

EFFECTS OF SYLLABIC POSITION IN THE PERCEPTION OF SPOKEN ENGLISH

Athanassios Protopapas, Steven Finney, and Peter D. Eimas
protopap@cog.brown.edu, sf@cog.brown.edu, eimas@cogvax.cog.brown.edu
Dept. of Cognitive & Linguistic Sciences
Box 1978, Brown University, Providence, RI 02912
U.S.A

ABSTRACT

We report a series of experiments that demonstrate an effect of syllabic position on reaction time during phoneme monitoring in English. In our stimuli, the third phoneme of each word was the target phoneme, and it belonged either to the coda of the first syllable (coda-type target) or to the onset of the second syllable (onset-type target). By manipulating the proportions of the target types, reaction times to targets of the expected syllabic structure were significantly affected. This effect was only present when words stressed on their second syllable were used; words with first-syllable stress yielded no effect of expectation. The difference may be attributed to the strongly ambisyllabic nature of consonants following stressed vowels. The results are consistent with the idea that prelexical syllabic segmentation always occurs, but is only evident when the syllabic boundaries of the materials are unambiguous. These and previous findings are discussed from a cross-linguistic perspective.

1. INTRODUCTION

An important question in speech perception research concerns the segmentation of the speech signal prior to lexical access [7, 14]. Various kinds of units that access the mental lexicon have been proposed, including abstract linguistic entities such as syllables and phonemes. Earlier research has shown that syllabic structure affects response time in that targets that match whole syllables are responded to faster than targets that match parts of syllables or that span syllabic boundaries. This effect is clear in French and Catalan, less clear in Dutch and Spanish, and totally absent in British and Australian English [2, 5, 9, 13, 16]. In fact, cross-language experimentation with French and British English speaking subjects and with bilinguals has led to the conclusion that French speakers employ a syllabic strategy in speech perception whereas English speakers do not [3, 4, 6]. This difference was attributed to differences in the “clarity” of syllabification in English and French.

An important aspect of these studies concerns the syllabic structure of the stimuli in the experimental conditions. In French, the CV/CV₁ word “balance” and the CVC/CV₁ word “balcon” exemplify open and closed syllables. The corresponding words “balance” and “balcony” have been used in the English experiments. Although “balcony”-type words have the same CVC/CV₁ syllabic structure as do corresponding French words, the English word “balance” and other words of the same form

do not have the French CV/CV₁ structure. Rather, they have a CV[C]V₁ structure, where the intervocalic consonant [C] is ambisyllabic, i.e., belonging to both the first and the second syllable [1, 8, 15]. Therefore the first syllable of the English word “balance” may be “bal” with the /l/ clearly belonging to it (and thus same as the first syllable of “balcony”), or the syllable may show an unclear or uncertain membership of the /l/.

Cutler *et al.* hypothesized that if ambisyllabic boundaries are unclear, then a syllabic segmentation strategy might be inefficient, a view supported by their finding of syllabic effects in French but not in British English. They concluded that language-specific segmentation strategies may be used to cope efficiently with the particular structures of languages [3, 5]. However, there are alternative interpretations of their findings. It is possible that failure to find a syllabic effect in English may be due to the relation of syllabic effects to stress. Since language has been confounded with stress pattern (with English words stressed on their first syllable and French words stressed on their final syllable), stimuli with matched stress patterns might possibly result in more nearly comparable results (cf. unstressed syllable effect in Catalan [13]). Alternatively, ambisyllabic segments may clearly belong to both syllables, as Cutler *et al.* noted. If so, syllabic segmentation would arguably produce closed syllables in both conditions (“bal” would unambiguously be the first syllable of “balance” and of “balcony”) and no effect of word type would be found because there is no difference between the two types to produce a mismatch in one case but not the other.

In more recent experiments, the attention of listeners was focused on particular sublexical units by inducing expectations about the syllabic position of a target phoneme [11, 12]. If it is possible to attend to a subsyllabic unit, it is reasonable to conclude that this unit and the syllable to which it belongs have psychological reality. If the effects of such attention are faster than lexical access and are present with nonword stimuli, it can be further concluded that the unit in question is derived prelexically, and perhaps en route to lexical access. In experiments where expectations about the position of the target phonemes were manipulated, effects of position within a syllable were shown in French and in Spanish [11] and it was claimed that this procedure was more sensitive to syllabic effects than the older syllable monitoring paradigm.

