Cross-linguistic differences in early word form segmentation: a rhythmic-based account

KARIMA MERSAD
LOUISE GOYET
THIERRY NAZZI

Abstract

The present paper reviews recent studies on the early segmentation of word forms from fluent speech. After having exposed the importance of this issue from a developmental point of view, we summarize studies conducted on this issue with American English-learning infants. These studies show that segmentation abilities emerge around 8 months, develop during the following months, and rely on infants' processing of various word boundary cues the relative weight of which changes across development. Given that these studies show that infants mostly use cues that are specific to the language they are acquiring, we underline that the development of these abilities should vary cross-linguistically, and raise the issue of the developmental origin of segmentation abilities. We then offer one solution to both the crosslinguistic differences (also observed in adulthood) and bootstrapping issues in the form of the early rhythmic segmentation hypothesis. This hypothesis states that infants rely on the underlying rhythmic unit of their native language at the onset of segmentation abilities: the trochaic unit for stress-based languages, the syllable for syllable-based languages. After the presentation of various elements on which this hypothesis relies, we present recent data on French infants offering a first validation of this proposal.

What does it mean, from a developmental point of view, to learn a lexicon? A word corresponds to the specific pairing between the mental representation of a sound pattern (word form) and an abstract representation (concept) of an object or event in the world that constitutes the meaning associated to that word form. The building of a lexicon will then rely on the development of three sets of abilities: the ability to elaborate and store word

forms, the ability to build concepts for the objects and events in the world, finally the ability to appropriately link word forms and concepts. Note that the elaboration of word forms and of concepts starts before the onset of lexical acquisition per se around the ages of 10 to 12 months. It is thus likely that infants in their first year of life constitute a store of word forms and concepts that are later paired to make words. In the second year of life, all these acquisitions could happen simultaneously (see Nazzi & Bertoncini, 2003, for a discussion of possible changes in word acquisition around 18 months of age).

The acquisition of the lexicon by infants and young children has been a research topic for decades. In what follows, we review studies focusing on a specific ability involved in word learning, namely the ability to extract the sound pattern of words from fluent speech (henceforward, word form segmentation). Word form segmentation constitutes a crucial step in speech processing, which allows infants as well as adults to determine the sequence of lexical units that constitute the utterances they hear. The ability to extract word forms from spoken speech might play a critical role for the acquisition of the lexicon. This hypothesis is supported by the finding of positive correlations between word segmentation performance and later vocabulary levels (Newman, Bernstein Ratner, Jusczyk, Jusczyk, & Dow, 2006), and by the demonstration that newly segmented words are easier to link to new objects at 17 months of age (Graf Estes, Evans, Alibali, & Saffran, 2007). It also appears that word form segmentation is a prerequisite for the acquisition of syntax, given that all theories of syntax acquisition presuppose that infants have access to the segmented sequence of words constituting the utterances they hear (Newman et al., 2006).

Accessing word forms would not be an issue if word boundaries were clearly marked at the acoustic level, or if words were (often) presented in isolation. First, numerous studies show that word boundaries are not clearly marked in adult-directed speech (Cole & Jakimik, 1978, 1980; Klatt, 1979, 1989). Second, two studies evaluated the presence of isolated words in the input to English-learning infants (Aslin, 1993; Brent & Siskind, 2001) and a Dutch/German bilingual infant (van de Weijer, 1998). Their results showed that infant-directed speech consists mostly of multi-word utterances, words pronounced in isolation making up less than 10% of all words present in the analyzed corpuses. It also appeared that words uttered in isolation might be easier to acquire. Indeed, infants' production of a given word was better predicted by the frequency of isolated tokens of that word heard a few months earlier, than by the total frequency of that word at that same age (Brent & Siskind, 2001). However, this link could only be tested on a very small subset of infants' early words (those that appeared *both* in early isolated and sentential parental input and in subsequent infant productions). Moreover, given that many words appearing in isolation correspond to fillers (*yes*, *hmm*, …), vocatives ("infant's first name," …) and social expressions (*hi*!, …), as

shown by van de Weijer (1998), segmentation procedures remain necessary for types of words that do not appear in isolation (especially grammatical words). Although there are few isolated words, and few pauses between consecutive words in the signal, there are, as discussed below, many more subtle linguistic cues that signal word boundaries or indicate that two sounds belong to the same lexical unit.

Infants' sensitivity to word boundary cues

The first kind of word boundary information consists of prosodic cues that exist at different levels of the language structure. At the sentence level, intonational phrase boundaries are perceived by very young infants (Christophe, Dupoux, Bertoncini, & Mehler, 1994) and might help find word boundaries as they align with some of them (Nespor & Vogel, 1986). At the word level, prosody is related to the way stress and intonation are affected by position within the word. For example, a majority of words are stressed in initial position in English (Cassidy & Kelly, 1991; Cutler & Carter, 1987; Kelly & Bock, 1988), while in French, syllable lengthening is observed at the end of words in phrase final positions (Delattre, 1966; Fletcher, 1991). Existing studies show that infants are sensitive to the prosodic word boundary markers specific to their native language. For example, it has been found that French newborns distinguish between two versions of the same bisyllabic sequence according to whether the prosodic boundary of a phrase is present or not between the two syllables (Christophe, Dupoux, Bertoncini, & Mehler, 1994). Moreover, a preference for English words with the predominant trochaic stress pattern (word-initial stress or strong-weak, such as *PORter*) over English words with the less frequent iambic stress pattern (word-final stress or weak-strong, such as *rePORT*) emerges between the ages of 6 and 9 months in English-learning infants (Jusczyk, Cutler, & Redanz, 1993; Turk, Jusczyk, & Cutler, 1995). New evidence from German- and French-learning infants suggest that this preference might actually emerge earlier (between 4 and 6 months) and results from exposure to the native language rather than from a general trochaic bias (Friederici, Friederich, & Christophe, 2007; Hohle, Bijeljac-Babic, Nazzi, Herold, & Weissenborn, 2009). Further support for language specific differences in stress perception comes from data showing that Spanish-learning 9-month-olds distinguish between trochaic and iambic pseudo-words, while French-learning infants of the same age show discrimination difficulties (Skoruppa, Pons, Christophe, Bosch, Dupoux, Sebastián-Gallés, Galluppa, & Pon, 2009).

Allophonic cues constitute a second kind of cues to word boundaries as the acoustic realization of some phonemes depends on whether they are at the border or inside a word. For example, in English, the realization of the phonemes /t/ and /r/ differ in the word *nitrate* and in the sequence *night rate* (Hohne & Jusczyk, 1994). Sensitivity to allophonic differences has been

found in infants as young as 2 months of age, as attested by their ability to discriminate between pairs such as *nitrate* and *night rate* (Hohne & Jusczyk, 1994).

Phonotactic constraints, that is, constraints regarding the phonetic sequences allowed at the lexical level, provide a third kind of cues to word boundaries. For example, the sequence /zt/ for English, or the sequences /kf/ or /vg/ for French, cannot be found within words. Their presence in the speech stream would thus signal the presence of a word boundary. Infants become sensitive to the phonotactic properties of their native language between the ages of 6 and 9 months (Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993; Jusczyk, Luce, & Charles-Luce, 1994; Mattys, Jusczyk, Luce, & Morgan, 1999; see also Friederici & Wessels, 1993, for Dutch; Sebastián—Gallés, & Bosch, 2002, for Catalan; Nazzi, Bertoncini, & Bijeljac-Babic, 2009, for French) showing a preference for legal (or frequent) sequences of phonemes over illegal ones (for example, in English, *chun* is a pseudoword with frequent phonotactics while *yush* is a pseudoword with rare phonotactics)

Fourth, statistical/distributional information regarding the ordering of consecutive phonemes (or consecutive syllables) within words provide information on the likelihood that two sounds belong to the same word. Indeed, some phonetic sequences are more frequent within words than others (Hockema, 2006; Jusczyk, Luce, & Charles-Luce, 1994), and transitional probabilities between two syllables are higher within words than across lexical boundaries (Curtin, Mintz, & Christiansen, 2005).

The existence of the above linguistic cues calls the following remarks. First, although none of these cues provide a systematic marking of word boundaries, their combination would provide sufficient information to allow the correct segmentation of the speech stream (Christiansen, Allen, & Seidenberg, 1998). Second, and more importantly*,* infants will have to learn how the above-mentioned cues mark word boundaries in the language spoken in their environment. This would be true for all cues except distributional information which might be domain- and species-general, as suggested for transitional probabilities by the fact that they were observed with non--linguistic auditory and visual stimuli (Saffran et al., 1999; Fiser et al., 2001), and found to be used by non-human mammals (Hauser et al., 2001; Toro et al., 2005).

In the next sections, we review pioneering studies on early word form segmentation by English-learning infants, the language that provides the largest body of research on this issue up to now.

