FORBEAR IS A HOMOPHONE: LEXICAL PROSODY DOES NOT CONSTRAIN LEXICAL ACCESS*

ANNE CUTLER

MRC Applied Psychology   Unit,   Cambridge

Because stress can occur in any position within an English word, lexical prosody could serve as a minimal distinguishing feature between pairs of words. However, most pairs of English words with stress pattern opposition also differ vocalically: OBject and obJECT, CONtent and conTENT have different vowels in their first syllables as well as different stress patterns. To test whether prosodic information is made use of in auditory word recognition independently of segmental phonetic information, it is necessary to examine pairs like FORbear - forBEAR or TRUSty - trusTEE, semantically unrelated words which exhibit stress pattern opposition but no segmental difference. In a cross-modal priming task, such words produce the priming effects characteristic of homophones, indicating that lexical prosody is not used in the same way as segmental structure to constrain lexical access.

Keywords: speech recognition, prosody, stress, lexical access

INTRODUCTION

Human speech recognition extracts meaning from acoustic waveforms. Psycholinguists attempt to discover the processes by which this result is so effortlessly achieved, and engineers strive to build machines which will perform speech recognition with even an approximation to the efficiency of humans. Central to the concerns of both is the process known as lexical access. Because the number of potential utterances a recognizer might be presented with is infinite, the recognizer cannot hold in its memory, for eventual match against an acoustic input, the meanings of entire utterances. Instead, what is stored must be the meanings of the discrete units of which utterances are composed. We may refer to these units simply as words (begging the question of whether they could as well be morphemes or phrases). The part of memory in which the sound of a word may be matched with its meaning is called the lexicon.

How does the recognizer access an individual entry in the lexicon? Is the access code composed of phonetic segments (as assumed by, e.g., Foss and Gernsbacher, 1983, or Marslen-Wilson, 1980), of syllables (Mehler, Dommergues, Frauenfelder and Segui, 1981) or of spectral templates (Klatt, 1979)? This is an issue of lively psycholinguistic debate. The present research extends the debate to the question: Does the lexical access code

contain prosodic information?

This question is of particular relevance for English, and for other languages which, like English, have "lexical stress." Not all languages have an opposition between stressed and unstressed syllables, and of those that do, the majority have stressed syllables occurring at a fixed position in the word. If stress can only occur in one position for a given string of phonetic segments, then stress can obviously not be used to distinguish between semantically distinct words of otherwise identical form; thus, while stress information could well be of use in the recognition of fixed-stress languages, e.g., to suggest where word boundaries are located, it cannot be of importance for lexical access. In English, however, the placement of stress is not fixed, i.e., English is a *lexical-stress* language. This term *lexical stress* itself suggests that in such languages stress pattern can have a lexically distinctive function.

Stress is a prosodic phenomenon;[1] the opposition between stressed and unstressed syllables is expressed in the prosodic dimensions of fundamental-frequency movements, segment and syllable duration, and intensity. However, it also has segmental consequences. Pairs of English words with stress-pattern opposition usually also differ vocalically. Thus *OBject* and *obJECT, CONtent* and *conTENT*[2] have different prosodic realizations of their first and second syllables, but they also have quite different vowels in their first syllables. In each case, the member of the pair with initial stress (strong-weak prosodic pattern or SW) has a full vowel in the first syllable, while the word with second-syllable stress (WS) has schwa in the first syllable. Hence stress variation is usually expressed simultaneously in the segmental and suprasegmental (prosodic) domains, and processing effects of stress may consequently reflect segmental as well as prosodic influences.

Research on the processing of stress has shown that stressed syllables are more readily perceptible than unstressed syllables. In many speech-perception tasks, therefore, stressed syllables show a clear processing advantage. For example, monosyllabic words spliced out of context are more recognizable if they were stressed (Lieberman, 1963); clicks which occur on stressed syllables are more accurately located than clicks which occur on unstressed syllables (Bond, 1971); likewise, mispronunciations in stressed syllables are far more likely to be detected than mispronunciations in unstressed syllables (Cole, Jakimik and Cooper, 1978); phoneme-monitoring response time is faster if the target phoneme begins a stressed syllable (Shields, McHugh and Martin, 1974; Cutler and Foss, 1977); slips of the ear are *least* likely to be made on stressed syllables — inaccurate perception characteristically occurs on unstressed syllables (Browman, 1978). These results have been explained in terms of the greater acoustic salience of stressed syllables.

A number of recent studies have explicitly investigated whether stress plays a role in

"Stress" is properly an abstraction; stressed syllables are marked for stress, and this marking may or may not be realized in physical differences between syllables in a particular utterance. For a more detailed discussion of the relationship between stress marking and the physical properties of utterances, as well as the theoretical assumptions implicit in choice of terminology, see Ladd and Cutler (1983). For the purposes of this paper, it need only be affirmed that all stress contrasts used in the experiments were realized as clear physical differences between stressed and unstressed syllables.

Upper case will be used to represent a stressed syllable.

word recognition. Cutler and Clifton (1984) found several apparently negative indications. For instance, stress information alone does not seem to facilitate recognition: Neither visual nor auditory lexical decision could be facilitated by explicit blocking of materials by stress pattern. Nor does lexical stress appear to play an indirect part in word identification via grammatical categorization. In English, there is a tendency for bisyllabic words with initial stress (strong-weak, or SW, pattern) to be nouns and bisyllables with final stress (WS) to be verbs. However, whether or not a word conforms to this pattern has no effect at all on how quickly or how accurately its grammatical category can be recognized. One clear result of several studies, however, is that mis-stressing is inhibitory. Bansal (1966) found that English listeners presented with English spoken by Indian speakers, who frequently applied word stress in a manner unorthodox by British English standards, tended to interpret what they heard to conform with the stress pattern, often in conflict with the segmental information. Lagerquist (1980) found that puns are unsuccessful if they require a stress shift. Deliberately mis-stressed words presented in recognition tasks by Bond and Small (1983) and Cutler and Clifton (1984) were responded to more slowly than correctly stressed words.

