Why not abolish Psycholinguistics?*

Anne Cutler

1 The problem with Psycholinguistics

I hereby propose the abolition of 'Psycholinguistics'. Please note the quotes - I am not advocating the abolition of my own job, nor indeed the job of anyone else, anywhere, who claims to be a psycholinguist. What I would like to get rid of is the *word* 'Psycholinguistics', and the reason is that it daily misleads and confuses.

The confusion arises because 'Psycholinguistics' is morphologically ambiguous. Consider other disciplines with similarly constructed names. On the one hand we have *astrophysics*, which is a branch of astronomy, *psychophysics* which, similarly, is a branch of psychology, as also is *psychophysiology*. On this model Psycholinguistics is clearly also a branch of psychology. But on the other hand *photophysics* is a branch of physics and *photochemistry* a branch of chemistry. And if *psychobabble* is anything, it's a type of babble. So these models lead us to conclude that Psycholinguistics is a variety of linguistics.

The ambiguity can be exemplified by imagining a new discipline. What could *pornolinguistics* be? Clearly, it could be either a branch of linguistics dealing with the language of pornography, or a type of pornography offering smutty texts in the guise of linguistic treatises - both interpretations are readily conceivable.

But does it matter that the term *Psycholinguistics* is potentially ambiguous? Surely psycholinguists know what they are, just as psychophysiologists know that they are more psychologists than physiologists. My argument here is that the ambiguity does indeed matter, because it is not just a potential ambiguity but an actual one: there is Psycholinguistics which is a variety of linguistics, and there is Psycholinguistics which is a branch of psychology. *This duality is rarely recognized*, which is why the existence

of a single term *Psycholinguistics* has caused and continues to cause confusion.

Consider the kind of Psycholinguistics I (and other psychologists) do. Our aim is to understand and explain the mental structures and processes which are involved in the use of language (speaking, listening, reading, writing, and also acquiring language or losing it). Since trying to understand and explain mental structures and processes is usually called *cognitive psychology*, it would be perfectly accurate to replace the name *Psycholinguistics* in our case by *language psychology*. But that would not be appropriate for other psycholinguists, whose studies of speaking, listening, reading, writing, and language acquisition or loss aim to understand and explain the structure of language itself. Since trying to understand and explain the structure of language is usually called *linguistics*, using evidence from language performance in this quest could be termed *performance linguistics* or possibly *laboratory linguistics* (in phonology, at least, this latter term is now preferred; see Kingston and Beckman 1990). The confusion that I mentioned above arises when evidence obtained in the service of one of these aims is mistaken for evidence relevant to the other aim. Psycholinguistics is an experimental discipline, and the key to conducting a successful experiment is to exercise strict control over variables which are irrelevant to the question at issue. This means that the answer one obtains from an experimental question is truly relevant to that question, and is not contaminated by irrelevant issues. But it also means that experimental results should not be generalized beyond the experimental question that motivated them. Thus a finding which illuminates a psychological question of processing may be quite irrelevant to questions about the structure of language, while a finding which illuminates a question of linguistic structure may equally be of no relevance to processing issues.

In the following sections I will describe some recent work from my own laboratory which, appropriately for the present volume, is centred upon concepts from phonology. These concepts seem to me *entirely* linguistic in nature (by which I mean that they are abstract, and not directly translatable into measurable physical properties of the speech signal, which is of course not to claim that linguistic concepts never correspond to physical entities). Nevertheless the work addresses questions of language processing, and hence falls into the category of language psychology. To be consistent with my claims above, I will suggest in the conclusion that there is an asymmetry in my relationship with phonology: phonology is central to my current work, but my current work is of relatively little importance to phonology.

