

# **Language, Brain, and Cognitive Development**

Essays in Honor of Jacques Mehler

edited by Emmanuel Dupoux

A Bradford Book  
The MIT Press  
Cambridge, Massachusetts  
London, England

**(2001)**

---

## The Roll of the Silly Ball

Anne Cutler, James M. McQueen, Dennis Norris, and A. Somejuan

What is the syllable's role in speech processing? We suggest that the answer to this question championed by Jacques Mehler some twenty years ago (see Mehler, 1981) is wrong. But our purpose is not to flog a dead horse. Rather, we would like to bring Jacques the comforting news that he was less wrong than we might once have thought he was about the syllable's role. On the basis of recent research, it appears that the syllable does have a language-universal part to play in speech processing, just as Jacques argued two decades ago. It is, however, not the role he proposed then.

The experimental study of spoken word recognition is barely three decades old. The syllable played a role in the theoretical debate from the earliest days (see, e.g., Savin and Bever, 1970), and Jacques has been and still is one of the syllable's staunchest advocates. The focus of the debate, as might be expected, has changed somewhat in those thirty years; but the syllable, like its trusty champion, is still an active contender.

### Why the Syllable Is Not the Unit of Perception

Below, we argue that the syllable in fact has two roles to play in speech processing. Crosslinguistic evidence suggests that one of these roles is language-specific, whereas the other is language-universal. But neither corresponds to the position adopted by Mehler (1981), which was that "the syllable is probably the output of the segmenting device operating upon the acoustic signal. The syllable is then used to access the lexicon" (p. 342). On this view, syllables are the "units of perception" which provide the interface between the speech signal and the mental lexicon. This

theory is no longer tenable. In many experiments carried out in the 1970s and early 1980s which compared monitoring times for phonemes and syllables, syllables were always detected faster than phonemes (e.g., see Savin and Bever, 1970; Segui, Frauenfelder, and Mehler, 1981; reviewed in Norris and Cutler, 1988). This might suggest that, in line with the theory that syllables are the mental representations used for lexical access, syllable perception is primary, and that the detection of phonemes depends on the perception of syllables. As pointed out by Foss and Swinney (1973), however, there is a flaw in this argument: the order of identification of units in a laboratory task need not reflect the order in which the corresponding units are processed during speech perception. More generally, there is no logical necessity that the units which listeners detect in laboratory tasks must correspond directly to representations used in normal lexical access (Cutler et al., 1987). Furthermore, the advantage which the syllable showed in identification times was found to be an experimental artifact. Norris and Cutler (1988) asked listeners to monitor for either phoneme or syllable targets in lists of words and nonwords. The materials contained foils: phonemes or syllables that were phonologically very similar to the targets. In contrast to earlier experiments, listeners were therefore obliged to analyze each stimulus fully. Under these conditions, phoneme targets were detected faster than syllable targets.

Stronger evidence for the syllable's role as the "unit of perception" was presented by Mehler et al. (1981). French listeners were faster to detect the sequence BA, for example, in "balance" than in "balcon" (where the target matches the syllabification of ba.lance), and faster to detect BAL in "balcon" than in "balance" (where the target matches the syllabification of bal.con). Although this crossover interaction could be taken as evidence that the speech signal is classified into syllables prior to lexical access (as Mehler, 1981, and Mehler et al., 1981, argued), there is an alternative explanation for this finding which we prefer. We discuss it later. For the moment, let us simply note that neither English nor Dutch listeners, presented with exactly the same French materials, showed this syllabic effect (Cutler, 1997; Cutler et al., 1983, 1986), and that the effect did not replicate when French listeners were asked to detect similar targets in French nonwords (Frauenfelder and Content, 1999). These findings challenge the hypothesis that the syllable is the universal prelexical unit.

A further major problem for this hypothesis is that lexical access appears to be continuous. The rapidity of spoken word recognition (Marslen-Wilson, 1973; Marslen-Wilson and Welsh, 1978) suggests that lexical access begins as soon as a word's first few sounds are available (Marslen-Wilson, 1987); the initiation of access is not delayed until, say, the end of the first syllable. This observation is incompatible with any model of lexical access in which there is a strict classification of the speech signal into syllables which intervenes between low-level acoustic analysis and the mental lexicon. Even more striking is the repeated demonstration that lexical activation changes with fine-grained changes in acoustic-phonetic information in the speech signal. For example, words are recognized more slowly when they contain subphonemic mismatches than when they do not (Marslen-Wilson and Warren, 1994; McQueen, Norris, and Cutler, 1999; Streeter and Nigro, 1979; Whalen, 1984, 1991).

