Real Words, Phantom Words and Impossible Words

Anne Cutler

Abstract

Because spoken language arrives as a continuous signal, recognising words in real speech requires segmentation of the speech stream. Current computational models of human spoken-word recognition conceive of this process as involving automatic activation of word candidates, with subsequent competition between them, and experimental evidence supports the proposal that recognition involves competition. Word representations are activated by incoming phonetic information, which can be segmental or suprasegmental. Although models of spoken-word recognition have concentrated primarily on the evaluation of matching and mismatching information at the segmental level, with little attention to the role of suprasegmental information, experimental evidence from a variety of prosodically differing languages now confirms that suprasegmental information is used in word recognition insofar as it is informative and, where it is exploited, is used as it becomes available. The activation and competition process is modulated by further constraints which aid in the disposal of "phantom" words. Studies of vocabulary structure across languages show that there is extensive embedding and overlap of words within the vocabulary, so that in principle phantom words in the input could pose a serious problem to the recogniser. However, the competition process is able to exploit language-specific cues to segmentation, i.e. to the most likely positions for word boundaries. Moreover, listeners are sensitive to what could and could not be a possible word of their language, and by rejecting impossible words, they further speed the recognition of real words.

1. Words in Speech

A visual representation - oscillogram or spectrogram - of a naturally spoken utterance reveals a nearly continuous acoustic signal. Units which are separated in a written text such as this one, i.e. in English, are not at all demarcated in the speech stream. There are discontinuities in the spoken signal; but they rarely correspond to the boundaries signalled by white spaces in text. Stop consonants cause a momentary cessation of the signal; but stop consonants can occur in the middles of words as well as at word edges. Word edges as such are not necessarily signalled at all. Speech is continuous; words run into one another with no discernable discontinuity. Yet the listener's subjective experience is one of effortlessly perceiving a sequence of separate words, one after the other. And the recognition of spoken language must include recognition (and subsequent integration) of individual words - or rather, recognition

of whatever individual units are scored in our lexical memory. It is clear that whole utterances cannot be stored in lexical memory, since most of the utterances we hear have not been part of our previous language experiences.
Thus the listener must effectively segment the continuous speech stream into its component units. The boundaries of these units are not reliably signalled in speech. Yet listeners have no apparent difficulty in experiencing speech input as a sequence of individual words. How is this apparently effortless solution of a formidable problem achieved?

The present paper describes psycholinguistic research on this issue. The study of spoken-word recognition by human listeners has led to the development of a wide array of useful laboratory tasks for investigating (particular aspects of) this process - Grosjean and Frauenfelder (1997) provide an overview of this repertoire. Typically these experiments use tasks in which listeners' accuracy and response time is measured; the response in question may be a decision as to whether the string as a whole is or is not a real word (lexical decision; Goldinger, 1996), the detection of real words embedded in otherwise nonsense input (word spotting; McQueen, 1996), or lexical decision to a visually presented string which has been preceded by a spoken word or word fragment (cross-modal priming; Zwitserlood, 1996). There is now a large body of relevant data from experiments using these tasks; and the evidence now supports a view of human spoken-word processing based, as the following sections describe, on automatic activation of word forms, and competition between activated words where the signal may be simultaneously compatible with more than one interpretation. Current models of spoken-word recognition in psycholinguistics have converged on this view, resulting in a greater degree of unanimity in this field than was previously the case.

2. Modelling the Recognition of Real Words

Psycholinguistics is an experimental science, and experiments using the laboratory tasks mentioned above produce a huge array of data on how human listeners recognise words in speech. But as in any science, experiments must be carried out in an intimate relation with theory; the theory generates predictions which are tested in the experimental laboratory, and the results of the experiments constrain the development of the theory. Theoretical models in psychology in general, including those in the area of spoken-word recognition, have become increasingly precise over the past decades; the early descriptive models gave way to explicit information-processing accounts, and these in turn were supplanted by computational models which allow computer simulations (usually in a considerably simplified form, it must be admitted) of psychological processes. This has had the advantage of making the connection between theory and experiment even easier to demonstrate, as (in the domain of spoken-word recognition, for instance) simulations can be conducted with the computational model using the very experimental materials with which an experiment is conducted, so that the predictions from the model and the results from the experiment can be compared directly.

