

Eight questions about spoken-word recognition

James M. McQueen

Max Planck Institute for Psycholinguistics

Nijmegen, The Netherlands

Why?

Why are psycholinguists interested in spoken-word recognition? Imagine a typical listening situation. The phone rings, and you find yourself being addressed by an unknown speaker. How do you come to realize that this person wants to sell you a carriage clock? There may be aspects of the situation and of the speaker's ingratiating conversational style that might help you deduce that this is a cold call, but it is only by recognising the words you hear that you can recover the speaker's full intentions (i.e., that they are selling carriage clocks and not car insurance). Because each sentence that you hear comes from an unlimited set of potential sentences, it would be impossible to derive what speakers mean by trying to recognize their utterances as wholes. But utterances are made from a limited set of words that, for fluent speakers of a language, will usually already be stored in long-term memory. So speakers' messages must be decoded via recognition of their parts.

Successful communication thus depends on word recognition. Since word recognition is at the heart of the language comprehension process, it has also always been a central topic in psycholinguistics. This chapter reviews the available evidence on spoken-word recognition, focussing on what I take to be the key aspect of the process: The way in which the listener derives from a spoken utterance a satisfactory lexical parse (i.e., an estimate of what word forms the speaker intended, in a plausible order). My assumption will be that this process entails abstraction, that is, a type of decoding in which the specific acoustic realization of any given utterance is

mapped onto stored knowledge about the phonological form of individual words. I make this assumption because it is only via recognition of specific tokens in the speech signal as instances of particular lexical types that semantic and grammatical knowledge about those words can be retrieved and used in comprehension.

I also assume that, in normal comprehension, the listener does not necessarily make explicit and categorical decisions about the word forms in each and every stretch of speech that they hear. The process is probably more stochastic. Thus, while certain psycholinguistic tasks may require listeners to make explicit decisions about the words they are hearing, word recognition in normal listening is more continuous; that is, there need be no definitive "magic moment" (cf. Balota, 1990) at which each word form is absolutely identified. Word recognition instead seems to involve the derivation of a number of different possible (sequences of) word forms, weighted by their likelihood of being correct. As we will see, as plausible hypotheses about the word forms in an utterance come available, their grammatical and semantic properties are retrieved, so that possible interpretations of the meaning of that utterance can be built.

What?

So, to begin, what information in the speech signal is used in word recognition? There are, broadly speaking, two classes of information that are extracted from the signal and used in lexical access: segmental information (i.e., that which distinguishes among speech sounds) and suprasegmental information (i.e., that which specifies the prosodic characteristics of words). I discuss each of these information sources in turn below.

Note that the ensuing discussion assumes a prelexical level of processing which acts as the interface between auditory and lexical processing. Scharenborg *et al.* (in press) argue that the

prelexical level is required to solve the “invariance problem” – the problem that the highly variable speech signal must be mapped onto discrete lexical representations. I further motivate the assumption of a speech-specific prelexical level in McQueen (2005); to summarize, “word recognition would benefit if at least part of the speech code could be cracked prelexically” (p. 264). I should mention here, however, that this assumption is controversial in many ways – about the nature of the representations at the prelexical level, about whether the processes operating there are speech specific, about the role of visual cues in speech processing, and even about whether there is a prelexical level (see Pisoni, this volume, Massaro, this volume, Stevens, 2002, and Diehl *et al.*, 2004, for further discussion).

Segmental information specifies which sounds are in an utterance, and hence must be the primary determinant of successful word recognition. It is thus not surprising that computationally implemented models of spoken-word recognition (TRACE, McClelland & Elman, 1986; Shortlist, Norris, 1994; the Distributed Cohort Model, DCM, Gaskell & Marslen-Wilson, 1997; ARTWORD, Grossberg & Myers, 2000; PARSYN, Luce *et al.*, 2000) all assume a prelexical stage of processing in which a representation of the segmental content in the input speech signal is constructed for use in subsequent lexical access. See Gaskell (this volume) for further discussion of models.

Given the indisputable role of segmental information in word recognition, many experiments on this issue have instead focussed on the effect of mismatch between the input and stored lexical knowledge. Marslen-Wilson and Zwitserlood (1989), for example, showed that mismatch on an initial phoneme (e.g., the Dutch word *honing*, honey, changed into *woning*, dwelling) appeared to be enough to block lexical access. In a cross-modal priming task, presentation of a prime such as *honing* facilitated responses to the target *bij*, bee, but the prime *woning* had no such effect. This

result confirms that segmental match is critical for lexical access, but also suggests that the word-recognition process is rather intolerant of any segmental mismatch.

Although this conclusion has, in its essence, stood the test of time, subsequent research has suggested a number of refinements. First, segmental mismatch is more disruptive of lexical access in word-initial than in word-final position (Allopenna *et al.*, 1998). This is because of the temporal nature of speech: A word with a final mispronunciation is, prior to the arrival of that mispronunciation, perfectly consistent with the intended word, but a word with an initial mispronunciation is immediately put at a disadvantage. Second, in spite of the initial perfect match of a word with a late mispronunciation, the mismatching sound, when it arrives, still interferes effectively with the word-recognition process, even in long words (Frauenfelder *et al.*, 2001; Soto-Faraco *et al.*, 2001). Such results suggest that mismatching evidence tends to weigh more heavily against a word than matching evidence weighs in its favour (Norris, 1994). Third, mismatching segments appear to be more disruptive in short than in long words. Thus, for example, there is stronger evidence that words with an initial mismatching phoneme can be accessed when they are polysyllabic (Connine *et al.*, 1993a) than when they are monosyllabic (Gow, 2001). Fourth, lexical neighbourhood plays a role. If the mismatch creates another word (e.g., *buns* becoming *guns*) no evidence that the source word has been accessed is found (Gow, 2001), but when the mismatch creates a nonword (e.g., *cat* becoming *gat*), there is evidence of access of the source word (Milberg *et al.*, 1988). Marslen-Wilson *et al.* (1996) also demonstrate that the presence/absence of similar-sounding words influences the effect of segmental mismatch.

Finally, phonetic distance (the similarity between the intended and mispronounced sounds) plays a critical role in determining the effect of segmental mismatch. The more dissimilar the mismatching sound is to the sound in the word's correct pronunciation, the more disruptive that sound is in

lexical access (Connine *et al.*, 1993a, 1997; Marslen-Wilson *et al.*, 1996, but see also Soto-Faraco *et al.*, 2001, for contradictory results, and Ernestus & Mak, 2004, for evidence that these effects further depend on the informational value of the mispronunciation). These studies all examined mismatches which involved discrete substitutions of one phoneme by another. More subtle effects of phonetic similarity have also been observed, when changes involving less than one phoneme are created. One way to achieve this is to cross splice stimuli (Dahan *et al.*, 2001b; Marslen-Wilson & Warren, 1994; McQueen *et al.*, 1999; Streeter & Nigro, 1979; Whalen, 1984, 1991). Cross splicing the initial consonant and vowel of *shop* with the final consonant of *shock*, for example, produces a stimulus which sounds like *shock*, but which contains (in the vocalic portion) acoustic evidence consistent with a final /p/. The interfering effects of these cross-splicings on lexical access depend on lexical factors (whether the entire sequence is a word or nonword and whether its components derive from words or nonwords). This suggests that the phonetic mismatch created by the cross-splicing influences not only prelexical but also lexical processing.

Disruptions of phonetic detail can also be created without cross splicing. Andruski *et al.* (1994), for example, artificially reduced the Voice Onset Time (VOT) of the initial stop consonants of words such as *pear*. VOT is a major phonetic cue in English to the distinction between voiceless stops (e.g., /p/, with long VOTs) and voiced stops (e.g., /b/, with short VOTs). In a cross-modal priming task, responses to targets (e.g., *fruit*) were faster after *pear* than after an unrelated word. But this priming effect became smaller as VOT was reduced (see also McMurray *et al.*, 2002; Utman *et al.*, 2000). These results suggest again that detailed information about segmental distinctions influences lexical processing, especially since these effects depend on lexical factors (whether the critical sequence with a voiceless or voiced stop is a word or not; van Alphen & McQueen, in press). The influence of this kind of phonetic detail on lexical access also depends on the usefulness of that detail in the making of lexical distinctions (van Alphen & McQueen, in

press): the stronger the phonetic cue (e.g., to whether a stop is voiced or voiceless), the greater the influence of that cue on lexical processing.

Segmental information thus strongly constrains lexical access. The presence of even a small amount of mismatching information (i.e., of even less than one whole phoneme) is enough to disrupt word recognition. But mismatch effects do not depend solely on phonetic similarity; they also depend on variables such as word length, position of the mismatch, lexical factors, and informational value, and on the interaction of all of these variables.