These studies show that the activation of lexical representations is modulated by changes in the speech signal which are not only subsyllabic but also subphonemic. Marslen-Wilson and Warren (1994) therefore argued that prelexical representations are featural. The subphonemic mismatch data, however, are in fact neutral with respect to the size of prelexical representations (McQueen et al., 1999; Norris, McQueen, and Cutler, 2000). It is true that these data rule out strictly serial models, in which information is only sent on to the lexical level in discrete chunks (whether those chunks are phonemic, syllabic, or whatever). If information cascades continuously up to the lexicon, however, subphonemic changes could modulate lexical activation, whatever the size of the units coding that information. The evidence of continuous uptake of information at the lexical level therefore rules out only the strongest form of the syllabic model, that of strict serial classification into syllables prior to lexical access.

Another argument against the syllable as the unit of perception comes from the analysis of the phonological structure of the world's languages. The syllable structure of some languages is very clear; in others it is not. For example, there is no consensus among English native speakers on the syllabification of English (Treiman and Danis, 1988). There is even no consensus on the syllabification of French by French speakers (Content, Dumay, and Frauenfelder, 2000; Frauenfelder and Content, 1999).

Phonologists (e.g., Kahn, 1980) have argued that some consonants in English (such as single intervocalic consonants following a lax vowel and preceding schwa, like the /l/ in *balance*) are ambisyllabic, that is, they belong to two syllables. A model of lexical access based on classification into syllables cannot work in a straightforward way in languages with ambisyllabicity. Unless one were willing to posit that the form of prelexical representations varies from language to language, such that only the speakers of languages with fully regular syllable structure perform lexical access on the basis of syllabic units, ambisyllabicity provides another challenge to the hypothesis that syllables are the units of perception. We believe that there are important crosslinguistic differences in speech processing, driven by differences in the phonological structures of languages, but we think it very unlikely that such differences would be so large that spoken word access would be based on different units in different languages.

The strong position advocated by Mehler (1981) can therefore be rejected. Note, however, that we are unable to rule the syllable out completely as a prelexical unit. The nature of the mental representations which provide the key to lexical access remains to be determined: they could be gestural (Lieberman and Mattingly, 1985), featural (Marslen-Wilson and Warren, 1994), phonemic (Fowler, 1984; Nearey 1997), or even syllabic. What is clear, however, is that if these representations do prove to be syllabic, they will have to operate in such a way that information can cascade continuously from them to lexical representations, and they would have to operate in a way which would deal with ambisyllabicity. In other words, there can be no strict classification of the speech signal into discrete and indivisible syllables before lexical access.

### **How the Syllable Is a Possible Word**

How, then, can we explain the findings of Mehler et al. (1981)? We will do so in the context of something which was not available in the early 1980s: Shortlist, a computational model of spoken word recognition (Norris, 1994; Norris et al., 1997). Shortlist does not provide an answer to the unit-of-perception question. The input in the current implementation of Shortlist is a string of phonemes. The key aspects of the model's

function, however, do not depend on the validity of this assumption. The domain which the model seeks to explain is one step higher in the processing sequence: lexical access itself, and how spoken word recognition is achieved, in particular word recognition in the normal case, continuous speech. This is difficult for two independent reasons. One is that continuous speech lacks reliable cues to word boundaries, so that listeners have to segment the continuous stream of sounds that a speaker produces into the words of the speaker's message. The second reason is the nonuniqueness of most speech: a given word onset is often consistent with many continuations, and many words have other words embedded within them. This further compounds the segmentation problem. Multiple lexical parses of a given input are indeed sometimes possible, for example in the sequence

The Shortlist model has two main properties. The first is that only the words most consistent with the information in the input (the shortlist of best candidates, hence the model's name) are activated. The second is that the activated candidate words compete with each other (via inhibitory connections). Competition between candidate words provides a solution to the segmentation problem. Words like "silly," "balls," and "roll," along with other words like "sill," "syllabub," and "roe" will be activated along with "syllable" when the listener hears /dasihbDlɜrɔl/. The speech signal itself can act to disfavor some of these alternative words (such as the mismatch between "syllabub" and the syllable /bolz/; see, e.g., Connine et al., 1997, and Marslen-Wilson, Moss, and van Halen, 1996, for discussion of the role of mismatching information in lexical access). Lexical competition, however, provides an efficient mechanism by which alternative candidates can be evaluated relative to each other even when there is no clear information in the signal to favor or disfavor particular alternatives (McQueen et al., 1995). The net result of the competition process is a lexical parse which is the optimal interpretation of the input. The parse "The silly balls roll" is better than "The silly balls roe," for example, since the latter parse leaves the final /l/ unaccounted for. The candidate "roe" will lose in its competition with the candidate "roll" since there is more evidence in favor of "roll."

