

6 Prosody and the Word Boundary Problem

Anne Cutler

MRC Applied Psychology Unit

THE WORD BOUNDARY PROBLEM

The problem with word boundaries lies in locating them. In most spoken language, few cues are available to signal reliably where one word ends and the next begins. However, understanding spoken language must be a process of understanding discrete words rather than utterances as indivisible wholes, because most complete utterances have never previously been experienced by the listeners to whom they are directed. To understand a spoken utterance, therefore, listeners must somehow, in the absence of explicit signals, locate the boundaries between the individual words (or more precisely, the lexically represented units, whatever these may be) of which the utterance is composed.

Models of spoken-word recognition have addressed the word boundary problem in several ways, but the proposed solutions fall into two principal classes: Those which incorporate some explicit mechanism for the location (or at least postulation) of word boundaries, versus those which avoid the need for explicit word boundary location, by proposing that boundary information simply falls out of the normal processes of word recognition. The former class, which we can term Explicit Segmentation models, is differentiated according to the principles that various models propose for guiding word boundary location. The latter class, Serendipitous Segmentation models, essentially contains two candidate solutions: one based on word recognition in strictly sequential order, and one based on competition between word candidates for recognition.

Explicit Segmentation models include the proposal that, in English, listeners apply a strategy of assuming that any strong syllable in the input is word-initial (Cutler & Norris, 1988; Cutler & Butterfield, 1992). This "Metrical Segmentation Strategy" (Cutler, 1990) is efficient in that most English lexical words do indeed begin with strong syllables, and most strong syllables in typical utterances are indeed word-initial (Cutler & Carter, 1987; see that article for more details on the proposal for English, including separation of the access of open- versus closed-class words). Similar prosodically based proposals for explicit segmentation exist for other languages, for example, the proposals that French listeners use a syllabic segmentation procedure (Mehler, Dommergues,

Frauenfelder, & Segui, 1981) or that Japanese listeners use a mora-based procedure (Otake, Hatano, Cutler, & Mehler, 1993).

Serendipitous Segmentation models see word boundary information as arising incidentally from the processing operations of recognition. Cole and Jakimik (1978) made a clear statement of such a model in the elaboration of their proposal that recognition of spoken utterances proceeds in strictly temporal order: "one word's recognition automatically directs segmentation of the immediately following word" (1978, p. 93). A similar sequential recognition proposal is embodied in Marslen-Wilson and Welsh's (1978) Cohort Model, which focused on the fact that some words become unique (distinct from all other words in the language) prior to their acoustic offsets; the claim here is that the recognizer can in such cases predict in advance where the current word will end and, by implication, where the next will begin. Competition among candidate words, as embodied in recent connectionist models such as TRACE (McClelland & Elman, 1986) and SHORTLIST (Norris, 1991, 1994), provides another mechanism by which segmentation emerges from independent processing operations. Given an input string, candidates matching any part of the string will be activated, and competition will occur among them. The competition will in general be won by any highly activated sequence of competing candidates that successfully accounts for the entire string without any leftover portions. This process permits no role for boundary detection processes, since words beginning *at any point* in the input can be activated.

A feature of the major classification into explicit versus serendipitous segmentation that is rarely acknowledged is that only the proposal that segmentation be explicit simultaneously addresses the word boundary problem from the point of view both of the adult language user and of the prelinguistic infant. As Mehler, Dupoux, and Segui (1990) spelled out in detail, the word boundary problem for an infant is substantially greater than that for an adult language user. For the prelinguistic infant, most speech input is continuous, just as it is for the adult listener; although caretakers will, in many communities, explicitly teach words, this caretaker behavior appears *after* the infant has begun to produce language (i.e., essentially after the initial segmentation problem has been solved; see Cutler, 1994, for a review). In solving the initial segmentation problem, though, the infant cannot rely on lexical knowledge at all; a lexicon must be constructed, and the construction process must be begun from the most minimal of bases. The most an infant can be born with in this respect might be, perhaps, the expectation that there will be words (i.e., that linguistic means will exist to express communicable knowledge in discrete memorizable chunks). Further than that there can be no specific expectations—for example, there can be no expectations regarding the *structure* of words, since this varies widely from language to language. Certainly the infant confronted with the first samples of speech input can have no preexisting stock of words or

word templates on which to build. For this simple reason, any proposal involving serendipitous segmentation (i.e., the emergence of word boundary information from the normal process of recognizing known words in speech input) offers no aid to the prelinguistic infant attempting to decompose continuous speech input into its component words.

