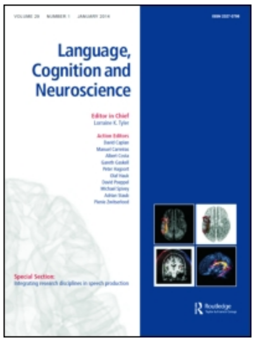

Language and Cognitive Processes

ISSN: 0169-0965 (Print) 1464-0732 (Online) Journal homepage: https://www.tandfonline.com/loi/plcp20

The status of the syllable in the perception of spanish and englisb

Dianne C. Bradley , Rosa M. Sánchez-Casas & José E. García-Albea

To cite this article: Dianne C. Bradley , Rosa M. Sánchez-Casas & José E. García-Albea (1993) The status of the syllable in the perception of spanish and englisb, Language and Cognitive Processes, 8:2, 197-233, DOI: 10.1080/01690969308406954

To link to this article: https://doi.org/10.1080/01690969308406954

LANGUAGE AND COGNITIVE PROCESSES, 1993, *8* (2), 197–233

The Status of the Syllable in the Perception of Spanish and English

Dianne C. Bradley, Rosa M. Sánchez-Casas
Department of Psychology, Monash University, Clayton, Victoria, Australia

José E. García-Albea
Universidad Complutense de Madrid, Madrid, Spain

A series of monitoring studies is reported, in replication of the cross-language research of Cutler, Mehler, Norris and Seguí (1983; 1986), which found evidence of language-specific perceptual routines. Monolingual speakers of Spanish and English detected CV and CVC target sequences in native and non-native materials. The replication succeeded only in the case of Spanish speakers and Spanish materials, where a cross-over interaction of target (CV *vs* CVC sequences) and carrier types (CV- *vs* CVC-syllabified words) gave evidence of a sensitivity to the input's syllabification; no such pattern emerged for Spanish speakers and English materials, nor for English speakers and materials in either language. For English speakers, the consistent finding was for faster performance with CVC targets, regardless of the structure of the carrier word. Whether or not this is to be interpreted as evidence of syllabified input representations is not clear. Analyses of English syllabification that are alternatives to that adopted by Cutler et al. exist, to weaken the original contrast drawn between syllable-favouring and syllable-disfavouring languages.

A final experiment examines monitoring performance in Spanish speakers who have become bilingual as a consequence of emigration to an English-speaking country; these subjects showed no syllable sensitivity for Spanish

Requests for reprints should be addressed to Dianne Bradley, Department of Psychology, Monash University, Clayton, Victoria 3168, Australia.

This research was supported in part by a grant from the Australian Research Grants Committee to the first author, and the experiments reported formed part of a doctoral thesis submitted to Monash University by the second author, who was supported by a Monash Graduate Scholarship. Rosa Sánchez-Casas is now at St Louis University (Madrid Campus). We wish to thank two anonymous reviewers for sharpening our appreciation of the controversy surrounding the analysis of English syllabification.

language materials. We speculate that factors outside the perceptual system may determine the basis on which responses are made in the monitoring task, and therefore conclude that the case for language specificity in perceptual routines has yet to be made.

INTRODUCTION

In speech recognition, given a physically continuous and richly detailed signal but a message constituted of discrete word types, we are confronted with a problem. What representational format mediates the mapping between waveform and lexicon? Traditional acoustic phonetics seems to have construed the speech recognition task as one involving the initial recovery of a phonetic transcription, and in pitching its perceptual routines at the level of phones has made evident the "lack of invariance" problem, the apparent unavailability of consistent acoustic cues to phonetic identity which is the outcome of co-articulation. Klatt's (1979) proposal abandons altogether the idea of a mediating representation which is at all abstract, and moves directly from acoustic feature matches to lexical identification via a network of spectral templates. Between these extremes lies the possibility that the speech recogniser initially seeks something larger than the phone but smaller than the word, and one candidate immediately suggesting itself is the syllable. This paper evaluates the evidence for a perceptual routine sensitive to proper syllabification, and explores the circumstances under which representations couched in the vocabulary of syllables might be exploited.

The standard argument in favour of syllable-based perceptual routines says, straightforwardly, that a recognition system might trade off the size of its representational inventory against control of variability in the acoustic realisations of speech gestures—and finds grounds for suggesting that the syllable is a "unit of perception" in the observation that the impact of co-articulation is less between syllables than within (e.g. Liberman & Studdert-Kennedy, 1978). However, early work highlighted the difficulty of experimentally evaluating claims about the perceptual status of one level of analysis versus another (see Mehler, Seguí & Frauenfelder, 1981b, for a comprehensive review). Savin and Bever (1970), for example, sought to make arguments based on the relative speed with which subjects could detect phoneme or syllable targets in sequences of nonsense syllables. Their finding that reaction times were faster for syllable-sized targets may indicate a primary status of the syllable in perception, and the extraction of phonemes from syllables as a second step, or it may not. Syllables may simply be more available to support detection responses than phonemes for reasons that have little to do with the time-course of perceptual processes. A good deal of work has established the naturalness of the concept of a syllable: Preliterate children (e.g. Liberman, Shankweiler, Fischer & Carter, 1974) and illiterate adults (e.g. Morais, Cary, Alegria & Bertelson, 1979), while generally unable to perform tasks requiring phonemic analysis, fare considerably better in dealing with syllables. But while performance requiring phonemic and even syllabic analyses is enhanced by literacy (Morais, Cluytens & Alegria, 1984; Morais, Bertelson, Cary & Alegria, 1986), there is no reason to suppose that the changes in metalinguistic skill that follow reading acquisition have any parallel in embedded perceptual routines, and the differences observed may reflect only what is most available to conscious inspection. It may also be the case that any larger unit (like the syllable), even if constructed from a sequence of smaller elements (like phonemes), permits greater certainty about the match between a monitoring target and its occurrence in an utterance.

An alternative means of demonstrating the relevance of syllabic analyses in speech recognition avoids these difficulties by using targets of a fixed type, and looking instead at whether monitoring performance reflects the syllabic status of the target's occurrence in a stimulus utterance. Thus in the study of Mehler, Dommergues, Frauenfelder and Seguí (1981a) with French-speaking subjects, consonant–vowel (CV) or consonant–vowel–consonant (CVC) targets sometimes coincided with the proper syllabic decomposition of a carrier word (e.g. *pa* in *pa'lace* or *pal* in *pal'mier*), but in the contrasting cases offered more or less than the initial syllable (e.g. *pal* in *pa'lace* or *pa* in *pal'mier*). The finding was that, indeed, monitoring responses were faster for targets corresponding to true syllables. Mehler et al. argued, plausibly, that an account of these data in terms of the match or mismatch between a subject's expectancy and the actual stimulus (cf. Mills, 1980) could be offered only if the mental representation of the stimulus was one that displayed its syllabification, and went further to suggest that this syllabically based representation was the code mediating access to the lexicon. A variety of evidence is compatible with the claim that monitoring performance in these circumstances does not require information represented lexically: Sheer speed of response in Mehler and co-workers' study suggested a pre-lexical basis of performance, and an associated study (Mehler et al., 1981b) found no evidence of lexicality effects when monitoring was examined across word and nonword carriers. Perhaps most striking are the data of Cutler, Mehler, Norris and Seguí (1983; 1986). When French speakers monitored for CV and CVC targets in lists of *English* words, and contrasts were set up as in the original study with French materials, the reaction time patterns were sufficiently similar to those found earlier to suggest to those researchers that a syllabified representation of spoken stimuli can be achieved without lexical support.

Does this body of evidence necessarily mean, as Cutler et al. (1983) would have it, that "in language perception ... incoming words are processed syllable by syllable" (p. 159)? The disquietening possibility must be that syllabic analyses are made available or given emphasis, whether pre- or post-lexically, precisely because the task demands it. If CV or CVC target sequences presented in isolation have their most natural mental representation as syllables rather than phoneme strings—to the extent that /p/-plus-/a /-plus-/l/ is unsayable, it is plausibly also unthinkable—it might be argued that a syllabified stimulus representation is no more than the code through which matches must be sought in this version of the monitoring task, so that a "syllabification effect" is the inevitable outcome. But this cannot be so. In parallel studies with speakers of English, Cutler et al. reported that monitoring was *not* sensitive to syllabic structure. Their English subjects, like the French, showed performance patterns which were stable over language materials, but differed critically in showing no advantage for cases where target sequences and carrier utterances matched syllabically over those where they did not.

The data for English speakers counter any argument that sensitivity to syllabic status reflects only what the monitoring task demands, and, taken together with the arguments against necessary lexical involvement, support a strong interpretation: For French speakers, the syllable is a basic perceptual unit. The very same data raise a mystery, since they say that the format in which inputs are represented as a preliminary to lexical processing is not fixed, but depends on the language spoken natively. We defer for the moment discussion of the problems which arise for the intended experimental contrasts in a language like English (but see below), to focus on the issue which for us makes these findings important. We take it that a working assumption in psycholinguistics has been that fundamental aspects of language processing are universal, determined in the first instance by the kind of mental object that language is—not any particular language, but human languages, taken generally. In these terms, it seems surprising that the routines employed by a listener in something so basic as segmenting the speech stream might be language-dependent, particularly if it is through syllable-sized units that the lack of invariance problem is to be overcome (as Mehler et al., 1981a, originally suggested). Cutler et al. (1986) recognised the difficulties of assuming that one language is somehow easier to perceive than another, and proposed instead that syllable-based analysis is one among a repertoire of strategies available to the newborn. Moreover, the incorporation of a particular strategy in the developing comprehension system is to depend on "the degree to which the structure of the language being acquired encourages the use of the strategy in question" (Cutler et al., 1983, p.160). They therefore looked for properties in the phonologies of French and English which might control the utility of analyses at the level of the syllable.