The mis-stressing used in these studies, however, was not simply a prosodic manipulation. As pointed out above, stress is a linguistic feature which has both supra-segmental *and segmental* consequences. The stress shifts which were tested also involved vowel alterations. Thus any observed effect of stress could, logically, be simply a segmental phonetic effect; an observed processing difference between, say, *CONtent* and *conTENT* could simply be due to the vowel difference in their initial syllable, in the same way that the vowel difference in *cot* and *cut* is lexically significant. The purely prosodic difference between the syllables could be lexically irrelevant.

Accordingly, to investigate prosodic effects on word recognition in a lexical-stress language, it is necessary to control for vowel structure. Although most unstressed syllables in English have a neutral (schwa) vowel, a reasonably large class of polysyllabic words with exclusively full vowels does exist. *Nutmeg* and *typhoon* are two such words. Some linguists have argued that the opposition between full and reduced vowels is proso-dically more important than the opposition between stressed syllables and others (e.g., Bolinger, 1981); from a processing point of view, however, words with only full vowels offer the only opportunity to assess lexical prosodic effects without confounding segmental phonetic effects. If prosody *can* be exploited in lexical access, then the fact that *nutmeg* is stressed on the first syllable, *typhoon* on the second - a clearly perceptible prosodic difference — should be lexically relevant.

In the mispronunciation experiment referred to above, Cutler and Clifton (1984) explicitly compared bisyllabic words in which the unstressed syllable has a neutral vowel *(wisdom, deceit)* with words like *nutmeg* and *typhoon.* Word recognition was clearly inhibited by mis-stressing for the former group. The words with full vowels, however, were only harder to recognize in a mis-stressed version if their citation form pronun-ciation was SW. That is, *nutMEG* was much harder to recognize than *NUTmeg;* but *TYphoon* was not significantly more difficult than *tyPHOON.*

Cutler and Clifton attributed this result to the "stress shift rule" of English, whereby the demands of sentence rhythm can cause stress to shift, but (a) only to a full syllable: and (b) only to a syllable earlier in the word than the syllable marked for citation-form

stress. By this rule, *unknown* is stressed on its second syllable in "he is unknown" but on its first in "the unknown soldier." Words like *typhoon,* Cutler and Clifton argued, are in practice encountered sufficiently often in SW form that their SW form has achieved the lexical status of an optional pronunciation. They went on to claim that the significant inhibition of recognition in the *nutMEG* case was nevertheless evidence that prosody was important in word recognition.

The process of word recognition, however, includes several subsidiary operations, and there are at least two ways in which prosody could be relevant in the recognition process. These correspond to the commonly drawn distinction between lexical access and lexical retrieval. On the one hand, lexical prosody, i.e., stress marking, could be an essential part of the access code by which lexical entries are located; on the other, however, it could be part of the phonological code listed for a word in the lexicon and consulted only in retrieval, i.e., once access has been achieved.

The Cutler and Clifton mis-stressing results do not distinguish between these two possibilities. If prosody is a constraining feature of the lexical access code, *nutMEG* could be hard to recognize because the initial access will fail on encountering no match, and successful access will only be achieved after the code has been recomputed. If prosody does not constrain access, however, *nutMEG* could be hard to recognize because the phonological information contained within the lexical entry for *nutmeg* will show a mismatch with the acoustic signal.

Note that two conditions appear indisputable: firstly, that stress marking is indeed part of the phonological lexical representation. This has long been argued, on the basis of various lines of evidence — tip-of-the-tongue phenomena (Brown and McNeill, 1966), recognition memory (Robinson, 1977), cued word production (Engdahl, 1978), slips of the ear (Browman, 1978), and slips of the tongue (Cutler, 1980). Secondly, it must be the case that prosodic differences are perceived and coded in the phonetic representation into which the acoustic information in the signal is transformed. If prosodic differences were imperceptible, mismatches between acoustic information and lexical phonology could not be registered. However, prosody is clearly perceptible and phonetically computable; we can all tell a mis-stressing when we hear one, irrespective of whether a vowel change has occurred. The undecided question is whether the prosodic component of the computed phonetic representation can actually constrain the process of lexical access.

As was pointed out above, in lexical-stress languages prosody is available as a minimal distinctive feature between words. If prosody constrains lexical access in much the same way that segmental phonetic information does, then pairs of words which differ only in prosody should generate quite distinct lexical access codes and be, in practice, not confusable. It is an interesting fact, however, that in all the world's lexical-stress languages such pairs are extraordinarily rare. In English, for instance, although stress oppositions between verb and noun forms of the same stem *{decrease, conduct, import)* are common, a concentrated search for pairs of lexically clearly distinct words whose (British English) pronunciation differed only in prosody produced fewer than a dozen. They include *forbear, forearm, retail, insight/incite.* Nonce formations can result in further pairs *(Ovalise,* a neologism from my own collection, forms a stress pair in British English with *overLlES),* but such words are clearly unlikely to have established lexical represen-

tations. Other lexical-stress languages also contain remarkably few such pairs.

If prosody could constrain lexical access as effectively as segmental structure, one would expect languages to exploit it, so the extreme rarity of such "minimal stress pairs" as *forbear* might already suggest that lexical stress is perhaps not a very valuable source of such constraint. However, prosodic information could be of use in word recognition even if it did not minimally discriminate between words. For example, guessing a polysyllabic word from presentation of its initial syllable could in some cases be very much easier if the stress value of that syllable were included. Take, for example, the English words beginning with the syllable *for - (fore-, four-).* The majority of such words in which the following phoneme is /f/ are stressed on the first syllable: *forefather, forefront, forfeit,* etc. Only *forfend,* and in certain contexts *four-footed,* have second syllable stress. Some models of auditory word recognition (e.g., Marslen-Wilson, 1980) assume that a spoken word can be effectively recognized before its end if there are no other potential word candidates with the same left-to-right phonetic structure; thus a listener hearing *forfend* would have many fewer alternatives to choose from on hearing the second /f/ if it were clear that the first syllable was not stressed. The reverse would be true if the second syllable began with /g/; most such words have second syllable stress — *e.g., forget, forgive* and their morphological relatives.