Suprasegmental information also constrains lexical access. This type of information goes beyond the segmental make-up of the speech signal, and specifies instead the prosodic structure of words (e.g., their syllabification and lexical-stress pattern) and the position of a word in the intonational structure of an utterance. The role of lexical-stress information in word recognition has received considerable attention (see Cutler, 2005, for an extensive review). The stress pattern of a word is the way in which its syllables differ in accentuation (compare, e.g., *trusty*, with stress on the first syllable, and *trustee*, with stress on the second syllable), and may be signalled by differences in pitch, duration and amplitude. As in studies on uptake of segmental information, experiments on lexical stress have focussed on the effects of mismatch. Experiments on the recognition of English minimal pairs such as *trusty/trustee* (Cutler, 1986) suggested that the acoustic differences between the members of such pairs did not influence lexical access: Cross-modal priming effects on targets which were associates of both members of the pair were found irrespective of which member of the pair was the preceding prime. Such minimal pairs are very rare, however. More recent experiments have used larger sets of materials consisting of words which begin with the same segments but have different stress patterns (e.g., *admiral/admiration*), and fragment priming (e.g., the prime *admi-* with either primary or secondary stress on the first syllable). These priming

experiments, and related studies using word fragments with other tasks, suggest that the incorrect stress pattern in a word fragment disrupts lexical access in Dutch (Cutler & Donselaar, 2001; Donselaar *et al.*, 2005) and Spanish (Soto-Faraco *et al.*, 2001), but also in English (Cooper *et al.*, 2002). Yet other studies, using a variety of techniques (e.g., Connine *et al.*, 1987; Cutler & Clifton, 1984) also suggest that listeners use lexical-stress information in lexical access.

There are, however, cross-linguistic differences in the role of lexical stress. Not all languages make lexical-stress distinctions. Clearly speakers of such languages cannot use stress information. But they may use other types of suprasegmental information instead. Japanese speakers, for example, appear to use the pitch-accent patterns of Japanese word in lexical access (Cutler & Otake, 1999), and Mandarin speakers use tone information (Fox & Unkefer, 1985). Furthermore, within languages that do have stress distinctions, some have free stress, where stress can be placed in different syllabic positions (e.g., English), and others have fixed stress (e.g., French). Listeners of fixed-stress languages are likely to use stress information less in lexical access, since it is predictable and therefore does not enhance lexical distinctions. Indeed, French listeners, for example, appear not to be sensitive to lexical-stress information (Dupoux *et al.*, 1997). Finally, even within free-stress languages, there is variability in how much value stress information appears to play. Cutler (2005) summarizes evidence that Dutch listeners tend to be more sensitive to stress information than English listeners. This may be because stress information per se has greater informational value in Dutch than in English. Change of lexical-stress pattern may be more often accompanied by a segmental change (reduction of the unstressed vowel to schwa, e.g., compare the first syllables of the noun *conduct* and the verb *conduct*) in English than in Dutch. The conclusion on the uptake of lexical-stress information thus resonates with that on the uptake of segmental information: Stress information is used in lexical access when it is available, but its usage appears to depend on its informational value.

Suprasegmental information which specifies the position of segments in the prosodic hierarchy is also used in word recognition. This type of information includes word-internal properties (the syllabification of segments within words), and properties referring to increasingly larger domains – the prosodic word, the phonological phrase, and the intonational phrase (see, e.g., Shattuck-Hufnagel & Turk, 1996, for review). Information about all these different levels of prosodic structure appears to constrain the lexical access process. In each case, experiments have involved the manipulation of fine-grained phonetic signatures of prosodic structure.

At the syllable level, fragment priming from spoken Italian sequences on responses to visual target words is stronger when the syllabification of the sequence (e.g., *si.l* or *sil.*, where a period indicates a syllable boundary) matches the syllabification of the target (e.g., *si.lenzio*, silence, or *sil.vestre*, sylvan; Tabossi, *et al.*, 2000). These syllabification differences appear to have been signalled by durational differences in the fragments. At the prosodic-word level, differences in the acoustic duration of consonants as a function of whether they are word initial or word internal (Gow & Gordon, 1995) or as a function of whether they are word initial or word final (Shatzman & McQueen, *in press*) modulate word recognition. Once again, it appears that acoustic cues to suprasegmental structure differ in their importance. Thus, although Shatzman and McQueen (*in press*) found that the signal contained several acoustic signatures that could potentially be used by listeners in lexical disambiguation of Dutch sequences such as *een staart / eens taart* (a tail / once tart), duration of the critical consonant (the /s/) appeared to be the most important cue.

There are also durational differences between monosyllabic words (e.g., *cap*) and the same sequence appearing as the initial syllable of a polysyllabic word (e.g., in *captain*). Such differences are used to determine the goodness of fit of the lexical representations of the shorter and longer

word (Davis *et al.*, 2002; Salverda *et al.*, 2003). Salverda *et al.* interpret this as evidence for uptake of information about a prosodic-word boundary (there is such a boundary after, e.g., *cap* in *cap tuck*, but not after the first syllable of *captain*). At the next level up, Christophe *et al.* (2004) found that lexical access in French was faster for the first word in a two-word sequence (e.g., *chat grincheux*, lit. cat grumpy) when there was a phonological phrase boundary between the two words than when the sequence fell within a phonological phrase. Finally, Cho *et al.* (submitted) have shown that the phonetic detail associated with intonational phrase boundaries (specifically, the acoustic correlates of the articulatory strengthening of segments in domain-initial position) is used to modulate word recognition.

The answer to the “what” question therefore appears to be this: The signal contains a great deal of information specifying the segmental and suprasegmental content of an utterance, and listeners appear to extract that information and use it in word recognition, to the extent that it is useful for lexical disambiguation. Stated this way, this conclusion may seem unsurprising. But it is worth bearing in mind for at least two reasons. First, as we will see below, although other constraints are also involved in word recognition, the primary determinant of recognition is the information in the signal itself. Second, the notion that the weight assigned to acoustic evidence in word recognition depends on its informational value may be critical in understanding why there can be differences in the apparent effects of different types of evidence (e.g., in comparisons across languages). Thus, while the speech signal is rich in information, some aspects of the signal are more important in word recognition than other aspects.

Where?

Where are the words in the continuous speech stream? The recognition process, at least when confronted with a multi-word utterance, must not only determine which words are in that input, but

also where they begin and end. This is because there are no fully reliable cues to the location of word boundaries in continuous speech (Lehiste, 1972; Nakatani & Dukes, 1977). Spoken-word recognition in normal listening therefore entails segmentation of a quasi-continuous signal into a discrete lexical parse.

The primary source of evidence that is used for solving this segmentation problem is the speech signal itself. Although there is no fully reliable word-boundary cue in speech – no auditory equivalent of the white spaces between words in a written English text like this – there are many less reliable cues, and listeners appear to use them when they are available. We have already seen one such set of cues: Suprasegmental information signalling prosodic structure helps to solve the segmentation problem. Thus, for example, durational evidence favouring *captain* over *cap* (Davis *et al.*, 2002; Salverda *et al.*, 2003) simultaneously signals that there is no word boundary after the /p/. Likewise, phonological-phrase boundary information can help French listeners segment *chat grinchaux* and reject a parse with no word boundary between the first two syllables (*chagrin ...*; Christophe *et al.*, 2004). The same kind of segmentation-based story can also be told for the other studies on prosodic structure reviewed above.

The speech signal contains other suprasegmental information that is used in segmentation. In particular, metrical structure provides evidence on where word boundaries might be. English listeners appear to be sensitive to the rhythmic distinction between strong syllables (those with full vowels) and weak syllables (those with reduced vowels such as schwa). They appear to have picked up on the fact that (content) words in English tend to begin with strong syllables (Cutler & Carter, 1987; Cutler & McQueen, 1995). Thus, when asked to spot real words in nonsense bisyllabic sequences, English listeners find it more difficult to do so in sequences with two strong

syllables than in strong-weak sequences (Cutler & Norris, 1988; Norris *et al.*, 1995; see also Cutler & Butterfield, 1992).

There are important crosslinguistic differences to consider here, however. Thus, Dutch listeners also show sensitivity to the metrical strong-weak difference (Vroomen & de Gelder, 1995; Vroomen *et al.*, 1996), but listeners of French and Japanese obviously cannot, because these languages do not make this metrical distinction. What appears to be true across languages, however, is the use of rhythmic information in segmentation. Thus, because Romance languages such as French, Catalan, Spanish and Italian have syllable-based rhythm, speakers of these languages appear to use syllabic information in segmentation (Cutler *et al.*, 1986; Pallier *et al.*, 1993; Sebastián-Gallés *et al.*, 1992, Tabossi *et al.*, 2000; but see also Content *et al.*, 2001). Similarly, because the rhythm of Japanese is based on the mora (a sub-syllabic structure), Japanese listeners use moraic information in segmentation (Cutler & Otake, 1994; Otake *et al.*, 1993; McQueen *et al.*, 2001).

