

INVITED ADDRESS

**THE ROLE OF RHYTHMIC CUES FOR SPEECH SEGMENTATION
IN MONOLINGUAL AND BILINGUAL LISTENERS**

Vincent GOETRY and Régine KOLINSKY
Université Libre de Bruxelles

Spoken word recognition involves the segmentation and identification of a continuous and highly complex stimulus. It has been proposed that, in segmenting speech, listeners apply a universal rhythmic strategy that has language-specific manifestations depending on the phonological characteristics of their native language (Cutler, Mehler, Norris, & Segui, 1983, 1986): While native listeners of Romance languages like French are said to rely on syllabic structures, native listeners of Germanic languages like English or Dutch would use metrical structures. In the first part of the present paper, these proposals are discussed with regard to speech segmentation in monolinguals. It will be argued that word stress may provide powerful cues to word boundaries in both French and Dutch. The second part of the present contribution addresses the issue of speech segmentation in bilinguals, and, in particular, the claim that bilinguals develop a single rhythmic segmentation procedure restricted to their dominant language (Cutler, Mehler, Norris, & Segui, 1992). It will be argued instead that the use of adapted rhythmic segmentation cues is a necessary component of second language acquisition, and, consequently, that bilinguals who attain a high level of proficiency in their second language are able to exploit the rhythmic structures of that language in speech segmentation.

A striking paradox characterises the fields of psycholinguistics. Whereas everyday interactions provide listeners with the subjective impression that the speech signal presents a set of clear and discrete linguistic units (i.e., words, syllables, phonemes), the objective analysis of the acoustic input

Vincent Goetry, Laboratoire de Psychologie Expérimentale; Régine Kolinsky, Unité de Recherches en Neurosciences Cognitives.

This research was supported by an *Action de Recherche Concertée* grant from the *Direction Générale de la Recherche Scientifique-Communauté Française de Belgique* (convention 96/01-203). The first author is a Research Fellow of the Belgian *Fonds National de la Recherche Scientifique* (F.N.R.S.), the last author is Research Associate of the F.N.R.S. The experimental data collected by Goetry et al. (in progress) that are mentioned in the present work were sponsored by a Belgian Psychological Society (BPS) award for the best undergraduate thesis in Psychology. We thank the President and the members of the BPS Committee for support to this research, as well as the Editor of *Psychologica Belgica*, Axel Cleeremans, who invited us to write the present paper. We also thank Jose Morais, Marc Brysbaert and Axel Cleeremans, for their helpful comments on an earlier version of the manuscript.

Correspondence concerning this article should be addressed to Vincent Goetry, Laboratoire de Psychologie Expérimentale, Université Libre de Bruxelles, C.P. 191, Av. F. D. Roosevelt, 50, 1050 Brussels. Electronic mail may be sent to vgoetry@ulb.ac.be.

reveals that it is actually a continuous and highly variable complex stimulus. That is, the speech stream does not contain obvious and constant cues indicating word boundaries (e.g., Lehiste, 1962; Nakatani & Dukes, 1977). Moreover, phonemes are not uttered in a pure sequential way but are co-articulated, and thus each of them also carries information on its bordering phonemes (e.g., Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Liberman, 1960). In addition to this constant overlap, the acoustic signal is also characterised by acoustic-phonetic variability, i.e. the acoustic properties that correspond to a particular segment vary as a function of phonetic context, speech rate, speakers accent, etc. (Liberman et al., 1967). Psycholinguists have been working for decades to understand the processes by which listeners segment and identify such an acoustic medley so efficiently and rapidly despite its variable and continuous nature.

The present paper focuses on the role of rhythmic cues in speech segmentation, in particular in French and Dutch. After a short introduction to the types of solutions that have been proposed to solve "the word boundary problem" (Cutler, Dahan, & van Donselaar, 1997, p. 147), two main issues are examined. The first part of the paper discusses the notion that, in monolingual listeners, the speech recognition processes are assisted by a syllabic segmentation strategy in Romance languages like French and by a metrical segmentation strategy in stress-based languages like English and Dutch. On the basis of a critical review of the available experimental evidence, it will be argued that both French and Dutch monolinguals use stress cues for speech segmentation in their respective native language. The second part addresses the issue of speech segmentation in bilingual listeners. In particular, the notion that bilinguals develop a single rhythmic segmentation procedure is challenged by recent data suggesting that highly proficient bilinguals are able to use adapted segmentation cues in their non-dominant or second language.

Introduction: Solutions to the Word Boundary Problem

According to the first models devoted to speech recognition, listeners identify words in a strictly sequential order. For example, in *Cohort* (Marslen-Wilson & Welsh, 1978; Marslen-Wilson, 1987), the unfolding speech signal generates on line a phase of lexical activation, followed by a phase of deactivation of the candidates that no longer match the gradually incoming input. The result of these processes is the isolation of a unique candidate completely matched with the input and theoretically corresponding to the word to recognise. Within this view, speech segmentation is an *implicit* consequence of word recognition, since the identification of each

word specifies the location of the following word onset.

The segmentation efficiency of such a speech recogniser lies in its anticipation of words offsets through the isolation of the correct candidate before all phonetic information concerning this candidate is available (see e.g. Mattys, 1997, for a comprehensive review). However, words are not always recognised before their end (Bard, Shillcock, & Altmann, 1988; Grosjean, 1985; Tabossi, Burani, & Scott, 1995). Moreover, many polysyllabic words contain shorter embedded words. For example, the French word *ravissant* (*ravening*) contains the shorter words *rat* (*rat*), *vit* (*lives*), *sans* (*without*), *ravi* (*delighted*), *vissant* (*screwing*), etc. (for statistical analyses, see Frauenfelder, 1991a, for Dutch; McQueen, Cutler, Briscoe, & Norris, 1995, for English). This embeddedness problem, which cannot be solved by strictly sequential models, requires decision about word identity to be delayed until enough information is available. Such a delayed commitment component has been introduced in subsequent speech recognition models like *TRACE* (McClelland & Elman, 1986) and *Shortlist* (Norris, 1994), which involve a competition between many lexical candidates beginning at different points of the signal and considered in parallel (for detailed comparison of these two models, see e.g., McQueen, Norris, & Cutler, 1994; McQueen et al., 1995; Norris, McQueen, & Cutler, 1995).

When considered alone, lexical competition does not solve the problem of speech segmentation, which requires a decision on the parts of the acoustical wave that correspond to the beginning of words (e.g., Frauenfelder, 1991b; Norris et al., 1995). Moreover, the highly complex and variable nature of the speech stream would frequently lead to erroneous recognition if speech parsing were merely relying on a simple matching process between parts of the input and stored representations. Instead, listeners are able to identify spoken words correctly and almost instantly, i.e. to discriminate each of them from amongst the tens of thousands of other words stored in their mental lexicon. As Cutler argued (1994), "a much more robust model is needed to account for what is obviously true, namely that human speech recognition is extremely successful even with background noise, distance between speakers, distortion of the speaker's vocal tract, foreign accent, slips of the tongue, etc." (p. 89).

To explain how the complex mapping between form and meaning can be successfully achieved, it was proposed that both segmentation and recognition processes could be best achieved by a prelexical component that exploits units or cues indicating where word boundaries are likely to occur. Within such a view, speech segmentation is an *explicit* process that precedes the listener's attempts to recognise words. Prelexical segmentation is actually not incompatible with lexical competition, as shown by several studies that have assessed the specific and additional effects of these two compo-

nents (McQueen et al., 1994; Norris et al., 1995; Vroomen & de Gelder, 1997a). The next part of the present paper focuses on the nature of the segmentation cues that have been proposed to assist the speech recognition process.

I. Speech Segmentation In Monolinguals

1.1. From the Search for the Universal Unit to the Notion of Language-Specific Implementations of a Universal Rhythmic Strategy

The idea that the human speech recogniser is provided with a separate, prelexical, segmentation component has prompted the search for a universal speech segmentation unit. Several candidates have been considered, ranging from temporally defined templates (e.g., Klatt, 1980) to abstract linguistic units. The most influential of these proposals was that speech recognition is assisted by a syllabic segmentation procedure (e.g., Mehler, 1981; for a review, see e.g. Frauenfelder & Kearns, 1996). In their seminal study, Mehler, Dommergues, Frauenfelder and Segui (1981) showed that French listeners detect faster a sequence of two or three phonemes when this matches exactly the first syllable of a subsequent auditory carrier word than when it does not. For example, *pa* was detected faster in *pa.lace* than in *pal.mier*, while *pal* was detected faster in *pal.mier* than in *pa.lace*. The authors concluded that "the syllable is probably the output of the segmenting device operating upon the acoustic signal. The syllable is then used to access the lexicon" (p.342).

However, a major conceptual change in theories of speech segmentation emerged with the cross-linguistic follow-up study of Cutler, Mehler, Norris and Segui (1983, 1986). As a matter of fact, they showed that while French listeners syllabified English words, English listeners did not show a syllabic effect with either French or English materials in the fragment detection task. This cross-linguistic difference was interpreted as a consequence of the specific phonological structures of the two languages. That is, French listeners segment the speech stream into syllables because their native language displays clear and little diversified syllabic structures. English listeners do not use such a strategy because English presents both widespread ambisyllabicity, i.e., consonants belonging to two syllables at once (see Kahn, 1980; Kager, 1989) and a larger variety of syllabic structures than French (see Goldman, Content, & Frauenfelder, 1996).

The idea that syllabic structures are inappropriate parsing units in English led Anne Cutler and her colleagues to consider other phonological properties of that language. An essential characteristic of English and other stress-

based languages like Dutch is the metrical distinction between strong syllables, which have full vowels, and weak syllables, which have reduced vowels (usually a schwa). Since most English words display word-initial strong syllables (Cutler & Carter, 1987), which is also the case in Dutch (Vroomen & de Gelder, 1995), Cutler and Norris (1988) proposed that an efficient strategy in such languages is to segment the speech stream and start a lexical access attempt at every strong syllable.

Laboratory-induced juncture misperceptions provide strong empirical evidence supporting the use of this *Metrical Segmentation Strategy* (henceforth, MSS) in both English (Cutler & Butterfield, 1992) and Dutch (Vroomen, van Zon, & de Gelder, 1996). These slips of the ear, which are word boundary localisation errors, are observed when listeners are asked to recognise spoken sequences on the basis of partial acoustic cues, for example when these sequences are presented at individual speech perception threshold. Both in the English and Dutch studies, such juncture misperceptions were significantly related to the metrical structure of the sentences: listeners erroneously inserted word boundaries mainly before strong syllables and deleted word boundaries mainly before weak syllables (e.g., *conDUCT asCENTS upHILL* perceived as *the DOCTOR SENDS her BILL*¹ in Cutler & Butterfield's study).

Word-spotting experiments, in which listeners are asked to detect any real word embedded in nonsense disyllabic strings (McQueen, 1996), also support the use of the MSS in English (Cutler & Norris, 1988; McQueen et al., 1994). For example, Cutler and Norris (1998) showed faster and more accurate detection of the words (e.g., *MINT*) when the second syllable of the string was metrically weak (e.g., in *MINTesh*, /mintəʃ/) than when it was metrically strong (e.g., in *MINTAYVE*, /mintɛif/). This was predicted by the MSS, according to which the second syllable triggers segmentation only when it is strong, i.e. only in *MINTAYVE*, thus requiring assembly of the speech material across a segmentation point for successful detection.