Explicit segmentation, on the other hand, does offer such aid. The infant has to start somewhere with the decomposition of continuous speech input, and an explicit segmentation proposal amounts to a claim that this *starting point* arises from an explicit procedure. Once a start has been made, and a small stock of known words exists, the way is of course open for serendipitous segmentation to begin operating; even then, however, serendipitous procedures will only work at the boundaries of those known items, so there will be limits on their effectiveness until a very much greater lexical stock has been built up. Thus explicit segmentation models offer advantages that may be overlooked if they are considered only in comparison with their rivals as accounts of the adult word recognition process.

EVIDENCE FROM ADULT PROCESSING

Adult word recognition evidence offers substantial support for the explicit segmentation position. As mentioned earlier, experiments in English have suggested that listeners segment speech at strong syllable onsets. For example, finding a real word in a spoken nonsense sequence is hard if the word is spread over two strong syllables (e.g., *mint* in [mintef]) but easier if the word is spread over a strong and a following weak syllable (e.g., *mint* in [mintef]; Cutler & Norris, 1988). The proposed explanation for this is that listeners divide the former sequence at the onset of the second strong syllable, so that detecting the embedded word requires recombination of speech material across a segmentation point, while the latter sequence offers no such obstacles to embedded-word detection as the non-initial syllable is weak and so the sequence is simply not divided. Similarly, when English speakers make slips of the ear that involve mistakes in word boundary placement, they tend most often to insert boundaries before strong syllables (e.g., hearing *by loose analogy* as *by Luce and Allergy*) or delete boundaries before weak syllables (e.g., hearing *how big is it?* as *how bigoted?*; Cutler & Butterfield, 1992). These findings prompted the proposal of the Metrical Segmentation Strategy for English (Cutler & Norris, 1988; Cutler, 1990), whereby listeners are assumed to segment speech at strong syllable onsets because they operate on the assumption, justified by distributional patterns in the input, that strong syllables are highly likely to signal the onset of lexical words.

Moreover, there is strong reason to believe that one of the serendipitous segmentation proposals simply could not work for English; the efficiency of strictly sequential word-by-word recognition is dependent upon words not being mistaken for one another, but it is clear that in

the case of the English vocabulary this criterion cannot be met. Words have other words embedded in them, and these are overwhelmingly at the beginning (McQueen & Cutler, 1992; McQueen, Cutler, Briscoe, & Norris, in press). Thus *fundamentalism* contains *fun*, *fund*, *fundament*, and *fundamental* (as well as *men*, *meant*, *mental*, and *mentalism*), *circumference* contains *succumb* (in British English), *chemotherapy* contains *key*, *battery* contains *bat* and *batter*, *startle* contains *star* and *start*, and so on. This problem is in fact not unique to English, since at least in Dutch the same pattern is found (Frauenfelder, 1991). The problem is not solved by relying, for instance, on syntactic disambiguation, since embedded words and the words in which they are embedded often match in syntactic class (McQueen, Cutler, Briscoe, & Norris, in press).

The effect of this embedding is that a sequential recognition model will often not be able to assume that an incoming string matching a lexical entry is indeed a token of that lexical entry until subsequent input has ruled out the possibility that the string is only part of a larger word; for example, *star* cannot be recognized as *star* until subsequent input has ruled out the possibility of *start*, *startle*, *starling*, and so forth. Indeed, Luce (1986) has computed that in typical speech contexts more than a third of all words are likely to be potentially continuable in this fashion. Thus listeners simply cannot capitalize on the apparent efficiency of strictly sequential word recognition, because the vocabulary itself does not meet the necessary uniqueness criterion.

Under these circumstances it is not surprising to discover that experimental evidence confirms that listeners indeed do not recognize words strictly sequentially. Using a gating task in which words were presented incrementally in fragments, Grosjean (1985) showed that many short words could not be recognized until some time after their offset. In a similar task, differing from Grosjean's in that the input was spontaneous speech and was presented whole word by whole word, Bard, Shillcock, and Altmann (1988) found that words were often recognized in groups—that is, a word was not recognized until the following word was itself recognized.

