ELSEVIER

Speech Communication 41 (2003) 233–243

SPEECH COMMUNICATION

www.elsevier.com/locate/specom

Perception and acquisition of linguistic rhythm by infants

Thierry Nazzi [a,*], Franck Ramus [b]

[a] *Laboratoire Cognition et Développement, CNRS, Université Paris 5, 71, Avenue Edouard Vaillant, 92774 Boulogne Billancourt Cedex, France*

[b] *Institute of Cognitive Neuroscience, University College London, UK*

Abstract

In the present paper, we address the issue of the emergence in infancy of speech segmentation procedures that were found to be specific to rhythmic classes of languages in adulthood. These metrical procedures, which segment fluent speech into its constitutive word sequence, are crucial for the acquisition by infants of the words of their native language. We first present a prosodic bootstrapping proposal according to which the acquisition of these metrical segmentation procedures would be based on an early sensitivity to rhythm (and rhythmic classes). We then review several series of experiments that have studied infants' ability to discriminate languages between birth and 5 months, in an attempt to specify their sensitivity to rhythm and the implication of rhythm perception in the acquisition of these segmentation procedures. The results presented here establish infants' sensitivity to rhythmic classes (from birth onwards). They further show an evolution of infants' language discriminations between birth and 5 months which, though not inconsistent with our proposal, nevertheless call for more studies on the possible implication of rhythm in the acquisition of the metrical segmentation procedures.

1. Introduction

The focus of the present paper is on the issue of the segmentation of fluent speech into words, and more particularly on the development of speech segmentation procedures during infancy. For adults, speech segmentation involves language-specific phonological procedures (see below, and Cutler, McQueen and Norris, this volume) that allow for the retrieval of the acoustic sound patterns of words from fluent speech, and the connecting of these sound patterns to the lexical representations stored in the lexicon. Hence for adults, this task could be facilitated by the lexicon itself, making it a problem of word recognition as well as (or rather than) a problem of word segmentation. However, the segmentation task has to be different for the infant who, starting with no lexicon and no language-specific phonological knowledge, has to discover, rather than recognize, the words in the input and learn the speech segmentation procedure appropriate to the language to be learnt. In the following, we present several studies that have recently investigated the issue of word segmentation in early infancy. These studies have first established that speech segmentation emerges between 6 and 7.5 months of age in English-learning American infants, hence months

* Corresponding author. Tel.: +33-1-5520-5992; fax: +33-1-5520-5985.

E-mail address: thierry.nazzi@psycho.univ-paris5.fr (T. Nazzi).

0167-6393/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0167-6393(02)00106-1

before the onset of lexical comprehension and production (Jusczyk and Aslin, 1995). They have then set out to specify the kind of information that infants rely on to postulate word boundaries in fluent speech.

Some of these studies have established that young American infants are sensitive to different kinds of potential word-boundary markers that are language-specific: allophonic information (the fact that the distribution of allophones within words is position-dependent), phonotactic information (the fact that some but not all phonemic sequences are legal at the lexical level), and prosodic/metrical information (the fact that lexical stress is predominantly word initial in English). A sensitivity to allophonic differences was found in infants as young as 2 months of age (Hohne and Jusczyk, 1994), as attested by their ability to discriminate between pairs such as "nitrate" and "night rate". Infants have also been found to become sensitive to phonotactic properties of their native language between 6 and 9 months of age (Friederici and Wessels, 1993; Jusczyk et al., 1993b, 1994; Mattys et al., 2001). This is shown by the emergence of a preference for legal or frequent sequences of phonemes in their native language with respect to illegal or infrequent ones (e.g., the frequent and infrequent non-word sequences "chun" and "yush"). Last but not least, a preference for words with the predominant English strong–weak stress pattern (e.g., "porter") over less frequent weak–strong words (e.g., "report") also emerges between these two ages (Jusczyk et al., 1993a; Turk et al., 1995), revealing the emergence of a sensitivity to native word stress patterns.

Other studies have investigated whether, once they are sensitive to these language-specific markers, infants actually use them to infer word-like units. First, some studies showed that although 9-month-old infants only use allophonic cues to locate familiar words in fluent speech when they are guided by distributional cues, 10.5-month-old can rely on these sole allophonic cues (Jusczyk et al., 1999a). Moreover, a recent study has shown that the fact of providing infants with word-boundary phonotactic information helps them extract familiarized words from fluent speech (Mattys and Jusczyk, 2001).

