Journal of Psycholinguistic Research, Vol. 22, No. 2, 1993

Phonological Cues to Open- and Closed-Class Words in the Processing of Spoken Sentences

Anne Cutler[1]

Evidence is presented that (a) the open and the closed word classes in English have different phonological characteristics, (b) the phonological dimension on which they differ is one to which listeners are highly sensitive, and (c) spoken open- and closed-class words produce different patterns of results in some auditory recognition tasks. What implications might link these findings? Two recent lines of evidence from disparate paradigms—the learning of an artificial language, and natural and experimentally induced misperception of juncture—are summarized, both of which suggest that listeners are sensitive to the phonological reflections of open- vs. closed-class word status. Although these correlates cannot be strictly necessary for efficient processing, if they are present listeners exploit them in making word class assignments. That such a use of phonological information is of value to listeners could be indirect evidence that open- vs. closed-class words undergo different processing operations.

INTRODUCTION

Words fall into a number of grammatical classes: nouns, verbs, adjectives, pronouns, articles, conjunctions, prepositions, and so on. These classes themselves comprise two main groups, sometimes referred to as "content words" vs. "function words," on the argument that members

Parts of the research reported in this paper were carried out in collaboration with Sally Butterfield and David Carter, and supported by the Alvey Directorate (United Kingdom). Jonathan Stankler's master's research was supported by the Science and Engineering Research Council (United Kingdom). Thanks to all of the above, and to Merrill Garrett, Mike Kelly, James McQueen, and Dennis Norris for further assistance.

[1] Address all correspondence to Dr. Anne Cutler, MRC Applied Psychology Unit, 15 Chaucer Rd., Cambridge CB2 2EF, United Kingdom.

0090-6905/93/0300-0109$07.00/0

of one group can be said to possess a context-independent meaningful content, while members of the other group principally exercise a syntactic function. A similar argument underlies the use of "lexical words" and "grammatical words" to express the same distinction; the implication here would be that only members of the former group have proper entries in a language user's main lexicon. A more neutral pair of terms is the "open class" and the "closed class"; this is based on the fact that, whereas the set of, for example, nouns in a language is in principle infinite, and one can invent a new noun with no trouble at all, the set of, say, conjunctions is closed in the sense that it would be very difficult to imagine inventing a new one. Nouns, verbs, adjectives, and adjectival adverbs are open classes; all the other classes are closed.

This operational definition based on class extendability seems preferable to the alternative terms because it is more neutral while the alternatives seem more theoretically loaded; but it also appeals on the grounds that the various categories within the two groups do in fact differ in some possibly relevant ways. Thus, for example, nouns form the largest of the open classes, and are also in a sense the "most open" in that new nouns are formed more frequently than new verbs (Kelly, 1992). Likewise, within the closed class, forms such as prepositions and pronouns are more numerous than, say, articles and complementisers, and the former are more likely than the latter to be semantically highlighted in a sentence by being deployed in a contrastive construction ("in the world but not of it")—although of course such constructions are far more common again with open-class words ("stirred, not shaken"). Whether any such distinction has a reflection in sentence processing is an empirical question.

In comparison with some other languages, English and its European relatives have a relatively large number of closed-class items which are realized as separate words. Most of them are short, but otherwise they are orthographically not particularly distinct from short open-class words; thus there is no obvious word-class clue in *bus* vs. *but*, *thin* vs. *this*, *whet* vs. *when*. Under these circumstances, it has been possible to conduct a large number of cleanly designed experiments in English (and other European languages) in which the processing of open- and closed-class words has been contrasted. The rationale for such studies is precisely the intuition underlying the distinctions described above: The two groups of words may play different roles in conveying linguistic messages, and therefore if we can establish exactly how they are processed, we may make progress towards an overall model of sentence processing.

Many such studies of open- vs. closed-class word processing have

been conducted in the visual modality, as of course has most work on sentence processing. For some topics in sentence processing, modality of presentation is unlikely to affect the results. Where clear reflections of a distinction are present in one modality but not in another, however, it is obviously worth considering results from each presentation modality separately. This paper examines the role of open- and closed-class words in the processing of spoken sentences—which is, for most people, a more frequent experience than the processing of written sentences. As the next section shows, the orthographic similarity of open- and closed-class words in English is not matched in the auditory domain.

PHONOLOGY AND THE OPEN/CLOSED-CLASS DISTINCTION

In English (and in phonologically similar languages) the open/closed-class distinction is directly reflected in sentence-level prosodic structure. English is a stress language, and English syllables are either strong (i.e., contain a full vowel) or weak (i.e., contain a reduced vowel, usually schwa). With virtually no exceptions, open-class words in spoken English sentences contain at least one strong syllable. Closed-class words, on the other hand, are frequently realised as weak syllables only. There is no orthographic reflection of this: The difference is present in speech only.

This asymmetry in phonological realization of the open and the closed classes has, in turn, a very great impact on the structure of spoken English. This is because of the relative frequency of occurrence of members of the two groups. There are only a few hundred closed-class items in the English vocabulary, so they form only a tiny fraction of the words any language user knows; among them, however, are all the most frequently used words in the vocabulary. In fact, less than 1% of the word *types* that speakers know manage to account for at least 50% of the word *tokens* which they hear in everyday speech.

Cutler and Carter (1987) examined the words making up a corpus of spontaneous British English speech—the *Corpus of English Conversation* (Svartvik & Quirk, 1980). There were nearly 200,000 tokens in this corpus, of which 59% were in fact closed-class words. Assuming that the latter were, where appropriate, realized in their most common (i.e., weak) phonological forms, it was possible for Cutler and Carter to examine the distribution of weak and strong syllables throughout the entire corpus. Although closed-class words made up well over half of

the word tokens, they accounted for just *less* than half (49%) of the syllable tokens (because, of course, many more open-class than closed-class words in the corpus were polysyllabic). Of all the strong syllables in the corpus, 86% occurred in open-class words and only 14% in closed-class words, while for weak syllables the pattern was quite different: 72% occurred in closed-class words, 28% in open-class words. The most likely realization of a strong syllable was as the sole or initial syllable of an open-class word (74% of all strong syllables), while the most likely realization of a weak syllable (69% of the total) was as the sole or initial syllable of a closed-class word.