Emergence of word form segmentation abilities in English-learning infants

Most of the studies on infants' use of word boundary cues to segment fluent speech rely on an adaptation of the headturn preference paradigm (HPP, see Figure 1) by Jusczyk and Aslin (1995).
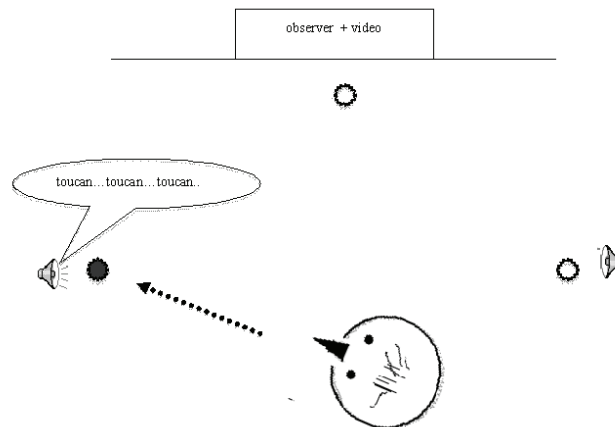
(1)
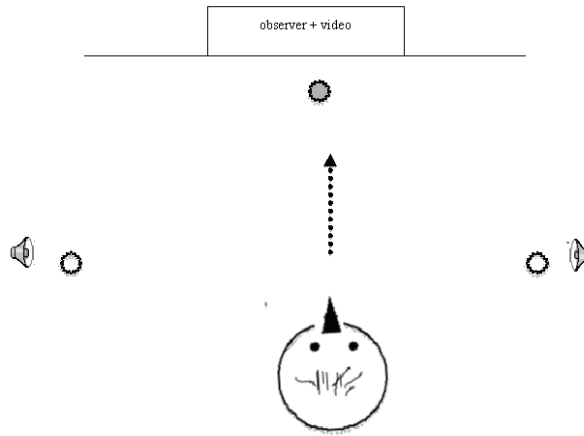
(2)

Figure 1. The headturn preference procedure (HPP) is based on the comparison of orientation/listening times to different kinds of stimuli. For each trial, the center light first blinks to bring infant's attention to the center of the display; when the infant looks at the center, (1) one of the side lights starts blinking, when the infant turns to that light, (2) the stimuli are presented from the loudspeaker on the same side and orientation/listening times are recorded. In segmentation studies, infants are first familiarized with two words, and then tested with passages containing or not those words (the reversed passage/word order is also used).

In the first experiment of that study, 7.5-month-old infants were familiarized with two monosyllabic words (*cup* and *dog*, or *bike* and *feet*) and then heard four passages, each passage being build around one of the four target words (each of these words was repeated 6 times in its corresponding passage). The results revealed a preference for the passages corresponding to the familiarized words, indicating that infants had recognized the target words, which in turn implied that they had segmented the words from the passages.[1] Failing to extend this result to 6-month-old infants, Jusczyk and Aslin (1995) concluded that word form segmentation abilities emerge between the ages of 6 and 7.5 months.

In further experiments, Jusczyk and Aslin (1995) established that the segmented word forms are phonetically specified: 7.5-month-olds do not show a segmentation effect when familiarized with the pseudoword *zeet* while the target word contained in the passages is the word *feet*. This result was later extended to the final consonant (*feek*, Tincoff & Jusczyk, 1996). The specificity of the vowel was not tested, and will have to be explored given recent results showing more specific use of consonantal over vocalic information by young word-learners (Nazzi, 2005; Nazzi & New, 2007; Havy & Nazzi, 2009). On a related issue, recognition of target words was also found to be affected, early in development, by acoustic distance such as that related to gender differences (Houston & Jusczyk, 2000, 2003) or speech affect (Singh, Morgan, & White, 2004; Thiessen, Hill, & Saffran, 2005).

While the first experiments familiarized infants with words and then tested them with passages, Jusczyk and Aslin (1995) showed that similar segmentation effects can be obtained when the order of presentation of the stimuli is reversed, that is, when infants are first familiarized with two passages (each containing a target word). In this situation, 7.5-month-olds listen longer, at test, to the two words that were contained in the familiarization passages, compared to two control words. This last design, which is more akin to the situation that infants have to face outside the laboratory, reinforces the claim that young infants are already proficient at segmenting word forms from native fluent speech (see also Jusczyk and Hohne, 1997).

To sum up, the importance of the study by Jusczyk and Aslin (1995) is two-fold. First, it established an experimental procedure adapted to the study of early speech segmentation. Second, it was the first study to reveal segmentation abilities at such an early age. Subsequent experiments explored the development of the use of the different segmentation cues.

---

[1] As will be discussed later, this interpretation of infants' preference/recognition in terms of segmentation is confirmed by later studies, starting with Jusczyk et al. (1999) finding that infants of 7.5 months do not recognize the sole stressed syllable of a trochaic bisyllabic word (such as *DOCtor*) while they recognize that whole word.

Cues used for word form segmentation by English-learning infants during development

From the studies on English-learning infants emerges the following developmental trajectory. Around 7 to 8 months of age, infants appear to use rhythmic information in order to segment speech into sequences of one or more syllables starting with a stressed syllable, that is, into trochaic units (Jusczyk, Houston & Newsome, 1999b; for further convergent data, see also Curtin, Mintz & Christiansen, 2005; Echols, Crowhurst, & Childers, 1997; Houston, Santelmann & Jusczyk, 2004; Johnson & Jusczyk, 2001; Morgan & Saffran, 1995; Nazzi, Dilley, Jusczyk, Shattuck-Hufnagel, & Jusczyk, 2005). Given that most English bisyllabic words have a trochaic stress pattern (Cassidy & Kelly, 1991; Cutler & Carter, 1987; Kelly & Bock, 1988), this rhythmic segmentation procedure (similar to the metrical segmentation procedure used by adults, see Cutler, Mehler, Norris, & Segui, 1986; Cutler & Norris, 1988; McQueen, Norris, & Cutler, 1994) would allow English-learning infants to appropriately segment, from a young age, most bisyllabic words. The research showing English-learning infants' early use of rhythmic information for segmentation can be illustrated by Jusczyk et al. (1999b) study. Their results (see Figure 2) show that 7.5-month-olds segment trochaic (strong-weak) words such as *DOCtor*, whether words or passages are presented first.
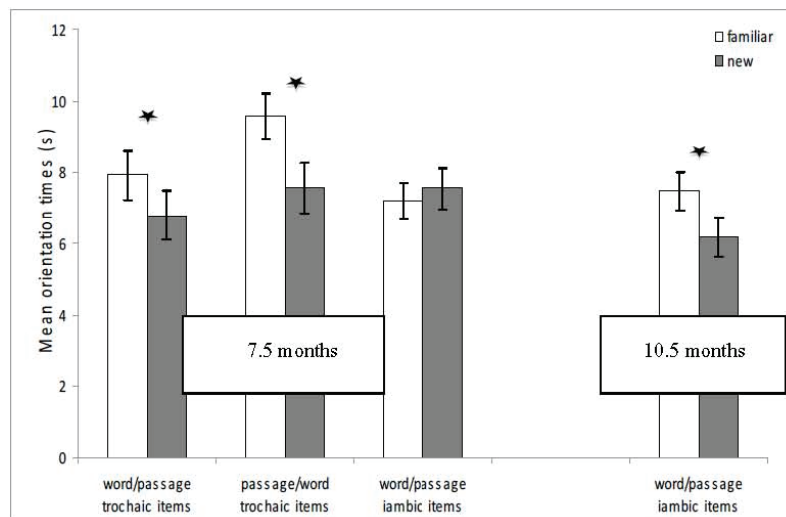

Figure 2: Mean orientation times (and standard error) to the test passages containing the familiar(ized) versus new items (for the word/passage order), or to the word-lists corresponding (familiar) or not (new) to the familiarized passages (for the passage/word order). Results are broken down by age (7- versus 10.5-month-olds) and item type (trochaic versus iambic word).

However, 7.5-month-olds, though not 10.5-month-olds, missegment iambic (weak-strong) words such as *guiTAR*, placing a word boundary between the initial/weak and final/strong syllables (e.g., *gui / TAR*). An advantage of trochaic words was also found for English verbs, even though most English bisyllabic verbs have an iambic stress pattern (Nazzi et al., 2005). This last finding suggests that although some acoustic and phonological properties distinguish nouns and verbs (Kelly, 1992), the trochaic bias is applied to all lexical categories in this language.

At 7.5 months, infants also use distributional regularities on the order of syllables in the speech signal (from now on: syllabic distributional information). For example, 7.5-month-old infants tested with passages containing trochaic words such as *DOCtor* showed a segmentation effect if they had been familiarized with the whole words, but not if they had been familiarized solely with its initial syllable *DOC* (Jusczyk et al., 1999b). Moreover, with an artificial language paradigm in which infants are presented with a continuous sequence made-up of randomly ordered repetitions of 4 trisyllabic pseudo-words, 8-month-olds were found to group syllables into cohesive word-like units on the basis of syllabic distributional information (Saffran, Aslin, & Newport, 1996; though see Perruchet & Vinter, 1998, for an alternative interpretation of these results, and Brent & Cartwright, 1996, Dahan & Brent, 1999, for an alternative model). Subsequent work has demonstrated infants' use of distributional information in other domains such as music (Saffran, Johnson, Aslin & Newport, 1999), and by non-human primates (Hauser, Newport & Aslin, 2001).

