

Listening to a second language through the ears of a first

Anne Cutler

Max Planck Institute for Psycholinguistics, The Netherlands

The processes by which listeners recognize spoken language are highly language-specific. Listeners' expectations of how meaning is expressed in words and sentences are formed by the lexicon and grammar of the native language; but the phonology plays an even more immediate role. Thus the native phoneme repertoire constrains listeners' ability to discriminate phonetic contrasts; and a further area in which such constraints arise is the segmentation of continuous speech into its component words. A large body of research is summarised here, motivating three conclusions: (1) In segmenting speech, speakers of different languages apply different heuristic procedures, efficiently exploiting the specific phonological structure of their various languages. (2) These procedures have become part of the listeners' processing system, to an extent that they are also applied when listening to nonnative languages, even though this may lead to inefficiency. (3) It may be impossible to acquire the use of multiple procedures of this kind; but it is possible to inhibit the misapplication of native procedures to other languages for which they are inefficient.

Keywords: Phonology, reception, language processing

1. Introduction: Language-specific listening

Understanding language is one of the most refined skills of the human mind. Much of the acquisition of this complex ability is, however, accomplished while we are lying around for want of the motor skills necessary to undertake anything more adventurous. That is, we learn to listen in the first year of our life.

At birth babies already know a little about the way language sounds. At only a few days of age they show preference for speech in the mother's language over

speech in another language (Moon, Panneton-Cooper & Fifer, 1993). The rhythmic patterns of speech can, it appears, be perceived in the womb before birth (Armitage, Baldwin & Vince, 1980; Abrams, Gerhardt, Huang, Peters & Langford, 2000), and young babies can discriminate between languages with different rhythmic structure but not between two rhythmically similar languages (Nazzi, Bertoncini & Mehler, 1998).

Sensitivity to speech sound contrasts is refined during the first year of life. In the early months babies can discriminate contrasts well (Eimas, Siqueland, Jusczyk & Vigorito, 1971; see Jusczyk, 1997, for a review). Importantly, their discrimination abilities are not sensitive to the categories of the environmental language, but embrace both contrasts which are relevant for the native language (e.g. for an English-learner, a contrast between /t/ and /d/, or between /r/ and /l/) and those which are not (e.g. a dental /t/ versus a retroflex /T/ which is relevant for Hindi but not for English, or /IT/ versus /I/ which is not relevant for Japanese). Thus infants can discriminate contrasts with which they have had no previous experience as well as those which are used in their environment. The literature on speech perception by English-learning infants of six months or younger includes demonstrations of discrimination for contrasts from Thai (Aslin, Pisoni, Hennessy & Perey, 1981), Czech (Trehub, 1976), and Hindi (Werker & Tees, 1984), none of which are found in the English phonological repertoire. Similarly, infants in a Japanese-learning environment have been shown to discriminate the English /r-l/ distinction (Tsushima, Takizawa, Sasaki, Siraki, Nishi, Kohno, Menyuk & Best, 1994).

This contrasts markedly with adult performance, since English adults do not appear to discriminate the same Thai, Czech and Hindi contrasts easily, and Japanese adults have notorious difficulty with /r/ versus /l/. Adult listeners are very bad at perceiving contrasts in foreign languages if they do not occur in the native language. At some point, therefore, the infants have to turn into adults; phonetically speaking, this occurs by the end of the first year of life. Classic work by Werker and her colleagues has demonstrated that contrasts which English-learning infants of six months could indeed learn to discriminate categorically are not discriminated by ten- to twelve-month-old infants (Werker, Gilbert, Humphrey & Tees, 1981; Werker & Lalonde, 1988). Infants who made a nonnative discrimination successfully at six months have lost the ability to make that same discrimination when tested four months later (Werker & Tees, 1984). Japanese-learning infants likewise cannot perform the /r-l/ discrimination by the end of their first year (Tsushima et al., 1994). At 10-12 months of age, in other words, listeners are effectively adult.

It should be noted that this developmental change is not a loss of perceptual acuity. Adult listeners can tell whether two exemplars from a single category are the same or different, even if they cannot learn to assign them reliably to different categories. And Best, McRoberts and Sithole (1988) showed that adult listeners can indeed learn to categorise nonnative contrasts which are nothing like anything in the native phonological repertoire. Children's striking ability to acquire further languages until their teenage years attests that multiple phonological systems can be acquired. However, the important point for the present discussion is the change that occurs in infants' speech sound categorisation during the second half of the first year of life. Contrasts which do not occur in the speech environment are not discriminated; infants have identified the distinctions that they need to pay attention to in order to learn words of the ambient language. It is at this point that children indeed give evidence of learning actual words, and gratify their parents with their first utterances. Thus the specialisation of speech sound perception may be crucially involved in helping vocabulary acquisition to begin. Irrelevant variation between pronunciations can be ignored; only the phonemically relevant contrasts indicate new words.

The downside of this efficiency is that it has far-reaching implications for the learning of multiple languages. There would perhaps be little problem if all languages used the same or nearly the same set of phonetic contrasts. But this is far from the case. If languages are acquired in childhood, the perceptual system is still flexible enough to learn new contrasts. By adulthood, however, the flexibility diminishes—foreign accent is the classic mark of the adult learner of a nonnative language, and this production deficit is accompanied by perceptual deficit. And just as auditory discrimination suffers, so is visual integration of phonetic cues highly efficient in the native language (Massaro, 1998) but far less so in second-language listening (Massaro, Cohen, Gesi & Heredia, 1993; Massaro, Cohen & Smeele, 1995). Adult listeners simply cannot discriminate nonnative contrasts as efficiently as native contrasts. The early specialisation which facilitated word learning, and thus enabled the beginning of communication abilities in the native language, has produced disabled second-language learners in later life.

Intensive training with nonnative contrasts can produce an improvement in discrimination performance (Lively, Pisoni, Yamada, Tohkura & Yamada, 1994). But there is still a world of difference between the hard work which an adult must engage in to produce a small improvement and the ease with which an infant in the crib masters the same discrimination perfectly. Language-specific listening — at the level of speech sound discrimination — is an essential component of human speech processing.