There is more to speech segmentation than lexical competition, however. While it is true that speech is continuous, and lacks reliable cues to

word boundaries, it is not true that the speech signal contains no word-boundary cues. An obvious cue is that provided by silence: before and after an utterance, or in pauses at clause boundaries within an utterance. There is evidence that listeners are sensitive to this cue, and indeed to many others, when they are available in the speech signal (see Norris et al., 1997, for a review).

In Norris et al. (1997), we sought to unify the evidence that listeners use this variety of segmentation cues with a theory of word recognition based on competition. The idea was that the cues indicate the likely locations of word boundaries, and that this information could be used to bias the competition process. What was required, therefore, was a mechanism by which each activated lexical hypothesis could be evaluated relative to the signals provided by the segmentation cues. Clearly, a candidate that is perfectly aligned with a likely word boundary (i.e., a word which begins or ends at the point signaled as a likely boundary cue) should be at an advantage relative to a candidate that is misaligned with a signaled boundary. But what should count as misaligned? The crucial factor in determining the goodness of fit of a lexical parse for a given input sequence is whether the entire input has been accounted for in a plausible way. The alignment constraint we proposed, therefore, was the possible word constraint (PWC). If there is an impossible word between a candidate word and a likely word boundary, then the parse including that word is highly implausible, and the candidate in question is disfavored (in the computational implementation of the PWC in Shortlist, the word's activation level is halved). Since silence at the end of the phrase indicates a likely word boundary after the final /l/, and "1" by itself is not a possible English word, the candidate "roe" can be penalized ("The silly balls roe 1" is a highly implausible utterance). Although, as discussed earlier, competition alone might favor "roll" over "roe," the PWC deals "roe" a deathblow (pushing it well down the ranking of shortlisted candidates).

Norris et al. (1997) presented experimental evidence in support of the PWC. Listeners were asked to spot English words embedded in nonsense sequences. They found it much harder to spot a word such as "apple" in "fapple" than in "vuffapple." According to the PWC, this is because in "fapple" the word "apple" is misaligned with the boundary cued by silence at the beginning of the sequence; between the word and the bound-

ary is a single consonant, "f," which is not a possible English word. In "vuffapple," however, there is a syllable, "vuff," between "apple" and the boundary, and this syllable, though not an actual English word, is a possible English word. The PWC penalty should therefore be applied in the former but not in the latter case. The simulations with Shortlist (also in Norris et al., 1997) showed not only that the PWC is required for the model to be able to fit these data (and other data on segmentation) but also that the model's recognition performance on continuous speech was better with the PWC than without it.

The PWC evaluates activated lexical hypotheses relative to likely word boundary locations no matter how those boundaries are cued. Multiple segmentation cues, when available in the speech stream, can therefore all bias the competition process via the same mechanism. It should be clear that these cues are language-specific. What counts as a permissible phonotactic sequence of phonemes varies from language to language. Vowel harmony, for example, can only provide a segmentation cue in a language with vowel harmony constraints (Suomi, McQueen, and Cutler, 1997). Other cues which must be language-specific are those related to rhythm: different languages have different rhythmic properties. A large body of crosslinguistic research suggests that listeners use the rhythm of their native language in speech segmentation (Cutler and Mehler, 1993; Cutler and Norris, 1988; Otake et al., 1993).

### **Where the Syllable Plays a Language-Specific Role**

With this crosslinguistic perspective, we can now look again at French, and the sequence monitoring results of Mehler et al. (1981). The reason why French listeners are faster to detect BA in "ba.lance" than in "bal.con," and BAL in "bal.con" than in "ba.lance," we suggest, is that in French syllables can serve to assist in speech segmentation for lexical access (see Frauenfelder and Content, 1999, and Content et al., 2000, for a similar suggestion).