For Cutler et al. (1986), the difference between French and English lay in the presumed efficiency with which syllables can be identified in the two languages, turning on two factors; these were the *regularity* of syllable structure and the *clarity* of syllable boundaries, the latter being given more emphasis. On the question of regularity, the idea seems to be that while French admits of a variety of syllable structures, there is in some sense a prototypical syllable shape; and, indeed, in the estimate of Dauer (1983) for colloquial French, some 75% of syllables are open, and some 75% of those open syllables are simple CV structures. English, on the other hand, disperses its syllable types more broadly. For the factor of boundary clarity, Cutler et al. focused on analyses of English concerned with word-level stress patterns and associated allophonic variation, and cited those which represent consonants in certain restricted environments as improperly bracketed (e.g. Anderson & Jones, 1974) or, in the terminology which now is more standard, ambisyllabic (Kahn, 1976).[1] Ambisyllabicity is widespread but certainly not free in English: A typical instance prompting such analyses involves a syllable which has been assigned no coda by the initial rules of syllabification; a later fast-speech rule (usually referred to as Right Capture) permits that syllable to share the onset consonant of a following stressless syllable. The phonological arguments here involve the distribution of position-sensitive processes such as aspiration, (pre)glottalisation and weakening, whose correct prediction is said to require that some English consonants be syllable-initial for the purposes of one rule but syllable-final for another (e.g. Gussenhoven, 1986).

Cutler and colleagues' assertion was that where the phonology assigns a consonant simultaneously to two syllables, there can be no clearly flagged syllable boundary in the speech signal. There may be an immediate difficulty here. While both modern and traditional treatments of the phonology of French apparently agree on the placement of syllable boundaries, the ambisyllabicity analysis is not the only one available for English. Selkirk (1982), for example, rejects outright the view that consonants can ever be ambisyllabic, on theory-internal grounds, and argues instead that

[1]For the record, we note that French (with putatively clear syllable boundaries) has not been immune to ambisyllabicity analyses. Prunet (1987), for example, accounts for different patterns of nasalisation in liaison environments by proposing that "floating" nasal segments become ambisyllabic under certain structural conditions. Thus, the /n/ of specifiers like *mon* is said to act both syllable-initially and syllable-finally in *mon ami*, permitting both liaison and nasalisation (cf. *bon ami*, with liaison but no nasalisation). Prunet's ambisyllabicity analysis for French is not restricted to external sandhi situations; for structurally parallel word-internal cases involving the prefixes *non-* and *in-*, he makes arguments to the same point.

re-syllabification rules can give an appropriate account of the relevant phenomena.[2] In the theory she proposes, the fact that a segment can act in ways requiring different syllabic positions is handled in terms of syllable assignments which differ between underlying and surface representations; however, those assignments are absolute, and permit no shared consonants. More than theory-internal consistency or the niceties of theoretical terminologies are at issue here, and especially so when we are in the business of translating from the phonological notion of the syllable to a claim about its status as a perceptual unit.

To illustrate the point, consider the /l/ of English *palace*: Where Kahn (1976) analyses this as shared between syllables, Selkirk (1982) has it assigned to the second syllable in underlying representation *but exclusively to the first syllable in surface representation*. Whether those different claims about surface syllabification lead to different psychological predictions depends on the (as yet unspecified) character of the device segmenting the speech stream into its syllabic units. Suppose, on the one hand, that a segmentation device acts in a brute force way, and responds to broad acoustic cues (energy minima, say) in picking out syllables, disregarding the detail of phonetic content. With this characterisation of the segmentation device, competing claims about surface syllabification have real consequences, since it is a surface representation which the phonology outputs to the system implementing speech gestures. Under a re-syllabification analysis (but not an ambisyllabicity analysis), English may differ from French not so much in the clarity of a syllable boundary in an utterance of *palace* as in where that boundary is placed—for French, after *pa*, and for English, after *pal*. But now suppose, on the other hand, that the segmentation device is driven by precisely the kind of allophonic detail which is the common ground for competing phonologies. (In general, it is not the facts about English which are in dispute, but only how they are to be captured in linguistic theory.) In this case, the contrast of hard-to-syllabify English and easy-to-syllabify French stands just as Cutler et al. proposed on just the grounds they suggested, whatever phonological theory wins the day. Boundaries for which some cues go one way and some cues another way are, by that very fact, unclear. We note, however, that the syllable segmentation device as proposed by Mehler, Dupoux and Seguí (1991) seems necessarily to have the first rather than the second of the characterisations outlined above. In the framework of their SARAH

---

[2]Gussenhoven (1986) points out some instances in American and British English where different schools of analysis make different predictions, and concludes that the evidence, though conflicting, weighs on the side of ambisyllabicity. Our impression, however, is that the inherent variability in the phenomena in question (e.g. aspiration, nasal deletion, stop insertion) make theory-neutral observations difficult.

(Syllable Acquisition, Representation and Access Hypothesis), that device is to be the bootstrap to the infant's acquisition of a phonology as well as an inventory of lexical forms—and in being in place prior to any language-specific phonological knowledge, can have no access to the relevant facts of allophonic variation. Thus, to the extent that the processing formats of the initial state are carried over into the adult system (and this is the spirit in which SARAH is proposed), competing analyses of surface syllabification in English complicate the picture, and may undermine the grounds on which different developmental outcomes for native speakers of French and English are to be expected.

In pointing out these difficulties in Cutler and co-workers' treatment of syllabification in English, it is not our intention to dismiss the claim that something about French encourages syllabic routines while English discourages them. We wish only to stress that the notion "syllable" has as much a place in the phonology of English as in the phonology of French, that the translation from phonological constructs to processing models is by no means straightforward, and that the availability of a re-syllabification account means there can be no guarantee that surface syllables are realised with different precision in French and English. Thus the case which can be made rests largely on the monitoring data themselves. Given the potential significance of these findings for models of basic speech recognition processes, the research presented here undertakes a replication.

The proposal of Cutler and Mehler and their colleagues is clearly intended to be general; it is to be languages like French which encourage perceptual strategies tuned to proper syllabification, and languages like English which discourage them. In putting together our experiments, we therefore allowed ourselves to depart from exact replication in the languages studied. Castillian Spanish replaced French as the language putatively encouraging syllabic routines, and Australian English replaced British English as the discouraging case; in the respects which have been identified as relevant, Spanish is a language like French, distinct from either variety of English. Dauer's (1983) sampling shows the simple CV syllable to be as prototypical for colloquial Spanish as for colloquial French, and analyses of Spanish claim clear syllabification (Agard, 1984; Harris, 1983). One point of difference deserves mention. Unlike French, Spanish shares with English the property of lexical stress. It does not, however, share its widespread stress-conditioned reductions, i.e. the tendency in English for centralisation of unstressed vowels, so that the difference between stressed and unstressed syllable durations in Spanish is minor compared with that in English (Delattre, 1966). Thus, typologies which classify languages in terms of their characteristic rhythms place French and Spanish together, as syllable-timed, in opposition to stress-timed English.

The first four of the five experiments that follow constitute the basic replication study. Experiments 1 and 2 examine patterns of performance when subjects are asked to detect CV and CVC targets in materials constructed in their native Spanish and English languages, respectively; these cases set up the basic contrast in speech-processing routines which has been the focus of attention in earlier work. The prediction is that only Spanish-speaking subjects will show a syllabification effect, the interaction of target type and carrier type reported in earlier work. Experiments 3 and 4 ask English and Spanish speakers, in turn, to detect those same target sequences in non-native materials; these cases are intended to assess the extent to which perceptual routines are stable for a given population of speaker-hearers, regardless of the particular properties of the stimulus language. The experiments are presented one by one, rather than as a subject population × stimulus language cluster, because in each instance there are departures from the data patterns reported by Mehler et al. (1981a) and Cutler et al. (1986) which require extended comment. The final experiment, although using the design and materials of those earlier in the series, seeks to make a different kind of argument about the stability of perceptual routines. Thus, Experiment 5 uses a subject population not previously examined, Spanish speakers with late acquisition of English (generally, post-puberty), and tests them with native-language stimulus materials. The performance patterns of such "ex-monolinguals" assesses the robustness of the perceptual routines developed through childhood exposure to a language favouring syllabic routines when, by hypothesis, the second language differs in type.

In all the experiments reported here, stimulus materials were designed around pairs sharing initial CVC sequences (cf. in the original studies, *palace–palmier* for French, *balance–balcony* for English), but for greater generality and a roughly four-fold increase in item power, were permitted to cover a wider range in the composition of those initial sequences. A final point of departure from exact replication was also motivated by considerations of experimental power. The central claim that the language spoken natively determines the availability of syllable-based perceptual analyses relies, in the end, on a clear demonstration that a syllabification effect is present for subjects of one language background and absent for another. But effects in the original French study were certainly not large (mismatching syllabic status costing some 20 msec), and findings in our pilot study with Spanish speakers and Spanish materials were of the same order. It seemed likely that the way in which targets were specified might be important, since the basic effect relies on the match or mismatch between what subjects expect to hear (their mental representation of the target sequence) and what they in fact hear in a spoken list of words. In our pilot, as in the experiments of Mehler et al. (1981a) and Cutler et al.

(1986), targets were presented as orthographic sequences, and it was left to the subjects to form an image of a sound shape—which might not always precisely be the one intended. On the face of it, written targets avoid any bias in favour of the syllabification effect, but the avoidance of bias is more apparent than real: CV or CVC sequences, whether imaged as sound shapes by naive subjects or spoken in isolation in the preparation of an experimental tape, can only be rendered as syllables. To specify the target with a spoken presentation is to do no more than can be presumed to happen, covertly, when subjects are left to generate their own phonological representations; and plausibly, spoken presentation of target sequences lessens the opportunity for idiosyncrasies of pronunciation to introduce variability between subjects or between items. Our experiments therefore used spoken target specifications in preference to written ones, when the subjects were tested with native language materials; for non-native materials, the targets were specified in spoken and written form, simultaneously, to give maximum support.