Hence the rarity of minimal-stress pairs is not in itself an argument against prosodic constraints in word recognition. And the fact that such pairs do exist, even if only in small numbers, allows a direct experimental test of the question of the precise location of prosodic effects in word recognition. For if prosody does *not* constrain the lexical access code, a very simple consequence ensues: Pairs such as *forbear* are homophonous.

All that is then necessary to test this hypothesis is a diagnostic test for homophony. Fortunately, an experimental paradigm developed in recent years allows just such a test. This is the cross-modal priming paradigm (Swinney, 1979; Swinney, Onifer, Prather and Hirshkowitz, 1979). In this task, a listener is presented with an auditory sentence; at some point during the sentence, a visual target (a string of letters) appears on a screen, and the subject must decide whether or not it is a word. That is, the subject performs a visual lexical-decision task while also engaging in auditory sentence comprehension. Using this task, Swinney (1979) established the following effect: If the visual target appears immediately after a homophone in the sentence, then subjects respond to it faster when it is related to *either meaning* of the homophone. For instance, if the sentence contains the homophone *bug,* then both "ant" and "spy" are accepted faster than a non-related control word such as "sew." Swinney explained this effect as access of both readings of the homophone from the lexicon; only after access was it possible to choose the contextually more appropriate reading. That this choice is made, and quickly, is shown by the transience of the homophone priming effect: Only three-quarters of a second after the end of the homophone in the sentence, Swinney found, priming effects are present only for the contextually related reading.

If *forbear* is effectively homophonous, then subjects listening to a sentence containing either its SW or its WS version should show lexical-decision facilitation for words related both to the SW (ancestor) version and to the WS (be patient) version. If it behaves just like other homophones, then priming effects to the inappropriate version should disappear very rapidly. If, however, lexical prosodic patterns can constrain lexical access

and *forbear* is not homophonous, then only words related to the appropriate stress version should be facilitated. Thus the cross-modal priming paradigm allows a simple and direct test of the role of lexical prosody in lexical access.

This paper reports two cross-modal priming experiments in which words like *forbear* were manipulated in the same way that Swinney (1979) manipulated homophones. Setting up such experiments was not, however, a simple matter. In particular, choosing words which were reliably associated to the two members of the minimal stress pairs presented difficulties; because most such words, like *forbear,* are homographs, written association norms could not be exploited. Thus it was necessary to collect auditory-association norms simply in order to provide associates of proven reliability. In fact, to make the cross-modal priming experiments possible, it was necessary to run three preliminary control studies; these are briefly described in the next section, and serve also as an account of the methodological preparation of the priming study.

I. CONTROL EXPERIMENTS

Experiment 1

Materials, Eleven pairs of phonetically identical but prosodically distinct bisyllabic words were chosen, in each of which there was a clear semantic distinction between the two members, i.e., it could reasonably be assumed that the two did not share a lexical entry. The pairs were: *DIScount(N)/disCOUNT(V), FORbear(N)lforBEAR(V), FORE-arm{K)lforeARM{W}, FOREgoing/forGOing, GOATy/goatEE, IMpress(N)/limPRESS(V), INsight/inCITE,   RElay(N)lreLAY{V},   REtail(A)/reTAIL(V),   TRUSty/trustEE,* and *UNderground(N)/underGROUND(A).[3]*

Two lists were prepared. Each contained one member of each pair (with stress pattern counterbalanced as far as was possible) plus 16 other words of similar length and frequency (e.g., *mishap, contrive, parentage).* During the first presentation to a group of subjects it was noticed that the speaker neutralized the initial vowel of *incite* and pronounced *goatee* with SW stress pattern. Subsequent informal testing established that these pronunciations are quite common in British English. This means that neither *goaty/goatee* nor *insight/incite* was likely to be a true minimal stress pair for the subject population to be tested in the cross-modal priming experiments. For many speakers, *GOATee* would be a complete homophone, and *goatEE* a nonword; similarly, for many speakers *inCITE* with a full vowel in the first syllable would be an unusual pronunciation. Hence such speakers would not have only lexical prosodic differences between these words. A further problem for *inCITE* is that it is homophonous with the phrase *in SIGHT.* Accordingly, these pairs were dropped from further testing.

Subjects. One hundred and forty-two members of the Applied Psychology Unit volunteer subject panel were tested, in six groups of varying size. Seventy subjects heard list 1 and 72 heard list 2.

"Underground" with initial stress is a British English noun referring to a subway system or train.

Procedure. Subjects were given a response sheet with a numbered blank for each word to be presented, and were instructed to write down "the very first word that springs to mind" for each stimulus word.

The words were read out separately to each group, with sufficient time between trials for each subject in the groups to complete a response. Administration of the entire experiment took approximately five minutes. The two lists were each presented once by the author and once by each of two male native speakers of British English. On two occasions a stimulus word was presented with the stress pattern intended for the other list; this was corrected by introducing a compensatory stress change in a subsequent presentation of the other list. Because each group contained different numbers of subjects, this meant that individual words received different numbers of total responses, varying from 57 to 72.

Results. The responses were divided into three groups: words clearly related to the stimulus word, words clearly related to the other member of the stress pair, and other responses (which included illegible words, responses which could conceivably be related to either member of the pair — e.g., "strong" to *forearm,* responses which were very general or bore no obvious relation to the stimulus or only a syntagmatic relation — e.g.. "me" to *impress,* and "klang" responses — e.g., "tail" to *retail).*

Table la summarizes these results, averaged across the nine pairs used in the experiment. The results are expressed as percentages because individual words received differing numbers of total responses. It can be seen that the first group of responses is, as expected, the largest; a x^2 test on the first two columns of Table la is significant at beyond the 0.001 level. However, the fact that any responses fall into the second column at all is indicative that on some occasions the stimulus word's stress partner was accessed by the subject; typical responses in this category are "elbow" to *foreARM,* "ignore" to *Discount.*

Underground received many more responses falling into the "other" category than the remaining pairs. Accordingly, *underground* was eliminated from further testing. Table lb shows the distribution of association responses for the eight pairs used in the later experiments.