Listeners also use phonological knowledge in segmentation. Knowledge about the phonotactic restrictions on syllable structure in a language (e.g., that the sequence /mr/ in English cannot occur within a syllable) could be used to indicate the location of likely word boundaries (e.g., between the /m/ and the /r/). Listeners appear to use this kind of absolute phonotactic knowledge in segmentation (Dumay *et al.*, 2002; McQueen, 1998; Warner *et al.*, 2005a; Weber & Cutler, submitted). Probabilistic phonotactic knowledge (i.e., knowledge that sequences of sounds vary in how likely they are to occur at a word boundary) is also used (van der Lugt, 2001). Vowel harmony (e.g., in Finnish; Suomi *et al.*, 1997) provides another source of word-boundary information. In Finnish, there are restrictions on which vowels can co-occur within a word; there are effectively two distinct sets of vowels that never both occur within the same word. Listeners

appear to have learned to use the knowledge that, if a sequence of speech contains vowels from these two sets, there must be a word boundary between those vowels (see also Vroomen *et al.*, 1998).

There is, however, a very different way in which the segmentation problem is solved: It appears that the manner in which the word-recognition process works provides a means of finding where the words are. As we have already seen in the discussion of the effects of mismatching information, multiple lexical hypotheses appear to be considered simultaneously as speech is heard. Thus, to take just one previous example, words which differ in their onsets or their offsets from the word in the input (e.g., *speaker* and *beetle* given *beaker*, Allopenna *et al.*, 1998) are considered in parallel with what ultimately proves to be the correct hypothesis. A common way of thinking about this process is in terms of activation – a concept derived in large part from Morton's (1969) logogen model. Representations of word forms that are consistent with the information in the current input are said to be activated, with their activation level reflecting their goodness of fit.

The concept of multiple lexical activation is supported by a large body of other evidence. Words beginning in the same way as other words are jointly considered (e.g., in Dutch, *kapitaal*, capital, when the onset of *kapitein*, captain, is heard, Zwitserlood, 1989; see also Moss *et al.*, 1997; Zwitserlood & Schriefers, 1995), as are onset-embedded words, as discussed previously. Words ending in the same way as other words are also activated when the longer word is heard (e.g., offset-embedded words such as *bone* in *trombone*; Isel & Bacri, 1999; Luce & Cluff, 1998; Shillcock, 1990; Vroomen & de Gelder, 1997), though the evidence is weaker than for onset embeddings (see, e.g., Luce & Lyons, 1999; Shatzman & McQueen, submitted-a). Furthermore, words straddling word boundaries in the signal (e.g., *tulips*, given the input *two lips*) also appear to be activated (Gow & Gordon, 1995; Tabossi *et al.*, 1995).

In addition to this evidence on multiple activation, there is also evidence for a form of competition among the activated candidate words. Thus, as the number and frequency of similar-sounding words increases, word recognition becomes harder (Cluff & Luce, 1990; Luce & Large, 2001; Luce & Pisoni, 1998; Vitevitch, 2002; Vitevitch & Luce, 1998, 1999). It appears that, as the number of words in the lexical neighbourhood increases, competition gets fiercer, and recognition is delayed. Gaskell and Marslen-Wilson (2002) show that the size and nature of priming effects arising from word or word-fragment primes depends on the number of words beginning in the same way as those primes. They interpret this as evidence for competition between the prime words and the co-activated words beginning in the same way. Words starting at different points in time also appear to compete. For example, listeners find it harder to spot the real word *mess* in a nonsense sequence which is the onset of a real word (e.g., *domes*) than in one which is not (e.g., *nemess*; McQueen *et al.*, 1994). The increased difficulty in the former case suggests competition between the two words (*mess* and *domestic*). The number of words beginning at a different point in the signal from the target word also influences target recognition (Norris *et al.*, 1995; Vroomen & de Gelder, 1995), again suggesting a competition process. Further evidence for competition comes from the research on mismatch reviewed earlier. With respect to segmental mismatch, the evidence that the influence of mismatching information depends on lexical factors suggests competition among activated lexical alternatives. With respect to suprasegmental mismatch, it appears that mismatching lexical-stress information produces inhibitory priming effects when the mismatch is consistent with another word (Donselaar *et al.*, 2005) but not when it is not consistent with another word (Cutler & Donselaar, 2001). This is presumably because the competition process is stronger in the former case.

We can now return to the segmentation problem. The process of multiple activation of lexical hypotheses, and in particular their relative evaluation through competition, provides a means to solve this problem. If word hypotheses have to fight for control of their parts of the input, then the result will tend to be a lexical parse with each part of the input accounted for by only one strongly activated hypothesis, and no parts left over. Competition thus finds word boundaries even when there are no cues to those boundaries in the speech signal (Norris, 1994, provides computational simulations confirming this behaviour in Shortlist, a model instantiating lexical competition).

Listeners therefore appear to work out where the words are in spoken utterances in three ways: they use information in the speech signal which specifies the location of likely word boundaries, they use phonological knowledge to assist in this process, and they rely on a process of competitive evaluation of multiple lexical hypotheses. There are two open issues. The first concerns the relative importance of these different sources of information. A start has been made to address this. Mattys *et al.* (in press) have recently argued, on the basis of experiments directly contrasting knowledge-driven and signal-driven cues for segmentation in English, that lexical cues (e.g., whether the context of a target word was a word or a nonword) tend to outweigh signal-based segmental cues (e.g., whether segments were coarticulated with their contexts), which in turn tend to outweigh signal-based suprasegmental cues (e.g. whether the stimuli began with strong or weak syllables; see also Mattys, 2004).

The second issue concerns how these different sources of information are integrated. Norris *et al.* (1997) have proposed an algorithm by which signal- and knowledge-based cues to likely word boundaries influence lexical competition. The idea is that activated candidate words are evaluated with respect to whether they are aligned with the boundaries signalled by, for example, metrical structure. If those words are found to be misaligned with those boundaries, their activation is

penalized. A word counts as being misaligned if there is no vocalic portion between the edge of the word (its beginning or its end) and the signalled likely word boundary location. The reason for this definition of misalignment is that a sequence without a vowel cannot be a possible word. A parse involving, for example, a candidate word followed by a vowelless residue and then a likely word boundary, is very improbable. Evidence for this simple vowel-based Possible Word Constraint (PWC) has now been found in a range of languages including English (Norris *et al.*, 1997, 2001), Dutch (McQueen, 1998), Japanese (McQueen *et al.*, 2001) and Sesotho (Cutler *et al.*, 2002). As has been argued in greater detail elsewhere (Cho *et al.*, submitted), it is however unlikely that the PWC penalty is the only mechanism by which word boundary cues modulate the competition process. Embedded words such as *cap* in *captain* are not misaligned with any likely word boundary, and yet their strength as competitors does appear to vary as a function of signal-based information (i.e., their duration; Davis *et al.*, 2002, Salverda *et al.*, 2003). Competition may therefore be modulated by boosts for aligned candidates as well as by penalties for misaligned candidates (see also Norris *et al.*, 1995).

Which?

Which words did the speaker intend? This question lies at the heart of the spoken-word recognition problem; it is the question the listener must answer. We already have a partial picture about how the listener does just that. Words that are consistent with the current input are activated, as a function of how good a match there is between the input and stored phonological knowledge about those words, and they compete with each other for recognition. Segmental and suprasegmental information in the signal modulates this competition, by indicating which words are in the input and where they begin and end.

But there is more to it than that. Phonological knowledge, beyond that about the words themselves, also plays an important role. We have already seen that word recognition is influenced by language-specific segment sequence constraints (e.g., those due to phonotactic and vowel-harmony restrictions). Knowledge about the alterations to the signal that are the result of phonological processes is also brought to bear. One such process is place assimilation. In English, for example, a coronal consonant such as /t/ at the end of a word such as *night* can sound like a bilabial /p/ when the following consonant has a bilabial place of articulation (e.g., in *night bus*). The evidence reviewed above suggests that *nipe* would be a poor match to *night*, and indeed, when presented in isolation, such assimilated forms do not strongly activate their source words (Marslen-Wilson *et al.*, 1995). Furthermore, when the assimilation process creates another word (e.g., *right* becoming *ripe*) but the following context that caused that assimilation is not presented, the altered input activates both words (*ripe* and *right*; Gow, 2002). A number of studies, however, have shown that when following context is present, and that context licenses the assimilation, the altered word is recognised correctly (Coenen *et al.*, 2001; Gaskell & Marslen-Wilson, 1996, 1998, 2001; Gow, 2001, 2002; Mitterer & Blomert, 2003).