2. Language-specific segmentation of speech

Acquiring an initial vocabulary presents infants with other challenges than those posed by the language-specific phonemic repertoire. Words must be extracted from their surrounding speech context. Even though utterances addressed directly to infants tend to be shorter than utterances spoken between adults, they rarely consist of only a single word (Van de Weijer, 1999). Furthermore, speakers do not pause between words of an utterance — natural speech consists of a continuous flow. No reliable signals tell the listener where word boundaries are to be found. For the infant who as yet knows no words, the boundaries of other known words also cannot help find new words. Segmentation of continuous speech into its component lexical constituents is a considerable task for the new language learner.

Yet this too is successfully achieved within the first year of life. Infants in this age range show impressive ability to exploit probabilistic patterning within the input (Aslin, Saffran & Newport, 1998). This sensitivity is undoubtedly of use in identifying recurring patterns which could potentially be words. And once a word pattern is identified, infants can recognise it in a continuous-speech context. Jusczyk and Aslin (1995) developed an experimental paradigm in which infants were first familiarised with a pair of target words in isolation. Then, in a second phase, the infants were presented with passages of short stories to listen to. Some passages contained multiple occurrences of one of the familiarized words. Jusczyk and Aslin measured how long infants maintained a head turn to listen to the passages; they found that infants preferred to listen to passages containing previously familiarised words rather than to passages containing multiple occurrences of words which had not been presented earlier. That is, the infants were able to detect the occurrences of the familiar words within the fluent speech.

Jusczyk and Aslin's findings have been replicated and extended by Jusczyk, Houston, and Newsome (1999); and Houston, Jusczyk, Kuijpers, Coolen and Cutler (2000) further showed that American-learning infants familiarised with Dutch words could recognise the familiar items in Dutch passages. Even though the familiarised words would never have occurred in the infants' previous listening experience, and even though the texts in which these words were presented were at a phonetic level clearly non-native, the infants succeeded in spotting the familiar words in the fluent contexts. These studies show that infants under a year of age are able to retain word patterns they have heard, and recognise those same patterns when they re-occur in a continuous-speech context.

Segmentation of fluent speech by adult listeners is in principle made easier by the availability of a well-stocked vocabulary. Adults can tell where a word begins by identifying the end of the preceding word, for instance. But languages complicate this approach by combining vocabularies in the tens or hundreds of thousands with phonetic repertoires of only a few dozen sounds, a circumstance which makes it inevitable that words resemble one another and occur embedded in other words. Even very short words may not be unambiguous; for instance, an utterance of *they may drink rum* contains phoneme sequences which match the words *aim*, *maid*, *aid*, *ring*, *rink* and *crumb*, as well as the four intended words. Current psycholinguistic theory accepts that speech input automatically activates all words with which it is compatible, and that word recognition occurs via a process of competition between the activated candidates (see Frauenfelder & Floccia, 1998, for a review). On this view, segmentation is a by-product of the word recognition process.

However, adult listeners also possess strategies which they can use to facilitate segmentation of continuous speech. One way to tell that this is so is to observe the errors that sometimes result. A pop song of some decades ago, for instance, contained the rather unlikely line *She's a must to avoid*, which was apparently interpreted by very many listeners independently as *She's a muscular boy*. In this case, the final syllable of the line (*void*), a stressed syllable, was interpreted as a new word (*boy*), while the two syllables before it (*to a-*), both unstressed syllables containing reduced vowels, were interpreted as non-word-initial, giving the trisyllabic word *muscular*. Remarkably, in mishearings by English listeners this pattern recurs. Cutler and Butterfield (1992) analysed natural slips of the ear and found a significant tendency to posit erroneous word boundaries before strong (stressed) syllables rather than before weak syllables, and a corresponding tendency to overlook boundaries before weak syllables. Typical examples of such errors included *by loose analogy* perceived as *by Luce and Allergy*, or *how big is it?* perceived as *how bigoted?*. Cutler and Butterfield followed up their analysis of natural slips with an experiment in which they generated mishearings in the laboratory by presenting listeners with very faint speech input. Exactly the same patterns emerged: insertion of boundaries where the input had none tended to occur before strong syllables, while deletion of boundaries in the input tended to occur before weak syllables. Thus *conduct ascents uphill* might be reported as *the doctor sends her bill*, or *sons expect enlistment* as *sons expectant listen*.

Cuddeback and Butterfield explained this pattern as evidence of a segmentation heuristic which English listeners applied to speech input: strong syllables are

likely to be word-initial. Analyses of the vocabulary of English and of corpora of natural English conversation (Cutler & Carter, 1987) suggested that the heuristic would work very well, since most content words in English in fact do begin with strong syllables. Analyses of Dutch (Schreuder & Baayen, 1993; Vroomen & De Gelder, 1995) suggested that the strategy would also work well for that language, and indeed a similar faint speech experiment by Vroomen, Van Zon and De Gelder (1996) showed that Dutch listeners apparently applied the strong syllable segmentation heuristic also.

Observation of slips of the ear is one of the more enjoyable ways to collect psycholinguistic data. However, laboratory techniques also exist for the study of speech segmentation. In the word-spotting paradigm, first developed by Cutler and Norris (1988), listeners hear short nonsense strings in which real words may be embedded, and their task is to press a button whenever they spot a known word, and then to report the detected word. For instance, in *omzel* or *clibthish* listeners would not be expected to find words, but in *bookving* and *vuffapple* they would be expected to find *book* and *apple* respectively. The technique allows researchers to minimize the segmentation problem, i.e. to compare the effects of segmenting a word from one minimal context versus another. Using this task, Order and Norris found further evidence supporting the strong syllable segmentation heuristic. Words like *mint were* harder to spot in, for instance, *mintayf* (with two strong syllables) than in *mintef* (with a weak second syllable); in *mintayf*, listeners apparently postulated a word boundary before *-tayf*, so that correct detection of *mint* required recombination of speech material across the segmentation point. The weak second syllable of *mintef*, in contrast, triggered no segmentation and hence had no adverse effect on detection of *mint*.