For American English, some relevant data are available in a small (1500-word) spoken corpus gathered by Waibel (1988). According to the definition of the open vs. closed classes used here, Waibel's corpus contained 941 open-class and 558 closed-class words. Waibel reported for each word class the proportion of tokens bearing (any level of) sentence stress; 87.9% of all open-class words were, by his definition, stressed, but only 24.4% of closed-class words. It is clear, then, that the phonological differences in the realization of open- and closed-class words in English are robust and consistent, and should be highly accessible to language users. A prerequisite for such accessibility, however, is that the distinction between strong and weak syllables should be salient to English language users. The next section reviews some evidence that this is indeed the case.

THE PROCESSING OF STRONG AND WEAK SYLLABLES

The distinction between strong and weak syllables, as outlined above, is based on vowel quality: Strong syllables are those which contain full vowels, while weak syllables are those which contain reduced vowels. Of course, this is almost the same as making a distinction between stressed vs. unstressed syllables. Consider the four words *generous*, *generic*, *generate*, and *generation*; their strong–weak patterns are SWW, WSW, SWS, and SWSW, respectively. It can be seen that stressed syllables are necessarily strong, and weak syllables are necessarily unstressed. However, it is in this case irrelevant that primary stress falls on the first syllable in *generate* but on the third in *generation*—the first and third syllable of each count as equal (i.e., strong in both cases). Vowel quality is the dimension upon which stressed syllables are acoustically more informative than unstressed syllables (Altmann & Carter, 1989). Evidence that the perceptual distinction is based on vowel quality was pro-

vided by a study by Cutler and Fear (1991), in which vowels were cross-spliced between syllables with primary stress (e.g., the initial syllable of *audience*), syllables with secondary stress (e.g., the initial syllable of *auditorium*), unstressed syllables with full vowels (e.g., the initial syllable of *audition*), and unstressed syllables with reduced vowels (e.g., the initial syllable of *addition*). Listeners' acceptability ratings of the cross-spliced forms grouped the crucial unstressed–unreduced case with the stressed vowels rather than with the unstressed–reduced vowels, i.e., the listeners preferred to make a distinction on the basis of vowel quality rather than of stress level.

Vowel quality, of course, is an absolute property of syllables, whereas stress level is necessarily relative—a syllable's stress level is determined only in relation to other syllables in a word or sentence. It is difficult and in some cases impossible for listeners to use relative distinctions "on line," i.e., in the course of processing a spoken sentence. Stress level, for example, could only be used to guide the lexical access process if listeners were to wait until the end of the word before beginning to process it, since only by the end of the word would it be clear what each syllable's stress level was. However, studies of spoken word recognition do not lend support to any notion that listeners wait; processing appears to begin as soon as the word begins to be heard (Marslen-Wilson & Welsh, 1978). Consistent with this, experimental evidence indicates that purely prosodic cues to lexical stress are ineffective in guiding lexical access (Cutler, 1986). There is no reason why vowel quality cues to stress should be similarly ineffective, however; that is, the vowel-quality-based distinction between strong and weak syllables, being an absolute property of syllables, requires no postponement of processing and hence is available for on-line use.

Indeed, recent evidence suggests that listeners do exploit the distinction between strong and weak syllables on line, namely, in confronting the task of speech segmentation. The recognition of spoken sentences presents listeners with a problem which does not arise in the recognition of isolated spoken words, and does not confront readers of written texts (at least in most orthographies). The act of recognition is the identification of an input as something we already know; what we already know is not the whole of an input utterance, because human memory is not infinite, and it would be impossible to store in our memories every complete utterance we might ever hear. Therefore the entries in our mental lexicon must be *discrete*, and recognition will involve finding these discrete lexical units as sound patterns in the speech signal input, and matching them to lexical entries in order to determine their meaning.

The problem for listeners arises in the fact that speech is continuous: Lexical unit boundaries are not reliably marked. Finding the boundaries between such units, i.e., segmenting the speech signal, is therefore a nontrivial task for listeners.

Consider, however, the finding of Cutler and Carter (1987), described above, that about three-quarters of *all* strong syllables in a speech sample were in fact the initial syllable of an open-class word. It would be quite a good bet for listeners to assume in the first instance that any strong syllable they encountered was the beginning of a new open-class word, and to use this heuristic as a first approach to speech segmentation. In fact, less than 10% of the open-class words in the corpus analyzed by Cutler and Carter began with weak syllables, so that even such a simple heuristic would have correctly located over 90% of open-class word onsets; this success rate is presumably high enough to make this heuristic very useful in the processing of spoken sentences.

Evidence that listeners do behave according to such a principle comes both from laboratory studies and from naturalistic observation. Firstly, Cutler and Norris (1988) asked listeners to perform a task called "word spotting," which consisted of deciding whether or not a nonsense bisyllable began with a real word. They found that a word like *mint* is harder to detect in *mintayf* (two strong syllables) than in *mintef* (a strong and a weak syllable). They explained this result by suggesting that the second strong syllable in *mintayf* triggers segmentation, so that detection of the embedded word requires assembly of speech material across a point at which speech has been segmented.

Cutler and Butterfield (1992) studied the pattern of errors which listeners make when they misperceive word boundaries in continuous speech. In spontaneous slips of the ear, they found, listeners significantly more often mistakenly insert word boundaries before strong syllables than before weak syllables; mistaken deletions of word boundaries, on the other hand, occur significantly more often before weak syllables than before strong. For example, the misperception of *bought a Mercedes* as *Mortimer Sadie's* involves deletion of the boundaries before the second and third syllables of the utterance (both of which were weak), and insertion of a boundary before the (strong) fourth syllable instead. Cutler and Butterfield also conducted an experiment in which they presented listeners with very faint speech; this elicited the same pattern of mistakenly inserting boundaries before strong syllables and mistakenly deleting boundaries before weak syllables (e.g., *conduct ascents uphill* was reported as *the doctor sends her bill*). Listeners gave clear signs of using

the distinction between strong and weak syllables in dealing with the problem of speech segmentation.