These results suggest that language-specific phonological knowledge about place assimilation is being used in word recognition, as it were, to undo the effects that the assimilation process had during speech production. Two comments on this conclusion need to be made, however. First, the assimilation process tends to be phonetically incomplete (e.g., the final consonant of the first word in *night bus* is not identical to a natural word-final /p/; it has phonetic features consistent with both a /p/ and a /t/; Gow, 2002). Listeners are sensitive to this fine-grained information, as they are to many other types of acoustic detail (as discussed earlier), and they use it in word recognition (Gow, 2002). Second, this kind of fine detail may be processed by low-level perceptual mechanisms and so recovery from at least some kinds of assimilatory processes may not depend on language-

specific knowledge. Thus, listeners who do not speak Hungarian nonetheless show similar sensitivity as native Hungarians to the effects of Hungarian liquid assimilation (Mitterer *et al.*, in press) and Hungarian voicing assimilation (Gow & Im, 2004). Other studies comparing native and nonnative listeners on their sensitivity to assimilation phenomena, however, have found effects of language-specific knowledge (Otake *et al.*, 1996; Weber, 2001). It is therefore probably the case that recovery from the effects of assimilation in word recognition will depend in some cases on language-universal perceptual mechanisms and in other cases on language-specific phonological knowledge.

Other types of phonological process which alter the realization of words in the speech signal and which have been examined for their effects on word recognition include resyllabification (Gaskell *et al.*, 2002; Vroomen & de Gelder, 1999), liaison (the combination of resyllabification and surfacing of latent consonants in e.g. French; Gaskell *et al.*, 2002; Spinelli *et al.*, 2003), neutralization (e.g., the realization of intervocalic stops in American English as flaps; Connine, 2004; McLennan *et al.*, 2003), reduction (the deletion of single or multiple segments from words; Ernestus *et al.*, 2002; LoCasto & Connine, 2002; Mitterer & Ernestus, in press; Utman *et al.*, 2000), and epenthesis (e.g., the insertion of the vowel schwa into the canonical form of a Dutch word; Donselaar *et al.*, 1999). Space restrictions prevent detailed discussion of these phenomena. It is worth noting, however, that, as with assimilation, word recognition in the context of such processes entails a combination of phonological knowledge (e.g., knowledge about where epenthesis is legal; Donselaar *et al.*, 1999), and use of fine-grained phonetic detail to help recover the speaker's intentions (e.g., in the resolution of ambiguities between *dernier oignon*, last onion [with liaison], and *dernier rognon*, last kidney [without liaison]; Spinelli *et al.*, 2003).

A key issue concerning all of these phonological processes, including assimilation, is the nature of the form-based lexical representations involved. Are citation forms as well as their variants (e.g., *pretty* with a [t] and *pretty* with a medial flap) stored in the mental lexicon? If so, then variant pronunciations could of course be recognized via retrieval of those forms. If not, they would have to be recognized through some kind of phonological inference that would map them onto their citation forms. The jury is still out on this issue. It may be that its conclusions will vary for different types of phonological process. Thus, for example, while the evidence suggests that assimilated forms are not stored in the mental lexicon (if they were, then, counterfactually, their recognition would not depend on following context), the evidence on neutralization suggests that both flapped and unflapped forms are stored (Connine, 2004; McLennan *et al.*, 2003). Frequency of occurrence of pronunciation variants may determine which forms are stored (Connine, 2004): Words such as *pretty* occurred in their flapped variants in 96% of tokens in a corpus of American conversation (Patterson & Connine, 2001).

Frequency of occurrence is certainly another constraint that is used by listeners during word recognition. The work by Luce and colleagues discussed earlier on the effects of lexical neighbourhood has shown that it is not just the number of similar-sounding words that determines ease of recognition, but also their frequency of occurrence. Other research with a variety of experimental paradigms also suggests that word frequency influences lexical activation (Connine *et al.*, 1990, 1993b; Dahan *et al.*, 2001a).

Listeners use contextual information to determine which words speakers intend. Swinney (1979), in a cross-modal associative-priming study, showed that both meanings of an ambiguous word (e.g., the insect and spying meanings of *bug*) were activated at the offset of that word, irrespective of whether the preceding context biased interpretation of the word in one or other direction, but that

shortly thereafter only the contextually appropriate meaning was still active. Similarly, Zwitserlood (1989) showed that while the speech signal was consistent with two Dutch words (e.g., *kapitein* and *kapitaal*, at the /p/ in *kapitein*), meanings associated with both words were activated, even in a strongly biasing context, but that as the speech signal unfolded, context influenced lexical activation. In strongly biasing contexts, priming effects were stronger for the appropriate meaning even before the signal provided disambiguating information (e.g., at the /t/ in *kapitein*), but, in more weakly biasing contexts, both meanings remained active until after the signal provided disambiguation. These classic studies suggest that contextual information is not used to determine which words are considered for recognition, but is used rapidly thereafter to select among the set of activated candidates.

Many other studies support this conclusion. Thus, results from a wide range of experimental paradigms suggests that multiple senses of ambiguous words are simultaneously active, even in contexts where semantic or syntactic constraints could bias interpretation in favour of one sense (Blutner & Sommer, 1988; Conrad, 1974; Lackner & Garrett, 1972; Lucas, 1987; Oden & Spira, 1983; Onifer & Swinney, 1981; Seidenberg *et al.*, 1982; Tanenhaus & Donenwerth-Nolan, 1984; Tanenhaus *et al.*, 1979; Whitney *et al.*, 1985; see Lupker, this volume, for further discussion). But several semantic-priming studies have indicated that context can bias meaning activation (most strongly when an ambiguous word has a dominant and a subordinate meaning, and the context favours the dominant meaning; e.g., Moss & Marslen-Wilson, 1993; Simpson, 1981; Tabossi, 1988a, 1988b, Tabossi & Zardon, 1993).

A critical distinction that needs to be made here, however, is that between representations of word form (phonological representations) and representations of word meaning (conceptual representations). Differences in identity (i.e., form-based) priming and associative (i.e., meaning-

based) priming on the same sets of materials (Gaskell & Marslen-Wilson, 2002; Norris *et al.*, submitted) or between identity priming and eye-tracking data (Shatzman & McQueen, submitted-a) support this distinction. Indeed, some type of form/meaning distinction must be made: Conceptual (and grammatical) knowledge must be stored so that interpretations of the meaning of utterances can be built, but that knowledge can only be accessed on the basis of phonological information. If one accepts this distinction, then the evidence from semantic priming on contextual biases in meaning activation does not speak to whether form-based representations were activated, because it is then possible that activation of phonological representations could occur without activation of conceptual representations. The data showing activation of multiple meanings in spite of contextual biases, however, do necessarily imply activation of form-based representations. The data on ambiguous words are thus consistent with the view that context does not determine which word forms are considered in the recognition process, but does influence selection, certainly among conceptual representations.

Research on the influence of sentential context on phonetic decision-making (van Alphen & McQueen, 2001; Borsky *et al.*, 1998; Connine, 1987; Connine *et al.*, 1991; Miller *et al.*, 1984; Samuel, 1981) suggests that context can act as a bias on decision making but does not influence prelexical processing. In the gating task (Grosjean, 1980), where listeners are asked to identify words on the basis of incremental fragments of those words (“gates”), listeners produce contextually inappropriate responses, primarily at earlier gates (Tyler, 1984; Tyler & Wessels, 1983). Both these sets of data once again suggest that spoken-word recognition is based on what has been termed the principle of bottom-up priority (Marslen-Wilson, 1987; Marslen-Wilson & Tyler, 1980): The signal is the primary means by which listeners recover speakers’ intentions, and context plays a secondary (but nonetheless strong and rapid) role. Data from a recent eye-tracking study are consistent with this view. Dahan and Tanenhaus (2004) show that verb-based thematic

constraints have a powerful influence on sentence interpretation. Thus, when Dutch listeners heard a sentence onset such as *Nog nooit krom een bok...* (Never before climbed a goat...), they looked at pictures in a concurrent visual display of a goat, but not of a bone (*bot*), presumably because bones are inanimate and thus cannot climb. But when phonetic information favouring *bot* was inserted (by splicing the initial consonant and vowel from *bot* into *bok*), listeners did look at the bone. Phonetic evidence thus, at least temporarily, overrode the contextual bias.

When?

The speech signal is temporal in nature, and thus it is important to ask when, as that signal unfolds over time, the phonological forms of words are recognized. Marslen-Wilson (1987) estimates, on the basis of data from a variety of tasks, that word recognition occurs about 200 ms after word onset. This is of course an estimate of average recognition time: Some words can be recognized very early, but others are recognized only after their acoustic offset (Bard *et al.*, 1988; Grosjean, 1985). The temporal structure of speech certainly imposes strong constraints on recognition. Several lines of evidence discussed earlier support this conclusion. We saw that mismatching information in word-initial position tends to have a stronger inhibitory effect on word-form activation than word-final mismatch, because words with initial mispronunciations have to recover from a poor start, while words with final mispronunciations can be highly activated before the mismatching material arrives. A similar argument can be made to explain why the phonological representations of onset-embedded words (e.g., *cap* in *captain*) appear to be more strongly activated than those of offset-embedded words (e.g., *bone* in *trombone*): The longer words have a greater advantage over offset- than over onset-embeddings (Luce & Lyons, 1999; Shatzman & McQueen, submitted-a). It was also suggested earlier that the recognition process is rather intolerant of mismatching information. The idea that evidence inconsistent with a lexical hypothesis may weigh more heavily than evidence supporting that word is consistent with the view

that the process of word recognition entails a continuous form of optimization. As material inconsistent with a given lexical hypothesis appears, it is rapidly used to disfavour that word in the form-based competition process.