Again, Vroomen et al. (1996) replicated this finding in Dutch. Thus for both English and Dutch, listeners apparently exploit the statistical likelihood that words begin with strong syllables by assuming that strong syllables in the speech signal are probably word-initial. Even though this is not always going to be a successful procedure, it is apparently useful enough to maintain its place in English and Dutch listeners' processing. Cutler (1990) termed the heuristic the Metrical Segmentation Strategy, since the patterning of strong and weak syllables constitutes the metrical rhythm of English (and Dutch).

Rhythm, however, differs across languages. By no means do all languages exhibit the stress-based rhythm, in which strong syllables contrast with weak, which is characteristic of English and Dutch. This line of research therefore leads already to a conclusion:

(1) Segmentation of speech relies on language-specific procedures.

There is abundant evidence that this is so. Rhythmic effects in the segmentation of speech have indeed been demonstrated in languages differing from English and Dutch in phonological structure — that is, in languages without stress-based metrical structure. The rhythmic structure of French and Spanish is syllable-based, and evidence of syllabically based segmentation was produced for French by Mehler, Dommergues, Frauenfelder and Segui (1981), Cutler, Mehler, Norris and Segui (1986), Kolinsky, Morais and Cluytens (1995) and Peretz, Lussier and Beland (1996), and for Spanish by Sebastian-Galles, Dupoux, Segui and Mehler (1992), Bradley, Sanchez-Casas and Garcia-Albea (1993), and Pallier, Sebastian-Galles, Felguera, Christophe and Mehler (1993). Evidence of syllabic segmentation is provided for instance by more rapid detection of targets which correspond exactly to a syllable than of targets which are larger or smaller than a syllable; thus *ba-* is easier for French listeners to detect in *balance* {*ba#lance*, where # indicates a syllable boundary) than in *balcon* (syllabified *bal#con*), while *bal-* is easier to detect in *balcon* than in *balance* (Mehler et al., 1981; Cutler et al., 1986). The rhythmic structure of Japanese is mora-based, and Otake, Hatano, Cutler and Mehler (1993) and Cutler and Otake (1994) produced evidence that Japanese listeners segment speech at the boundaries of morae (these are subsyllabic units consisting of a syllabic nucleus plus optional onset, or of a syllable coda; thus *pokemon* contains four morae: *po-ke-mo-n*).

These are segmentation effects, not recoding effects; they do not necessarily imply that French listeners represent speech at some stage in processing as a sequence of syllables, nor that Japanese listeners represent speech as a sequence of morae. Recent evidence has emphasised that it is only the boundaries of the rhythmic units which are important for segmentation, not the units themselves as representational elements (Content, Meunier & Frauenfelder, 2001, and Content, Meunier, Kearns & Frauenfelder, 2001, for French; Cutler & Otake, 2002, and McQueen, Otake & Cutler, 2001, for Japanese).

Not only language-specific rhythmic structure, but also other aspects of language-specific phonology were shown to play a role in segmentation. Thus Finnish has a vowel harmony rule whereby certain vowels may not co-occur in a word. Two successive syllables containing such incompatible vowels should therefore be separated by a word boundary, and listeners can exploit this fact in speech recognition. Suomi, McQueen and Cutler (1997) used the word-spotting task to demonstrate this: they found that a word such as *palo* 'fire' was

easier to detect in *kypalo* than in *kupalo*. The Finnish vowel *yl* cannot occur in the same word as /a/ or /o/, while the vowel /u/ can; thus a word boundary must occur in *kypalo* but need not occur in *kupalo*. Clearly this exploitation of vowel harmony in segmentation is a strategy specific to languages which have this phonological feature.

Another aspect of phonological structure which is clearly language-specific is phonotactic sequencing. For instance, syllables in English cannot begin with /pi/, although this onset is fine in German. Similarly, voiced stops may be syllable-initial but not syllable-final in Dutch and German, but can be either in English. McQueen (1998) showed that listeners can exploit this phonological characteristic of language structure also, to help them locate word boundaries. Dutch listeners more easily spotted *rok* 'skirt' in *fjemrok* (/mʀ/ cannot be syllable-initial or -final; therefore a boundary must occur within the sequence) than in *fiedrok* (/dr/ must be syllable-initial and /d/ cannot be syllable-final; therefore a boundary cannot occur within /dr/). Since the latter effect does not hold for English or many other languages, this segmentation effect is again language-specific. Similar demonstrations of listener sensitivity to phonotactics in segmentation have been provided for Dutch, German and Cantonese respectively by Van der Lugt (2001), Weber (2000) and Yip (2000).

3. The locus of language-specificity

The phonological structure of the various languages discussed above clearly fosters exploitation of differing types of information in segmenting speech. Listeners efficiently make use of the phonological effects, be they rhythmic structure, vowel co-occurrence restrictions or phonotactic constraints. But to what extent are these strategies a necessary part of listeners' processing? Can listeners, for instance, choose to make use of a particular strategy in one situation but abandon it in another?

Two types of evidence, from the body of research on rhythmic effects, indicate that language-specific segmentation does not arise in response to a given input, but is part of the listener's processing competence. First, it is clearly not the case that listeners are sensitive to any and all levels of phonological structure in terms of which speech may be described. Morae and syllables, for instance, are phonological terms which can be applied to any language; but they play a role in segmentation only for native speakers of those languages in which they also form the basis of the characteristic linguistic rhythm. The rhythm of

English is not syllable-based, and English listeners do not use syllables in segmentation, irrespective of whether the input encourages syllabic segmentation or not (Cutler et al., 1986); the same is true of Japanese, and Japanese listeners do not segment syllabically either (Otake et al., 1993; Otake, Hatano & Yoneyama, 1996). Similarly, neither English nor French has moraic rhythm, and in consequence neither English listeners (Cutler & Otake, 1994) nor French listeners (Otake et al., 1993) use morae in segmentation. Rhythmic effects in segmentation are restricted to the language-specific rhythmic structure to which listeners are attuned in the native language.

