COMPETITION AND SEGMENTATION IN SPOKEN WORD RECOGNITION

Dennis Norris*, James McQueen and Anne Cutler**.**

* Medical Research Council, Applied Psychology Unit, 15 Chaucer Road, Cambridge, England CB2 2EF.
** Max-Plank-Institute for Psycholinguistics, Nijmegen, The Netherlands

ABSTRACT

This paper describes recent experimental evidence which shows that models of spoken word recognition must incorporate both inhibition between competing lexical candidates and a sensitivity to metrical cues to lexical segmentation. A new version of the Shortlist [1][2] model incorporating the Metrical Segmentation Strategy [3] provides a detailed simulation of the data.

I. INTRODUCTION

The TRACE model of McClelland and Elman [4] popularised the idea that spoken word recognition involves a process of competition between alternative overlapping lexical candidates. In TRACE, and the Shortlist model of Norris [1][2], competing candidates are connected together by inhibitory links in an interactive activation network and the final parsing of the input into words is determined by the result of the competition between these overlapping candidates. According to this view, the process of segmenting continuous speech into words is an emergent property of the competition process which is not dependent on the availability of any phonetic or phonological cues as to the location of word boundaries.

A contrasting view of the process of lexical segmentation has been presented by Cutler and Norris [3]. They propose that listeners make use of metrical or rhythmic information to determine the likely locations of word onsets. Speakers of English, for example, will capitalise on the fact that the most likely locations for word boundaries are the onsets of strong syllables. According to Cutler and Norris's Metrical Segmentation Strategy (MSS) listeners segment the input and initiate a new lexical access attempt at the onset of strong syllables. The primary source of evidence for the MSS comes from a word-spotting experiment in which subjects listened to a list of bisyllabic nonsense words and were required to press a button as soon as they heard a word embedded in one of the nonsense words. Cutler and Norris found that subjects were slower and less accurate when identifying a CVCC word like 'mint' in a strong-strong CVCCVC nonsense string like /minteif/ than in a strong-weak string like /mintef/. They argued that identification of 'mint' in /minteif/ was harder because it involved recombining information accross a segmentation boundary (before the second strong syllable in /minteif/). According to the MSS there is no segmentation boundary in the strong-weak case.

II. LEXICAL COMPETITION

Although there is strong experimental support for the MSS, until recently there was little evidence for the kind of lexical competition posited by TRACE and Shortlist. However, using a word spotting task, McQueen, Norris and Cutler [5] have shown that words like 'mess' are harder to identify in nonsense strings like /demes/ than in strings like /names/. In strings like /dames/ the target word 'mess' is inhibited by competition from words like 'domestic'. In /names/ there are no such overlapping competitors. Similar results were also found for nonsense words with a strong-weak stress patterns. The presence of competitors like 'sacrifice' made it harder to identify 'sack' in /saekref/ than /saekrek/. These results are important because not only do they provide clear evidence of competition, but they also show that

competing words need only overlap in the input, they do not have to begin at the same point in the input. McQueen et al. also found that words in weak-strong nonsense strings were easier to identify than words in strong-weak nonsense strongs. They argued that this was due to the beneficial effect of the MSS segmenting the weak-strong strings at the onset of the target word. In other words, a complete account of these findings demands a model which incorporates both lexical competition and metrically-based segmentation.

III THE SHORTLIST MODEL

McQueen et al. presented detailed simulations of their data using the Shortlist model [1][2], Like TRACE, Shortlist employs competition between overlapping lexical candidates. However, Shortlist has a more plausible architecture, has, an entirely bottom-up information flow, and, unlike TRACE, can perform simulations using realistically sized vocabularies.

