DRAFT DO NOT QUOTE
Submitted to JEP:LMC

Are Stress Units Used in Prelexical Processing in English?*

EMMANUEL DUPOUX
France Telecom, Paris & EHESS-CNRS, Paris.

MICHAEL HAMMOND
University of Arizona

August 18, 1994

*Part of this work was done when the first author was a reseach fellow at the University of Arizona and when the second author was an invited scientist at CNRS-EHESS. We thank the Human Frontier Science Program, CNET, Fyssen Foundation, Inria, CNRS, NIH, CIES and the Cognitive Science program at the University of Arizona for financial support. We thank Ken Forster for useful comments and access to testing facilities. We thank Tom Craig, Diane Meador, and Diane Ohala for discussion and assistance with the experiments, and members of the psycholinguistics lab at UA and at CNRS for useful comments on several aspects of this work. We especially thank Anne Cutler, Jacques Mehler, Anne Christophe, and several anonymous reviewers for constructive comments on earlier versions of this paper.

Address for correspondence: LSCP CNRS-EHESS, 54 Boulevard Raspail, F-75270 Paris CEDEX 06, France. E-mail: dupoux@lscp.msh-paris.fr

Abstract

Previous investigations have claimed that speech perception uses language-specific strategies and that, in particular, English does not use a strategy based on syllables (Cutler, Mehler, Norris, and Segui, 1983; 1986). This conclusion is based on a failure to replicate the interaction between target type and word type in *fragment monitoring* experiments that was originaly found in French (Mehler *et al.*, 1981) with English subjects and materials. Here, we explore the possibility that this might be due to one of three related hypotheses: i) syllable boundaries in English depend on the stress value of the following syllable; ii) English listeners use the foot instead of the syllable in speech perception; iii) English subjects posit a perceptual boundary before an unreduced (or "*strong*") syllable but not before a reduced, or "*weak*" one. These three hypotheses can all be tested on the basis of the same contrasts and so we group them together under the rubric of "stress-sensitive strategies" (SSSs). In Experiment 1, we find some incomplete support for a SSS, but the effect is not replicated in subsequent Experiments 2 (with slowed-down subjects) and 3 (with a different set of materials). An associated off-line task (Experiment 4) reveals that, according to subjects' intuitions, syllables have a rather different structure than we assumed at the outset. We conclude by rejecting these SSSs as the major source of the difference between French and English. In the final section, we discuss the possibility that the English perceptual system might still be based on the syllable, but not a stress-sensitive one.

How do we break up the continously varying speech stream into discrete units? Early investigations proposed a universal or preferred perceptual strategy based on syllables (Mehler, Dommergues, Frauenfelder and Segui, 1981; Mehler, 1981; Segui, 1984). For example, it was claimed that a word in French like *balance* is broken up into two parts, or syllables, *ba* and *lance* in the course of speech perception. Thus in "fragment monitoring" experiments, French subjects respond more quickly that the fragment *ba* occurs at the onset of a word like *ba.lance* than they do with the fragment *bal*. Conversely, they respond more quickly that *bal* starts *bal.con* than *ba*, where the syllables are *bal* and *con* (Mehler *et al.*, 1981).

This original study has since been replicated in a number of other languages, e.g. Spanish (Sanchez-Casas, 1988; Sebastian *et al.*, 1992), Catalan (Sebastian *et al.*, 1992) and Dutch, (Zwitserlood, Schriefers, and Lahiri, 1993). All these results, plus others with newborns (Bertoncini and Mehler, 1981; Bijeljak-Babic, Bertoncini and Mehler, 1993) suggest that our perceptual system naturally segments the speech stream into syllable-sized chunks. Such a segmentation would be consistent across languages and would also help young infants to perceive and internalize speech before they start to acquire words of the ambient language (Mehler, Dupoux, and Segui, 1990).

However, English is a clear and unequivocal problem for this rosy picture. Fragment monitoring experiments with English-speaking subjects have consistently failed to replicate syllabic effects using materials superficially similar to the ones used in French (Cutler *et al.*, 1983; 1986; Bradley *et al.*, 1993). Moreover, English subjects fail to display syllabic effects when they are tested on French materials, even when they are highly proficient French-English bilinguals (Cutler *et al.*, 1989; 1992). These results in English force one to the conclusion that different languages use very different perceptual strategies. French, Catalan, Spanish, and Dutch use syllables, but English uses a very different but still unknown unit. Along the same lines, Japanese, a "mora-timed" language, has been recently argued to use yet another kind of unit based on morae (Otake *et al.*, 1993). On this view, the processing architecture put to use in speech perception could vary rather substantially according to the phonological/rythmic properties of the language (see Cutler and Mehler, 1993).

The fact that contrasting results are found across languages is uncontroversial. However, what is the difference between English and French that gives rise to these results? What processing units are really used in English? The direction that we explore in this paper is that, unlike in French, stress plays a very important role in English. There are a number of reasons for believing that this should be so. First, unlike in French, English words display different stress patterns. Second, English has widespread vowel reduction that correlates with stress. Finally, English phrasal timing is widely taken to be based on stress–English is described as a "stress-timed" language–whereas French is described "as syllable-timed".

In this paper, we consider three hypotheses that we group together under the rubric of "stress-sensitive strategies" (SSSs). They all hold that the perceptual chunking that is done on English words depends on their stress pattern, or on a factor that correlates with stress. The hypotheses are the following: a) English uses syllables as in French, but the syllable boundaries depend on stress (stress-sensitive syllables), b), English uses a unit similar to the metrical foot (discussed below), and c) English uses units that are defined in terms of unreduced versus reduced syllables.

Stress-sensitive Syllables

The first hypothesis is that English uses syllabic units just like French, but that these units are defined differently. In French, syllables are subject to the Maximal Onset Principle (Pulgram, 1970; Anderson, 1982) which states that other things equal, intervocalic consonants tend to affiliate to the right (onset position) rather than to the left (coda position), provided they do not create ill-formed syllables[1]. So, *balance* would be syllabified [ba.lãs], *metro*, [me.tro], and *balcon*, [bal.kõ] (since /lk/ is not a legal syllable onset). Such a parsing into syllables is congruent with naive intuitions and with the results obtained in the fragment detection studies (Mehler *et al.*, 1981).

[1] The wellformedness of a syllable is independently determined by the possible clusters at word edges.

In their original study, Cutler *et al.*(1983) used English materials very similar to the ones used in the French study: word pairs like *balance/balcony*, which are like the ones used in French (*balance/balcon*) in terms of the vowel quality in the first syllable, the identity of the immediately following consonant, and in terms of the supposed syllable structure. If English subjects used maximal onset syllables like in French, there should have been an effect just like in French. The absence of effect with these materials means one of two things: a) English does not use the syllable as a processing structure (the interpretation of Cutler *et al.*, 1983), b) English subjects use syllables, but syllables that depart from maximal onset[2].

We want to explore the second alternative. This is warranted by the fact that, first, it is quite clear that the syllable plays a role in the phonology of English (Hooper, 1972; Kahn, 1976; Borowsky, 1986). Moreover, a great deal of evidence suggests that something like a syllable is accessible to subjects' intuitions in metalinguistic games (Treiman, 1983; Treiman and Danis, 1988; Derwing, 1992; among others). The question then becomes: what is the definition of syllabic structure relevant to this language? In the linguistic literature, there are accounts of the English syllables that depart from the Maximal Onset Principle. They hold that in VCV sequences the intervocalic consonant gets drawn into the first syllable when the second vowel is stressless (Kahn, 1976; Selkirk, 1982), which is precisely the case of words like *balance*. The difference between these theories is that for Selkirk (1982), the intervocalic consonant belongs to the first syllable only (*resyllabification*), whereas for Kahn (1976), it belongs to both (*ambisyllabicity*). But both theories would predict that the first syllable of *balance* is *bal*.

There is in fact some psychological evidence that the first 'chunk' of a word like *balance* is *bal* and not *ba* as Maximal Onset would have it. For example, Treiman and Danis (1988) had subjects reverse the two syllables of a disyllabic word; Derwing (1992) had them judge the appropriateness of word-internal pauses. We also conducted some experiments asking subjects directly to write down what they thought the first syllable of a given word was (Dupoux, Hammond, Meador, and Ohala, in preparation). The finding that emerges from all these studies (see Table 1[3]) is that the first chunk of *balance* is *bal* and not *ba*.