The source of the syllable's role in lexical segmentation in French is its role in the rhythmic structure of that language. Likewise, the reason that the syllable apparently plays no role in lexical segmentation by users of some other languages is that rhythm, in those languages, is not syllabically based. In English and Dutch, for instance, rhythm is based on stress;



in Japanese, rhythm is based on a subsyllabic unit called the mora. Rhythmic structure, whatever its (language-specific) nature, provides cues to the likely position of word boundaries, and these cues in turn are exploited by the PWC. Note that these cues need not be deterministic—it does not matter that, for instance, resyllabification in French makes some syllable boundaries not correspond to word boundaries, nor does it matter that some English words begin with unstressed syllables. Lexical access is based on continuous incoming information, and is unaffected by the rhythmic structure. But the rhythmic cues, where they are available, combine with the other available cues to assist the PWC in getting rid of at least some of the activated words.

In other words, this view of lexical processing differs substantially from that taken in the debate on units of perception (see Norris and Cutler, 1985). No strong claim about prelexical access representations is made, nor needs to be made. Information from language rhythm (and from other cues) simply indicates where in the signal likely word boundaries occur. This information is then used by the PWC in the evaluation of lexical hypotheses. The claim that in French syllabic structure marks the locations of likely word boundaries does not depend on any particular assumptions about lexical access units. The reason why BA is easier to detect in *ba.lance* than in *bal.con*, therefore, is not that the unit for /ba/ is extracted prior to lexical access given *ba.lance* (but not given *bal.con*), but that BA is perfectly aligned with the likely word boundary cued by syllable structure before the /l/ in *ba.lance* but misaligned with the likely word boundary cued by syllable structure after the /V/ in *bal.con*.

We prefer this explanation because it has crosslinguistic validity. It explains why speakers of different languages appear to segment their native languages in different ways. What is common across languages is that native language rhythm provides segmentation cues. This account also explains why speakers of different languages also segment non-native languages in different ways. French listeners segment English syllabically, whereas English listeners segment neither English nor French syllabically (Cutler et al., 1983, 1986). Like English listeners, Dutch listeners do not segment French syllabically (Cutler, 1997). French listeners also segment Japanese syllabically, unlike Japanese listeners (Otake et al., 1993). Finally, even French-English bilinguals do not segment their two languages

in the same way. Cutler et al. (1989, 1992) found that some bilinguals segmented French but not English syllabically; the other bilinguals showed no evidence of a syllabic strategy in either language; instead, they segmented English (but not French) using a stress-based strategy based on English rhythm. A theory in which the syllable is the universal "unit of perception" cannot account for this crosslinguistic variability.

### **How the Syllable Also Plays a Language-Universal Role**

Jacques might well be disappointed if he were to decide that the syllable might therefore have only a language-specific role to play in speech processing. But this conclusion would be premature. Our recent research has suggested that the syllable also has a language-universal role.

The PWC depends on multiple simultaneous activation of candidate words, and on cues to the location of likely word boundaries, but it also depends on a third factor: a clear definition of what counts as a "possible word," that is, an acceptable portion of speech between a candidate word and a boundary location. In Norris et al. (1997), consonantal portions, like the "f" in "fapple," were compared with syllabic portions, like the "vuff" in "vuffapple." As described above, the syllable here is clearly a possible word of English; the consonant is not. In the implementation of the PWC, therefore, a word is considered to be an acceptable candidate if the stretch of speech between its edge and a likely word boundary contains a vowel (i.e., if it is a syllable) and is considered to be an implausible candidate if the stretch of speech does not contain a vowel.

This simple implementation of the PWC could work across all languages. But does it? Languages differ with respect to what constitutes minimal possible words. Thus, in English, for example, consonant-vowel (CV) syllables in which the vowel is lax (e.g., /vae/, /te/) are not well-formed words. But in French these syllables are well-formed words ("va" and "te" are in fact both words). If the current implementation of the PWC is correct, CV syllables with lax vowels should be treated as acceptable words in the segmentation of any language. In a word-spotting task, therefore, words should be easier to spot in open-syllable contexts with lax vowels than in single-consonant contexts. If, however, the PWC respects the fact that open syllables with lax vowels violate the

phonological constraints of English, word spotting should be as hard in lax-vowel contexts as in consonantal contexts in English. Norris et al. (2001) tested this prediction. English listeners spotted "canal," for example, more rapidly and accurately in a CV context with a lax vowel (zekanael) than in a consonantal context (skanael).