Consider the utterance *ship inquiry* (example from Norris, 1994). This will be pronounced, by a British native speaker, with stress on the second syllable of *inquiry,* and assimilation of the nasal to the velar place of articulation of the following stop, hence: [Sipinkwaieri]. This string is compatible with a number of possible (British English) words: *ship, shipping, choir, ink, inquire, inquiry, why, wire, wiry;* in other words, there are a number of possible paths through the lattice of interpretations of the string, as Figure 1 makes apparent.

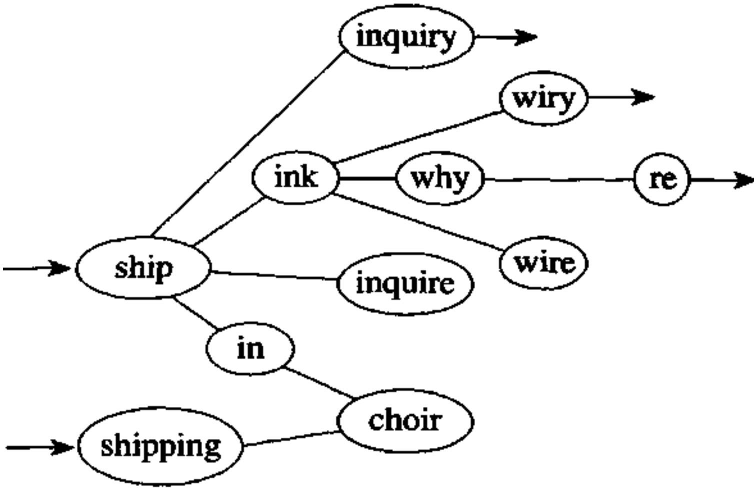

Figure 1. Possible paths through the string *ship inquiry.*

Current models of human spoken-word recognition (such as TRACE: McClelland & Elman, 1986; Shortlist: Norris, 1994; or the latest version of the Cohort model: Gaskell & Marslen-Wilson, 1997) assume that words which are compatible with the input are automatically activated and compete with one another for recognition. The manner in which the competition is instantiated differs from model to model; here we will use the framework of the Shortlist (Norris, 1994) model. The principal reason for choosing Shortlist is in fact that the model operates with a realistically sized lexicon of tens of thousands of words, as well as the full phonemic inventory of the language (although the work will be described via simulations with English, note that Shortlist has in fact been implemented with other languages also, and could in principle be implemented with any language; the only limiting factor at present is the availability of computationally tractable phonetically transcribed lexical databases, which so far exist for only very few of the world's languages). In Shortlist, therefore, the results of an experiment, or the processing of a particular input string, can be simulated, and the simulations will reflect the real availability of competition in the phoneme repertoire and vocabulary of a normal adult user of the language. Figure 2 shows an example of the kind of output produced by simulations with Shortlist; the input is specified phoneme by phoneme, and the output plots the activation levels of candidate words at each point in the input.