The metrical effect observed by Cutler and Norris (1988) was only replicated on correct responses, but not on reaction times, in a similar word-spotting experiment conducted in Dutch (e.g., detection of *MELK* in either /melkəs/ or /mɛlkos/, Vroomen et al., 1996). This suggests that the rhythmic cues used for speech segmentation in Dutch differ from those used in English, which might be a consequence of the different linguistic characteristics of these two languages (for a review, see Cutler & van Donselaar, in press). In particular, whereas vowel quality differences determine the metrical status of syllables in English (Fear, Cutler, & Butterfield, 1995), many weak syllables contain unstressed unreduced vowels in Dutch (Quené, 1993;

¹ Upper case signals strong syllables, lower case weak ones.

Quené & Smith, 1992; Quené & Koster, 1998; Vroomen & de Gelder, 1995). For example, the weak syllables of words like *cigar* or *cobra*, which nearly always contain reduced vowels in English, display unreduced vowels in Dutch (Cutler & van Donselaar, in press; Koster & Cutler, 1997). Therefore, vowel quality is not a reliable predictor of the metrical status of the syllable in Dutch. Since Dutch word-initial strong syllables often bear primary stress (e.g. Kager, 1989; van der Hulst, 1984), the most obvious acoustic characteristics other than vowel quality that predict metrical structure are the suprasegmental manifestations of primary stress, i.e., longer duration, higher intensity, and a flatter spectrum (Sluijter, 1995).

Vroomen and de Gelder (submitted) provided support for the idea that Dutch listeners use a *Stress Based Segmentation strategy* (henceforth, SBS), according to which word boundaries are better signalled by the degree of the syllable stress than by the occurrence of a metrically strong syllable. These authors showed indeed that the spotting of a Dutch disyllabic word like *KRAter* (meaning *crater* and pronounced as /'krətər/²) embedded in trisyllabic strings was faster when the first syllable of the target word was realised as primary stress (e.g., /,pɔ'krətər/) than when it was realised as secondary stress (e.g., /'pɔ,krətər/). This difference cannot be accounted for by the MSS, which attributes the same segmentation power to any syllable bearing a full vowel. Nor can it be attributed to acoustic differences between the two sets of trisyllabic carriers, since the effect was still observed after these differences were factored out in the response times. Vroomen, Tuomainen and de Gelder (1998) further showed that, in a learning task where listeners had to recognise new "words" previously presented in the context of an artificial continuous speech stream (cf. Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996), Dutch listeners (but not French listeners) performed better at recognising initial-stressed than non-initial stressed "words" of materials containing full vowels exclusively. Taken together, these results strongly suggest that stress is a determinant cue for speech segmentation in Dutch, and that vowel quality is less important in Dutch than in English.

Recent research has further suggested that stress may also play an important role in speech segmentation and lexical access in English. Fine-grained stress discrimination, which is a prerequisite for being able to use stress cues in speech segmentation, can be performed by English listeners in forced-choice discrimination (Mattys, 2000) or cross-modal fragment priming (Cooper, 2000). Mattys (2000; see also Mattys & Samuel, 2000) further argued that a SBS strategy would be more efficient than the MSS in English, since nearly half of the English words contain at least two strong syllables.

² The symbols ' and , indicate primary and secondary stress, respectively.

With such words, a recogniser provided with the MSS would postulate at least one erroneous word boundary, and would therefore trigger at least one erroneous lexical attempt. Mattys (2000) presented lexical statistics showing that most English words start with a primary stressed syllable, at least when weighted by frequency of occurrence, to argue that the SBS would improve the recogniser's accuracy in comparison with the MSS.

The use of stress cues in segmenting English has also been supported by studies on very young English learners. As a matter of fact, Vroomen and de Gelder (submitted) pointed out that stress pattern was confounded with metrical pattern in some of the studies showing infants' preference for the predominant strong-weak metrical pattern of English words (Jusczyk, Cutler, & Redanz, 1993; see also Echols, Crowhurst and Childers, 1997; Morgan, 1996). For instance, Morgan (1996) showed that nine-month-olds find trochaic sequences more cohesive than iambic sequences when the material included only full vowels, i.e. differed only in duration pattern (long-short vs. short-long). Together with other similar findings (Mattys, Jusczyk, Luce and Morgan, 1999), this suggests that 9-month English-learners use the distinction between primary and secondary stress, rather than the metrical distinction between full and reduced vowels, to compute word boundaries. However, since other studies showed that, by 7.5 months, English learners exploit the metrical strong/weak distinction in extracting words from connected speech (Jusczyk, Houston and Newsome, 1999; see also Jusczyk & Aslin, 1995), this issue deserves further investigation.

Whatever the exact basis (stress or metrics) of the rhythmic strategy applied in Dutch and English, the important point is that this strategy has been considered to be similar to the syllabic strategy observed in French, since "both stress in English and the syllable in French are the basis or rhythmic structure in their respective language" (Cutler et al., 1997, p. 147; see also Cutler, Mehler, Norris, & Segui, 1992). This led to the hypothesis that listeners apply a universal rhythmic solution to the word-boundary problem, by exploiting whatever rhythmic structure characterising their native language. Language-specific implementations of the universal rhythmic segmentation strategy would root in specific capacities developing from infancy on (Cutler et al., 1992; Cutler, 1994; Cutler & Mehler, 1993): Infants enter the world with a "periodicity bias" (Cutler & Mehler, 1993, p.17) that enables them to pick out the smallest recurring rhythmic regularities from the speech stream, which then allows them to progressively develop discrete lexical entries from the continuous signal.

This view is reminiscent of the early typological work of linguists like Pike (1945) or Abercrombie (1967), who classified languages into rhythmic classes. These authors proposed to distinguish between stress-timed languages (like English and Dutch), which were said to display regular inter-

stress intervals, and syllable-timed languages (like French and other Romance languages), in which syllables, rather than only stressed vowels, were said to recur at regular intervals of time establishing temporal organisation. This way to cluster languages, though criticised as too simplistic (e.g., Bertettino, 1989; Dauer, 1983; Nespor, 1990) and not fitting different temporal regularities³, is still useful. Indeed, the impression of different rhythmic types may be the by-product of the specific phonological properties of languages. In particular, syllable complexity (Bertinetto, 1981; Dauer, 1983) seems to have reliable acoustic/phonetic correlates in speech, like the vowel/consonant temporal ratio (Ramus, Nespor, & Mehler, 1999).

The rhythmic segmentation hypothesis proposed by Cutler and colleagues implies that if a language presents a rhythmic structure based on some phonological construct other than the syllabic or stress pattern, this construct should be used in segmentation. This has been illustrated with Japanese, which belongs to a third category of "mora-timed" languages⁴ (Ladefoged, 1975; see also Port, Dalby, & O'Dell, 1987; Ramus & Mehler, 1999; Ramus et al., 1999). As a matter of fact, Otake, Hatano, Cutler and Mehler (1993) observed in a Japanese fragment detection task that (mora) targets (e.g., *ta*) were detected as easily in CVC carriers (e.g., *tan.shi*) as in CV carriers (e.g., *ta.ni.shi*). In addition, Japanese listeners often missed the CVC target (*tan*) in CV carriers, presumably because this matching requires an internal segmentation of the second mora (*ni*).

1.2. *Problems for the Notion of Contrasted, Language-Specific, Rhythmic Strategies*

Various problems linked to the notion of language-specific speech segmentation strategies have already been discussed by Kolinsky (1998). First, rhythmic classes do not account for all results on syllabic sensitivity. Indeed, according to the notion of language-specific segmentation routines, all (Romance) languages presenting clearly bounded regular syllables, like Spanish, Catalan or Portuguese, should induce syllabification effects, whereas all (Germanic) languages with irregular and ambiguous syllables, like English, Dutch or German, should not (Cutler et al., 1986). Yet, this is not

³ Linguistic analyses have constantly failed to provide empirical evidence for the notion of basically different temporal regularities in "stress-timed" and "syllable-timed" languages (e.g., Delattre, 1966; Flechter, 1991; Nakatani, O'Connor & Aston, 1981; Roach 1982; Wenk & Wioland, 1982; Williams & Hiller, 1989).

⁴ The mora is a unit that can be roughly described as intermediate between the syllable and the phoneme.

the case. On the one hand, a syllabic effect is not always observed in Romance languages (e.g., Dupoux, 1993; Kearns, 1994; Frauenfelder & Content, 1999; Meunier, Content, Frauenfelder, & Kearns, 1997; Sebastián-Gallés, Dupoux, Segui and Mehler, 1992). On the other hand, some syllabic effects are observed in Germanic languages (e.g., for English: Allopenna, 1995; Bruck, Treiman, & Caravolas, 1995; Finney, Protopapas, & Eimas, 1996; Smith & Pitt, 1995; for German: Hölne & Schriefers, 1995; for Dutch: Cutler & van der Lugt, reported in Cutler, 1997; Frauenfelder, Rietveld, & van Til, reported in Kearns, 1994; van Donselaar & Stoujesdijk, 1994; Zwitserlood, Schriefers, Lahiri, & van Donselaar, 1993; Vroomen & de Gelder, 1997a).

Second, as Kolinsky (1998) further argued, the fragment detection task may tap post- rather than pre-lexical representations (see also Frauenfelder & Content, 1999; Meunier et al., 1997). If this were the case, there is a possible intervention of literacy-induced metaphonological or orthographic representations in the fragment detection task (see also Dupoux & Mehler, 1992). Japanese provides the clearer demonstration of this fact. Indeed, Kolinsky (1998) suggested that the "moraic" effect observed by Otake et al. (1993) could alternatively result from the application of an orthographic strategy based on the written kana characters that coincide regularly with the mora structure⁵. Recent evidence supporting this interpretation has been provided by Inagaki, Otake and Hatano (2000). Testing Japanese adults as well as children of various levels of kana literacy with a Japanese version of the fragment-detection task, these authors observed that the mora-based segmentation pattern was strongly associated to kana reading level. This result led the authors to conclude that, as children acquire kana literacy, the Japanese segmentation unit changes from a mixture of syllable- and mora-based to a strictly mora-based one.

We may of course accept the notion that literacy affects speech segmentation at an early processing level. Yet, a more conservative view, compatible with other experimental evidence (see discussion in Kolinsky, 1998, and in Morais & Kolinsky, 1994), would be that the fragment detection task taps later, post-lexical representations. This further complicates the empirical verification of the notion that Romance and Germanic languages induce radically different segmentation strategies. As a matter of fact, up to now, while the use of a syllabic segmentation routine has been investigated by means of the fragment detection task, the MSS was assessed using other experimental techniques, mainly the word-spotting task.

⁵ Regarding the CVC targets, the effect observed by Otake et al. (1993) may also result from phonetic/phonotactic mismatch between targets and carriers (Nakamura, Kolinsky, Spagnoletti & Morais, 1998).

These two tasks differ according to several parameters. Whereas listeners are required both to segment words from continuous speech and to access the lexical entries corresponding to these words in word spotting, this is not the case in fragment detection. First, the experimenter has essentially solved the segmentation problem when he/she presents the listeners with isolated carrier words and targets matching the onsets of these carriers (Frauenfelder & Content, 1999). As Kearns (1994) pointed out, no conclusion can be taken from these data concerning the idea that French listeners apply a syllabic segmentation procedure to the continuous speech input in everyday contexts. Second, no lexical access attempt is required to perform the fragment detection task⁶. More importantly, as it will be discussed next, the two tasks actually reflect important conceptual differences between the underlying assumptions of the MSS and the original conception of word segmentation that led to develop the fragment detection task.