By 10.5 months of age, infants start to use other word boundary cues, such as allophonic (Jusczyk, Hohne, & Bauman, 1999a) and phonotactic (Mattys & Jusczyk, 2001a) information. English-learning infants have also been found to use other prosodic cues than the rhythmic cue discussed so far: sentence edges by 8 months (Seidl & Johnson, 2006), phrase boundaries by 10.5 months (Gout, Christophe & Morgan, 2004; Nazzi et al., 2005),[2] pitch accent information between 10 and 13 months (Nazzi et al., 2005). Moreover, coarticulation information starts to play a role in segmentation by the age of 8 months (Johnson & Jusczyk, 2001). Lastly, words starting with a consonant have been found to be easier to segment than words starting with a vowel for English-learning infants between the ages of 8 and 13 months (Mattys & Jusczyk, 2001b; Nazzi et al., 2005). Note that the delay in segmenting vowel--initial words is also reported for French-learning infants, usually in relation to the acquisition of "liaison," which is a phenomenon that refers to the appearance of a consonant at the juncture of two words, when the second

---

[2] The use of these prosodic unit boundaries might allow infants to cut full utterances into smaller-size sequences, clauses or phrases, within which other segmentation procedures would apply (c.f. Hirsh-Pasek, Kemler Nelson, Jusczyk, Wright Cassidy, Druss, & Kennedy, 1987; Jusczyk, Hirsh-Pasek, Kemler Nelson, Kennedy, Woodward, & Piwoz, 1992; Nazzi, Kemler Nelson, Jusczyk, & Jusczyk, 2000; Soderstrom, Seidl, Kemler Nelson, & Jusczyk, 2003).

word begins with a vowel (Chevrot, Dugua, & Fayol, 2005; Wauquier-
-Gravelines, 2002). Difficulties for vowel-initial words probably stem from
coarticulation and resyllabification processes that blur word onset boundaries.

So far, the discussion has focused on infants' use of bottom-up
segmentation procedures relying on the presence of acoustic/phonetic cues.
However, we have remarked earlier that infants do hear a few (though not
many) words in isolation. It is conceivable that these isolated words are
stored, and later used to perform another kind of segmentation: top-down
segmentation. The Incdrop model (Brent & Cartwright, 1996) was proposed
in that perspective. The model states that infants will memorize an incoming
utterance as a whole unit (e.g., *dopuneribo*) unless it contains a sequence that
has been previously memorized (e.g., *ne*). In this case, the memorized item is
used to segment the new utterance, resulting in the memorization of the
complementing units (e.g., *dopu* and *ribo*). Evidence for this model was
obtained through computer simulation (Brent & Cartwright, 1996) and
through studies looking at adults' acquisition of artificial languages (Dahan &
Brent, 1999). More recently, one study provided the first piece of evidence
that known words can facilitate word form segmentation by infants. Indeed,
English-learning infants were found to segment unfamiliar words by 6 months
of age (as opposed to 7.5 months in Jusczyk and Aslin, 1995) if these words
were preceded by very familiar words, such as the infants' name, or the word
mommy (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005). If this type of
segmentation necessarily plays a limited role at the onset of lexical acquisition
when infants have only memorized a few words, its role is likely to grow in
parallel with the increase in the size of infants' vocabulary (see also data
suggesting the onset of function words use for segmentation by 11-to-13
months of age, Shi & Lepage, 2008; Shi, Werker, & Cutler, 2006).

Relative weight of segmentation cues in English

Recent research brings some information with respect to the relative weight of
various segmentation cues at different points in development, in particular for
the prosodic/rhythmic and syllabic distributional information cues, the two
cues that appear to play crucial roles in early word form segmentation.

Data obtained using the Jusczyk and Aslin (1995) paradigm suggest that
English-learning infants first segment speech according to prosodic information,
syllabic order information being then used within the rhythmically-defined units
(Jusczyk et al., 1999b). Indeed, the prosodic boundary placed between the two
syllables of an iambic word (e.g., *guiTAR*) appears to block 7.5-month-olds'
use of distributional information (i.e., the fact that *gui* and *tar* always
appeared consecutively) resulting in the segmentation of the sole strong
syllable. Similarly, if a weak-strong word is always followed by the same
weak syllable (e.g., *guiTAR_is*), 7.5-month-olds place a word boundary
between the first two syllables, and group together the last two syllables,

resulting in an incorrect segmentation (e.g., *gui / TARis*). These findings suggest that English-learning 7.5-month-olds use rhythmic information to perform a first-pass parsing of continuous speech into smaller units that constitute the basis of further analyses of the signal. They also show that 10.5--month-olds, contrary to 7.5-month-olds, segment iambic words correctly and no longer missegment trochaic units that span word boundaries. An interpretation is that by 10.5 months, infants use distributional information to detect the cohesiveness of two consecutive syllables even when they cross a rhythmically-placed boundary. This suggests that by this age, infants weight syllabic distributional information more heavily than rhythmic cues, and/or take benefit of the other segmentation cues mentioned earlier (e.g., allophonic variations, phonotactic constraints, …).

The above conclusions of a precedence of prosodic information are further supported by data from Johnson & Jusczyk (2001) showing in an artificial language paradigm that when prosody and syllabic distributional information are pitted against one another, 8-month-olds give more weight to rhythmic information. However, this conclusion was later challenged by Thiessen & Saffran (2003). Indeed, while these authors did replicate Johnson and Jusczyk's (2001) finding with 9-month-olds, they found the opposite pattern at 7 months, which led them to the conclusion that distributional information is used earlier than rhythm. However, even though Thiessen and Saffran's (2003) results show that infants can track syllabic distributional information and use it after a few minutes of exposure to a very simplified language, it is unclear whether infants would benefit from syllabic order information at such a young age in the context of a natural language made up of thousands of words of varied syllabic length. But again, recent studies on this issue bring contradictory evidence, with some data suggesting that this may actually not be the case (Johnson & Tyler, 2010), and some others suggesting it may be possible (Pelucchi et al., 2009).

In summary, the studies above reveal developmental changes in the use of syllabic distributional information and prosodic cues across development, while leaving still unclear the exact pattern of changes in their relative weight. Future studies will have to continue investigating this issue, in both infants and adults, using as a theoretical framework the hierarchical model of word segmentation proposed by Mattys, White and Melhorn (2005) to account for adults' segmentation abilities (although this model does not yet discuss the use of distributional information). This hierarchical model postulates three tiers of segmentation cues. Tier 1 refers to sublexical suprasegmental cues such as prosody. Tier 2 refers to sublexical segmental cues such as phonotactics, allophony and coarticulation. Tier 3 refers to lexical cues (which broadly include semantic, syntactic and pragmatic information). In optimal listening conditions, adults were found to rely on the lexical level. However, they were found to rely on segmental information when the speech signal was degraded, and to rely on prosodic information when the speech signal was

severely degraded. The data on early word form segmentation presented earlier support the notion that word segmentation information from Tier 1 (rhythmic information) becomes available before information from Tier 2 (allophonic and phonotactic information). Information from Tier 3 (lexical level), which appears to be possible to use as early as 6 months of age for a very limited set of highly familiar words, would become more crucial for segmentation as the size of infants' vocabulary increases. Recently, Mersad and Nazzi (submitted) have proposed that transitional probabilities are situated at the bottom of the hierarchical model of Mattys et al. (2005). This implies that transitional probabilities should have a crucial role in early infancy when no other cues are available, but also that they should be outranked in adulthood when pitted against other cues, a prediction supported by recent data (Shukla et al., 2007; Finn & Kam 2008).

Explaining crosslinguistic differences: the *early rhythmic segmentation hypothesis*

The developmental pattern described in the previous sections was established for English-learning infants. However, linguistic cues (prosody, phonotactic, allophonic cues, distributional information..) are instantiated differently across languages. In order to better understand the way segmentation procedures are put into place, it is crucial to evaluate how their emergence varies with different linguistic inputs.

A crosslinguistic approach should allow us, in particular, to investigate how infants start using the different segmentation cues. Indeed, even though the studies on English have led to the identification of some of the cues used for early segmentation, they left open the issue of how infants can rely, to start segmenting word forms, on cues that are mostly (except distributional information) defined at the lexical level (stress pattern/trochaic unit; phonotactic constraints; allophonic variations)? In other terms, what are the roots of the word form segmentation procedures used by infants for the cues that are language-specific?