As Bradley *et al.*, (1993) and Zwitzerlood *et al.*(1993) argue, given these facts about *balance*, it is not very surprising that no interaction between target type and word type was found, since *balance* and *balcony* end up not differing in their first part. Moreover, it has been found in some of the on-line English experiments (Bradley *et al.*, 1993; Cutler *et al.*, 1989) that CVC targets yield faster responses than CV targets irrespective of word type (see Table 1). This

[2]Such a case is not without precedent. Cross-linguistic investigation has shown that languages differ in the way strings are parsed into constituant syllables. In principle, it is possible to find a pair of languages where the same string of phonemes would be parsed into different syllables. For example, in Iraqi and Cairene Arabic, the same strings are syllabified differently. These different syllabifications also emerge when speakers of these different dialects acquire English as a second language. See Broselow, 1980.

[3]Fallows (1978) report similar results obtained with 4-10 year old children, but the data are not presented in a fashion that makes it possible to include them in the table.

is exactly what one would expect given the syllabic structure of these words. Thus, one might argue that it is not that processing in English is different from French, it is just that inappropriate English materials were chosen to test the difference.

Table 1: Metalinguistic and on-line evidence for the first syllable of words like *balance* (lax vowel in the first syllable, no stress in the second syllable, intervocalic liquid consonant, except in Bradley *et al.*(1993)). The first part of the table includes the metalinguistic studies and gives the percent of responses consistent with either CV or CVC as the first syllable of *balance*. The second part of the table includes the fragment monitoring studies and gives reaction times for detecting either CV or CVC in words like *balance*.

	Off-Line tasks	CV	CVC	
Treiman & Danis (1988).	Syllable reversal	33%	66%	
Derwing (1992).	Pause break judjment	13%	87%	
Dupoux, Hammond, Meador, & Ohala (in preparation).	First syllable elicitation, words	11%	88%	
	First syllable elicitation, nonwords	14%	84%	
	On-Line tasks	CV	CVC	CVC-CV
Cutler, Mehler, Norris & Segui (1986).	Exp.1; words:	456	448	8
	Exp.3; nonwords:	391	370	21
Cutler, Mehler, Norris & Segui (1992).	Exp. 1&2, French-English bilinguals:	513	493	20
Bradley, Sanchez-Casas, & Garcia-Albea (1993).	Exp. 2; various intervocalic consonants:	419	398	21

The proposal that we want to test in this study is that appropriate materials can only be found when stress is taken into account. In particular, we manipulate the stress of the second syllable in CVCV stimuli to test whether English uses stress-sensitive syllables.

Foot-like units

The second hypothesis that we test is that English listeners use a different structure from the one posited in the other languages: the metrical foot (e.g. Hayes, 1987). The metrical foot is defined in terms of a single stressed syllable optionally followed by one stressless syllable. In terms of foot structure, *balance* and *balcony* do not constitue a contrasting pair, since both words are made up of a single foot that emcompasses both *ba* and *bal* fragments. If this structure were the one used by subjects, one would not have expected a difference in the detection time to *ba* and *bal* fragments. The only way to get a difference in foot structure is to compare and contrast words that have a stressed-stressless pattern (one foot) to words that have a stressed-stressed pattern (two feet).

Expletive Infixation in English provides evidence that such a contrast has some psychological relevance. Expletive Infixation is a natural process of English whereby subjects insert an expletive like *fucking* or *bloody* inside a word, as in *[Minne]-fucking-[sota]*. It is well known that the infixation is only possible when it does not break a foot, that is, the infix can only be positioned between two feet (McCarthy, 1982; Hammond, 1991). Other language games in English have been show to be sensitive to feet (Hammond, 1990). It remains to be seen whether such units could be the basis of a perceptual chunking strategy similar to syllables in French and morae in Japanese.

Reduction-based structures

Finally, the third hypothesis that we explore posits that what counts in English is not stress per se, but the correlated phenomenon of vowel reduction, that is, the fact that unstressed syllables have a reduced vowel. Cutler and Norris, (1988) have argued that vowel reduction plays an important role in auditory lexical access. When subjects have to discover a word embedded in a nonword carrier (the *word spotting task*), the presence of vowel reduction in the following syllable can interact with access. Thus subjects can discover *mint* in [míntàyf] (unreduced-unreduced) more slowly than in [míntə̆f] (unreduced-reduced). The interpretation is that the processing system posits a boundary before each unreduced syllable, and that this boundary is used to initiate a lexical access attempt. Further evidence for such reduction-based boundaries have been found using other tasks such as identification of words in noise (Cutler & Butterfield, 1992).

Our hypothesis is that such reduction-based structures or boundaries (hereafter: ρ-boundaries, for reduction-based-boundaries[4]) have a status equivalent to syllabic structures or boundaries in French, and hence should also be found with the fragment monitoring paradigm. After all, the fragment detection and the word spotting tasks are quite similar; in both cases, subjects have to extract a fragment embedded in a larger stimulus, only in the former the fragment is given in advance, and in the latter, subjects have to discover it. So, if there is evidence for ρ-boundaries with the one task, one might also expect to find it with the other.

Note that under such a story, we would expect no difference between items like *balance* and *balcony* since the second syllable is reduced in BOTH cases, and hence there are no stimulus-internal ρ-boundaries in either case. Note also that the reduction hypothesis is similar although not identical to the foot hypothesis outlined above. Even though vowel reduction and absence of stress are highly correlated, there are cases like the final syllable in *window* which is arguably stressless, but contains an unreduced vowel (see Chomsky & Halle (1968) for an extensive discussion.). Thus, there exist materials for which the foot hypothesis and the ρ-boundary hypothesis make different predictions.

[4] We take this terminology from the ρ-feet posited by McCarthy (1979) to deal with vowel reduction in Tiberian Hebrew.

Stress-Sensitive Strategies

To summarize thus far, there are a number of reasons why the *balance/balcony* contrast may not have resulted in a significant effect in English (see Table 2). First, it might be that English listeners use stress-sensitive syllables instead of maximal onset syllables. Second, English subjects may be using a unit related to the foot instead of the syllable. Third, English subjects may be sensitive to ρ-boundaries defined by alternation of reduced and unreduced syllables. The three hypotheses posit no difference in structure between the two word classes previously tested and hence are consistent with the absence of an interaction in these materials. Moreover, they are all plausible candidates given the widespread belief that stress plays a predominant role in the processing of English. For example, Cutler and Mehler (1993) have proposed the "periodicity bias hypothesis" which is quite consistent with any of our three SSSs. The hypothesis states that language users choose processing strategies that are best suited to their language rhythm. French, a syllable-timed language would use syllables, Japanese, a mora-timed language would use the mora, and English a stress-timed language would logically use some kind of stress unit.

At this point the question becomes: what materials might be more appropriate for English? We propose to test all three hypotheses simultaneously by using words where stress and vowel reduction are confounded. That is, every stressed syllable has an unreduced vowel (a necessity in English), and every unstressed syllable is also reduced (not always the case). Keeping in line with established terminology in psychology, we will call the former *strong* and the latter *weak* syllables[5].

In our study we will thus compare and contrast C(C)VCVX words that are strong-strong and where both syllables are stressed (like *methane* or *climax*) to C(C)VCVX words that are strong-weak (like *method* or *climate*) (see Table 2)[6].

For strong-strong words, all three hypotheses predict that the first perceptual chunk should be CV because i) stress forces the first syllable to be open, ii) the CV fragment corresponds to a whole foot, or iii) there is a salient boundary just before the second unreduced syllable. So in all the cases, the intervocalic consonant will be affiliated with the second part of that word (as in French), and CV targets should be detected faster than CVC targets. In contrast, for strong-weak words, the first perceptual chunk is a CVC for the stress-based syllable hypothesis, and the entire word for the other two. So our SSSs predict either that CVC should be detected faster than CV or that they are detected equally fast. In any event, if one of the above hypotheses is correct, there should be a structural difference between strong-strong and strong-weak words which

[5] Note that this does NOT correspond to the use of these terms in linguistic theory where they have a different meaning! See Hayes (1987) for discussion.

[6] We do NOT test strong-weak versus weak-strong stimuli like *canada* versus *canoe*. The reason we do this is because we want to keep constant the initial part of the stimulus in order to use the 2×2 design of the original French study. Moreover, *canada* versus *canoe* displays a contrast not only in the stress of the second syllable but also in the stress and hence vowel quality of the first (full versus reduced). Since Treiman and her colleagues (e.g., Treiman and Danis, 1988) have shown that vowel quality in the first syllable has a great effect, the difference between *canada* and *canoe* could be due in part to differences in vowel quality alone.

should show up as an interaction between word type and target type.