1.3. From Uncovering the Nature of Classification Units to Determining the Cues Indicating Word Boundaries

Direct comparison of Romance and Germanic languages using the same technique is rather scarce. Actually, Cutler and colleagues (reported in Cutler & Norris, 1988, p. 114) did perform such a comparison by examining the English alternative to the syllable hypothesis, which would consist in classifying the speech input into feet⁷. Yet, they did not observe the corresponding effects: for English speakers fragment detection was not faster when the target, for example *GAR*, corresponded exactly to a foot, as in *GARGOYLE*, which includes two feet, than when it was smaller than a foot, as in *GARgle*, which constitutes one foot (but see Echols et al., 1997).

The word-spotting task was then designed "to put on a test, in a way that directly measures speech recognition processes, the hypothesis that segmentation for lexical access occurs at strong syllables" (Cutler and Norris, 1988, p. 114). This citation illustrates how the focus of research shifted from uncovering the nature and size of the prelexical unit(s) towards investigating the cue(s) that may best indicate where word boundaries are likely to occur in the speech stream. Indeed, the fragment detection technique was aimed at uncovering the classification representations that are computed from the

⁶ Note that even if all the (visually presented) syllabic targets used in the study of Mehler et al. (1981) were in fact words (e.g. /ba/, meaning *low* or *sock*; /bal/, meaning *ball*, etc.), only three of them actually corresponded to orthographic word forms (*pal*, *car*, *bal*).

⁷ In English the foot is a rhythmic unit that contains a strong syllable plus, optionally, one or more following weak syllables.

auditory input to contact the lexical representations (e.g., Frauenfelder & Tyler, 1987). Under this view, a prelexical representation of the signal is constructed as a sequence of specific units (e.g., feet, syllables, etc.). On the contrary, the word-spotting task was aimed at testing the MSS, which is a segmentation device that does not investigate classification processes. Since its role is merely to indicate *where* in the speech stream lexical access must be initiated, the MSS is compatible with models of speech perception involving classification as well as with models involving no prelexical units at all.

Despite the different nature and focus of the fragment detection and word-spotting tasks, the syllabic segmentation strategy and the MSS were repeatedly considered as equivalent (although language-specific) solutions applied to the word boundary problem (Cutler et al., 1992, 1997). It was indeed argued that even if the fragment detection task does not directly address the issue of segmentation, the classification of the speech input into any set of units is logically entailed by a segmentation process at the boundaries of these units (Norris & Cutler, 1985; Cutler & Norris, 1988).

If syllables were important segmentation cues in French, syllabic information should help French listeners to detect word boundaries efficiently in fluent speech. This is what has been shown for metrical cues to support the MSS. This model has been underpinned by computer simulations showing the independent contribution of metrical cues in predicting the accurateness of the speech recognition process (see Norris et al., 1995; Cutler, Norris & McQueen, 1996) and by distributional analyses performed on the English and Dutch vocabulary showing a systematic relationship between metrical patterns and word boundary locations (Cutler & Carter, 1987; Cutler & McQueen, 1995; Vroomen & de Gelder, 1995; see also Cutler, 1994). These analyses led to the conclusion that the false alarm rate of the MSS would "be low in comparison with a lexical segmentation procedure that considered each phoneme or syllable to be a potential word onset location" (Cutler & Norris, 1988, p. 114). As has been already discussed, the same arguments were used to support the SBS (e.g., Mattys, 2000). There is to our knowledge no similar argument as regards the syllable in French or other Romance language. On the contrary, a rough look at the distribution of French words according to their number of syllables reveals that no less than 93% of the 37000 words included in the BRULEX database (Content, Mousty, & Radeau, 1990) are polysyllabic, the majority being di- or tri-syllabic. When only the 1000 most frequent words of this database are taken into account, about two third of them include at least two syllables. Thus, a syllabic segmentation procedure in French would lead to a huge number of erroneous lexical attempts. In fluent speech, a pure syllabic segmentation strategy would sometimes run into trouble even for monosyllables: because of fre-

quent phenomena like *resyllabification* and *liaison*⁸, syllabic boundaries do not always coincide with word boundaries (e.g., Dejean de la Batie & Bradley, 1995).

One is thus faced with a rather asymmetrical account of speech segmentation, since a syllabic segmentation procedure would be obviously less efficient for French than the MSS (or SBS) for English and Dutch. Does one have to conclude that French is intrinsically harder to parse than English or Dutch? We know of no evidence suggesting that this is the case. Rather, as pointed out by Kolinsky (1998), we should question the notion that cues used for segmentation are necessarily isomorphic to classification units. Alternatively, we hypothesise that segmentation in French is triggered by (at least partially) other cues than syllable boundaries, even if the classification format may be syllabic. In the next section, it will be argued that French listeners exploit stress cues in locating word boundaries, as was already suggested for Dutch and English listeners.

1.4. *Rhythmic Regularities in French*

French has final fixed stress. More precisely, syllables are grouped into right-headed prosodic units that generally do not exceed two or three syllables (Wenk & Wioland, 1982). Consequently, most polysyllabic words with non-schwa final syllable bear stress on their last syllable. Stressed syllables are mainly characterised by a sizeable lengthening in comparison with unstressed syllables (Flechter, 1991; Garde, 1968; Tranel, 1987; Vaissière, 1983, 1991; Wenk & Wioland, 1982). Although this final lengthening is often accompanied by a falling or rising F0 movement, it is perceptually salient because the durational increase widely exceeds the perception threshold of duration differences (Rossi, 1972). This led Wenk and Wioland (1982) to characterise French as being "trailer-timed" rather than "syllable-timed": its basic rhythmic unit is an iambic foot characterised by a short-long durational pattern.

As Cutler et al. (1997) proposed, "it might be imagined that fixed stress could provide an excellent cue to word-boundary location" (p. 146, see also Cutler, 1990; Vaissière, 1983; 1991). Such a segmentation device based on

⁸ Resyllabification occurs when a final-word consonant becomes the onset of the syllable of a following word beginning with a vowel (e.g. *par ici* syllabified *pa.ri.ci*). Liaison is a particular instance of resyllabification within which the first word ends in a normally silent consonant, like the /t/ of the French word *petit*. This consonant surfaces when it is followed by a vowel-initial word (like in *petit air*), resulting in the resyllabification of the surfaced latent consonant over word boundaries (*pe.ti.tair*).

stress may even be more effective in fixed-stress languages than in variable-stress languages like English or Dutch, in that it constitutes a determinist cue for word boundary location only in the former case.

Several studies have provided support for the idea that stress helps word boundary localisation in French. A pioneering research by Rietveld (1980) showed that French listeners produce reliable suprasegmental contrasts when asked to pronounce phonologically ambiguous sequences like "*le couple est complet*" (*the pair is complete*) vs. "*le couplet complet*" (*the complete verse*). At the durational level, the critical sequence presented a trochaic (long-short: /ku:plɛ/) or an iambic (short-long: /kuplɛ:/⁹) pattern, respectively. In a subsequent perception experiment in which French listeners were asked to choose between the two semantic interpretations, the durational contrasts provided the best predictor of the participants' responses. Banel and Bacri (1994) addressed more directly the use of syllabic lengthening as word boundary marker in French. When asked report the number of perceived words in ambiguous disyllabic strings, French listeners more often perceived a single word (e.g., /kɔrbol/, *crow*) when the strings presented the typical French iambic pattern, but two words (e.g., /kɔr/, *body*, and /bo/, *beautiful*) when the strings presented the unusual trochaic pattern. In a word-spotting task, Banel and Bacri (1997) further observed faster detection times of monosyllabic words (e.g., *lampe*, /lāp/) in trochaic sequences (/lā:pzɔk/) than in iambic sequences (/lāpzɔ:k/). In the same vein, Banel, Frauenfelder and Perruchet (1998) showed better recognition performance in the learning of an artificial language for trisyllabic "words" that displayed the typical iambic pattern of French word (short-short-long) than for "words" that displayed a trochaic (long-short-short) or neutral (long-long-long) rhythmic pattern.

Taken together, these studies support the idea that French listeners treat lengthened syllables as word-offset markers, thus postulating a word boundary after stressed syllables. However, the durational contrasts used by Banel and co-workers reproduced the ratio observed between syllables of words pronounced in isolation. Such differences (e.g., 520 milliseconds between short and long syllables) are less likely to occur in connected speech where temporal contrasts are far less contrasted (e.g., Klatt, 1975).

To further examine the use of syllabic lengthening in the segmentation of continuous and natural speech in French, we presented French listeners with Cutler and Butterfield's (1992) "juncture misperception" method (Goetry, Van de Velde, & Kolinsky, in progress). Participants were asked to write down what they thought they had heard of part of sentences presented at

⁹ The symbol : indicates vowel lengthening.

individual speech perception threshold. The relationship between the rhythmic pattern and the word boundaries of identical phonemic strings (e.g., /livrɛ/) was manipulated across sentences. In one type of sentence (henceforth, *trochaic 2-words*), the sequence included two (parts of) words and formed a trochaic pattern, e.g., /li:vr#ɛ/¹⁰ in "j'ai vu que les livres, essentiels à notre époque, se vendaient très mal" ("I saw that the books, essential in our epoch, were sold very badly"). In another sentence type (henceforth, *iambic 2-words*), the phonemic string also contained two words but formed an iambic pattern, e.g., /li#vrɛ:#/ in "j'ai vu que les lits vrais, sans cesse vantés, sont plus confortables que ces paillasses" ("I saw that the true beds, unceasingly praised, were more comfortable than those pallets"). In a third sentence type (henceforth, *iambic 1-word*), the sequence contained a single disyllabic word and displayed an iambic pattern, e.g. /livrɛ:#/ in "j'ai vu que les livrets, censés détenir les comptes du commerce, étaient vides" ("I saw that the booklets, supposed to contain the accounts of the firm, were empty"). According to the idea that in French a trochaic pattern should trigger segmentation while a iambic pattern should not, we predicted that the sequences including two words would be correctly segmented when trochaic but would be often misperceived as a single disyllabic word when iambic. In addition, few segmentation errors were expected for the iambic sequences including one word, since their stress pattern should not trigger segmentation. The results confirm largely these predictions by showing that the French listeners produced much *more* segmentation errors for the iambic 2-words sequences than for the two other sequence types. This supports and extends the idea that stress cues assist speech segmentation in French by showing that rhythmic effects emerge for temporal contrasts displayed in continuous and natural speech.

1.5. *Rhythmic Regularities in Dutch: A Direct Comparison With French*

The basic iambic rhythmic structures of French words differs fundamentally from those of "leader-timed" languages (cf. Wenk & Wioland, 1982) like Dutch or English, in which stress usually falls on the initial syllables of words (for Dutch, see Vroomen & de Gelder, 1997b; Vroomen et al., 1998; for English, see Mattys, 2000). These languages present basic trochaic, left-headed, rhythmic structures. As already mentioned, direct evidence for the use of stress as a word-onset marker has been provided in Dutch (Vroomen & de Gelder, 1997b, submitted; Vroomen et al., 1998).

¹⁰ The symbol # indicates a word boundary.

