THE PROSODIC STRUCTURE OF INITIAL SYLLABLES IN ENGLISH

Anne Cutler* and David Carter+

ABSTRACT

Studies of human continuous-speech recognition suggest that listeners use a strategy of postulating a word boundary, and initiating a lexical access procedure, at each metrically strong syllable. The likely success of this strategy was here estimated against the characteristics of the English vocabulary. Computerised dictionaries of English were found to list approximately three limes as many words beginning with strong syllables (i.e. syllables containing a full vowel) as beginning with weak syllables (i.e. syllables containing a reduced vowel). Furthermore, the mean frequency of occurrence of words beginning with strong syllables is nearly twice as great as that of words beginning with weak syllables. These findings motivated an estimate for everyday speech recognition that approximately 85% of lexical words (i.e. excluding function words) will begin with strong syllables. In fact, in a large corpus of spontaneous conversation 90% of lexical words were found to begin with strong syllables.

INTRODUCTION

Word recognition in continuous speech is complicated by the absence of reliable word boundary correlates. Human listeners nevertheless recognise words in running speech at least as efficiently as they recognise words in isolation, if not more efficiently (ref 1). Recent studies of human speech processing have suggested that listeners may use heuristic strategies for overcoming the absence of word boundary information. Such strategies may allow listeners to guide their lexical access attempts by postulating word onsets at what linguistic experience suggests are the most likely locations for word onsets to occur.

Cutler and Norris (ref 2) have proposed such a strategy based on prosodic structure. In a stress language like English, syllables can be either strong or weak; strong syllables contain full vowels, while weak syllables contain reduced vowels (usually schwa). Culler and Norris found thai listeners were slower to detect the embedded real word in *mintayf* (in which the second vowel is strong) than in *mintef* (in which the second vowel is schwa). They suggested that listeners were segmenting *mintayf* prior to the second syllable, so that detection of *mint* therefore required combining speech material from parts of the signal which had been segmented from one another. No such difficulty would arise for the detection of *mint* in *mintef,* since the weak second syllabic would not be segmented from the preceding material. Culler and Norris proposed that, in English, listeners use strong syllables as the basis for a segmentation strategy in continuous speech processing. Strong syllables are taken to be likely word onseis, and the continuous speech stream is segmented at strong syllables so that lexical access attempts can be initiated.

* MRC Applied Psychology Unit, 15 Chaucer Rd., Cambridge CB2 2EK

+ Computer Laboratory, University of Cambridge, Corn Exchange St., Cambridge CB2 3QG.

The success rate of such a strategy, however, depends at least in part on how realistically it reflects the characteristics of the vocabulary. Hypothesising that strong syllables may be word onsets is unlikely to be a very efficient strategy for detecting actual word onsets if most actual words do not begin with strong syllables. The present study estimates the likely success rale of the strategy proposed by Culler and Norris against the characteristics of the English vocabulary, and then tests it on an actual corpus of English conversation.

WORD-INITIAL SYLLABLES IN ENGLISH

The MRC Psycholinguistic Database (ref 3) is a lexicon of over 98000 words, based on the Shorter Oxford Dictionary. Over 33000 entries have phonetic transcriptions. Fig. 1 shows the prosodic characteristics of the initial syllables of the transcribed words, divided into four categories; monosyllables (such as *bone* or *splint),* polysyllables with primary stress on the first syllable (such as *lettuce* or *splendour),* polysyllables with secondary stress on ihe first syllable (such as *trombone* or *polysyllabicity),* and polysyllables with weak initial syllables(in which the vowel in the first syllable is usually schwa, as in *annoy ox trapeze* but may also be a reduced form of another vowel, as in *investor external).* Any of the first three categories would satisfy the segmentation strategy proposed by Cutler and Norris. It can be seen that these categories together account for 73% of the words analysed.

Since the proposed strategy is aimed at the efficient initiation of lexical access, however, it is reasonable to exclude from our analysis those words whose interpretation in a speech context relies not upon a lexical lookup but upon strictly contextual factors; that is, it is reasonable to exclude grammatical words (such as articles, conjunctions and pronouns). The distribution of the prosodic characteristics of the initial syllables of lexical words (nouns, verbs, adjectives and most adverbs) in the MRC Database is, however, virtually identical to Fig. 1, since exclusion of grammatical words reduced the total corpus size by less than 1%.

