboratoire de Science Cognitive et Psycholinguistique, CNRS-EHESS, Paris, France, and Ecole Nationale Superieure de Télécomunications, Paris, France

JUAN SEGUÍ

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

AND

JACQUES MEHLER

Laboratoire de Science Cognitive et Psycholinguistique, CNRS-EHESS, Paris, France

The role of syllabic structure and stress assignment in the perceptual segmentation of Catalan and Spanish words is studied. Previous research suggested that the syllable is the segmentation unit for languages with clear syllabic structure. In Experiment 1, we found that syllabification effects are found in Catalan but only in unstressed first syllable word-targets. No syllabification is obtained when the first syllable is stressed. In Experiment 2, we failed to find any syllabification effect in Spanish, regardless of stress in word-targets. Nonetheless, Experiment 3 shows that syllabification effects emerge in Spanish when subjects are made to respond to 250 ms slower than in Experiment 2. On the basis of these results, a modified version of the original syllabic hypothesis is proposed. We propose that both task demands and language specific parameters play a role in the presence or absence of syllabification effects in segment detection.

An important step in understanding speech perception is to discover how words are recognized. Many models of speech perception postulate that the acoustic signal maps directly onto word representations (Klatt, 1989). Other models however, posit that there is a prelexical code or intermediary level between the lexicon and the signal (Liberman, 1970; Elman & McClelland, 1986). Mehler (1981) and Mehler et al. (1990) have argued that the latter kinds of models are biologically more realistic.

Psycholinguists have focused much of their attention on two sublexical units: the phoneme and the syllable. Mehler, Dommergues, Frauenfelder, and Seguí (1981) found that French subjects respond faster in a syllable monitoring task when the target corresponds to the first syllable of the stimulus word than when it corresponds to a longer or shorter segment than the first syllable. That is, subjects respond faster when monitoring for BA in BALANCE than for BA in BALCON, and for BAL in

This research was supported by grants from CNET (Convention numéro 00 790 92 45 DIT), from CAYCIT (PB87-0601, GR88-0166, HF90, and SCI-54/88) and from HFSP program ("Contrasting language phonologies" project). The authors express their thanks to Jean Luc Aucouturier, Perfecto Herrera, Sergi Pérez, and Teodora Felguera for their technical assistance. Requests for reprints should be addressed to the first author at Dept. de Psicologia Bàsica, Universitat de Barcelona, Campus de Pedralbes, 08028 Barcelona, Spain.

0749-596X/92 $3.00

BALCON than in BALANCE. In other words, they found an interaction between word type and target type. From these studies, Mehler et al. (1981) concluded that the speech signal is segmented into syllables prior to lexical access. Cutler, Mehler, Norris, and Seguí (1983, 1986) tried to extend the generality of the original French experiments to English, but failed to replicate the main interaction of word by target type. Cutler et al. (1983, 1986) found an overall reaction time advantage for CV words over CVC words in their experiments with English.[1]

In order to explain the different response patterns in French and English, Cutler et al. (1986) proposed that for languages with clear syllabic boundaries, such as French, listeners spontaneously use a syllabic representation to segment speech prior to lexical access; while for languages with widespread ambisyllabicity, as in English, such a representation is not used. According to the authors, in these cases, subjects use a nonsyllabic unit whose detailed nature has yet to be uncovered. If these authors are right, the sublexical units used to process and recognize speech are not universal, but vary according to the phonological properties of the language.

In the Cutler et al. (1983, 1986) study, the aim was to contrast two types of languages, one with clear syllabic boundaries, and the other presenting ambisyllabicity.[2] Unfortunately, the characteristics of English and French make it difficult to select items where only the ambiguity of syllable boundaries is manipulated while all other properties are kept constant. In English, syllabic boundaries are primarily unclear in words that carry stress in the first syllable; while in French, a fixed stress language, stress is always carried by the last syllable in a word. As a result, in the Cutler et al. (1983, 1986) experiments, while French subjects monitored segments in unstressed syllables, English subjects found their targets in stressed syllables. This discrepancy makes the interpretation of the above cited experiments more difficult, since the French and English materials differ in more than just the cues to syllabic boundary. In fact, stressed and unstressed syllables differ widely in their acoustic characteristics. It could be argued that when the first syllable is unstressed, the segments are more coarticulated, forcing subjects to rely on a syllabic representation, even when they have to identify subparts of a syllable. In contrast, when the first syllable is stressed, the segments are more transparently realised and can be identified on the basis of a more peripheric phonetic code. By acoustic–phonetic transparency we refer to the ease with which a segment can be identified as distinct from another competing candidate in a language. Could it be that the contrasting behaviors of the French and English subjects observed by Cutler et al. arise from the nature of the stimuli used and not from the nature of the languages?

In fact, the disparity in the distribution of stress in the two sets of materials may not entirely explain the observed pattern of results. The authors found that the presence or absence of an interaction between targets and words types was not a function of the experimental list but rather a function of the subject's native language. Indeed, French subjects showed a strong syllabic effect, even when they were tested with the English materials. Similarly, English subjects did not display syllabic effects even when they were tested with the French material. The authors claim, therefore, that the results cannot be an artifact of the experimental material but are instead derived from perceptual strategies that subjects de-

[1] However, the C in all the CV words had an ambisyllabic C, so the CV words should be labelled CV[C] rather than plain CV.

[2] Stress languages, such as English, tend to have unclear syllable boundaries. In some cases, e.g., in English, intervocalic consonants that precede an unstressed vowel belong to two syllables at once (Anderson & Jones, 1974; Kahn, 1976). For example, in the English word *palace*, the /l/ may be part of the first and second syllables. Accordingly the syllabic structure of this word is [pa[l]ace].

velop to process their native language. Such strategies are then used even when listening to a foreign language. Still, the exact role of stress in segment monitoring remains to be explored in greater detail. To assess the role of stress in speech segmentation experiments one must study languages that allow first syllables to be either stressed or unstressed. This is why we chose Catalan and Spanish, a pair of languages that allow stress to appear in different parts of words.