Sequential recognition was initially justified by the possibility of exploiting early uniqueness—that is, identifiability of a word prior to its offset. However, the statistical analyses described earlier have shown that true cases of early uniqueness are rare, and the experimental evidence has shown that, without early uniqueness, word recognition may not even be strictly sequential. Therefore sequential recognition models should probably be abandoned as a class.

Competition models do not, however, suffer from the same problems. At the present time competition is a serious contender in the word recognition field, and the competition-based version of serendipitous segmentation built into models such as TRACE (McClelland & Elman, 1986) and SHORTLIST (Norris, 1991, 1994) offers a potential alternative to explicit segmentation accounts. In a direct test of competition and ex-

plicit segmentation (the Metrical Segmentation Strategy version for English), McQueen, Norris, and Cutler (1994) demonstrated, however, that evidence can simultaneously be found for both. When listeners were presented with a word-spotting task in which some words were embedded in strings with a preceding weak syllable (e.g., *mess* in [names], *sack* in [klesaek]) while others were embedded in strings with a following weak syllable (e.g., *mess* in [mestem], *sack* in [saekrek]), the former set proved easier to spot than the latter. This is exactly as predicted by the Metrical Segmentation Strategy: Segmentation at the onsets of strong syllables would place a boundary prior to *mess* in [names], for example, while no boundary would be placed between the syllables of a strong-weak string such as [mestem], so segmentation could not affect recognition in this case. However, it was also the case that words that were embedded in strings containing no potential competitor word (*mess* in [names], *sack* in [saekrek]) were detected more easily than words embedded in potentially competing strings (e.g., *mess* in [dames], which is the onset of *domestic*, or *sack* in [saekref], which is the onset of *sacrifice*). This is exactly as predicted by competition models, and indeed, as McQueen et al. (1994) pointed out, it constitutes more direct evidence for competition than had previously been available in the literature. Abundant evidence had been available for simultaneous activation of potential word candidates consistent with a given input (e.g., Goldinger, Luce, & Pisoni, 1989; Zwitserlood, 1989; Shillcock, 1990; Cluff & Luce, 1990; Goldinger, Luce, Pisoni, & Marcario, 1992). But words may be simultaneously active without actively competing with one another. The inhibition of *mess* by *domestic* in the input [dames] (in comparison with [names]) in the McQueen et al. study, though, seems to provide clear support for actual competition—*mess* was less easily recognized when *domestic* was actively contending for recognition.

Following the McQueen et al. finding, a subsequent study by Norris, McQueen, and Cutler (in press) confirmed the joint influence of metrical segmentation and competition in word recognition by demonstrating effects of the *number* of competitors for a given input string. In a word-spotting task, the disadvantage for detecting a word in a string of two strong syllables (e.g., *mint* in [mintef] in comparison with a strong-weak string (e.g., [mintef]) was larger when there were many potential words in the vocabulary beginning with the final consonant of the target word (here, the /t/ of *mint*) and the following vowel in the stimulus string than when there were few potential competitor words. In a cross-modal priming task (in which activation of lexical candidates by spoken input is measured via facilitation of recognition of simultaneously presented visual input) an analogous effect of number of competitors was demonstrated for Dutch by Vroomen and de Gelder (1995).

LANGUAGE-SPECIFICITY OF ADULT SEGMENTATION PROCEDURES

The fact that English listeners showed BOTH explicit segmentation and competition effects in the same recognition situation further suggests that competition cannot "explain away" the evidence that supports models involving explicit segmentation. Competition exists, but it does not offer a complete account of human word recognition, because the serendipitous segmentation that it allows does not appear to be the whole solution to the word boundary problem. Listeners use explicit segmentation as well. Given the additional dimension that explicit segmentation offers in the form of a potential handle on the infant's initial word boundary problem, it is particularly interesting that it appears to be used by adult listeners for whom it might have been thought to be potentially dispensable. As was described previously, however, the Metrical Segmentation Strategy version of explicit segmentation for English is undoubtedly efficient; such efficiency may be all that is needed to maintain the use of a segmentation procedure in adult recognition.

On the other hand, efficiency has not necessarily been demonstrated for explicit segmentation in the processing of languages other than English. Yet there is experimental evidence that clearly supports explicit segmentation in other languages. Unsurprisingly, though, the form that explicit segmentation takes in other languages is not exactly the form it takes in English; the Metrical Segmentation Strategy for English is founded on the opposition between strong and weak syllables that is such an important feature of English phonology, but other languages may have quite different phonologies, in which no such opposition can be drawn.