2 Metrical prosody and speech segmentation

The work that I will describe addresses the question of how we determine where one word ends and the next begins when we are listening to natural continuous speech. Alas for the listener's convenience, word boundaries in continuous speech are not reliably marked by any acoustic cue or signal. Cutler and Norris (1988) have proposed that listeners may use heuristic strategies for overcoming the absence of word boundary information, specifically that they guide their lexical access attempts by postulating word onsets where linguistic experience suggests that word onsets are most likely to occur. The proposal accounts for the results of an experiment in which Cutler and Norris found that listeners were slower to detect the embedded real word in *mintayf* (in which the second vowel is [e]) than in *mintef* (in which the second vowel is schwa). That is, these two bisyllables differ in their metrical structure: one has two strong syllables, the other a strong and a weak syllable. Cutler and Norris suggested that in a stress language like English, metrical structure could provide the basis of a segmentation heuristic, whereby strong syllables could be treated as likely to be the initial syllables of new (lexical) words. In effect, listeners would employ a strategy of segmenting speech signals at the onset of each strong syllable. In the experiment, therefore, *mint* would be relatively difficult to detect in *mintayf* because listeners were segmenting *mintayf* prior to the second syllable, so that detection of *mint* in this case required combining speech material from parts of the signal which had been separated from one another by segmentation. No such difficulty would arise for the detection of *mint* in *mintef,* since the weak second syllable would not be segmented from the preceding material. Cutler and Norris' proposal for English (and metrically similar languages) will henceforth be called the Metrical Segmentation Strategy (MSS).

Note at this point that the distinction between strong and weak syllables (and indeed the whole concept of metrical structure) is properly a phonological one. It is *not* possible to provide an acoustic definition such that any syllable can be unambiguously characterized as strong or weak on physical properties alone. In real speech, there is a continuum of vowel 'strength', with long tense vowels and diphthongs at one end, schwa at the other, and in between all possible realizations of vowel quality and duration. (Cutler and Norris suggested that in practice a recognizer could base a decision about vowel strength on an arbitrary durational criterion, rate-normalized individually for a given speech sample. But the details of how the strategy can be implemented constitute a separate question from

whether it is promising enough to be worth implementing.) In fact, it is no accident that a proposed solution to the segmentation problem in speech recognition involves abstract phonological concepts. If it were possible to couch such a solution in terms of easily detectable physical characteristics of the signal, it would have been done years ago. It is precisely *because* no such clear distinctions can be drawn in the speech signal itself that it is necessary to turn to phonology for the conceptual framework within which a potential solution may be sought.

3 Metrical prosody of English speech

The success rate of the Metrical Segmentation Strategy for real-life listening performance, of course, depends on how realistically it reflects the structure of the language. Hypothesizing that strong syllables are likely to be lexical word onsets, but weak syllables are not, will only prove to be an efficient strategy for detecting actual word onsets if most lexical words indeed begin with strong syllables and not with weak syllables. Fortunately, research on the lexical statistics of English has shown that the MSS is indeed well adapted to the characteristics of the vocabulary. Cutler and Carter (1987) examined the metrical structure of word-initial syllables in English. First, they looked at the metrical structure of words in the English vocabulary. The MRC Psycholinguistic Database (Coltheart 1981; Wilson 1988) is a lexicon (based initially on the Shorter Oxford English Dictionary) which contains over 33,000 words in phonetic transcription. In this lexicon, Cutler and Carter found, about 12% of the words were monosyllables (such as *camp* or *lodge),* just over 50% were polysyllables with primary stress on the first syllable (such as *camphor* or *cycle),* a further 11% were polysyllables with secondary stress on the first syllable (such as *campaign* or *psychological),* while the remaining 27% were polysyllables with weak initial syllables (in which the vowel in the first syllable is usually schwa, as in *camellia,* but may also be a reduced form of another vowel, as in *illogical).* Any of the first three categories would satisfy the MSS, and these categories together account for 73% of the words in the list.

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 (Kucera and Francis 1967) are listed in the MRC Database, and Cutler and Carter found that the mean frequency for the four metrical word-categories did indeed

differ. Firstly, monosyllables occur on average far more frequently than any type of polysyllable. (Note that to analyse frequency of occurrence, Cutler and Carter considered only the lexical, or content, words in the database, and excluded the grammatical, or function, words, which accounted for less than 1% of the phonetically transcribed words. These were overwhelmingly monosyllabic and of high frequency; their inclusion would have inflated the mean frequency of monosyllables still further.) Secondly, within the set of polysyllables, words with strong initial syllables occur somewhat more frequently than words with weak initial syllables. If the type counts reported above are multiplied by their mean frequencies, one can estimate that 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. Moreover, only about 17% of lexical tokens in most speech contexts should begin with weak syllables.