EXPERIMENT 1

Subjects

Twenty-four students from the Universidad Complutense de Madrid participated in the experiment, none of whom had taken part in any pilot study. All of them were monolingual speakers of the Castillian dialect of Spanish.

Materials and Design

Forty-four carrier Spanish words were selected, as 22 pairs.[3] All of them were trisyllabic nouns with stress on the penultimate syllable, the regular stress pattern for this language. The elements of each pair began with a common CVC sequence (e.g. *<u>pal</u>oma, <u>pal</u>mera*), but differed in syllabic structure, one having only a CV subsequence as its initial syllable (e.g. *pa'loma*), while for the other the full CVC constituted the initial syllable (e.g. *pal'mera*). Unlike the materials of Mehler et al. (1981a), this carrier set covered a wide range in the composition of its word beginnings. That is, carriers could begin with a stop consonant, a fricative, a liquid or a nasal; the first vowel could be any one of the monophthongs permitted by

[3]Twenty-four word pairs were originally selected and employed in the experiment, but an error in the construction of the list materials meant that two pairs had to be excluded from the analysis. The error was that the specified target appeared among the filler words preceding the carrier, though in a non-initial position. The excluded items become, in effect, catch trials of an additional type.

Spanish, and the second consonantal position could be a liquid (/l/, /r/), a nasal (/n/, /m/) or the fricative /s/. (The set of item pairs is listed in Appendix 1.) In contrast, the five carrier pairs of the original French study were all stop-initial, and had only /l/ or /r/ in the second consonant position, with a uniform use of the vowel /a/.

Two target sequences were assigned to each of the 44 words, corresponding to the initial syllable proper (i.e. the CV *pa* in *pa'loma* and the CVC *pal* in *pal'mera*), or having one phoneme more (*pal* in *pa'loma*) or one phoneme less (*pa* in *pal'mera*). The experimental design, in short, factorially combined a contrast of target type (CV and CVC sequences) with a contrast of carrier type (words properly syllabified as CV- and CVC-initial).

Each of the 44 words carrying a target sequence was placed as the final item in a list, preceded by filler words. The elements of each carrier pair appeared in length-matched lists, placed from the second to the fifth position. The fillers themselves differed for each element of a pair, and could be nouns, verbs or adjectives, with from two to four syllables. Trisyllabic nouns were included in greatest proportion, since the carriers were of this type. No fillers included the CV or CVC sequence which began their list-final carriers.

In addition to the 44 experimental lists, a set of 24 distractor lists was constructed. These also varied in length from two to five words. The words which appeared in final position were trisyllabic nouns similar in structure to the carriers, but did not bear the target assigned to those lists. To encourage the subjects to base their monitoring responses on the detection of complete target sequences, 16 of the 24 distractor lists were designed to constitute catch trials. That is, the specified target had either the first consonant of the list-final word (e.g. *cul/colina*) or its first vowel (e.g. *tar/madera*). A set of 12 practice lists was constructed along the same lines as the experimental and distractor lists. List-final carriers bearing the target sequence were included in only eight of the practice items.

The lists of words, ordered pseudorandomly, were recorded by a male Spanish speaker (Castillian dialect) on one channel of a two-channel tape-recorder (Revox B77). Reading from a word-by-word computer screen presentation with list elements displayed at 1.5-sec intervals, the speaker used the intonation appropriate for isolated words. He had knowledge neither of the target which was to be specified with any list nor of how many words it included; thus, no special emphasis could be given to carrier words, and the intonational line had no sharp fall, list-finally. Before each list, the following sequence was recorded: first, "responde al sonido" to focus the subject's attention on the target specification for that list and, secondly, "atención" as a signal for the beginning of the list, following 1 sec later.

The experiment was constructed in two versions. The same tape of the materials set was presented in both, with target type being manipulated for experimental items between versions; there was thus no repetition of materials to any subject. For half of the words within each carrier type within each version, the target sequence corresponded to its proper initial syllable; for the other half, it was the longer or shorter non-syllable sequence. Across versions of the experiment, both target types were combined with each of the carriers. Distractor lists, however, were identical in the two versions. Spoken tokens of the targets were recorded separately by the same Spanish speaker, then dubbed into the already recorded list sequence so that the targets were heard between the instruction "responde al sonido" and the ready signal "atención". Two separate tapes were mixed in this way, since the critical targets for experimental items differed in each of the two versions of the experiment.

Timing clicks were placed on the tape's second channel, just before the onsets of carriers. These clicks, inaudible to the subject, started a millisecond clock which would be stopped by the subject's key-press response. With a Nicolet digital oscilloscope, a millisecond adjustment was measured for each click placement, so that reaction times were in effect taken from the onset of carrier words.

Procedure

The subjects were tested individually, 12 in each of the two versions of the experiment. They were instructed to listen to each list of items, and to listen for a sequence specified separately for each list, responding with a button-press only when the target occurred. The instructions emphasised both speed and accuracy. Reaction times were recorded manually by the experimenter, as were instances of false alarm. The duration of the experiment was approximately 30 min, with a break of 5–10 min halfway through.

The data from seven subjects making more than four omissions or false alarms were rejected, and those subjects were replaced. For the reaction time data (positive targets), cut-off points were established to moderate the influence of extreme reaction times. First, absolute limits were established at 200 and 2000 msec, and reaction times outside these limits (constituting some 3% of the total) were excluded from the analysis. Secondly, boundary values were set for each subject at two standard deviations on either side of his or her overall mean reaction time, and reaction times lying beyond these limits were replaced by the limiting value. Means for each condition (by subjects, collapsing over items, for the subject-based analysis, and by items, collapsing over subjects, for the item-based analysis) formed the data matrices for the analyses of variance

TABLE 1
Mean Monitoring Reaction Times for Spanish Speakers and Spanish Materials

Item Type	*Example*	*Reaction Time (msec)*
CV carrier words		
CV targets	*pa/paloma*	556
CVC targets	*pal/paloma*	609
CVC carrier words		
CV targets	*pa/palmera*	602
CVC targets	*pal/palmera*	565

which examined the effects of two factors, type of carrier and type of target, and their interaction. In interpreting analysis outcomes, we assumed significance at the 10% level for the min F' statistic in which subject and item F-values were combined, since Monte Carlo studies have shown it to be somewhat conservative (Forster & Dickenson, 1976). Min F' was calculated only when the component analyses themselves reached conventional levels of significance.

Results and Discussion

The mean reaction times for each type of carrier word and target sequence are presented in Table 1. The data show a syllabification effect for Spanish speakers which is both sizeable and robust. As predicted by Cutler et al. (1986) for speakers of languages like French, monitoring performance was faster when the target corresponded to the initial syllable of the carrier than when it was one phoneme more or less, leading to a cross-over interaction of carrier and target type [min $F'(1,53) = 8.91$, $P < 0.01$]. Subanalyses showed the components of the interaction to be themselves strong. For CV-syllabified words, the advantage of a CV target over a sequence which had one phoneme too many was 53 msec [min $F'(1,41) = 8.39$, $P < 0.01$]; for CVC-syllabified words, the corresponding difference between a CVC target and one which had one phoneme too few was 37 msec [min $F'(1,41) = 3.16$, $P < 0.10$]. A tendency for responses to be faster overall for CV targets (579 msec) than CVC targets (587 msec) was significant neither by subjects [$F_1(1,22 = 1.01$, $P > 0.25$], nor by items ($F_2 < 1$), and the effect of carrier type did not reach significance in either analysis (both Fs < 1).

It seems, as we predicted, that the procedural change from written to spoken target specification acted to limit variability, with the result that the syllabification effect is roughly doubled in size from our pilot[4] and Mehler and co-workers' (1981a) study; mismatching syllabic status now costs some 45 msec, on average. Overall, our finding of a syllabification effect replicates Mehler and colleagues' (1981a) result, and by generalising to Spanish materially strengthens the case that syllable-based representations of spoken inputs are made perceptually relevant by languages with particular phonological characteristics. We note, however, that the subjects' reaction times in this experiment (averaging 583 msec) were slower by some 200 msec than in the French study. Because of the weight given to fast responding in shaping Mehler and co-workers' interpretation of their data, this difference in outcome requires comment; first, about how it might have arisen and, secondly, about how it might constrain our arguments.

What underlies differences in speed of response cannot be decided definitively, although a comparison with our pilot study makes it clear that the shift from written to spoken target specification is not a factor; if anything, performance was slower still with written targets. A number of likely possibilities remain. The experiment reported here employed catch trials, whereas the French experiment did not, and Mills (1980) reports that monitoring responses become slower (and presumably more accurate) in these circumstances. Substantial trial-to-trial variation in the composition of stimulus materials (cf. the relative uniformity of target shapes in the French study) means, first, that the subjects' task was inherently a more difficult one. It also means that the acoustic duration of the relevant portions of the carrier stimuli was greater, overall—only 59% of carrier onsets were stops (cf. stop onsets, uniformly, in the French study)—and where carriers take longer to hear, reaction times will be elevated, directly. Finally, it may be relevant that our Spanish-speaking subject population had had no prior experience in speeded response tasks.

Whatever the reason for slower responding, the argument around sheer speed which allowed Mehler et al. (1981a) to claim that syllable-sensitive performance in French subjects was based on a pre-lexical representation cannot be made for our Spanish subjects. Indeed, Seguí, Dupoux and

[4]The pilot experiment used the master recording of Experiment 1 and subjects drawn from the same population, but presented the target for each list in written form on a card which the subjects themselves turned up on hearing the instruction "responde al sonido". No spoken version of the target was offered. The data showed a cross-over interaction of carrier and target type [min $F'(1,61) = 2.83$, $P < 0.10$], with CV target detection (618 msec) faster than CVC target detection (642 msec) for CV-syllabified words, and a reverse effect, CVC detection (636 msec) faster than CV detection (647 msec) for CVC-syllabified words. There was no effect of carrier type [$F_1(1,22) = 2.87$, $P > 0.10$; $F_2 < 1$] or of target type [$F_1 < 1$; $F_2(1,40) = 1.38$, $P > 0.10$].