One potential solution to this issue is the *early rhythmic segmentation hypothesis* (Nazzi et al., 1998a, 2000, 2006). This hypothesis is based on the findings that there are different rhythmic classes of languages, and that adults segment speech according to their native language rhythmic class. In addition, infants have been found to be sensitive from birth to the rhythmic properties of languages (see details in the following sections). The early rhythmic segmentation hypothesis states that infants could learn the rhythmic segmentation procedure appropriate to the rhythmic unit of their native language independently of the lexical level. This procedure will develop on the basis of newborns' and young infants' sensitivity to rhythm as attested by language discrimination abilities (Mehler, Dupoux, Nazzi & Dehaene--Lambertz, 1996; Nazzi, Bertoncini & Mehler, 1998a; Nazzi, Jusczyk &

Johnson, 2000).[3] Importantly, these analyses at the prosodic level do not require prior phonological acquisitions, as sensitivity to prosodic information has been found to extend to stimuli in foreign languages, and to stimuli with reduced phonetic information such as lowpass filtered stimuli (Jusczyk et al., 1993a; Nazzi, Bertoncini & Mehler, 1998a). Once acquired, the appropriate rhythmic segmentation procedure (maybe in conjunction with distributional information) would allow infants to start segmenting their first sound patterns (some being erroneous, as shown by Jusczyk et al., 1999b) and then progressively start specifying other language-specific word boundary cues (allophonic, phonotactic, …). This *early rhythmic segmentation hypothesis* thus offers a solution to the issue of the emergence of different segmentation abilities for different linguistic backgrounds. Crucially, it predicts different developmental trajectories of segmentation abilities for languages of different rhythmic types. Before presenting the studies on early word form segmentation in French that have started directly evaluating this hypothesis, we review in more details the data on linguistic rhythm and its influence on speech processing by adults and young infants, on which the *early rhythmic segmentation hypothesis* was initially based.

Linguistic rhythm and its impact on speech processing

Specifically, the *early rhythmic segmentation hypothesis* relies on a series of findings suggesting (a) the existence of rhythmic cues in the speech signal, (b) that appear to influence speech processing by adults and (c) that are perceived by infants from birth onwards. In the following, we review the evidence regarding these three points, and discuss how they converge towards the idea that language specific segmentation procedures might be a consequence of the processing of linguistic rhythm. As will become clear, the early rhythmic segmentation hypothesis proposes an integrated explanation for a range of linguistic and psychological phenomena that has the advantage of offering a parsimonious developmental account for the differences in segmentation abilities found in adulthood.

Rhythmic classes

The idea that there are different rhythmic classes of languages goes back many decades (Abercrombie, 1967; Pike, 1945). Three rhythmic classes have usually been considered in the literature: the stress-timed class (including

---

[3] Note that this proposal is in line with numerous other proposals about infants' early sensitivity to prosodic information, and the importance of prosody in early lexical acquisition (Fernald & Kuhl, 1987; Jusczyk et al., 1993a, 1993b; Karzon & Nicholas, 1989; Kuhl & Miller, 1982; Nazzi et al., 1998a, 1998b; Spring & Dale, 1977).

languages such as English, Dutch, German, …), the syllable-timed class (including French, Spanish, Korean …) and the mora-timed class (including Japanese, Telugu, …). An underlying rhythmic unit is associated to each of these language classes: the stress unit, the syllable and the mora[4] respectively. Importantly, there is a hierarchical relationship between these three units: stress units are made up of syllables that are themselves made up of morae.

   The initial definition of these rhythmic classes stated that all tokens of the rhythmic unit underlying a given language were of (roughly) similar duration. For example, it was predicted that for stress-timed languages, all stress units should be of relatively identical duration (or at least, should show a tendency towards similar durations in milder version of this isochrony definition) independently of their number of syllables or their position within the sentence. If that definition proved to be incorrect, more recent studies have started specifying a more subtle acoustic signature of the different rhythmic classes that brings new support to this notion (Arvaniti, 1994; den Os, 1988; Fant, Kruckenberg & Nord, 1991; Ramus, Nespor, & Mehler, 1999; Shafer, Shucard, & Jaeger, 1999). For example, the analyses conducted by Ramus et al. (1999) on utterances produced in eight different languages have identified two acoustic measures (proportion of vocalic intervals duration and standard deviation of consonantal intervals duration) that define a bidimensional space in which these languages fall according to the three rhythmic classes: stress-based English, Dutch, and Polish; syllable-based French, Italian, Spanish and Catalan; mora-based Japanese.

Rhythmic classes and crosslinguistic differences in adult speech processing

The notion of rhythmic classes has proved useful to explain crosslinguistic differences in the way that adults segment speech and access the lexicon, by suggesting a link between the adults' behavior and the global rhythmic properties of their native language.

   First, the syllable appeared as the segmentation unit for adults speaking several syllable-based languages: French (Mehler, Dommergues, Frauenfelder, & Segui, 1981, for Parisian French; Peretz, Lussier & Béland, 1998, for Canadian French; but see Content, Meunier, Kearns, & Frauenfelder, 2001[5]), Spanish and Catalan (Sebastián-Gallés, Dupoux, Segui,

---

[4] The mora is a rhythmic unit that can either be syllabic or subsyllabic. In English, a mora roughly corresponds to a CV syllable with a short vowel (e.g. "the" as opposed to "thee", which has a long vowel). In Japanese, CV syllables with long vowels and syllables with final nasals (like the first syllable in "Honda") or final geminate consonants (like the first syllable in "Nissan") have two morae.

[5] The study by Content et al. (2001) revealed that syllabic effects are not always obtained in French, such effects being particularly robust only when a liquid consonant is present in intervocalic position. However, in spite of these relative

& Mehler, 1992), and Korean (Kim, Davies, & Cutler, in press). Second, the segmentation procedure used in English (Cutler et al., 1986; Cutler & Norris, 1988; McQueen et al., 1994) and Dutch (Vroomen, van Zon, & de Gelder, 1996) appeared to rely on the trochaic stress unit (with most words in these languages following this stress pattern). A third pattern was found for Japanese and Telugu speakers, these adults relying on the mora (Otake et al., 1993; Murty, Otake & Cutler, 2007). Given the existence of a hierarchical relation between these three rhythmic units, the acquisition of the native language procedure might proceed from the specification of the level of rhythmic unit most appropriate for the language in acquisition.

There is also evidence that adult segmentation procedures are deeply embedded in their native language-specific abilities, and thus are likely to have been acquired at an early age. This claim is based on findings showing that the procedure used by adults is determined by the rhythm of their native language rather than by the rhythm of the language they are actually listening to. Thus, once the procedure appropriate to the native language has been acquired, adults use it when listening to stimuli in a foreign language (Cutler et al., 1986; Otake et al., 1993). It has even been shown that proficient bilinguals are dominant in one of their languages when it comes to speech segmentation, and that they have developed a specific rhythmic segmentation procedure in only one of their two languages (Cutler et al., 1992).

The interpretation offered to these findings is that there are different kinds of rhythmic segmentation procedures, each being optimally adapted to the processing of one rhythmic class of languages (Cutler & Mehler, 1993; Otake et al., 1993; see also Sebastián-Gallés et al., 1992; Vroomen et al., 1996), even if more minor differences can be found for languages within a given rhythmic class. This finding that rhythmic properties impact adults' speech processing procedures has influenced the way infants are thought to acquire the segmentation procedures appropriate for their native language. Accordingly, Mehler, Dupoux, Nazzi and Dehaene-Lambertz (1996) proposed that the emergence of rhythmic segmentation relies on infants' early sensitivity to prosody, and more precisely, on linguistic rhythm defined at a non-lexical level (see also Nazzi & Ramus, 2003). In what follows, we present further experimental data relevant to this proposal.

Rhythmic classes and the acquisition of rhythmic properties: studies on early language discrimination

The hypothesis of an early sensitivity to rhythmic classes relies on a number of studies exploring young infants' ability to discriminate between utterances

---

limitations of the syllabic effect in French, this study does not question the existence of crosslinguistic differences in these experimental tasks given that no syllabic effects were found for English, including when using intervocalic liquid consonants.

from different languages. The first studies on this issue concluded that newborns' language discrimination abilities stemmed from their familiarity with their native language (Mehler, Jusczyk, Lambertz, Halsted, Bertoncini & Amiel-Tison, 1988; Moon, Panneton-Cooper & Fifer, 1993). However, later studies have explored whether early language discriminations are in fact based on newborns' sensitivity to the rhythmic properties of utterances, inducing language categorization into a limited number of rhythmic classes.