A critical factor determining when a word can be recognized is its Uniqueness Point (UP) – the point as one moves left-to-right through a word at which the information in the signal uniquely specifies that word. Many short words do not become unique before they end (Luce, 1986). It is these words that tend not to be recognized until after their offset. But the UP in longer words is often before offset. For such words strong relationships have been found between UPs and various measures of recognition time. These measures include phoneme-monitoring latency to phonemes varying in serial position (Marslen-Wilson, 1984), recognition point in gating (Marslen-Wilson, 1987; Tyler & Wessels, 1983; recognition point is an operational definition of when subjects can identify a target word in the gating task correctly and confidently; Grosjean, 1980), shadowing latency (i.e., speed to repeat spoken words; Radeau & Morais, 1990) and gender decision times (e.g., deciding whether French nouns are masculine or feminine; Radeau *et al.*, 1989). The UP effect, at least in the latter two tasks, may however depend on speaking rate: The effect in both shadowing and gender decision tends to be larger at slower speaking rates (Radeau *et al.*, 2000). Furthermore, we have already seen that sentential context can influence the recognition process prior to a word's UP (e.g., Zwitserlood, 1989). Confirming evidence comes from electrophysiological studies (van Berkum *et al.*, 2003; van den Brink *et al.*, 2001; Van Petten *et al.*, 1999) which have shown that, before a word's UP has been reached, event-related brain potentials vary depending on the contextual appropriateness of that word.

It is thus not the case that timing of word recognition is determined completely by when words become unique. Nevertheless, it is clear that word recognition is strongly influenced by the two

factors underlying the UP concept: the information in the speech signal, and when that information is taken up. Analysis of how the pattern of responses in the gating task changes as more of the speech signal is heard also suggests that acoustic detail is taken up rapidly and continuously (Lahiri & Marslen-Wilson, 1991; Marslen-Wilson & Warren, 1994; McQueen *et al.*, 1999; Smits *et al.*, 2003; Warner *et al.*, 2005b; Warren & Marslen-Wilson, 1987, 1988). Analyses of how the lexical activation pattern changes over time, using priming paradigms (e.g., Davis *et al.*, 2002; Gaskell & Marslen-Wilson, 2002; Zwitserlood, 1989), eye-tracking paradigms (e.g., Allopenna *et al.*, 1998; Dahan *et al.*, 2001a,b; Salverda *et al.*, 2003), and, most recently, the tracking of hand movements directing a computer mouse towards a display on a computer screen as spoken words are heard (Spivey *et al.*, 2005), all confirm that there is continuous modulation of the lexical competition process as the speech signal unfolds.

This kind of data thus suggests that word recognition tends to be as early as it can be, given available constraints. As I argued at the outset, however, there may be no magic moment at which a word's phonological form is definitively recognized. Psycholinguistic tasks which require explicit judgements about what words have just been said provide discrete estimates of recognition time. While it is critical to consider the extent to which these measures reflect task-specific processing, it is perhaps just as critical to bear in mind that there may be no equivalent of these explicit judgements, and thus no instantaneous "recognition time", in normal speech comprehension.

There are two reasons why it appears to be wrong to think of the recognition of phonological word forms as a serial and categorical process. The first is the evidence already reviewed on activation of lexical conceptual representations. Much of the evidence on the activation of multiple lexical candidates used the cross-modal associative priming task (e.g., Tabossi *et al.*, 1995; Shillcock,

1990; Zwitserlood, 1989) or measures of eye movements to pictures of objects mentioned in spoken instructions, or pictures of their phonological competitors (e.g., Allopenna *et al.*, 1998; Salverda *et al.*, 2003). The effects measured with both these paradigms indicate that there is rapid spread of information to the conceptual level of processing, and, critically, that this information spreads before a unique lexical form has been identified. Secondly, there are cases of form-based ambiguities which the signal might never be able to resolve (e.g., in oronyms such as *tulips/two lips*; Gow & Gordon, 1995). Although we have seen that there are subtle acoustic cues which help solve these ambiguities, it appears that at least in some cases (e.g., Spinelli *et al.*, 2003) these cues are not powerful enough to resolve them fully. In these cases, alternative form-based parses must be passed forward for resolution (using contextual information) at higher levels of processing. These arguments thus suggest that word-form recognition is probabilistic and incremental. The recognition system tends to settle on one most likely lexical parse of the phonological word forms in an utterance, but does not always do so. It tends to settle fast, but information is passed continuously to processes responsible for deriving an utterance interpretation.

How?

How, then, are words recognised? I will answer this question, and summarize the previous review, by discussing the representations and the processes that appear to be involved in word recognition. I have described three representational types: prelexical, word-form and word-meaning representations. Though this three-way distinction is not uncontroversial, it does tend to be made (albeit in different ways) in models of spoken-word recognition (Gaskell, this volume). More detailed discussion about the nature of prelexical and lexical-conceptual representations, and about the way in which morphologically complex words might be represented at the form and meaning levels, can be found in other chapters in this volume.

The focus has instead been on the recognition of the phonological form of words. But I have said little about the nature of form-based representations. This is largely because this issue is far from resolved. Perhaps words are represented only in their canonical pronunciations, or perhaps multiple pronunciation variants are stored. As already noted, the answer to this question is likely to depend on the frequency of occurrence of different pronunciations. The content of form representations (however many there are for each word) is also not yet resolved. Content may be very restricted (to the minimal abstract phonological specifications required to derive the word's pronunciation; Lahiri & Marslen-Wilson, 1991). Alternatively, it may be more fully specified but still phonologically abstract (e.g., strings of phonemes in TRACE, McClelland & Elman, 1986, and Shortlist, Norris, 1994), or richly specified, including speaker- and situation-specific detail (e.g., in a model where particular episodes of words that the individual listener has heard are stored, Goldinger, 1998). Furthermore, phonological knowledge might be stored in localist representations (e.g., as in TRACE and Shortlist) or in a distributed fashion (e.g., as in the DCM, Gaskell & Marslen-Wilson, 1997, 2002).

The picture is much clearer about the processes involved in word recognition. We have seen data showing that form-based recognition entails the parallel evaluation of multiple lexical hypotheses and a process of competition among those hypotheses. It appears that processing is cascaded (McClelland, 1979), both from the prelexical level to the word-form level, and from word forms to word meanings. Processing is cascaded with respect to information flow and with respect to time. Thus, the evidence that fine phonetic detail influences lexical processing shows that there is continuous flow of information from the prelexical level to the word-form level (e.g., McQueen *et al.*, 1999), and likewise the evidence of activation of lexical meaning before the speech signal can uniquely specify a word's identity (e.g., Zwitserlood, 1989) shows that there is graded flow of information from form to meaning representations. This means that, in the temporal dimension,

processing is incremental: As the speech signal unfolds, lexical hypotheses are continually updated, leading usually, but not always, to only one very probable lexical parse of the input utterance.

We have also seen that word recognition involves the evaluation of multiple information sources: segmental and suprasegmental information in the speech signal (modulated by its usefulness), frequency of occurrence biases, phonological knowledge, and contextual constraints. Particular algorithms have been proposed for how these sources of information are integrated. Lexical competition is the primary algorithm, but the PWC (Norris *et al.*, 1997) has been proposed as a means by which cues to likely word boundaries can modulate the competition process. It is possible that the competition process is influenced only by positive information in the input (i.e., through bottom-up facilitation), but an additional algorithm of bottom-up inhibition (Norris, 1994) may be the means by which mismatching information has a stronger effect on lexical activation than matching information. Finally, the bottom-up priority restriction imposes strong constraints on the recognition process: Only information in the signal can determine which word-form representations are considered.

One final “how” question that has not yet been touched on is that concerning feedback: In addition to bottom-up flow of information from the prelexical level to the word-form level, is there also top-down information flow back to the prelexical level? Norris *et al.* (2000) review the large literature on lexical involvement in phonetic decision-making. They argue that all of these effects can be explained without feedback, and that data from Pitt and McQueen (1998) suggest that there is no feedback. The debate has continued since then, however, with arguments for feedback (Magnuson *et al.*, 2003a; Mirman *et al.*, 2005; Samuel, 2001; Samuel & Pitt, 2003) and against it (McQueen, 2003). A factor that undoubtedly will be involved in resolving this debate is the plasticity of the speech-recognition system. Norris *et al.* (2000) argued that feedback could be of no benefit to on-

line word recognition (passing lexical decisions back to the prelexical level will not improve those lexical decisions). But feedback could be of benefit in perceptual learning. Norris *et al.* (2003) indeed show that listeners use lexical knowledge to adjust their interpretation of a speech sound that was spoken in an unusual way. Since it thus appears that there is feedback for learning, the question for future research will be whether apparent demonstrations of feedback in on-line processing (i.e., feedback as a word is being heard) are in fact the result of longer-term learning effects, or are indeed true on-line effects that might arise epiphenomenally, that is, as a consequence of the need for feedback for perceptual learning.