It is clear, then, that listeners can exploit the distinction between strong and weak syllables; thus they possess the ability which would be needed for extracting information about word classes from this prosodic dimension. Would such information, however, be of any use to listeners in the process of understanding? To answer this question, we need to know whether there are any differences in the way open- vs. closed-class words are processed in spoken sentences; the following section examines the available evidence.

THE PROCESSING OF SPOKEN OPEN- AND CLOSED-CLASS WORDS

A few studies of the recognition of spoken sentences have suggested that there may indeed be processing distinctions between open- and closed-class words. In certain aphasic syndromes patients have particular difficulty with closed-class words, and this difficulty extends to auditory recognition (Swinney, Zurif & Cutler, 1980; Zurif & Blumstein, 1978). Kean (1978) has argued that the auditory recognition difficulty does not necessarily reflect the functional differences between the two word classes, but may arise solely from the phonological differences described in the section on phonology above.

Studies with normal listeners include a gating experiment conducted by Bard, Shillcock, and Altmann (1988), in which listeners heard samples of spontaneous conversational speech in successively larger sequences, each sequence incremented by a whole word in comparison to the previous sequence. Approximately 20% of words correctly recognized under these conditions were not recognized on their first presentation, but only after subsequent context had been heard (and, on average, more than one word later). The majority of these late-recognized words were short closed-class words such as *is*, *a*, and *on*.

Salasoo and Pisoni (1981) found that noise bursts which replaced open-class words in a spoken text were far more detectable than noise bursts which replaced closed-class words; this was true even when the duration and amplitude of the noise replacing each type of word was held constant. However, they did not control the *relative* salience of noise bursts, and this could well differ as a function of word context—

for instance, bursts replacing open-class words might be surrounded by prosodically weak closed-class words, while bursts replacing closed-class words might be surrounded by prosodically strong open-class words, making the latter group of bursts relatively less salient. Although Salasoo and Pisoni interpreted their finding as evidence that listeners pay more attention to understanding open-class words in speech recognition, their result should at present be viewed with some caution.

A result originally reported from visual lexical decision by Bradley (1978) was replicated in auditory lexical decision by Matthei and Kean (1989): Nonwords beginning with a real word (*lostner*) take longer to reject in the lexical decision task than nonwords with no embedded word (*vostner*), except that embedded words which are members of a closed class (*mostner*) produce no such interference. Bradley interpreted her original result as evidence of differing lexical access processes for the open vs. the closed class, but Matthei and Kean ascribed the result to the postaccess consequences of lexical retrieval of grammatical category.

Grosjean and Gee (1987), reviewing the evidence to that time on recognition of open- and closed-class words, suggested that apparent auditory processing differences for the two classes may (as Kean, 1978, suggested for the aphasic deficits) reflect purely phonological factors. Closed-class words, being phonologically weaker, are simply harder to perceive in general. Grosjean and Gee propose a continuum of phonological salience, from open-class monosyllables in accented position through deaccented open-class words, closed-class items of various kinds, affixes, down to unstressed syllables in polysyllabic words. They pointed out that many closed-class words can be accented and hence can become salient in appropriate contexts ("I want you IN the bath"; "Are you THE Madonna?")—such realizations may well be processed differently from the more common phonologically weak forms. The effects of open/closed-class status and sentence accent were, in fact, manipulated orthogonally in a phoneme-monitoring study by Cutler and Foss (1977). Phoneme targets occurred on either open-class [as in example (1) below] or closed-class words [as in example (2) below], which could be accented or not [as in the (a) and (b) examples of each, respectively]:

(1) target: /k/ (a) Does John really want to KEEP that old van?
(b) Does John really WANT to keep that old van?

(2) target: /b/ (a) I'm not sure Shakespeare's plays are even BY Shakespeare.
(b) I'm not sure Shakespeare's plays are EVEN by SHAKEspeare.

Response times were faster to accented than to unaccented targets, but

were not affected separately by open/closed-class status. This result was replicated by Swinney, Zurif, and Cutler (1980), and by Cutler and Swinney (1987), in each case with the word-monitoring task. Friederici (1985), using closed-class words which were very predictable (as measured by a sentence completion task), found that word-monitoring response times to closed-class targets were actually faster than to open-class targets. (In Friederici's materials, which were in German, the closed-class targets were *not* phonologically weak.)

The evidence from auditory recognition studies is, as can be seen, quite sparse. Moreover, explicit comparisons of processing in the visual and auditory modalities are lacking. However, it is probably fair to conclude that there is no evidence that processing of closed-class words presents difficulty to the listener, despite their typical realization in phonologically weak form. The gating study by Bard, Shillcock, and Altmann (1988) found that prosodically weak closed-class words were only recognized well after their acoustic offset; but so were short open-class words (see also Grosjean, 1985). That is, closed-class words may in practice be hard to perceive, but in principle they are not otherwise hard to process.

In fact it would be highly surprising if they *were* hard to process; recall that closed-class words make up more than 50% of all word tokens occurring in typical speech samples (Cutler & Carter, 1987). If this high a proportion of all words we hear were to cause processing difficulty, then at the very least one might feel that our processing mechanism was not functioning optimally. There is even a sense in which closed-class words could be thought of as enjoying something of a processing advantage, since it is always possible to realize them as strong syllables, while it is not usually possible to realize open-class words as weak syllables. From the study by Cutler and Foss (1977), we know that closed-class words are not put at any perceptual disadvantage by being spoken with contrastive sentence stress. The reverse manipulation, however (realizing open-class words as weak syllables), *would* presumably cause processing difficulty. Although such a study has not been carried out, its results can easily be predicted; since open-class words have at least one strong syllable, realizing such words as only weak syllables would necessarily result in vowel quality alteration, and we know from studies of misstressing that precisely alterations of vowel quality have the most deleterious effects on recognition of spoken words (Bond & Small, 1983; Cutler & Clifton, 1984). Similarly, investigations of elliptic speech (i.e., speech which has been systematically distorted—for instance, by replacing all voiced stops with unvoiced stops, or the like) show that the

manipulation which most impairs intelligibility is to alter vowels in stressed syllables (Bond, 1981).