In the following we report a series of experiments where the effects of syllabic position are investigated using the procedure of Pallier *et al.* [11], first with words stressed on their initial syllable, as most English words are, and then with words stressed on their second syllable. We predicted that there would be little or no syllabic effect with the former because of the ambisyllabic consonant that follows stressed syllables. In contrast, we predicted a syllabic effect with the latter because of the clearer syllabification following unstressed syllables.

¹Throughout this report, C stands for consonant and V for vowel. The slash (/) indicates a syllabic boundary and [C] indicates an ambisyllabic consonant. An underscore () denotes arbitrary continuation of the word.

2. EXPERIMENT 1: FIRST-SYLLABLE STRESS

In this experiment two groups of listeners monitored English words stressed on their initial syllable for the occurrence of phonemes. Target phonemes always occurred in the third (serial) position in the word, but belonged to either the coda of the first syllable only (e.g., /p/ in “captive”) or to both the coda of the first syllable and the onset of the second one (e.g., /b/ in “fabric”). Each subject’s experimental word list contained a higher proportion of words with one of the two possible syllabic structures so that, if attention was implicitly focused to syllabic position, reaction times would be facilitated for the words with the most frequent structure. An important aspect of this design is that reaction times to the same target phonemes in the same words are compared across subjects and the experimental manipulation involves only the contextual stimuli that serve to induce expectation.

2.1. Methods

2.1.1. Subjects

Twenty Brown University Students participated in this experiment, ten in the ONSET and ten in the CODA induction condition (see *Stimuli*). An additional set of ten subjects were presented with the CONTROL experimental list for a baseline test of the TARGET items and in order to provide a means for a cost and benefit analysis. However, costs and benefits are not considered in this report.

2.1.1. Stimuli

Three sets of words beginning with a CVCC_ sequence and stressed on their initial syllable were constructed. In each set, half of the words had a closed first syllable (coda-type words, e.g., “magnet”) and half had a first syllable with an ambisyllabic coda segment that was also the onset of the second syllable (onset-type words, e.g., “juggler”). The TARGET set contained 32 words (16 coda-type and 16 onset-type) that had a labial or velar stop consonant in the third position. The INDUCTOR set contained 100 words with no restriction on the consonant of the third position (50 onset-type, e.g., “zebra” and “custom,” and 50 coda-type, e.g., “kidney” and “normal”). The DISTRACTOR set contained 32 words (16 onset-type and 16 coda-type), also without restriction on the third-position consonant. A target phoneme was selected for each word so that it occurred in the third position for words in the TARGET and INDUCTOR sets and did not occur in any position for words in the DISTRACTOR set.

A list of all 164 words, in random order, was recorded by a male speaker of American English and sampled into a computer at 20 KHz using 12-bit quantization. For each TARGET word, a digital marker was positioned to coincide with the noise burst that signified the closure release of the target phoneme (always a stop consonant).

Three different experimental lists were constructed, each containing all 32 TARGET and 32 DISTRACTOR items, but only half the INDUCTOR items, as follows: The ONSET experimental list contained all 50 onset-type and no coda-type INDUCTOR items, the CODA list contained all 50 coda-type and no onset-type INDUCTOR items, and the CONTROL list contained 25 onset-type and 25 coda-type INDUCTOR items. Thus

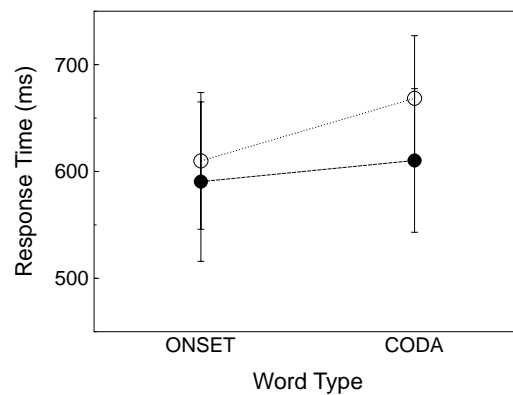


Figure 1. Results of Experiment 1 (with first-syllable stress words): Reaction times to the two word types in the ONSET (○) and CODA (●) induction conditions. Error bars show standard error.

each experimental list comprised 114 trials and the percentage of positive-response onset-type words was 80% for the ONSET list, 50% for the CONTROL list, and 20% for the CODA list. The order of the items within the three lists was randomly chosen with the restrictions that no TARGET items occur in the first eight trials (the initial induction and warm-up period) and that at least one INDUCTOR precede each TARGET item (following Pallier *et al.*).