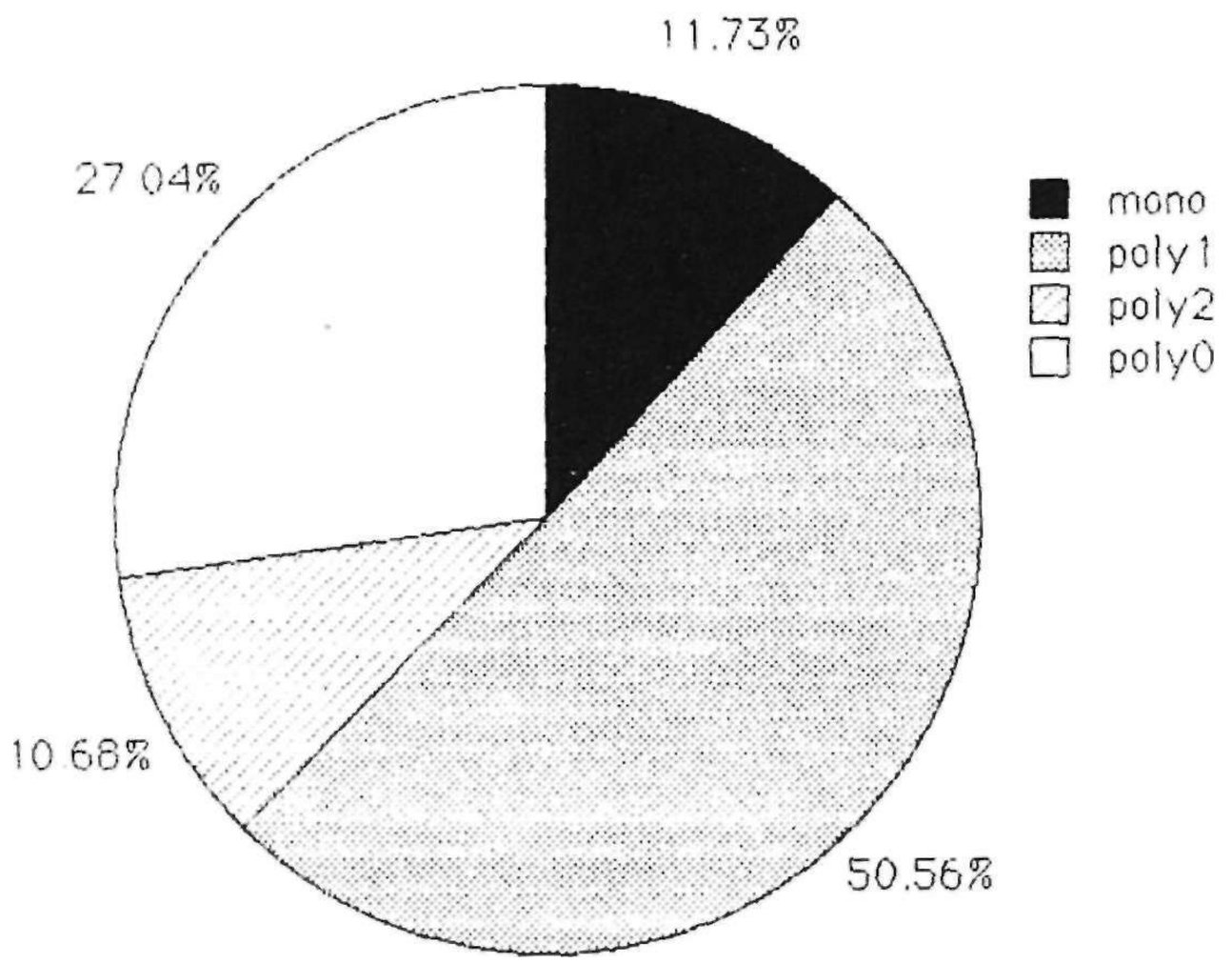

Fig. 1. Prosodic categories as proportions of the MRC Database.