Differences in difficulty of processing may, then, be dismissed; but the question as to whether there exist differences in *type* of processing between open- and closed-class words is still open. This question will be considered further below, in the light of evidence, to be presented in the following two sections, that listeners can indeed exploit phonology to make decisions about word class in processing speech.

EVIDENCE FROM THE LEARNING OF AN ARTIFICIAL LANGUAGE

A first question that may be posed is whether native speakers can generalize from their underlying knowledge of the relationships between phonology and word class in English. In other words, would English language users apply to new linguistic material the knowledge which they have abstracted from their linguistic experience in this respect? Would there, as a consequence, be a processing advantage for linguistic material which conformed to the English pattern of word class phonology, and a processing disadvantage for material which did not conform?

Hypotheses about relative learnability of different types of linguistic material can be tested by presenting language users with input in an artificial language. Recent studies of artificial language learning have shown, for instance, that explicit markers of syntactic structure assist the learning of syntactic patterns (in fact, without explicit markers the patterns are virtually unlearnable; Green, 1979); the usefulness of the markers is greater the higher their frequency of occurrence relative to other items of the language's vocabulary (Valian & Coulson, 1988).

A master's thesis by Jonathan Stankler (1991) at the University of Cambridge investigated whether subjects learning an artificial language would be helped by the presence of a pattern analogous to the English relationship between prosodic structure and open- vs. closed-class status. To do this it was first necessary to ensure that some distinction similar to that between open- and closed-class words was built into the artificial language's structure. In fact this was a simple matter, because the work on syntax markers, described above, had used just such a distinction: Markers are typically a small set of high-frequency items (i.e., the equivalent of a closed class), while the rest of the artificial language's vocabulary amounts to a larger set of items all of which have a lower frequency of occurrence (the equivalent of the open class). Thus the grammar used

by Green (1979), and also by Valian and Coulson (1988) incorporated a distinction between two such word classes as its principal feature. The grammar consisted of the following five rules:

(2) S → P
S → P dD
P → aA cC
P → cC aA
P → aA bB cC

Each lower-case terminal element had exactly one realization, whereas upper-case elements had from three to six possible realizations.

In the version of the language implemented by Stankler (1991), all terminal elements were monosyllabic. The language input to the subjects was presented auditorily from prerecorded tapes. In one version of the language, "Weak–Strong", all the terminal elements represented in (2) by lower-case letters, i.e., the "closed-class" elements, were spoken as weak syllables, while all the upper-case, "open-class" elements were produced as strong syllables. This version of the language thus incorporated the mapping of prosody to word class characteristic of English, and was therefore predicted to be easy for English listeners to learn.

In contrast, another version was predicted to be hard to learn, because it contained no consistent mapping between prosody and word class at all. In this version, "Random", both the "open" and the "closed" classes were composed of some elements which were realized as strong syllables and some which were realized as weak syllables. Listeners presented with this version would simply not be able to use prosodic information, so they would have one source of information less than listeners presented with a consistently mapped version.

Two further versions of the language were prepared, which acted as control conditions. If the "Weak–Strong" version indeed proved relatively easy to learn, this could be because it reflected the prosody to word class mapping of English, or it could be simply because it embodied a mapping which was consistent. If the latter were the case, then any consistent mapping should work as well—for example, one which reversed the English mapping by realizing "closed-class" elements as strong syllables and "open-class" elements as weak syllables. Accordingly the third version, designed to test this hypothesis, embodied a "Strong–Weak" mapping.

One difference between these three versions lay in the number of individual vocabulary items realized as strong vs. weak syllables. "Weak–

Strong" had four weak syllables (all the "closed-class" items) and 18 strong syllables (all the "open-class" items); "Random" had 11 weak and 11 strong, distributed equally across classes; "Strong–Weak" had 18 weak and 4 strong. It could be that having more vocabulary items realized as strong syllables would increase the ease with which a language could be learned; if so, such an effect would favor the "Weak–Strong" version, predicted to be easy for other reasons. As a further control, then, the fourth version of the language, "Strong–Strong", realized every element in both classes as a strong syllable.

Although Stankler (1991) used the grammar of Green (1979) and Valian and Coulson (1988), he could not use their methodology because those previous studies had used visual presentation. Auditory presentation of artificial language input had been used in several investigations (e.g., Braine et al., 1990; Morgan, Meier, & Newport, 1987). Of these, the study by Morgan et al., which compared a consistently mapped prosodic variable with inconsistent mapping in the same dimension, seemed closest to Stankler's study, so he chose to mimic the methodology of that study. Thus the experiment consisted of four iterations of the following sequence: (a) presentation to the subject of 40 grammatical spoken sentences of the language, each accompanied by visual presentation on a VDU screen of a written version of the same sentence plus a "reference field," i.e., graphic symbols associated with each vocabulary item; (b) administration of a written vocabulary test in which pairing of the graphic symbols with their associated vocabulary items was tested; (c) administration of a sentence structure test, in which subjects were asked to choose between pairs of strings, with only one of each pair being a grammatical sentence of the language.

Morgan et al. (1987; see also Morgan & Newport, 1981) used both types of test to control for intergroup ability differences. They found performance differences between the groups which had had the consistent vs. the inconsistent versions of the prosodic variable which they manipulated (intonational cues to phrase boundaries); however, these differences showed up only in the test which measured subjects' learning of sentence structure. No intergroup differences appeared in the vocabulary test results, a finding which they used to argue that the difference in results on the sentence structure test was not simply a function of different levels of learning ability in the separate groups, but really reflected differences in the difficulty of learning structural relations in the grammar.

Consistent with this, Stankler (1991) also found no significant intergroup differences on the vocabulary test he administered. On the sen-

tence structure test, however, the groups differed: The Weak–Strong group performed significantly better than all the other three groups (which did not differ significantly from one another). The results are shown in Fig. 1.

Stankler (1991) concluded that the subjects in his experiment had found the Weak–Strong version of the language easiest to learn precisely because it embodied a consistent mapping of prosody to word class, and this mapping was in the same direction as that found in English. In other words, his study provided evidence that listeners both know and can exploit the correlations found in the English vocabulary between open vs. closed word class and prosodic structure.