2.1.2. Procedure

Subjects were seated in front of a computer monitor wearing headphones. Each trial started with a 1 s visual presentation of a letter that indicated the sound the subject was to listen for. Explicit instructions were given to the subjects to ensure phonemic and not orthographic monitoring. The screen was cleared and 500 ms later a word was presented over the headphones. The subject had to press a button labeled “YES” if the sound that was represented by the letter was present in the word or a button labeled “NO” if it was not. Subjects were instructed not to wait until the end of the word if they heard the sound. The time between the target marker (at stop release) and the subject’s response was measured and stored for analysis. Each subject heard only one of the experimental lists. The entire experiment (114 trials) lasted about 15 minutes.

2.2. Results

Only responses to TARGET words in the ONSET and CODA induction conditions are considered here. Incorrect responses (5%) were excluded from the analyses. Of the correct responses, those with reaction times less than 100 ms or greater than 2 s (0.8%) were also excluded. The remaining data are shown in Figure 1. In a 2×2 analysis of variance (Induction condition × Word type) there were no significant main effects or interactions, indicating that syllabic position had no effect, in agreement with results previously reported with English speakers.

3. EXPERIMENT 2: SECOND-SYLLABLE STRESS

The unclear syllabification (or clear ambisyllabicity) of the target segment in Experiment 1 makes it difficult to argue conclusively against the occurrence of syllabic effects. It might be the case

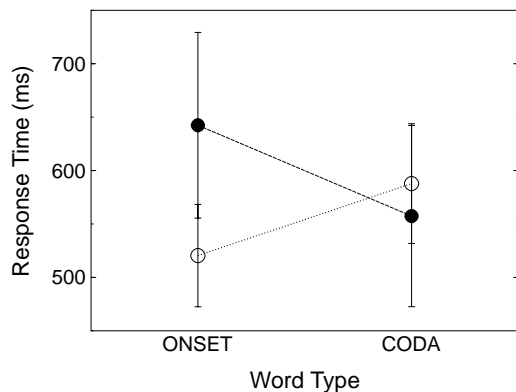


Figure 2. Results of Experiment 2 (with second-syllable stress words): Reaction times to the two word types in the ONSET (○) and CODA (●) induction conditions. Error bars show standard error.

that expectation is indeed induced, always to the coda of the first syllable, but since in the onset-type words the target phoneme also belongs to the coda of the first syllable no differences can be detected. In the second experiment we used words stressed on their second syllable; these have clearer syllabification, so the target phonemes unambiguously belong to either the coda of the first syllable (coda-type words, e.g., “segmental,” “narcotic”) or the onset of the second (onset-type words, e.g., “seclude,” “bestow”). A lack of syllabic effects in this experiment would be strong evidence for language-dependent models of prelexical syllabification. If, however, an effect of syllabic position is found, then a syllabic segmentation strategy may be hypothesized that holds in all cases but is only evident when the syllabic boundaries of the materials are unambiguously marked.

3.1. Methods

The methods of Experiment 1 were used, except that words stressed on their second syllable were chosen to construct the lists. The same constraints were applied in the selection of the words. There were again a TARGET set and a DISTRACTOR set (each with 16 onset-type words and 16 coda-type words) and an INDUCTOR set (with 50 onset-type words and 50 coda-type words). Three experimental lists were constructed using in each all the TARGET and DISTRACTOR items, and half of the INDUCTOR items, as previously. Each subject was tested in one induction condition only.