Mehler (1991), commenting on an earlier report of these data (Sánchez-Casas, 1988), warn that the representational code engaged in performance of the task might be a lexical one. While this possibility cannot be denied, data reported by Zwitserlood (1991) argue against any necessary assumption of a lexical basis when responses are slow. Using Dutch, Zwitserlood contrasted CV and CVC detections for materials with unambiguous syllabification (e.g. *dee'ling*, *deel'baar*) in a design like that of Experiment 1, and found the same cross-over interaction. A complementary study indicated quite clearly that performance was sensitive to the syllabification cues of the speech signal; here, materials were cross-spliced versions of the original stimuli (yielding reversed surface syllabifications, e.g. *deel'ing*, *dee'lbaar*) and produced an exact mirror reversal of the outcome. In short, the data followed the signal. What makes Zwitserlood's study relevant to the issue of response speed is that the Dutch subjects' reaction times were of the order of 600 msec. Clearly, slow responding does not rule out the use of pre-lexical codes.

With the appropriate caveats, our conclusion is that Experiment 1 has indeed extended the case made by Mehler et al. (1981a), for French, to a language with similar phonological properties. And, since the claim for language specificity in the availability of syllable-sensitive routines depends as much on the absence of a syllabification effect for English speakers as it does on its presence for French and Spanish speakers, we proceed in Experiment 2 to examine the situation for English. As a preliminary, however, we outline what seems to us an almost intractable problem in the construction of materials for English which are in any strict sense parallel to those for Spanish.

Close replication commits us to comparing monitoring performance in pairs like *palace* and *palpitate*, for which Cutler et al. (1983; 1986) claimed contrasting syllabifications as [pa[l]ace], with ambisyllabic /l/ and a syllable boundary which was therefore unclear, and [pal][pitate], with a well-defined initial syllable. But, as we have already noted, Selkirk's (1982) analysis of English proposes that words like *palace* have a surface syllabification as [pal][ace], and there may be a further complication for the paired cases with a medial consonant–consonant sequence which the experimental design calls for. Cutler and co-workers' treatment of words like *palpitate* as having an initial CVC syllable, unambiguously, reflects a focus on the status of intervocalic consonants in the linguistic literature on ambisyllabicity. But note the possibility of Flapping, a litmus for ambisyllabicity in Kahn's (1976) analysis, in the pronunciations of words like *winter*, *shelter* and *center* in a General American dialect; and, similarly, the occurrence of the glide /y/ in words like *pendulum* in the same dialect (which allows no glide in words like *duel*), signalling a re-syllabification in the analysis of Borowsky (1986). Accounting for these and other intricately interrelated facts in British and American English, Gussenhoven (1986) proposes that ambisyllabicity can arise not only intervocalically, but also where the "captor" syllable has a coda consonant, provided the resulting cluster is a legal one. And, just as Gussenhoven allows Extended Right Capture to incorporate more material than does Kahn with Right Capture, so Anderson and Jones (1974) represent the medial /p/ of *palpitate* as ambisyllabic, and the analysis offered by Selkirk permits its re-syllabification to give [palp] as the surface initial syllable.

What then is being contrasted, when subjects are asked to detect *pa* or *pal* in *palace* or *palpitate*? In all that follows, we refer to these two word types as CVC(V) and CVC(C), the bracketed segment specifying only the right context for an initial CVC sequence without any commitment to patterns of syllabification. Under the ambisyllabicity analysis, the inclusion of CVC(C) words among those permitting ambisyllabicity might amount to no more than a strengthening of the claim that unclear syllable boundaries are rife in English, if it were not for the complication they create in the prediction of experimental outcomes.

Cutler et al. outlined three data patterns which they supposed might provide evidence of real syllable sensitivity in speakers of English. The first was simply the cross-over pattern which had been observed for French (and now, for Spanish). This outcome would arise if underlying rather than surface syllabification was relevant (or, as the authors put it, if ambisyllabicity was ignored). The second pattern which they entertained, reflecting their focus on ambisyllabicity in intervocalic contexts, saw no difference between target types for CVC(V) words but an advantage to CVC targets for the CVC(C) words which they supposed to be clearly syllabified. (And this is just the outcome they observed for French-speaking subjects with English materials.) Clearly, if both word classes are permitted ambisyllabicity—and if ambisyllabicity means unclear boundaries—we might be in the unhappy position of having a materials design for which no difference between target types is predicted for English, anywhere. Cutler and co-workers' third pattern allowed faster performance for CVC targets in ambisyllabic CVC(V) words as well as in CVC(C) words, because that sequence does constitute the initial syllable, if not uniquely. An ambisyllabicity analysis (or, indeed, Selkirk's re-syllabification) recognising more than intervocalic contexts muddies the waters here, since CVC(C) now constitutes the initial syllable in the case which had been supposed to have CVC, unambiguously. With surface syllables maximised, CVC targets would match the initial syllable in CVC(V) words but fall short in CVC(C) words, so that syllable sensitivity would be reflected in an advantage to CVC targets for CVC(V) words only, and no difference for CVC(C) words (i.e. the opposite of the outcome observed for French speakers and English materials). We note, however, that CV is in no sense a possible syllable of CVC(C) words; in the cases we are considering, the medial consonant sequence (e.g. /lp/ in *palpitate*) is always an illegal onset, and that fact alone might be sufficient to produce a performance advantage for CVC detections—though not one that reflects syllabification, strictly.

Given this maze of possibilities, it was a quirk in the data of Cutler et al. (1986) which allowed them to argue, strongly, for non-syllabic routines in the early stages of comprehension in English speakers. None of the patterns outlined above was observed, but only uniformly faster target detections in CVC(V) words. The authors sought an account of the finding in terms of *sub*-syllabic identifications. Citing earlier studies examining environmental effects on the identification of different segment types (Liberman, Delattre, Cooper & Gerstman, 1954, for consonants; Strange, Verbrugge, Shankweiler & Edman, 1976, for vowels), they supposed that strings made up of alternating consonant–vowel sequences (as in *palace*) had their elemental segments identified more easily than non-alternating sequences (as in *palpitate*). Since an explanation of this kind assumes that English speakers characteristically examine speech inputs in terms of their constituent segment types and not any larger unit, the unexpected finding was cast as positive evidence for perceptual routines pitched at the level of phonemes, in English speakers.

Against this background, we present Experiment 2 in replication of the critical experiment which, for Cutler et al., laid the foundation of a claim that the perceptual routines through which speech inputs are first represented depend on the language spoken natively. English speakers monitor for CV and CVC targets in stimulus materials selected with the same properties as those of the original experiment; but, as in Experiment 1, these sample the range of the word beginnings permitted by the language, more broadly.

EXPERIMENT 2

Subjects

A total of 24 students from Monash University acted as paid participants in the experiment. All of them were monolingual speakers of English.

Materials and Procedure

Forty-eight English words were selected, as 24 pairs (Appendix 2 lists this item set). For the closest possible parallel to the Spanish materials of Experiment 1, the words making up any pair began with a common CVC sequence but then diverged, one type continuing the initial sequence with a vowel (e.g. *palace*) and the other with a consonant (e.g. *palpitate*). These critical carrier words were polysyllabic nouns or verbs (mean number of syllables 2.5, range 2–4), which all had primary stress on the initial syllable, and a stressless second syllable. The range of segment types making up the common CVC sequences was as broad as possible, consistent with identity across the pairs; thus in materials construction, no constraints were placed on the first consonant (only 50% were stops) and the sequence could continue with any of the short vowels /æ/, /ɛ/, /ɪ/ or /ɔ/. The second consonant varied among the liquid (/l/), the nasals (/n/, /m/) and the fricative /s/; in Australian English, as in British English, /r/ is not possible without altering vowel quality. Two target sequences were specified for each of the 48 words, for a design factorially combining a contrast of target type (CV and CVC) with a contrast of carrier type [words beginning CVC(V) and words beginning CVC(C)].

The carrier words were placed as last items in lists of words occupying from the second to the fifth position; the words of each pair were placed in length-matched lists. The fillers were similar in structure to the carrier words, but did not contain any specified target. An equivalent set of 24 distractor lists was constructed; none had words containing their specified target, but 16 included catch trials designed in the same way as those in the Spanish experiment. That is, the CV and the CVC targets in the catch trials could contain either the same consonant as the last word of the list (e.g. *ba/bundle*) or the same first vowel (e.g. *ta/gallery*). In total, the subjects were presented with 72 list sequences: 48 experimental lists with carrier words and 24 distractor lists without them. These were preceded by 12 practice lists of similar distribution.

The word lists were recorded in a pseudorandom sequence by a male speaker of General Australian English (the dialect of the subject population); the circumstances of recording were identical to those in the preparation of the tape for Spanish materials, as was the protocol establishing the timing of his utterances. The within-item sequence of Experiment 1 was followed, so that each trial began with the instruction "respond to the sound", continued with the alerting signal "ready", and went, 1 sec later, to the list of words. Target sequences were recorded on separate tapes and then dubbed into the place to produce, as before, two versions of the experiment. In recording targets, the speaker was careful to pronounce all vowels as short; extensive coaching ensured that the CV sequences, whose short vowel makes them illegal in English as isolated syllables, were as natural as possible.

The experiment followed the procedures and data treatment used previously, except that the instructions to the subjects at the beginning of the experimental session were explicit in pointing out that CV and CVC targets would always be encountered with short vowels.