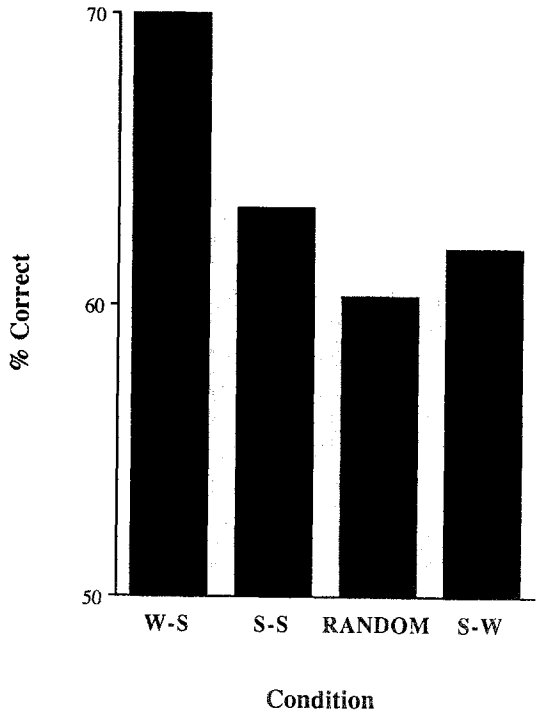

Fig. 1. Mean percentage of correct responses in the sentence structure test for four prosodically variant versions of a miniature artificial language. W – S = Weak–Strong version; S – S = Strong–Strong version; S – W = Strong–Weak version. (From Stankler, 1991.)

EVIDENCE FROM JUNCTURE MISPERCEPTION

A reasonable next question is whether listeners can exploit their knowledge of the phonological correlates of word class on line. This question has not, it appears, been put to direct experimental test. Some indirect evidence, however, can be obtained from further analysis of the results of Cutler and Butterfield (1992) described above in the section on the processing of strong and weak syllables. Cutler and Butterfield undertook an explicit analysis of word class assignments in listeners' mis-segmentations. The rationale for this analysis was provided by the segmentation algorithm which Cutler and Carter (1987) proposed on the basis of their statistical analysis of the English vocabulary. Recall that they found that strong syllables were most likely to be the sole or initial syllables of open-class words, while weak syllables were most likely to be closed-class words. Cutler and Carter suggested the following main features of a segmentation algorithm, in which exploitation of prosodic cues to word class was explicitly envisaged:

(1) 1.1 The main lexicon contains only open-class words; closed-class words constitute a separate list.
1.2 An initial segmentation process scans the input and places markers at the onset of each strong syllable.
1.3.1 If the initial string of the current input is not preceded by a marker, it is submitted to the closed-class list; if it is preceded by a marker, it is submitted to the open-class list.
1.3.2 The lookup process in both lists returns the longest candidate consistent with the input, *except that* the occurrence of a marker indicating the beginning of a strong syllable will terminate the current lookup process and initiate a new lookup process in the open-class list.

One of the predictions tested by Cutler and Butterfield (1992) against their corpus of spontaneous and laboratory-elicited mis-segmentations concerned the word class mapping of strong vs. weak syllables in cases where a word boundary was mistakenly inserted (i.e., the input contained no boundary at that point, but a listener reported a boundary). The above algorithm, they argued, would predict that, where a word boundary was mistakenly inserted before a strong syllable, the following word should be an open-class item, but where a word boundary was mistakenly inserted before a weak syllable, the following word should be a member of the closed class. Figures 2 and 3 show the results of their analysis for the spontaneous and the laboratory-induced boundary insertion errors, respectively.

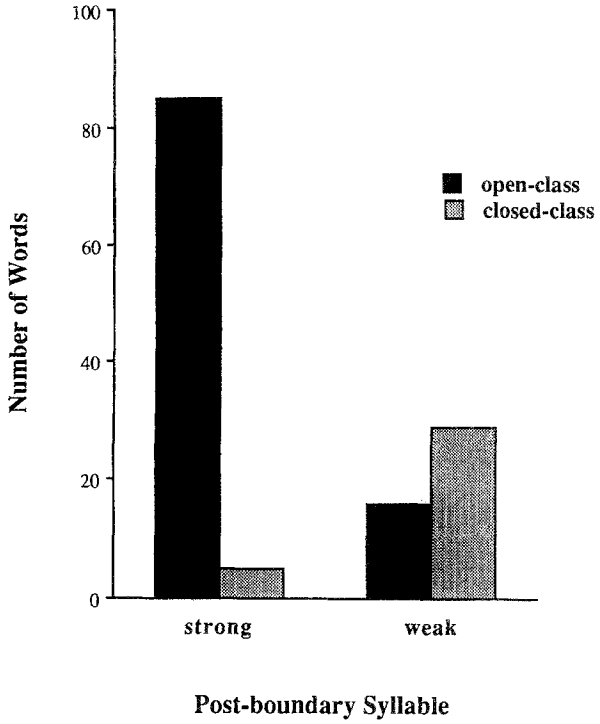

Fig. 2. Occurrence of open-class vs. closed-class words following inserted word boundaries before strong versus weak syllables in spontaneous slips of the ear. (From Cutler & Butterfield, 1992.)