Specific segmentation procedures across languages should be assessed with similar tasks (see sections 1.2 and 1.3). The fact that French and Dutch present contrasted stress structures provides a very good testing ground for examining the notion that listeners develop specific segmentation devices upon the recurrent rhythmic or stress regularities displayed in their native language (Cutler, 1985; Cutler et al., 1983, 1986, 1992). To this aim, we used a cross-linguistic design in which Dutch listeners were presented with a Dutch material matched to the French one (Goetry et al., in progress). If Dutch listeners use syllabic lengthening (together with other stress cues) as a word-onset marker, they should display opposite segmentation patterns for the three sequence types compared to those observed for the French listeners. That is, the Dutch listeners should produce much less segmentation errors for the iambic 2-words than for the trochaic 2-words sequences, as well as many segmentation errors for the iambic 1-word sequences. Indeed, the more prominent second syllable of the iambic 2-words sequences was expected to correctly trigger segmentation, thus leading to the correct detection of the medial word boundary. On the contrary, the less prominent second syllable of the trochaic 2-words sequences should not trigger segmentation and thus should be perceived together with the preceding one as a single disyllabic word. Likewise, we predicted the more prominent second syllable of the iambic 1-word sequences to erroneously trigger segmentation and induce the perception of two monosyllabic words.

The results computed on both the total set of responses and the relative proportions of segmentation errors corroborated these predictions. As a matter of fact, they showed the reversed rhythmical effect for the two sequence types containing two words, as compared to what was observed for the French listeners. Moreover, about 85% of the segmentation errors made by the Dutch listeners were distributed across the trochaic 2-words and iambic 1-word sequences, in which the rhythmic patterns are incongruent with word boundaries.

A control experiment showed that the difference between the French and Dutch listeners could not be attributed to some acoustical mismatch between the two material sets. Indeed, another group of French listeners presented with the Dutch material displayed similar segmentation patterns as those observed for the French participants tested in their native language, thus differing significantly from the Dutch listeners presented with the same Dutch material.

Taken together, these cross-linguistic results show that rhythmically matched materials lead to opposite segmentation patterns in French and Dutch listeners, thus suggesting that stress cues are exploited in an opposite way in these two languages. In French, a "trailer-timed" language, disyllabic sequences are interpreted as a single word when they display an iambic

pattern but as two monosyllabic words when they display a trochaic pattern. In a "leader-timed" language like Dutch, stress cues would indicate the beginning of words, thus leading to perceive a disyllabic trochaic sequence as a single lexical item but an iambic sequence as two monosyllabic words.

Since the different rhythmic structures of French and Dutch seem to induce native monolingual listeners to use opposite stress-based segmentation strategies, one may wonder what happens in bilinguals who master these two languages: may opposite strategies coexist in the same individual? This issue will be discussed in the final part of the present paper.

II. Speech Segmentation in Bilinguals

Although one may consider that "genuine" bilingualism refers to a native and equal competence in more than one language (e.g., Thiery, 1976, 1978), most researchers consider different degrees of bilingualism to range along a continuum (e.g., Baetens Beardsmore, 1986; de Groot & Kroll, 1997; Grosjean, 1982; Romaine, 1995; Schaerlaekens, 1998). It should be noted that whereas few (if any) people represent the bilingual extremity of this continuum (see e.g., Flege, Yeni-Komshian, & Liu, 1999; Pallier, Sebastián-Gallés & Colomé, 1999), the opposite extreme, that is, pure monolingualism, is no more representative of the human linguistic knowledge. Indeed, it has been estimated that more than one person out of two uses at least two languages in everyday interactions (Grosjean, 1982; Harris & Nelson, 1992). Any general model of speech processing should thus be able to account for the understanding of more than one language.

Paradoxically, studies relating speech segmentation to bilingualism are scarce, and, as it will be argued, most of them provide us with an incomplete picture of speech processing in bilinguals since they focus on the availability of a syllabic segmentation procedure in these listeners.

II.1. The Hypothesis of Restricted, Mutually Exclusive, Segmentation Routines

To our knowledge, Cutler, Mehler, Norris and Segui (1992) conducted the only study that explored the availability of specific segmentation procedures in bilinguals' two languages. These authors presented French-English bilinguals with the experimental situations that had been used previously to demonstrate syllabic parsing in French monolinguals (that is, fragment detection, Mehler et al., 1981) and metrical parsing in English monolinguals (that is, word-spotting, Cutler & Norris, 1988). The participants had learned

both languages from the earliest stages of acquisition, spoke both languages daily, and were considered by monolinguals as native speakers in each language.

The fragment detection results showed that neither for the English nor for the French material did the bilingual group as a whole produce results analogous to those of either previously studied monolingual group. To look further for a reflection of their findings with monolinguals, the authors subdivided the bilinguals according to several criteria - including their country of residence, parents' native language, and preferred language -. The factor dividing the bilinguals according to their preferred language (called "dominant" by the authors), which amounted to a decision as to which of their two languages they would be most sorry to lose if they had to, produced interpretable data. Indeed, the "English-dominant" bilinguals showed no syllabic effect in either language, while the "French-dominant" bilinguals produced a syllabic effect in French but not in English. This was interpreted as evidence for the development of a restricted segmentation procedure based on the bilinguals' "dominant" language, in this case, syllabic for the "French-dominant" bilinguals and non-syllabic for the "English-dominant" bilinguals. Cutler et al. found confirmation of the idea that bilinguals develop an adapted segmentation procedure restricted to their "dominant" language with the English word-spotting experiment. In this task, significantly faster target detection in the strong-weak than in the strong-strong items, indicative of the use of the MSS, was observed for the "English-dominant" but not for the "French-dominant" bilinguals.

From the entire set of results, Cutler et al. (1992) concluded that the restricted speech segmentation procedures are mutually exclusive. In other words, bilinguals would behave functionally as monolinguals in some aspects of their processing, hence developing a single restricted segmentation procedure adapted to their "dominant" language.

In the next sections, the two main issues related to Cutler et al.'s (1992) influential claims will be examined more closely, namely the relationship between language dominance and the restricted segmentation procedure (section II.2.) and the mutual exclusivity of the restricted segmentation procedures (sections II.3, II.4 and II.5).

II.2. Is Language Dominance Actually Predictive of the Restricted Segmentation Procedure?

Cutler et al. (1992) were more interested in knowing whether balanced bilinguals can process their two languages in the same way as the respective monolingual groups than in characterising their absolute competence levels.

These authors nevertheless attempted to establish whether the participants' way of responding in the two languages was related to their language dominance. However, we have no idea of the relationship that may exist between the actual competence-related language dominance of these bilinguals and Cutler et al.'s classification criteria assessing the bilinguals' "dominance", which was based on a single forced-choice preference question. Moreover, the participants examined by Cutler et al. were very reluctant to state any preference for one language over the other. In fact, they often claimed to prefer French for some purposes but English for other purposes, hence the necessity to present a forced-choice question (see also Kearns, 1994, for similar concerns). As Grosjean (1998) pointed out, one does not know on which criteria the participants have answered to this difficult question, and, therefore, what kind of variables underlay the different patterns of results observed in the "French-dominant" and "English-dominant" subgroups.

Consequently, the replication of these results with other participants may be very difficult. Kearns (1994, Experiment 5) examined French-English balanced bilinguals comparable to those tested by Cutler et al. (1992), using similar fragment detection tasks and classifying them on the basis on the same question. Surprisingly, she observed a syllabic effect for the French material in the "English-dominant" subgroup but not in the "French-dominant" subgroup (neither of the two groups syllabified the English material). Kearns further reasoned that dominance in balanced bilinguals might be better assessed on the basis of the participants' linguistic background and linguistic habits rather than on the basis of their preferred language. She computed for each bilingual an "Englishness score" derived from a questionnaire related to the acquisition and use of the two languages. Contrary to what was predicted, correlational analyses showed a *positive* relationship between this Englishness score and the syllabic effect in French. Multiple regressions revealed that the first spoken language was the best predictor of the syllabic effect in French: the bilinguals were more likely to show a syllabic effect if they had spoken English first. Further analyses performed on the French data showed a clear syllabic effect for the bilinguals who had spoken English first, but not for those who had spoken French first.

Thus, even if balanced bilinguals seem to separate into two groups of which only one shows a syllabic effect in French, the "dominant" language - as defined either in terms of preferred language or in terms of "most native" language - is not predictive of the use of a syllabic segmentation procedure. What would induce only some balanced bilinguals to parse the French input into syllabic units has so far escaped clear understanding.

This problem is further complicated by the extreme difficulty to assess the multi-dimensional nature of dominance in this population, if dominance is a meaningful concept at all for such bilinguals (e.g., Grosjean, 1982; Kearns,

1994; Romaine, 1995; Schaerlaekens, 1998). As Schaerlaekens (1998) pointed out, "bilingualism is a loose concept that covers many quantitative variants, and within each variant there are also many qualitative nuances" (p. 131). While traditional competence-based measures of dominance may be unsuitable for early balanced bilinguals, some authors have argued that differences in performances in the two languages may nevertheless be revealed under specific conditions. For example, Baetens Beardsmore (1986) and Dornic (1978, 1981) both pointed out that stress and fatigue can show differences in ease of use of two languages even in early balanced bilinguals. Kearns (1994, Experiments 6 and 7) showed that early balanced bilinguals selected on the same criterion as those used by Cutler et al. (1992) displayed unequal performances in their two languages, as compared to monolinguals, when required to recognize speech under difficult listening conditions (i.e., faint speech or speech in noise). Recent studies on early bilinguals have reported subtle differences between their performances in their two languages as regards both phonetic perception (Bosch, Costa, & Sebastián, 2000; Pallier, Bosch, & Sebastián, 1997; Sebastián-Gallés & Soto-Faraco, 1999) and lexical activation (Pallier et al., 1999). For example, Pallier et al. (1999) examined the abilities of early Spanish-Catalan bilinguals to identify words differing on a vocalic contrast which exists in Catalan (i.e., /e/ vs. /ɛ/) but not in Spanish (which has only /e/). In a lexical decision task, the Spanish-dominant bilinguals treated Catalan minimal pairs like /netə/ (*grand-daughter*) - /netə/ (*clean*) as if they were homophones. That is, they showed a priming repetition effect, suggesting that they did not correctly perceive the Catalan vocalic contrast. By contrast, the Catalan-dominant bilinguals did not show any priming repetition effect for these minimal pairs. Interestingly, no significant overall difference in reaction times and error rates was observed between the two bilingual groups, which thus displayed equivalent competence at the vocabulary level.

It thus seems possible to reveal fine-grained differences between the performances in the two languages of early bilinguals. Yet, neither Cutler et al. (1992) nor Kearns (1994, Experiment 5) did present their bilingual participants with tests assessing competence in their two languages. It seems worth to examine in further bilingual studies whether or not there is a relationship between the performance level of bilinguals in their two languages and the monolingual-like speech segmentation procedure they may apply to these languages.

Whether defined in terms of competence or in terms of "nativeness", the dominant language of many other (unbalanced) bilinguals, who have acquired their second language many years after their native language, is much less problematic to determine. Bradley, Sanchès-Casas and García-Albea (1993) turned to native Spanish speakers who subsequently learned

English through immersion in an English-speaking community for a long period (namely, for 18 years on the average). According to the idea that listeners' segmentation routines depend on the rhythmic regularities encoded during childhood (cf. e.g., Cutler & Mehler, 1993), these unbalanced bilinguals were predicted to show a syllabic effect in their native language. Unexpectedly, this group showed no trend to syllabify Spanish. It is difficult to accept the idea that Spanish-English bilinguals abandon their native processing routine when we have no evidence that they have ever applied this routine to their native input. Even if a syllabic effect had been observed with the same material in Spanish monolinguals (Bradley et al., 1993), other experiments have not replicated these results unless listeners were artificially slowed down (Sebastián-Gallés et al., 1992). Moreover, the bilinguals' results of Bradley et al. have been replicated in another study conducted by Kearns (1994, Experiment 4) on French native speakers acquiring English as second language. That is, these listeners did not show any sign of syllabification of the French material in a fragment detection task.