Second, nonnative listeners do not segment speech the way native speakers of a given language do. English listeners presented with French do not segment it syllabically (Cutler et al., 1986), nor do Dutch listeners (Cutler, 1997), nor do Japanese listeners (Otake et al., 1996). French and English listeners do not segment Japanese at mora boundaries (Otake et al., 1993; Cutler & Otake, 1994); Japanese listeners do not use syllables in segmenting Spanish (Otake et al., 1996). Thus it is not the nature of the input which triggers the use of a particular strategy of segmentation. Instead, the strategies are part of the listening competence of a speaker of a given language, and cannot be commanded unless the requisite linguistic experience has been enjoyed. The second conclusion motivated by the body of cross-linguistic segmentation studies is, therefore:

- (2) Language-specific segmentation is in the listener, not in the speech signal.

This is still not the whole story. For not only do listeners not segment a nonnative language in the way that native speakers of that language do, but they may apply their native strategy to the nonnative language. Again there is evidence from the studies of rhythmic structure. French listeners segment Japanese input at syllable boundaries even though Japanese listeners do not (Otake et al., 1993); likewise, they segment English syllabically even though English listeners do not (Cutler et al., 1986). Japanese listeners similarly are sensitive to the moraic structure of English words even though English listeners are not — thus for English listeners detection of the target /n/ is equally fast in (for example) *candy* and *canopy*, while Japanese listeners detect /n/ faster in *candy* (in which it is a syllable coda and hence moraic) than in *canopy*.

Clearly this application of native procedures to nonnative input cannot promote listening efficiency. If the strategies in question were useful and efficient, then native speakers would use them; but they do not, they use other strategies which the nonnative listener does not command and apparently cannot call into being.

Nor are rhythmically based segmentation procedures the only such processing routines to be inappropriately applied. Phonotactics of the native language can also be misapplied to nonnative input. Weber (2000) carried out a word-spotting study in English in which the materials included nonsense words such as *moysluck* and *moysluck*. Both contain the word *luck*, which should be detected by any native English-speaking listener, or any nonnative listener with adequate listening competence in English. Moreover, because no native English words begin with the sequence *shl-*, while many begin with *si-*, *luck* should be easier to find in *moysluck* than in *moysluck*, because its onset should be clearer in the former. Weber found that this effect indeed occurred with native English listeners. However, she also conducted the same experiment with native German listeners with very high competence in English (students of interpreting). These listeners were easily able to perform word-spotting in English, and accurately found the words within the nonsense items. However, they did not find it easier to detect *luck* in *moysluck* than in *moysluck*. Instead, they showed the reverse effect — *moysluck* proved easier. In fact this aspect of the performance of the German listeners reflected patterns of phonotactic segmentation constraints not in the language they were listening for, but in their native language: in German, no words begin *si-* but many begin *shl-*.

All these effects suggest that segmentation strategies resemble patterns of sensitivity to phonetic contrasts; listeners command a repertoire of procedures appropriate for their native language and not only cannot call at will upon new procedures appropriate to input in a new language but perforce apply the native procedures to the new input irrespective of whether these act to facilitate processing or to render it inefficient. This in turn motivates an extension of the second conclusion:

- (2a) Language-specific segmentation is part of the listener's processing even when the speech signal discourages it.

Thus the efficiency with which we can segment speech in our native language into its component words brings with it a penalty: reduced efficiency in listening to a second language learned later in life, at least in so far as the first and the second language differ in those aspects of phonological structure used in speech segmentation. Listening to a second language is notoriously hard, especially in the early stages of acquisition; it is known for instance to involve the language user in greater processing load (Takano & Noda, 1993). Moreover, it is often paradoxically asymmetric with other language abilities. For instance, the skills required to listen to speech in different languages are presumably

more closely alike than the skills required to read alphabetic versus non-alphabetic orthography; but learners often report that even when reading is fluent in a second language with a different orthography from that used by their native language, listening to naturally spoken language (lectures, conversations) can remain difficult. Even highly proficient second-language users are less efficient than native speakers at listening under degraded conditions, e.g. with computer-generated speech (Mack, 1988). All these well-known effects are presumably in part to be laid at the door of language-specific listening: listening to our native language is rendered highly efficient by reliance on strategies which are specifically tailored to the native phonology, but the strategies make listening to other languages with differing phonology more problematic than it in principle needs to be.

4. Escape from language-specific listening

Language-specific listening is a widespread problem for listeners confronted with more than one language — from the impression that speakers of unfamiliar languages with a differing phonology talk faster than speakers of one's native language, to the subtle listening deficits of even highly proficient late learners. Nonetheless, there are many speakers of more than one language who seem to experience no greater inefficiency in one language than in another. Certainly it would be hard to detect obvious signs of listening difficulty in people who have learned more than one language very early in life. An investigation of early bilinguals by Cutler, Mehler, Norris and Segui (1989,1992) revealed, however, a surprising result. For the rhythmically based segmentation procedures, it appeared that these bilinguals could command no more than one of the differing procedures which were characteristic of their languages.

This group of bilinguals had acquired both English and French from a very early age, and were to all intents and purposes native speakers of both languages (most of them had been raised bilingually by one French- and one English-speaking parent). As we saw in Section 2 above, French listeners typically give evidence of segmentation at syllable boundaries, while English listeners typically use a stress-based segmentation strategy. The bilinguals took part in syllable detection experiments in both English and French (using the materials of Cutler et al., 1986), and in a word-spotting experiment in English (using the materials of Cutler and Norris, 1988). For any one experiment, the results of the group as a whole gave an unclear pattern, very unlike any of the original findings when

the same experiments had been conducted with monolingual groups. However, this turned out to be because the bilinguals were not a homogeneous group. Cutler et al. partitioned the group and examined the comparative performance of sub-groups in each experiment. The first partitioning they chose was perhaps the most obvious, namely a partitioning by country of residence; approximately half of the group were resident in the United Kingdom, half in France. However, this again failed to produce results which resembled those found with monolinguals. Similar failure attended a partitioning of the group by language of either parent.

One further partitioning was tried, which was based on the bilinguals' answer when they were asked to express a preference for one of their languages. Although all averred that they spoke each language equally happily, they were willing to fantasise in response to the question: "suppose you had a dread disease, and the only way to save your life was a brain operation which would result in the loss of one of your languages; which one would you keep?" Cutler et al. deemed their answer to be their preferred language. When the subjects were then divided into two groups according to which language they had preferred, their results fell into a pattern which clearly resembled the previous findings. Bilinguals who had chosen English performed like English monolinguals with both the English and the French materials; in the word-spotting experiment, they showed stress-based segmentation, while in the syllable detection experiments they did not show syllabic segmentation. Bilinguals who had chosen French, in contrast, showed no stress-based segmentation in the word-spotting experiment but did show syllabic segmentation when performing syllable detection in French.