It can be seen that the prediction is strongly confirmed: Where a boundary is inserted before a strong syllable, what follows is more likely to be an open-class word (such as in *economists* heard as *the communist*); but where a boundary is inserted before a weak syllable, what follows is more likely to be an closed-class word (e.g., *dusty senseless drilling* heard as *thus he sent his drill in*). Cutler and Butterfield (1992) concluded that listeners processing spoken English operate on the assumption that strong syllables are highly likely to be the initial syllables of open-class words, whereas weak syllables are more likely to be closed-class words. The strong phonological reflections of word class in the English vocabulary are part of language users' competence and, moreover, have been incorporated into their everyday processing procedures. Furthermore, the evidence is consistent with on-line exploitation of this knowledge of the

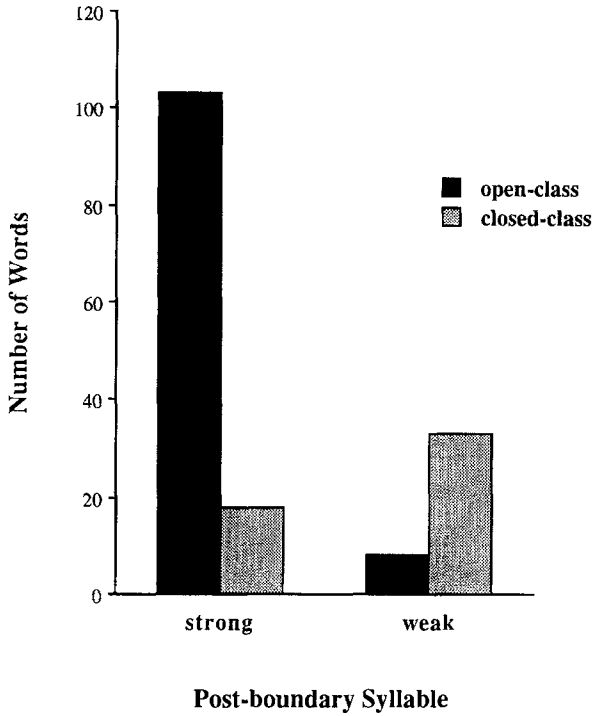

Fig. 3. Occurrence of open-class vs. closed-class words following inserted word boundaries before strong versus weak syllables in faintly heard speech. (From Cutler & Butterfield, 1992.)

prosodic correlates of word class. The studies concerned listeners who were presented with ordinary speech under conditions in which, for naturally occurring or laboratory-manufactured reasons, perception was made difficult; such conditions occur quite frequently for most listeners. The results showed that under these conditions listeners base word class judgments on prosodic characteristics of the input in such a way that their judgments accord with the prosodic probabilities in English speech. The laboratory mis-segmentations and the spontaneous mis-segmentations produced exactly the same pattern. Although there is as yet no direct demonstration that listeners use prosody to determine word class on line, the juncture misperception findings surely offer quite strong indirect evidence.

WORD CLASS DECISIONS IN SENTENCE PROCESSING

The final question must be: If listeners *do* use prosodic cues to open- vs. closed-class status on line, why? To what use is the knowledge put in the course of processing a spoken sentence?

Issues of both storage and processing are raised by this question. First, are there separate stores for open- vs. closed-class words? Separate stores were proposed in the model put forward by Cutler and Carter (1987) and Cutler and Butterfield (1992), and have also been espoused by other researchers (e.g., Bradley, 1978; Friederici, 1985; Garrett, 1978). On-line categorization of words into open versus closed classes is consistent with separate storage, but it does not force such a conclusion. [Matthei and Kean (1989) discuss this issue at some length, and opt for a postaccess explanation of their finding that closed-class words embedded in nonwords do not slow nonword rejection, while embedded open-class words do, although Bradley (1978) had taken the same result in the visual domain as evidence of separate stores.] On the other hand, separate storage may be warranted by differences in the type of processing to which open- vs. closed-class words must be subjected.

Second, then, in what way could there be such processing differences? One hypothesis, based on Cutler and Butterfield's (1992) approach, could be that the lexical access process is more complex for open-class than for closed-class words. Open-class words have context-independent lexical meanings in a way closed-class words do not; their meanings are also more complex (and in practice may relate in complex ways to sentence context, so that the contextually appropriate meaning may have to be selected from many). The most extreme case of this is homophony, and evidence from the processing of spoken homophones (Swinney, 1979; Tanenhaus, Leiman, & Seidenberg, 1979) suggests that sentence context does not necessarily determine in advance which meaning will be appropriate; instead, all meanings are momentarily accessed and must be checked for appropriateness against the context. The sooner this potentially time-consuming processing can begin, the better. Closed-class words, on the other hand, have mainly syntactic information associated with their stored entries, and relatively little contextually conditioned variation in this. Thus reason enough to distinguish closed- from open-class words on line could be to locate the onset of open-class words and hence begin their access in the lexicon as soon as possible. Location of open-class word onsets was the explanation which Cutler and Norris (1988) and Cutler and Butterfield (1992) offered for their segmentation

findings. This is essentially a Semantic Hypothesis, which focusses on the amount of semantic work caused by open- and by closed-class words, respectively, in spoken-sentence processing.

A conceptually similar but in practice rather different proposal was made by Garrett (1978). Garrett's suggestion, a Syntactic Hypothesis, also proposed that processing of open-class words is more complex than of closed-class — but that it is *syntactically* more complex. Ambiguity of category membership within the open-class is, he pointed out, more the rule than the exception (especially among the more frequently used words). Again, evidence from studies of spoken homophones suggests that syntactic context cannot necessarily determine in advance which category to select from the lexicon; alternative categories are briefly activated (Lucas, 1987; Onifer & Swinney, 1981) and the appropriate one selected on the basis of matching to the context. So far, the Semantic and the Syntactic Hypothesis seem to agree that the purpose of distinguishing open- from closed-class words might be to locate the harder-to-process open-class items as soon as possible. However, Garrett went on to point out that the context which is vitally necessary for determining which of the possible syntactic forms of a given open-class word is appropriate is in fact provided by the adjacent closed-class words. Therefore, he argued, the main point of distinguishing open- from closed-class words in sentence processing would be to identify the closed-class words quickly and reap the benefit of the essential parsing information they provide.

It is interesting to ask at this point whether there is any alternative source of (prelexical) information which might indicate whether a particular open-class word is a noun, verb, or adjective. In fact there are certain phonological correlates of grammatical category within the open class in English: Nouns tend to be longer than verbs, and more likely to bear initial stress; nouns are more likely to contain back vowels, and verbs front vowels. In a series of studies Kelly (Cassidy & Kelly, 1991; Kelly, 1988, 1992; Kelly & Bock, 1988) has shown that listeners are more likely to treat nonwords which are longer, or bear initial stress, as nouns, while shorter nonwords and nonwords with unstressed initial syllables are more likely to be treated as verbs. Thus listeners are able to use their linguistic experience of phonological regularities within the open class to make word class assignments based on distributional probabilities within the vocabulary. Would such information be useful on line?