In addition there are other reasons to try to replicate the Mehler et al. (1990) and Cutler et al. (1983, 1986) studies in other languages. If speech segmentation procedures vary from one language to another, it is important to establish how the language specific parameters constrain the segmentation procedures. Cutler et al. propose that the use of a syllabic representation is determined by the nature of syllable boundaries in the language. Languages with clear syllable boundaries always use syllables for segmentation, but languages with widespread ambisyllabicity do not. We might question such a strong claim, given that French and English differ in many aspects that are not directly related to differences in syllabic boundaries. For instance, in English, stress can have contrastive value (*con*tract and con*tract* are two different lexical items) whereas stress is fixed in French. The presence of contrastive stress in a language may turn out to be quite important for speech segmentation: for example, one might argue that in such languages lexical access cannot occur unless the stress pattern of the word being processed is available. The processing system might then need to include several syllables in its window to compute the relative value of stress prior to lexical access. In French, since stress is entirely predictable, the processing system might rely on single syllables as access units without paying attention to stress. Another dimension where English and French differ relates to the amount of vowel reduction. In English, vowel reduction[3] is widespread, whereas in French it is uncommon and limited to the vowel /e/. Again, the abundant presence of reduced vowels in a language may diminish the information content of single syllables. Thus larger units (such as feet) might be needed for reliable lexical access.

All these considerations make it necessary to further investigate the claim made by Cutler et al. True enough, English and French differ in the presence or absence of clear syllabic boundaries; but they also differ in the distributional properties of the main stress, the contrastive value of stress, and the amount of vowel reduction. Since all these factors could potentially account for the French–English results, studies using a different set of languages are clearly needed. In this paper, we explore the issue of speech segmentation with better controlled material using two relatively unexplored languages: Catalan and Spanish.

Catalan and Spanish have phonological systems that are somewhat intermediate between those of French and English (see Table 1) (Academia Española, 1973; Mascaró, 1978). Both languages, like French, have clear syllabic boundaries. Following Cutler et al.'s (1983, 1986) hypothesis, one would predict that Catalan *and* Spanish show a strong syllabic effect in segment detection. However, in other respects, Spanish and Catalan differ from French and are more like English. For instance, both have variable stress. This property will allow us to present words starting with stressed and unstressed syllables, and hence to study the role of stress value within a single language. We will also see if having variable stress affects the use of syllables in speech segmentation. Furthermore, Spanish has contrastive stress and no vocalic reduction whereas the reverse is true for Catalan. This contrast will allow us to sort out the

[3] Languages bearing vowel reduction do not have the same vowels in stressed and unstressed positions. In unstressed syllables, full vowels are replaced by reduced vowels (in English, [ɪ] and schwa.)

TABLE 1
COMPARISON OF SOME FRENCH, ENGLISH, SPANISH, AND CATALAN PHONOLOGICAL PARAMETERS

	French	English	Spanish	Catalan
Variable stress	—	+	+	+
Stress localization	Always final	Mostly initial	Mostly initial	Indifferent
Contrastive stress	—	+	+	—?
Ambisyllabicity	—	+	—	—
Vowel reduction	—	+	—	+
Number of vowels	16	14	5	8

relevant language dependent parameters that may modulate the use of syllabic representation in speech processing.

EXPERIMENT 1

Catalan has clear syllabic boundaries like French. Following Cutler et al.'s claim (1983, 1986), one might predict a syllabic effect (i.e., an interaction between word type and target type) in this language regardless of any other parameter. However, like English, Catalan has variable stress and widespread vowel reduction. If these parameters are the critical ones, then no syllabic effect should be obtained in Catalan. Of course, it might also turn out that syllabic effects depend on the position of stress in the experimental word. As mentioned above, syllabic effects might be more robust when the experimental word starts with an unstressed syllable rather than a stressed one. To test for this possibility, half of the experimental words were stressed on the first syllable and the other half on the second syllable.

Method

Subjects

Forty-four subjects whose first language was Catalan were used in this experiment; they were all undergraduate Psychology students at the University of Barcelona. Most of them, although dominant Catalan speakers[4] also spoke Spanish fluently.

[4] Catalan was their maternal language and they only learned Spanish when they were at school.

Material

Five quadruplets of bisyllabic words were selected. Each quadruplet was composed of words sharing the same initial CVC phonemes. The first syllable of two words in each quadruplet was stressed, while in the other two words the second syllable was stressed. The initial syllable of two words in each quadruplet had CV structure while the other two had CVC structure. In this way, each word in each quadruplet was a combination of ± stress and CV/CVC first syllable. We used the vowels /i/ and /u/, as they are the only ones which do not reduce when they appear in unstressed syllables. All words in the Catalan quadruplets started with one of the following CVC combinations: PUR, CUR, TIM, TIN, and VIS. Experimental words are shown in Table 2.

Two blocks, A and B, of 30 lists each were built. Lists ranged from one to six words in length. The targets always appeared in the last position of each list, except in the case of six word lists, where the target did not appear at all. Experimental words always appeared in three and four words long lists (eight and 12 lists, respectively, in each block). All members of each quadruplet always appeared in the same sequence position. In the remaining 10 lists of each block, lengths were distributed in the following way: two lists of one word, three lists of two words, three lists of five words, and two lists of six words. The same experimental words were used in the two blocks, but the target was changed from one block to the other. That is, if subjects had to mon-

TABLE 2
QUADRUPLETS USED IN EACH LANGUAGE

Spanish				Catalan			
CV		CVC		CV		CVC	
Stressed	Unstr.	Str.	Unstr.	Str.	Unstr.	Str.	Unstr.
Cala	Calor	Calma	Calcar	Cura	Curat	Cursi	Cursar
Cana	Canal	Canto	Candor	Pura	Puré	Purga	Purgant
Bala	Balón	Balsa	Balcón	Tina	Tinent	Tinta	Tinter
Paro	Pared	Pardo	Partir	Timid	Timó	Timbre	Timbal
Paso	Pasión	Pasta	Pastel	Viso	Visó	Vista	Viscós

itor PU in PURA in block A, they had to monitor PUR in PURA in block B. Thus, as in the Mehler et al. (1981) and Cutler et al. (1983, 1986) experiments, each stimulus word occurred twice (although preceded by different filler words) in the experiment. The two appearances of any experimental word shared the same position in each of the two blocks.