Thus, one is still left with a puzzle. What seems clear is that dominance is not a good predictor of the use of a syllabic segmentation procedure, since late unbalanced bilinguals whose native language is Spanish or French seem not to use a syllabic segmentation strategy to process that language. As Kearns (1994) pointed out, more variables may influence the way in which bilinguals process speech, in comparison to monolinguals, since the former have a wider variety of linguistic knowledge to apply to the input than the latter. For example, bilinguals may not have exactly monolingual-like phonemic categories (Elman, Diehl, Buchwald, 1977; Flege & Eefting, 1987; Hazan & Boulakia, 1993) and phonological representations (Bosch et al., 2000; Sebastián-Gallés & Soto-Faraco, 1999; Pallier et al., 1997; 1999). According to Kearns (1994), this may hold true for other factors that influence the phonological information about syllable boundaries in the speech stream. For example, within the frame of parameter-setting (e.g., Roeper & Williams, 1987), Flege (1988) argued that bilinguals establish mean values between the syllabic structures of the first language and those of the target language, which would afford them to display a single system adapted to both languages. The bilinguals' different amount of phonological knowledge, in comparison with monolinguals, may have been responsible, at least partly, for the puzzling picture of results obtained with the fragment detection task. It is therefore difficult to draw conclusions regarding the basic mechanisms underlying speech segmentation in bilinguals on the basis of these sole results.

11.3. What is the Evidence for Arguing That Restricted Segmentation Procedures Are Mutually Exclusive?

Neither Bradley et al. (1993) nor Kearns (1994) did address the issue of speech processing in the bilingual participants' second language. Thus, these two studies do not provide information regarding Cutler et al.'s (1992) strong claim that bilinguals develop a single restricted segmentation procedure, i.e. that speech segmentation procedures are mutually exclusive. This argument was based on the fact that, in Cutler et al.'s study, the bilinguals who fulfilled their criterion for French "dominance" syllabified the French fragment detection material but did not show evidence of a stress-based segmentation in the English word-spotting experiment, while the bilinguals who fulfilled their criterion of English "dominance" did not syllabify the French fragment detection material but showed evidence for the use of a MSS in the English word-spotting task. However, it should be noted that, in this study, only six participants (out of 24) who took part in the fragment detection study also took part in the word-spotting study. In other words, for most of the sample, different bilingual participants have been examined with the experimental situations assessing the use of a specific segmentation strategy in French and English. It thus could be the case, as suggested by Kearns (1994), that at least some of the bilinguals who were showing evidence for syllabification in French may also have shown evidence of stress-based segmentation in English, and vice versa. This possibility will be further discussed in the two next sections, in the light of some studies conducted on language development in bilingual infants as well as additional speech segmentation data collected on bilingual adults.

11.4. Against Mutual Exclusivity of Restricted Segmentation Procedures: Language Development in Early Bilinguals

Cutler et al. (1992) claimed that individuals learning two languages simultaneously from birth develop a single restricted segmentation procedure based on the rhythmic regularities of their "dominant" language. If this were the case, one would expect to observe an asymmetrical development of the knowledge of the two languages in bilingual infants, the "dominant" language being prioritised over the other. Bosch and Sebastián-Gallés (in press) did not find support for this view. Using the familiarisation-preference procedure (Jusczyk & Aslin, 1995), they observed that four-month-old infants from bilingual Catalan-Spanish environments were able to discriminate between Spanish and Catalan, even though these two languages share many prosodic features. These results suggest an early capacity to distinguish lan-

guages in simultaneous bilingual exposure.

Moreover, bilinguals do not differ from the "monolingual" infants in showing within-rhythmic class language discrimination abilities by four or five month of age (Bosch & Sebastián-Gallés, 1997; see also Nazzi, Jusczyk, & Johnson, 2000). This suggests that early, simultaneous, exposure to two languages does not delay language discrimination and lexicon pre-compiling abilities. When subsequently faced with the necessity of compiling two lexicons simultaneously, bilingual babies seem to progressively understand and speak their two languages without any difficulty (for reviews, see e.g. de Houwer, 1990; Genesee, Nicoladis, & Paradis, 1995).

According to the notion that the languages rhythmic regularities play a determinant role in lexical acquisition, it is reasonable to predict that bilingual infants need to acquire appropriate rhythmic segmentation cues to parse their two languages. Cutler et al.'s (1992) claim that bilinguals would develop a single restricted segmentation procedure fails to explain how these infants are able to successfully develop two lexicons simultaneously.

II.5. A Direct, Within-Subject But Between-Languages Comparison: French and Dutch Segmentation in Bilingual Adults

As Cutler (1994) pointed out, "the scale of the segmentation problem in the structure of the input is remarkably similar for the infant and for the adult" (p. 87). In other words, adults acquiring a second language are faced with the same auditory input as the infants, from which they must pick out discrete chunks that have to be mapped onto stored representations for recognition. As already argued, universal processes such as implicit segmentation or competition, when considered alone, may not be powerful enough to allow a correct retrieval of the lexical units from the highly variable acoustic signal (Frauenfelder, 1991b; Norris et al., 1995). Nevertheless, language learners become progressively successful at recognising auditory words from their non native language, and some of them become highly proficient bilinguals even if they have acquired their second language post-puberty (e.g., Birdsong, 1992; 1998; Coppieters, 1982).

On the basis of the importance of rhythmic cues for speech segmentation in both infants and adults, one may hypothesise that all language learners, including bilinguals, must progressively develop rhythmic solutions adapted to the target language, which may be at least partially similar to the one used by monolingual speakers. Indeed, if some bilinguals were reaching very high proficiency levels in a second language without the help of these rhythmic cues, there is no *a priori* reason to believe that monolinguals need to use these cues, or even that these cues play any role in speech segmentation.

Such a view is at odds with Cutler et al.'s (1992) proposal that bilinguals do not develop a restricted segmentation procedure to parse their "non-dominant" language. However, as already mentioned, the evidence on this issue is limited to Cutler et al.'s study in which different bilingual participants took part in the fragment detection and word-spotting experiments. As acknowledged by the authors themselves (p. 407), the different processes tapped by the two tasks (see sections I.2. and I.3) further complicate the comparison between the two sets of results.

This is why we tried to address the issue of the coexistence of specific rhythmic segmentation cues in bilinguals by presenting them with identical tasks in their two languages (Goetry et al., in progress). We presented French-Dutch bilinguals, all French-dominant, with *both* the French and the Dutch materials that had induced opposite segmentation patterns in monolinguals tested in their respective native language (see sections I.4 and I.5). We reasoned that the bilinguals should have developed adapted segmentation cues to correctly segment their two languages. Yet, since French and Dutch monolinguals use stress cues in an opposite way, in bilinguals it may be the case that the stress-based strategy used in the (French) dominant language interferes with the use of an adapted strategy in the (Dutch) non-dominant language. As we expected this influence to be greater if Dutch is acquired later on, we tested two groups of bilinguals: early bilinguals who had acquired French and Dutch before the age of four, and late bilinguals who had French as native language and had acquired Dutch during adolescence.

The results for the French materials showed no significant difference between the segmentation patterns of the two bilingual groups and those of the French monolinguals. That is, both bilingual groups produced more segmentation errors for the iambic 2-words sequences than for the two other sequence types. This is coherent with the idea that bilinguals and monolinguals relied to a similar extent on the typical rhythmic structures of French to segment that language.

For the Dutch materials, the two groups of bilinguals *also* displayed similar segmentation patterns as those found for the Dutch monolinguals. These results suggest that both bilingual groups were highly familiar to the typical trochaic rhythmic structures of Dutch words and exploited these rhythmic cues to locate word boundaries in the same way as the Dutch monolinguals did. This is even more remarkable if we consider the fact that all the bilinguals, although displaying general listening abilities similar to those of Dutch monolinguals (as assessed by a speeded listening comprehension task), were nevertheless clearly French-dominant. This was documented by the participants' self-ratings as well as by their lower recognition rates of the experimental sentences (presented at speech perception threshold) and their

lower performance in a lexical decision task, as compared to the Dutch monolinguals.

The surprising absence of difference between early and late bilinguals' segmentation patterns for the Dutch material may be related to the fact that the late bilinguals displayed Dutch listening comprehension abilities roughly equivalent to those of the early bilinguals, both in normal and in difficult listening conditions (assessed through a speeded listening comprehension task and through the recognition rate of the experimental sentences, respectively). This absence of difference between early and late bilinguals may reflect the ability of both groups to rely on adapted rhythmic cues. While further investigation is required, our results thus suggest that the use of rhythmic cues may be little affected by age of acquisition for language learners who have reached a certain proficiency threshold in that language.

Sanders, Yamada and Neville (1999) provided convergent evidence supporting the idea that, in highly proficient bilinguals, the late acquisition of a second language does not prevent them from exploiting the rhythmic regularities of that language. These authors compared the reaction times and event-related potentials (ERPs) of English monolinguals and Japanese speakers who learned English after the age of 12, but who became highly fluent in English. The participants were presented with normal English sentences or sentences in which words were replaced with pseudowords displaying similar stress patterns as the original words. They were asked to detect target phonemes as well as to locate their position (word-initial vs. word-medial). Both groups of listeners detected initial and non-initial target phoneme (e.g., /n/) significantly faster in words and pseudowords when these displayed the typical English strong-weak stress pattern (e.g., *NEctar*, *WITness*) than when these displayed the infrequent weak-strong stress pattern (e.g., *neGLECT*, *igNITE*). This indicates that the bilinguals were able to exploit the English stress pattern to perform the detection task, even though they had not been exposed to it before the age of twelve. Thus, monolinguals and bilinguals seem to rely on the same prosodic cues, although the ERPs suggest that bilinguals did so using different neural systems than English native speakers.

The notion that attained proficiency may be more important than age of acquisition in bilinguals' second language processing has also received support from a series of PET studies. Perani et al. (1998) examined two groups of highly proficient bilinguals: Italian-English bilinguals who acquired their second language after the age of 10, and Spanish-Catalan bilinguals who acquired their second language before the age of four. The cortical responses of these bilinguals listening to stories were highly similar for their first and second language, regardless of age of acquisition. This was not the case for low proficiency, late Italian-English bilinguals, who showed very differ-

ent patterns of cortical activity when listening to stories in the first vs. second language (Perani et al., 1996; see also Dehaene, Dupoux & Mehler, 1997).

Taken together, the three sets of results (Goetry et al., in progress; Perani et al., 1998; Sanders et al., 1999) thus show little effect of age of acquisition on some aspects of second language processing. Yet, it should be noted that these results do not question the notion that age of acquisition is a major determinant of attained proficiency in the second language of bilinguals (e.g., Johnson & Newport, 1989). Rather, they suggest that when proficiency in the second language is kept constant, age of acquisition *per se* does not have an impact on second language processing and (macroscopic) brain representations.