Note that, regardless of the effect of stress and/or reduction, other factors may play a role in the fragment monitoring paradigm. For instance, it could be the nature of the intervocalic consonants is relevant. In the original English experiment, only intervocalic liquids were used. Since these are relatively sonorous consonants, and increased sonority has been shown to enhance the likelihood of leftwards affiliation with metalinguistic tasks (Treiman and Danis, 1992, Derwing and Neary, 1992), it could be that just using different consonants, like stops in *tacky* [tǽki] versus *tactic* [tǽktɪk], would show a syllabic effect just like in French[7]. Another hypothesis altogether would build on the observation that syllables like /bæ/ with a lax vowel never occur at word ends. So, vowel quality (or vowel length) could also be a variable that influences syllable structure (Treiman and Danis, 1992, Derwing and Neary, 1992; see Zwitzerlood *et al.*, 1993, for a case like this in Dutch). In the following, we will control these other variables to examine the effect of stress alone. We return to these other variables in our final discussion.

Table 2: Hypothetical chunking of words into units in French and English under different hypotheses.

		French	
	balance	balcon	ba/bal interaction
Max Onset σ	[ba][lɑ̃s]	[bal][kõ]	Yes
		English	
	balance	balcony	ba/bal interaction
Max Onset σ	[bǽ][ləns]	[bǽl][kə][ni]	Yes
Stress sensitive σ	[bǽl][ləns]	[bǽl][kə][ni]	No
ρ-boundary	!bǽləns	!bǽlkə!ni	No
Foot	[bǽləns]	[bǽlkə]ni	No
	climax	climate	cli/clim interaction
Max Onset σ	[kláy][mæ̀ks]	[kláy][mət]	No
Stress sensitive σ	[kláy][mæ̀ks]	[kláym][mət]	Yes
ρ-boundary	!kláy!mæ̀ks	!kláymət	Yes
Foot	[kláy][mæ̀ks]	[kláymət]	Yes

Experiment 1:

In the following, we select strong-strong words like *climax* [kláymæ̀ks], where both syllables are stressed such that the intervocalic consonant should not be drawn into the first unit. We contrast this case with strong-weak words like

[7]Some pilot work on English conducted by Toby Mintz at the CNRS-EHESS lab failed to find any effect with such materials. Also, Bradley *et al.* (1993) using a mixture of liquid, nasals, and obstruents, replicated the absence of interaction in English.

climate [kláymət], where the second syllable is stressless and reduced. Under all versions of the SSS, the intervocal consonant should affiliate to the left. If the syllabification of intervocalic consonants is controlled by stress, these word types should exhibit different syllabifications. If the affiliation of the intervocal consonants is controlled by reduction, again these word types should show different affiliations. Finally, if the foot structure is what is relevant in English, again these word types should differ with the intervocalic consonant falling in the second foot in the first type *climax* and in the first (and only) foot in the second type *climate*.

We also carefully control the vowel quality of the first syllable so that both CV (consonant-vowel) and CVC (consonant-vowel-consonant) are possible at word ends. This was done so that the effects of vowel quality would not obscure the picture[8]. Also, the intervocalic consonants in both strong-strong and strong-weak words were matched in order to avoid a confound with sonority.

We also included a second phonological condition that had not been tested before: cases where the words did not start with a main stress (like *humanity/humanistic* [hỳumǽnɪDi]/[hỳumənɪ́stɪk]). Like the *climax/climate* type opposition, these materials were controlled for vowel quality, consonant type, and contrasted in stress such that they should result in a surface difference in syllabification. However, unlike *climax/climate*, these new materials involved an initial secondary stress as opposed to an initial primary stress.

These additional materials were added to provide for a more general examination of the SSSs. Also, they constitute, in a sense, a more powerful test of our hypotheses. First, the primary-stress-first words and the secondary-stress-first words exhibit a different degree of reduction of the second syllable of the strong-weak words (Pérez, 1992). Thus the second syllable of a word like *human* generally is not "as reduced" as the second syllable of a word like *humanistic*. Second, the primary-stress-first words and the secondary-stress-first words exhibit a different degree of stress in the strong-strong condition. Thus the second syllable of *climax* has less stress (secondary) than the second syllable of *climactic* (primary). Both factors go in the same direction, that is, that the secondary-stress-first items show an enhanced contrast between the strong-weak and the strong-strong conditions than the climax/climate items. If reduction or "strength" is what is really playing a role in these data, we would expect an enhanced interaction in the secondary-stress-first words.

Method

Materials

Fourteen pairs of polysyllabic words were selected. Each pair was composed of words sharing the same CVC (or CCVC) segments and the initial stress value but differing in the stress/reduction value of their second syllable (e.g. *climax/climate*): in one word, the second syllable was stressed and unreduced (like in *climax*), whereas in the other, it was unstressed and reduced (like in *climate*). All words had a tense first vowel [i,

[8] Bradley *et al.*observe that the CV response to items like *balance* [bǽləns] may have been hindered by the fact that [bæ] is an impossible word in English.

u, e, o, ay][9], and a single intervocalic consonant which was not a stop, to avoid the allophonic effects of stress (Kahn, 1976). In seven of the pairs, the first syllable had primary stress (*climax/climate*); in the other seven, the first syllable had secondary stress (*humanity/humanistic*).

Each of these 28 experimental words was used to make two distinct short lists of three or four items. The experimental words always appeared in the last position of each list (e.g. *nature, pirate, climax*). There were hence a total of 56 such lists. Each of the list was associated with a *target fragment* that correponded either to the C(C)V or the C(C)VC initial portion of the experimental word (e.g. one of the *'climax'* list was associated to the target *cli* while the other list to the target fragment *clim*).

Thirty-two additionnal distractor lists were constructed. There were one item long (6 lists), two items long (10 lists), 5 items long (10 lists), 6 or 6 items long (6 lists). Fourty of these lists were associated to a target fragment with similar properties to the experimental lists. The remaining 12 lists had a target fragment that never appeared in any of the words of the list.

The total of 88 lists were splitted into two blocks in such a way that each block contained only one repetition of each experimental word, and an equal number of words in each experimental condition. The non-experimental words were chosen to match globally the average frequency and length of the experimental items. None of the non-experimental words in a list started with the same consonant as an experimental item. A set of ten training lists was also constructed conforming to the same specifications.

Lists were recorded by a male American English speaker at a rate of one word every 1.5 sec. The speaker read the items from a master list containing all the lists merged together. The speaker did not know what target corresponded to what list, or how many words each list had. This was done in order to prevent the speaker from producing any intonational fall in final-position items. The target syllables were recorded by a female speaker as follows: A phonetic transcription of each target syllable was provided together with a word example. The speaker was instructed to read the isolated syllable "as it would appear in the example word". Each target syllable was recorded twice, once with an example of the *climax* type, once with an example of the *climate* type. All the materials were digitized on a 386 computer at 16kHz and edited with the BLISS waveform editorprovided by John Mertus..

For counterbalancing, two versions of the Experiment were constructed (A and B). Version B was similar to version A except that all CV targets were replaced by their associated CVC targets and vice versa.

Procedure

The experiment was run on each subject individually in an experimental booth equipped with headphones, a computer screen, and response buttons. The digitized stimuli were presented binaurally via headphones by the 386 computer. The two different exemplars of each target fragment were presented auditorily, followed by the short list of words. At the onset of the experimental word, an inaudible digital pulse was sent to each of the test booth machines which recorded subjects' responses and latencies[10]. Feedback was provided through the computer screen when the responses were too long (over 2000ms).

Subjects were instructed to monitor for the occurrence of the auditorily presented targets in the spoken words and were told that in some sequences there might be no

[9] The 'tensity' of the low vowels [a] and [ɔ] in English is questionable, and the diphtongs [aw] and [oy] are too unfrequent to be used.

[10] The experiment was carried out by using the DMASTR/VMASTR package of Ken and Jonathan Forster.

word matching the specified target. Subjects were told to respond only if the specified target was detected in the auditory materials. The experiment lasted about 20 minutes.

Subjects

44 subjects whose first language was American English were used in this experiment. They were undergraduate or graduate students of psychology at the University of Arizona and received course credit for their participation. Half of the subjects were run with version A of the Experiment and half with version B.

Results

Mean reaction times for each subject and experimental condition were computed. Cutoff points were established at 100 ms. and 1500 ms, and responses outside this range were counted as errors. Each subject's mean and standard deviation was computed. Reaction times above and below two standard deviations for each subject were replaced by the upper (or lower) value of the two standard deviations.