Fillers and nonexperimental targets were also bisyllabic words, half of them bearing stress on the first syllable and half of them on the second. Most of them were nouns, while the rest were verbs. None of the fillers started with the same consonant as the experimental item. A set of 10 training lists was constructed conforming to the experimental specifications.

Procedure

Lists were recorded by a female Spanish–Catalan bilingual speaker at a rate of one word every 2.0 s. The speaker read the items from a word-by-word computer screen presentation for a list of elements presented at 2-s intervals. The speaker did not know either what target corresponded to any list, or how many words each list contained. This was done in order to prevent the speaker from producing a fall in intonation in last position items. The experiment was run on each subject individually in a soundproof room. The experimental tape was binaurally presented via headphones. Targets were visually specified on small 8 × 5 cm cards. Before hearing each list, subjects heard two beeps and had 5 s to read the next card. Subjects were instructed to monitor for sound associated with the visually presented targets in the spoken word. They were told that in some sequences there would be no word carrying the specified target. Subjects were told to respond only if the specified target was detected in the auditory materials. At the onset of the experimental word an inaudible signal started a clock on an IBM PS-50 microcomputer. Subjects responded by pressing a response button which stopped the clock.

The experiment lasted about 20 min. Half of the subjects heard block A first and then block B, while the other half first heard block B, followed by block A.

Results

Mean reaction times for each subject and experimental condition were computed. Cutoff points were established at 100 ms and 1000 ms. Each subject's mean and standard deviation were computed. Reaction times above and below two standard deviations for each subject were excluded from the analyses. These procedures excluded less than 2.5% of the total responses. The mean response times for each condition are shown in Fig. 1. Overall analyses of variance were performed on means for subjects and items. These analyses showed that none of the main factors (stress, target type, and word type) introduced any significant effect. The interaction between target

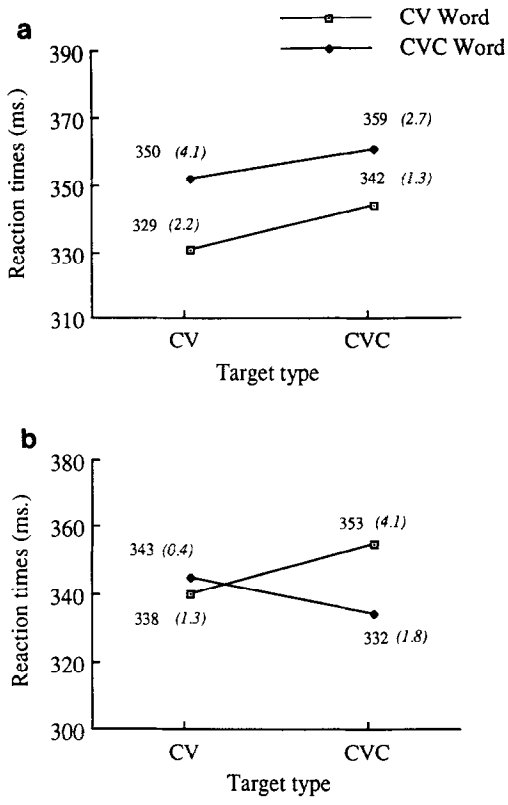

FIG. 1. Mean reaction times as a function of target size and word type in Catalan: (a) stressed segments; (b) unstressed segments. (In parenthesis, the percentage of missing rates of each condition.)

type and word type was only marginal in the subjects analysis ($F_1(1,43) = 3.63$, $p < 0.06$). No other interaction was significant. To explore in more detail the interaction between target and word type, we conducted planned comparisons for words whose first syllable was stressed with words whose first syllable was unstressed. For words with an initial unstressed syllable we found the interaction between word and target type was significant, both in the subject and item analysis ($F_1(1,43) = 4.95$, $p < 0.03$; $F_2(1,8) = 5.57$, $p < .04$). This interaction was due to the fact that for CV words, CV targets were detected faster than CVC targets (15 ms; $t(43) = 2.04$, $p < .04$), whereas a trend in the opposite direction was found for CVC words (11 ms; $t(43) = 1.41$, $p > .10$). Additionally, the word type factor introduced a significant effect in the subject analysis, $F_1(1,43) = 4.5$, $p < .03$ ($F_2 < 1$), CVC words were responded to faster than CV words.

In contrast, for words with an initial stressed syllable no significant interaction between word and target type was observed (all Fs < 1). However, the two main factors had a significant effect in the subject analysis: CV words were responded to faster than CVC words ($F_1(1,43) = 12.69$, $p < .0009$ but $F_2(1,8) = 1.6$, $p > .10$); CV targets were responded to faster than CVC targets ($F_1(1,43) = 4.84$, $p < .03$, but $F_2(1,8) = 2.37$, $p > .10$).

Discussion

The results of this experiment suggest that the position of stress in experimental words determines the subjects' response patterns. When the first syllable of the experimental word is unstressed, results are similar to those obtained for French; that is, a word by target interaction is obtained. In contrast, when the first syllable of the experimental word is stressed (something that cannot be studied in French) no such effect is observed.

The fact that a word by target interaction *can* be observed in the case of unstressed syllables shows that Catalan subjects are able to represent speech sounds in terms of syllables and use this representation to perform the segment detection task. This implies that the use of syllables can be demonstrated even in a language that has variable stress assignment and widespread vowel reduction. In this sense, the results are compatible with the hypothesis of Cutler et al. (1983, 1986) that the important parameter is the amount of ambisyllabicity in the language. Catalan has no ambisyllabicity and therefore syllables are used in speech segmentation.