This pattern suggested that the bilingual listeners, despite their exceptional mastery of both English and French, could call upon only one rhythmic segmentation procedure — either the procedure typical of French, or that typical of English, but not both. If nothing else, this underscores the rather exceptional status of the rhythmic segmentation strategies in listening: clearly, effective listening does not crucially rely upon use of these strategies. Models of spoken-word recognition can simulate recognition of continuous utterances via activation of multiple lexical forms, and competition between simultaneously active candidates, without necessarily invoking explicit segmentation procedures. And bilinguals with only one rhythmic segmentation procedure at their disposal can nonetheless faultlessly comprehend two languages with different rhythmic structures. The strategies have been developed as add-ons (so to speak) to the normal process of recognition via activation and competition; they

clearly serve a useful purpose but they are not vital to the success of the recognition process and any speaker apparently has only one such strategy available.

Of course, an obvious implication of the availability of any procedure is that it ought to be equally applicable to any language for which it is useful. Thus if a bilingual commands two languages with the same rhythmic structure, an available segmentation procedure should be usable with both. Indeed, studies by Van Zon (1997) with Dutch-English bilinguals, all of whom were in fact clearly dominant in Dutch, showed that they used stress-based segmentation procedures when listening to English or Dutch. The same author however also studied French-dominant French-Dutch bilinguals and found, in line with the results of Cutler et al. (1992), that these listeners did not use stress-based segmentation with Dutch. Bradley et al. (1993), who studied Spanish-English bilinguals, and Kearns (1994), whose experiments involved French-English bilinguals similar in background to those studied by Cuder et al. (1989, 1992), also failed to find evidence of simultaneous command of both syllable-based and stress-based segmentation procedures.

Cutler et al. (1992) speculated that the rhythmic segmentation procedures might have their origin in infants' initial vocabulary acquisition. It has long been known that infants are capable of discrimination of rhythms (Demany, McKenzie & Vurpillot, 1977); and as mentioned in the introduction, above, the rhythmic characteristics of speech signals may be perceptible to the unborn infant in the womb (Armitage et al., 1980; Abrams et al., 2000). In the past few years, moreover, evidence has accrued that rhythmic sensitivity in infancy can effect precisely the discrimination between groups of languages which would result from a classification in terms of segmentation strategies based on this aspect of phonological structure. Thus Nazzi et al. (1998) found that newborn infants could discriminate between languages from different rhythmic classes (English, Japanese) but not between languages with the same metrical structure (English, Dutch). Christophe and Morton (1998) also observed failure of discrimination for English and Dutch, this time with two-month-olds. On the other hand, Nazzi, Jusczyk and Johnson (2000) found that English-learning five-month-olds could indeed discriminate English from Dutch, as well as they could discriminate unfamiliar languages from different rhythmic classes (Japanese, Italian). However, they could not discriminate between pairs of unfamiliar languages from the same rhythmic class, irrespective of whether the rhythmic class in question was a nonnative one (Italian, Spanish) or the native class (Dutch, German).

All these findings suggest that infants are well attuned to the rhythmic characteristics of language from a very early stage, and thus support the speculation first proposed by Cutler et al. (1992) that infants exploit language rhythm to assist in segmenting speech into words. That bilinguals have only one rhythm-based procedure at their disposal, Cutler et al. suggested, could indicate that such an initial launching of segmentation is unique — the assistance is effectively only needed once. Note, however, that a test of this explanation would require discrimination experiments with infants exposed in bilingual environments to rhythmically differing languages, and to date no such studies have been conducted.

The apparent limitation on rhythm-based procedures is not the only result of the studies with bilinguals, however. For the present purposes, a separate aspect of the results is perhaps more important. Recall that Cutler et al. (1992) conducted three sorts of experiments with their bilingual listeners. Not only did they test syllable detection in French (the experiment in which French monolinguals show evidence of syllabic segmentation) and word spotting in English (the experiment in which English monolinguals show evidence of stress-based segmentation), they also tested syllable detection in English. In the original experiments of this type by Cutler et al. (1986), French listeners produced evidence of syllabic segmentation while English listeners simply showed a null effect. The responses of the bilingual listeners of Cutler et al. (1992) in this respect patterned like those of the English listeners in the earlier study — no evidence of syllabic segmentation appeared.

This finding has important implications. Although the bilinguals with a preference for French did command a syllabic segmentation procedure and did make use of it when listening to French, they did not apply it inappropriately to their other language, English. This contrasts with the inappropriate use of syllabic segmentation when listening to English by the French listeners in Cutler et al.'s 1986 study, who did not have fluent mastery of English. Bradley et al. (1993) similarly showed that Spanish-English bilinguals did not use syllabic segmentation with English, and Kearns (1994) replicated the French-English bilingual result with experiments using sentence stimuli. All these results thus motivate a third conclusion:

- (3) Inappropriate language-specific segmentation is avoidable.

That is, the bilingual listeners, although they might command only one segmentation procedure, did not apply this procedure willy-nilly where its use was inefficient. Clearly procedures which are developed to exploit the rhythmic

characteristics of a particular language will be inefficient when applied to a language with a different rhythmic structure, and sufficient experience with such a language can apparently lead to avoidance of such inefficiency. The procedure is available for use, but its application is inhibited with input in that language in which, as the bilinguals' experience has taught them, the procedure does not pay off. If adequate experience shows that the procedure does not improve listening efficiency, then listeners can learn to abandon its use.

The question of how much experience is "adequate" is, then, still an open one. Bilinguals can learn to inhibit inappropriate use of a segmentation procedure where its use would not be conducive to efficient listening; but the bilinguals in these experiments had had effectively native exposure to each language. It is tantalizing to speculate in the light of these findings, however, that significantly lower levels of exposure than this might in fact suffice. If lower levels of exposure were indeed to prove adequate for inducing abandonment of inefficient segmentation strategies, this body of research might have eventual implications for the development of improved methods for teaching the art of listening to nonnative languages. Might it be possible for even moderately proficient users of a second language to learn not to listen with the ears of their first?