After the experiment was run, we discovered that the speaker had produced the item pair *chirographic/chirography* incorrectly so that it did not match the target, and this pair was removed from any further analysis[11].

The mean response times for each condition are shown in Table 2. Overall analyses of variance were performed on means for subjects and items. These analyses showed that only the target type factor introduced a significant effect; CV were detected faster than CVC ($F1(1,42)=4.47$, $p<.05$; $F2(1,11)=5.58$, $p<.04$). None of the other factors (stress and word type) introduced any significant effect ($p>.1$). None of the two-way interactions between these factors was significant ($p>.1$) and the three-way interaction was only marginal in the subject analysis ($F1(1,38)=3.01$, $.05<p<.1$; $F2(1,11)=2.26$, ns).

The marginal three-way interaction suggested that the two types of materials (primary stress versus secondary stress) might behave differently. We conducted a separate analysis for the two types of materials. For the words starting with a primary stress (*climax/climate*) the main factors did not attain significance ($p>.1$). However, there was a significant interaction between target type and word type ($F1(1,42)=6.34$, $p<.02$, $F2(1,11)=12.65$, $p<.005$). This interaction can be accounted for by the fact that CV targets were detected faster than CVC targets in words like *climax* (21ms $F1(1,42)=5.38$ $p<.03$; $F2(1,6)=7.28$, $p<.04$). This was not the case in words like *climate* where responses to CV vs CVC did not differ significantly (-2ms, all $Fs<1$).

For the words starting with a secondary stress, the picture was different. There was a significant Target effect in the Subjects analysis ($F1(1,42)=4.84$ $p<.04$; $F2(1,5)=6.13$, $.5<p<.1$). This difference was only marginal when analysed separately on words like *humanistic* ($F1(1,42)=3.29$, $.5<p<.1$, $F2(1,5)=2.59$, ns) or on words like *humanity* ($F1(1,42)=3.36$, $.5<p<.1$, $F2(1,5)=2.19$, ns). None of the other factors nor any interaction among them approached significance.

[11]Inspection of the responses showed a 7% miss rate on these items which was 6 times greater than on the other items!

An analysis of the missing rate did not reveal any significant effect.

Table 3: Reaction times for detecting CV or CVC fragments in strong-strong versus strong-weak words in Experiment 1.

	CV targets	CVC targets	CVC-CV
	Primary-stress-first		
CLIMAX	397 [17.1] (0.3%)	419 [19.2] (0.0%)	22*
CLIMATE	403 [19.8] (0.0%)	401 [18.5] (0.0%)	-2
	Secondary-stress-first		
HUMANITY	399 [19.9] (0.8%)	416 [17.8] (0.0%)	17
HUMANISTIC	394 [18.2] (0.4%)	411 [19.0] (0.4%)	17

Note. (*) Difference significant by subject and by item at .05
Standard Error is shown between square brackets.
Missing Rate is shown between parentheses.

Discussion

This experiment failed to find an overall signifiant target-by-word interaction, failing to support the SSSs. However, looking separately at the effects in our two categories of words revealed that there was a significant interaction in the expected directions for words starting with a primary stress (*climax/climate*), but not for words starting with a secondary stress (*humanity/humanistic*). Yet, the fact that the three-way interaction was not significant does not allow us to conclude that the two types of words show a qualitatively different pattern. At this point, we have three possible interpretations of this: 1) primary-stress-first and secondary-stress-first words have different structures, but experiment 1 was not powerful enough to capture the difference; 2) primary-stress-first and secondary-stress-first items have the same structure, but some processing factor prevented the effect from arising in the latter case; and 3) primary-stress-first and secondary-stress-first words are really the same, and the apparent difference found in Experiment 1 was just a statistical artifact. We consider each of these possibilities in turn.

The first possibility is that primary-stress-first and secondary-stress-first words have different structures. This difference would have to apply to the first unit of the strong-strong condition such that, in some sense, the first syllable or unit of words like *climax/climate* are more distinct than the first syllable or unit of words like *humanity/humanistic*. However, our three SSSs predicted exactly the opposite: in terms of degree of stress, the contrast between very reduced and primary stress (the *humanity/humanistic* case) is greater than the contrast between moderately reduced and secondary stress (the *climax/climate* case). So, some variable other than the ones posited in the SSSs would have to be worked out in order to account for the proposed dissociation between primary and secondary stress. Note that our two types of words did not differ only with respect to the stress value of the first syllable. They also were not matched in vowel quality and consonant type (see appendix 1 for a complete listing). Experiment 3 addresses this point.

The second possibility is that primary-stress-first and secondary-stress-first items have the same structure, but some processing effect prevents the effect from arising in the latter case. This has a precedent in Spanish, French, and Catalan (Sebastian, *et al.*, 1992; Dupoux, 1993), where syllabic effects can be made to disappear in some structures when subjects respond with very short latencies. However, with slower reaction times, syllabic effects reappear. Yet, in this case too, we might, if anything, have predicted just the opposite effect. Primary-stress-first words are mostly disyllabic and hence have a longer initial syllable than secondary-stress-first which are mostly 3 or 4 syllables long. As a consequence, subjects would have had more chances to get to the stress value second vowel by the time they respond in *humanity/humanistic* than in *climax/climate* cases. Hence, one would have expected a greater effect of stress in the former than in the latter. The next experiment will more directly address the issue of the speed of response.

Finally, it could be the case that secondary-stress-first and primary-stress-first words are really the same, and the apparent difference found in Experiment 1 was just a statistical artifact. If this were the case, this would suggest that the true picture is somewhere between what we found for the *climax*-type words and the *humanity*-type words. This would mean, however, that the effect of stress/reduction in these tasks is quite reduced, since an overall interaction between stress and target type was not found.

In order to tease apart these different alternatives, we conducted additional experiments. Experiment 2 examines the effect of speed on these effects, using exactly the same materials. Experiment 3 uses much more controlled materials to examine the possible difference between primary first and secondary first words. These two experiments will be helpful to address the first question we raised as well, namely the generality of the interaction found with *climax/climate*.

Experiment 2:

In the preceding experiments, the mean reaction time was around 400ms, and assuming approximately 150ms for the motor response, subjects had heard only 250ms of signal when they trigger their response. In *climax/climate*, this corresponds to the end of the first vowel and in *humanity/humanistic* to the end of the intervocalic consonant. So, unless there was some kind of allophonic information correlating with the stress of the following syllable (which we tried to avoid), one should not expect much effect to start with!

This is similar to what has been described as a *truncation* phenomenon (Dupoux, 1993) whereby subjects have access to information regarding only a subpart of a syllable when they are in a speeded reaction time experiment. However, when they are slowed down, more information becomes available to the decision system and syllabic effects become much stronger. For example, Sebastian *et al.*(1992) found that syllabic effects in Spanish were much more likely to emerge when the responses were slow than when subjects responded very rapidly. Similarily, Dupoux (1989) argued that early responses allow subjects to ignore the syllable and answer before the syllable is completed. When the time pressure is alleviated, syllable effects seem to be obligatory. In this experiment we use the same materials, but alleviate the time pressure, using a secondary task that induces subjects to slow their responses (Sebastian *et al.*, 1992). If the SSS hypothesis is correct, a larger interaction should be found for both primary-first and secondary-first words.

Method

Materials

The same materials as in Experiment 1 were used.

Procedure

As in Experiment 1, subjects were instructed to monitor for the occurrence of the auditorily presented targets in the spoken words and were told that in some sequences there might be no word matching the specified target. A lot of emphasis was put on accuracy of the responses. In addition, a secondary task was added whereby subjects had to monitor for the meaning of the words. They were told that the lists contained semantic associates (of the bread-butter sort) and that they would have to memorize these associates, in order to fill in a questionnaire about the semantic content of the lists after the experiment. This is the same secondary task used in Sebastian *et al.*(1992) in their Experiment 3.

Subjects

48 subjects whose first language was American English were used in this experiment. They were undergraduate or graduate students of psychology at the University of Arizona and received course credit for their participation. Half of the subjects were run in version A and half in version B.

Results

Mean reaction times for each subject and experimental condition were computed. Cutoff points were established at 100 ms. and 1500 ms and responses outside this range were counted as errors. Each subject's mean and standard deviation were computed. Reaction times above and below two standard deviations for each subject were replaced by the upper (or lower) value of the two standard deviations.

As in the previous experiment, the pair *chirographic/chirography* was removed from any further analysis.