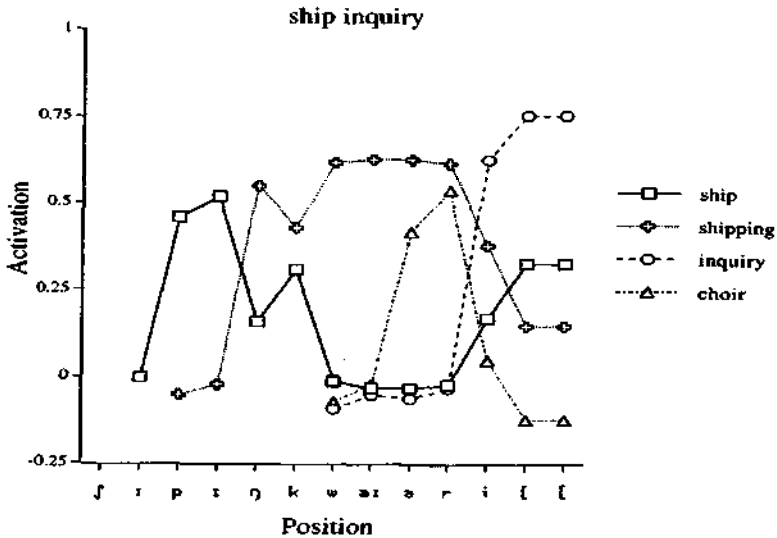

Figure 2. Word activation patterns in Shortlist, given the input *ship inquiry.*

Competition is modelled in Shortlist by interactive activation, including lateral inhibition between units at the same level. At the word level, any candidate word will compete with other words which incorporate portions of the same input - thus in the above example *shipping* will compete with *ship* for its initial portions, and with *ink, inquire* and *inquiry* for its last part. The more a given word is activated by the incoming speech, the more it is able to compete with - inhibit - other activated words; as words are inhibited, they lose activation and therefore become less able to inhibit other words. This process will eventually lead to only one successful competitor for each part of the input. As Figure 2 shows, *ship* is initially strongly activated by the input *ship inquiry,* but as the following syllable [in] arrives, *shipping* becomes more strongly activated and succeeds in inhibiting *ship.* However, the next syllables then arrive and as the string is now also compatible with *inquire* and *inquiry* and *choir,* these words become activated; at first, *choir* is able to win the competition because *shipping* is inhibiting *inquire* and *inquiry* (by competing for the [irj]), but is not inhibiting *choir*. When the final syllable has arrived, though, *inquiry* has such strong support from the input that it is able to inhibit *choir* and *inquire,* and also *shipping.* Once *shipping* is no longer strongly activated, it can no longer inhibit *ship,* with the result mat activation of *ship* is able to rise; at the end of the input, therefore, the two most highly activated words are *ship* and *inquiry,* corresponding to the actual content of the input string.

3. Activating Real Words

Because Shortlist is able to operate with a realistically sized lexicon, and the full phoneme set of a language, it can support sensitive simulations of the results of experiments on word recognition by human listeners. Such experiments, in particular with the cross-modal priming task, have produced abundant evidence of automatic activation of multiple alternative words. For instance, Zwitserlood (1989) showed that a word beginning which could fit either of two words would produce activation of both (an English example would be *paral-* activating *parallel* and *paralyse).* Words embedded in longer words can be activated as well as the words in which they occur (Cluff & Luce, 1990; Shillcock, 1990; Gow & Gordon, 1995). Ambiguous strings consistent with two words lead to activation of both (e.g. *-oast* beginning with something which could be a [k] or a [g] activates both *coast* and *ghost,* Connine, Blasko & Wang, 1994).

Of course, the phonetic structure which distinguishes one word from another can be segmental or suprasegmental. Models of spoken-word recognition have concentrated almost exclusively on the evaluation of matching and mismatching information at the segmental level, with little attention to the role of suprasegmental information in the activation of and selection between lexical candidates. Partly this is because of the inevitable English-centredness of most research in this area. English is a stress language; that is, in every polysyllabic word one of the syllables will be *more* salient than the others. Thus the first syllable is stressed in *petrol,* the second syllable is stressed in *patrol;* the same difference holds in *nature/mature, simply/imply, senate/sedate* and numerous other pairs. The stress differences are expressed in the suprasegmental structure, in that waveforms of any such pair will show that the stressed syllables are longer and louder than the unstressed syllables, and pitch traces will reveal more pitch movement on the stressed than on the unstressed syllables. However, the suprasegmental differences are accompanied by segmental correlates of the stress differences as well. Although the spelling of the word pairs reveals no hint of differences in the vowel quality, the second syllables of *petrol* and *patrol* in fact contain quite different vowels. The same is true of the second syllables of *simply* and *imply,* both syllables of *nature* and *mature* and of *senate* and *sedate,* and so on; in every case the unstressed syllable has a weak (reduced) vowel (schwa in *petrol, nature, senate, mature* and *sedate,* short [i] in *simply)* whereas the stressed syllable contains a full (i.e. unreduced) vowel.