Another constraint on English words is that weak syllables (containing the reduced vowel schwa) are not well-formed content words. Words from the closed set of functors can have schwa as their only vowel (the, a, of, etc.), but words from the open set of content words all have at least one full vowel. Norris et al. (2000) also tested whether the PWC is sensitive to this constraint on the well-formedness of English words. Listeners could spot targets like "bell" faster in /belfef/ than in /belf/. This difference was equivalent to the difference between syllabic and consonant contexts with the same set of target words in Norris et al. (1997), where, for example, "bell" was easier to spot in contexts with syllables containing full vowels (/belfig/) than in consonantal contexts (/belf/).

Syllables with open lax vowels or with schwa therefore do not appear to be treated by the PWC like single consonants. This suggests that there is no correspondence between what constitutes a possible word in a language according to abstract phonological constraints and what constitutes a viable portion of speech in the computation of acceptable lexical parses during continuous speech recognition. Another demonstration of this comes from a word-spotting experiment in Sesotho, a language spoken in southern Africa. Bantu languages like Sesotho have the phonological constraint that any surface realization of a content word must be minimally bisyllabic. Monosyllables are thus not well-formed words in Sesotho. Cutler, Demuth, and McQueen (submitted) compared Sesotho speakers' ability to spot a word like "alafa" (to prescribe) in "halafa" (where the single consonant context "h" is an impossible word) and "roalafa" (where the monosyllabic context "ro" is not a word, and not even a well-formed Sesotho word). Listeners were able to spot words in the monosyllabic contexts faster and more accurately than in the consonantal contexts. It would appear that the fact that "ro" is not a possible word in the Sesotho vocabulary does not make "ro" unacceptable as part of the parse "ro + alafa."

All of these results point to the idea that, across the world's languages, possible parses of continuous speech consist of chunks no smaller than a syllable. In other words, the segmentation procedure instantiated by the PWC does not vary from language to language, depending on the well-formedness of words in each language. Instead, the PWC mechanism appears to respect a simple language-universal constraint: if the stretch of speech between a likely word boundary and the edge of a candidate word is a syllable, that candidate word is a viable part of the parse; if the stretch of speech is not syllabic (a single consonant or a string of consonants), the candidate word is not part of a plausible parse, so its activation is reduced. Thus, across many languages, including English and Sesotho, listeners find it harder to spot words in nonsyllabic contexts than in syllabic contexts: in Dutch, "lepel" (spoon) is harder to spot in /blepel/ than in /selepel/. McQueen and Cutler (1998) note that weak syllables are also not possible content words in Dutch; and in Japanese, "ari" (ant) is harder to spot in /rari/ than in /eari/ (McQueen, Otake, and Cutler, 2001). The size or nature of the syllable does not appear to matter: any syllable will do.

The syllable therefore does appear to have a central role to play in speech processing. Syllables are not, as Mehler (1981) suggested, units of perception into which listeners classify spoken input prior to lexical access. Instead, the syllable appears to be the measuring stick against which viable and unviable parses of continuous speech are judged. Syllables form acceptable chunks in the ongoing lexical parse of the speech stream; nonsyllabic sequences do not. *La syllabe est morte; vive la syllabe!*

### **Acknowledgments**

The order of authors (C, M, N, S) is not arbitrary. We thank Peter Gipson for suggesting the title to us, in the form of a most memorable reprint request, which long adorned the wall of Jacques's office.

### **References**

Connine, C. M., Titone, D., Deelman, T., and Blasko, D. (1997). Similarity mapping in spoken word recognition. *Journal of Memory and Language*, 37, 463-480.

- Content, A., Dumay, N., and Frauenfelder, U. H. (2000). The role of syllable structure in lexical segmentation: Helping listeners avoid mondegreens. In A. Cutler, J. M. McQueen, and R. Zondervan (Eds.), *Proceedings of the SWAP Workshop* (pp. 39-42). Nijmegen, Netherlands: Max Planck Institute for Psycholinguistics.
- Cutler, A. (1997). The syllable's role in the segmentation of stress languages. *Language and Cognitive Processes*, 12, 839-845.
- Cutler, A., and Mehler, J. (1993). The periodicity bias. *Journal of Phonetics*, 21, 103-108.
- 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.
- Cutler, A., Demuth, K., and McQueen, J. M. (submitted). Universality versus language-specificity in listening to running speech.
- Cutler, A., Mehler, J., Norris, D., and Segui, J. (1983). A language specific comprehension strategy. *Nature*, 304, 159-160.
- Cutler, A., Mehler, J., Norris, D., and Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, 25, 385-400.
- Cutler, A., Mehler, J., Norris, D., and Segui, J. (1987). Phoneme identification and the lexicon. *Cognitive Psychology*, 19, 141-177.
- Cutler, A., Mehler, J., Norris, D., and Segui, J. (1989). Limits on bilingualism. *Nature*, 340, 229-230.
- Cutler, A., Mehler, J., Norris, D., and Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology*, 24, 381-410.
- Foss, D. J., and Swinney, D. A. (1973). On the psychological reality of the phoneme: Perception, identification and consciousness. *Journal of Verbal Learning and Verbal Behavior*, 12, 246-257.
- Fowler, C. A. (1984). Segmentation of coarticulated speech in perception. *Perception and Psychophysics*, 36, 359-368.
- Frauenfelder, U. H., and Content, A. (1999). The role of the syllable in spoken word recognition: Access or segmentation? In *Proceedings of the Second Journées d'Etudes Linguistique* (pp. 1-8). Nantes, France: Université de Nantes.
- Kahn, D. (1980). Syllable-structure specifications in phonological rules. In M. Aronoff and M.-L. Kean (Eds.), *Juncture* (pp. 91-105). Saratoga, CA: Anna Libri.
- Liberman, A. M., and Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1-36.
- Marslen-Wilson, W. D. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, 244, 522-523.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25, 71-102.