The mean response times for each condition are shown in Table 3. Overall analyses of variance were performed on means for subjects and items. These analyses showed that all the three main factors introduced a significant effect; CV was detected faster than CVC ($F1(1,46)=5.07$, $p<.03$; $F2(1,11)=7.69$, $p<.02$). Strong-strong words were responded to slower than strong-weak words ($F1(1,46)=21.1$, $p<.001$, $F2(1,11)=5.06$, $p<.05$). Primary-stress-first words tended to be responded to faster than Secondary-stress-first words ($F1(1,46)=5.97$, $p<.02$; but $F2(1,11)<1$). None of the interactions between these factors was significant ($p>.1$).

However, the advantage of CV targets (over CVC targets) was significant for strong-strong words (25ms; $F1(1,48)=6.63$, $p<.02$; $F2(1,11)=6.07$, $p<.04$). The same contrast failed to reach significance for strong-weak words (6ms; all $Fs<1$). Similarly, the avantage for strong-strong words (over strong-weak words) was only significant for CVC targets (31ms; $F1(1,48)=15.3$, $p<.001$; $F2(1,11)=8.49$, $p<.02$) but not for CV targets (12ms; $F1(1,48)=3.37$, $.05>p>.1$; $F2(1,11)<1$).

As in Experiment 1, we conducted a separate analysis for the two types of materials. For the primary-stress-first word types (*climax/climate*) there was a significant advantage of CV targets ($F1(1,46)=6.65$, $p<.02$; $F2(1,6)=9.74$, $p<.03$). The effect of word type did not attain significance ($p>.1$), nor did the interaction between word type and target type ($p>.1$). For the secondary-stress-first words, strong-strong words tend to be responded to faster than strong-weak words ($F1(1,48)=17.2$, $p<.01$; but $F2(1,5)=3.23$, $p>.1$). No other effect or interaction was significant.

An analysis of the missing rate did not reveal any significant effect.

Table 4: Reaction times for detecting CV or CVC fragments in strong-strong versus strong-weak words in Experiment 2.

	CV targets	CVC targets	CVC-CV
	Primary-stress-first		
CLIMAX	537 [22.9] (1.2%)	568 [25.4] (0.9%)	31*
CLIMATE	532 [19.9] (0.6%)	546 [22.5] (0.3%)	14
	Secondary-stress-first		
HUMANITY	566 [27.4] (0.7%)	585 [26.9] (1.0%)	19
HUMANISTIC	546 [25.1] (0.0%)	545 [27.8] (0.7%)	-1

Note. (*) Difference significant by subject at .05
Standard Error is shown between square brackets.
Missing rate is shown between parentheses.

Discussion

The manipulation of subjects' response time by means of a secondary task did not increase the effects observed in Experiment 1. Even when subjects were slowed down, this experiment failed to uncover a significant interaction between target type and word type. Moreover, the interaction found in Experiment 1 with words like *climax/climate* was no longer significant. A significant interaction for words like *climactic/climatology* did not emerge either. In fact, the two categories of words look quite similar now, which suggests that the apparent differences found in Experiment 1 might have been an accident. Finally, as in Experiment 1, there is a significant upward trend, CV targets yielding faster reaction times than CVC targets. This effect was more pronounced for strong-

strong words than for strong-weak words, which goes in the direction predicted by the SSSs, but was too small to generate a reliable interaction.

Our failure to find a strong enough effect of stress in Experiment 1 & 2 could be accounted for in a number of ways. One is that stress really governs the processing units of English, but our experiments were not powerful enough to reflect it. A second possibility is that, SSSs notwithstanding, there is really very little effect or even no effect of stress on the relevant processing structures.

In order to address the issue of experimental power, we replicated Experiment 1, using a larger materials set. As noted above, to check the behavior of words with primary-stress-first versus secondary-stress-first, we made matched quadruplets of words like *climax/climate/climactic/climatology*, where vowel quality and consonant sonority were held constant across the diferent stress patterns.

In addition, in Experiment 4 we ask subjects to decide whether a given syllable is the first syllable of a word. This metalinguistic task should inform us what subjects think is the syllable structure of the words that they are presented with in this experiment, and whether this structure is stress-sensitive or not.

Experiment 3: A Replication

Method

Materials

Ten quadruplets of the form *climax-climate-climactic- climatology* were selected. These quadruplets are produced by crossing two parameters: presence/absence of stress in the second syllable and degree of stress of the first syllable (primary or secondary). In each quadruplet, the words were matched minimally with respect to the first vowel and the intervocalic consonant. As in the first experiment, the first vowel was tense [i, u, e, o, ay] and the intervocalic consonant was not a stop. These materials and their accompanying targets were recorded and edited as in Experiment 1.

Instead of presenting a target syllable and then a list of one to six words, only one word is presented immediately after each target. This variant of the fragment detection technique has the advantage of allowing more responses to be collected for each subject, and hence allows us to increase the number of items tested. All 40 test items were presented twice, once with a CV target, once with a CVC target. Fifty-two NO filler items and 24 YES filler items were added as well.

Two versions of the experimental list were made to control for serial effects. The same experimental words were used in the two versions, but the target was changed from one list to the other. That is, if subjects had to monitor for *cli* in *climax* in version A, they had to monitor for *clim* in *climax* in version B. Again, as in the Mehler *et al.*and Cutler *et al.*experiments, each stimulus word occurred twice (although preceded by different filler words) in the experiment. Finally, a set of ten training items was constructed conforming to the experimental specifications.

Procedure

Subjects were required to listen to the target fragment and to respond as soon as possible, if the following word started with that fragment. Very strong emphasis was put on accuracy. Subjects were discouraged from any kind of guessing, and the point was made that they had to detect the full fragment before responding.

Subjects

Fifty-eight subjects whose first language was American English were used in this experiment. They were undergraduate or graduate students of psychology at the University of Arizona and received course credit for their participation. Half of the subjects were run in version A and half in version B.

Results

Mean reaction times for each subject and experimental condition were computed. Cutoff points were established at 100 ms. and 1500 ms, and responses beyond these cutoffs were counted as errors. Each subject's mean and standard deviation were computed. Reaction times above and below two standard deviations for each subject were replaced by the upper (or lower) value of the two standard deviations. The mean response times for each condition are shown in Table 4. Overall analyses of variance were performed on means for subjects and items.

There was a significant word type effect, strong-weak words globally giving rise to slower reaction times than strong-strong words ($F1(1,54)=12.85$, $p<.001$; $F2(1,18)=4.43$, $p<.05$). There was also a significant target effect in the subjects analysis, CV targets being responded to faster than CVC targets ($F1(1,54)=6.13$, $p<.02$; $F2(1,18)=1.64$, $p>.1$). Finally, there was a significant stress type effect, with primary-stress-first words being faster than secondary-stress-first words ($F1(1,54)=16.7$, $p<.0001$; $F2(1,18)=4.54$, $p<.05$).

Stress type interacted with word type in the subject analysis ($F1(1,54)=10.93$, $p<.002$). The interaction between word type and target type failed to reach significance in the subjects analysis ($F1(1,54)=3.26$, $.1>p>.05$, $F2(1,18)=2.49$; $p>.1$). No other two-way or three-way interaction by subject or by item reached significance ($p>.1$).

In the following, we ran a series of contrasts to evaluate whether the target effect or the word effect was significant in more specific conditions.

Target effect. The Target effect was significant for strong-strong words taken globally ($F1(1,54)=8.35$, $p<.006$; $F2(1,18)=6.63$, $p<.002$). For strong-strong words, this effect was significant for primary-stress-first and secondary-stress-first separately in the subjects analysis (primary-stress-first: $F1(1,54)=4.28$, $p<.05$; $F2(1,9)=4.59$, $.1>p>.05$, secondary-stress-first: $F1(1,54)=4.50$, $p<.04$; $F2(1,9)=2.63$, $p>.1$). For strong-weak words, the target effect was not significant either globally or on the primary or secondary cases separately ($p>.1$).

Word type effect. The word type effect was much stronger for CV targets ($F1(1,54)=12.3$, $p<.001$; $F2(1,9)=5.73$, $p<.03$), than for CVC targets ($p>.1$). The net word type effect was stronger in secondary stress cases ($F1(1,54)=16.7$, $p<.0001$; $F2(1,9)=8.65$, $p<.02$) than in primary stress cases ($Fs<1$). This last aspect of the data accounts for the interaction between stress type and word type reported above.

A similar analysis was performed on missing rates. The only significant effect was an interaction between stress and word type in the subjects analysis but not in the items analysis ($F1(1,54)=5.86$, $p<.02$; $F2(1,18)=2.8$, $p>.1$).