There are only a handful of word pairs in English which differ in stress but have the same segmental structure - *the forbear* vs. *to forbear, the foregoing* vs. *forgoing, trusty* vs *trustee,* and about a dozen more. Thus English listeners can identify all but a tiny fraction of their language's vocabulary by reference to segmental structure alone, ignoring suprasegmental structure. Of course, very many words in English (as in most languages) have more than one meaning - *match, terminal* and so on have quite distinct meanings but absolutely no difference between these meanings in how they are pronounced. If listeners actually do not bother to refer to the suprasegmental structure when activating the stored representations of English words, then *foregoing* and *forgoing,* for instance, would both be activated by the initial processing of the segmental structure, just as the processing of [maetf] would activate multiple meanings *of match.* And indeed, there is evidence from a cross-modal priming study by Cutler (1986) that presentation of either member of a minimal stress pair such as *foregoing/forgoing* does momentarily activate words related to both members of the pair. Cutler argued that the number of homophones *(match, terminal* etc) in English is so large that adding a few extra *[forbear, trusty/trustee* etc.) is an insignificant cost compared with the enormous saving of not having to take suprasegmental structure into account in the process of initial activation of word forms.

Yet to speakers of tone languages such concentration on the segmental information must seem very foolish. Investigations of the role of lexical tone in spoken-word recognition suggest that the contribution of tonal information to word recognition is analogous to that of segmental information. For example, Cutler and Chen (1995; Chen & Cutler, 1997) conducted studies on phonological similarity effects in Cantonese, in which subjects decided whether or not the second of a pair of two spoken two-syllable items was a real word of Cantonese. The crucial item pairs were phonologically related (but semantically unrelated), with the phonological difference in either the tone or the rime of one syllable. In both experiments alterations of rime and of tone between prime and target had exactly parallel effects, implying that the contribution of tonal information and segmental information to auditory word recognition is comparable.

Pitch accent in Japanese resembles tone in that contrasts are realized via pitch variation. Cutler and Otake (1999) studied the role of pitch accent in the recognition of Japanese words. Listeners could reliably tell whether a single syllable *ka* had been extracted from a word in which it bore high (H) accent or a word in which it bore low (L) accent. Identification was more accurate for initial than for final syllables, suggesting that pitch accent information is realized most clearly in just the position where it would be of most use for listeners in on-line spoken-word recognition. A subsequent gating experiment confirmed that listeners do use pitch accent information effectively at an early stage in the presentation of a word, and use it to constrain selection of lexical candidates. Pairs of Japanese words such as *nimotsu/nimono,* beginning with the same CVCV sequence but with the accent pattern of this initial CVCV being HL in one word and LH in the other, were presented, in increasingly large fragments, to native speakers of Japanese. After presentation of each fragment, which was incremented in each case by one phoneme transition from the previous fragment, listeners recorded a guess regarding the word's identity and a confidence rating for that guess. The results showed that the accent patterns of the word guesses corresponded to the accent patterns of the actually spoken words with a probability significantly above chance from the second fragment onwards - i.e., from the middle of the first vowel (ni-).

This pattern of evidence might seem to warrant the conclusion that suprasegmental structure constrains word activation in tone languages and pitch accent languages (in both of which just one suprasegmental dimension, namely pitch, encodes the relevant distinctions), but not in stress languages (in which the suprasegmental encoding of stress involves more than one dimension). However, it is now clear that even other stress languages do not pattern like English. Dutch, for instance, although it resembles English very strongly in its overall prosodic structure, differs from it in the strength of the link between stress and vowel quality. Just as in English, stressed Dutch syllables must contain lull vowels and reduced vowels must be in unstressed syllables. But full vowels in unstressed syllables, which occur relatively rarely in English, are quite common in Dutch. Consider for instance the Dutch words *octopus* and *oktober,* which mean exactly the same as their English cognates. In the English words, the vowels in the second syllables differ - *octopus* has schwa in its second syllable. In the Dutch words the vowels in the second syllables are the same; in fact both words begin with the same sequence of four segments *octo-*. Priming experiments in Dutch (Donselaar & Cutler 1997) have shown that presentation of the first two syllables of *octopus* activates *octopus* but not *oktober,* while presentation of the first two syllables of *oktober* activates *oktober* but not *octopus*. Other experiments (Koster & Cutler, 1997; Cutler & Donselaar, submitted) have confirmed that Dutch listeners can use suprasegmental information to constrain lexical activation, and the priming effects also appear in another stress language, Spanish (Soto, Sebastian & Cutler, submitted).