In French, for example, a contrast between strong and weak syllables is not a salient feature of phonological structure. Evidence from a wide variety of experimental tasks in French favors explicit segmentation into syllable-sized units (Mehler, Dommergues, Frauenfelder, & Segui, 1981; Segui, Frauenfelder, & Mehler, 1981; Cutler, Mehler, Norris, & Segui, 1986; Kolinsky, 1992; Pallier, Sebastian-Galles, Felguera, Christophe, & Mehler, 1993). Confirming evidence suggests that syllabic segmentation can be observed under certain conditions in other languages also—for instance, in Spanish (Sebastian-Galles, Dupoux, Segui, & Mehler, 1992; Bradley, Sanchez-Casas, & Garcia-Albea, 1993), in Catalan (Sebastian-Galles et al., 1992) and in Dutch (Zwitserslood, Schriefers, Lahiri, & van Donselaar, 1993). Syllabic segmentation is by no means the same process as the stress-based segmentation proposed, in the form of the Metrical Segmentation Strategy, for English.

Yet in one sense the procedures that have been experimentally demonstrated for English and for French are closely parallel. Both stress in English and the syllable in French are the basis of rhythmic structure in their respective languages. This parallelism prompted the hypothesis (see e.g., Cutler, Mehler, Norris, & Segui, 1992) that listeners might in

fact adopt a universally applicable solution to the word boundary problem, in that to solve it they exploit whatever rhythmic structure happens to characterize their language. This universal rhythmic segmentation hypothesis in turn led to the proposal that where a language has a rhythmic structure based on some phonological construct other than stress or the syllable, it should be possible to find evidence for exploitation of this construct in speech segmentation.

Japanese is such a language; its rhythm is described in terms of a subsyllabic unit, the mora. A mora can be a CV structure, or a single vowel, or a syllabic coda (usually a nasal consonant); thus *Honda*, for example, has three moras: Ho-n-da. (This is not a completely exhaustive list of mora structures, but it covers the vast majority.) Otake, Hatano, Cutler, and Mehler (1993) undertook to test the hypothesis that Japanese listeners should exhibit evidence of mora-based segmentation. They presented listeners with spoken words and required them to detect CV (consonant-vowel) or CVC targets within these words. For instance, detection responses were compared for TA- versus TAN- targets in *tanshi* (mora structure ta-n-shi) and *tanishi* (mora structure ta-ni-shi). In *tanshi* the TA- target corresponds to the first mora and the TAN- target to the first two moras. Subjects had no difficulty detecting either target (though they were faster detecting the one-mora than the two-mora target). In *tanishi*, TA- corresponds to the first mora; subjects detected it readily, and with the same speed that they detected the same target in *tanshi*. TAN- in *tanishi*, however, constitutes all the first mora and part of the second; that is, it does not correspond exactly to mora structure at all. In fact, subjects simply did not respond in this case—TAN- targets were overwhelmingly not detected in *tanishi*. Otake et al. interpreted this response pattern as evidence of mora-based segmentation by Japanese listeners.

Subsequent experiments extended this finding to other mora structures; detection of single-phoneme targets is faster and more accurate if they constitute exactly a mora (e.g., O in *aoki*, a-o-ki, N in *kanko*, ka-n-ko) than if they are part of a mora (e.g., O in *tokai*, to-ka-i, N in *kanoko*, ka-no-ko; Cutler & Otake, 1994).

The universal rhythmic segmentation hypothesis was thus supported: Explicit segmentation procedures are language-specific but only insofar as rhythmic structure is language-specific. Note that the presence of a particular rhythmic structure in the input does not of itself produce segmentation based on that structure. English listeners show no evidence of syllabic segmentation with French input, for example (Cutler et al, 1986), and neither do Japanese listeners (Otake, 1992); English listeners likewise show no evidence of mora-based segmentation of Japanese input (Otake et al., 1993; Cutler & Otake, 1994), nor do French listeners (Otake et al, 1993). The segmentation procedures are, instead, part of the processing repertoire of the listener rather than an input-driven phenomenon. Indeed, given the opportunity, listeners will apply their native language-specific procedures to foreign language input,

even in cases where the procedures may not operate efficiently at all. Thus French listeners apply syllabic segmentation to English input (Cutler et al., 1986) and to Japanese input (Otake et al., 1993), and Japanese listeners apply moraic segmentation where possible to English input (Cutler & Otake, 1994).