Table 5: Reaction times for detecting CV or CVC fragments in strong-strong versus strong-weak words in Experiment 3.

	CV targets	CVC targets	CVC-CV
	Primary-stress-first		
CLIMAX	538 [23.3] (1.6%)	554 [23.1] (1.6%)	16
CLIMATE	551 [23.0] (1.4%)	548 [21.3] (1.1%)	-3
	Secondary-stress-first		
CLIMACTIC	549 [24.9] (0.0%)	569 [29.0] (1.1%)	20
CLIMATOLOGY	590 [25.9] (0.9%)	592 [29.8] (1.8%)	2

Note.
Standard Error is shown between square brackets.
Missing rate is shown between parentheses.

Discussion

In this experiment, we carefully matched the phonological properties of words with primary stress first and words with secondary stress first.

The pattern of results we got for these two types of words is remarkably uniform, which suggests that the apparent divergent pattern found in Experiment 1 was just an accident.

Second, it is not because of lack of power that we do not find a significant interaction in experiments 1 & 2. Indeed, this experiment had more items and more subjects than in Experiment 1. Yet, no significant interaction between target type and word type was uncovered, either globally or when primary-stress-first or secondary-stress-first words were analysed separately. In fact, this experiment was much more powerful than the experiments run either in French or in Spanish. We used 20 pairs of items, whereas only 5 were used in the original French study (and two times 5 in the Spanish study) and we used 50% more subjects, which makes our study 6 times more powerful than the French study. Yet, no significant interaction was found.

What was found instead was a trend for CV targets to give faster reaction times than CVC targets. This trend was actually significant for strong-strong words and not for strong-weak words, which goes in the expected direction for SSSs. However, just like in Experiment 2, this effect is not very large in strong-strong words to start with (18 ms), and fails to provoke a target-by-word interaction.

We can conclude that none of the SSSs are supported by the fragment monitoring experiments we have performed.

However, there is still another possibility that we have not considered yet. It could be that while the stress/reduction-based structures we have considered are not appropriate for the fragment monitoring results, some other structure dependent on other phonological variables is. That is, just because the SSSs are not supported, it does not follow that a purely phonemic/segmental process is required. In particular this is true of the syllabic hypothesis. What we have rejected is the hypothesis that stress-sensitive syllables are used as on-line processing units, not that syllables (under some form or another) are used as on-line processing units. That is, there remains the possibility that syllables have a definition in English that conforms neither to Maximal Onset nor to stress-based theories. In order to check this, we conducted the next experiment, directly asking for subjects' intuitions regarding the syllable structure of our materials.

Experiment 4: Subjects' intuitions

In this experiment, we test whether stress-based syllabification is correct as far as subjects' intuitions go. If stress-based resyllabification holds in English at this level, subjects should report that the first syllable of *climate* is *clim* and not *cli* and vice versa for *climax*. If this were the case, we could conclude that English subjects do not use syllables on-line, or for that matter, any kind of large-sized unit. In contrast, if some other kind of syllable is at hand in English, this leaves open the issue of its relevance in on-line processing.

Method

Materials

The 10 experimental quadruplets of Experiment 3 were used. To these forty words, 10 fillers selected from the fillers of Experiment 3 were added. Using these 50 words, two counterbalanced lists were made. In List A, half of the words were paired with a CV target and half with a CVC target (each type of word was paired equally often with each target type). List B was identical to List A except that the words paired with a CV target were now paired with a CVC target and vice versa. So, contrary to Experiment 3, in a given list, a given experimental word was only presented once (say, paired with a CV target in List A, and with a CVC target in List B). A warm-up series of six filler items was further added at the onset of each list. The exact same stimuli as in Experiment 3 were used. Only the protocol file governing the order of presentation of the various items was changed. In this experiment, a trial consisted in the presentation of a Word (e.g. *climate*), and after an interval of 700ms a Target syllable (e.g. *clim*). Each trial was separated with an interval of 5 seconds.

Procedure

Each subject was presented with a uniquely randomized version of one of the experimental lists (A or B). The task was to press the Yes button if the Target syllable was the first syllable of the Word, and the No button otherwise. The notion of syllables was exemplified in the instructions by simply saying that the word '*dog*' has one syllable and the words '*balance*' or '*police*' have two. Subjects were told to answer according to their intuitions and that they had ample time to make their decision.

Subjects

The same 58 subjects that were used in the previous experiment ran this experiment immediately after having completed the fragment detection task.

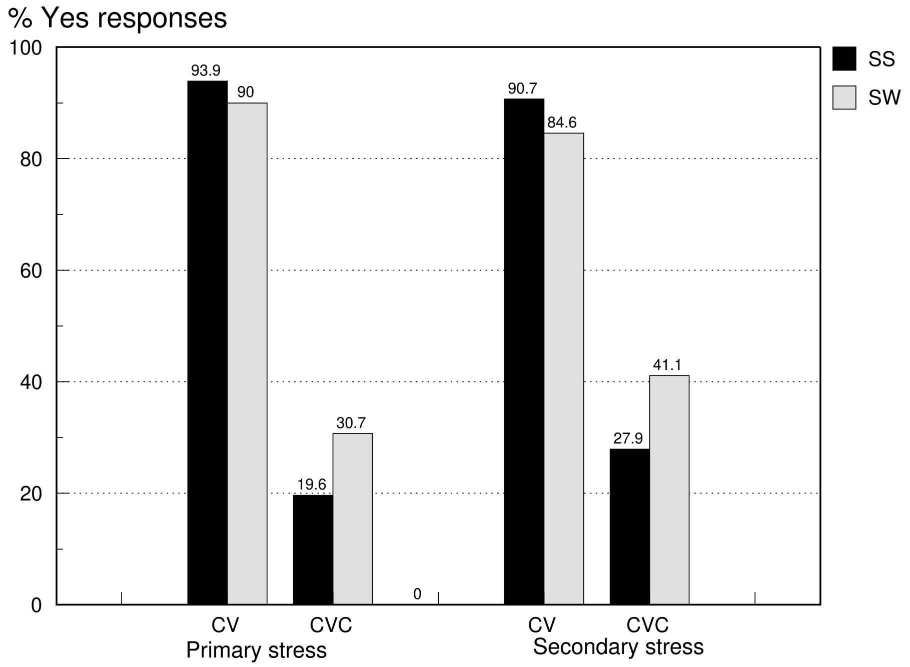

Figure 1: Percent Yes response to CV or CVC targets in a Syllable Judgment task (Experiment 4), in Strong-Strong versus Strong-Weak words that start with a Primary or a Secondary stress.

Results

The percent Yes response was subjected to an analysis of variance with Subjects and Items as random variables. There was a very strong effect of Target, CV targets eliciting much more Yes responses than CVC targets ($F1(1,54)=240$, $p<.0001$; $F2(1,8)=351$, $p<.0001$). This effect was significant for all four categories of words: *climax* ($F1(1,54)=339$, $p<.0001$; $F2(1,8)=1017$, $p<.0001$), *climate* ($F1(1,54)=162$, $p<.0001$; $F1(1,8)=94$, $p<.0001$), *climactic* ($F1(1,54)=185$, $p<.0001$, $F2(1,8)=120$ $p<.0001$) and *climatology* ($F1(1,54)=55.3$, $p<.0001$; $F2(1,8)=55.2$, $p<.0001$). However, as one can see in Figure 1, this effect was stronger for strong-strong words than strong-weak words, generating a significant interaction Target and Word ($F1(1,54)=30.3$, $p<.0001$; $F2(1,8)=14.4$, $p<.006$). Also, the Target effect was somewhat weaker in the primary-stress-first than in the secondary-stress-first words generating a Target by Stress interaction ($F1(1,54)=12.7$, $p<.001$; $F2(1,8)=6.1$, $p<.05$).

In terms of the responses for CV versus CVC targets, simple chi- squares with subjects as a random variable were run. They revealed that there were more Yes responses than No responses to CV targets for the four types of words ($p<.0001$). In contrast, there were more No responses than Yes responses for CVC targets ($p<.05$ at least).

Discussion

The overwhelming result of Experiment 4 is that subjects accept CV and reject CVC syllables, irrespective of the stress pattern (see Figure 1). This is consistent with the notion that the first syllable of both *climax* and *climate* is *cli*, despite the SSSs. It is not the case, however, that stress had no effect at all. Indeed, there was a significant target-by-word interaction in subjects' yes responses. However, this interaction was NOT a cross-over interaction in that it did not reverse the syllable structure for strong-strong versus strong-weak words. Rather, the effect of stress was merely to make subjects less confident about the syllable structure of strong-weak words, hence they responded with a less clear-cut pattern than with strong-strong words. Thus, even though stress had a significant effect in this Experiment 4, the effect was an order of magnitude smaller than what the SSS hypothesis would predict.