No direct evidence from sentence processing bears on this question; but some studies have compared listeners' reaction time to categorize spoken words as "noun" or "verb." In one such experiment, Sereno and Jongman (1990) found that categorization responses were faster when

the words contained back or front vowels as characteristic of the two classes. Similarly, Kelly and Martin (in press) found that listeners responded faster when the words conformed to the prosodic patterns characteristic of their class. On the other hand, Cutler and Clifton (1984) found that the same prosodic manipulation had no effect on how quickly listeners could judge the grammatical acceptability of a two-word phrase (*to apple, the borrow, to await*, etc.). It might be argued that Cutler and Clifton's acceptability judgment task is closer to the sentence processing situation than explicit category judgment is, suggesting that listeners may not use prosodic cues to category membership within the open class on line. Additional support for this conclusion comes from a *post hoc* analysis of the boundary insertion errors in Cutler and Butterfield's (1992) corpus. In this analysis only insertion errors which produced open-class words were considered; the question at issue was whether the open-class words produced when the boundary was inserted before a strong syllable (i.e., strong-initial open-class words) were more likely to be nouns, while open-class words produced when the boundary was inserted before a weak syllable (i.e., weak-initial open-class words) were more likely to be verbs (or adjectives, which resemble verbs prosodically). The analysis revealed that among the strong-initial open-class words produced by subjects' erroneous boundary insertions 64% were nouns and 36% were verbs and adjectives, while of the weak-initial open-class words 59% were nouns and 41% verbs and adjectives. In neither case was the ratio significantly different from the categorial distribution of open-class words in the English vocabulary (roughly: 63% nouns, 37% verbs and adjectives). Although the weak-initial set is so small (22 words) that the comparison is in fact of uncertain worth, it surely does not reveal strong effects of prosodic structure on noun/verb assignments in this natural perceptual situation, in agreement with the conclusion of Cutler and Clifton that noun/verb prosodic differences are not exploited on line.

In fact, it is doubtful whether the phonological correlates of noun/verb status *could* be usefullly exploited on line. One of the distributional probabilities is vowel fronting: Verbs are more likely to have front vowels than back. But this probability could only be useful on line if the reverse implicature held, i.e., if a front vowel was more likely to occur in a verb. Because of the predominance of nouns in the vocabulary, this is not the case. Another noun/verb asymmetry involves length: Nouns tend to be longer than verbs. But word length is not a candidate for on-line exploitation for the same reason that purely prosodic cues to stress are not: Listeners would have to wait till the word had ended to know its length. The third main difference is stress pattern, and, as was argued

above, only the vowel-quality reflection of this can be effectively exploited. But in speech contexts, the noun/verb effects on the distribution for strong and weak syllables are overwhelmingly swamped by the closed-class effects. Cutler and Carter's (1987) corpus statistics show that a weak syllable is *fourteen times* more likely to be a closed-class word than the beginning of an open-class word. Even if a listener were to be sure that a word belonged to an open class, a weak initial syllable would not suffice to give strong evidence that the word was a verb. Because of the predominance of nouns in the vocabulary, there are in fact almost twice as many weak-initial English nouns as verbs. And because of the great preponderance of strong-initial words in the open class, the majority (73%) of verbs are in fact strong-initial. Only in the case of disyllabic words could vowel quality in the initial syllable be of some use: A weak-initial disyllable is twice as likely to be a verb as a noun. But of course listeners cannot compute information about number of syllables on line, for the familiar reason that this would require a delay in processing until the end of the word.

If direct phonological cues to category membership within the open class are ineffective on line, then the processing proposed by the Syntactic Hypothesis could indeed be best achieved by exploiting cues to open- vs. closed-class status. The difference between the Semantic and the Syntactic Hypotheses therefore boils down to whether the point of distinguishing open- from closed-class words is in the first instance to speed processing of the former or of the latter. Some support for the Semantic Hypothesis comes from the mis-segmentation findings of Cutler and Butterfield (1992); although it was the case that word boundaries inserted before weak syllables tended to produce closed-class words, it will be recalled that the main finding of Cutler and Butterfield's study was that boundaries hardly ever *were* inserted before weak syllables—it was strong syllables which were most often taken to be word-initial. This is very much in line with the claim of the Semantic Hypothesis that the primary reason to distinguish open- from closed-class words is to initiate open-class processing as rapidly as possible, and it would seem counter to the claim of the Syntactic Hypothesis that locating closed-class words is of primary importance. However, as has been pointed out before, the mis-segmentation evidence only sheds indirect light onto issues of on-line processing. The balance of the evidence at present cannot decide the issue; but it is reasonable to conclude that there are at least two defensible explanations for why listeners should find it useful to distinguish open- from closed-class words in on-line sentence processing.

CONCLUSION

What, finally, is the reality of the distinction between the open and the closed class? As the introductory section to this paper outlined, the distinction as so formulated is an operational one, and the question of whether it maps to a processing distinction must be decided by empirical evidence. It should be clear from the discussion in the sections on juncture misperception and word class decisions in sentence processing that in the present author's view, the fact that listeners find it useful to exploit phonology in such a way as to distinguish open- from closed-class words in sentences offers indirect evidence that a processing distinction between two major classes of words exists. These classes may not, however, be exactly coextensive with the operationally defined open and closed classes. In line with what has been termed the Semantic Hypothesis, the main basis for a distinction between two classes might be semantic complexity: words which could be ambiguous, polysemous, or contextually variant in semantic effect vs. words which have minimal variability of this kind. Such a distinction in fact accords quite well with the open/closed dichotomy. The so-called Syntactic Hypothesis might draw the line rather on the grounds of syntactic constraint: words which were highly syntactically predictive of following context vs. words which were minimally predictive. A distinction of this kind might restrict the syntactically predictive class to only a subset of closed-class words. Interestingly, the closed-class words which would be strongest candidates for inclusion in that restricted set., i.e., which are most syntactically predictive, are exactly those which are also most likely to be realised as weak syllables—*the*, *a*, and *to*, for example. Thus this latter distinction might seem to accord quite well with the phonological differences to which listeners attend. Thus when direct empirical evidence is available to decide the issue, it may turn out that the relevant distinction is not best captured by whether or not a class can be extended. For the time being, however, the open/closed distinction remains the most neutral and hence the most useful one.