Cutler and Carter tested this estimate against a natural speech sample, the *London-Lund Corpus of English Conversation* (Svartvik and Quirk 1980), using the frequency count of this corpus prepared by Brown (1984). The London-Lund corpus consists of approximately 190,000 words of spontaneous British English conversation. The distribution of metrical categories for lexical words in this corpus was indeed markedly different from the simple distribution of metrical types in the vocabulary: almost 60% were monosyllables, only 28% polysyllables with initial primary stress, and less than 3% polysyllables with initial secondary stress. Most noticeably, perhaps (especially when one considers that a relatively high proportion of the speech in this corpus came from learned academic conversation!), less than 10% of the lexical words had weak initial syllables. In other words, the three categories with strong initial syllables accounted, together, for 90% of the tokens. Thus the MSS would correctly locate the onsets of 90% of the lexical words in this sample.

However, this lexical word count disguises one important fact: lexical words actually comprised only 41% of all words in the London-Lund corpus. The majority of words in the corpus were, in fact, grammatical words. But because hardly any grammatical words had more than one syllable, the lexical word total nevertheless accounts for 51% of all *syllables*. In fact, with some reasonable assumptions it was possible to compute the probable distribution of syllables in this speech sample. Cutler and Carter assumed that grammatical words such as *the* and *of* were in general realized as weak syllables. In that case, the most likely distribution of syllables was as given in Table 1. It can be seen that about three-

Table 1. Distribution of Strong and Weak Syllables in the 'London-Lund Corpus of English Conversation'

	Strong	Weak
Sole or initial syllable of lexical word:	74%	5%
Non-initial syllable of lexical word:	12%	23%
Sole or initial syllable of grammatical word:	11%	69%
Non-initial syllable of grammatical word:	3%	3%
Total number of syllables	93,989	145,888
	(39%)	(61%)

quarters of all strong syllables in the sample were the sole or initial syllables of lexical words. Of weak syllables, however, more than two-thirds were the sole or initial syllables of grammatical words.

Thus a listener encountering a strong syllable in spontaneous English conversation would seem to have about a three to one chance of finding that strong syllable to be the onset of a new lexical word. A weak syllable, on the other hand, would be most likely to be a grammatical word. It would appear, therefore, that English speech indeed provides an adequate basis for the implementation of a segmentation strategy such as Cutler and Norris' MSS, whereby strong syllables are assumed to be the onsets of lexical words.

4 Missegmentations and metrical prosody

The next step in this attack on the segmentation problem was to look for evidence that listeners may indeed use the MSS in processing natural continuous speech. Cutler and Norris' (1988) experiment, which motivated the MSS proposal, presented listeners only with nonsense bisyllables. Butterfield and Cutler (1988) extended the testing of the MSS to listeners' perception of continuous speech. They focused on the listeners' segmentation *errors*, i.e. the way in which word boundaries tend to be misperceived.

The absence of reliable cues to word boundary location makes it seem that misperception of word boundary location in speech should be fairly probable. Butterfield and Cutler reasoned that if listeners are indeed assuming strong syllables to be word-initial and weak syllables to be not word-initial, then it should be the case that word boundary misperceptions will be very unequally distributed across the four possible types of error.

Specifically, erroneous insertion of a boundary before a strong syllable and erroneous deletion of a boundary before a weak syllable should prove to be relatively common, whereas erroneous insertion of a boundary before a weak syllable and erroneous deletion of a boundary before a strong syllable should be relatively rare.

Psycholinguists have for many years collected and analysed the slips of the ear that occur in conversation, and many such slips contain word boundary misplacements. Butterfield and Cutler examined all the errors listed in published studies of slips of the ear (such as Browman 1978; Garnes and Bond 1975), plus all the slips of the ear included in a speech error collection assembled over several years by the present author. Among these slips over one hundred involved misplacement of a word boundary across at least one syllabic nucleus. Some example errors are shown in Table 2. Butterfield and Cutler found that in this set of naturally occurring errors precisely the pattern predicted by the MSS appeared: insertions of a word boundary before a strong syllable (e.g. 'phonological' -» 'formal logical') and deletions of a word boundary before a weak syllable (e.g. 'the chance that was wasted' -» 'the Chancellor's waistcoat') outnumbered by more than two to one insertions of a boundary before a weak syllable (e.g. 'The Biggin sector' -» 'the big inserter') or deletions of a boundary before a strong syllable (e.g. 'once a fortnight' -» 'one support night').