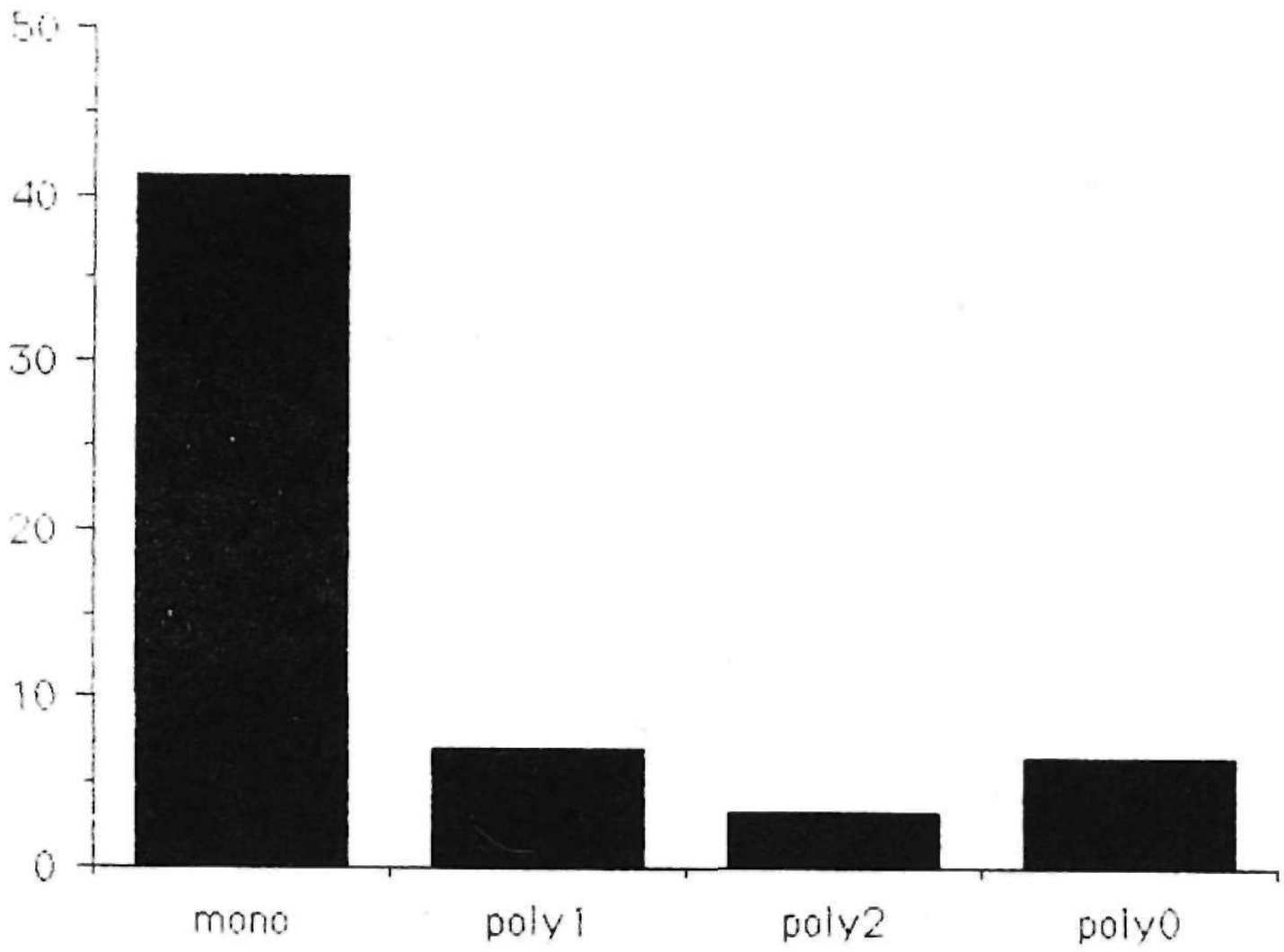

Fig. 2. Mean frequency of occurrence for lexical items by prosodic category.

WORD PROSODY AND FREQUENCY OF OCCURRENCE

The most common word *type* in English is clearly a polysyllable with initial stress. However, individual word types differ in the frequency with which they occur. Frequency of occurrence statistics (ref 4) are listed in the MRC Database. Fig. 2 shows the mean frequency for the four prosodic word-categories (lexical words only). It can be seen that monosyllables occur on average far more frequently than other prosodic types. Thus although there are more than seven times as many polysyllables in the language as there are monosyllables, average speech contexts are likely to contain almost as many monosyllables as polysyllables. Fig. 3 shows an estimate of the likely distribution of prosodic categories in a real speech context, derived from a combination of the data in Figs. 1 and 2; this suggests that only 17% of lexical tokens will begin with weak syllables.