Word recognition in Shortlist is essentially a two- stage process. An initial stage of lexical access generates a set of lexical candidates, a shortlist, which are roughly consistent with the input. In the current version of Shortlist this process is simulated by an exhaustive lexical search procedure using a large machine readable dictionary. Condidates in the shortlist are then entered into an interactive activation network in which all overlapping candidates are connected together by inhibitory links. The strength of the inhibitory links is proportional to the number of phonemes by which the candidates overlap. The interactive activation network is constructed dynamically to contain only those candidates in the shortlist. A crucial feature of Shortlist is that the size of the network is assumed to be limited. In all current simulations no more than 30 candidates may begin at each phoneme in the input. This limit of 30 candidates is more than generous as the model performs adequately with as few as two candidates when presented with unambiguously transcribed input. By employing a restricted candidate set which is constructed entirely on the basis of a data-driven analysis Shortlist avoids the computational

explosion which limits both the plausibility and practicality of TRACE. In TRACE every word in the lexicon acts as a candidate at all points. There must therefore be an entire lexical network aligned at each point where a word might begin, and all overlapping words in these networks must be connected together by inhibitory links. Consequently, TRACE would require more than one billion times more inhibitory connections than Shortlist to perform the simulations described here.

Although Shortlist is an interactive activation model, it differs from most other interactive activation networks in that the flow of information within the model is completely bottom-up. Interaction takes place only within nodes in the candidate set. There is no feedback from these lexical nodes to earlier levels of analysis. Shortlist is therefore an modular bottom-up model in the tradition of the race model of Cutler and Norris [6]. In fact, Shortlist can be thought of as an implementation of the lexical access component of Cutler and Norris's race model.

IV. COMPETITION AND SEGMENTATION

McQueen et al's demonstration that spoken word recognition involves both competition and metrical segmentation raises the question of how these two processes should be combined in a single model. To provide some empirical constraint on the development of an integrated model Norris, McQueen and Cutler [7] collected data from a modified version of the original Cutler and Norris word-spotting experiment. The crucial variation from Cutler and Norris's experiment was that Norris et al. varied the number of competitors beginning at the onset of the second syllable. Whereas items like /staempid3/ and /staemped3/ had many competitors beginning at the /p/, items like /mintaup/ and /mintep/ had few competitors beginning at the /t/.

This design serves two important functions. First, as Cutler and Norris did not control for possible effects of number of competitors, it provides a replication of the metrical segmentation effect with number of competitors in strong- strong and strong-weak conditions equated. Second, it enables

us to determine the relationship between segmentation and competition. Do they interact, or does segmentation have a constant effect independent of the number of competitors ? Will the effect of segmentation be the same when there are many as when there are few competitors ?

The results for the CVCC targets in the experiment are shown in Table 1. In the errors there is a significant interaction between number of competitors and stress pattern, with the effect of metrical segmentation being significant only when there are many competitors. A similar effect also emerges in the reaction times, but here it is significant only for the least accurate half of the subjects. The efects of metrical segmentation are clearly modulated by the number of competitors. However, these data also represent an important extension of the earlier work on competition. McQueen at al. had demonstrated an effect of competition attributable to a single highly active competitor. These new data show that competition effects increase with an increasing number of competitors.

Stress	Many		Few	
	SW	WS	SW	WS
RT	696	607	650	610
Errors	22%	13%	19%	17%

Table 7 Results of word-spotting experiment in strings with many and few competitors

Norris et al. simulated their results in Shortlist by allowing metrical information to modulate the activation levels of lexical candidates. Shortlist makes use of both match and mismatch information in computing the degree of fit between the input and lexical candidates. Not only is the fit increased by matching evidence but it is decreased by mismatching evidence. The MSS also has two components: segmentation at strong onsets and initiation of a lexical access attempt at strong onsets. Clearly these two components of the MSS cannot be implemented literally in a model like Shortlist where lexical access is actually taking place continuously to generate and update the candidate set. However, we can achieve the same goals that Cutler and Norris were aiming at by allowing the MSS to influence computation of the degree of fit between input and candidates. To this end, activation levels were decreased for words that did not have a lexically marked strong onset when there was a strong onset in the input (segmentation) and were increased for words beginning at a strong onset in the input ('initiating a lexical access attempt'). The results of these simulations are shown in Figure 1. Note that the simulations without the MSS actually show a strong-weak - weak-strong difference in the opposite direction from that observed in the data. Norris et al. also show that the modified version of Shortlist simulates the full pattern of data reported by McQueen et al. In other simulations they also show that these data can only be adequately simulated by incorporating both components of the MSS. Either component alone fails to give a complete account of the data from either this experiment or those of McQueen et al.