Similarly, the findings summarised above have implications for the practice of interpreting. They suggest on the one hand that it is possible that even the most experienced nonnative listeners may be hindered by inefficient listening strategies; but on the other hand, they suggest that such inappropriate listening strategies can be discarded. Again, however, no research has as yet examined how the transition from the former to the latter state can best be brought about. As described in the following section, moreover, the segmentation effects discussed here do not all pattern similarly in second-language listening; whether any or all of them might be manipulable via specifically targeted listening instruction, and what form such instruction should then take, are without doubt suitable subjects for subsequent investigation.

5. Conclusion: Language-specificity in the model of the listener

This overview has concentrated on the body of work on language-specific effects in the segmentation of speech into its component words. There are of course many other language-specific effects in listening. The different types of language-specificity may have varying origins and may differ quite substantially in how strongly they affect second-language listening, in how persistent they are

even when they bring no benefit in a second language, and in how hard they are to overcome. The segmentation findings are however a case study of how listening may be constrained by the native language, how the native constraints may render second-language listening less efficient than it might be, but how it is nonetheless possible under certain circumstances to abandon the use of inappropriate and inefficient procedures.

Even within the reported segmentation effects, however, there is variation with respect to effects in a second language. First, there is variation in how easily inappropriate use of a segmentation strategy may be avoided. Consider the case of Finnish listeners' use of vowel harmony in segmenting their language; this is a strategy which is clearly language-specific in that most languages do not exhibit such vowel harmony effects. The French word *avenue* with /a/ in the first syllable and /y/ in the last would violate Finnish vowel harmony rules, but it is a perfectly acceptable French word, and has also been adopted in its French form in Dutch, and in a very similar form in English and many other European languages (though not in Finnish!). Finnish listeners with undergraduate competence in English show no sign of sensitivity to vowel sequencing effects in listening to English, however, as unpublished studies by the present author with K. Suomi and J. McQueen have shown. The vowel harmony effect appears to be relatively easily inhibited in a nonnative language in which its use would be inappropriate; certainly it seems to be less stubborn than the rhythm-based effects described in the preceding sections.

Second, segmentation effects may vary in whether more than one may be at a speaker's command. Although the evidence from bilingual speakers described in the preceding section suggests that speakers may be restricted to only one rhythmic segmentation procedure, the use of phonotactic sequencing constraints in segmentation may not be similarly limited. Weber's (2000) studies with highly competent German-native speakers of English suggested that they could exploit both their native German sequence restrictions and also English constraints. Recall that listeners in this study detected (for example) *luck* in *moysluck*, *moyshluck* or *moyfluck*. Both *fl-* and *shl-* are possible onsets in German, and both *fl-* and *si-* in English. Although, as described in Section 3 above, the German listeners found *luck* easier to detect in *moysluck* than in *moyshluck*, their detection responses were nevertheless faster given *moyshluck* than given *moyfluck*. This difference also appeared in the English listeners' responses (though for them *moyfluck* and *moysluck* were equally difficult). The overall pattern of Weber's findings therefore suggests that the German listeners maintained sensitivity to their native sequencing constraints (inappropriately)

in listening to English, but had also acquired sensitivity to English sequencing constraints. Again, this double sensitivity contrasts with the apparent pattern for rhythmically based segmentation.

Third, there are factors which play a role in speech segmentation and are not necessarily subject to cross-language variation. Norris, McQueen, Cutler and Butterfield (1997) describe an effect whereby words are harder to detect if they leave a residue which could not itself be a word; for instance, English listeners find it harder to detect *apple* in *fapple* than in *vuffapple*. Neither/nor *vuffis* actually a word of English, but while *vuff* might have been (both its first two and its last two segments occur in existing words: e.g. *vulture*, *tough*), *f* could never be a word, because no English word consists of just a single consonant. This effect, which Norris et al. dubbed the Possible-Word Constraint, has been replicated in many languages (see e.g. McQueen & Cutler, 1998; McQueen et al., 2001; Cutler, Demuth & McQueen, 2002). However, it appears to have essentially the same form in all languages in which it has been investigated so far. That is, although the constraints on possible vocabulary items differ across languages, the operation of the Possible-Word Constraint appears not to. In English, for example, an open syllable with a short full vowel such as [e] or [æ] could not be a word (whereas open syllables with short vowels are perfectly acceptable words in Japanese or French). Norris, McQueen, Cutler, Butterfield and Kearns (2001) showed that English listeners detected words as easily in contexts consisting of syllables which could not be English words as in contexts which might be English words. Only single consonant contexts made detection difficult, and this was true in other languages also.

Thus listening is (unsurprisingly) a mix of language-universal and language-specific processes. Listeners exploit whatever aspects of linguistic structure they can to make listening to speech as efficient as it can be; many of these structural factors are specific to a particular language or group of languages, while others are universal phenomena.

How should this mix of factors be incorporated into a psycholinguistic model of speech recognition? In fact, this is less of a problem than it might seem. The current framework for spoken-word recognition modelling is eminently suited to provide the necessary flexibility. All current models in this area propose, in one form or another, automatic activation of word forms and a competition process from which recognition ensues. This framework we can assume to be the universal backbone of the model. Language-specific effects can then be added in a variety of ways. In Shortlist (Norris, 1994), for example, the rhythmic and phonotactic cues to boundaries can be assumed to provide

markers which can be used in the modulation of activation strength of candidate words. The Possible-Word Constraint was, for instance, instantiated in Shortlist by Norris et al. (1997) as a filter which checks the context between the edges of each activated word and the nearest known boundary in each direction; the word's activation is reduced if only consonantal material is found in such a portion of the signal. This manner of operation is universal. However, the Possible-Word Constraint can refer to language-specific factors, in that what constitutes a hypothesised boundary will differ across languages. Thus in German a sequence [si] will contain a boundary but a sequence [pf] might not, while in English the reverse will be the case. The various processes in the model are universal, while the instantiations of different types of information, or even the kind of information itself (e.g. in the case of vowel harmony), can differ across languages. The types of boundary information may also vary in force, from obligatory (e.g. silence at onset or offset of an utterance, or word-internally impermissible phoneme sequences) through degrees of probability (e.g. the various rhythmic unit boundaries). Note that within this framework it is not assumed that such boundary information is used to constrain lexical activation (such that lexical access is triggered by the presence of a boundary), but to modulate it, by defining the domains within which support for activated words is reduced when they result in implausible residues. The amount by which support is modulated can vary with the force of the different sources of boundary information.