Whither?

Whither spoken-word recognition? I end with a few more remarks on future directions in word recognition. The flexibility of the recognition system will need to be considered more fully, both with respect to prelexical processing, as we have just seen, and with respect to lexical processing. Researchers have recently been asking how new words are learned and integrated into the word-recognition system (Gaskell & Dumay, 2003; Magnuson *et al.*, 2003b), and how prior phonological knowledge constrains the recognition of novel words (Shatzman & McQueen, submitted-b). A special case of word learning is in second-language acquisition (Dijkstra, this volume): How do listeners learn the words of a new language, and what consequences does this have for lexical representation and process? One direction research will therefore undoubtedly take in the next few years will be the development of more dynamic accounts of spoken-word recognition.

Current investigations using the Norris *et al.* (2003) perceptual learning paradigm suggest that, at least under some circumstances, detail about how a specific speaker makes a phonetic contrast is stored by listeners (Eisner & McQueen, 2005; Kraljic & Samuel, in press). These findings are consistent with other results showing talker-specific effects in memory for words and in speech

processing (reviewed in Goldinger, 1998; Eisner & McQueen, 2005). It is not yet clear, however, where or how talker-specific detail (and other episodic detail) is stored; most models of spoken-word recognition have been silent on this issue (see McLennan & Luce, 2005, for recent discussion). Spoken-word recognition research will therefore need to address how specificity effects can be reconciled with the need for phonological abstraction (i.e., recognition of specific tokens in the signal as instances of particular lexical types).

The review of uptake of fine-grained segmental and suprasegmental information made clear that the speech signal is not just a sequence of phonemes. Prelexical processing appears to involve the extraction of a segmental representation of an utterance, but this representation is not sufficient for word recognition. Prelexical processing also seems to entail the extraction of rich sources of information specifying suprasegmental structure. Current models of spoken-word recognition do not fully specify how this might be done. It is to be hoped that this situation will change in the future.

Pause for thought should you ever be pestered by a carriage-clock telesales person. A complex process of continuous competitive evaluation of candidate word forms will ensue, the inner workings of which you, as a listener, will largely be unaware. Nevertheless, you will probably settle rather quickly on a unique lexical parse of the word forms in the sales spiel. That is why, in spite of the fact that you have never heard the speaker before and never previously been bothered in this way, you will be able to slam the phone down rapidly and confidently. Unless of course your cold-call or time-keeper predilections are different from mine.

Who?

Allopenna PD, Magnuson JS and Tanenhaus MK (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, **38**, 419-39.

Alphen P van and McQueen JM (2001). The time-limited influence of sentential context on function word identification. *Journal of Experimental Psychology: Human Perception and Performance*, **27**, 1057-71.

Alphen P van and McQueen JM (in press). The effect of Voice Onset Time differences on lexical access in Dutch. *Journal of Experimental Psychology: Human Perception and Performance*.

Andruski JE, Blumstein SE and Burton M (1994). The effect of subphonetic differences on lexical access. *Cognition*, **52**, 163-87.

Balota DA (1990). The role of meaning in word recognition. In DA Balota, GB Flores d'Arcais and K Rayner eds. *Comprehension processes in reading*, pp. 9-32. Erlbaum, Hillsdale, NJ.

Bard EG, Shillcock RC and Altmann GTM (1988). The recognition of words after their acoustic offsets in spontaneous speech. *Perception & Psychophysics*, **44**, 395-408.

Berkum JJA van, Zwitserlood P, Brown CM and Hagoort P (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research*, **17**, 701-18.

Blutner R and Sommer R (1988). Sentence processing and lexical access: The influence of the focus-identifying task. *Journal of Memory and Language*, **27**, 359-67.

Borsky S, Tuller B and Shapiro LP (1998). "How to milk a coat": The effects of semantic and acoustic information on phoneme categorization. *Journal of the Acoustical Society of America*, **103**, 2670-6.

Brink D van den, Brown CM and Hagoort P (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience*, **13**, 967-85.

Cho T, McQueen JM and Cox EA (submitted). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. Manuscript submitted for publication.

Christophe A, Peperkamp S, Pallier C, Block E and Mehler J (2004). Phonological phrase boundaries constrain lexical access: I. Adult data. *Journal of Memory and Language*, **51**, 523-47.

Cluff MS and Luce PA (1990). Similarity neighborhoods of spoken two-syllable words: Retroactive effects on multiple activation. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 551-63.

Coenen E, Zwitserlood P and Boelte J (2001). Variation and assimilation in German: Consequences for lexical access and representation. *Language and Cognitive Processes*, **16**, 535-64.

Connine CM (1987). Constraints on interactive processes in auditory word recognition: The role of sentence context. *Journal of Memory and Language*, **26**, 527-38.

Connine CM (2004). It's not what you hear but how often you hear it: On the neglected role of phonological variant frequency in auditory word recognition. *Psychonomic Bulletin & Review*, **11**, 1084-9.

Connine CM, Blasko D and Hall M (1991). Effects of subsequent sentence context in auditory word recognition: Temporal and linguistic constraints. *Journal of Memory and Language*, **30**, 234-50.

Connine CM, Blasko DG and Titone D (1993a). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language*, **32**, 193-210.

- Connine CM, Clifton CE and Cutler A (1987). Effects of lexical stress on phonetic categorization. *Phonetica*, **44**, 133-46.
- Connine CM, Mullennix J, Shernoff E and Yelen J (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **16**, 1084-96.
- Connine CM, Titone D, Deelman T and Blasko D (1997). Similarity mapping in spoken word recognition. *Journal of Memory and Language*, **37**, 463-80.
- Connine CM, Titone D and Wang J (1993). Auditory word recognition: Extrinsic and intrinsic effects of word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **19**, 81-94.
- Conrad C (1974). Context effects in sentence comprehension: A study of the subjective lexicon. *Memory & Cognition*, **2**, 130-8.
- Content A, Meunier C, Kearns RK and Frauenfelder UH (2001). Sequence detection in pseudowords in French: Where is the syllable effect? *Language and Cognitive Processes*, **16**, 609-36.
- Cooper N, Cutler A and Wales R (2002). Constraints of lexical stress on lexical access in English: Evidence from native and nonnative listeners. *Language and Speech*, **45**, 207-28.
- Cutler A (1986). *Forbear* is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech*, **29**, 201-20.
- Cutler A (2005). Lexical stress. In DB Pisoni and RE Remez eds. *The handbook of speech perception*, pp. 264-89. Blackwell, Oxford.
- Cutler A and Butterfield S (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, **31**, 218-36.
- Cutler A and Carter D (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, **2**, 133-42.

- Cutler A and Clifton CE (1984). The use of prosodic information in word recognition. In H Bouma and DG Bouwhuis eds. *Attention and performance X: Control of language processes*, pp. 183-96. Erlbaum, Hillsdale, NJ.
- Cutler A, Demuth K and McQueen JM (2002). Universality versus language-specificity in listening to running speech. *Psychological Science*, **13**, 258-62.
- Cutler A and Donselaar W van (2001). *Voornaam* is not a homophone: Lexical prosody and lexical access in Dutch. *Language and Speech*, **44**, 171-95.
- Cutler A and McQueen JM (1995). The recognition of lexical units in speech. In B de Gelder and J Morais eds. *Speech and reading: A comparative approach*, pp. 33-47. Erlbaum, Hove.
- Cutler A, Mehler J, Norris D and Seguí J (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, **25**, 385-400.
- 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**, 113-21.
- Cutler A and Otake T (1994). Mora or phoneme? Further evidence for language-specific listening. *Journal of Memory and Language*, **33**, 824-44.
- Cutler A and Otake T (1999). Pitch accent in spoken-word recognition in Japanese. *Journal of the Acoustical Society of America*, **105**, 1877-88.
- Dahan D, Magnuson JS and Tanenhaus MK (2001a). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology*, **42**, 317-67.
- Dahan D, Magnuson JS, Tanenhaus MK and Hogan EM (2001b). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes*, **16**, 507-34.
- Dahan D and Tanenhaus MK (2004). Continuous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **30**, 498-513.