Experiment 2

Materials. For each of the eight stress pairs a related association response was chosen from the Experiment 1 data. This was the most frequent response except in two cases: "giving up," the most frequent response to *forgoing,* was rejected because only one-word stimuli were wanted; and "bank," the most frequent response to *trustee* was rejected because it is highly ambiguous. Each was replaced by the second most frequent response.

Since the words associated to the two members of a stress pair differed in length and frequency, each word was assigned a separate control word, matched as closely as possible on length in letters, frequency of occurrence (Kucera and Francis. 1967), word class and number of meanings. Three related/control pairs *(money-field, please-travel,* and *ancestor-dictator)* were not precisely matched on number of syllables or stress pattern. but these factors have no effect on visual lexical decision response times when frequency is controlled (Cutler and Clifton, 1984; Forster and Chambers, 1973). All words are listed

TABLE 1

Association Responses

a. Proportions of association responses to nine stress pairs

		Related to stimulus	Related to stimulus word's stress partner	Other
	SW	70.6	14.3	15.1
Stimulus Word				
	WS	53.3	26.0	20.7

b. Proportions of association responses to eight stress pairs used in later experiments

		Related to stimulus	Related to stimulus word's stress partner	Other
	SW	71.2	14.3	14.5
Stimulus Word				
	WS	56.7	26.6	16.7

TABLE 2

Lexical decision response times (msec), Experiment 2

	Stress pattern of source prime word	
	SW	WS
"Related"	602	594
"Control"	597	620

in the Appendix.

The 32 words to be used in the cross-modal priming experiments were randomly mixed with 27 nonwords in a lexical decision experiment. Twenty practice trials and six warm-up trials, in each of which half were non-words, preceded the experimental trials.

Subjects. Sixteen subjects drawn from the APU volunteer subject panel and the APU community took part in the experiment; subject panel members were paid for their participation.

Procedure. The words were presented centrally on a VDU screen; presentation and response timing were under the control of a PDP 11/23 computer. Display of each item was terminated by the subject's response. The next item was presented after a delay of 1.5 seconds.

Results. Lexical decision response times are presented in Table 2. The results were analysed in terms of the factors by which the words had been chosen: "related" vs. "control" words, separately for SW and WS prime words. As expected, an analysis of variance produced no significant results. Thus, any difference between related and control words when they are presented for lexical decision while being cross-modally primed from a sentence context may be attributed to effects of the context.

Although this result was as expected, a regression of RT on frequency of occurrence, separately for "related" and "control" sets, produced somewhat different functions: The "related" set showed a correlation coefficient of —0.407 and a slope of -0.214, the "control" set a correlation coefficient of —0.228 and a slope of -0.037. Thus, although these two sets are matched for frequency (and length), the matching is apparently imperfect. This is not unexpected since several items are of low frequency, and frequency ratings are notoriously unreliable in the lower ranges (Gernsbacher, 1984). This result strongly suggests that the lexical decision response times in the priming experiments should be corrected for the unprimed lexical decision times in this experiment.

Experiment 3

Materials. For each of the eight stress pairs, a pair of sentences was constructed. Each pair of sentences had an identical initial portion, to within one or two syllables of the occurrence of the prime word; after the prime words, the sentences diverged, but were of approximately the same length in words and syllables. The sentence pairs are listed in the Appendix. The initial portions were intended to be unconstraining with respect to following syntax or semantics; to check this, a completion experiment was performed on these portions of the sentences.

Subjects. Thirty-one subjects, 21 members of the APU subject panel and 10 APU staff and students, participated; the panel members were paid.

Procedure. The subjects were presented with the eight initial sentence fragments (the portions before the braces in the Appendix) and asked to supply a completion of approximately the same length (i.e., 6-12 words) in their own words. No time limit was set.

Results. No completion included a word which was actually one of the prime words in Experiments 4 and 5. The only response which was at all semantically related to any of the complete sentences used in Experiments 4 and 5 was "the long-awaited book finally appeared and reviews confirmed that it was a masterpiece," from one subject. A syntactic analysis of the subjects' completions showed that only in one case was one of the syntactic structures used in the later experiments significantly preferred over the other: Sentence 3 in the Appendix produced 84% continuations with *to* + infinitive as opposed to 10% *for* + noun and 6% other structures.

It was concluded that the chosen sentence frames were not semantically biased towards one prime word or the other. Only in one case was a syntactic bias detectable. Studies by Prather and Swinney (1977) (using the cross-modal priming task) and Tanenhaus, Leiman and Seidenberg (1979) (using a naming task) have shown that homophone priming effects resist syntactic constraints. It was therefore felt that this item would not behave significantly differently than the others.

II. CROSS-MODAL PRIMING EXPERIMENTS

In the next two experiments, the cross-modal priming technique was employed to test the degree to which the SW and WS members of a stress pair prime words related to their independent readings. The prime words were the eight stress pairs, embedded in the sentences of Experiment 3; the lexical decision target words were those generated by Experiment 1 and tested in Experiment 2.

In the first priming experiment, the target occurred immediately after the prime word. Under these circumstances Swinney (1979) found that all readings of homophones were primed, irrespective of context. Three possible outcomes for this experiment were hypothesized. If prosodic information is sufficient to constrain lexical access, then priming should occur only when the target word is related to the prime word. Thus *FORbear* should prime "ancestor" (in comparison with "dictator") and *forBEAR* should prime "tolerate" (in comparison with "simulate"), but not vice versa. If, on the other hand, prosodic information does not constrain lexical access, then *forbear* should be a homophone; both *FORbear* and *forBEAR* should prime both "ancestor" and "tolerate." A third possibility is that while SW words can be accessed by SW pronunciations only, WS words can be accessed either by WS or by SW pronunciations, as suggested by Cutler and Clifton (1984). In this case *FORbear* would prime both "ancestor" and "tolerate," but *forBEAR* would prime only "tolerate."