The further extension of such a framework to input from more than one language is not in principle problematic, though it offers interesting challenges. Note that it is not proposed that bilingual listening be modelled in any way differently from the standard model. If input in more than one language is to be processed, then more than one lexicon of word forms must be accessible, and more than one set of language-specific constraints can come into play, but the basic model is unaltered. The challenges arise in simulating the different ways in which different effects — e.g. rhythmic versus phonotactic — play a role in second-language listening. Further, it will be a challenge to develop a model which progresses, as the empirical evidence suggests that listeners do, from extensive use of native strategies in nonnative listening to later abandonment of native strategies which are inappropriate, and moreover to incorporate such developmental progression in a different manner as warranted by the evidence pertaining to each type of segmentation effect.

The present overview has motivated three conclusions: that segmentation of speech relies on language-specific procedures, that these procedures are part

of the listener's processing, rather than being consequential upon aspects of the speech signal, and that although they may therefore be applied irrespective of their appropriateness, such inappropriate language-specific segmentation is avoidable. As the evidence demonstrates, the tailoring of language processing to the native language comes in a multiplicity of forms, many of which have their origins in the earliest stages of language acquisition. Some such language-specific constraints may be directly applicable to certain nonnative languages, others may be easy to adapt and to operate in more than one version, while with others it may be the case that the best a language user can hope to achieve is avoidance of inappropriate application. Although the extreme efficiency with which we understand language is an immense achievement of human cognition, it can have as a side-effect a paradoxical inefficiency in processing of nonnative languages; the next challenge is to understand how such inefficiency can most effectively be eliminated.

Submission date: 10 January 2001; *Accepted:* 20 April 2001; *Published:* July 2002

References

- Abrams, R. M., Gerhardt, K. J., Huang, X., Peters, A. J. M., & Langford, R. G. (2000). Musical experiences of the unborn baby. *Journal of Sound and Vibration*, 231, 253-258.
- Armitage, S.E., Baldwin, B. A., 8s Vince, M. A. (1980). The fetal sound environment of sheep. *Science*, 208,1173-1174.
- Aslin, R.N., Saffran, J.R., & Newport, E.L. (1998). Computation of conditional probability statistics by 8-month old infants. *Psychological Science*, 9, 321-324.
- Aslin, R.N., Pisoni, D.B., Hennessy, B.L., & Perey, A. J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child Development*, 52,1135-1145.
- Best, C.T., McRoberts, G.W., 8c Sithole, N.M. (1988). Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception & Performance*, 14, 345-360.
- Bradley, D.C., Sanchez-Casas, R.M., 8c Garcia-Albea, J.E. (1993). The status of the syllable in the perception of Spanish and English. *Language and Cognitive Processes*, 8,197-233.
- Christophe, A., & Morton, J. (1998). Is Dutch native English? Linguistic analysis by 2-month-olds. *Developmental Science*, 1, 215-219.
- Content, A., Kearns, R. K., & Frauenfelder, U. H. (2001). Boundaries versus onsets in syllabic segmentation. *Journal of Memory and Language*, 45,177-199.
- Content, A., Meunier, C, Kearns, R. K., & Frauenfelder, U. H. (2001). Sequence detection in pseudowords in French: Where is the syllable effect? *Language and Cognitive Processes*, 16,609-636.

- Cutler, A. (1990). Exploiting prosodic probabilities in speech segmentation. In G. Altmann (Ed.) *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives* (pp. 105-121). Cambridge, MA: MIT Press.
- Cutler, A. (1997). The syllable's role in the segmentation of stress languages. *Language and Cognitive Processes*, 12, 839-845.
- Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, 31, 218-236.
- Cutler, A., & Carter, D.M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language*, 2, 133-142.
- Cutler, A., Demuth, K., & McQueen, J.M. (2002). Universality versus language-specificity in listening to continuous speech. *Psychological Science*, 13, 258-262.
- Cutler, A., Mehler, J., Norris, D.G., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, 25, 385-400.
- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1989). Limits on bilingualism. *Nature*, 340, 229-230.
- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology*, 24, 381-410.
- Cutler, A., & Norris, D.G. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 113-121.
- Cutler, A., & Otake, T. (1994). Mora or phoneme? Further evidence for language-specific listening. *Journal of Memory and Language*, 33, 824-844.
- Cutler, A., & Otake, T. (2002). Rhythmic categories in spoken-word recognition. *Journal of Memory and Language*, 46, 296-322.
- Demany, L., McKenzie, B., & Surpillot, E. (1977). Rhythm perception in early infancy. *Nature*, 266, 718-719.
- Eimas, P.D., Siqueland, E.R., Jusczyk, P.W., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171, 303-306.
- Frauenfelder, U.H., & Floccia, C. (1998). The recognition of spoken words. In A.D. Friederici (ed.) *Language comprehension: A biological perspective* (pp. 1-40). Berlin: Springer.
- Houston, D.M., Jusczyk, P. W., Kuijpers, C, Coolen, R., & Cutler, A. (2000). Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin and Review*, 7, 504-509.
- Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.
- Jusczyk, P.W., & Aslin, R.N. (1995). Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology*, 29, 1-23.
- Jusczyk, P. W., Houston, D., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39, 159-207.
- Kearns, R. K. (1994). *Prelexical speech processing in mono- & bilinguals*. PhD thesis, University of Cambridge.
- Kolinsky, R., Morais, J., & Cluytens, M. (1995). Intermediate representations in spoken word recognition: Evidence from word illusions. *Journal of Memory and Language*, 34, 19-40.
- Lively, S.E., Pisoni, D.B., Yamada, R.A., Tohkura, Y., & Yamada, T. (1994). *Journal of the Acoustical Society of America*, 96, 2076-2087.
- Lugt, A.H. van der (2001). The use of sequential probabilities in the segmentation of speech. *Perception & Psychophysics*, 63, 811-823.