However, Butterfield and Cutler found that the contextual information available for these errors was insufficient to determine what opportunities the listeners had for word boundary misplacement. Thus the statistical significance of the asymmetric distribution of the natural slips was difficult

Table 2. Slips of the ear

the chance that was wasted	-»	the Chancellor's waistcoat
phonological coding	-»	formal logical coding
the chief economist	-»	the chief communist
'Antenna' asked	-»	a tender-arsed
Coke and a Danish	-»	Coconut Danish
my gorge is rising	-»	my gorgeous ...
by tonight	-»	butter knife
she'll officially	-»	Sheila Fishley
she's a must to avoid	-»	she's a muscular boy
once a fortnight	-»	one support night
is burgling the garage	-»	a bird in the garage
the Biggin sector	-»	the big insecter
three in ten derive	-»	three intend arrive

to determine. Therefore Butterfield and Cutler followed up their analysis of spontaneous misperceptions with an experiment involving deliberately induced misperceptions. In this study, unpredictable utterances (e.g. 'soon police were waiting') were presented to listeners at a level minimally above their threshold for speech reception (which was determined separately for each listener in an extensive pre-test). The subjects' task was to write down what they thought was said.

Overall, about 20% of all the responses contained analysable word boundary misplacements. Some responses contained more than one boundary misplacement, so the total number of errors available for analysis was in the end 257. The distribution of these errors across the four possible error classes is shown in Table 3. It can be seen that exactly the pattern predicted by the proposed strategy emerged: erroneous insertions of a word boundary before a strong syllable and deletions of a word boundary before a weak syllable greatly outnumbered insertions of a boundary before a weak syllable or deletions of a boundary before a strong syllable.

Because the opportunity for each type of error could in this case be determined exactly, the difference could be evaluated statistically, and Butterfield and Cutler found that it was indeed significant. A comparison of the frequency of lexical words in the target and in the error showed no overall preference for higher frequency responses, and no significant difference in frequency effects across those responses which were predicted by the strategy and those which were not. This suggests that listeners were not simply producing plausible responses, since a preference for plausible utterances would lead to relatively high-frequency responses.

One further characteristic of the error pattern in this experiment is worthy of note. Although word boundary insertions before weak syllables, which are predicted to be relatively uncommon, are indeed the second rarest type of error, they nevertheless occur four times as often as the rarest type of error, boundary deletions before strong syllables. From

Table 3. Response frequencies in the faint speech experiment

1.	Insertion before a strong syllable (sons expect enlistment -» some expect a blizzard)	144	>	196
2.	Deletion before a weak syllable (achieve her ways instead -» a cheaper way to stay)	52		
3.	Deletion before a strong syllable (soon police were waiting -* soon to be awakened)	13	>	61
4.	Insertion before a weak syllable (dusty senseless drilling -» thus he sent his drill in)	48		

Cutler and Carter's examination of natural speech, one may formulate a prediction about the prosodic probabilities of weak syllables and hence the way they are most likely to be misperceived. In the spontaneous speech corpus which Cutler and Carter examined, more than two-thirds of all weak syllables were monosyllabic grammatical words. Thus one might predict that a weak syllable in faintly perceived speech is most likely to be perceived as a monosyllabic function word. A subsidiary prediction about the misperception data might be, then, that erroneous insertions of word boundaries before weak syllables should tend to involve erroneous report of monosyllabic function words.

This is indeed the case. Exactly two-thirds of the boundary insertions before weak syllables (32 out of 48 cases) involved monosyllabic function words (such as 'dusty senseless drilling' -» 'thus he sent his drill in'). Examination of the natural slips of the ear showed that a large number of the erroneous insertions of word boundaries before weak syllables in that corpus also involved monosyllabic function words (e.g. 'is burgling the garage' -» 'a bird in the garage'). Word boundary misplacements by human listeners therefore seem to reflect the prosodic probabilities of English remarkably accurately. The initial statement of the MSS, which referred only to lexical word boundaries, may be an under-estimate of the degree to which the segmentation of continuous speech is driven by prosodic probability.