However, the effect of syllables on segment detection times does not generalize to the whole language. No interaction has been found when words start with stressed syllables. This result does not fit at all with Cutler et al's claim. Indeed their claim pre-

supposes that the units that get used are the same across all the items in the language. Furthermore, there is no reason to suspect that stressed syllables would be more ambisyllabic than unstressed ones. Could it be thought that syllables are used when their vowel is unstressed, but not when it is stressed? Although this is certainly a logical possibility, it seems rather unlikely. From a processing point of view it seems strange to use one type of unit sometimes and an entirely different one at others. An alternative possibility is that a syllabic representation is always extracted from the signal regardless of stress, but subjects do not necessarily use it to perform the task.

This proposal is indeed quite compatible with other observations which show that in segment detection tasks fast subjects can bypass the syllabic representation, that is, to initiate their response before the underlying syllable has been fully processed (Miller & Dexter, 1988; Dupoux, 1989). In fact, subjects may have access to a stage in prelexical processing where information concerning the first part of the syllable is available. However, this subsyllabic processing stage is quite unstable, and when the response is not made fast enough, subjects can only have access to the full syllabic representation. It could then be that in the case of stressed syllables, the speech signal is quite strong and clear, and subjects can indeed bypass the syllabic representation. In the case of an unstressed syllable, however, the speech signal is much shorter and lower in intensity, making it difficult to access this subsyllabic representation.

In short, this interpretation posits that in Catalan, the speech stream is segmented into syllable-sized units, like in French. However, such a syllabic organization does not always gives rise to an interaction between word and target type. When the segments are more transparently realized in the signal, as in stressed syllables, subjects can bypass the syllabic organization and respond on the basis of a more peripheral representation. When they are realized in an unstressed environment, however, subjects use the full syllabic representation.

Given the data reported for Catalan, the study of Spanish is informative. Indeed, like Catalan and French, Spanish has clear syllabic boundaries. However, like English, Spanish has contrastive stress. Thus Spanish will allow us to determine whether the use of syllables is determined by the nature of the syllabic boundaries or by the presence of lexical items in the language that are identical in their phonemic content, except for the stress value.

EXPERIMENT 2

Spanish is a language that is ideally suited to discover whether the use of syllables arises in languages that separate one syllable from another by unambiguous boundaries or in languages that use lexical stress. Spanish is also interesting because, compared to the other languages that have been studied, it uses a smaller number of vowels. Indeed, Spanish has only five distinctive vowels. Moreover, in Spanish there is hardly any vowel reduction. These properties make Spanish words more phonetically transparent than either French, Catalan or English words. If the hypothesis formulated above to explain the absence of syllabification for Catalan words stressed in the first syllable is correct, one might predict the absence of a syllabic effect for all Spanish words. Indeed, given the limited number of vowels in the language and the fact that these vowels are never reduced, Spanish segments may be so distinctive that subjects might be able to bypass the syllabic representation in all cases.

However, previous experiments conducted in Spanish by Sánchez-Casas (1988) uncovered a strong tendency to syllabify in Spanish speakers. This result suggests that the above prediction is unwarranted. Still, in the Sánchez-Casas study only unstressed first syllable words were used. In view of the Catalan results, it was therefore important to carry out a new Spanish study controlling the position of stress. We still

would expect that the syllabic effect would disappear, or be greatly attenuated, for Spanish words with stress on the first syllable.

Method

Subjects

Thirty-six undergraduate psychology students at the University of Barcelona served as subjects; their first language was Spanish. Most of the subjects understood spoken Catalan.

Material and Procedure

The same specifications concerning stress and syllabic structure were used as in Experiment 1. However, in the Spanish quadruplets it was possible to use the same vowel, /a/ (as in the Mehler et al. (1981) and Cutler et al. (1983, 1986) experiments). All words in the Spanish quadruplets started with one of the following CVC combinations: CAL, CAN, BAL, PAR, or PAS. Experimental words are shown in Table 2.

Results

Mean reaction times for each subject and experimental conditions were computed. Cutoff points were established at 100 ms and 1000 ms. Each subject's mean and standard deviations were computed. Reaction times above and below two standard deviations for each subject were excluded from the analyses. These criteria resulted in the exclusion of less than 2.5% of the responses. The mean response times for each condition are shown in Fig. 2. Overall analyses of variance were performed on subjects' and items' means. These analyses showed that both the main effects of stress and word type were statistically significant in the subjects' analysis, although they failed to reach significance in the items' analysis ($F_1(1,35) = 7.25, p < .01, F_2(1,16) = 1.39, p > .10$ for the stress factor, and $F_1(1,35) = 7.67, p < .01$ and $F_2(1,16) = 1.25, p > .10$ for the word type factor).

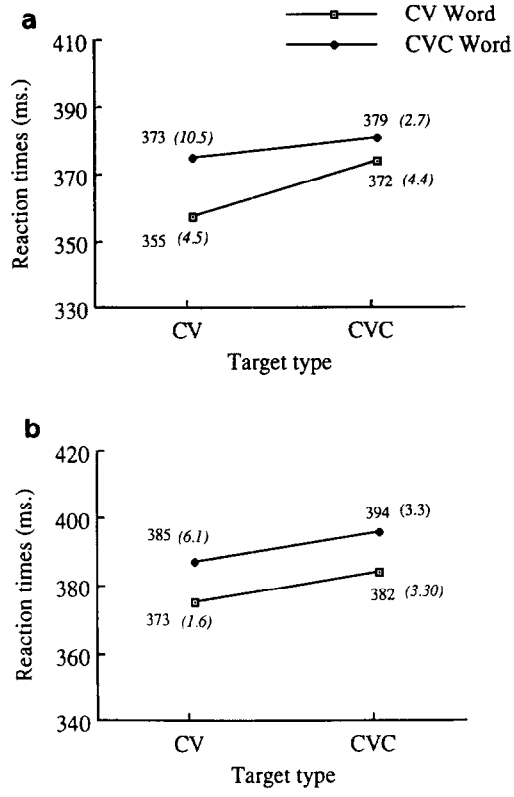

FIG. 2. Mean reaction times as a function of target size and word type in Spanish: (a) stressed segments; (b) unstressed segments. (In parenthesis, the percentage of missing rates of each condition.)