3.2. Results

Only responses to TARGET words in the ONSET and CODA induction conditions are considered here. Incorrect responses (3%) were excluded from the analyses. Of the correct responses, those with reaction times less than 100 ms or greater than 2 s (0.5%) were also excluded. The remaining data are shown in Figure 2. Note the crossing of the two lines, indicating the relative decrease of reaction time in the expected word type (or relative increase of reaction time in the unexpected word type). A 2×2 analysis of variance (Induction condition × Word type) showed no significant main effects. However, the interaction between induction condition and word type was highly significant both by subjects and by items. ($F_1(1,18)=17.11, p<0.001$; $F_2(1,30)=19.29, p<0.001$). Investigating simple effects, we found that onset-type words were responded to faster than coda-

type words in the ONSET induction condition ($F_1(1,9)=5.68, p=0.04$; $F_2(1,30)=3.89, p=0.06$), but slower in the CODA induction condition ($F_1(1,9)=12.96, p=0.006$; $F_2(1,30)=8.03, p=0.008$).

There was clearly an induction of expectation to a position within the word that cannot simply be serial order. Although our manipulation involved only syllabic position, alternative interpretations are possible and must be excluded before we can confidently conclude that syllabification occurred.

4. ALTERNATIVE EXPLANATIONS

In order to accept the syllabic segmentation hypothesis on the basis of Experiment 2, it must be established that the syllabic position caused the observed differences between the ONSET and CODA induction groups. One possible confounding factor turned out to be the time from word onset to the target mark. Measurements indicated that the target phonemes in the onset-type words of Experiment 2 occurred about 40 ms earlier than target phonemes in the coda-type words. Therefore, subjects may have been attuned to the time a target phoneme was expected to occur following word onset. In order to investigate this possibility, all TARGET and INDUCTOR items were edited to equate the lengths of the first syllables. A waveform editor was used to repeat voiced periods and noise segments in the onset-type words and to excise corresponding acoustic segments in the coda-type words. In a control experiment subjects were unable to discriminate between the original and the spliced stimuli ($d' < 0.5$). We repeated Experiment 2 using the edited stimuli and found the same significant interaction between Induction condition and Word type, thus rejecting the temporal induction hypothesis (detailed results of this and the next experiment may be obtained from the authors).

An additional possible concern is whether the demonstrated syllabic segmentation occurs postlexically, i.e., after lexical access. Because of the relatively long reaction times in our experiments, compared to those reported from previous studies, it is reasonable to question the immediate, i.e., prelexical, availability of the syllabic structure evidenced by the data. An alternative interpretation would involve lexical access first and then syllabic and phonemic analysis that would then be used to match the target. Clearly, such a strategy would be impossible with non-words, i.e., phonotactically acceptable strings that do not exist in the listeners’ lexicon and are thus not available for lexical access (and postlexical syllabic segmentation). To test this, new TARGET and DISTRACTOR sets were constructed by altering the initial segments of the words used in Experiment 2 to create nonwords. Using the same procedure, we again found a significant interaction between item type and induction condition, strongly suggesting that the significant effects in Experiment 2 were not postlexical. Additional support comes from analysis of the data of the fastest subjects in all experiments, which showed a significant interaction between Word type and Induction condition with response times much lower than those generally considered necessary for lexical access.

5. GENERAL DISCUSSION

In the experiments reported here it was shown that, using American English materials, subjects’ attention can be focused on syllabic position by inducing expectations about the position of the

target phoneme. This expectation is created by the statistical properties of the task and not by any explicit mention of syllables in the instructions or elsewhere. In fact, the instructions given to the listeners might have led them to adopt phoneme-based strategies; the syllabic effect we found can only be explained if a speech parsing strategy is hypothesized that segments speech into syllables prior to other analyses. Evidence for such a strategy has already been provided in other languages, such as French, Catalan, Dutch, and Spanish. The absence of evidence for this strategy in English was previously attributed to the clarity of syllabification of those languages as opposed to a more “obscure” syllabification in English.

However, the lack of syllabic effects with English materials in the literature may have been due to insensitivity of the procedure or a confound with word stress. In light of the present results, a replication of the previous studies is required to resolve the matter, using materials with a wide range of syllabic (and stress) characteristics.² In addition to using second-syllable stress words, ambisyllabicity may be investigated by contrasting words with “long” vs. “short” vowels in their first syllable. The critical segments in the former (e.g., /k/ in “microbe”) are believed to be less ambisyllabic than those in the latter (e.g., /k/ in “macro”) [10].