EVIDENCE FROM LANGUAGE ACQUISITION

The pattern of results summarized earlier suggests that development of a rhythmically based segmentation procedure is part and parcel of development of one's native language. Exactly how such procedures arise cannot as yet be illuminated by direct experimental evidence. One hypothesis might be that they arise as a result of fairly extensive exposure to the input language and the consequent acquisition of accurate models of the statistical probabilities of input patterns. For instance, as Cutler and Carter (1987) showed, stress-based segmentation is an extremely efficient strategy for solving the word boundary problem in English. However, there is evidence that adjustment to statistical probability patterns as a result simply of exposure to the language is *not* the source of segmentation procedures. If it were, then language users who have experienced extensive exposure to two languages with differing rhythmic structures should develop the segmentation procedures appropriate to both. However, they do not. Cutler, Mehler, Norris and Segui (1992) studied a group of balanced English-French bilinguals (i.e., speakers who were equally in command of both languages to indistinguishable native levels); these speakers, they found, commanded only one such procedure—either syllabic segmentation (characteristic of French) or stress-based segmentation (characteristic of English). A measure of language preference determined which procedure was available—if on this measure a subject was classed as "English-dominant," he or she used stress-based segmentation with English but did not use syllabic segmentation with French. If a subject was "French-dominant," then syllabic segmentation was used with French but stress-based segmentation was not used with English.

The explanation that Cutler et al. proposed for this finding was one based in the earliest stages of language acquisition. They proposed that explicit segmentation really does link the infant and adult processing situations. The beginning language user needs only one starting point; and whichever one the bilingual infant happens to get, that is the one that remains available throughout life. (Of course, establishing the precise parameters that determine the infant's options at this point is necessary to complete this explanation.) In consequence, the explicit segmentation procedures used by any adult have their source in that early experience of beginning lexical acquisition.

The Cutler et al. proposal is supported by evidence that infants are indeed highly sensitive to language rhythm (see Cutler, 1994, for a review of this evidence). Already in the first days of life, infants can make

durational discriminations between sets of bisyllables with versus without an internal word boundary (Christophe, Dupoux, Bertoncini & Mehler, 1994). Furthermore, there is also evidence that at the age at which infants are building up a receptive vocabulary without yet being in a position to use it, they are sensitive to the very prosodic characteristics of words that explicit segmentation procedures exploit. For instance, Jusczyk, Cutler, and Redanz (1993) showed that 9-month-olds in an English language environment prefer to listen to lists of words beginning with strong syllables than to lists of words beginning with weak syllables. These 9-month-old subjects were not yet producing any language; and the words in the input lists were largely low-frequency words to which they were unlikely ever to have been exposed. Accordingly their preferences seem likely to have been determined by development of a concept of what phonological form English words are most likely to take. Six-month-olds, Jusczyk et al. found, did not exhibit any preferences among the experimental word lists; so the 9-month-olds' performance seems likely to have been based on their recent experience with the initial stages of (as yet passive) vocabulary acquisition. That the 9-month-olds' preference was actually for a prosodic structure rather than for specific words was confirmed by the fact that the preference also emerged when the input was low-pass filtered to produce input with clear prosody but no discernible segmental structure.

Further supporting evidence appears in a series of studies by Morgan and his colleagues (Goodsitt, Morgan, & Kuhl, 1993; Morgan, 1994; Morgan & Saffran, in press). Infants show a preference for distributionally regular over irregular sequences, and in English-acquiring 9-month-olds, this preference extends to trochaic (e.g., strong-weak) over iambic (weak-strong) sequences. Moreover, the 9-month-olds showed evidence of integrating the prosodic and the segmental structure of the input, lending support to the Jusczyk et al. claim that infants at this age have acquired a concept of *word-level* prosody. (In contrast, 6-month-old infants in Morgan's studies showed no such evidence of integration.)