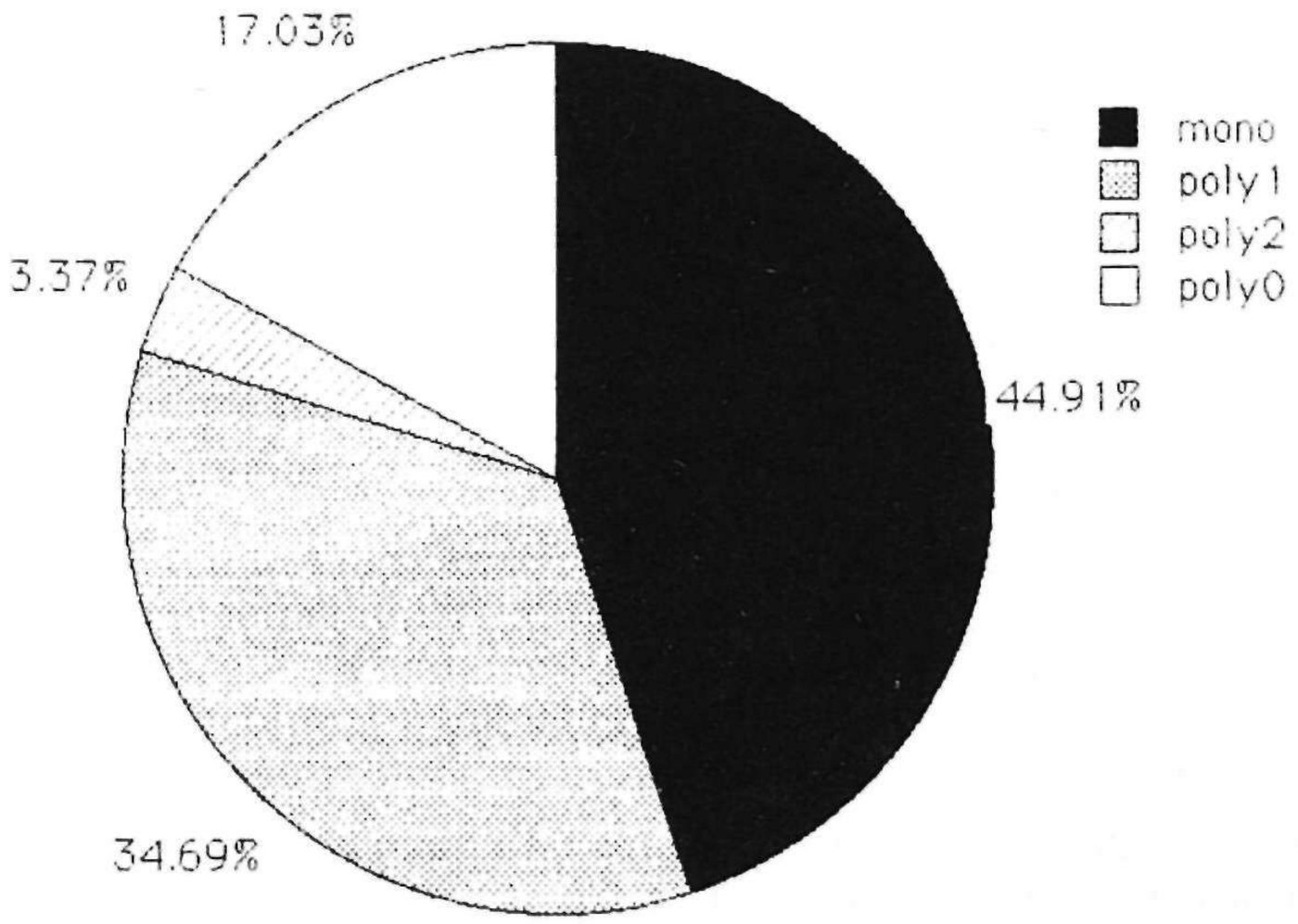

Fig. 3. Predicted distribution of prosodic categories in real speech.

WORD PROSODY IN A NATURAL SPEECH SAMPLE

We tested the estimate shown in Fig. 3 against a natural speech sample, the *London-Lund Corpus of English Conversation* (ref 5), using the frequency count of this corpus prepared by Brown (ref 6). The London-Lund corpus consists of approximately 190,000 words of spontaneous British English conversation. Fig. 4 shows the distribution of prosodic categories for lexical words in this corpus. The three categories with strong initial syllables account for 90% of the tokens-, only 10% of the lexical words have weak initial syllables.