Stressed segments were responded to faster than unstressed ones (369 vs 383 ms) and CV words were responded to faster than CVC words (369 vs 382 ms). The target type introduced a marginal effect in the subjects analysis ($F_1(1,35) = 3.4, p < .07$), although it was significant in the item analysis $F_2(1,16) = 5.31, p < .03$, with CV targets being responded to faster than CVC targets (371 vs 381 ms). No interaction reached significance levels in any analysis. Planned comparisons were conducted for words with initial stressed and unstressed syllables. In both sets of analyses, the only marginally significant factor was the type of word, and only in the subjects analysis ($F_1(1,35) = 3.91, p < .055$ for stressed segments, and $F_2(1,35) = 3.53, p < .06$ for unstressed segments).

Discussion

In this second experiment, words with stress on their first syllable ellicited faster responses than words whose first syllable was unstressed. In addition, responses to CV words were faster than to CVC words, and CV targets were detected faster than CVC targets. But, in contrast with previous reports by Sánchez-Casas, no interaction between word type and target type was found for either kind of stress position.

The global effect of stress in Spanish is interesting given that no such effect was found in Catalan. It might be possible to attribute this effect to low level factors. Stressed syllables convey more energy and have larger formant excursion, yielding more information than unstressed syllables. An alternative explanation, might be derived from the fact that in Spanish, most bisyllabic words start with stressed syllables. Thus conceivably Spanish listeners might give priority to stressed versus unstressed syllables (see Cutler & Norris, 1988, for a similar proposal in English).[5]

The main result of this experiment is the complete lack of interaction between word and target type. This result is at variance with the Cutler et al. (1983, 1986) claim that a syllabic effect should be found in a language with clear syllabic boundaries. This result is also at variance with the hypothesis we formulated in Experiment 2, namely, that a larger syllabic effect should be found in unstressed syllables because they are less transparent. In fact, no trace of any syllabic effect was found, irrespective of the stress pattern. Finally, and more problematic, these results are at variance with the results of Sánchez-Casas's (1988) who did find a significant syllabic effect in Spanish.

How is it that different results are obtained within the same language? It must be remembered that in Experiment 1 we stated that the absence of syllabic effects does not imply that a syllabic representation is not extracted from the signal. It just means that subjects did not use this representation to perform the task. There are a number of factors that may induce subjects not to use a full syllabic representation in a speeded task, even though the processing system extracts it. One of these factors is *speed*. In a variety of experimental situations, it has been reported that the processing levels that a segment detection or classification task can tap are radically different for fast and slow responses (see, for instance, Fox, 1984; Miller & Dexter, 1988; Dupoux, 1989). In particular, Dupoux (in preparation) argues that in segment detection tasks, subjects are more likely to respond before the syllable has been wholly identified when their latencies are short. In contrast, when the latencies are long, responses are much more likely to reflect influences from a syllabic level of processing. As a matter of fact, the average speed of response between the two Spanish experiments is very different. In our Experiment 2 the mean reaction time was 376 ms, while in the Sánchez-Casas' experiment it was 636 ms. Such a difference is large enough that it ought to be taken into consideration. One possible account for the large differences in reaction times between the two experiments could be due to the fact that in Sánchez-Casas' experiments there were catch trials, while we did not use such sequences. That is, in Sánchez-Casas' material, there were sequences from which the specified target had either the first consonant of the list-final non-experimental word (e.g. cu/colina), or its first vowel (tar/madera); "with the same purpose some filler words in the experimental lists shared either a vowel or a consonant with the target." Under these conditions, Sánchez-Casas' subjects could have delayed their responses until a higher-level unit (such as

[5] Notice that, in Catalan, there is an equal distribution of stressed and unstressed first syllables in bisyllabic words. In fact, no main effect of stress was significant in this language.

the syllable) became available. In contrast, our subjects were never presented with catch trials, and thus, they might have based their responses on subsyllabic units.[6]

In brief, the discrepancy between the results obtained by Sánchez-Casas and our Experiment 2 could very well be due to task demands caused by differences in the material. In Sánchez-Casas's experiment, reaction times were quite long, and subjects had the possibility of using a full syllabic representation to perform the task. In contrast, in our experiment, reaction times were much faster and subjects might have responded on the basis of an incomplete syllabic representation, thus yielding no target by word type interaction. In other words, the two results are not incompatible, but may result from tapping two separate stages in a single processing system. Thus, before going into any further interpretation of the Spanish results, another experiment is necessary, using the same procedure as in Experiment 2, but trying to slow down subjects' responses.

EXPERIMENT 3

Experiment 3 was designed to test the hypothesis that the average speed of response influences the pattern of data in syllable monitoring. Specifically, the prediction is that if subjects are made to respond with relatively long latencies, a syllabic effect should reemerge in Spanish.

In order to get slower reaction times, subjects' instructions were modified in the present experiment. Following Dupoux (1989), subjects were told to pay attention to the semantic relations that could arise between words in the same list. Of course, our lists had been designed not to allow evident semantic relations between words. Still, subjects were told that at the end of the experiment they would be given a recognition test of some of the words presented in the experiment. This manipulation turns out to be quite effective in slowing down reaction times. Dupoux (1989) noted up to 200-ms increase in reaction time using otherwise identical experimental materials and procedure.

Method

Subjects

Thirty-six undergraduate psychology students at the University of Barcelona served as subjects. They conformed to exactly the same characteristics of subjects in the previous experiment. None of them were subjects in the previous experiments.