Accordingly, Nazzi, Bertoncini et Mehler (1998a) presented newborns with different combinations of languages unfamiliar to them. The stimuli were lowpass filtered such as to degrade phonetic information while preserving prosodic information. The rhythmic distance between the languages was systematically manipulated, the contrasted languages belonging either to two different rhythmic classes (stress-based English vs. mora-based Japanese) or to the same class (stress-based English and Dutch). Discrimination was found only for the languages from different classes. Nazzi et al. (1998a) further showed that if newborns are familiarized with utterances from two languages, and then tested with utterances from two other languages, discrimination is found if the languages are arranged in congruence with the rhythmic classes (stress-based English and Dutch vs. syllable-based Spanish and Italian), but not if they are arranged orthogonally to the rhythmic classes (for example: stress-based English and syllable-based Italian vs. stress-based Dutch and syllable-based Spanish). These results were later generalized to the Dutch/Japanese contrast, and stimuli that had been resynthesized to neutralize possible phonetic differences (Ramus, Hauser, Miller, Morris, & Mehler, 2000). Ramus (2002) also showed that the F0 contour plays a marginal role in language discrimination, thus putting further emphasis on the role of rhythm.

Further studies have investigated the changes in language discrimination ability across development. It was found that by 4 or 5 months of age, infants' language discrimination abilities have improved for contrasts involving the native language. Indeed, these infants can, contrary to newborns, discriminate two languages from the same rhythmic classif (a variant of) the native language is presented (stress-based British-English vs. Dutch, and British- vs. US-English for US-English-learning infants, c.f. Nazzi, Jusczyk & Johnson, 2000; syllable-based Spanish vs. Catalan for Spanish-learning infants, c.f. Bosch & Sebastián-Gallés, 1997). If the native language is not presented, then infants fail to discriminate two languages from the same class, whether or not they belong to the native language rhythmic class (stress-based Dutch vs. German, and syllable-based Italian vs. Spanish for English-learning infants, c.f. Nazzi et al., 2000). These changes in language discrimination ability are compatible with the hypothesis that infants progressively tune their rhythmic perceptual skills to the unit of their native language. Recent work in this domain further supports this hypothesis. A study by Hohle et al. (2009) showed for German-learning infants, a preference for trochaic over iambic words at 6 months but not at 4 months, suggesting that this bias emerges

between the two ages. However, the same study failed to find preference for either stress patterns by French-learning 6-month-old infants, albeit they could discriminate the two stress patterns. Although still unknown, the mechanisms that could allow infants to exploit rhythmic differences to specify the segmentation unit appropriate to their native language are currently under investigation (see research based on an adaptive dynamical model, McLennan, 2005).

Evaluation of the early rhythmic segmentation hypothesis

We predict that early rhythmic segmentation will induce the extraction of multisyllabic sequences starting with a stressed syllable in stress-based languages such as English, and isolated syllables in syllable-based languages such as French. Accordingly, we predict that a French bisyllabic word (e.g., *toucan*) will initially be segmented as two independent syllabic units (as opposed to smaller moraic units or larger stress units). In the following, we first review what was known of early segmentation in languages other than English when we started our project, and then present work that we have been conducting on this issue in French, and some ideas for new experiments to be conducted in the future.

Early word form segmentation in light of the *early rhythmic segmentation hypothesis*

When we started working on this issue, data supporting the rhythmic hypothesis was mainly limited to evidence of early trochaic segmentation in English-learning infants (Jusczyk et al., 1999b). Studies investigating languages other than English had focused on the ages at which segmentation appears rather than on the mechanisms underlying segmentation, or the types of words segmented.

With respect to other stress-based languages, Dutch-learning infants had been found to start segmenting trochaic words between 7.5 and 9 months of age (Houston & Jusczyk, 2000; Kooijman, Hagoort & Cutler, 2005; Kuijpers, Coolen, Houston & Cutler, 1998). However, we postulated, according to our hypothesis, that these infants should start segmenting iambic words, at a later age. The finding with German-learning infants of segmentation effect of monosyllabic words between 6 and 8 months of age (Höhle & Weissenborn, 2003), did not seem to allow an evaluation of our hypothesis. Nevertheless, more recent research suggests the existence of a trochaic bias in German, similar to the one previously found for English: German learning infants have been found to segment trochaic words by 9 months, while they still fail to segment iambic words by 11 months (Höhle & Weissenborn, 2005).

With respect to syllable-based languages, only two studies had been conducted investigating French: one with Parisian infants (Gout, 2001) and

one with Canadian-French infants (Polka & Sundara, 2003). Gout (2001) study had shown that Parisian infants segment monosyllabic words at 7.5 months but could not establish that they segment bisyllabic words between 7.5 and 11 months of age, bringing initial support for syllabic segmentation in syllable-based languages. In order to test more specifically this hypothesis, and in particular the prediction that bisyllabic words are first segmented as independent syllables and only later as whole units, Nazzi, Iakimova, Bertoncini, Frédonie and Alcantara (2006) conducted the following study.

Syllables and early segmentation in French

Nazzi et al. (2006) investigated how bisyllabic words, inserted in fluent passages, are segmented by infants during development (following the procedure used in previous studies on English and French). Remember that the prediction was that when segmentation abilities emerge, the rhythmic procedure should place boundaries between each two consecutive syllables, no other segmentation cue (e.g., distributional regularities of syllable order) being yet available to attach the consecutive syllables of bisyllabic words. Accordingly, Parisian French-learning infants were tested with four passages each containing a different target (each passage was made up of 6 sentences, each sentence containing one occurrence of the target word). The target words were *putois*, *toucan*, *bandeau* and *guidon*. Two of these passages corresponded to the items previously presented in the familiarization phase. The familiarization items differed across experiments. Infants were either familiarized with two bisyllabic words (e.g., *putois* and *toucan*), or with their final syllables (e.g., *tois* and *can*) or their initial syllables (e.g., *pu* and *tou*). Infants were tested at 8, 12 and 16 months of age in the first two conditions, and at 12 months of age in the initial syllable condition (see Figure 3).

First of all, no segmentation effects were obtained at 8 months, whether infants had been familiarized with whole words or with their final syllables. These results suggested a possible delay in the emergence of segmentation abilities in French compared to English, a question further discussed below.

However, the other results corroborate the hypothesis of early syllable--based rhythmic segmentation in French. At 12 months, no segmentation effect was found following whole word familiarization, while a segmentation effect was found following final syllable familiarization and, under certain conditions, initial syllable familiarization.[6] Moreover, the opposite pattern was

---

[6] A segmentation effect for the initial syllable was found only when infants were familiarized with syllables spliced-out from the test passages, while there was only a non-significant trend when infants were familiarized with initial syllables recorded in isolation. This difference in results underlines one crucial element enabling the finding of segmentation results with the present paradigm: the need to match the acoustically different familiarization and test targets (see also Houston & Jusczyk, 2000). Such a match was probably made difficult in the isolated initial

observed at 16 months. A segmentation effect emerged after whole word familiarization but could not be found any longer after final syllable familiarization. Taken together, these results show that at 12 months, French-
-learning infants independently segment the two syllables of a bisyllabic word, in spite of the fact that these syllables always appear consecutively in the signal (see Figure 3).
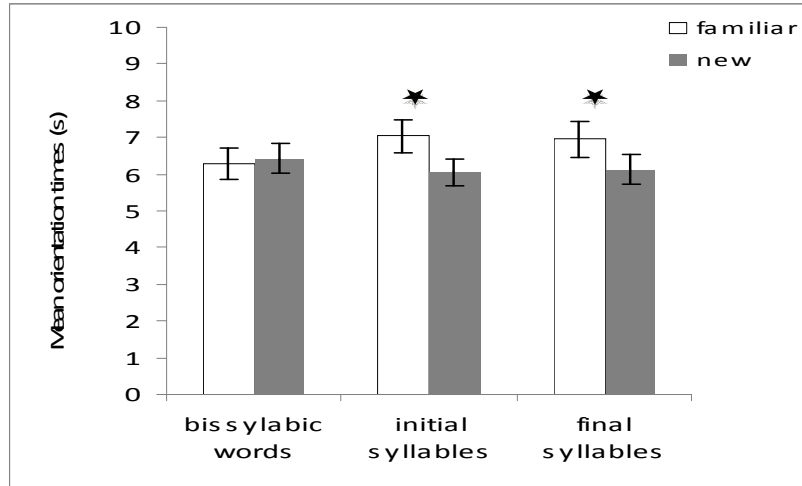

Figure 3: Mean orientation times (and standard error) to the test passages containing the familiar(ized) versus new items for 12-month-old French-learning infants, broken down according to the type of familiarization items: whole bisyllabic word (left panel); initial syllables (central panel); final syllables (right panel).

Yet, by 16 months, this distributional information (and probably other segmentation cues, to be explored in future research) seems to be taken into account, when infants recognize the whole words but now fail to recognize separately their final syllables. This developmental change in segmentation pattern is similar to the one found between 7.5 and 10.5 months for iambic (though not trochaic) words in English (Jusczyk et al., 1999b), except that it extends to the independent segmentation of both syllables in French, rather than just the final stressed syllable in English.