The Cutler et al. proposal remains as yet in need of direct test. One open question, for example, is the time span over which infants might develop the ability to segment speech input. Is the acquisition of a segmentation procedure a sudden, one-off experience? Or is it a process that takes place over a certain critical period? In the latter case, could statistical properties of the input *during that period* play a determining role (even though they may not be sufficient to induce segmentation procedures in the absence of critical-period sensitivity)? Cutler and Mehler (1993) referred to the infant's rhythmic sensitivity as an instance of a more general "periodicity bias"; their aim, the integration of the rhythmic segmentation proposals into a more general picture of the initial stages of language acquisition, depends upon empirically established answers to questions such as these.

THE WORD BOUNDARY SOLUTION

This chapter has attempted to give a summary overview of a range of research addressing the recognition of words in continuous speech. The central argument has been that the word boundary problem is a real one: For infants, it consists of dividing continuous speech input into lexically significant chunks to be stored, while for adults it consists of identifying known lexical items in the continuous input stream. Models of adult speech recognition exist that claim that there is no word boundary problem once one is in possession of a lexicon—segmentation occurs serendipitously, as a by-product of the normal recognition processes either of sequential processing or of competition. However, statistical analyses of vocabulary and speech corpora indicate that the distributional assumptions embodied in sequential models are unjustified; moreover, experimental evidence indicates that recognition is often not sequential. Although competition, on the other hand, is supported by experimental evidence from adult processing, the same findings also indicate that it co-exists with explicit segmentation, for which extensive evidence now exists across many languages.

Explicit segmentation has the strong theoretical advantage that it offers a solution to the word boundary problem both for the adult and for the infant listener. The second major argument presented here has been that the nature of this solution is in fact the same for adult and infant listeners, namely exploitation of prosodic structure. Summarized evidence from English, French, and Japanese suggests that the explicit segmentation procedures used by native speakers of these languages differ, but in a very systematic way. English listeners exploit stress patterns in speech segmentation, French listeners exploit the syllable as a unit, and Japanese listeners exploit mora structure; but underlying the language-specific realization of segmentation procedures is a universal similarity in that all three of these procedures can be interpreted as exploitation of the characteristic rhythm of the language. Prosodic structure, in the form of language rhythm, allows adult listeners to increase recognition efficiency via the application of explicit segmentation.

The claim that infant listeners use the same explicit segmentation procedures is based on less direct evidence, but is supported by a number of independent arguments. First, adult evidence from balanced bilinguals indicates that adult listeners can command only one rhythmically based segmentation procedure. This suggests that adult segmentation procedures are not developed simply as a result of extensive exposure to statistical properties of the native language, but may instead originate in a single learning experience. Second, evidence exists that infants are highly sensitive to rhythmic structure in language. And third, experimental studies have shown that at the stage that infants are developing a passive vocabulary (but before they show evidence of an active vocabulary), they are already sensitive to the prosodic probabilities of word structure in the language they are acquiring. Together

these strands of evidence motivate the claim that the explicit segmentation procedures used by adult listeners may in fact have their origin in the infant's exploitation of rhythmic structure to solve the initial word boundary problem.

It is no accident that explicit segmentation is satisfying at a theoretical level in that it offers an integration of adult and infant models of recognition: Adults use explicit segmentation precisely because infants do. Prosody bootstraps lexical segmentation and offers a solution to the word boundary problem. The prosodic option is in fact all that the infant has to rely on; but it is still sufficiently practical that the adult listener retains it in the repertoire of recognition processes.

ACKNOWLEDGMENTS

Much of the research reported in this chapter was supported by grants from the Joint Councils' Initiative in Cognitive Science (UK) and from the Human Frontier Scientific Program. Thanks to all the colleagues with whom the work was jointly conducted, especially James McQueen, Dennis Norris, and Takashi Otake. Further thanks to James McQueen, Jacques Mehler, Jim Morgan, and Katherine Demuth for comments on an earlier version of the manuscript.