Materials and Procedure

The materials and general procedure employed were exactly the same as in the previous experiment. Subjects' instructions were modified in order to get slower reaction times. Subjects were told to pay attention to the words in the lists, since they were going to be asked to perform a recognition test on them at the end of the experiment. They were also told that in order to make the recognition test easier, sometimes words within a list could be semantically related.

Results

Mean reaction times for each subject and experimental conditions were computed. Cutoff points were established at 300 ms and 1300 ms. Each subject's mean and standard deviation were computed. Reaction times above and below two standard deviations were excluded from the analyses. These criteria resulted in the exclusion of less than 5% of the responses. The mean response times for each condition are shown in Fig. 3. Overall analyses of variance were performed on subjects' and items' means. These analyses showed that the interaction between target type and

[6] The longer reaction times in the CVC target condition compared with the CV target one found in our experiment prevents us from concluding that subjects did not follow the instructions and that, since there were no catch trials, they only monitored for the first C or CV phonemes of each target.

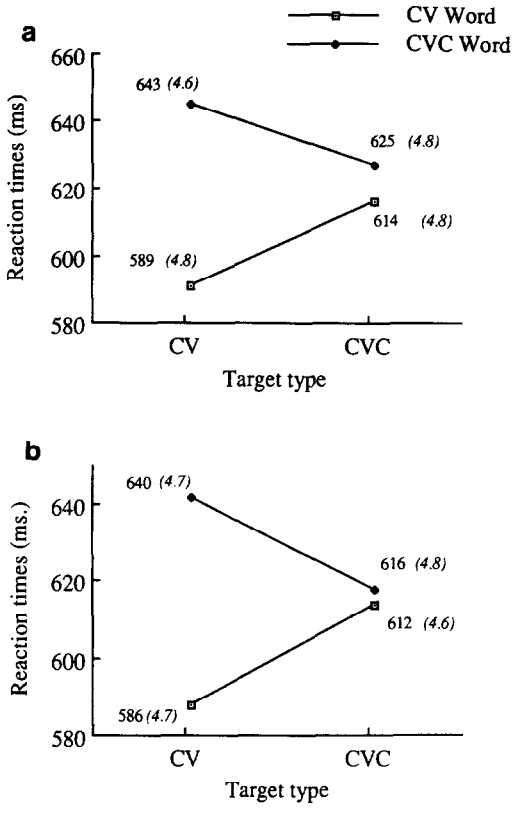

FIG. 3. Mean reaction times as a function of target size and word type in Spanish: (a) stressed segments; (b) unstressed segments. (In parenthesis, the percentage of missing rates of each condition.)

word type reached significance levels both in the subjects' and items' analyses ($F_1(1,35) = 12.51, p < 0.001$ and $F_2(1,16) = 6.80, p < 0.02$.) The type of word factor also reached significance levels both in the subjects' and items' analyses ($F_1(1,35) = 27.62, p < 0.0001$ and $F_2(1,16) = 5.82, p < 0.03$.) No other main factor, or interaction reached significance levels either in the subjects' or items' analysis.

As in the previous experiments, separate analyses of variance were performed for both initial stressed and unstressed segments. For words with initial unstressed syllable we found that the interaction between type of word and target type was significant both in the items' and subjects' analyses ($F_1(1,35) = 10.08, p < 0.003$ and $F_2(1,8) = 5.78, p < 0.04$.) The type of word factor was significant in the subjects' analysis, although it did not reach significance levels in the items' analysis ($F_1(1,35) = 14.83, p < 0.0005$ and $F_2(1,8) = 2.26, p > 0.10$). For words with initial stressed syllables, we found that the interaction between type of word and target type was only significant in the subjects' analysis ($F_1(1,35) = 6.39, p < 0.01$ and $F_2(1,8) = 2.32, p > 0.16$). The type of word factor was significant in the subjects' analysis, but only marginally significant in the items' analysis ($F_1(1,35) = 17.46, p < 0.0002$ and $F_2(1,8) = 4.36, p < 0.07$).

Discussion

In this experiment, reaction times were about 250 ms longer than in Experiment 2. As in Sánchez-Casas' (1988) experiment a strong interaction between word type and target type was uncovered. Moreover, the pattern of the interaction did not differ significantly between stressed and unstressed segments.

The upshot of this experiment is that changing the characteristics of the task can have a drastic impact on the pattern of result obtained in segment detection tasks. The experimental materials were identical to the ones used in Experiment 2; the only difference was the introduction of a secondary task. This secondary task could be seen as relatively innocuous and nondemanding, since subjects were only asked to recognize a few words at the very end of the session. Yet this modification introduced a powerful effect on average reaction time and a qualitative change in the pattern of results.

To assess that the task change introduced a qualitative change and that the results are not just due to random subject fluctuation between Experiments 2 and 3, we ran a conjoint analysis of variance by declaring task as a between subject factor. There was a three way interaction between word type, target type, and task ($F_1(1,70) = 5.56, p < .02$; $F_2(1,16) = 4.26, p < 0.05$). This interaction showed that the task of Experiment 3 significantly induced subjects to behave

more syllabically than the task of Experiment 2. The task factor also showed a reliable interaction with the stress factor in the subject analysis ($F_1(1,70) = 6.70$, $p < .01$; but $F_2(1,16) = 2.90$ p > .1). This interaction was due to the fact that stressed first words were responded to faster than unstressed first words in Experiment 2, but no such effect was found in Experiment 3.

As previously mentioned, the present results replicate previous findings of Sánchez-Casas et al. (1988) and they also support the explanation previously given of the apparent discrepancies between data from our Experiment 2 and Sánchez-Casas' in terms of tasks differences, tapping different stages of the processing system. Of course, this explanation suggests that the results of syllable detection experiments should be interpreted with great caution when it comes to inferring the properties of the processing system across languages. In fact, different results across various languages could be due to task differences rather than bona fide processing differences. Could such effects explain the difference observed between Catalan and Spanish (Experiments 1 and 2)? The response is no. In our Catalan and Spanish experiments we used identical methods and procedures in both languages, and obtained very similar latencies. Nonetheless, Spanish responses (Experiment 2) showed no trace of syllabification while the Catalan responses did (Experiment 1). In fact, Catalan responses were even slightly faster than Spanish responses; thus, if anything, we would expect less syllabification in Catalan than in Spanish. Accordingly, the results obtained in our first two studies cannot be accounted for in terms of differences in response times or method. They have to be accounted for in terms of language-specific processing characteristics.