- Marslen-Wilson, W., and Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review*, 101, 653-675.
- Marslen-Wilson, W. D., and Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10, 29-63.
- Marslen-Wilson, W., Moss, H. E., and van Halen, S. (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 1376-1392.
- McQueen, J. M., and Cutler, A. (1998). Spotting (different types of) words in (different types of) context. In *Proceedings of the Fifth International Conference on Spoken Language Processing*. Vol. 6 (pp. 2791-2794). Sydney: Australian Speech Science and Technology Association.
- McQueen, J. M., Cutler, A., Briscoe, T., and Norris, D. (1995). Models of continuous speech recognition and the contents of the vocabulary. *Language and Cognitive Processes*, 10, 309-331.
- McQueen, J. M., Norris, D., and Cutler, A. (1999). Lexical influence in phonetic decision making: Evidence from subcategorical mismatches. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1363-1389.
- McQueen, J. M., Otake, T., and Cutler, A. (2001). Rhythmic cues and possible-word constraints in Japanese speech segmentation. *Journal of Memory and Language*, 45, in press.
- Mehler, J. (1981). The role of syllables in speech processing: Infant and adult data. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 295, 333-352.
- Mehler, J., Dommergues, J.-Y., Frauenfelder, U. H., and Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, 20, 298-305.
- Nearey, T. (1997). Speech perception as pattern recognition. *Journal of the Acoustical Society of America*, 101, 3241-3254.
- Norris, D. G. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52, 189-234.
- Norris, D., and Cutler, A. (1985). Juncture detection. *Linguistics*, 23, 689-705.
- Norris, D., and Cutler, A. (1988). The relative accessibility of phonemes and syllables. *Perception and Psychophysics*, 43, 541-550.
- Norris, D., McQueen, J. M., Cutler, A., Butterfield, S. and Kearns, R. (2001). Language-universal constraints on speech segmentation. *Language and Cognitive Processes*, 16, in press.
- Norris, D., McQueen, J. M., Cutler, A., and Butterfield, S. (1997) The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology*, 34, 191-243.

- Norris, D., McQueen, J. M., and Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, 23, 299-325.
- Otake, T., Hatano, G., Cutler, A., and Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, 32, 258-278.
- Savin, H. B., and Bever, T. G. (1970). The non-perceptual reality of the phoneme. *Journal of Verbal Learning and Verbal Behavior*, 9, 295-302.
- Segui, J., Frauenfelder, U., and Mehler, J. (1981). Phoneme monitoring, syllable monitoring and lexical access. *British Journal of Psychology*, 72, 471-477.
- Streeter, L. A., and Nigro, G. N. (1979). The role of medial consonant transitions in word perception. *Journal of the Acoustical Society of America*, 65, 1533-1541.
- Suomi, K., McQueen, J. M., and Cutler, A. (1997). Vowel harmony and speech segmentation in Finnish. *Journal of Memory and Language*, 36, 422-444.
- Treiman, R., and Danis, C. (1988). Syllabification of intervocalic consonants. *Journal of Memory and Language*, 27, 87-104.
- Whalen, D. H. (1984). Subcategorical phonetic mismatches slow phonetic judgments. *Perception & Psychophysics*, 35, 49-64.
- Whalen, D. H. (1991). Subcategorical phonetic mismatches and lexical access. *Perception & Psychophysics*, 50, 351-360.