Results and Discussion

The mean reaction times are presented in Table 2 for target sequences of each type, detected in each type of word. The data here are as clear as they were for the Spanish speakers in Experiment 1, but fall into a strikingly different pattern: There was no interaction of target and carrier type ($F < 1$ in both subject and item analyses). Nor is the pattern like that observed by Cutler et al. for English speakers: Detections in CVC(V) words (409 msec) were not reliably faster than in CVC(C) words (412 msec) ($F < 1$ in both analyses). Only an effect of target type was shown, with responses to CVC targets faster (398 msec) than to CV targets (422 msec) [min $F'(1,49) = 7.35$, $P < 0.01$].[5]

TABLE 2
Mean Monitoring Reaction Times for English Speakers and English Materials

Item Type	*Example*	*Reaction Time (msec)*
CVC(V) carrier words		
CV targets	*pa/palace*	419
CVC targets	*pal/palace*	398
CVC(C) carrier words		
CV targets	*pa/palpitate*	426
CVC targets	*pal/palpitate*	397

The absence of the interaction which has been the evidence for syllable-based analysis in French and Spanish speakers is not itself problematic, for, as has been noted, the indicators of syllable sensitivity may differ between Spanish and English just as the facts of surface syllabification do. But the absence of any advantage for CVC(V) words over CVC(C) requires explanation, and especially so since, for Cutler et al. (1986), this finding was the basis for a suggestion about the character of a *non*-syllabic input representation in English speakers.

[5]Differences between the data pattern in Experiment 2 and that of Cutler et al. (1986) cannot be due to the difference between spoken and written specifications of the monitoring targets. An identical pattern emerged when the same materials were run with a new group of 24 subjects and a visual presentation of targets. Again, neither the interaction nor the carrier type effect reached significance ($F < 1$ in all analyses), but CVC targets (368 msec) were detected faster than CV targets (399 msec) [min $F'(1,65) = 9.61$, $P < 0.01$].

Differences in materials seem likely to be important here, and the variety of CVC initial sequences in the critical pairs of Experiment 2 is to be compared with Cutler and co-workers' uniform use of a single vowel /æ/ and a single medial consonant /l/. In an examination of materials effects on what they called "phonological expectation", Cutler, Norris and Williams (1987) asked subjects to give polysyllabic completions for the target sequences of their earlier study, and found that words beginning CVC(V) were produced roughly four times as often as words beginning CVC(C), whether the word fragment was specified as CV or CVC. That is, the word shape which was most typical in the completion task was exactly the type for which an advantage was observed in the monitoring task. However, the preponderance of CVC(V)-word responses was found to reduce markedly when completions were sought for a more varied set of sequences. We therefore undertook a similar exploration of "expected" word shapes, employing only the CVC targets of Experiment 2 to guarantee the appropriate short vowel pronunciation. Fifteen volunteers who had not taken part in Experiment 2 were asked to produce polysyllabic completions; among the responses produced, only 50.7% began CVC(V), so that CVC(V) words were not more likely than CVC(C) for these materials. In short, there is a case to be made that the presence or absence of an overall advantage to CVC(V) words in the monitoring task reflects an expectancy bias set up by the experimental materials—and hence that the effect is not likely to be relevant to the central question of the pre-lexical representation of spoken inputs.

Experiment 2's second point of departure from earlier findings is one not so readily set aside, because faster monitoring performance with CVC targets is one of the three data patterns which, for Cutler et al. (1986), could signal the use of syllable-based routines in English speakers. If the CVC advantage is interpreted as evidence of syllable sensitivity in English speakers, what demands an explanation is why Cutler et al. (1986) found no such effect in three separate experiments. Arguments around experimental power are not likely to be useful, since in two of those three experiments, Cutler and co-workers' target type non-effect ran, if anything, in favour of CV targets and thus against the putative syllabification effect.

When a response time advantage is tied to a particular syllable type, factors having nothing to do with syllabification are potentially in play, and it was for exactly this reason that the original French study used different kinds of carrier words, half with CV and half with CVC initial syllables. Earlier work in which monitoring performance was compared across targets at different levels of analysis (phoneme, syllable, word) raised the possibility that sequence size itself might determine the certainty with which the match between a target sequence and an utterance could be asserted—roughly, more stuff, more certainty—and hence control the speed of response. The finding that responses are faster for three-element sequences (CVC targets) than two-element sequences (CV) is open to an interpretation in these terms.[6]

Beyond sequence size, CVC and CV targets are likely to differ in the accuracy with which a mental model displays the target intended by the experimenter, with the disadvantage decidedly on the side of CV. Our use of an auditory specification of the monitoring targets was intended in the first instance to reduce between-item and between-subject variability in the faithfulness with which target representations were generated. It might also go some way, we thought, to overcoming an unavoidable difference between CVC and CV targets. The simple illegality, in English, of an isolated open syllable with a short vowel—the difficulty experienced by our model speaker in producing CV sequences, prior to training, testifies to their unnaturalness—stands in contrast to the legality of CVC. But though a spoken target might provide a firm basis on which to *form* an accurate mental representation of a CV sequence, it might not be sufficient to *maintain* that representation veridically. In the experiments reported here, the target was presented at the beginning of any list, between 4 and 9 sec prior to the critical carrier word; thus, the subject's representation of a target had to be maintained beyond the period over which acoustic detail can be held in memory (estimated at around 400 msec; e.g. Howell & Darwin, 1977). Where the phonotactics weigh against a short vowel interpretation for open syllables, any drift over time in the now abstract representation is likely to be towards a legal sound shape, and away from the intended target. If we assume that problems in maintaining the representation of illegal shapes are likely to be greater when the vowels vary (as here) than when the vowel is fixed (as in the studies of Cutler and her colleagues), a difference in experimental outcomes may again find its explanation in a materials difference.

If Experiment 2's finding of faster responses to CVC targets is open to several interpretations—as a genuine syllabification effect reflecting the facts of surface syllabification for English, or as a certainty-plus-legality advantage for those longer sequences—the same will not be true where the pattern of surface syllabification is not itself at issue. Experiment 3

[6]Even a finding that CVC targets do not take *longer* to detect than CV targets (Mehler et al., 1981a; our Experiment 1) requires explanation. It must surely be the case—co-articulation notwithstanding—that the sensory information supporting a CVC hypothesis (whether as a syllabic unit, or as a sequence of phones) takes longer to *hear* than that for CV. Some notion like certainty is undoubtedly called for. What is special in the case we are considering is that patterns of reaction time run so strongly against the facts of the signal's deployment over time.

therefore aims to clarify the situation by examining the performance of English speakers with the CV- and CVC-syllabified Spanish materials of Experiment 1. Here, a sensitivity to the syllabification cues provided by the signal can only be reflected in the interaction of target and carrier types seen previously.

EXPERIMENT 3

Subjects

Twenty-four undergraduate and graduate students from Monash University participated in this experiment. All of them were native English speakers with no knowledge of Spanish. Most had no experience with any other Romance language, and for the very few who did, knowledge of the language in question was very rudimentary.

Materials and Procedure

The Spanish language materials were those of Experiment 1, in which classes of trisyllabic nouns were syllabified either with CV as initial syllable (e.g. *pa'loma*) or with CVC as initial syllable (e.g. *pal'mera*). Thus, any processing advantage for targets having real syllable status in the carrier utterance will be revealed by faster performance with CV targets than with CVC targets, for CV words, and the reverse pattern for CVC words. The recordings of the earlier experiment were used, unaltered, so that for each list our monolingual English speakers heard in sequence: the instruction "responde al sonido", an utterance of the appropriate target (e.g. "pa"), the alerting signal "atención" and, finally, the Spanish word list.

One difference in procedure was designed to fix the description of the non-native phoneme sequence as securely as possible for subjects who had absolutely no prior knowledge of Spanish. (In Australia, Spanish is not a prominent community language, nor a diplomatic or business language, and is rarely encountered at school.) For additional support of the spoken target, the subjects themselves turned up a card on hearing the instruction "responde al sonido" to see the target also in written form, and kept that card in front of them until the completion of the list. The orthographic representation of target sequences was an English rather than a Spanish one; in particular, targets spelled in Spanish with "c" were written as TH before front vowels, and Spanish "v" was written as B. The subjects were coached at the beginning of the experimental session to familiarise them with the sound shapes they would hear. In all other respects, the procedures of the earlier experiments were followed.

Results and Discussion

Table 3 summarises the data for English speakers monitoring for targets in Spanish language materials. The pattern here exactly reproduces that found in Experiment 2 for the same population of speakers. There was no evidence of the interaction between carrier and target type which characterised the performance of Spanish speakers ($F < 1$ in both subject and item analyses), nor any overall effect of word type [$F_1(1,22) = 3.39$, $P > 0.05$; $F_2 < 1$]; however, target type did have an effect [min $F'(1,60) = 5.52$, $P < 0.05$]. Responses for CVC targets (382 msec) were again faster than for CV targets (404 msec). Though the subjects commented on the difficulty of the task, it is to be noted that reaction times here (grand mean 393 msec) are comparable to those of English speakers in the earlier experiment (410 msec). The number of false alarms showed a slight elevation, and their distribution suggests its cause; though more than half were made on distractor items which shared either a first consonant or a first vowel with the target (e.g. pa in pintura), as had happened for English language materials, false alarms also occurred where a filler item shared only features with a target (e.g. mo in *botella*, cor in *golpe*). Given that the details of the phonetics of Spanish are unavailable to English speakers, just this kind of difficulty in discrimination might be expected. Against these considerations, the performance of the English speakers was impressively fast and accurate.

The findings here rule out the possibility which Experiment 2's outcome left live, that English speakers construct input representations which respect the syllabification of an utterance. For Spanish materials, a uniform advantage for CVC targets (for example, for *pal* in *pa'loma* just as in *pal'mera*) can have no interpretation in terms of a routine which is driven by the surface cues to syllabification, and thus our disagreement with

TABLE 3
Mean Monitoring Reaction Times for English Speakers and Spanish Materials

Item Type	*Example*	*Reaction Time (msec)*
CV carrier words		
CV targets	*pa/paloma*	397
CVC targets	*pal/paloma*	379
CVC carrier words		
CV targets	*pa/palmera*	411
CVC targets	*pal/palmera*	385

Cutler and her colleagues about the facts need not translate to a disagreement about the claim to be made: Speakers of English are unmoved by syllables.