General Discussion

In Experiment 1, we found a syllabic effect similar to that reported for French (Mehler et al., 1981), when we used Catalan words with unstressed first syllables. In contrast, no syllabic effect was observed for words whose initial syllable was stressed. This led us to speculate that the acoustic phonetic transparency of the target may enable subjects to respond without showing a syllabic effect, i.e., a response can be triggered before having been influenced by full syllabic analysis. These results induced us to study a language that has clear syllabic structure, but considerable acoustic phonetic transparency. Thus, in Experiment 2 we tested Spanish subjects with Spanish words. In that experiment we failed to find a syllabic effect regardless of whether the first syllable in word targets was stressed or unstressed. These results contrasted with previous ones in which syllabic effects have been obtained in Spanish. However, there was a large difference in reaction times between our results and the previous ones. In Experiment 3 subjects' instructions were modified in order to slow down subjects' responses. The results obtained showed an overall syllabic effect.

The fact that syllabic effects *can* be observed in Catalan and Spanish show that syllables play a role in speech processing in these languages. It can be argued that, as in French or Portuguese, speakers spontaneously tend to represent and segment speech signals in terms of syllables. By and large, the original hypothesis by Cutler et al. (1983, 1986) has been sustained.

However, the present study shows that the picture is clearly more complex than previous investigators have proposed. Although syllabic effects can be obtained, subjects can also respond without taking into account the full syllabic structure of the utterance. The ease with which subjects have access to this "subsyllabic" stage depends on many parameters, the most obvious one being task characteristics. As we have shown in Experiments 2 and 3, syllabic effects are very likely to disappear if subjects are made to respond extremely fast. In fact, subjects behave as if there were a stage in processing, where informa-

tion concerning parts of the syllable is accessible to trigger a response. This subsyllabic stage, however, is transient in nature, so that only extremely fast responses can pick it up. Slow subjects do not have access to this early stage and respond on the basis of the full syllabic representation (Dupoux, 1989).

This shows that extreme caution is to be taken when segment detection tasks are used to investigate speech processing. Task demand, however, is not the only factor playing a role. More importantly, there also seem to be factors that depend on language characteristics. The case of Catalan illustrates that subsyllabic responses are more likely to be obtained when the target occurs in a stressed syllable. In unstressed syllables, however, a significant impact of a syllabic representation was uncovered. Stress remains a very important parameter to be kept in mind when studying speech processing. However, this cannot explain the full picture. For instance, in Experiment 2 in Spanish no syllabic effect was found in unstressed syllables. In fact, the stress value of syllables seems to play a minor role in Spanish, at least, with respect to syllabification effects. We argued that Spanish's results were due to the fact that this language only uses a restricted set of highly distinctive vowels. As a result, even unstressed segments are quite transparently realized and can promptly be categorized without the need to analyze a larger representation.

In brief, Spanish and Catalan use syllables for speech segmentation, but subjects' responses can be unaffected by this process by either of two conditions: when the syllable is stressed, or when the language uses a very limited number of highly distinctive vowels. We can try to somewhat simplify this statement by replacing the disjunctive condition by a single notion that we alluded to in the introduction: "acoustic transparency." This notion refers to the ease with which a system might quickly come to a decision concerning the various segments present in the signal. Research in vowel perception has shown that subjects use both coarticulation information at the onset and at the coda of CVC items to determine the nature of the vowel (see Strange, 1989a,1989b; Strange et al., 1983). Thus, it might be that in languages with very few vowel types, coarticulation at the onset of the syllable is sufficient to allow for the unambiguous determination of the nature of the vowel. In contrast, in languages where the vocalic space is filled with fine-grained distinctions, the perceptual system may have to integrate information across the whole syllable to determine the identity of the vowel. Accordingly, in languages with a large number of vowels, subjects may rely on a syllabic representation, especially if the target occurs in an unstressed syllable.

We should highlight the fact that acoustic–phonetic information is more redundant in stressed than unstressed syllables. Presumably, stressed syllables carry more energy and transmit a very clear stimulus to the perceptual system compared to their unstressed counterparts. Thus stressed syllables tend to be more transparent than unstressed ones. To add some twist to this argument, languages with a large number of vowels usually allow for reduced vowels in unstressed position. Such reduced segments, and specifically the schwa, are characterized by a high degree of ambiguity with respect to their acoustic–phonetic identity. Languages allowing reduced vowels are thus acoustically less transparent and native speakers of such languages probably develop specific strategies to cope with this. Spanish has only five vowels and does not allow for vowel reduction. Of the above mentioned languages, Spanish is acoustically the most transparent of all. In a segment monitoring task, subjects may then be able to bypass the syllabic representation, knowing that no ambiguity will arise, even in unstressed syllabic contexts. Catalan, in contrast, has eight vowels and vowel reduction is widespread. In French, although vowel reduction is exceptional,

the number of different vowels is rather large (14). Thus, both languages can be said to be less transparent than Spanish, and the subjects are reluctant to abandon a syllabic level of processing to perform a segment detection task. Unsurprisingly, in unstressed syllables a syllabic effect is obtained in both languages.

This interpretation accounts nicely for languages with clear syllabic structure as French, Spanish, and Catalan. It falls short, however, to account for the English case. English has 14 different vowels and widespread vowel reduction. According to our standards, English is then the least "transparent" language that has been tested so far. One should then expect a strong syllabic effect in this language. So far no such effects have been reported, although only stressed syllables have been tested. It might turn out that when unstressed syllables are used, a syllabic effect emerges. It might also turn out that no syllabic effect is ever found in English because, as has been proposed by Cutler et al. (1983, 1986), this language uses a different type of unit. At any rate, subsequent studies ought to take into account the two powerful variables that we have uncovered in Experiments 1, 2, and 3: stress value and average speed of response. Unless these factors are carefully controlled, cross linguistic comparisons of speech segmentation strategies are going to be impossible.