Experiment 4

Materials. Two tapes were recorded, each containing one version of each of the sentences listed in the Appendix, with stress pattern of prime word counterbalanced across tapes.

Four lexical decision lists were prepared, each containing one target word from each of the sets in the Appendix, with target-word condition counterbalanced across lists.

The experiment was conducted in two separate administrations, in that it was

embedded within the materials for two separate much larger cross-modal priming studies. The first of these contained 80 trials, on 40 of which the target was a non-word. Half of the word-target trials had related primes. The primary manipulation in this first matrix experiment was position at which the visual target appeared, so that many targets occurred earlier or later than the sentence-medial position used for the present study. The experimental trials were preceded by 30 practice trials, half of which had non-word targets.

In the second administration, the matrix experiment contained 34 trials, on 18 of which the target was a nonword, and on six of which a word target was preceded by a related prime. Targets occurred at various sentence positions, although in this case no target occurred either very early or very late in a sentence. The primary manipulation in this experiment was type of relation between target and prime. Complete descriptions of the matrix experiments can be found in Williams (1986).

The variable of first versus second administration was included in the analysis of variance as a between-subjects (unequal N) variable.

A timing mark, inaudible to the subjects, was placed on the tape coincident with the offset of each prime word.

Partway through the first administration, it was noticed that the target words for the matrix experiment included the non-word *roney,* and that this occurred prior to, and was likely to interfere with, the word *money,* which was a target item in one of the four lists of the present experiment. This problem did not occur in the second administration.

Subjects. Forty-eight members of the APU volunteer subject panel were tested in the first administration. Each combination of tape and lexical decision list was presented to six subjects. Forty undergraduate members of Cambridge University were tested in the second administration, five per condition. All subjects were paid for their participation. No subject had participated in any of the control experiments.

Procedure. Subjects were tested individually. They were instructed to pay attention to the content of the sentences but also to perform the lexical decision task as rapidly as possible. Attention to content was tested by a single-sentence recognition test after every eighth trial in the first administration, every sixth trial in the second administration. The timing mark on the tape triggered presentation of the lexical decision target word, in lower case, centrally on a VDU screen; presentation was terminated by the subject's response. Timing and data collection were under the control of a PDP 11/23 computer for the first administration, a microcomputer for the second.

Results. The three crucial effects are the main effect of relatedness (the overall priming effect), the interaction of relatedness with whether the target is matched to the prime word in the sentence context or to its stress partner, and the three-way interaction of these two effects with stress pattern of prime word. A main effect of relatedness and no interactions would support the second hypothesis, according to which stress pairs are homophonous. A relatedness effect when the target was matched to the prime word *(FORbear - "ancestor")* but not when it was matched to the prime word's stress partner *(FORbear - "tolerate")* would support the first hypothesis, according to which prosody discriminates stress pairs. Finally, the third hypothesis, postulating different effects

TABLE 3

Mean lexical decision response times (msec) from Experiment 4.
The numbers in parenthesis are raw response times adjusted for unprimed
baseline lexical decision time

		Related to prime word in sentence		Related to prime word's stress partner	
		Related	Control	Related	Control
Stress pattern of prime word in sentence	SW	710(108)	728(132)	678(84)	768(149)
	WS	677(83)	762(143)	696(94)	716(120)

for SW and WS members of stress pairs, predicts a significant three-way interaction - priming in all conditions except SW-matched targets preceded by WS primes (*forBEAR* - "ancestor").

The raw means for the eight conditions are shown in Table 3. An initial analysis of variance performed on these RTs showed a main effect of relatedness ($F1(1,86)$ = 12.39, p < 0.001). However, it was argued above that a more accurate estimate of priming effects in sentence context would be obtained if the raw RTs were adjusted for unprimed lexical decision times, since Experiment 2 showed some differences between response times to individual "related" targets and their controls. In fact, there was reason to believe that intrinsic differences between the lexical decision target words were having an effect in this experiment. There was markedly less priming effect (control RT — related RT) in the case of SW primes followed by words related to them, and WS primes followed by words related to SW primes (i.e., the top left and bottom right quadrants of Table 3) than for the other conditions. That is, the words related to SW primes produced less priming than the words related to WS primes. Table 2 shows that in Experiment 2 words related to SW primes were responded to slightly more *slowly* than their controls, while words related to WS primes were responded to *faster* than their controls. Hence intrinsic differences between the words even in the absence of priming context could be inflating the "priming" effect for WS-related words and reducing it for SW-related words. Accordingly it was decided to adjust the RTs for the unprimed baselines determined in Experiment 2.

Since cross-modal priming is a dual-task paradigm, it characteristically produces longer lexical decision RTs than does unprimed lexical decision. The present study was no exception: The grand mean for Experiment 2 was 603 msec, for Experiment 4, 717 msec. Adjusting for the unprimed baseline response was accomplished by subtracting from each response the mean RT for the same item from Experiment 2. The resulting

TABLE 4

Mean priming effect (adjusted control RT - adjusted related RT),
excluding first administration responses to "money" and its control, Experiment 4

		Related to prime word in sentence	Related to prime word's stress partner
Stress pattern of prime word in sentence	SW	36	65
	WS	60	35

means for the eight conditions are shown in parenthesis in each cell of Table 3. An analysis of variance on these adjusted RTs showed that the main effect of relatedness was again significant (F1(1, 86) = 8.54, *p* < 0.005). No other main effect or interaction (and hence neither of the critical interactions) reached significance; although the mean reaction time for the undergraduate subjects in the second administration was 30 msec faster than the mean for the first administration subjects, this difference was not significant (F1(1, 86) = 1.73).