Thus English would appear to be an isolate: an atypical language in which a source of information which proves valuable for constraining word activation in closely related languages is ignored because it contributes relatively little in comparison to its cost. This is the conclusion reached by Cutler, Dahan and Donselaar (1997). In many languages, however, distinctions between words will be realised via both suprasegmental and segmental information, and both types of information will constrain the candidate words which are activated by spoken input.

4. Disposing of Phantom Words

The incoming segmental and suprasegmental information leads, as we saw above, to automatic activation of candidate words compatible with the input. But are these activated words only the words really uttered by the speaker? Consider the fact that every natural human language has a vocabulary running into the tens of thousands, yet no language has a phonemic inventory running into even the low hundreds. Languages with a wide variety of tonal or other suprasegmental contrasts can add further dimensions of distinction; but even a "wide" variety of tonal contrasts is only a handful (six in Cantonese, for example). The inevitable consequence of this situation is that the words of any language resemble one another strongly, and are often found to be embedded within one another. Thus most speech signals will contain quite a high proportion of "phantom words", and the amount of potential competition represented by these spurious candidates for recognition should not be underestimated. Studies of the vocabulary have shown that most longer words contain shorter embedded words (McQueen & Cutler, 1992; McQueen, Cutler, Briscoe & Norris, 1995; Frauenfelder, 1991).

The research referred to in the preceding section also shows that such phantom words are indeed activated - words embedded in other words, for instance. Cluff and Luce (1990) examined cases like *mad* in *madcap;* Shillcock (1990) cases like *bone* in *trombone;* Gow and Gordon (1995) cases like *tulips* in *two lips*. Activation of the embedded words was observed in each case. Even words embedded across the boundaries of other words can be activated - Tabossi, Burani and Scott (1995) presented Italian listeners, in a cross-modal priming study, with sentences containing phrases like *visi tediati* (>bored faces=), and found activation of the embedded word *visite* (>visits=). So phantom words in the input certainly play a real role in the process of spoken-word recognition. How do listeners manage to settle on the real words and dispose of the phantom words?

The experimental evidence suggests that this is done by active competition between the simultaneously activated words. Studies with the word-spotting task, for instance, have shown that such competition occurs. For instance, words are harder to spot if the remainder of the string partially activates a competing word. Thus listeners presented with the auditory input [names] will correctly and rapidly spot the presence of a real English word *(mess)* in that string; if, however, the string is [dames], word-spotting will be slower and less accurate, because the latter string, although it also contains *mess,* is the beginning of another word - *domestic* - and presumably activates that word as well as *mess,* leading to competition (McQueen, Norris & Cutler, 1994). The same result can be observed in Dutch - *zee* ('sea') is easier to spot in *luzee,* which activates no competitor in the Dutch vocabulary, than in *muzee,* which activates the Dutch word *museum* (Cutler & Donselaar, submitted).