Nevertheless, not all aspects of the linguistic knowledge of a second language are equally easy to master, and age of acquisition may have different effects on different types of linguistic abilities. For example, whereas lexical acquisition and semantic processing seem to occur normally even in late learners (e.g., Long, 1990; Weber-Fox & Neville, 1996), other aspects of the linguistic knowledge show stronger effects of late acquisition. This would be the case for accent (Oyama, 1976; Flege, Munro, & McKay, 1995), unknown phonemic contrasts (Bosch et al., 2000; Pallier et al., 1997, 1999; Sebastián-Gallés & Soto-Faraco, 1999), and complex grammatical structures (Johnson & Newport, 1991; Newport, 1990; but see Flege et al., 1999, for an alternative account). Similarly, it might be the case that age of acquisition is more critical for acquiring segmentation cues other than the rhythmic one examined in our study. What our results only suggest is that, once a high level of proficiency has been reached in a second language, the learner would be sufficiently attuned to the typical stress pattern of that language to be able to exploit it in speech segmentation.

General Discussion

The central problem of spoken word recognition is to understand how listeners segment continuous speech into discrete portions so efficiently despite the lack of reliable acoustic cues signalling the beginning of words.

Two different proposals have been made regarding the information that assists the speech segmentation processes. The first involves universal mechanisms like lexical competition (e.g., Norris, 1994). However, lexical competition fails as a segmentation procedure in many contexts (e.g. for very short or embedded words), and is unable to explain how infants begin to extract meaningful units from the signal in the absence of lexical knowledge. The second involves some specific knowledge of the phonological and

rhythmic characteristics of the native language. According to this proposal, the listeners' speech segmentation routines are critically influenced by the basic rhythmic regularities of their native language (e.g. stress-, syllable- or mora-based). Recent studies have provided support for the notion that stress may also indicate word boundaries and/or initiate lexical access in languages like Dutch (Vroomen & de Gelder, 1997b, submitted) or English (Mattys, 2000; Mattys & Samuel, 1997, 2000).

Such language-specific adaptations may be traced back to an innate rhythmic sensitivity (Nazzi, Bertoncini, & Mehler, 1998; Ramus, Hauser, Miller, Morris, & Mehler, 2000; see also Mehler & Christophe, 2000), which is rapidly incremented by language-specific prosodic knowledge of the mother tongue (Bosch & Sebastián-Gallés, 1997; Nazzi et al., 2000). In English, the sensitivity to the predominant metrical and/or stress pattern of words has been shown to develop between six and nine months (Jusczyk et al., 1993; Mattys et al., 1999; Morgan, 1996; Turk, Jusczyk, & Gerken, 1995). By 7.5 months, English learners already rely on this information to extract whole words (rather than just the salient strong syllables) from fluent speech (Jusczyk et al., 1999).

As pointed out by several authors (e.g., Cutler, 1990; Cutler et al., 1997; Jusczyk et al., 1999; Vaissière, 1983; 1991) learners of languages other than English may also develop a stress-based segmentation strategy. As far as French is concerned, Jusczyk et al. (1999) suggested that "French learners could possibly use information about lengthening at the ends of words as a marker of word offsets" (Jusczyk et al. 1999, p. 201). Obviously, data on French infants' sensitivity to the stress pattern of their native language are critically lacking, and future research should address this issue.

In the present paper, we discussed studies conducted on adult French listeners which strongly suggest that one of the main acoustic correlates of stress, namely vowel lengthening, may act as a powerful word-offset marker in that language. In addition, we discussed data suggesting that, because of the differing rhythmic structures of their language, French and Dutch adult listeners display opposite segmentation patterns when presented with matched material sets in a similar experimental situation (Goetry et al., in progress). Similar stress cues (vowel lengthening) were used as word-initial boundary markers by Dutch listeners and as word-offset boundary markers by French listeners. This suggests that listeners of both languages use stress cues in locating word boundaries, although they rely on the rhythmic structure typical of their mother tongue, namely on the iambic (short-long) pattern for French listeners and on the trochaic (long-short) pattern for Dutch listeners. These results thus extend, with a direct cross-linguistic comparison, the idea that listeners adopt a universal rhythmic segmentation strategy exploiting the specific regularities of their native language.

The parallel drawn between the use of stress cues in French and Dutch does not deny the fact that stress has a different role in lexical access in these two languages. As a consequence of its fixed final position, stress has no distinctive function in French, so that coding stress in the representation of the words would be completely uninformative. Therefore, it is likely that stress is not represented lexically in French. By contrast, although rare, minimal pairs that only differ by their stress patterns can be found in both Dutch (e.g., *'voornaam* - first name - vs. *voor'naam* - respectively -) and English (e.g., *'forbear* vs. *for'bear*). At least in Dutch, stress has been shown to constrain lexical activation. For example, Cutler and van Donselaar (in press; see also Koster & Cutler, 1997) showed no facilitation from one member of a minimal stress pair to the other in auditory lexical decision, although reliable repetition priming occurred. In the same study, word-spotting results further showed that mismatching suprasegmental information reduced word activation (e.g., *mu'seum* received a greater degree of activation from the fragment *mu'zee* than from the fragment *'muzee*). In English, stress is rather redundant to vowel quality, and therefore would not constrain the initial stages of word activation (Cutler & van Donselaar, in press). For example, Cutler (1986) showed that presentation of either of the two members of a minimal stress pair, e.g. either *'forbear* or *for'bear*, primed lexical decision to words associated to both of them, like *ancestor* and *tolerate* (see also Bond & Small, 1983; Small, Simon, & Goldberg, 1988). Yet, other data suggest that stress information does play an important role in lexical activation in English. Using the migration paradigm (Kolinsky & Morais, 1996), Mattys and Samuel (1997) showed that the vowels of secondary stressed syllables were less likely to migrate in words than in matched nonsense words, whereas no such lexical effect was observed with the vowels of primary stressed syllables. This difference was interpreted as support for the idea that a primary stressed syllable would be processed more autonomously (i.e. with less assistance from the lexicon) than secondary stressed syllable, and thus that primary stress may play a critical role in lexical access in English (see also Mattys & Samuel, 2000). This issue thus clearly deserves further investigation.

In any case, the important point is that the role of stress in lexical access and its potential use as a word-boundary marker are two different issues. Such a view has also been held by Vrooomen et al. (1998), who examined word segmentation in Finnish. Finnish has fixed word stress on the initial syllable, and thus stress is probably not represented in the lexical representations of Finnish words. Nevertheless, using both word spotting and a task in which participants had to segment an artificial language, Vrooomen et al. showed that a stressed syllable is a reliable segmentation cue for Finnish listeners. Similarly, to our view the fact that stress is probably not part of the

lexical representations in French, while it is in Dutch, is not incompatible with the idea that the acoustic correlates of word stress can signal likely word boundaries in both languages.

Relying on rhythmic or stress cues alone would however sometimes lead to incorrect segmentations. For example, while vowel lengthening may constitute a much better word boundary predictor in French than a (pure) syllabic segmentation strategy, it would lead to missegment words ending with a schwa-syllable. Likewise, many Dutch (and English) words do not bear word-initial primary stress or do not begin with a strong syllable. In these cases, it seems likely that listeners might take advantage of other cues that can be indicative of likely word boundaries in the language input.

Indeed, as proposed for example by Church (1987), there are several other potential sources of language-specific information that listeners could draw on in segmenting words from fluent speech, such as allophonic, phonotactic, and distributional cues. Phonotactics refers to the constraints on the possible ordering of phonetic segments within morphemes, syllables and words in a language. Similarly, different phonetic realisations (allophones) of the same phoneme are often restricted in terms of the position that they can appear within a word. For example, /t/ will be aspirated at the beginning of English words but not at their end (e.g., Umeda & Cocker, 1974). The knowledge of co-occurrence relations between syllables may also be exploited to extract words from fluent speech (e.g., Brent & Cartwright, 1996). Recognising and segmenting words has been shown to be helped by these additional cues, including phonotactics (McQueen, 1998; McQueen & Cox, 1995; Vitevitch & Luce, 1999), distributional regularities (Saffran, Newport, & Aslin, 1996), vocalic harmony (Suomi, McQueen, & Cutler, 1997), and allophonic variation (Yerkey & Sawusch, 1993; see also the studies conducted by Frauenfelder et al. reported in Kearns, 1994, and by van Donselaar and Stoujesdijk, 1994, who showed that syllabic parsing may be critically affected by the phonetic make-up of the material).

Such an approach to the word boundary problem, relying on several sources of information that would be extracted and integrated from infancy on, might provide a realistic account of listeners' accurateness in recognising continuous speech (e.g., Mattys et al., 1999; Morgan & Saffran, 1995; Myers et al., 1996; Saffran, Newport & Aslin, 1996). Computational models even suggest that a multiple-cue integration approach is more powerful than what would be predicted on the basis of the individual contribution of each cue (Christiansen, Allen, & Seidenberg, 1997).

A major challenge for future research will be to characterise the time course at which these different cues become available and the way in which they are combined. The few adult studies in which multiple speech segmentation cues were systematically manipulated suggest that rhythmic cues are

prioritised: listeners would rely on other potential word boundary markers only when rhythmic cues are absent. For example, Banel and Bacri (1994) observed syllable frequency effects in parsing when stress cues were absent, i.e. in spondees displaying a long-long rhythmic pattern. Likewise, Vroomen et al. (1998) showed that vowel harmony was used for segmentation in Finnish only when stress cues were unavailable. The infants data are also compatible with this view. For example, Mattys et al. (1999) showed that, in case of conflicting rhythmic and phonotactic information, nine-month-olds preferentially relied on prosodic cues to locate word boundaries.

These investigations may also shed some light on the processes involved in the acquisition of a second language. In the second part of the present paper, we argued that the acquisition of a language-specific rhythmic strategy is a necessary component of second language learning. In particular, it was hypothesised that the typical stress pattern of words, which has been shown to provide a strong cue for word boundary location in monolinguals, should be used by second language learners to allow a certain level of proficiency in their second language. This hypothesis seems supported by our data (Goetry et al., *in progress*), which show that the typical word stress pattern of the second language seems to be actually exploited in speech segmentation by bilingual listeners who have attained a high proficiency level in that language. Future research should address this issue more systematically as well as the potential impact of proficiency level, for example by examining the sensitivity to the specific rhythmic structures of their second language of several groups of bilinguals displaying various levels of second language comprehension abilities.

We further suggested that, for highly proficient bilinguals, the use of rhythmic cues for speech segmentation in a second language may show weaker effects of age of acquisition than learning other aspects of the linguistic system characterising that language. Yet, this might hold true only for pairs of languages where similar rhythmic word boundary cues can be used. Given the considerable implications of this issue for foreign-language learning, it would be critical to determine exactly what aspects of language-specific segmentation strategies are (or are not) strongly depending on age of acquisition.