The study by Nazzi et al. (2006) brings the first direct piece of evidence supporting the hypothesis that infants learning French initially rely on the rhythmic unit of French (the syllable) to segment fluent speech, as infants learning English rely on the rhythmic unit of English (the trochaic unit). Note

syllable condition given large acoustic differences between the familiarization and test items (differences that were much larger than for the whole word and final syllable conditions).

that the observation of (1) a syllabic effect for monosyllables at 7.5 months, and (2) no syllabic effect for individual syllables of bisyllabic words at 8 months, might be the result of increased coarticulation between the two syllables of the bisyllabic words. It is then possible that syllabic segmentation in French is not as delayed as suggested by Nazzi et al (2006) results.

Regarding the studies with infants learning Canadian French, segmentation of bisyllabic words was evaluated at 8 months in two separate experiments, by presenting stimuli either in Canadian or Parisian French (Polka & Sundara, 2003). Segmentation was tested by familiarizing infants with two bisyllabic words, and then presenting them two passages containing occurrences of the familiarized words, and two passages containing other bisyllabic words. Similar segmentation results were obtained with both dialects, as attested by the observation of a significant preference for the passages with the familiarized words. This study therefore suggested that Canadian French infants segment whole bisyllabic words as early as 8 months of age. These results appear at first sight in contradiction with those by Nazzi et al. (2006), given that the data on Canadian French infants established segmentation of bisyllabic words as early as 8 months when presented with stimuli recorded either by a Canadian or a Parisian French speaker. However, it is important to note that syllabic segmentation was not evaluated in Canadian French infants, leaving open the possibility for syllable segmentation at, or before, 8 months of age.

The above possibility is actually supported by recent experiments in which Canadian infants were tested on the original stimuli by Nazzi et al. (2006). Results show that Canadian French infants actually perform similarly to the Parisian French infants: whole word segmentation is found to emerge between 12 and 16 months, while 12-month-olds show evidence of segmenting the final syllables (Polka et al., 2008). A parallel investigation (Nazzi, Mersad, Iakimova, Sundara, Polka, 2008), in which Parisian infants were tested on the Canadian and Parisian French stimuli used by Polka & Sundara (2003), confirmed the failure to find whole word segmentation effects at 8 months when infants were familiarized with words and then tested with passages (word-passage order). However, when 8-month-olds were tested in the reversed passage-word order, whole word segmentation was found for the Parisian stimuli after 30 seconds of familiarization with the passages containing the bisyllabic words. Success in this condition might result from the fact that hearing the passages during the familiarization gave infants more time, after performing an initial syllable-based segmentation, to compute some distributional analysis of syllable order, allowing them to succeed in grouping together the syllables that co-occur in the signal. Note also that when 8-month-old infants were tested in the passage-word order on the Canadian French stimuli, a segmentation effect was found after 45 seconds of familiarization, while no effect emerged after 30 seconds. This suggests a cost in adjusting to a different dialect of the native language.

These new crossdialectal studies establish that the segmentation advantage of Canadian French infants over Parisian infants is less important than

suggested by Polka and Sundara (2003) and Nazzi et al. (2006) results. The advantage for Canadian French infants might result from more intonation variability in Canadian French (Ménard, Ouellon, & Dolbec, 1999) which could help locate word boundaries. This would be compatible with earlier data suggesting that pitch variations affect segmentation (Nazzi et al., 2005). Second, these new studies importantly confirm the role that the syllable plays in early speech segmentation, both for Canadian and Parisian French infant populations. Lastly, they also show that the HPP technique is highly sensitive to small methodological changes, such as the order of presentation of the stimuli (word-passage versus passage-word) or the duration of the familiarization phase.

Conclusions and perspectives

The studies reviewed so far demonstrate that infants learning French start segmenting speech by relying on the rhythmic unit of their native language, that is, the syllable. Similarly, infants learning English start segmenting by relying on the rhythmic unit of their native language, that is, the trochaic unit. This cross-linguistic pattern of early speech segmentation, underlines, in the one hand, the universality of segmentation procedures (use of rhythmic information) and, on the other hand, the fact that the rhythmic segmentation procedure is instantiated differently in different languages (use of different rhythmic units). We also presented evidence suggesting that the acquisition of the rhythmic unit takes place before the emergence of segmentation abilities, thus independently of the lexical level, and might be the element enabling the emergence of rhythmic segmentation procedures. So far, the available data support the *early rhythmic segmentation hypothesis* (Nazzi et al., 1998a, 2000, 2006).

At this point, we would like to conclude by briefly discussing three pending issues regarding the emergence of segmentation abilities. First, research will need to conduct further crosslinguistic exploration of the emergence of segmentation abilities. While there is evidence of early trochaic segmentation from different stress-based languages, namely English, Dutch and German, there is only evidence of syllable-based segmentation from French-learning infants, and no data on mora-based segmentation. Evaluation of the rhythmic-based segmentation hypothesis (and potential modifications of this proposal) will require more data from a broader range of languages. European Portuguese would be an interesting case study. Indeed, while rhythmic cues point to both *syllable-timing* and *stress-timing* properties (Vigário et al., 2003), these authors have proposed, on the basis of adult discrimination data and analyses of syllabic structure, that European Portuguese clusters with the syllable-based languages. If correct, then syllabic segmentation effects should be obtained in that language.

Moreover, the comparison of early segmentation abilities by Parisian and Canadian French infants revealed crossdialectal differences that would be

worth exploring in other languages. Preliminary evidence from British English infants suggest that these infants might show a different developmental pattern than American English infants. Indeed, 7.5-month-old British English infants tested with British English stimuli (from Nazzi, Paterson, & Karmiloff-Smith, 2003) failed to show evidence of segmenting strong-weak words, an effect found at the same age in American English--learning infants (Vihman, de Paolis & Nazzi, in preparation).

The second issue is related to the fact that so far, the studies on French have focused on the sole use of rhythmic information. The studies on English have revealed that infants can use different cues (rhythm, syllabic distribution information, allophony, phonotactics, known words, …) for segmentation, and that the relative weight given to these cues changes with development. Given the likelihood that the developmental trajectory of use of these cues will vary across languages, it will be important to study their use by French-learning infants. This is even more important as we have argued that the changes observed by Nazzi et al. (2006) between 12 and 16 months of age are due to decreased weight given to rhythmic cues over this developmental period. Future studies will thus have to continue tracing the pattern of emergence of segmentation cues in French, testing in doing so the hypothesis by Mattys et al. (2005) according to which infants start using sublexical suprasegmental cues (prosody), followed by sublexical segmental cues (phonotactics, allophony, coarticulation) followed by lexical (semantic, syntactic, pragmatic) cues.

The third issue is more methodological, but has theoretical implications. We presented evidence showing that HPP is a method sensitive to small methodological changes, and that this sensitivity might even vary according to the language tested: reversing the order of presentation of the isolated words and the passages did not affect performance in English (Jusczyk et al., 1999) but critically did so for Parisian French infants (Nazzi et al., 2008). Thus, it would be important to assess segmentation abilities using different methods, and ERPs have recently started to be used to explore this issue. A series of studies on Dutch-learning infants (Kooijman, 2007; Kooijman, Hagoor, & Cutler, 2005) established that both 7- and 10-month-old infants are able to segment bisyllabic strong-weak words, even though no segmentation evidence could be found using HPP in 7-month-olds. Moreover, ERPs further revealed that when hearing bisyllabic weak-strong words, Dutch-learning infants are reacting to the onset of the strong syllable, rather than to the onset of the word-initial weak syllable. These results are compatible with the use of the trochaic unit for word segmentation in this stress-based language. Similar studies conducted on French-learning infants at 12 months revealed a very different pattern: these infants were found to react to the onset of both the first and the second syllable of bisyllabic words, a pattern of data compatible with the use of the syllabic unit for word segmentation in this syllable-based language (Goyet, de Schonen & Nazzi, 2010). Interestingly, although the unit to which the infants responded in the two languages was different (the

trochaic versus the syllabic unit), the responses observed were similar, and involved a more negative deflection for the familiarized items around 350-500 ms of stimulus onset. So to conclude, the ERP data (like the HPP data) so far support the predictions of the *early rhythmic segmentation hypothesis*. But, as mentioned earlier for HPP, future ERP research will have to be extended to the exploration of more languages from the different rhythmic classes.

Acknowledgements:

References

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: University of Edinburgh Press.

Arvaniti, A. (1994). Acoustic features of Greek rhythmic structure. *Journal of Phonetics,* **22**, 239-268.

Aslin, R.N. (1993). Segmentation of fluent speech into words: learning models and the role of maternal input. In B. de Boysson-Bardies et al. (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life*, pp. 305-315 Kluwer.

Bortfeld, H., Morgan, J.L., Golinkoff, R.M., & Rathbun, K. (2005). Mommy and me: familiar names help launch babies into speech-stream segmentation. *Psychological Science,* **16,** 298-304.

Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4--month-old infants from monolingual and bilingual environments. *Cognition,* **65**, 33-69.

Brent, M.R., & Cartwright, T.A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition,* **61**, 93-125.

Brent, M.R., & Siskind, J.M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition,* **81**, B33-B44.