Moreover, the more competitors there are, the more they will interfere with spotting an embedded word. This conclusion can be drawn from a study by Norris, McQueen and Cutler (1995). They built upon an earlier study by Cutler and Norris (1988), who found that words such as *mint or jump* were harder to spot when followed by a full vowel *(mintayf, jumpoove)* than when followed by the reduced vowel schwa *irnintef, jumpev).* Cutler and Norris explained this as the effect of a segmentation strategy employed by English-speakers whereby syllables with full vowels were assumed to be word-initial. Thus the second syllable of *min-tayf* would be segmented from the first and detection of the embedded word would be slowed by the necessity of recombining material across a point at which the segmentation procedure had applied. Norris et al. (1995) then argued that segmentation for lexical access should result in greater interference if there were more competitors beginning at the segmentation point that if there were few; *mask* should be harder to find in *mas-kuck* (because there are lots of English words beginning *ku- - cut, cup, company* etc.) than *hint* in *hin-towp* (because there are relatively few words beginning *tow-).* This was exactly the result they found (and a parallel finding from a cross-modal priming experiment in Dutch was reported by Vroomen and de Gelder, 1995).

The evidence is thus compelling that spoken-word recognition by human listeners involves automatic activation plus a process of competition via which the real words which were spoken triumph over spuriously activated competitors. In other words, the evidence from experiments with human listeners supports the kind of model which most psycholinguists now subscribe to: like Shortlist, one involving competition. All the results from the experiments described in this section have indeed been reproduced in Shortlist simulations with, in each case, the full set of materials used in the experiment. Figure 2, that is to say, captures something of the true pattern of activity in the lexical processing system of a human listener. Nevertheless the competition process efficiently produces the right outcome in nearly every case - only very rarely do listeners become aware of the phantom words embedded in the strings of real words which they hear. The possible paths and possible alternative word candidates represented in Figures 1 and 2 do not cause noticeable difficulty to the human listener. Recent research, described in the next section, has identified a constraint which effectively applies to reduce the activation of at least some potential phantom words, undoubtedly contributing to the eventual victory of the real words.

5. Rejecting Impossible Words

In a word-spotting study by Norris, McQueen, Cutler and Butterfield (1997), English listeners were presented with words like *egg,* embedded in nonsense strings *like/egg* and *maffegg.* In *fegg,* the added context [f] is not a possible word of English - there are no English lexical items consisting of a single consonant. In contrast, the added context *maff* in *maffegg,* although it is actually not a word of English, might conceivably have been one - *mat, muff and gaff are* all English words. Listeners were faster and more accurate in detecting real words embedded in possible-word than in impossible-word contexts, whether the context preceded (*fegg, maffegg)* or followed the target *(sugarth, sugarthig);* in other words, the listeners in this study found it hard to detect a word if the result of recognising it was to leave a residue of the input which could not itself be parsed into potential words.

In Shortlist, these data were simulated by reducing the activation of any candidate word which leaves no vocalic segment between the edge of the word and the nearest known boundary in the input. In this experiment, of course, the nearest known boundary is the silence at each end of the stimulus string. But the constraint can be implemented in a more general way to capture other boundary effects known to be exploited by listeners, such as segmentation at the onset of syllables with a full vowel (Cutler & Norris, 1988), or at phonologically mandatory syllable boundaries (McQueen, 1998). The Shortlist simulations, described in detail in Norris et al. (1997), accurately captured the pattern revealed in the data from a wide range of listening experiments.

Further investigations of the possible-word constraint addressed its language-specificity versus universality. Languages differ in the precise constraints which apply to what may or may not be a word. Yet all human listeners, whatever their language, deal effectively with the phantom words and partially activated words which occur, in any speech input, as the inevitable result of a large vocabulary constructed from a small phonemic repertoire. They do so by drawing on highly effective mechanisms which allow words to compete with one another for the input, as well as on further constraints which include an early filter to rule out any segmentation which would postulate an impossible word.

Thus it seems that the prevalence of phantom words in a real-word input, inevitable consequence though it is of the structure of human languages, poses no real problem for the human listener. Spoken-word recognition in utterance contexts involves a combination of automatic activation by all relevant dimensions of phonological structure in the input, suprasegmental as well as segmental, plus effective competition between the activated candidate words; the competition is constrained by the immediate rejection of impossible words and handicapping of phantom word candidates which imply such impossible words. Together these procedures allow us to listen, recognise and understand.