An analysis of variance by materials was also carried out (although in this case the stimulus materials could be said to constitute the entire population of words meeting the relevant criterion in the language). Despite the very small number of item pairs, the adjusted RTs from Experiment 4, excluding *discount,* showed a significant effect of relatedness (F2(1, 13)= 7.0, *p* < 0.025). No other *F* ratios were significant.

The response times to the target word "money" in the first administration were inspected, and as expected they seemed to indicate an inhibition from the previously occurring non-word target "roney." The mean adjusted RT to "money" was approximately 130 msec *slower* than that to its control "field" both when primed *by Discount* and by *disCOUNT* in this administration, but on average 32 msec *faster* in the second administration. It was concluded that this item was not allowing an unbiased test of the hypothesis. Mean priming effects (adjusted control RT minus adjusted related RT) excluding first administration responses to "money" and its control are shown in Table 4.

Post hoc comparisons were carried out on the priming effect in the four conditions of Table 4 (i.e., excluding "money" and its control from the first administration). The SW primes produced significant priming effects when the related target matched the WS prime (r(87) = 2.55, p < 0.01), and near-significant priming effects when the related target matched the SW prime itself (r(87) = 1.40, *p* = 0.08). The WS primes produced significant priming effects with their own matched related words (f(87) = 2.55,p < 0.01) and near-significant priming when the related target matched their stress partner (t(87) = 1.37,p<0.09).

The pattern of results suggests that priming is present irrespective of the stress pattern of the prime word in the sentence. There is no indication that the priming effect is in general significantly weaker when the target does not match the sentence prime, nor is there any indication that SW and WS prime words differ in whether they exercise priming. In other words, the results support the hypothesis that stress pairs are effectively homophonous.

Experiment 5

Swinney (1979) found that when the target occurred a little later in the sentence than the prime word, only the contextually appropriate meaning of the prime word facilitated responses to associated words; priming effects of contextually inappropriate readings, which were present immediately after the prime word, disappeared. The next cross-modal priming experiment used delayed target presentation in the same way as Swinney's original study.

Materials. The tapes and lists were identical to those of Experiment 4 except that the click which triggered presentation of the visual target was placed, in accordance with the procedure of Swinney (1979), 750 msec (i.e., three to four syllables) after the offset of the prime word.

Again the experiment was administered twice, embedded within two separate larger experiments. The first of these was a follow-up study to that which formed a matrix for the first administration of Experiment 4, and contained a subset of the same sentences and visual targets, except that the nonword "roney" was replaced. The target-position manipulation was varied so that in this case a majority of targets occurred early in the spoken sentence. The experimental lists were preceded by 25 practice trials. The second embedding experiment was the identical experiment to that within which the second administration of Experiment 4 was embedded. In each case the change from Experiment 4 to 5 was effected by altering the computer program to introduce a 750 msec delay into the interval between timing mark occurrence and target presentation; thus, the recording and timing mark placement was controlled across experiments.

Subjects. Forty-eight members of the APU subject panel were tested in the first administration (six subjects per condition). Data from a further three subjects were discarded because of a high number of errors on the recognition trials. Thirty-two undergraduate members of Cambridge University took part in the second administration (four per condition). All subjects were paid. No subjects had participated in any of Experiments 1—4.

Procedure. The procedures were identical to those of Experiment 4.

Results. If *forbear* and similar pairs behave exactly as did monosyllabic homophones in the Swinney (1979) study, then we would expect in this delayed-target condition that priming would occur when the related word matches the prime word *(FORbear -* "ancestor" *forBEAR* - "tolerate") but not when it matches its stress partner *(FORbear* - "tolerate"; *forBEAR* - "ancestor"). Mean response times (and response times adjusted for baseline lexical decision time) for all conditions are shown in Table 5. It can be seen that precisely this predicted effect was found: With the adjusted RTs, the advantage

TABLE 5

Mean lexical decision times (msec), with lexical decision times adjusted for
unprimed baseline lexical decision time in parenthesis, from Experiment 5

		Related to prime word in sentence		Related to prime word's stress partner	
		Related	Control	Related	Control
Stress pattern of prime word in sentence	SW	703(101)	774(177)	727(133)	741(122)
	WS	709(98)	730(135)	724(126)	722(114)

for related words matching the prime over their controls was 76 msec for SW, 37 msec
for WS;[4] for related words matched to the prime's stress partner, however, there was
a small negative difference in each case. Analyses of variance on the adjusted RTs showed
that the main effect of relatedness did not reach significance ($F1(1, 78) = 3.9$, $p <
0.1$; $F2(1, 14) = 1.96$, $p < 0.2$), in contrast to the predicted interaction of relatedness
with matching vs. non-matching prime ($F1(1, 78) = 8.99$, $p < 0.005$; $F2(1, 14) = 5.39$,
$p < 0.04$). The difference between first and second administration was significant
($Fl(l, 78) = 9.47$, $p < 0.005$; $F2(l, 14) = 39.31$, $p < 0.001$), with the all-undergraduate
population of the second administration producing a mean response time 69 msec faster
than the first-administration subjects. However, this effect did not interact with any
other effect. No other main effect or interaction reached significance in either analysis.

Again, post hoc comparisons were carried out. When the "related" target was
associated with the prime word in the sentence, priming effects were significant (SW:
$f(79) = 3.01$, $p < 0.005$; WS: $r(79) = 1.91$, $p < 0.05$). When the "related" target was
associated with the prime word's stress partner, there was no significant difference
between response times to related and control words *($p > 0.5$ in both cases)*.

Thus the results were consistent with the conclusion drawn from the preceding
experiment: Stress pairs like *forbear* are functionally homophones.

It was expected that priming to the contextually related item might actually increase
further into the sentence, as has been reported by Shillcock (1982), and appears to
occur, if weakly, in Swinney's (1979) data. The increase from an overall relatedness
effect of 44 msec in Experiment 4 to the 56 msec effect in the matched prime con-
dition of Experiment 5 is not as large as the increase reported by Shillcock.