The evidence described in this paper has shown that listeners are aware of the phonological correlates of a major distinction between word classes in English. It is clear that these correlates are not strictly necessary for efficient processing, because stressed closed-class words suffer no processing disadvantage. Nevertheless, if the correlates are there, listeners can exploit them in making categorial assignments. The balance of the evidence, moreover, is consistent with on-line exploitation of phonological cues to open- vs. closed-class status in sentence processing.

REFERENCES

Altmann, G. T. M., & Carter, D. M. (1989). Lexical stress and lexical discriminability: Stressed syllables are more informative, but why? *Computer Speech and Language, 3*, 265–275.

Bard, E. G., Shillcock, R. C. , & Altmann, G. T. M. (1988). The recognition of words after their acoustic offsets in spontaneous speech: Effects of subsequent context. *Perception and Psychophysics, 44*, 395–408.

Bond, Z. S. (1981). Listening to elliptic speech: pay attention to stressed vowels. *Journal of Phonetics, 9*, 89–96.

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

Bradley, D. C. (1978). *Computational distinctions of vocabulary type.* Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Braine, M. D. S., Brody, R. E., Brooks, P. J., Sudhalter, V., Ross, J. A., Catalano, L., & Fisch, S. M. (1990). Exploring language acquisition in children with a miniature artificial language: Effects of item and pattern frequency, arbitrary subclasses, and correction. *Journal of Memory and Language, 29*, 591–610.

Cassidy, K. W., & Kelly, M. H. (1991). Phonological information for grammatical category assignments. *Journal of Memory and Language, 30*, 348–369.

Cutler, A. (1986). *Forbear* is a homophone: lexical prosody does not constrain lexical access. *Language and Speech, 29*, 201–220.

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

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language, 2*, 133–142.

Cutler, A., & Clifton, C. E. (1984). The use of prosodic information in word recognition. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 183–196). Hillsdale, NJ: Erlbaum.

Cutler, A., & Fear, B. (1991). Categoricality in acceptability judgments for strong versus weak vowels. *Proceedings of the ESCA Workshop on Phonetics and Phonology of Speaking Styles, Barcelona*, 18.1–18.5.

Cutler, A., & Foss, D. J. (1977). On the role of sentence stress in sentence processing. *Language and Speech, 20*, 1–10.

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

Cutler, A., & Swinney, D. A. (1987). Prosody and the development of comprehension. *Journal of Child Language, 14*, 145–167.

Friederici, A. D. (1985). Levels of processing and vocabulary types: Evidence from online comprehension in normals and agrammatics. *Cognition, 19*, 133–166.

Garrett, M. F. (1978). Word and sentence perception. In R. Held, H. J. Leibowitz, & H. L. Teuber (Eds.), *Handbook of sensory physiology, Volume VIII: Perception* (pp. 611–625). Berlin: Springer.

Green, T. R. G. (1979). The necessity of syntax markers: Two experiments with artificial languages. *Journal of Verbal Learning and Verbal Behavior, 18*, 481–496.

Grosjean, F. (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception and Psychophysics, 38*, 299–310.

Grosjean, F., & Gee, J. (1987). Prosodic structure and spoken word recognition. *Cognition, 25*, 135–155.

Kean, M.-L. (1978). The linguistic interpretation of aphasic syndromes: Agrammatism in Broca's aphasia, an example. *Cognition, 5*, 9–46.

Kelly, M. H. (1988). Phonological biases in grammatical category shifts. *Journal of Memory and Language, 27*, 343–358.

Kelly, M. H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review, 99*, 349–364.

Kelly, M. H., & Bock, J. K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 389–403.

Kelly, M. H., & Martin, S. (1993). Domain general abilities applied to domain specific tasks: Sensitivity to probabilities in perception, cognition and language. *Lingua* (in press).

Lucas, M. M. (1987). Frequency effects on the processing of ambiguous words in sentence contexts. *Language and Speech, 30*, 25–46.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10*, 29–63.

Matthei, E. H., & Kean, M.-L. (1989). Postaccess processes in the open vs. closed class distinction. *Brain and Language, 36,* 163–180.

Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology, 19*, 489–550.

Morgan, J. L., & Newport, E. L. (1981). The role of constituent structure in the induction of an artificial language. *Journal of Verbal Learning and Verbal Behavior, 20*, 67–85.

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency-of-meaning and contextual bias. *Journal of Verbal Learning and Verbal Behavior, 17*, 225–236.

Salasoo, A., & Pisoni, D. B. (1981). *Perception of open and closed class words in fluent speech* (Research on Speech Perception, Progress Report 7, pp. 187–195). Bloomington: Speech Research Laboratory, University of Indiana.

Sereno, J. & Jongman, A. (1990). Phonological and form class relations in the lexicon. *Journal of Psycholinguistic Research, 19*, 387–404.

Stankler, J. (1991). *Phonological distinctions as morphological signals.* Unpublished master of philosophy thesis, Department of Engineering, University of Cambridge, Cambridge, England.

Svartvik, J., & Quirk, R. (1980). *A Corpus of English Conversation,* Lund: Gleerup.

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. A., Zurif, E. B., & Cutler, A. (1980). Effects of sentential stress and word class upon comprehension in Broca's aphasics. *Brain and Language, 10*, 132–144.

Tanenhaus, M. K., Leiman, J. M., & Seidenberg, M. S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior, 18*, 427–440.

Valian, V. V., & Coulson, S. (1988). Anchor points in language learning: The role of marker frequency. *Journal of Memory and Language, 27*, 71–86.

Waibel, A. (1988). *Prosody and speech recognition.* London: Pitman.

Zurif, E. B., & Blumstein, S. (1978). Language and the brain. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 229–245). Cambridge, MA: MIT Press.