In this paper, we primarily discussed the presence or absence of an interaction between target type and word type in different languages. However, segment detection tasks can potentially provide far richer data that ought to be taken into consideration in future studies. For instance, we have noted that in Experiment 2, Spanish stressed syllables yield faster reaction time than unstressed syllables. Such an effect, however, was not found either in Catalan, nor in Experiment 3, where Spanish subjects were slowed down. So effects of stress seem to go in complementary distribution with syllabic effects. Further experiments are clearly needed to sort out this phenomenon and to understand in more detail the role of stress in speech segmentation. Another puzzling result relates to the data obtained for CV targets: we have found throughout the experiments that CV words yield faster detection times than CVC words. This effect is similar to the syllabic complexity effect obtained in phoneme monitoring: the more complex the syllable, the longer it takes to respond to an initial phoneme (see Seguí et al., 1990). This effect, however, seems to be obtained, even in the case where no syllabic effect is found. The fact that the syllabic structure of the utterance can influence the detection time of CV sequences is interesting in itself. If true, this would mean that coarticulatory/allophonic information relevant to the structure of the syllable is available to the early processing system even though subjects may not use a full syllabic representation to initiate their responses.

Finally, more research is needed to examine whether the notion of "acoustic transparency" is a property of the signal or stems from the listener's perceptual system. For instance, it would be important to know whether monolingual Catalan or Spanish subjects syllabify a nonnative language. It would also be interesting to use this technique to study tone languages like Mandarin Chinese, which has very few vowels, but whose tonal characteristics force the perceptual system to make a great many fine-grained distinctions. At any rate, our results emphasize the fact that future research into speech processing should explore different language types to determine the effect of syllabic structure, stress distribution, and other phonological properties on lexical access. Such models will also clarify what the child must extract from his/her linguistic environment in order to construct relevant speech processing routines.

REFERENCES

ACADEMIA ESPAÑOLA (1973). *Esbozo de una nueva gramática de la lengua española.* Madrid: Espasa-Calpe.

ANDERSON, J., & JONES, C. (1974). Three thesis con-

cerning phonological representations. *Journal of Linguistics*, **10**, 1–26.

CUTLER, A. (1990). Exploiting prosodic probabilities in speech segmentation. In G. Altman (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. Cambridge, MA: MIT Press.

CUTLER, A., MEHLER, J., NORRIS, D., & SEGUÍ, J. (1983). A language specific comprehension strategy. *Nature (London)*, **304**, 159–160.

CUTLER, A., MEHLER, J., NORRIS, D., & SEGUÍ, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, **25**, 385–400.

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

DUPOUX, E. (1989). *Identification des mots parlés: Détection de phonème et unité prelexicale*. Unpublished doctoral dissertation, EHESS, Paris.

DUPOUX, E. (in preparation). Prelexical processing: The syllabic hypothesis revisited. Submitted to G. Altmann and R. Shillcock (Eds.) *Cognitive Models of Speech Processing*.

ELMAN, J. L., & MCCLELLAND, J. L. (1986). Exploiting the lawful variability in the speech wave. in J. Perkell and D. Klatt. (Eds.). *Invariance and variability in speech processes*. Hillsdale, NJ: Erlbaum.

FOX, R. A. (1984). Effects of lexical status on phonetic categorisation. *Journal of Experimental Psychology: Human Perception and Performance*, **10**, 526–540.

JENKINS, J. J., STRANGE, W., & EDMAN, T. R. (1983). Identification of vowels in vowelless' syllables. *Perception and Performance*, **34**, 441–450.

KAHN, D. (1976). *Syllable based generalisations in English phonology. Doctoral dissertation*, MIT, Cambridge, Mass.

KLATT, D. H. (1989). Review of selected models of speech perception. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process*. Cambridge, MA: MIT Press.

LIBERMAN, A. M. (1970). The grammars of speech and language. *Cognitive Psychology*, **1**, 301–323.

MASCARÓ, J. (1978). *Catalan phonology and the phonological cycle*. Bloomington, IN: Indiana University Club.

MEHLER, J. (1981). The role of syllables in speech processing: Infant and adult data. *Philosophical Transactions of the Royal Society, London, Ser. B*, **295**, 333–352.

MEHLER, J., DOMMERGUES, J., FRAUENFELDER, U., & SEGUÍ, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, **20**, 298–305.

MEHLER, J., DUPOUX, E., & SEGUÍ, J. (1990). Constraining models of lexical access: The onset of word recognition. In G. Altman (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. Cambridge, MA: MIT Press.

MILLER, J. L., & DEXTER, E. R. (1988). Effects of speaking rate and lexical status on phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 369–378.

SÁNCHEZ-CASAS, R. M. (1988). *Access representation in visual word recognition*. Unpublished Doctoral dissertation. Monash University, Melbourne, Australia.

SEGUÍ, J., DUPOUX, E., & MEHLER, J. (1990). The role of the syllable in speech segmentation, phoneme identification and lexical access. In G. Altman (Ed.). *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. Cambridge, MA: MIT Press.

STRANGE, W. (1989a). Evolving theories of vowel perception. *Journal of the Acoustic Society of America*, **85**, 2081–2087.

STRANGE, W. (1989b). Dynamic specification of coarticulated vowels in sentence context. *Journal of the Acoustic Society of America*, **85**, 2135–2153.

STRANGE, W., JENKINS, J. J., & JOHNSON, T. L. (1983). Dynamic specification of coarticulated vowels. *Journal of the Acoustic Society of America*, **74**, 695–705.

(Received January 7, 1991)
(Revision received May 15, 1991)