- Davis MH, Marslen-Wilson WD and Gaskell MG (2002). Leading up the lexical garden-path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **28**, 218-44.
- Diehl RL, Lotto AJ and Holt LL (2004). Speech perception. *Annual Review of Psychology*, **55**, 149-79.
- Donselaar W van, Koster M and Cutler A (2005). Exploring the role of lexical stress in lexical recognition. *Quarterly Journal of Experimental Psychology*, **58A**, 251-73.
- Donselaar W van, Kuijpers C and Cutler A (1999). Facilitatory effects of vowel epenthesis on word processing in Dutch. *Journal of Memory and Language*, **41**, 59-77.
- Dumay N, Frauenfelder UH and Content A (2002). The role of the syllable in lexical segmentation in French: Word-spotting data. *Brain and Language*, **81**, 144-61.
- Dupoux E, Pallier C, Sebastián-Gallés N and Mehler J (1997). A distressing deafness in French. *Journal of Memory and Language*, **36**, 399-421.
- Eisner F and McQueen JM (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics*, **67**, 224-38.
- Ernestus M, Baayen H and Schreuder R (2002). The recognition of reduced word forms. *Brain and Language*, **81**, 162-73.
- Ernestus M and Mak WM (2004). Distinctive phonological features differ in relevance for both spoken and written word recognition. *Brain and Language*, **90**, 378-92.
- Fox RA and Unkefer J (1985). The effect of lexical status on the perception of tone. *Journal of Chinese Linguistics*, **13**, 69-90.
- Frauenfelder UH, Scholten M and Content A (2001). Bottom-up inhibition in lexical selection: Phonological mismatch effects in spoken word recognition. *Language and Cognitive Processes*, **16**, 583-607.

- Gaskell MG and Dumay N (2003). Lexical competition and the acquisition of novel words. *Cognition*, **89**, 105-32.
- Gaskell MG and Marslen-Wilson WD (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, **22**, 144-58.
- Gaskell MG and Marslen-Wilson WD (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*, **12**, 613-56.
- Gaskell MG and Marslen-Wilson WD (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, **24**, 380-96.
- Gaskell MG and Marslen-Wilson WD (2001). Lexical ambiguity resolution and spoken word recognition: Bridging the gap. *Journal of Memory and Language*, **44**, 325-49.
- Gaskell MG and Marslen-Wilson WD (2002). Representation and competition in the perception of spoken words. *Cognitive Psychology*, **45**, 220-66.
- Gaskell G, Spinelli E and Meunier F (2002). Perception of resyllabification in French. *Memory & Cognition*, **30**, 798-810.
- Goldinger SD (1998). Echoes of echoes?: An episodic theory of lexical access. *Psychological Review*, **105**, 251-79.
- Gow DW (2001). Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language*, **45**, 133-59.
- Gow DW (2002). Does English coronal place assimilation create lexical ambiguity? *Journal of Experimental Psychology: Human Perception and Performance*, **28**, 163-79.
- Gow DW and Gordon PC (1995). Lexical and prelexical influences on word segmentation: evidence from priming. *Journal of Experimental Psychology: Human Perception and Performance*, **21**, 344-59.

- Gow DW and Im AM (2004). A cross-linguistic examination of assimilation context effects. *Journal of Memory and Language*, **51**, 279-96.
- Grosjean F (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, **28**, 267-83.
- Grosjean F (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception & Psychophysics*, **38**, 299-310.
- Grossberg S and Myers CW (2000). The resonant dynamics of speech perception: Interword integration and duration-dependent backward effects. *Psychological Review*, **107**, 735-67.
- Isel F and Bacri N (1999). Spoken-word recognition: The access to embedded words. *Brain and Language*, **68**, 61-7.
- Kraljic T and Samuel AG (in press). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*.
- Lackner JR and Garrett MF (1972). Resolving ambiguity: Effects of biasing context in the unattended ear. *Cognition*, **1**, 359-72.
- Lahiri A and Marslen-Wilson W (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition*, **38**, 245-94.
- Lehiste I (1972). The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America*, **51**, 2018-24.
- LoCasto PC and Connine CM (2002). Rule-governed missing information in spoken word recognition: Schwa vowel deletion. *Perception & Psychophysics*, **64**, 208-19.
- Lucas MM (1987). Frequency effects on the processing of ambiguous words in sentence contexts. *Language and Speech*, **30**, 25-46.
- Luce PA (1986). A computational analysis of uniqueness points in auditory word recognition. *Perception & Psychophysics*, **39**, 155-8.

- Luce PA and Cluff MS (1998). Delayed commitment in spoken word recognition: Evidence from cross-modal priming. *Perception & Psychophysics*, **60**, 484-90.
- Luce PA, Goldinger SD, Auer ET and Vitevitch MS (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception & Psychophysics*, **62**, 615-25.
- Luce PA and Large NR (2001). Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes*, **16**, 565-81.
- Luce PA and Lyons EA (1999). Processing lexically embedded spoken words. *Journal of Experimental Psychology: Human Perception and Performance*, **25**, 174-83.
- Luce PA and Pisoni DB (1998). Recognizing spoken words: The Neighborhood Activation Model. *Ear and Hearing*, **19**, 1-36.
- Lugt AH van der (2001). The use of sequential probabilities in the segmentation of speech. *Perception & Psychophysics*, **63**, 811-23.
- Magnuson JS, McMurray B, Tanenhaus MK and Aslin RN (2003a). Lexical effects on compensation for coarticulation: The ghost of Christmash past. *Cognitive Science*, **27**, 285-98.
- Magnuson JS, Tanenhaus MK, Aslin RN and Dahan D (2003b). The time course of spoken word learning and recognition: studies with artificial lexicons. *Journal of Experimental Psychology: General*, **132**, 202-27.
- Marslen-Wilson WD (1984). Function and process in spoken word-recognition. In H Bouma and DG Bouwhuis eds. *Attention and performance X: Control of language processes*, pp. 125-50. Erlbaum, Hillsdale, NJ.
- Marslen-Wilson WD (1987). Functional parallelism in spoken word-recognition. *Cognition*, **25**, 71-102.
- Marslen-Wilson W, Moss HE and van Halen S (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, **22**, 1376-92.

- Marslen-Wilson WD, Nix A and Gaskell MG (1995). Phonological variation in lexical access: Abstractness, inference and English place assimilation. *Language and Cognitive Processes*, **10**, 285-308.
- Marslen-Wilson WD and Tyler LK (1980). The temporal structure of spoken language understanding. *Cognition*, **8**, 1-71.
- Marslen-Wilson W and Warren P (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review*, **101**, 653-75.
- Marslen-Wilson W and Zwitserlood, P (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, **15**, 576-85.
- Mattys SL (2004). Stress versus coarticulation: Towards an integrated approach to explicit speech segmentation. *Journal of Experimental Psychology: Human Perception and Performance*, **30**, 397-408.
- Mattys SL, White L and Melhorn JF (in press). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General*.
- McClelland JL (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, **86**, 287-330.
- McClelland JL and Elman JL (1986). The TRACE model of speech perception. *Cognitive Psychology*, **10**, 1-86.
- McLennan CT and Luce PA (2005). Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **31**, 306-21.
- McLennan CT, Luce PA and Charles-Luce J (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **29**, 539-53.
- McMurray B, Tanenhaus MK and Aslin RN (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition*, **86**, B33-42.

- McQueen JM (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language*, **39**, 21-46.
- McQueen JM (2003). The ghost of Christmas future: Didn't Scrooge learn to be good? Commentary on Magnuson, McMurray, Tanenhaus and Aslin (2003). *Cognitive Science*, *27*, 795-9.
- McQueen JM (2005). Speech perception. In K Lamberts and R Goldstone eds. *The handbook of cognition*, pp. 255-75. Sage, London.
- McQueen JM, Norris D and Cutler A (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **20**, 621-38.
- McQueen JM, 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**, 1363-89.
- McQueen JM, Otake T and Cutler A (2001). Rhythmic cues and possible-word constraints in Japanese speech segmentation. *Journal of Memory and Language*, **45**, 103-32.
- Milberg W, Blumstein SE and Dworetzky B (1988). Phonological factors in lexical access: Evidence from an auditory lexical decision task. *Bulletin of the Psychonomic Society*, **26**, 305-8.
- Miller JL, Green K and Schermer T (1984). On the distinction between the effects of sentential speaking rate and semantic congruity on word identification. *Perception & Psychophysics*, **36**, 329-37.
- Mirman D, McClelland JL and Holt LL (2005). Computational and behavioral investigations of lexically induced delays in phoneme recognition. *Journal of Memory and Language*, **52**, 416-35.
- Mitterer H and Blomert L (2003). Coping with phonological assimilation in speech perception: Evidence for early compensation. *Perception & Psychophysics*, **65**, 956-69.
- Mitterer H, Csépe V and Blomert L (in press). The role of perceptual integration in the perception of assimilation word forms. *Quarterly Journal of Experimental Psychology*.