For questions of the kind of input analysis which is in fact undertaken by English speakers, whether listening to native or non-native stimulus materials, the data here are in a sense negative. That is, they tell us what English speakers are not doing—they are not syllabifying utterances—rather than what they are doing. A more precise test of Cutler and co-workers' claim that the comprehension strategy rests with the speaker rather than with the language of the test materials is to be found where positive evidence of syllabic routines is likely. Experiment 4 completes this replication series by looking at the performance of monolingual Spanish speakers with English language materials, the converse of Experiment 3. As we have noted, the controversy surrounding the analysis of surface syllabification in English makes the prediction of an outcome difficult; nevertheless, Cutler et al. were able to demonstrate that French speakers showed at least one of the patterns compatible with syllabic routines for English materials, faster detection of CVC than CV targets with CVC(C) words, and no more than a trend in the opposite direction for CVC(V) words. If these French data were replicable, syllable-based routines in Spanish subjects should carry over to English language materials.

EXPERIMENT 4

Subjects

A naive group of 24 students from the Universidad Complutense de Madrid participated successfully in the experiment. All had Spanish (Castillian dialect) as their native language, with at most a very basic knowledge of English, but more typically no knowledge. For the majority of these subjects, English was not only not studied, but very rarely heard, there being little English language material in the popular broadcast media at the time the data were collected.

Materials and Procedure

The English language materials and the recordings of Experiment 2 were used and, as in Experiment 3, the spoken monitoring targets were supported by written versions on cards that were turned up by the subject as the instruction "respond to the sound" was heard. The subjects were also coached, prior to the experiment, in the sound values to be encountered in an unfamiliar Australian English.

The criteria for acceptable subject performance were set more loosely than in previous experiments, with data from five subjects who made more than 10 omissions or false alarms being replaced in the analysis. It was simply the case that, for no clear reason, our Spanish subjects had greater difficulty in performing the monitoring task with English materials than did the English-speaking subjects with Spanish materials. All other data procedures, however, were the same as those previously described.

Results and Discussion

Table 4 presents the mean reaction times in the monitoring task for each type of carrier and target sequence. The data show that syllable-sensitive routines in Spanish subjects do not carry over to English language materials, in contrast to the pattern previously reported for French speakers; there was no evidence of an interaction which might indicate syllabification of the English language input ($F < 1$ in both subject and item analyses). The remaining pattern which might support a claim for syllable sensitivity has a uniform advantage for CVC over CV targets, but here the data show at best a trend. The difference between CVC (518 msec) and CV (531 msec) target types was in the same direction as that obtained for English-speaking subjects with both English and Spanish materials, but was half the magnitude seen earlier and did not reach significance in either analysis [$F_1(1,22) = 2.70$, $P > 0.10$; $F_2 < 1$]. There was no reliable effect of word type [$F_1(1,22) = 4.47$, $P < 0.05$; $F_2(1,44) = 1.28$, $P > 0.25$], though the trend was for slower performance on CVC(V) words (534 msec) than on CVC(C) words (515 msec).

Thus a subject population for whom the syllabification effect was very robust on native language materials failed to show evidence of syllabification on non-native materials. We have noted that false alarms occurred in

TABLE 4
Mean Monitoring Reaction Times for Spanish Speakers and English Materials

Item Type	*Example*	*Reaction Time (msec)*
CVC(V) carrier words		
CV targets	*pa/palace*	541
CVC targets	*pal/palace*	527
CVC(C) carrier words		
CV targets	*pa/palpitate*	521
CVC targets	*pal/palpitate*	509

greater number (for the final set of subjects, on 7.0% of lists), but this hardly seems common enough to wash out effects, of itself, and the data obtained in this experiment apparently constitute a strong disconfirmation of Cutler and co-workers' claim that listeners process foreign language inputs with the routines which characterise treatment of native language inputs. Given the almost four-fold increase in the power of the experiment relative to the earlier report for French, the failure in replication cannot be dismissed lightly. We conclude that a syllable-sensitive routine may indeed reflect a Spanish listener's usual approach to the representation of speech inputs, but that its application is not independent of the stimulus language.

In speculating about what might differ between French and Spanish subjects to explain stability in perceptual routines for one group but not for the other, the thing which strikes us as most obvious is prior experience with English. Though the details are sketchy, the French subjects were reported only to be non-fluent, some having secondary school study of English as a foreign language; thus, Cutler et al. had only to remind their subjects of its grapheme–phoneme correspondences, whereas our Spanish subjects' minimal knowledge of English made a good deal more work necessary for accurate target specification. The original point of foreign language monitoring was a demonstration that materials drawn from a particular language do not induce a sensitivity to syllabification, for the nonce, but rather that particular language backgrounds encourage listeners to seek syllabic units in whatever inputs they hear. But this may be over-ambitious. An unfamiliar language may simply be hard to analyse with the categories of the native language, whether syllabic or non-syllabic; and with the limited detail provided by the monitoring data, there can be no guarantee that target/utterance matches are made with the same accuracy as for native inputs. Different patterns may emerge in foreign language monitoring performance for reasons that are entirely uninformative about the routines which characterise comprehension of native language inputs.

Perhaps worst of all, a finding of similar patterns for foreign and native language materials could be misleading, because they could arise in different ways. In illustration of this point, consider the findings of Cutler et al. (1986) for French speakers and English materials. They observed a 27-msec advantage for CVC over CV targets in the CVC(C) words which they supposed to be unambiguously CVC-syllabified and a non-significant 9-msec difference in favour of CV targets for the CVC(V) words which they supposed to be ambisyllabic. They argued that this reflected the employment of a syllabic routine with easily syllabified words and, simultaneously, a non-syllabic routine where ambisyllabicity blurred the cues to boundary. Suppose, though, that there is an across-the-board ambisyllabicity in materials of this kind, blurring all boundaries, as some analyses of English have suggested. Then, an alternative interpretation of these French data must be sought. For us, the best-fitting (though by no means satisfactory) solution supposes that a real target type difference for CVC(V) words has simply been underestimated at 9 msec [cf. Mehler and co-workers' (1981a) report of an effect of only 19 msec for the same subject population with French pronunciations of CVC(V) words], so that the pattern does not depart at all from that found with native language materials. This could only mean that the subjects were not constructing syllable-level representations which reflected the input's acoustic cues to boundary, but were imposing syllabifications in accord with the phonotactics of their own language. A picture only slightly different emerges from the analysis which, through re-syllabification, maximises all surface initial syllables bearing primary stress. Now the difference between CV and CVC target types found with CVC(V) cases in the data for French speakers runs *against* the input's supposed syllabification; and since neither target sequence matches the surface initial syllable proposed for CVC(C) cases (the CVC sequence falls short, though less drastically than does the CV), it cannot be the input's syllabification which accounts for the difference found for these cases, either. The problem in fitting the data is a different one, but the solution we see is identical to that offered when ambisyllabicity is assumed: The syllabifications which are at issue are imposed, top-down, and the subjects are not responding at all to the input's boundary cues.

For the general point which Cutler et al. wish to make, that French speakers characteristically represent spoken inputs in syllabic terms while English speakers do not, it probably does not matter whether input syllabifications are shaped by the phonotactics of the listener's language or driven by whatever acoustic cues signal syllable boundaries, any more than it matters that listeners typically categorise foreign sounds in terms of the segmental inventory of their native language. At the same time, it probably does matter for a more detailed consideration of the way in which a syllabically based routine might operate. In particular, a syllabification which operates top-down is likely to have its fit resolved in terms of what parsing of segment sequences is legal, where the assessment of legality crucially turns on segment content. What may be lost, then, is the claim that syllables are the primary units to be extracted—for foreign language inputs, at least.

Questions of the way in which non-native materials might be analysed have complicated, unnecessarily, what was intended to be a straightforward demonstration of the stability of syllable-based perceptual routines for speakers of languages like French and Spanish. The final study we report therefore takes a different tack on the issue of stability, and considers whether a strategy encouraged by a syllable-favouring language survives despite later syllable-disfavouring experiences. The experiment examines syllable-monitoring performance in speakers whose first language is Spanish but who have learned English later through immersion in an English-speaking community. For such speakers, the phonology of the native language typically remains dominant, at least in production. Omaya (1975), for example, comments that there are "age-related limits on the ability to acquire a second phonological system . . . and all the prosodic trimmings that go with it" (p. 275). Plausibly, what is true in production will be true for the comprehension strategies which by hypothesis are closely tied to the phonology, and thus bilingual speakers of this type provide a new way of evaluating the central claim of the earlier contrast of French and English. If syllabically based input representation is incorporated into the developing comprehension system through experience of a language in which syllables can be easily identified, and maintained as a perceptual strategy by continuing use of that language, Spanish–English bilinguals should not differ from monolingual speakers of Spanish; a syllabification effect should be evident for CV and CVC target monitoring in Spanish language materials.

EXPERIMENT 5

Subjects

The subjects were 24 bilingual speakers of Castillian Spanish and English drawn from the Spanish community of Melbourne, representing a broad age range (17–63 years, mean = 37 years). All of them had resided in an English-speaking country for a minimum of 3 years (mean = 18 years) but spoke Spanish as their first language; two-thirds had learned English post-puberty and had received all or part of their education in Spanish. Eleven subjects had received education to the tertiary level, seven to the secondary level and only six had not progressed beyond primary school. However, all of the subjects could read and write and did both things well, at least for Spanish.

A questionnaire surveying the emphasis put on Spanish and English in different domains (in the family circle, in broader social contexts and at work) revealed that although English was important socio-economically, Spanish remained the significant cultural vehicle. A question relating to language preference ("¿Se siente más cómodo usando Inglés, Español, o lo mismo con Inglés y Español?") was answered with the age-related bias which would be expected on sociolinguistic grounds—the subjects under the age of 30 tended to favour English as much as or more than Spanish—whereas patterns of language use were less determined by age. Self-ratings of competence in English suggested, for the large majority of the subjects, a reasonable to high level of competence in both speech comprehension and production, though some subjects had less confidence in reading and writing English.