5 Conclusion

The work described in the preceding sections tells us (I hope) something about how language is processed, namely that listeners can exploit their knowledge of the prosodic structure of their language to help them overcome the segmentation problem which arises from the absence of reliable, consistent word boundary cues in continuous speech. In English, the relationship between prosodic structure and word onsets can be expressed in terms of concepts from phonology: the metrical prosodic dichotomy of strong versus weak syllables. But although this phonological distinction is central to the claims I have made, my experiments were designed to address issues of processing alone, and the only conclusions that can properly be drawn concern the mental structures and processes involved in the act of speech recognition. The work belongs firmly in the realm of language psychology.

Elsewhere (Cutler 1987) I have described language-psychological research on the role of the syllable and the role of lexical stress in speech recognition. Both syllables and lexical stress are phonological constructs. As it happened, the experiments suggested that neither construct plays a role in the prelexical stages of English speech recognition. Nevertheless it would be a mistake to draw conclusions for phonology from these results, tempting as such a course might be for a phonologist concerned about the 'psychological reality' of either construct (note, incidentally, that 'psychological reality' is a *linguistic* concept, not a psychological one; psychologists' concepts of mental structures and processes are far more specific than this). It would be a mistake on the one hand because plenty of evidence exists that the two constructs play a role in other aspects of language performance, and on the other hand because questions of 'psychological reality' can only properly be answered by experiments designed directly to test 'psychological reality'.

Analogously, I believe that phonology would be unwise to seek lessons from my arguments for the importance of metrical prosody in speech segmentation. And perhaps if I succeed in abolishing the term *Psycholinguistics* the temptation will no longer exist. Language psychologists will continue to accumulate evidence about language processing, while evidence of relevance to the proper account of phonological structure will continue to be produced by the many thriving laboratories of performance linguistics. But without an ambiguous common name, the two will no longer be confused.

Notes

* The work reported in this paper was supported by a grant (MMI-069) from the Alvey Directorate, U.K. to Cambridge University, the Medical Research Council, and STL Technology Ltd.

1 Exactly the same ambiguity applies, of course, to similar terms, such as *neurolinguistics* and *sociolinguistics*. In the latter case the ambiguity has been remarked upon by Trudgill (1978).

References

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

Brown, G.D.A. 1984. A frequency count of 190,000 words in the London-Lund Corpus of English Conversation. *Behavior Research Methods, Instrumentation and Computers* 16: 502-532.
Butterfield, S. and Cutler, A. 1988. Segmentation errors by human listeners: Evidence for a prosodic segmentation strategy. *Proceedings of the Federation of Acoustic Societies of Europe, Edinburgh:* 827-833.
Coltheart, M. 1981. The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology* 33A: 497-505.
Cutler, A 1987. Phonological structure in speech recognition. *Phonology Yearbook* 3: 161-178.
Cutler, A and Carter, D.M. 1987. The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language 2:* 133-142.
Cutler, A and Norris, D. 1988. The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance* 14: 113-121.
Games, S. and Bond, Z.S. 1975. Slips of the ear: Errors in perception of casual speech. *Proceedings of the Eleventh Regional Meeting Chicago Linguistic Society:* 214-225.
Kingston, J. and Beckman, M. (eds.). 1990. *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech.* Cambridge: Cambridge University Press.
Kucera, H. and Francis, W.N. 1967. *Computational Analysis of Present-Day American English.* Providence: Brown University Press.
Svartvik, J. and Quirk, R. 1980. *A Corpus of English Conversation.* Lund: Gleerup.
Trudgill, P. 1978. Sociolinguistics and sociolinguists. In P. Trudgill (ed.), *Sociolinguistic Patterns in British English.* London: Arnold.
Wilson, M.D. 1988. MRC Psycholinguistic Database: Machine-usable dictionary, Version 2.0. *Behavior Research Methods, Instrumentation and Computers* 20: 6-10.