- Mitterer H and Ernestus M (in press). Listeners recover /t/'s that speakers reduce: Evidence from /t/-lenition in Dutch. *Journal of Phonetics*.
- Morton J (1969). The interaction of information in word recognition. *Psychological Review*, **76**, 165-78.
- Moss HE and Marslen-Wilson WD (1993). Access to word meanings during spoken language comprehension: Effects of sentential semantic context. *Journal of Experimental Psychology: Learning, Memory & Cognition*, **19**, 1254-76.
- Moss HE, McCormick SF and Tyler LK (1997). The time course of activation of semantic information during spoken word recognition. *Language and Cognitive Processes*, **10**, 121-36.
- Nakatani LH and Dukes KD (1977). Locus of segmental cues for word juncture. *Journal of the Acoustical Society of America*, **62**, 714-9.
- Norris D (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, **52**, 189-234.
- Norris D, Cutler A, McQueen JM and Butterfield S (submitted). Phonological and conceptual activation in speech comprehension. Manuscript submitted for publication.
- Norris D, McQueen JM and Cutler A (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 1209-28.
- Norris D, McQueen JM and Cutler A (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, **23**, 299-325.
- Norris D, McQueen JM and Cutler A (2003). Perceptual learning in speech. *Cognitive Psychology*, **47**, 204-38.
- Norris D, McQueen JM, Cutler A and Butterfield S (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology*, **34**, 191-243.
- Norris D, McQueen JM, Cutler A, Butterfield S and Kearns R (2001). Language-universal constraints on speech segmentation. *Language and Cognitive Processes*, **16**, 637-60.

- Oden GC and Spira JL (1983). Influence of context on the activation and selection of ambiguous word senses. *Quarterly Journal of Experimental Psychology*, **35**, 51-64.
- Onifer W and Swinney DA (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition*, **9**, 225-36.
- Otake T, Hatano G, Cutler A and Mehler J (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, **32**, 358-78.
- Otake T, Yoneyama K, Cutler A and van der Lugt A (1996). The representation of Japanese moraic nasals. *Journal of the Acoustical Society of America*, **100**, 3831-42.
- Pallier C, Sebastián-Gallés N, Felguera T, Christophe A and Mehler J (1993). Attentional allocation within the syllable structure of spoken words. *Journal of Memory and Language*, **32**, 373-89.
- Patterson D and Connine CM (2001). Variant frequency in flap production: A corpus analysis of variant frequency in American English flap production. *Phonetica*, **58**, 254-75.
- Pitt MA and McQueen JM (1998). Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language*, **39**, 347-70.
- Radeau M and Morais J (1990). The uniqueness point effect in the shadowing of spoken words. *Speech Communication*, **9**, 155-64.
- Radeau M, Morais J, Mousty P and Bertelson P (2000). The effect of speaking rate on the role of the uniqueness point in spoken word recognition. *Journal of Memory and Language*, **42**, 406-22.
- Radeau M, Mousty P and Bertelson P (1989). The effect of the uniqueness point in spoken-word recognition. *Psychological Research*, **51**, 123-8.
- Salverda AP, Dahan D and McQueen JM (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition*, **90**, 51-89.
- Samuel AG (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General*, **110**, 474-94.

Samuel AG (2001). Knowing a word affects the fundamental perception of the sounds within it. *Psychological Science*, **12**, 348-51.

Samuel AG and Pitt MA (2003). Lexical activation (and other factors) can mediate compensation for coarticulation. *Journal of Memory and Language*, **48**, 416-34.

Scharenborg O, Norris D, den Bosch L and McQueen JM (in press). How should a speech recognizer work? *Cognitive Science*.

Sebastián-Gallés N, Dupoux E, Seguí J and Mehler J (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language*, **31**, 18-32.

Seidenberg MS, Tanenhaus MK, Leiman JM and Bienkowski M (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, **14**, 489-537.

Shattuck-Hufnagel S and Turk AE (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, **25**, 193-247.

Shatzman KB and McQueen JM (in press). Segment duration as a cue to word boundaries in spoken-word recognition. *Perception & Psychophysics*.

Shatzman KB and McQueen JM (submitted-a). The activation of offset-embedded words: Evidence from eye-tracking and identity priming. Manuscript submitted for publication.

Shatzman KB and McQueen JM (submitted-b). Prosodic knowledge affects the recognition of newly-acquired words. Manuscript submitted for publication.

Shillcock RC (1990). Lexical hypotheses in continuous speech. In GTM Altmann, ed. *Cognitive models of speech processing: Psycholinguistic and computational perspectives*, pp. 24-49. MIT Press, Cambridge, MA.

Simpson GB (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning and Verbal Behavior*, **20**, 120-36.

- Smits R, Warner N, McQueen JM and Cutler A (2003). Unfolding of phonetic information over time: A database of Dutch diphone perception. *Journal of the Acoustical Society of America*, **113**, 563-74.
- Soto-Faraco S, Sebastián-Gallés N and Cutler A (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language*, **45**, 412-32.
- Spinelli E, McQueen JM and Cutler A (2003). Processing resyllabified words in French. *Journal of Memory and Language*, **48**, 233-54.
- Spivey MJ, Grosjean M and Knoblich G (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences*, **102**, 10393-8.
- Stevens KN (2002). Toward a model for lexical access based on acoustic landmarks and distinctive features. *Journal of the Acoustical Society of America*, **111**, 1872-91.
- Streeter LA and Nigro GN (1979). The role of medial consonant transitions in word perception. *Journal of the Acoustical Society of America*, **65**, 1533-41.
- Suomi K, McQueen JM and Cutler A (1997). Vowel harmony and speech segmentation in Finnish. *Journal of Memory and Language*, **36**, 422-44.
- Swinney D (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, **18**, 645-59.
- Tabossi P (1988a). Effects of context on the immediate interpretation of unambiguous nouns. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **14**, 153-62.
- Tabossi P (1988b). Accessing lexical ambiguity in different types of sentential contexts. *Journal of Memory and Language*, **27**, 324-40.
- Tabossi P, Burani C and Scott D (1995). Word identification in fluent speech. *Journal of Memory and Language*, **34**, 440-67.
- Tabossi P, Collina S, Mazzetti M and Zoppello M (2000). Syllables in the processing of spoken Italian. *Journal of Experimental Psychology: Human Perception and Performance*, **26**, 758-75.

Tabossi P and Zardon F (1993). Processing ambiguous words in context. *Journal of Memory and Language*, **32**, 359-72.

Tanenhaus MK and Donenwerth-Nolan S (1984). Syntactic context and lexical access. *Quarterly Journal of Experimental Psychology*, **36A**, 649-61.

Tanenhaus MK, Leiman JM and Seidenberg MS (1979). *Journal of Verbal Learning and Verbal Behavior*, **18**, 427-40.

Tyler LK (1984). The structure of the initial cohort: Evidence from gating. *Perception & Psychophysics*, **62**, 1297-1311.

Tyler LK and Wessels J (1983). Quantifying contextual contributions to word-recognition processes. *Perception & Psychophysics*, **34**, 409-20.

Utman JA, Blumstein SE and Burton MW (2000). Effects of subphonetic and syllable structure variation on word recognition. *Perception & Psychophysics*, **62**, 1297-1311.

Van Petten C, Coulson S, Rubin S, Plante E and Parks M (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **25**, 394-417.

Vitevitch MS (2002). Influence of onset density on spoken-word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **28**, 270-8.

Vitevitch MS and Luce PA (1998). When words compete: Levels of processing in spoken word recognition. *Psychological Science*, **9**, 325-9.

Vitevitch MS and Luce PA (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, **40**, 374-408.

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

Vroomen J and de Gelder B (1997). Activation of embedded words in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **23**, 710-20.

Vroomen J and de Gelder B (1999). Lexical access of resyllabified words: Evidence from phoneme monitoring. *Memory & Cognition*, **27**, 413-21.

Vroomen J, Tuomainen J and de Gelder B (1998). The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language*, **38**, 133-49.

Vroomen J, van Zon M and de Gelder B (1996). Cues to speech segmentation: Evidence from juncture misperceptions and word spotting. *Memory & Cognition*, **24**, 744-55.

Warner N, Kim J, Davis C and Cutler A (2005a). Use of complex phonological patterns in speech processing: Evidence from Korean. *Journal of Linguistics*, **41**, 353-87.

Warner N, Smits R, McQueen JM and Cutler A (2005b). Phonological and statistical effects on timing of speech perception in Dutch: Insights from a database of Dutch diphone perception. *Speech Communication*, **46**, 53-72.

Warren P and Marslen-Wilson W (1987). Continuous uptake of acoustic cues in spoken word recognition. *Perception & Psychophysics*, **41**, 262-75.

Warren P and Marslen-Wilson W (1988). Cues to lexical choice: Discriminating place and voice. *Perception & Psychophysics*, **43**, 21-30.

Weber A (2001). Help or hindrance: How violation of different assimilation rules affects spoken-language processing. *Language and Speech*, **44**, 95-118.

Weber A and Cutler A (submitted). First-language phonotactics in second-language listening. Manuscript submitted for publication.

Whalen DH (1984). Subcategorical phonetic mismatches slow phonetic judgments. *Perception & Psychophysics*, **35**, 49-64.

Whalen DH (1991). Subcategorical phonetic mismatches and lexical access. *Perception & Psychophysics*, **50**, 351-60.

Whitney P, McKay T, Kellas G and Emerson WA (1985). Semantic activation of noun concepts in context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **11**, 126-35.

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

Zwitserslood P and Schriefers H (1995). Effects of sensory information and processing time in spoken-word recognition. *Language and Cognitive Processes*, **10**, 121-36.