Materials and Procedure

The experiment employed the Spanish language materials of Experiment 1, and exactly the procedures used with the monolingual Spanish speakers of that earlier experiment.

Results and Discussion

The mean monitoring reaction times of CV and CVC targets carried in words of different syllable structure are presented in Table 5. In contrast to our findings for Spanish monolingual speakers, these data for Spanish–English bilinguals showed no evidence of syllable-based input representation, even in their syllable-favouring native language; the interaction between target and carrier type was not significant [$F_1(1,22) = 1.04$, $P > 0.25$; $F_2(1,40) = 1.49$, $P > 0.10$].[7] And, unlike monolingual English speakers, there was no effect of target type [$F_1 < 1$; $F_2(1,40) = 2.27$, $P > 0.10$], the trend favouring CV targets (497 msec) over CVC targets (506 msec), if anything. An unexpected finding was an effect of carrier type, with detections being faster overall in CV-syllabified words (479 msec) than in CVC-syllabified words (524 msec) [min $F'(1,57) = 3.48$, $P < 0.10$]. This is the pattern found by Cutler et al. (1986) for English speakers and English materials, and taken by those authors as positive evidence of the use of phoneme-based input representations.

Rounding out our consideration of the evidence, we examined the extent to which patterns of language use and preference in Experiment 5's sample of Spanish–English bilinguals controlled the syllabification effect, exploiting the subjects' varied experience with English in correlational analyses. For each subject, an estimate of the critical effect can be achieved by calculating the average response time advantage of the target/carrier match conditions (CV and CVC targets for CV- and CVC-syllabified words, respectively) over the remaining mismatch conditions; this is exactly the

[7]Two separate replication experiments differed from Experiment 5 in using written rather than spoken targets, for a closer correspondence to the procedure of Mehler et al. (1981a). One tested subjects whose first language was Castillian, and the other—using a new recording of the stimulus materials—tested speakers of Uruguayan Spanish. All of the subjects in these replications had learned English post-puberty as a consequence of their emigration to Australia. In neither experiment did we observe the critical interaction which is the hallmark of syllabic routines; in both instances, $F < 1$ for the interaction term, in both the subject- and item-based analyses.

TABLE 5
Mean Monitoring Reaction Times for Spanish–English Bilingual Speakers and Spanish Materials

Item Type	*Example*	*Reaction Time (msec)*
CV carrier words		
CV targets	*pa/paloma*	469
CVC targets	*pal/paloma*	489
CVC carrier words		
CV targets	*pa/palmera*	524
CVC targets	*pal/palmera*	523

interaction term of the factorial design. These estimates of the syllabification effect can then be correlated with whatever questionnaire information plausibly might make one language more salient than the other. There were two directly quantitative measures, the age of learning English and the number of years of residence in an English-speaking country. If the extent of exposure to English controlled the probability that Spanish-style syllabification would be abandoned (crudely, earlier or more English leading to less syllabification), these measures of exposure should correlate with the tendency to syllabify, the first positively and the second negatively. For age of learning English, there was at most a trend in the direction predicted ($r = +0.245$, $P > 0.10$), and for the period of immersion in an English-speaking community, not even that ($r = +0.030$, $P > 0.25$).

The other ways of assessing experience with English were more qualitative, and related to a pattern of language use at home and in the community, the language which felt more natural (i.e. *se siente más cómodo*), the language in which education was received, and self-assessed competence in English comprehension, production, reading and writing. For these seven variables, questionnaire responses were quantified on a 3-point scale, under an arbitrary convention giving higher values to greater English experience. The left-hand side of Table 6 presents correlations between these qualitative variables with the tendency to syllabify, estimated for each subject, and shows that there was no statistically reliable relation between the magnitude of the syllabification effect and any of the language history variables which might measure the impact of English.

There are undoubtedly range restrictions limiting the power of this kind of test, but these alone cannot account for the null findings. The right-hand side of Table 6 presents correlations between the language use and language preference variables and the remaining five variables estimating

TABLE 6
Correlations (Left) Between the Size of the Syllabification Effect, for Individual Subjects, and Seven Language History Variables Estimating English Experience, and Correlations (Right) Between Language Use, Preference Variables and Self-assessed English Competence

	Size of Effect[a]	*Language Use*[b]	*Language Preference*[b]
Language use	+0.004		
Language preference	−0.091		
Language of education	−0.045	+0.589	+0.791
English comprehension	−0.186	+0.750	+0.461
English production	−0.025	+0.662	+0.428
English reading	−0.026	+0.652	+0.360
English writing	−0.234	+0.771	+0.673

Note: Critical value of r ($P < 0.05$, one-tailed) is 0.379; questionnaires were incomplete for four subjects.
[a]Predicted to be negative; [b]predicted to be positive.

competence in English. Here, all correlations were in the direction predicted (use of, and preference for the second language tracking with competence) and only one of the ten failed to reach significance. So although our quantification of experience with English cannot be called sophisticated, it is sufficient to support an internal validation. The correlations with the syllabification effect in Table 6 therefore stand, suggesting that the extent of English experience does not determine the degree to which Spanish ex-monolinguals hold to the strategies which are characteristic for their native language. If strategies developed in childhood are discarded only with some threshold exposure to a syllable-disfavouring language like English, that threshold has been passed in the 3-year residence period which we set as our floor.

If the subjects in Experiment 5 have detected CV and CVC targets on a basis altogether different from their monolingual counterparts, it cannot be the stimulus language itself which brings about the abandonment of syllabic routines, since Spanish is syllable-favouring by the evidence of Experiment 1. Nor can there be any problem with materials which are unfamiliar, since the language of the experiment was the native language of this subject group, whose day-to-day experience with English was never to the exclusion of Spanish. For these reasons, the modification revealed in the performance of Spanish–English bilinguals constitutes a striking confirmation of the instability suggested by the performance of monolingual Spanish speakers in Experiment 4. It seems that the representational basis on which responses are issued in the monitoring task is not a fixed aspect of a listener's comprehension system.

A recent report by Sebastián-Gallés, Dupoux, Seguí and Mehler (1992) makes essentially this suggestion, though for different reasons. With the now standard experimental design for this kind of investigation, they found different data patterns for Spanish speakers and Spanish materials, depending on the speed of response. Only when responses were slowed by an additional task requirement—the subjects were asked to pay attention to the semantic relations among words on a list—did the interaction pattern of Experiment 1 emerge to signal the availability of syllable-sensitive input representations as a basis for CV and CVC detections; in the normal, unloaded task when responses were fast, no such pattern was evident. The authors concluded that, for Spanish speakers, there is a transient stage of input processing at which sub-syllabic information is accessible, presumably before a syllabified representation has been fully established.[8]

Given the findings of Sebastián-Gallés et al., we must ask whether the difference in outcomes between Experiments 1 and 5 is properly attributed to the difference in language history which was our focus, or whether response speed is once more an issue; after all, monolingual subjects (interaction present) responded more slowly than did bilingual subjects (interaction absent). We note that the naturally arising speed contrast of our experiments (583 *vs* 501 msec) was markedly less than that which Sebastián-Gallés et al. induced by task variation (616 *vs* 377 msec), but this does little more than raise an imponderable question: How *much* difference in response speed is needed to shift the basis of response, on speed grounds alone? Our aim must be to establish whether the two subject groups belong to the same "speed class", language history differences aside. That aim would be achieved if we could offer an instance in which monolingual Spanish speakers showed sensitivity to proper syllabification in the monitoring task, and yet were faster in performance of the

[8]When very fast responses fail to produce an effect which is evident when responses are slower, an uninteresting interpretation is clearly invited—the subjects may simply not have complied with the experimental instructions. In this particular case, we should worry whether the subjects have in fact monitored for CV and CVC sequences in full, or whether they have downgraded the task to something like phoneme monitoring. The relevant guarantee of performance accuracy cannot be given in Sebastián-Gallés and co-workers' experiments, in the absence of catch trials; in their experimental materials, no filler item ever began with the initial consonant of the target sequence nominated for that trial. On this point, the authors draw comfort from the 10-msec trend (reliable in an item- but not a subject-based analysis) for responses to CV targets to be faster, overall, than to CVC targets.

task than either of the subject groups in Experiments 1 and 5. Fortunately, this is the case: Sánchez-Casas and García-Albea (1990) reported an experimental series with materials of the kind used here, and found, in two cases, the required interaction of target and carrier type even when reaction times were of the order of 450 msec.

Issues around response speed do not, therefore, complicate the comparison of monolingual and bilingual speakers, and we may conclude that differences in post-childhood language experience *per se* are responsible for the shift in experimental outcomes. A natural account of these data might make the assumption that, for a bilingual speaker whose languages differ in type, language processing requires a uniform approach to input analysis, and a non-syllable strategy is no more than the lowest common denominator. But a report from Cutler, Mehler, Norris and Seguí (1989) apparently rules this possibility out. In examining syllable sensitivity in bilingual speakers, they sampled subjects who had acquired both French and English in early childhood, and in adulthood spoke both languages regularly and with native fluency and accent; language-dominance subgroups were defined by responses to the question "If you had to lose one of your languages to save your life, which would you keep?" Critically, French-dominant bilinguals gave evidence of syllabic routines for French language materials but not for English language materials, and the authors interpreted these data in terms of a simultaneous command of different strategies. A syllabic routine is not necessarily discarded following experience with English.

Cutler and co-workers' English-dominant bilinguals, however, showed no syllabification effect with French materials, just like the Spanish–English bilinguals with Spanish in Experiment 5. Replication could be claimed only if the subjects in Experiment 5 (see also note 7) could be called English-dominant, in some sense other than a purely redescriptive one. And, setting aside the doubts outlined earlier about the informativeness of performances over materials which are absolutely foreign, a parallel could be drawn between Spanish *monolinguals* and Cutler and co-workers' French-dominant bilinguals; both groups sometimes syllabify (Spanish, French materials) and sometimes do not (English materials).