An important feature of these simulations is that a proper account of the competition effect depends entirely on using a modification to Shortlist suggested by Norris [2] to overcome the problem of the network latching onto a single interpretation of the input and failing to revise it in the light of subsequent input. In its default mode of operation Shortlist, like TRACE, begins processing each new phoneme of input with all nodes in the network having the level of activation achieved after processing the previous phoneme. This means that after processing /mint/, the node corresponding to 'mint' has such a high level of activation that it completely supresses activation of candidates beginning at the final /t/. These candidates are inhibited before they have a chance to get off the ground and consequently have no impact on the target word. In this mode of operation candidates arriving late on in a word therefore have little or no impact on its identification. In general, once the network has reached one interpretation of the input it is difficult to revise that interpretation because activation of competing analyses is supressed. This problem can be remedied by allowing the network to continuously recompute a new and optimised
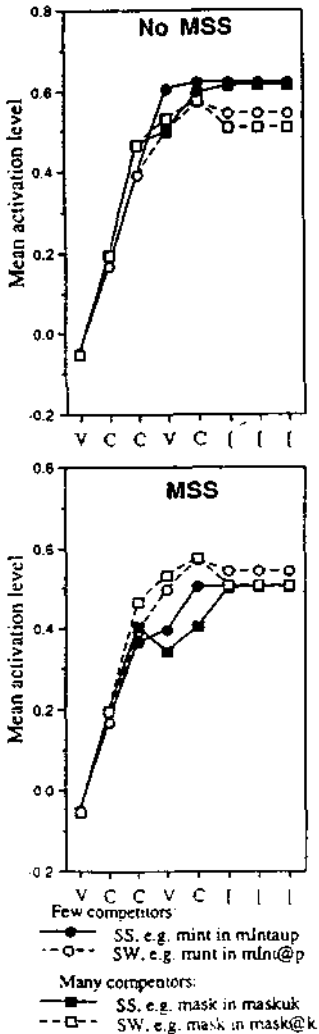
Figure 1. Shortlist simulations of data shown in Table 1 with and without the MSS.

recomputing a new optimised interpretation of the input. Operating in this mode the model shows appropriate sensitivity to the effects of late arriving competitors and is also able to revise early interpretations in the light of new and contradictory data.

V. CONCLUSIONS

The data described here provide convincing evidence for the operation of both lexical competition and metrically guided segmentation. We now have evidence that competition occurs between words beginning both earlier and later than the target word and that the effects of competition increase as the number of competitors increases. Competition and segmentation effects interact such that the effects of metrical segmentation are seen only at points where there are many competitors. All of these effects can be accurately simulated in a modified version of Shortlist incorporating the MSS. In this new model metrical information acts to modify the goodness of fit metric according to the match between the metrical structure of the input and the lexical representation of the word.

REFERENCES

[1] Norris, D.G. (1991). Rewiring lexical networks on the fly. In *Proceedings of Eurospeech* '97, Genova, 1, 117-120.
[2] Norris, O.G. (in press). Shortlist: A hybrid connectionist model of continuous speech recognition. *Cognition.*
[3] Cutler, A. & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance,* 14, 113-121.
[4] McClelland, J.L. & Elman, J.L. (1986). The TRACE model of speech perception. *Cognitive Psychology,* 18, 1 -86.
[5] McQueen, J.M., Norris, D.G. & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. Journal of *Experimental Psychology: Learning, Memory, and Cognition,* 20, 3,621-628.
[6] Cutler, A. and Norris. D. (1 979) Monitoring Sentence Comprehension. In W.E. Cooper and E.C.T. Walker (Eds.)
[7] Norris, D.G., McQueen, J.M. and Cutler, A. (submitted) Competition and Segmentation in Spoken Word Recognition

interpretation of the input which is uninfluenced by earlier analyses. In practice this is acheived by setting the activation levels of lexical nodes to zero on presentation of each new phoneme and then