- Mack, M. (1988). Sentence processing by non-native speakers of English: Evidence from the perception of natural and computer-generated anomalous L2 sentences. *Journal of Neurolinguistics*, 3, 293-316.
- Massaro, D.W. (1998). *Perceiving talking faces: From speech perception to a behavioral principle*. Cambridge, MA: MIT Press.
- Massaro, D.W., Cohen, M.M., Gesi, A., & Heredia, R. (1993). Bimodal speech perception: An examination across languages. *Journal of Phonetics*, 21, 445-478.
- Massaro, D.W., Cohen, M.M., & Smeele, P.M.T. (1995). Cross-linguistic comparisons in the integration of visual and auditory speech. *Memory and Cognition*, 23, 113-131.
- McQueen, J.M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language*, 39, 21-46.
- McQueen, J.M., & Cutler, A. (1998). Morphology in word recognition. In A.M. Zwicky & A. Spencer (eds.) *The Handbook of Morphology* (pp. 406-427). Oxford: Blackwell.
- McQueen, J.M., Otake, T., & Cutler, A. (2001). Rhythmic cues and possible-word constraints in Japanese speech segmentation. *Journal of Memory and Language*, 45, 103-132.
- Mehler, J., Dommergues, J.-Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning & Verbal Behavior*, 20, 298-305.
- Moon, C, Panneton-Cooper, R., & Fifer, W.P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, 16, 495-500.
- Nazzi, X, Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perceptual Performance*, 4, 756-766.
- Nazzi, T, Jusczyk, P.W., & Johnson, E.K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, 43, 1-19.
- Norris, D.G. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52, 189-234.
- Norris, D.G., McQueen, J.M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology*, 34, 191-243.
- Norris, D.G., McQueen, J.M., Cutler, A., Butterfield, S., & Kearns, R.K. (2001). Language-universal constraints on speech segmentation. *Language and Cognitive Processes*, 16, 469-490.
- Otake, T, Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, 32, 258-278.
- Otake, X, Hatano, G., & Yoneyama, K. (1996). Speech segmentation by Japanese listeners. In T. Otake & A. Cutler (Eds.) *Phonological Structure and Language Processing: Cross-Linguistic Studies* (pp. 183-201). Berlin: Mouton de Gruyter.
- Pallier, C, Sebastian-Galles, N., Felguera, T, Christophe, A., & Mehler, J. (1993). Attentional allocation within the syllabic structure of spoken words. *Journal of Memory and Language*, 32, 373-389.
- Peretz, I, Lussier, I., & Beland, R. (1996). The roles of phonological and orthographic code in word stem completion. In T. Otake, & A. Cutler (Eds.), *Phonological Structure and Language Processing: Cross-Linguistic Studies* (pp. 217-226). Berlin: Mouton de Gruyter.
- Schreuder, R., & Baayen, R.H. (1994). Prefix stripping re-revisited. *Journal of Memory and Language*, 33, 357-375.

- Sebastian-Galles, N., Dupoux, E., Segui, J., & Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language*, 31, 18-32.
- Suomi, K., McQueen, J.M., & Cutler, A. (1997). Vowel harmony and speech segmentation in Finnish. *Journal of Memory and Language*, 36, 422-444.
- Takano, Y., & Noda, A. (1993). A temporary decline of thinking ability during foreign language processing. *Journal of Cross-Cultural Psychology*, 24, 445-462.
- Trehub, S.E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development*, 47, 166-172.
- Tsushima, T., Takizawa, O., Sasaki, M., Siraki, S., Nishi, K., Kohno, M., Menyuk, P., & Best, C. (1994). Discrimination of English *li-ll* and */w-y/* by Japanese infants at 6-12 months: Language specific developmental changes in speech perception abilities. *Proceedings of the Third International Conference on Spoken Language Processing* (pp. 1695-1698). Yokohama, Japan.
- Van de Weijer, J. (1999). *Language input for word discovery*. PhD thesis, University of Nijmegen, The Netherlands.
- Van Zon, M. D.C.M. (1997). *Speech processing in Dutch: A cross-linguistic approach*. PhD thesis, Tilburg University, The Netherlands.
- Vroomen, J., & de Gelder, B. (1995). Metrical segmentation and lexical inhibition in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 98-108.
- Vroomen, J., van Zon, M., & De Gelder, B. (1996). Cues to speech segmentation: Evidence from juncture misperceptions and word spotting. *Memory and Cognition*, 24, 744-755.
- Weber, A. (2000). The role of phonotactics in the segmentation of native and non-native continuous speech. In *Proceedings of SWAP (Workshop on Spoken Word Access Processes)* (pp. 143-146). Nijmegen, The Netherlands.
- Werker, J.F., Gilbert, J.H., Humphrey, K., & Tees, R.C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, 52, 349-355.
- Werker, J.F., & Lalonde, C.E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology*, 24, 672-683.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*, 7, 49-63.
- Yip, M.C. W. (2000). Recognition of spoken words in continuous speech: Effects of transitional probability. In *Proceedings of the Sixth International Conference on Spoken Language Processing*, (pp. 758-761). Beijing, China.

Author's address

Max Planck Institute for Psycholinguistics
 P.O. Box 310
 6500 AH Nijmegen, The Netherlands
 E-mail: anne.cutler@mpi.nl

About the author

Anne Cutler is one of four directors of the Max-Planck-Institute for psycholinguistics in Nijmegen, the Netherlands, where she is responsible for work in the area of language comprehension. Before moving to Nijmegen in 1993 she had studied in Australia, Germany and the US, and had worked in the UK at the University of Sussex and, from 1982 to 1993, at MRC Applied Psychology Unit in Cambridge. Her research centres on the recognition of spoken language, in particular the role of prosody in comprehension, and this in turn (since prosodic structure varies so widely across languages) led her into cross-linguistic work, with current projects involving over a dozen languages from different families.