References

Chen, H.-C, & Cutler, A. (1997). Auditory priming in spoken and printed word recognition. In H.-C. Chen (Ed.), *The Cognitive Processing of Chinese and Related Asian Languages* (pp. 77-81). Hong Kong: Chinese University Press.

Cluff, M.S., & Luce, P.A. (1990). Similarity neighborhoods of spoken two-syllable words: Retroactive effects on multiple activation. *Journal of Experimental Psychology: Human Perception & Performance,* 16, 551-563.

Connine, C.M., Blasko, D.G., & Wang, J. (1994). Vertical similarity in spoken word recognition: Multiple lexical activation, individual differences, and the role of sentence context. *Perception & Psychophysics,* 56,624-636.

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., & Chen, H.-C. (1997). Lexical tone in Cantonese spoken-word processing. *Perception & Psychophysics,* 59, 165-179.

Cutler, A., & Chen, H.-C. (1995). Phonological similarity effects in Cantonese word recognition. *Proceedings of the Thirteenth International Congress of Phonetic Sciences, Stockholm,* Vol. 1, 106-109.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech,* 40, 141-201.

Cutler, A., & van Donselaar, W. (submitted). *Voornaam* is not a homophone: Lexical prosody and lexical access in Dutch.

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., & Otake, T. (1999) Pitch accent in spoken-word recognition in Japanese. *Journal of the Acoustical Society of America,* 105, 1877-1888.

van Donselaar, W., & Cutler, A. (1997). Exploitation of stress information in spoken-word recognition in Dutch. *134th Meeting, Acoustical Society of America, San Diego,* December 1997.

Frauenfelder, U.H. (1991). Lexical alignment and activation in spoken word recognition. In J. Sundberg, L. Nord, & R. Carlson (Eds.), *Music, Language, Speech and Brain* (pp. 294-303). London: Macmillan.

Gaskell, M.G., & Marslen-Wilson, W.D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes,* 12, 613-656.

Goldinger, S.D. (1996). Auditory lexical decision. *Language and Cognitive Processes,* 11,559-567.

Gow, D. W., & Gordon, P. C. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception and Performance,* 21, 344-359.

Grosjean, F., & Frauenfelder, U.H. (1997). *A Guide to Spoken Word Recognition Paradigms.* Psychology Press, Hove (UK).

Koster, M., & Cutler, A. (1997). Segmental and suprasegmental contributions to spoken-word recognition in Dutch. *Proceedings ofEurospeech 97, Rhodes,* 2167-2170.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology,* 18,1-86.

McQueen, J.M. (1996). Word spotting. *Language and Cognitive Processes, 11,* 695-699

McQueen, J.M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language,* 39, 21-46.

McQueen, J.M., & Cutler, A. (1992). Words within words: Lexical statistics and lexical access. *Proceedings of the Second International Conference on Spoken Language Processing, Banff, Canada,* Vol. 1,221-224.

McQueen, J.M., Cutler, A., Briscoe, T., & Norris, D.G. (1995). Models of continuous speech recognition and the contents of the vocabulary. *Language and Cognitive Processes,* 10,309-331.

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,* 20,621-638.

Norris, D. G. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition,* 52,189-234.

Norris, D.G., McQueen, J.M., and Cutler, A. (1995). Competition and segmentation in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 21, 1209-1228.

Norris, D.G., McQueen, J.M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology,* 34, 191-243.

Shillcock, R.C. (1990). Lexical hypotheses in continuous speech. In G.T.M. Alrmann (Ed.), *Cognitive Models of Speech Processing* (pp. 24-49). Cambridge MA". MIT Press.

Soto, S., Sebastian-Galles, N., & Cutler, A. (submitted). Segmental and suprasegmental cues for lexical access in Spanish.

Tabossi, P., Burani, C, & Scott, D. (1995). Word identification in fluent speech. *Journal of Memory and Language,* 34,440-467.

Vroomen, J., & de Gelder, B. (1995). Metrical segmentation and lexical inhibition in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance,* 21, 98-108.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition, 32,*25-64.

Zwitserlood, P. (1996). Form priming. *Language and Cognitive Processes, 11,* 589-896.