General Discussion

We began this paper with a hypothesis to account for why on-line syllable effects occur in languages like French, but not in English: an intervocalic consonant before a weak syllable affiliates to the left in English. (In French, or in English before a stressed syllable, a consonant in analogous positions affiliates to the right.) This hypothesis is inspired by evidence from linguistics, off-line tasks, and stress effects in the processing of English and is consistent with three different views of what the precise structure is that English speakers are using to perform this task. The SSS hypothesis predicted an interaction between target type and word type with strong-strong words like *climax* and strong-weak words like *climate*. It also predicted that the stress pattern of the words should change subjects' intuitions about the syllable structure.

First, we did find an interaction in Experiment 1, for cases like *climax*, but not for cases like *humanity*. However, this interaction did not resurface in Experiment 2 with slowed-down responses, whereas in French and Spanish, such a manipulation quite effectively increases the influence of syllable structure on reaction time data. Further, no interaction was found in the more powerful Experiment 3[12]. True enough, we always got trends of an interaction in the predicted direction. However, in terms of the sheer size of the trends we get (6 ms in Experiment 1; 9 ms in Experiments 2 and 3) we never reach the effects reported in French (20ms, Mehler *et al.*, 1981) and Spanish (Sanchez-Casas, 1988; Experiment 1, 17ms; Experiment 2, 45ms; Sebastian *et al.*, 1992, Experiment 3, 23ms)[13]. Finally, the off-line Experiment 4 showed only a very small effect of stress with respect to explicit syllabification judgments. Indeed, the main finding was that as far as subjects' intuitions go, both *climate* and *climax* words start with the syllable *cli* and not *clim*. This finding is consistent with the reaction time data showing a consistent trend for CV targets being

[12]This is consistent with some unpublished data by Cutler and her collaborators (Cutler *et al.*, personal communication) reporting a failure to get interactions with strong-strong versus strong-weak materials like *gargle-gargoyle*.

[13]The "size" of the interaction is just the average between CV advantage for CV words, and CVC advantage for CVC words. It reflects the average cost of a mismatch between target structure and word structure.

faster than CVC targets (Experiments 1, 2 & 3). Such a target effect was more pronounced for strong-strong words than for strong-weak, just as the subjects' intuitions were more clear-cut for strong-strong than for strong-weak. In brief, both intuitions and reaction times suggest that stress of the following vowel has only a second-order effect on syllable structure (or whatever structure is used in the on-line task)[14]. So, it cannot be that stress alone accounts for the discrepancy between French and English. This allows us to reject the three SSSs in one fell swoop: as far as fragment detection goes, the English processing system does not use stress-sensitive syllables, it does not use feet, and it does not use reduction-based structures.

So, are we back to square one? What is the difference between English and French? What is the on-line structure used in English? We address the substantive issue first and then turn to some methodological issues.

First, a general comment relates to our rejection of the three SSSs. Our SSSs were motivated by the general claim that languages use processing strategies consistent with their rhythm (e.g., Cutler and Mehler, 1993). English being described as stress-timed, it seemed most natural to seek processing units that would be sensitive to stress. Our experiments have provided no evidence for such stress units. Does this mean that the general language rhythm claim is wrong? No. It is our impression that in broad strokes, such hypothesis taps into a plausible and promising avenue of research. Moreover, the precise domain of application of the language rhythm hypothesis is a debated question, and more research is needed to determine at what level of processing it is relevant.

Second, as far as the units of on-line processing are concerned, one possible account, discussed in Cutler *et al.*(1986) is that no syllable-like or foot-like unit is used at all in English, but rather, single phonemes constitute the only unit of processing (see also Pisoni, and Luce 1987). Such a unit would be an option selected by speakers, presumably because the large-sized structures of their language are sufficiently murky so as to preclude their reliable use in processing. This would mean that languages can use radically different processing units, since French would use a rather large unit (the syllable) and English a very small small one (the phoneme). Although this is a logical possibility, it strikes us as a rather implausible one. It also misses the fact that there seems to be differential target effects in on-line tasks in *balance/balcony* as opposed to *climax/climate* (see below). Finally, it disregards the fact that subjects do have reliable intuitions regarding how to break up speech utterances into chunks, and that these intuitions are consistent with the on-line data. Of course, this is far from providing a knock-down argument especially since the alleged correlation between intuitions and reaction times stems mostly from a failure to find interactions. It would be more convincing to find direct positive evidence in favor of such a correlation as well as for a large sized unit of some form or another. More research is needed to examine these points.

Another possibility is that a large-sized unit is used in on-line processing in English, but, the language rhythm hypothesis notwithstanding, it differs from

[14]It remains to be seen, however, whether stress could have a bigger effect in continuous or fast speech where vowel reduction is more pronounced.

the French syllable because of some property of English *other* than stress. The key to this puzzle may be to consider in some more detail the off-line data, and in particular the difference between *balance* and *climate*. Both have a strong-weak stress pattern, hence, they should have the same structure according to the SSS. However, they turn out to be quite different. *Climate* is predominantly judged to have an open first syllable (Exp. 4) whereas *balance* has more of a closed one (Table 1). Similarily, as was noted above, on-line studies have generally found that CVC targets were faster than CV targets with *balance*, and we found that for *climate* the difference was zero or slightly in the opposite direction. Assuming that these effects reflect some perceptual chunking of the same nature as the syllabic effects in French, what would be the resulting unit for English?

We can see two differences between the two word types. The first one is that *balance* has an intervocalic liquid, whereas *climate* has a nasal. As indicated in the introduction, liquids are more sonorous than nasals and hence could adhere more easily to the preceding syllable (Treiman and Danis, 1988; Derwing, 1992; Derwing and Neary, 1991)[15]. More importantly, *balance* has a lax vowel in the first syllable whereas *climate* has a tense vowel. This last difference could be the most important one, since vowel tenseness interacts with the distribution of English syllable-final consonants. That is, lax vowels **never** appear in open monosyllables or in open syllables at the end of a word (other than schwa). Also, vowel tenseness has been shown to have a quite powerful effect in metalinguistic judgments (Treiman and Danis, 1988; Derwing, 1992; Dupoux *et al.*, in preparation).

One could envision the possibility that English uses syllabic-like structures based on vowel quality. More specifically, an intervocalic consonant in English would affiliate to the left if the preceding vowel is one that cannot otherwise terminate a word. The first vowel of *balance* cannot terminate a word in English, but the first vowel of *climate* can, e.g. in *pie*. If the syllabification of intervocalic consonants is as we have indicated, then relevant examples from our experiments and Cutler/Bradley's are syllabified as illustrated in Table 6.

Table 6: Hypothetical syllabification based on vowel quality

Tense	Lax
kláy.mət	bǽl.ləns
kláy.mæ̀ks	bǽl.kə.ni

If this is so (see Pulgram, 1970, and Hammond, to appear, for the linguistic assessment of such an analysis), this would lead us to expect no interaction in either case. However, we should find a CV advantage for *climax/climate* and a CVC advantage for *balance/balcony*. Again, in order to establish this point empirically, the comparison should be performed in a single experiment to

[15]In our experiments, we had 88% nasals and obstruents. In the Cutler *et al.*(1983) experiments, only liquids were used. However, in Bradley *et al.*(1993), there was a mixture of liquid, nasals and obstruents.

neutralize possible effects of methodological differences (filler items and catch trials). Note, however that such an experiment would NOT be comparable to the French *balance/balcon* study. We are in the unfortunate position that it will always be possible that the effects obtained are due to some property of the target alone or the word alone. In fact, if we are right about our interpretation of the data, there are simply NO materials in English that would be similar to the French ones. With lax vowels, no CV/CVC contrast is possible to obtain, because, hypothetically, the vowel attracts the following consonant. We would then have to turn to tense vowels, however it turns out that there are very few (if any!) cases of monomorphemic CVC.CV words with a first tense vowel. What remains is to try either to use ill-formed nonwords, or morphologically complex words which entails a series of control Experiments. Alternatively, one might use other techniques that do not require the minimal pairing that is used in the standard fragment detection paradigm. In any of theses cases, a series of experiments, possibly involving other languagues, would be needed in order to cross-validate the necessary change in paradigm before it would be usable in English.