DISCUSSION

The finding that lexical access routines do not draw upon the information available from word prosody in English is in a sense surprising. It is unusual to find a source of potential information which is *not* exploited in speech recognition; in general, the speech processor appears to be capable of using any and every type of information available to it. It might be argued that little is lost by failing to discriminate between minimal stress pairs, since they are so few in number — another dozen or so homophones cannot make much difference to the language, after all; but as was pointed out in the introduction, the value which could potentially be extracted from prosodic information, in terms of reduction of number of possible word candidates, is far greater than simply distinguishing a few minimal pairs. It was also pointed out in the introduction, though, that in previous experiments prior knowledge of prosodic pattern has not facilitated lexical access at all; the present finding is, at least, consistent with those earlier results.

It is also consistent with other research on the pre-lexical stages of speech recognition, and in particular with the importance at this stage of cues to segmentation. Consider that in order to know the lexical prosodic structure of a word, the recognizer must know how many syllables the word has. This may seem utterly trivial, but in fact it is far from trivial in the recognition of continuous speech. In order to know how many syllables a word has, the recognizer must know where the word begins and ends; in other words, it must be able to locate word boundaries. Information about word boundaries is however usually not available in speech.

This greatly complicates the process of lexical access, since the recognizer has to decide what portions of the speech signal are appropriate candidates to match against lexical entries. Engineering solutions to word recognition in continuous speech have nearly always involved some form of matching of stretches of the acoustic signal against stored lexical templates; if such a template-matching system has no boundary information, it must start potential lexical searches at arbitrarily determined points in the signal, and the vast majority of these searches will be fruitless.

Psycholinguistic models of speech recognition have postulated intermediate levels of representation between the acoustic input and the lexicon — representation in terms of phonemes, for example (Foss and Gernsbacher, 1983; Marslen-Wilson, 1980), or syllables (Mehler, Dommergues, Frauenfelder and Segui, 1981). In comparison with simple acoustic template-matchers, such intermediate representations drastically reduce the number of potential lexical strings to be considered, but they do not directly address the problem of locating boundaries between lexical items.

Recently Norris and Cutler (1985) have argued that effective boundary location could, in fact, speed lexical access without the need for complete recoding of the input in terms of intermediate representational units. They further proposed (Cutler and Norris, in press) that in a stress language like English a strategy of boundary location might be based upon the prosodic structure of the signal. Specifically, listeners might treat metrically strong syllables (any syllable containing a full vowel) as potentially word-initial, while metrically weak syllables (any syllable containing schwa) would be assumed to be non-initial. Lexical access would be initiated from the onset of each strong syllable.

In support of this hypothesis, Cutler and Norris showed that words were more difficult

to detect in nonsense-syllable matrices when they belonged to two strong syllables (e.g., *mintayf)* than when they belonged to a strong-weak sequence *(mintef,* with schwa in the second syllable). They argued that this result reflected interference from postulation of a potential boundary prior to a second strong syllable; when the second syllable was weak, however, no boundary would be assumed and hence there would be no equivalent interference with detection of the embedded word. Similarly, Taft (1984) showed that listeners tend to segment ambiguous phonetic strings in such a way that strong syllables are word-initial.

In English, at least, the strategy would have a high rate of success: All of the most common lexical prosodic patterns in English have a full vowel in the first syllable (Carlson, Elenius, Granstrom and Hunnicutt, 1985).

Thus, there is some evidence that the prosodic structure of speech is indeed important during the prelexical stages of speech recognition. However, what is important is not *lexical* prosody (i.e., which syllable is marked for primary stress) but *metrical* prosody (which syllables are strong and which are weak). In terms of metrical structure, *FORbear* and *forBEAR* have identical patterns: strong-strong. Lexical prosody, i.e., the fact that *FORbear* is stressed on the first syllable and *forBEAR* on the second, would appear, from the results of the present experiments, to be irrelevant in the lexical access process. This may simply reflect the fact that at this stage of recognition lexical prosody *cannot* be unambiguously computed if word boundaries are not certainly known. Prosody may, however, be the route by which word boundaries are hypothesized and lexical access is most efficiently directed; that is, metrical prosody may be crucially important at precisely the stage at which we have found lexical prosody to be irrelevant.

There is certainly evidence that prosodic structure alone contains only rudimentary stress information. In a classic experiment, Lieberman (1965) presented phonetically trained listeners with speech which had been electronically modified to remove all segmental (and hence all syntactic and semantic) information. The full complement of prosodic information, however, remained. The trained listeners proved unable to identify different levels of stress. What they could identify was, in effect, metrical prosody - that is, they could effectively distinguish between stressed and unstressed syllables. Similarly, Nakatani and Shaffer (1978) found that listeners could not very successfully distinguish between primary and secondary stress levels in reiterant speech, though they could distinguish these from unstressed syllables. Two levels of salience appear to be all that the prosody alone can signal. Two levels of salience are what metrical prosody consists of, and metrical prosody offers the framework within which it can be demonstrated that prosody *does* play a role in lexical access, despite the fact that the lexical prosody of lexical-stress languages does not directly constrain the access process.

REFERENCES

BANSAL, R.K. (1966). The intelligibility of Indian English. Ph.D. thesis, London University.

BO LINGER, D.L. (1981). *Two Kinds of Vowels, Two Kinds of Rhythm.* Bloomington: Indiana University Linguistics Club.

BOND, Z.S. (1971). Units in speech perception. *Ohio State University Working Papers in Linguistics,* 9,1-112.

BOND, Z.S. and SMALL, L.H. (1983). Voicing, vowel and stress mispronunciations in continuous speech. *Perception & Psychophysics,* 34, 470-474.

BROWMAN, C.P. (1978). Tip of the tongue and slip of the ear: Implications for language processing. *UCLA Working Papers in Phonetics,* 42.

BROWN, R. and MCNEILL, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior,* 5, 325-337.

CARLSON, R., ELENIUS, K., GRANSTROM, B. and HUNNICUTT, S. (1985). Phonetic and orthographic properties of the basic vocabulary of five European languages. *Speech Transmission Laboratory: Quarterly Progress and Status Report,* 1/1985,63-94.