Cassidy, K.W., & Kelly, M.H. (1991). Phonological information for grammatical category assignments. *Journal of Memory and Language,* **30***,* 348-369.

Chevrot, J. P., Dugua, C., & Fayol, M. (2005). Liaison et formation des mots en français: un scénario développemental. *Langages***, 158**, 38-52.

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.

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America,* **95**, 1570-1580.

Cole, R.A. & Jakimik, J. (1978). Understanding speech: How words are heard. In G. Underwood (Ed.), *Strategies of information processing*. New York, NY: Academic Press.

Cole, R.A., & Jakimik, J. (1980). How are syllables used to recognize words? *Journal of the Acoustical Society of America,* **67**, 965-970.

Content, A., Meunier, C., Kearns, R.K., & Frauenfelder, U.H. (2001). Sequence detection in pseudowords in French: Where is the syllable effect? *Language and Cognitive Processes,* **16**, 609-636.

Curtin, S., Mintz, T.H., & Christiansen, M.H. (2005). Stress changes the representational landscape: evidence from word segmentation. *Cognition,* **96**, 233-262.

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. & Mehler,(1993). The periodicity bias. *Journal of Phonetics,* **21**, 103-108.

Cutler, A., Mehler, J., Norris, D. G., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, **25**, 385-400.

Cutler, A., Mehler, J., Norris, D. G., & 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.

Dahan, D., & Brent, M. R. (1999). On the discovery of novel wordlike units from utterances: An artificial-language study with implications for native-language acquisition. *Journal of Experimental Psychology: General,* **128**, 165-185.

Delattre, P.C. (1966). A comparison of syllable length conditioning among languages. *International Journal of Applied Linguistics,* **4**, 182-198.

den Os, E. (1988). *Rhythm and tempo of Dutch and Italian: a contrastive study*. Utrecht, Holland: Drukkerij Elinkwijk BV.

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.

Fant, G., Kruckenberg, A., & Nord, L. (1991). Durational correlates of stress in Swedish, French and English. *Journal of Phonetics*, **19**, 351-365.

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development,* **10**, 279-283.

Finn, A. & Kam, C. (2008). The curse of knowledge firstlanguage knowledgeimpairs adult learners'use of novel statistics for word segmentation. *Cognition*, **108**, 477-499.

Fletcher, J. (1991). Rhythm and final lengthening in French. *Journal of Phonetics, 19*, 193-212.

Fiser, J., & Aslin, R.N. (2002). Statistical learning of higher-order temporal structure from visual shape-sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **28**, 458-467.

Friederici, A.D., Friedrich, M., & Christophe, A. (2007). Language experience affects brain responses in 4-month-old infants. *Current Biology,* **17**, 1208–1211.

Friederici, A.D. & Wessels, J.M.I. (1993). Phonotactic knowledge and its use in infant speech perception. *Perception & Psychophysics,* **54**, 287-295.

Gout, A., (2001). *Etapes précoces de l'acquisition du lexique*. Unpublished Dissertation, Ecole des Hautes Etudes en Sciences Sociales, Paris, France.

Gout, A., Christophe, A., & Morgan, J.L. (2004). Phonological phrase boundaries constrain lexical access II. Infant data. *Journal of Memory and Language,* **51**, 548--567.

Goyet, L., de Schonen, S. & Nazzi, T. (2010). Syllables in word segmentation by French-learning infants: an ERP study. *Brain Research*, **1332**, 75-89.

Graf Estes, K., Evans, J.L., Alibali, M.W. & Saffran, J.R. (2007). Can infants map meaning to newly segmented words? *Psychological Science,* **18**, 254-260.

Hauser, M.D., Newport, E.L., & Aslin, R.N. (2001). Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition,* **78**, B53-B64.

Havy, M., & Nazzi, T. (2009). Better processing of consonantal over vocalic information in word learning at 16 months of age. *Infancy*, **14**, 439-456.

Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Wright Cassidy, K., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition,* **26**, 269-286.

Hockema, S.A. (2006). Finding Words in Speech: An Investigation of American English. *Language Learning and Development,* **2**, 119-146.

Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). The development of language specific prosodic preferences during the first half year of life: evidence from German and French. *Infant Behavior and Development,* **32**, 262-274.

Höhle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science,* **6,** 122-127.

Höhle, B., & Weissenborn, J. (2005). Word segmentation in German-learning infants. Paper presented at the International workshop "Early word segmentation: a crosslinguistic approach taking advantage of Europe's linguistic diversity," Paris, France, February 25-26.

Hohne, E.A., & Jusczyk, P.W. (1994). Two-month-old infants' sensitivity to allophonic differences. *Perception & Psychophysics,* **56**, 613-623.

Houston, D.M. & Jusczyk, P.W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance,* **26,** 1570-1582.

Houston, D.M. & Jusczyk, P.W. (2003). Infants' long-term memory for the sound patterns of words and voices. *Journal of Experimental Psychology: Human Perception and Performance*, **29,** 1143-1154.

Houston, D.M., Jusczyk, P.W., Kuijpers, C., Coolen, R., & Cutler, A. (2000). Cross--language word segmentation by 9-month-olds. *Psychonomics Bulletin & Review,* **7**, 504-509.

Houston, D.M., Santelmann, L.M., & Jusczyk, P.W. (2004). English-learning infants' segmentation of trisyllabic words from fluent speech. *Language and Cognitive Processes,* **19**, 97-136.

Johnson, E.K. & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, **44**, 1--20.

Johnson, E.K. & Tyler, M.D. (2010). Testing the limits of statistical learning for word segmentation. *Developmental Science*, **19**, 339-345.

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. (1993a). Preference for the predominant stress patterns of English words. *Child Development,* **64***,* 675-687.

Jusczyk, P.W., Friederici, A.D., Wessels, J., Svenkerud, V.Y. & Jusczyk, A.M. (1993b). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language,* **32***,* 402-420.

Jusczyk, P. W., Hirsh-Pasek, K., Kemler Nelson, D. G., Kennedy, L., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology,* **24**, 252-293.

Jusczyk, P.W. & Hohne, E.A. (1997). Infants' memory for spoken words. *Science,* **277**, 1984-1986.

Jusczyk, P.W., Hohne, E.A., & Bauman, A. (1999a). Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics,* **62**, 1465-1476.

Jusczyk, P.W., Houston, D.M., & Newsome, M. (1999b). The beginning of word segmentation in English-learning infants. *Cognitive Psychology,* **39**, 159-207.

Jusczyk, P.W., Luce, P.A. & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language,* **33***,* 630-645.

Karzon, R.G., & Nicholas, J.G. (1989). Syllabic pitch perception in 2- to 3-month-old infants. *Perception and Psychophysics,* **45,** 10-14.

Kelly, M. H. (1992). Using sound to solve syntactic problems: the role of phonology in grammatical category assignment. *Psychological Review*, **99**, 349-364.

Kelly, M.H., & Bock, J.K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance,* **14***,* 389-403.

Kim, J., Davis, C., & Cutler, A. (2008). Perceptual tests of rhythmic similarity: II. Syllable rhythm. *Language and Speech,* **51**, 343-359.

Klatt, D.H. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics,* **7***,* 279-312.

Klatt, D.H. (1989). Review of selected models of speech perception. In W. Marslen--Wilson (Ed.), *Lexical Representation and Process,* pp. 169-226. Cambridge, MA: MIT Press.

Kooijman, V. (2007). Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. PhD-Thesis. Radboud University, Nijmegen

Kooijman, V., Hagoort, P., & Cutler, A. (2005). Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. *Cognitive Brain Research,* **24**, 109-116.

Kuhl, P. K., & Miller, J.D. (1982). Discrimination of auditory target dimensions in the presence or absence of variation in a second dimension by infants. *Perception and Psychophysics,* **31**, 279-292.

Kuijpers, C., Coolen, R., Houston, D., & Cutler, A. (1998). Using the head-turning technique to explore cross-linguistic performance differences. *Advances in Infancy Research (Vol. 12),* (pp. 205-220). Stamford, CT: Ablex.

Mattys, S., & Jusczyk, P.W. (2001a). Phonotactic cues for segmentation of fluent speech by infants. *Cognition,* **78**, 91-121.

Mattys, S., & Jusczyk, P.W. (2001b). Do infants segment words or recurring contiguous patterns? *Journal of Experimental Psychology: Human Perception and Performance,* **27,** 644-655.

Mattys, S., 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., White, L., Melhorn, J.F. (2005). Integration of Multiple Speech Segmentation Cues: A Hierarchical Framework. *Journal of Experimental Psychology: General,* **134**, 477-500.

McLennan, S. (2005). A dynamic adaptive model of linguistic rhythm. In *Proceedings of the I.S.C.A. Workshop on Plasticity in Speech Perception* pp. 130. London, U.K., June 15-17 2005.

McQueen, J.M., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **20**, 621-638.

Mehler, J., Dommergues, J. Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior,* **20**, 298-305.