Before closing, we would like to make some methodological comments. Whether our interpretation is correct or not, our results have important implications about the interpretation of the various tasks used in speech processing studies.

First, our results suggest that in cross-linguistic investigations, it may not be wise to try and choose superficially similar items in order to test for the same unit in two different languages. In fact, the French and the English *balance* although superficially similar, may receive quite a different perceptual analysis by the English and French listeners because the two languages have different phonologies. However, it may not be wise to rely on phonological theory either. Two hypotheses that were derived from linguistic theory, namely, the stress-sensitive syllable and the metrical foot did not pan out in our study. So far, it looks like the best predictor of the on-line results is subjects' **naive intuitions** regarding how to break up utterances up into parts.

Hence, it may be that metalinguistic tasks and on-line fragment detection tasks are more related to one another than what might be expected a priori. This is indeed what Dupoux and Mehler (1992) argue should be the case. The argument is that on-line matching tasks have to be performed on the basis of representations that are otherwise accessible in the metalinguistic sense. So, the representations that are spontaneously activated in speeded matching tasks, are also the representations that can be explicitely manipulated in word games. This would predict that the parallelism between judgments and fragment detection tasks should go through across languages. If so, this opens up a whole area of cross-linguistic investigation that would first probe naive intuitions for what kind of structure might be in use in the language, and then check whether they have any on-line reality.

The second methodological point that our data suggest is that not all on-line tasks tap into the same levels. As we have shown, the fragment monitoring technique is not very sensitive to the difference between strong-weak and strong-strong words. In contrast, tasks such as the word spotting task has been shown to be **very** sensitive to this same contrast (Cutler and Norris, 1988).

The difference betwen our non significant 6-9ms trend and the massively significant 100-200ms strong-weak/strong-strong effects in the word spotting task is overwhelming. Of course, it could be that the difference is due to the specific materials used[16]. But if this pattern of data holds up in a more controlled setup, it suggests a dissociation between tasks that can be performed at a non-lexical level of representation (fragment or segment detection tasks) and tasks that require lexical information (word spotting, lexical decision, etc). Only the latter would show some sensitivity to stress. This would imply that lexical and prelexical processing rely on different types of phonological structures. These issues will be further explored in future work.

Appendix A: Materials for Experiment 1 & 2

Strong-Strong	Strong-Weak
Primary-Stress-First	
primate	primal
climax	climate
primary	primacy
typhoid	typhus
labile	label
donate	donor
collate	colon
Secondary-Stress-First	
licentious	license
premedical	premature
devalue	deviation
humanity	humanistic
momentously	momentarily
prolixity	proletarian

Appendix B: Materials for Experiment 3

Primary-Stress-First		Secondary-Stress-First	
Strong-Strong	Strong-Weak	Strong-Strong	Strong-Weak
primate	primal	primarily	primatology
climax	climate	climactic	climatology
finite	fining	dynastic	dinosauric
finance	final	dynamic	dynamometer
gyrate	gyral	gyration	gyroscopic
cyborg	Cyber	hibiscus	hibernation
female	femur	hemography	hemolytic
cremate	creamer	cremation	crematorium
sumach	Sumer	numerical	numerology
rhubarb	ruby	rubidium	rubicundity

[16] Cutler and Norris (1988) used stimuli with a lax vowel and two medial consonants (*mintaif*), whereas we used words with a tense vowel and a single medial consonant (*climax*).

References

Anderson, S. R. (1982). The analysis of French schwa, or how to get something for nothing, *Language*, **58**, 534–573.

Bertoncini, J., & Mehler, J. (1981). Syllables as units in infant perception, *Infant Behavior and Development*, **4**, 247–260.

Bijeljak-Babic, R., Bertoncini, J., & Mehler, J. (1993). How do 4-day-old infants categorize multisyllabic utterances? *Developmental Psychology*, **29**, 711-721.

Borowsky, T. (1986). *Topics in the Lexical Phonology of English*, Ph.D. dissertation, U. Mass.

Bradley, D.C., Sanchez-Casas, R.M., Garcia-Albea, J.E. (1993). Language specific segmentation strategies: I. Performance over native language materials. *Language and Cognitive Processes*, **8**, 197–233.

Broselow, E. (1980). Syllable structure in two Arabic dialects, *Studies in the Linguistic Sciences*, **10**, 13–25.

Chomsky, N. and M. Halle (1968). *The Sound Pattern of English*, Harper & Row, New York.

Cutler, A., & Butterfield, S. (1992). Rythmic cues to speech segmentation: evidence from juncture misperception. *Journal of Memory and Language*, **31**, 218-236.

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

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

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

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1989). Limits on bilinguism. *Nature*, **320**, 229–230.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1992). Limits on bilinguism. *Cognitive Psychology*, **24**, 381–410.

Cutler, A., & Mehler, J. (1993). The periodicity bias. *Journal of Phonetics*, **21**, 103-108.

Derwing, B.L. (1992). A 'pause-break' task for eliciting syllable boundary judgments from illiterate and illiterate speakers: preliminary results for five diverse languages. *Language and Speech*, **35**, 219–235.

Derwing, B.L., & Neary, T.M. (1992). The 'vowel-stickiness' phenomenon: three experimental sources of evidence. *Proceedings of the 12th International Congress of Phonetic Sciences*, **3**, Aix-en-Provence, France.

Dupoux, E. (1989). *Identification des mots parlés: détection de phonèmes et unité prélexicale*. PhD dissertation, EHESS, Paris, France.

Dupoux, E. (1993). Prelexical processing: the syllabic hypothesis revisited, In G. T. M. Altmann and R. Shillcock (Ed), *Cognitive models of speech processing: The second sperlonga meeting*, pp81–114, LEA: Hove, East Sussex, UK.

Dupoux, E., Hammond, M., Meador, D., & Ohala, D. (in preparation). Some properties of the English metalinguistic syllable.

Hammond, M. (1990) The 'name game' in English and onset simplification, *Phonology*, **7**.

Hammond, M. (1991) Poetic meter and the arboreal grid, *Language* **67**, 240–259.

Hammond, M. (to appear) Resyllabification in English, *proceedings of FLSM*.

Hayes, B. (1987) *A Metrical Theory of Stress*. New York: Garland Press. (1980 MIT doctoral dissertation)

Hooper, J. (1972). The syllable in phonological theory, *Language*, **48**, 525–540.

Kahn, D. (1976). *Syllable-based Generalizations in English Phonology*. MIT doctoral dissertation, distributed by IULC.

McCarthy, J. (1982). Prosodic structure and expletive infixation, *Language*, **58**, 574–590.

McCarthy, J. (1979). *Formal problems in Semitic phonology and morphology*, MIT doctoral dissertation.

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

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

Mehler, J., Dupoux, E., & Segui, J. (1990). Learning constraints on models of speech perception. In J. Altmann (Ed.), *Cognitive Models of Speech Processing*, pp236–263, Cambridge, Mass: MIT Press.

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable ? Speech segmentation in Japanese. *Journal of Memory and Language*, **32**, 258–278.

Pérez, P. E. (1992). *Syncope in English*, manuscript, U. of Arizona.

Pisoni, D.B., & Luce, P.A. (1987). Acoustic-phonetic representations in word recognition, *Cognition*, **25**, 21–52.

Pulgram, E. (1970). *Syllable, Word, Nexus, Cursus.* The Hague: Mouton.

Sanchez-Casas, R.M. (1988). *Access representation in visual word recognition.* PhD dissertation, Monash University, Australia.

Scullen, M. E. (1993). *The Prosodic Morphology of French.* Indiana University doctoral dissertation.

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

Segui, J. (1984). The syllable: A basic perceptual unit in speech perception ? In H. Bouma & D.G. Bouwhuis (Eds.), *Attention and Performance X, Control of language processes*, Hillsdale, N.J.: Erlbaum.

Selkirk , E. (1982). The syllable. In H. van der Hulst & N. Smith (Eds), *The Structure of Phonological Representations, volume II*, Foris, Dordrecht, 337–383.

Treiman, R., & Danis, C. (1988). Syllabification of intervocalic consonants. *Memory and Language*, **27**, 87–104.

Treiman, R., & Zukowski, A. (1990). Toward an understanding of English syllabification. *Journal of Memory and Language*, **29**, 66–85.

Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition*, **15**, 49–74.

Zwitserlood, P., Schriefers, H., Lahiri, A., & van Donselaar, W. (1993). The role of the syllable in the perception of spoken Dutch. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **19**, 260–271.