Given these parallels, the conclusion which seems appropriate is a woefully moderated one, and that is that Spanish is a language encouraging syllabification but not to the exclusion of phonemic analysis, so that from the outset an alternative routine is a viable part of the listener's repertoire. In the monolingual, the balance must favour a syllabified representation as the basis for response in the monitoring task, since our data for monolingual Spanish speakers and Spanish materials have shown the clearest syllabification effect yet reported in this kind of work. In the bilingual, the balance might be said to have tipped. It cannot be questions of the relative efficiencies of syllable- and phoneme-based routines which are at issue in setting that balance; plausibly, a perceptual routine has whatever efficiency it has over a given set of materials. What seems more likely to differ between speaker populations is the availability to conscious report of input representations of different kinds.

GENERAL DISCUSSION

In this attempted replication of the ground-breaking work of Mehler et al. (1981a) and Cutler et al. (1983; 1986), an underlying assumption has been that Spanish, like French, is a language likely to promote syllable-based input representation. The phonological properties of French which lend plausibility to the general claim—the prototypicality of the simple CV syllable and the clarity of syllable boundaries—apply equally to Spanish, by available analyses, and indeed our initial finding of a robust syllabification effect for Spanish speakers and Spanish materials strongly supports earlier findings. But in one crucial respect there has been failure to generalise: monolingual Spanish speakers apparently abandon syllabification when dealing with English materials, while French speakers do not.

In the framework established by the researchers whose work we have followed, two possibilities suggest themselves. Spanish may differ from French in some critical way (e.g. in having lexical stress), in which case we have yet to specify what aspects of a language's phonology determine whether a syllable-based routine will be the predominant form of input analysis, and why those aspects are critical. Alternatively, for French as much as for Spanish, the efficiencies of routines of different types might not uniformly favour syllabic over phonemic, and we have yet to specify what properties of spoken inputs underlie the application of one or the other, on particular occasions. Surveys across a variety of languages are unlikely to be useful since, as we have seen, native language performance in monolinguals does not ordinarily reveal a possible instability in the application of a syllabic routine.

It seems that progress is most likely to come if future research is directed at spelling out just how a syllabic routine operates. If, as Mehler (1981) has it, the syllable is the output emitted by a device that scans the acoustic signal, what does the device scan for? Are there acoustic properties that functionally mark syllable boundaries, so that it is juncture detection which drives the perceptual routine, bottom-up (cf. Norris & Cutler, 1985)? Or do we reverse figure and ground, and suppose that the scan of the signal seeks matches against an inventory of syllabic templates, with boundaries located as a by-product of an exhaustive syllabification? Either approach is compatible with previous researchers' portrayal of the syllabic routine as a segmentation strategy.

Real detail here would go some way to resolving the problems raised by our attempted replication. In particular, if context-free cues to syllable boundary exist, there could be some resolution of the controversy we have outlined in the analysis of English; in our view, the contrast of the phonologies of French and English is by no means as clear as it has been taken to be. Are syllable boundary cues in fact missing or muddy in English, as has been stipulated, or are they simply differently placed? As the remaining half of our replication has demonstrated, the performance data for English speakers cannot itself sort out the difficulty. The experiments have suggested that different materials selection controls the appearance of one effect (faster detection responses in words beginning CVCV) interpretable as evidence of a routine pitched at the level of phonemes, and the appearance of another (faster detections of CVC targets) possibly interpretable as evidence of a syllabic routine. And just this kind of detail is required to set aside the possibility that syllabifications imposed top-down might give rise to the data patterns for foreign language materials which were the basis of Cutler and co-workers' (1983; 1986) claim that a language-specific segmentation routine is a property belonging to listeners, not to experimental situations.

Whether or not these points of conflict can be sorted out by further research, the data we have reported for bilingual speakers of Spanish and English present a different kind of problem. If monolingual speakers' abandonment of syllabification strategies is temporary, or to be explained in terms of the processing load induced by an altogether unfamiliar language, the same is not true of bilinguals. The evidence we have presented suggests that an exposure to English which leads to its acquisition as a second language modifies the approach to input representation (at least, as it is revealed in the monitoring task), even for native materials. The phonology of the languages experienced thus has an influence beyond establishing perceptual strategies in the developing comprehension system of the newborn. In the end, the critical question must be whether it is perceptual strategies that are labile, or whether performance differences arise from factors whose significance is only task-specific.

Certainly, metalinguistic skill is an important contributor to performance in the monitoring task. For example, Morais et al. (1989) have shown, in their study of Portuguese-speaking illiterates and ex-illiterates (those learning to read only as adults), that literacy brings dramatic increases in the success with which sentences can be monitored for CV and CVC targets, and attribute these changes to greater awareness of the internal structure of words. Presumably, it is by a special kind of attentional distortion that the perceptual products which are unavailable in normal process come into focus. There is no suggestion that metalinguistic skill, overall, differs between monolingual Spanish speakers and Spanish–English bilinguals; but metalinguistic awareness is no monolith, any more than the phonology of a language is. Both have separate but intricately cross-linked levels of analysis. It may be, therefore, that a late acquisition of English in a native Spanish speaker does no more than boost the availability of a phonemic representation as a basis for monitoring responses, relative to a syllabic representation. Carried with this is the implication that the patterns revealed in the monitoring task are set by something other than the natural course of language processing, so that the appearance of language specificity in segmentation strategies could be a mistaken attribution of the source of experimental effects.

Manuscript received April 1991
Revised manuscript received December 1992

REFERENCES

Agard, F.B. (1984). *A course in Romance linguistics. Vol. 1: A synchronic view*. Washington, DC: Georgetown University Press.

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

Borowsky, T.J. (1986). *Topics in the lexical phonology of English*. Unpublished doctoral dissertation, University of Massachusetts.

Cutler, A., Mehler, J., Norris, D. & Seguí, J. (1983). A language-specific comprehension strategy. *Nature*, *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. & Williams, J.N. (1987). A note on the role of phonological expectations in speech segmentation. *Journal of Memory and Language*, *26*, 480–487.

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

Dauer, R.M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, *11*, 51–62.

Delattre, P. (1966). A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics*, *4*, 183–198.

Forster, K.I. & Dickenson, R.G. (1976). More on the language-as-a-fixed-effect fallacy: Monte Carlo estimates of error rates for F_1, F_2, F', and min F'. *Journal of Verbal Learning and Verbal Behavior*, *15*, 135–142.

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

Harris, J.W. (1983). Syllable structure and stress in Spanish: A nonlinear analysis. *Linguistic Inquiry Monographs*, *8*.

Howell, P. & Darwin, J.C. (1977). Some properties of auditory memory for rapid format transitions. *Memory and Cognition*, *5*, 700–708.

Kahn, D. (1976). *Syllable-based generalizations in English Phonology*. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Klatt, D. H. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, *7*, 279–312.

Liberman, A.M. & Studdert-Kennedy, M. (1978). Phonetic Perception. In R. Held, H. Leibowitz & H.L. Teuber (Eds), *Handbook of sensory psysiology: Vol. VIII: Perception*. Heidelberg: Springer.

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., Fischer, F.W. & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, *18*, 201–212.

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

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

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

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

Mills, C.B. (1980). Effects of the match between listener expectancies and coarticulatory cues on the perception of speech. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 528–535.

Morais, J., Cary, L., Alegria, J. & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, *7*, 323–331.

Morais, J., Cluytens, M. & Alegria, J. (1984). Segmentation abilities of dyslexics and normal readers. *Perceptual and Motor Skills*, *58*, 221–222.

Morais, J., Bertelson, P., Cary, L. & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition*, *24*, 45–64.

Morais, J., Content, A., Cary, L., Mehler, J. & Seguí, J. (1989). Syllabic segmentation and literacy. *Language and Cognitive Processes*, *4*, 57–67.

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

Omaya, S. (1975). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, *5*, 261–283.

Prunet, J.-F. (1987). Liaison and nasalization in French. In C. Neidle & R.A. Nuñez Cedeño (Eds), *Studies in romance languages*. Dordrecht: Foris.

Sánchez-Casas, R.M. (1988). *Access representations in visual word recognition*. Doctoral dissertation, Monash University, Australia.

Sánchez-Casas, R.M. & García-Albea, J.E. (1990). The syllable monitoring task: How do the subjects respond? Paper presented to the *4th Conference of the European Society of Cognitive Psychology*, Como, Italy.

Savin, H.B. & Bever, T.G. (1970). The nonperceptual reality of the phoneme. *Journal of Verbal Learning and Verbal Behaviour*, *9*, 295–302.

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

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

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

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.

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

APPENDIX 1

Spanish carrier pairs in Experiments 1, 3 and 5 (Note: in all cases, stress falls on the penultimate syllable)

balada/baldosa	coraje/corbata	camisa/campana
paloma/palmera	corona/corteza	limosna/limpieza
bolero/bolsillo	joroba/jornada	casino/castillo
volante/voltaje	morera/mortero	peseta/pestaña
culebra/cultura	muralla/murmullo	cosecha/costilla
garaje/garganta	ceniza/censura	rosario/rosquilla
cereza/cerveza	linaje/linterna	
pereza/perfume	moneda/montaña	

APPENDIX 2

English carrier pairs in Experiments 2 and 4 (Note: in all cases, primary stress falls on the initial syllable and the following syllable is unstressed)

balance/balcony	willow/wilderness	senator/sentence
calendar/calcium	dollar/dolphin	tenement/tendency
palace/palpitate	canopy/cancel	vinegar/vindicate
salad/salvage	lanolin/lantern	tonic/tonsil
talon/talcum	sanity/sandal	hammer/hamper
helicopter/helmet	menace/mention	similar/simple
pelican/pelvis	penalty/pencil	massacre/masculine
filament/filter	penetrate/pendulum	desecrate/destitute