Mehler, J., Dupoux, E., Nazzi, T. & Dehaene-Lambertz, G. (1996). Coping with linguistic diversity: The infant's viewpoint. In J.L. Morgan & K. Demuth (Eds.), *Signal to syntax* (pp. 101-116). Mahwah, N.J.: Lawrence Erlbaum Associates.

Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, **29**, 144-178.

Ménard, L., Ouellon, C., & Dolbec, J. (1999). Prosodic markers of regional group membership: The case of French of Quebec versus France. In *Proceedings of the XIVth International Congress of Phonetic Sciences,* San Francisco pp. 1601-1604.

Mersad, K., Nazzi, T. (submitted). Transitional probabilities and phonotactics in a hierarchical model of speech segmentation.

Moon, C., Panneton-Cooper, R. & Fifer, W.P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development,* **16**, 495-500.

Morgan, J.L. & Saffran, J.R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development,* **66**, 911-936.

Murty, L., Otake, T., & Cutler, A. (2007). Perceptual tests of rhythmic similarity: I. Mora rhythm. *Language and Speech,* **50**, 77-99.

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: differences between consonants and vowels. *Cognition*, **98**, 13-30.

Nazzi, T., & Bertoncini, J. (2003). Before and after the vocabulary spurt: two modes of word acquisition? *Developmental Science,* **6**, 136-142.

Nazzi, T., Bertoncini, J., & Bijeljac-Babic, R. (2009). A perceptual equivalent of the labial-coronal effect in the first year of life. *Journal of the Acoustical Society of America*, **126**, 1440-1446.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998b). 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., Dilley, L.C., Jusczyk, A.M., Shattuck-Hunagel, S., & Jusczyk, P.W. (2005). English-learning infants' segmentation of verbs from fluent speech. *Language and Speech,* **48**, 279-298.

Nazzi, T., Floccia, C., & Bertoncini, J. (1998b). Discrimination of pitch contours by neonates. *Infant Behavior and Development,* **21**, 779-784.

Nazzi, T., Iakimova, I., Bertoncini, J., Frédonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: emerging evidence for crosslinguistic differences. *Journal of Memory and Language,* **54**, 283-299.

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.

Nazzi, T., Kemler Nelson, D.G., Jusczyk, P.W., & Jusczyk, A.M. (2000). Six-month- -olds' detection of clauses in continuous speech: Effects of prosodic well- -formedness. *Infancy,* **1,** 123-147.

Nazzi, T., Mersad, K., Iakimova, G., Sundara, M., & Polka, L. (2008). Differences in the development of speech segmentation abilities in two French dialects. Paper presented at the *XIth International Congress for the Study of Child Language (IASCL)*, Edinburgh, United Kingdom, July 28 – August 1 2008.

Nazzi, T. & New, B. (2007). Beyond stop consonants: consonantal specificity in early lexical acquisition. *Cognitive Development,* **22**, 271-279.

Nazzi, T., Paterson, S. & Karmiloff-Smith, A. (2003). Early word segmentation by infants with Williams syndrome. *Infancy,* **4**, 251-271.

Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication,* **41**, 233-243.

Nespor, M., & Vogel, I. (1986). Prosodic Phonology. Dordrecht: *Foris Publication.*

Newman, R., Bernstein Ratner, N., Jusczyk, A.M., Jusczyk, P.W., & Dow, K.A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. *Developmental Psychology,* **42**, 643-655.

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language,* **32**, 258-278.

Pelucchi, B., Hay, J.F, & Saffran, J. (2009), Statistical Learning in a Natural Language by 8-month-Old Infants. *Child Developmen*t, **80**, 674-685.

Peretz, I., Lussier, I., & Béland, R. (1998). The differential role of syllabic structure in stem completion for French and English. *European Journal of Cognitive Psychology,* **10**, 75-112.

Perruchet, P., & Vinter, A. (1998). PARSER: a model for word segmentation. *Journal of Memory and Language,* **39**, 246-263.

Pike, K. (1945). *The intonation of American English*. Ann Arbor, MI: University of Michigan Press.

Polka, L., Proulx, J., Mersad, K., Iakimova, G., Sundara, M., & Nazzi, T. (2008). Speech segmentation in French-learning infants: Language-specific and dialect- -specific patterns. Paper presented at the *16th Biennal Meeting of the International Society for Infant Studies*, Vancouver, Canada, March 27-29 2008.

Polka, L. & Sundara, M. (2003). Word segmentation in monolingual and bilingual infant learners of English and French. *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Spain, 1021-1024.

Ramus, F. (2002). Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues. *Annual Review of Language Acquisition, 2*, 85--115.

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., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*, 265-292.

Saffran, J.R., Aslin, R.N. & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science, 274,* 1926-1928.

Saffran, J.R., Johnson, E., Aslin, R.N. & Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, 70*, 27-52.

Sebastián-Gallés, N. & Bosch, L. (2002). Building phonotactic knowledge in bilinguals: role of early exposure. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 974-989.

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.

Seidl, A. & Johnson, E.K. (2006). Infant word segmentation revisited: edge alignment facilitates target extraction. *Developmental Science, 9*, 565 –573.

Shafer, V. L., Shucard, D. W., & Jaeger, J. J. (1999). Electrophysiological indices of cerebral specialization and the role of prosody in language acquisition in three--month-old infants. *Developmental Neuropsychology,***15**, 73-109.

Shi, R., & Lepage, M. (2008). The effect of functional morphemes on word segmentation in preverbal infants.*Developmental Science*. **11**, 407-413

Shi, R., Werker, J.F., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy,* **10**, 187-198.

Shukla, M., Nespor, M., and Mehler, J. (2007) An interaction between prosody and statistics in the segmentation of fluent speech*. CognitivePsychology* **54,** 1-32.

Singh, L., Morgan, J.L., & White, K.S. (2004). Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language, 51*, 173-189.

Skoruppa, K., Pons, F., Christophe, A., Bosch, L., Dupoux, E., Sebastián-Gallés, N., Alves Limissuri, R., & Peperkamp, S. (2009). Language-specific stress perception by 9-month-old French and Spanish infants. *Developmental Science*, **12**, 914-919.

Soderstrom, M., Seidl, A., Kemler Nelson, D. G., & Jusczyk, P.W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language, 49*, 249-267.

Spring, D. R. & Dale, P. S. (1977). Discrimination of linguistic stress in early infancy. *Journal of Speech and Hearing Research, 20*, 224-232.

Swingley, D. (2005). Statistical clustering and the contents of the infant vocabulary. *Cognitive Psychology, 50*, 86-132.

Thiessen, E.D., Hill, E.A., & Saffran, J.R. (2005). Infant-directed speech facilitates word segmentation. *Infancy, 7,* 53-71.

Thiessen, E.D. & Saffran, J.R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology, 39,* 706-716.

Tincoff, R., & Jusczyk, P. W. (1996). Are word-final sounds perceptually salient for infants? Paper presented at the *Fifth Conference on Laboratory Phonology*, Evanston, IL, July, 1996

Toro, J.M., & Trobalon, J. (2005). Statistical computations over a speech stream in a rodent. *Perception and Psychophysics,* **67**, 867-875.

Turk, A.E., Jusczyk, P.W., Gerken, L. (1995). Do English-learning infants use syllable weight to determine stress? *Language and Speech,* **38**, 143-158.

van de Weijer, J. (1998). *Language input for word discovery.* Doctoral dissertation, University of Nijmegen (MPI Series in Psycholinguistics, 9).

Vigário, M., Frota, S., & Freitas, M.J. (2003). From signal to grammar: Rhythm and the acquisition of syllable structure. *Proceedings of the 27th Annual Boston University Conference on Language Development. Dommerville, Mass.: Cascadilla Press*, 809-821

Vihman, M., de Paolis, R., & Nazzi, T. (in preparation). Early word segmentation by British English infants.

Vroomen, J., van Zon, M., & de Gelder, B. (1996). Metrical segmentation and inhibition in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance,* **21**, 98-108.

Wauquier-Gravelines, S. (2002). Segmentation de la chaîne parlée et determination des frontiers gauches dans l'acquisition du lexique précoce. *Rencontres Internationales du GDR Phonologies*, Grenoble, 6-8 juin 2002.

Karima Mersad
Laboratoire Psychologie de la Perception,
Université Paris Descartes – CNRS
Institut Biomédical des Saints Pères
45 rue des Saints-Pères
75006 Paris, France
karima.mersad@etu.parisdescartes.fr

Louise Goyet
Laboratoire Psychologie de la Perception,
Université Paris Descartes – CNRS
Institut Biomédical des Saints Pères
45 rue des Saints-Pères
75006 Paris, France
louise.goyet@gmail.com

Thierry Nazzi
Laboratoire Psychologie de la Perception,
Université Paris Descartes – CNRS
Institut Biomédical des Saints Pères
45 rue des Saints-Pères
75006 Paris, France
thierry.nazzi@parisdescartes.fr