COLE, R., JAKIMIK, J. and COOPER, W.E. (1978). Perceptibility of phonetic features in fluent speech. *Journal of the Acoustical Society of America,* 64,44-56.

CUTLER, A. (1980). Errors of stress and intonation. In V.A. Fromkin (ed.), *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen and Hand* (pp. 67-80). New York: Academic Press.

CUTLER, A. and CLIFTON, C.E. (1984). The use of prosodic information in word recognition. In H. Bouma and D.G. Bouwhuis (eds.), *Attention and Performance X* (pp. 183-196). Hillsdale, N.J.: Erlbaum.

CUTLER, A. and FOSS, D.J. (1977). On the role of sentence stress in sentence processing. *Language and Speech,* 20,1-10.

CUTLER, A. and NORRIS, D.G. (in press). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance.*

ENGDAHL, E. (1978). Word stress as an organizing principle for the lexicon. *Papers from the Parasession on the Lexicon at the Fourteenth Regional Meeting, Chicago Linguistic Society* (pp. 138-147).

FORSTER, K.I. and CHAMBERS, S.M. (1973). Lexical access and naming time. *Journal of Verbal Leamingand Verbal Behavior,* 12,627-635.

FOSS, D.J. and GERNSBACHER, M.A. (1983). Cracking the dual code: Toward a unitary model of phoneme identification. *Journal of Verbal Learning and Verbal Behavior,* 22, 609-632.

GERNSBACHER, M.A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness and polysemy. *Journal of Experimental Psychology: General,* 113, 256-281.

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

KUCERA, H. and FRANCIS, W.N. (1967). *Computational Analysis of Present-Day American English.* Providence, R.I.: Brown University Press.

LADD, D.R. and CUTLER, A. (1983). Models and measurements in the study of prosody. In A. Cutler and D.R. Ladd (eds.), *Prosody: Models and Measurements* (pp. 1-10). Heidelberg: Springer.

LAGERQUIST, L.M. (1980). Linguistic evidence from paronomasia. *Papers from the Sixteenth Regional Meeting, Chicago Linguistic Society* (pp. 185-191).

LIEBERMAN, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech,* 6, 172-187.

LIEBERMAN, P. (1965). On the acoustic basis of the perception of intonation by linguists. *Word,* 21, 40-54.

MARSLEN-WILSON, W. (1980). Speech understanding as a psychological process. In J.C. Simon (ed.), *Spoken Language Generation and Understanding* (pp. 39-67). Dordrecht: Reidel.

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

NAKATANI, L.H. and SHAFFER, J.A. (1978). Hearing "words" without words: Prosodic cues for word perception. *Journal of the Acoustical Society of America,* 63, 234-245.

NORRIS, D.G. and CUTLER, A. (1985). Juncture detection. *Linguistics,* 23, 689-705.

PRATHER, P. and SWINNEY, D.A. (1977). Some effects of syntactic context upon lexical access. Paper presented at the American Psychological Association Convention, San Francisco.

ROBINSON, G.M. (1977). Rhythmic organization in speech processing. *Journal of Experimental Psychology: Human Perception and Performance,* 3, 83-91.

SHIELDS, J.L., MCHUGH, A. and MARTIN, J.G. (1974). Reaction time to phoneme targets as a function of rhythmic cues in continuous speech. *Journal of Experimental Psychology,* **102,** 250-255.

SHILLCOCK, R.C. (1982). The on-line resolution of pronominal anaphora. *Language and Speech,* 25, 385^*01.

SWINNEY, D.A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior,* 18, 645-659.

SWINNEY, D., ONIFER, W., PRATHER, P. and HIRSHKOWITZ, M. (1979). Semantic facilitation across sensory modalities in the processing of individual words and sentences. *Memory and Cognition,* 7, 159-165.

TAFT, L. (1984). Prosodic constraints and lexical parsing strategies. Ph.D. thesis, University of Massachusetts.

TANENHAUS, M.K., LEIMAN, J.M. and SEIDENBERG, M.S. (1979). Evidence for multiple stages in the processing of ambiguous words in sentence contexts. *Journal of Verbal Learning and Verbal Behavior,* 18,427-440.

WILLIAMS, J.N. (1986). The effective context for priming word recognition. Ph.D. thesis, University of Cambridge.

APPENDIX

Sentences and lexical-decision stimuli used in Experiments 2 to 5. For each sentence, the upper continuation is for the SW, the lower for the WS member of the stress pair; the words in parenthesis are 1) the word related to the SW prime, 2) the control for (1), 3) the word related to the WS prime, 4) the control for (3).

trusty old servant who
had worked for her father

1. The person that she was hurrying to see was the

trustee who was managing
her father's estate

(faithful, stubborn, guardian, anecdote)

with his forearm in plaster
and bruises on his face

2. The company secretary arrived at the board meeting

forearmed with knowledge
of the planned takeover

(elbow, folly, prepare, protect)

for discounts to be obtained
on members' personal purchases

3. The working party reported that it was impossible

to discount the rumours of
infiltration by radicals

(money, field, ignore, exceed)

that her forbears had been
hardy pioneer types

4. Gritting her teeth, she reminded herself

to forbear to mention her
grievance

(ancestor, dictator, tolerate, simulate)

' the relays should be dropped
from the schedule next year

5. The committee discussed the problem of whether

to relay the findings to
Head Office at once

(race, date, send, gain)

v.

ƒ with the impress of a little-
known specialist publisher

6. The long-awaited book finally appeared

and impressed everyone who
took the trouble to read it
thoroughly

(stamp, scrap, please, travel)

the foregoing sections would
have to be changed

7. After struggling for hours to get it right,
he decided

that forgoing breakfast had
been a bad idea

(previous, opposite, without, against)

into the retail trade instead
of taking up a profession

8. The couple was worried that their son might
be going

to retail the story of their
fight around the neighbourhood

(shop, wind, tell, keep)