References

- Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
- Alloppenna, P. (1995). *A re-examination of the syllable's role in the on-line segmentation of English*. Unpublished doctoral dissertation, Brown University.
- Baetens Beardsmore, H. (1986). *Bilingualism: Basic principles*. Clevedon: Tieto.
- Banel, M.-H., & Bacri, N. (1994). On metrical patterns and lexical parsing in French. *Speech Communication*, 15, 115-126.
- Banel, M.-H., & Bacri, N. (1997). Reconnaissance de la parole et indices de segmentation métriques et phonotactiques. *L'Année Psychologique*, 97, 77-112.
- Banel, M.-H., Frauenfelder, U. H., & Perruchet, P. (1998). Contribution des indices métriques à l'apprentissage d'un langage artificiel. *Actes des XXIIIèmes Journées d'Etudes sur la Parole* (pp. 29-32). Martigny, Suisse.
- Bard, E. G., Shillcock, R. C., & Altmann, G. T. M. (1988). The recognition of words after their acoustic offset in spontaneous speech: Effects of subsequent context. *Perception & Psychophysics*, 44, 395-408.
- Bertinetto, P. M. (1989). Reflections on the dichotomy "stress" vs. "syllable-timing". *Revue de Phonétique Appliquée*, 91-93.
- Birdsong, D. P. (1992). Ultimate attainment in second language acquisition. *Language*, 68, 706-755.
- Birdsong, D. P. (1998). *New perspectives on the critical period hypothesis for second language acquisition*. Hillsdale, NJ: Erlbaum.
- Bond, Z. S., & Small, L. H. (1983). Voicing, vowel and stress mispronunciations in continuous speech. *Perception & Psychophysics*, 34, 470-474.
- Bosch, L., Costa, A., & Sebastián, N. (2000). First and second language vowel perception in early bilinguals. *European Journal of Cognitive Psychology*, 12, 189-221.
- Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in four-month-old infants from monolingual and bilingual environments. *Cognition*, 65, 33-69.
- Bosch, L., & Sebastián-Gallés, N. (in press). Evidence for early language discrimination abilities in infants from bilingual environments. *Infancy*.
- Bradley, D. C., Sanchez-Casas, R. M., & Garcia-Albea, J. E. (1993). The status of the syllable in the perception of Spanish and English. *Language and Cognitive Processes*, 8, 197-233.
- Brent, M. R., & Cartwright, T. A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition*, 61, 93-125.
- Bruck, M., Treiman, R., & Caravolas, M. (1995). Role of the syllable in the processing of spoken English: Evidence from a nonword comparison task. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 469-479.
- Christiansen, M. H., Allen, J., & Seidenberg, M. S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes*, 13, 221-268.
- Church, K. W. (1987). Phonological parsing and lexical retrieval. *Cognition*, 25, 53-69.
- Content, A., Mousty, P., & Radeau, M. (1990). Brulex. Une base de données lexicale informatisée pour le français écrit et parlé. *L'Année Psychologique*, 90, 551-566.

- Cooper, N. (2000). Native and non-native processing of lexical stress in English word recognition. *Proceedings of the Workshop on Spoken Word Access Processes* (pp. 163-166). Nijmegen, The Netherlands.
- Coppieters, R. (1987). Competence differences between native and fluent non-native speakers. *Language*, 63, 544-573.
- Cutler, A. (1985). Cross-language psycholinguistics. *Linguistics*, 23, 659-667.
- Cutler, A. (1986). Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech*, 29, 201-220.
- Cutler, A. (1990). Exploiting prosodic probabilities in speech segmentation. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing* (pp. 105-121). Cambridge, MA: MIT Press.
- Cutler, A. (1994). Segmentation problems, rhythmic solutions. *Lingua*, 92, 81-104.
- 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 and Language*, 2, 133-142.
- Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40, 141-201.
- Cutler, A., & McQueen, J. (1995). The recognition of lexical units in speech. In B. de Gelder & J. Morais (Eds.), *Speech and reading: A comparative approach* (pp. 33-48). Trowbridge, Wiltshire: Erlbaum (UK) Taylor & Francis.
- Cutler, A., & Mehler, J. (1993). The periodicity bias. *Journal of Phonetics*, 21, 103-108.
- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1983). A language specific comprehension strategy. *Nature*, 304, 159-160.
- Cutler, A., Mehler, J., Norris, D., & 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. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology*, 24, 381-410.
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 113-121.
- Cutler, A., Norris, D., & McQueen, J. (1996). Lexical access in continuous speech: Language-specific realisations of a universal model. In T. Otake & A. Cutler (Eds.), *Phonological structure and language processing: Cross-linguistic studies* (pp. 227-242). Berlin: Mouton de Gruyter.
- Cutler, A., & van Donselaar, W. (in press). Voornaam is not a homophone: lexical prosody and lexical access in Dutch.
- Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 11, 51-62.
- de Groot, A. M. B., & Kroll, J. F. (1997). *Tutorials in bilingualism: Psycholinguistic perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Dehaene, S., Dupoux, E., & Mehler, J. (1997). Anatomical variability in the cortical

- representation of first and second languages. *Neuroreport*, 8, 3809-3815.
- de Houwer, A. (1990). *The acquisition of two languages from birth: A case study*. Cambridge: Cambridge University Press.
- Dejean de la Batie, B., & Bradley, D. C. (1995). Resolving word boundaries in spoken French: Native and non-native strategies. *Applied Psycholinguistics*, 16, 59-81.
- Delattre, P. (1966). A comparison of syllable-length conditioning among languages. *IRAL*, 7, 295-325.
- Dornic, S. (1978). The bilingual's performance: Language dominance, stress, and individual differences. In D. Gerver & H. Sinaiko (Eds.), *Language interpretation and communication* (pp. 259-271). New York: Plenum Press.
- Dornic, S. (1980). Language dominance, spare capacity and perceived effort in bilinguals. *Ergonomics*, 23, 369-377.
- Dupoux, E. (1993). The time course of prelexical processing: The syllable hypothesis revisited. In G. T. M. Altmann & R. Shillcock (Eds.), *Cognitive models of speech processing* (pp. 81-114). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dupoux, E., & Mehler, J. (1992). Unifying awareness and on-line studies of speech. A tentative framework. In J. Alegria, D. Holender, J. Junça de Morais, & M. Radeau (Eds.), *Analytic approach to human cognition* (pp. 59-75). Amsterdam: North-Holland.
- Echols, C. H., Crowhurst, M. J., & Childers, J. B. (1997). The perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language*, 36, 202-225.
- Elman, J. L., Diehl, R. L., & Buchwald, S. E. (1977). Perceptual switching in bilinguals. *Journal of the Acoustical Society of America*, 62, 971-974.
- Fear, B. D., Cutler, A., & Butterfield, S. (1995). The strong/weak distinction in English. *Journal of the Acoustical Society of America*, 97, 1893-1904.
- Finney, S. A., Protopapas, A., & Eimas, P. D. (1996). Attentional allocation to syllables in American English. *Journal of Memory and Language*, 35, 893-909.
- Flechter, J. (1991). Rhythm and final lengthening in French. *Journal of Phonetics*, 19, 193-212.
- Flege, J. E. (1988). The production and perception of foreign language speech sounds. In H. Winitz (Ed.), *Human communication and disorders: A review* (pp. 224-401). Nordwood, NJ: Ablex.
- Flege, J. E., & Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, 6, 185-202.
- Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97, 3125-3134.
- Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, 41, 78-104.
- Frauenfelder, U. H. (1991a). Lexical alignment and activation in spoken word recognition. In J. Sundberg, L. Nord, & R. Carlson (Eds.), *Music, language, speech and brain: Wenner-gren International Symposium Series* (pp. 294-303). London: Macmillan.
- Frauenfelder, U. H. (1991b). Une introduction aux modèles de reconnaissance des

- mots parlés. In R. Kolinsky, J. Morais, & J. Segui (Eds.), *La reconnaissance des mots dans les différentes modalités sensorielles: études de psycholinguistique cognitive* (pp. 7-36). Paris: Presses Universitaires de France.
- Frauenfelder, U. H., & Content, A. (1999). The role of the syllable in spoken word recognition: Access or segmentation? *Proceedings of the IIIèmes Journées d'Etudes Linguistique* (pp. 1-8). Nantes: Université de Nantes.
- Frauenfelder, U. H., & Kearns, R. K. (1996). Sequence monitoring. In F. Grosjean & U. H. Frauenfelder (Eds.), *Spoken word recognition paradigms. A special issue of Language and Cognitive Processes* (Vol. 11, pp. 665-673).
- Frauenfelder, U. H., & Tyler, L. K. (1987). The process of spoken word recognition: An introduction. *Cognition*, 25, 1-20.
- Garde, P. (1968). *L'accent*. Paris: Presses Universitaires de France.
- Genesee, F., Nicoladis, E., & Paradis, J. (1995). Language differentiation in early language development. *Journal of Child Language*, 22, 611-631.
- Goetry, V., Van de Velde, H., & Kolinsky, R. (in progress). Rhythmic cues for speech segmentation in French, Dutch, and bilingual listeners: The role of stress.
- Goldman, J.-P., Content, A., & Frauenfelder, U. H. (1996). Comparaison des structures phonologiques syllabiques en français et en anglais. *Proceedings of the XXIIèmes Journées d'Etudes sur la Parole* (pp. 315-318). Avignon, France.
- Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Cambridge, MA: Harvard University Press.
- Grosjean, F. (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception & Psychophysics*, 38, 299-310.
- Grosjean, F. (1998). Studying bilingualism: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, 1, 131-149.
- Harris, R. J., & Nelson, E. M. M. (1992). Bilingualism: Not the exception any more. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 3-14). Amsterdam: North-Holland.
- Hazan, V., & Boulakia, G. (1993). Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech*, 36, 17-38.
- Hölne, B., & Schriefers, H. (1995). *Ambisyllabizität in Deutschen: Psycholinguistische Evidenz*. Paper presented at the Akten des 29. Linguistischen Kolloquiums. Tübingen: Niemeyer.
- Inagaki, K., Hatano, G., & Otake, T. (2000). The effect of Kana literacy acquisition on the speech segmentation unit used by Japanese young children. *Journal of Child Experimental Psychology*, 75, 70-91.
- Johnson, J., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60-99.
- Johnson, J., & Newport, E. L. (1991). Critical period effects on universal properties of language: The status of subadjacency in the acquisition of a second language. *Cognition*, 39, 215-258.
- Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive psychology*, 29, 1-23.
- Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress pattern of English words. *Child Development*, 64, 675-687.
- Jusczyk, P. W., Houston, D., & Newsome, M. (1999). The beginnings of word seg-

- mentation in English-learning infants. *Cognitive Psychology*, 39, 159-207.
- Kager, R. (1989). *A metrical theory of stress and destressing in English and Dutch*. Dordrecht: Foris.
- Kahn, D. (1980). *Syllable-based generalizations in English phonology*. New York: Garland.
- Kearns, R. (1994). *Prelexical speech processing by mono- and bilinguals*. Unpublished doctoral dissertation, University of Cambridge, Cambridge.
- Klatt, D.H. (1975). Linguistic uses of suprasegmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59, 1208-1221.
- Klatt, D. H. (1980). Speech perception: A model of acoustic-phonetic analysis and lexical access. In R. A. Cole (Ed.), *Perception and production of fluent speech*. Hillsdale, NJ: Erlbaum.
- Kolinsky, R. (1998). Spoken word recognition: A stage-processing approach to language differences. *European Journal of Cognitive Psychology*, 10, 1-40.
- Kolinsky, R., & Morais, J. (1996). Migrations in speech recognition. In F. Grosjean & U. H. Frauenfelder (Eds.), *Spoken word recognition paradigms. A special issue of Language and Cognitive Processes* (Vol. 11, pp. 611-619).
- Koster, M., & Cutler, A. (1997). Segmental and suprasegmental contributions to spoken-word recognition in Dutch. *Proceedings of the Fifth European Conference on Speech Communication and Technology* (pp. 2167-2170). Rhodes, Greece.
- Ladefoged, P. (1975). *A course in phonetics*. New York: Harcourt, Brace Jovanovich.
- Lehiste, I. (1962). The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America*, 51, 2018-2024.
- Lieberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431-461.
- Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America*, 32, 451-454.
- Long, M. (1990). maturational constraints on language development. *Studies in Second Language Acquisition*, 12, 251-285.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25, 71-102.
- Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive psychology*, 10, 29-63.
- Mattys, S.L. (1997). The use of time during lexical processing and segmentation: A review. *Psychonomic Bulletin & Review*, 4, 310-329.
- Mattys, S. L. (2000). The perception of primary and secondary stress in English. *Perception & Psychophysics*, 62, 253-265.
- Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, 38, 465-494.
- Mattys, S. L., & Samuel, A. G. (1997). How lexical stress affects speech segmentation and interactivity: Evidence from the migration paradigm. *Journal of Memory and Language*, 36, 87-116.
- Mattys, S. L., & Samuel, A. G. (2000). Implications of stress pattern differences in spoken word recognition. *Journal of Memory and Language*, 42, 571-596.

- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18, 1-86.
- McQueen, J. (1996). Word spotting. In F. Grosjean & U. H. Frauenfelder (Eds.), *Spoken word recognition paradigms. A special issue of Language and Cognitive Processes* (Vol. 11, pp. 695-699).
- McQueen, J. M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language*, 39, 21-46.
- McQueen, J., & Cox, E. (1995). The use of phonotactic constraints in the segmentation of Dutch. *Proceedings of Eurospeech '95* (Vol. 3, pp. 1707-1710). Madrid: Universidad Politécnica de Madrid.
- McQueen, J. M., Cutler, A., Briscoe, T., & Norris, D. (1995). Models of continuous speech recognition and the contents of the vocabulary. *Language & Cognitive Processes*, 10, 309-331.
- McQueen, J. M., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 20, 621-638.
- Mehler, J. (1981). The role of syllable in speech processing: Infant and adult data. *Philosophical Transaction of the Royal Society London*, B295, 333-352.
- Mehler, J., & Christophe, A. (2000). Acquisition of languages: Infant and adult data. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences*. Cambridge, MA: MIT Press.
- Mehler, J., Christophe, A., & Ramus, F. (in press). What we know about the initial state for language. In A. & Y. M. Marantz, & W. O'Neil (Eds.), *Image, language, brain: Papers from the first Mind-Articulation Project symposium*. Cambridge, MA: MIT Press.
- Mehler, J., Dommergues, J.-Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, 20, 298-305.
- Meunier, C., Content, A., Frauenfelder, U. H., & Kearns, R. (1997). The locus of the syllable effect: Prelexical or lexical? *Proceedings of the '97 Eurospeech conference* (Vol. 5, pp. 2851-2854). Patras, Greece.
- Morais, J., & Kolinsky, R. (1994). Perception and awareness in phonological processing: The case of the phoneme. *Cognition*, 50, 287-297.
- Morgan, J. L. (1996). A rhythmic bias in preverbal speech segmentation. *Journal of Memory and Language*, 35, 666-688.
- Morgan, J. L., & Saffran, J. R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development*, 66, 911-936.
- Nakamura, M., Kolinsky, R., Spagnoletti, C., & Morais, J. (1998). Phonemic awareness in alphabetically literate Japanese adults: The influence of the first acquired writing system. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 17, 417-450.
- Nakatani, L. H., & Dukes, K. D. (1977). Locus of segmental cues for word juncture. *Journal of the Acoustical Society of America*, 62, 714-719.
- Nakatani, L. H., O'Connor, K. D., & Aston, C. H. (1981). Prosodic aspects of American English speech rhythm. *Phonetica*, 38, 84-106.
- Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns:

- Towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 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.
- Nespor, M. (1990). On the rhythm parameter in phonology. In I. M. Roca (Ed.), *Logical issues in language acquisition* (pp. 157-175). Dordrecht: Foris.
- Newport, E. L. (1990). maturational constraints on language learning. *Cognitive Science*, 14, 11-28.
- Norris, D., & Cutler, A. (1985). Juncture detection. *Linguistics*, 23, 689-705.
- Norris, D., Mc Queen, J., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology*, 21, 1209-1228.
- Norris, D. G. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52, 189-234.
- Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, 32, 258-278.
- Oyama, S. (1976). A sensitive period for the acquisition of a non-native phonological system. *Journal of Psycholinguistic Research*, 5, 261-283.
- Pallier, C., Bosch, L., & Sebastián, N. (1997). A limit on behavioral plasticity in vowel acquisition. *Cognition*, 64, B9-B17.
- Pallier, C., Sebastián-Gallés, N., & Colomé, A. (1999, September). *Phonological representations and repetition priming*. Proceedings of the Eurospeech '99 Conference (Vol. 4, pp. 1907-1910). Budapest, Hungary.
- Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, E., Fazio, F., & Mehler, J. (1996). Brain processing of native and foreign languages. *Neuroreport*, 7, 2439-2444.
- Perani, D., Paulesu, E., Sebastián-Gallés, N., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S. F., Fazio, F., & Mehler, J. (1998). The bilingual brain: Proficiency and age of acquisition of the second language. *Brain*, 121, 1841-1852.
- Pike, K. L. (1945). *The intonation of American English*. Ann Arbor, MI: University of Michigan Press.
- Port, R. F., Dalby, J., & O'Dell, M. (1987). Evidence for mora timing in Japanese. *Journal of the Acoustical Society of America*, 81, 1574-1585.
- Quené, H. (1993). Segment durations and accent as cues to word segmentation in Dutch. *Journal of the Acoustical Society of America*, 94, 2027-2035.
- Quené, H., & Koster, M. L. (1998). Metrical segmentation in Dutch: Vowel quality or stress? *Language and Speech*, 41, 185-202.
- Quené, H., & Smith, Y. (1992). On the absence of word segmentation at "weak" syllables. *Proceedings of the International Conference of Spoken Language Processing* (Vol. 1, pp. 213-215). University of Alberta, Alberta.
- Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, 288, 349-351.
- Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *Journal of the Acoustical Society of America*, 105, 512-521.
- Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the

- speech signal. *Cognition*, 73, 265-292.
- Rietveld, A. C. M. (1980). Word boundaries in the French language. *Language and Speech*, 23, 289-296.
- Roach, P. (1982). On the distinction between stress-timed and syllable-timed languages. In D. Crystal (Ed.), *Linguistic controversies* (pp. 73-79). London: Edward Arnold.
- Roeper, T., & Williams, E. (1988). *Parameter setting*. Dordrecht: Reidel.
- Romaine, S. (1995). *Bilingualism*. Oxford: Blackwell.
- Rossi, M. (1972). Le seuil différentiel de durée. *Papers in memory of Pierre Delattre* (pp. 435-450). The Hague: Mouton.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926-1928.
- Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35, 606-621.
- Sanders, L., Yamada, Y., & Neville, H. J. (1999). Speech segmentation by native and non-native speakers: An ERP study. *Society for Neuroscience*, 25, 358.
- Schaerlaekens, A. (1998). Bilingualism in children: An introduction. In M. Brysbaert (Ed.), *Bilingualism. Special issue of Psychologica Belgica* (Vol. 38, pp. 129-148).
- Sebastián-Gallés, N., Dupoux, E., Segui, J., & Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language*, 31, 18-32.
- Sebastián-Gallés, N., & Soto-Faraco, S. (1999). On-line processing of native and non-native phonemic contrasts in early bilinguals. *Cognition*, 72, 112-123.
- Sluijter, A. M. C. (1995). *Phonetic correlates of stress and accent*. Unpublished doctoral dissertation. Rijksuniversiteit Leiden, The Netherlands.
- Small, L. H., Simon, S. D., & Coldberg, J. S. (1988). Lexical stress and lexical access: Homographs versus nonhomographs. *Perception & Psychophysics*, 44, 272-280.
- Smith, K. L., & Pitt, M. A. (1995). Is syllabic information used in auditory word recognition?, Abstracts of the 36th Annual Meeting of the Psychonomic Society (p. 22). Los Angeles, CA.
- Suomi, K., McQueen, J. M., & Cutler, A. (1997). Vowel harmony and speech segmentation in Finnish. *Journal of Memory and Language*, 36, 422-444.
- Tabossi, P., Burani, C., & Scott, P. (1995). Word identification in fluent speech. *Journal of Memory & Language*, 34, 440-467.
- Thiery, C. (1976). Le bilinguisme vrai. *Etudes de Linguistique Appliquée*, 24, 52-63.
- Thiery, C. (1978). True bilingualism and second-language learning. In D. Gerver & H. Sinaiko (Eds.), *Language interpretation and communication*. New York: Plenum Press.
- Tranel, B. (1987). *The sounds of French: An introduction*. Cambridge: Cambridge University Press.
- Turk, A. E., Jusczyk, A. M., & Gerken, L. A. (1995). Do English-learning infants use syllable weight to determine stress? *Language and Speech*, 38, 143-158.
- Umeda, N., & Cocker, C. H. (1974). Allophonic variation in American English. *Journal of Phonetics*, 2, 1-5.
- Vaissière, J. (1983). Language-independent prosodic features. In A. Cutler & D. R. Ladd (Eds.), *Prosody: Models and measurements* (pp. 53-66). Berlin: Springer-

Verlag.

- Vaissière, J. (1991). Rhythm, accentuation and final lengthening in French. In L. Nord & R. Carlson (Eds.), *International Symposium Series* (Vol. 2, pp. 108-120). New York: MacMillan Press.
- Van der Hulst, H. (1984). *Syllable structure and stress in Dutch*. Dordrecht: Foris.
- van Donselaar, W., & Stoutjesdijk, O. (1994). The acoustic/phonetic factors behind the "syllable effect". *OTS Yearbook 1993*, 25-43.
- Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighbourhood activation in spoken word recognition. *Journal of Memory and Language*, 40, 374-408.
- 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., & de Gelder, B. (1997a). Activation of embedded words in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 710-720.
- Vroomen, J., & de Gelder, B. (1997b). Trochaic rhythm in speech segmentation. *Abstracts of the 38th Annual Meeting of the Psychonomic Society* (Vol. 2, pp. 73), Philadelphia.
- Vroomen, J., & de Gelder, B. (submitted). Stress is a cue to word onset: Stress based speech segmentation.
- Vroomen, J., Tuomainen, J., & de Gelder, B. (1998). The role of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language*, 38, 133-149.
- Vroomen, J., van Zon, M., & de Gelder, B. (1996). Cues to speech segmentation: Evidence from juncture misperception and word spotting. *Memory and Cognition*, 24, 744-755.
- Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neurosciences*, 8, 231-256.
- Wenk, B. J., & Wioland, F. (1982). Is French really syllable-timed? *Journal of Phonetics*, 10, 193-216.
- Williams, B., & Hiller, S. (1989). Investigating randomness in foot timing patterns in English. *Journal of the Acoustical Society of America*, 85, S59.
- Yerkey, P. N., & Sawush, J. R. (1993). The influence of stress, vowel and juncture cues on segmentation (abstract). *Journal of the Acoustical Society of America*, 94, 1880.
- Zwitsers, P., Schriefers, H., Lahiri, A., & van Donselaar, W. (1993). The role of syllables in the perception of spoken Dutch. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 260-271.

Received September, 2000

Accepted October, 2000