We argue that the low-level speech segmentation strategies operating in other languages are also those used by English speakers and that differences in speech perception are probably not fundamental differences in processing but, rather, a function of physical (acoustical) language-specific properties in the production of the signal. Why would some languages require one strategy and others a different one? The multitude of today’s languages have evolved from some common ancestor language and are spoken by the same kind of human beings, so it is at least arguable that the same basic mechanisms are involved. Moreover, all languages have syllables that conform to a very specific (if in part language dependent) structure. Syllables not only make speech possible (how can you utter a stop consonant without a vowel?), their rigid structure may also guide perception by providing the necessary constraints for parsing. Given the universality of the syllable, it is reasonable that its nature and role in speech perception and language in general are also universal.

To conclude, it is clear that prelexical syllabic effects can be obtained in the perception of English if the experimental materials are designed to make unambiguous syllabic distinctions between the experimental conditions. A language-independent speech processing strategy hypothesis is more parsimonious than a host of language-specific proposals and should be preferred unless direct evidence can be provided against it. Further research is certainly required to resolve the issue; such research should first identify the basic, general, mechanisms, and only later look into the particulars of each language.

Acknowledgements

We thank Sarah McKenna and Carolyn Healy for testing subjects, John Mertus for help with the software, and Emily R. Pickett for comments on the manuscript.

²Such studies are currently under way in our laboratory; preliminary results indicate that prelexical syllabification occurs but further controls are necessary.

6. REFERENCES

- [1] J. Anderson and C. Jones. Three theses concerning phonological representations. *Journal of Linguistics*, 10:1–26, 1974.
- [2] D. C. Bradley, R. M. Sánchez-Casas, and J. E. García-Albea. The status of the syllable in the perception of Spanish and English. *Language and Cognitive Processes*, 8:197–233, 1993.
- [3] A. Cutler, J. Mehler, D. Norris, and J. Seguí. A language-specific comprehension strategy. *Nature*, 304:159–160, 1983.
- [4] A. Cutler, J. Mehler, D. Norris, and J. Seguí. Limits on bilingualism. *Nature*, 340:229–230, 1989.
- [5] A. Cutler, J. Mehler, D. Norris, and J. Seguí. The syllable’s differing role in the segmentation of French and English. *Journal of Memory and Language*, 25:385–400, 1986.
- [6] A. Cutler, J. Mehler, D. Norris, and J. Seguí. The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology*, 24:381–410, 1992.
- [7] E. Dupoux. The time course of prelexical processing: The syllabic hypothesis revisited. In G. T. M. Altmann and Richard Shillcock, editors, *Cognitive Models of Speech Processing: The second Sperlonga meeting*, pp. 81–114. Lawrence Erlbaum, Hove, UK, 1993.
- [8] D. Kahn. *Syllable-based generalizations in English phonology*. Outstanding dissertations in linguistics. Garland Publishing, New York, 1980.
- [9] J. Mehler, J. Y. Dommergues, and U. Frauenfelder. The syllable’s role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, 20:298–305, 1981.
- [10] S. Myers. Vowel shortening in English. *Natural Language and Linguistic Theory*, 5:485–518, 1987.
- [11] C. Pallier, N. Sebastián-Gallés, T. Felguera, A. Christophe, and J. Mehler. Attentional allocation within the syllabic structure of spoken words. *Journal of Memory and Language*, 32:373–389, 1993.
- [12] M. A. Pitt and A. G. Samuel. Attentional allocation during speech perception: How fine is the focus? *Journal of Memory and Language*, 29:611–632, 1990.
- [13] N. Sebastián-Gallés, E. Dupoux, J. Seguí, and J. Mehler. Contrasting syllable effects in Catalan and Spanish. *Journal of Memory and Language*, 31:18–32, 1992.
- [14] J. Seguí, E. Dupoux, and J. Mehler. The role of the syllable in speech segmentation, phoneme identification, and lexical access. In Gerry T. M. Altmann, editor, *Cognitive models of speech processing*, pp. 263–280. The MIT Press, Cambridge, MA, 1990.
- [15] R. Treiman and C. Danis. Syllabification of intervocalic consonants. *Journal of Memory and Language*, 27:87–104, 1988.
- [16] P. Zwitserlood, H. Schriefers, A. Lahiri, and W. van Don-selaar. The role of syllables in the perception of spoken Dutch. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 19(2):260–271, 1993.