REFERENCES

- 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 & Psychophysics*, *44*, 395-408.
- Bradley, D. C., Sanchez-Casas, R. M., & Garcia-Albea, J. E. (1993). The status of the syllable in the perception of Spanish and English. *Language & Cognitive Processes*, *8*, 197-233.
- Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America*, *95*, 1570-1580.
- Guff, 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.
- Cole, R. A., & Jakimik, J. (1978). Understanding speech: How words are heard. In G. Underwood (Ed.), *Strategies of information processing* (pp. 67-116). London: Academic Press.
- Cutler, A. (1990). Exploiting prosodic probabilities in speech segmentation. In G. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 105-121). Cambridge, MA: MIT Press.
- Cutler, A. (1994). Segmentation problems, rhythmic solutions. *Lingua*, *92*, 81-104.
- Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory & Language*, *31*, 218-236.
- Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language*, *2*, 133-142.
- Cutler, A., & Mehler, J. (1993). The periodicity bias. *Journal of Phonetics*, *21*, 103-108.
- Cutler, A., Mehler, J., Norris, D. G., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory & Language*, *25*, 385-400.
- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology*, *24*, 381-410.
- Cutler, A., & Norris, D. G. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception & Performance*, *14*, 113-121.
- Cutler, A., & Otake, T. (1994). Mora or phoneme? Further evidence for language-specific listening. *Journal of Memory & Language*, *33*, 824-844.

- 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.
- Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbours of spoken words: Effects of competition and inhibition. *Journal of Memory & Language*, 28, 501-518.
- Goldinger, S. D., Luce, P. A., Pisoni, D. B., & Marcario, J. K. (1992). Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 18, 1211-1238.
- Goodsitt, J. V., Morgan, J. L., & Kuhl, P. K. (1993). Perceptual strategies in prelingual speech segmentation. *Journal of Child Language*, 20, 229-252.
- Grosjean, F. (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception & Psychophysics*, 38, 299-310.
- Juszyk, P., Cutler, A., & Redanz, N. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, 64, 675-687.
- Kolinsky, R. (1992). Conjunction errors as a tool for the study of perceptual processing. In J. Alegria, D. Holender, J. Morais & M. Radeau (Eds.), *Analytic approaches to human cognition* (pp. 133-149). Amsterdam: North Holland.
- Luce, P. A. (1986). A computational analysis of uniqueness points in auditory word recognition. *Perception & Psychophysics*, 39, 155-158.
- Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10, 29-63.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18, 1-86.
- 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, E. J., & Norris, D. G. (in press). Models of continuous speech recognition and the structure of the vocabulary. *Language and Cognitive Processes*.
- 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 & Cognition*, 20, 621-638.
- Mehler, J., Dommergues, J.-Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning & Verbal Behavior*, 20, 298-305.
- Mehler, J., Dupoux, E., & Segui, J. (1990). Constraining models of lexical access: The onset of word recognition. In G. Altmann (Ed.) *Cognitive models of speech processing: Psychological and computational perspectives* (pp. 236-262). Cambridge, MA: MIT Press.
- Morgan, J. L. (1994). Converging measures of speech segmentation in preverbal infants. *Infant Behavior and Development*, 17, 389-403.
- Morgan, J. L., & Saffran, J. R. (in press). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development*.
- Norris, D. G. (1991). Rewiring lexical networks on the fly. *Proceedings of EUROSPEECH 91*, Genoa; Vol. 1, 117-120.
- Norris, D. G. (1994). SHORTLIST: A connectionist model of continuous speech recognition. *Cognition*, 52, 189-234.
- Norris, D. G., McQueen, J. M., & Cutler, A. (in press). Competition and segmentation in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory & Cognition*.
- Otake, T. (1992, July). *Morae and syllables in the segmentation of Japanese*. Paper presented at the XXV International Congress of Psychology, Brussels.
- Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory & Language*, 32, 358-378.
- Pallier, C., Sebastian-Galtes, N., Felguera, T., Christophe, A., & Mehler, J. (1993). Attentional allocation within the syllabic structure of spoken words. *Journal of Memory & Language*, 32, 373-389.

- Sebastian-Galles, N., Dupoux, E., Segui, J., & Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory & Language*, 31, 18-32.
- Segui, J., Frauenfelder, U. H., & Mehler, J. (1981). Phoneme monitoring, syllable monitoring and lexical access. *British Journal of Psychology*, 72, 471-477.
- Shillcock, R. C. (1990). Lexical hypotheses in continuous speech. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing* (pp. 24-49). Cambridge, MA: MIT Press.
- 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.
- Zwitserslood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, 32, 25-64.
- Zwitserslood, P., Schriefers, H., Lahiri, A., & van Donselaar, W. (1993). The role of syllables in the perception of spoken Dutch. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 19, 260-271.