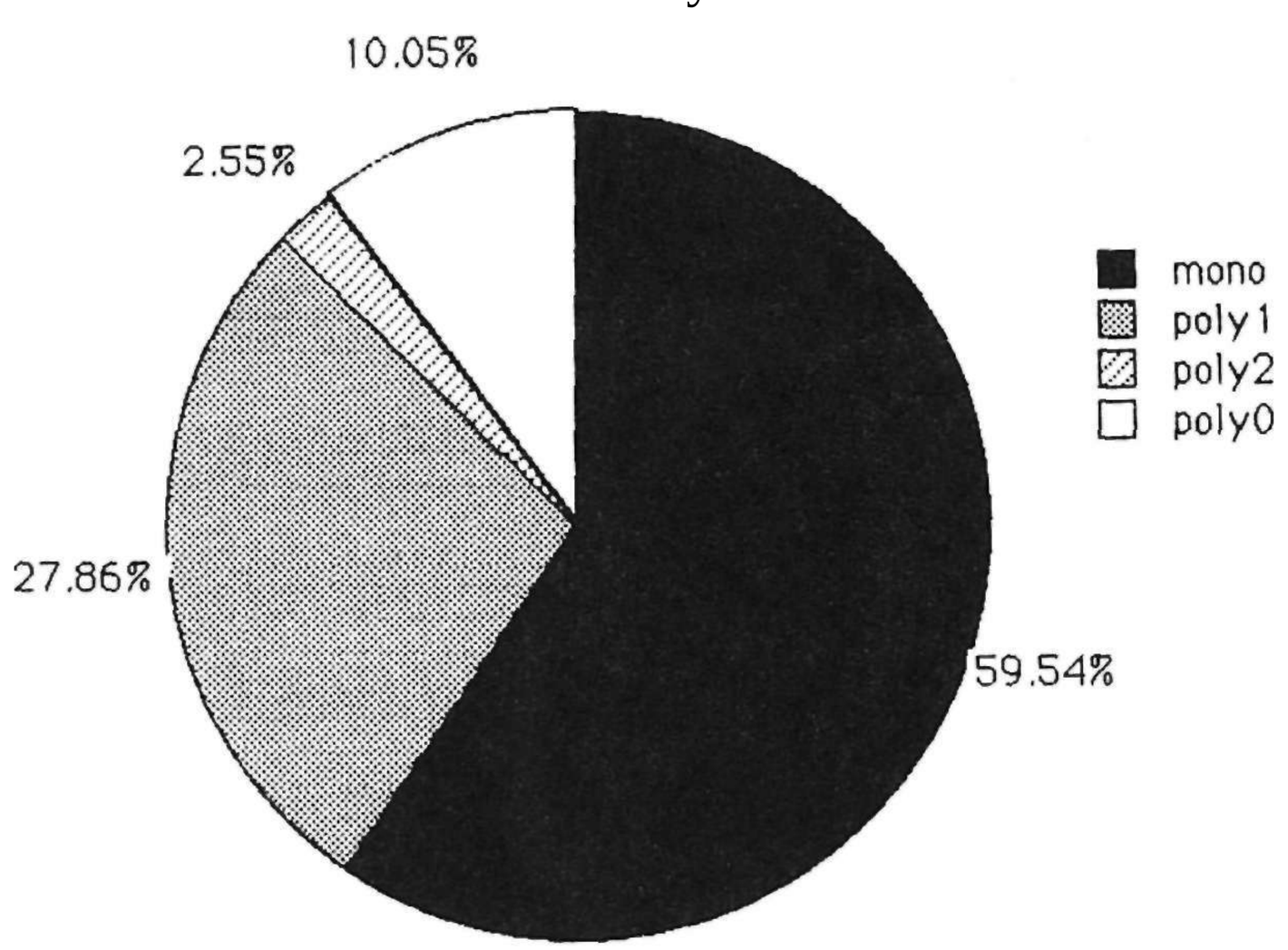

Fig. 4. Distribution of prosodic categories in the *Corpus of English Conversation.*

CONCLUSION

The distribution of word types in the English vocabulary, combined with relative frequency of occurrence across types, provides an adequate basis for the implementation of a segmentation strategy in continuous speech recognition whereby strong syllables are assumed to be the onsets of lexical words.

ACKNOWLEDGEMENTS

This research was supported by a grant from the Alvey Directorate (MMI-069) to Cambridge University, the Medical Research Council and Standard Telecommunications Laboratories. We thank Gordon Brown for making available the machine-readable version of his frequency count of the London-Lund corpus.

REFERENCES

1. EX. Schwab, H.C. Nusbaum & D.B. Pisoni, Hum. Factors, *27,* 395 (1985).
2. A. Cutler & D.G. Norris, J. Exp. Psychol. Hum. Perc. Perl, (in press).
3. M. Coltheart, Quart. J. Exp. Psychol. *33A,* 497 (198 1).
4. H. Kucera & W.N. Francis, Computational Analysis of Present-Day American English (Brown Univ. Press, Providence, 1967).
5. J. Svartvik & R. Quirk, A Corpus of English Conversation (Gleerup, Lund, 1980).
6. G.D.A. Brown, Beh. Res. Meth. Instr. &. Comp. *16,* 502 (1984).