Finally, it was shown that 9- to 10.5-month-old American infants rely on the typical stress pattern of English to group syllables into word-like units (Morgan and Saffran, 1995) and to remember familiar sequences of syllables (Echols et al., 1997). The ability to retrieve familiar words from fluent speech was also found to depend on their stress pattern. Indeed, Jusczyk et al. (1999b) found that infants begin segmenting strong–weak nouns (e.g., "doctor" and "candle") from fluent speech at 7.5 months, but begin segmenting weak–strong nouns (e.g., "guitar" and "beret") only at 10.5 months. These authors suggested that this processing advantage of strong–weak words might result from the specification of the predominant stress pattern of English, and the emergence by 7.5 months of a segmentation procedure, appropriate for English, and based on its metrical properties. Following this procedure, infants would place a word boundary before the occurrence of every strong syllable in the speech stream, allowing for the detection of strong–weak but not weak–strong words. Note that this procedure is language specific, as it would not be appropriate for the acquisition of French in which the metrical structure is different. [1] Later, that is between 7.5 and 10.5 months, infants would become sensitive to the (language-general) distributional properties of syllables in the speech stream, which would then allow them to segment weak–strong words by grouping these two syllables together. Note that further support that infants can use statistical regularities in the order of syllables forming a continuous sequence to build cohesive word-like units comes from a study by Saffran et al. (1996) on 8-month-old infants.

Hence, the studies above provide good evidence that the very onset of word learning/segmentation is under the influence of language-specific segmentation procedures, particularly a procedure based on the prosodic/metrical properties of the

[1] Unfortunately, it is not possible to specify the kind of segmentation procedures developed by infants growing up in a French-speaking environment, given that at this point all the studies on the emergence of speech segmentation were done with infants learning (American) English, with the exception of one study with Dutch-learning infants (Kuijpers et al., 1998).

native language (the strong–weak segmentation procedure for English). At this point, and in spite of its importance, this finding is problematic unless we explain how infants come to specify these phonological properties of their native language. Indeed, if we are trying to explain how infants come to start segmenting fluent speech, we cannot say that they use some 'knowledge' of the fact that words are stressed initially in their language unless we can explain how they discovered that fact independently of segmentation: a classical bootstrapping problem. In this paper, we propose the existence of such an independent mechanism for the acquisition of the metrical segmentation procedure. This proposal was initially inspired by data from the adult speech segmentation literature and the phonetic literature, which we review in the following. Then, we propose a prosodic bootstrapping account of the acquisition of the metrical segmentation procedures, and then turn to recent data that we have started to gather in support of this proposal.

Several studies have looked at the way adults segment fluent speech. These studies indicated that speech segmentation is influenced by the metrical system of the native language, such that adults speaking French, English and Japanese use different metrical segmentation procedures. It further appeared that each procedure is based on the metrical unit characteristic of a particular language. Hence, the syllable appeared to be the unit of segmentation used by French-speaking (Mehler et al., 1981), Spanish- and Catalan-speaking (Sebastián-Gallés et al., 1992), and Portuguese-speaking (Morais et al., 1989) adults. However, the segmentation procedures used in English (Cutler et al., 1986; Cutler and Norris, 1988) and Dutch (Vroomen et al., 1996) are apparently guided by information about typical word stress patterns, which involve an alternation of strong and weak syllables. A third pattern was found for Japanese adults, who appeared to rely on the mora (Otake et al., 1993). [2] Finally, adults' segmentation procedures appear to be deeply embedded in their language-specific competence, and acquired at a very young age. This is attested by the fact that the metrical procedure they use is determined by their native language rather than by the language they are listening to: once they have mastered a particular language, adults rely on procedures appropriate to that language even when listening to a foreign language (Cutler et al., 1986; Otake et al., 1993). Moreover, it has also been shown that even very proficient bilinguals are dominant in one of their languages, and have developed specialized metrical segmentation procedures in only one of their languages (Cutler et al., 1992).

It has been suggested that each of the different types of metrical segmentation procedures is optimally adapted to the processing of a particular rhythmic class of languages (Cutler and Mehler, 1993; Otake et al., 1993; see also Sebastián-Gallés et al., 1992; Vroomen et al., 1996), even if minor processing differences can be found within a class. [3] This proposal is based on a three-way classification of languages according to their predominant rhythmic structure (Abercrombie, 1967; Pike, 1945). By this classification, most Germanic languages (e.g., English, Dutch, German) have a rhythm based on the stress unit (i.e., the foot), most Romance languages (e.g., French, Italian, Spanish) have a rhythm based on the syllable, while languages such as Japanese have a mora-based rhythm. Note that these rhythmic units are hierarchically related at the phonological level, as feet are made up of syllables that are made up of morae.

The view that the rhythmic properties of a language shape adults' processing procedures has influenced views of how infants develop efficient procedures for segmenting native language utterances. Mehler et al. (1996) have proposed that the emergence of metrical segmentation procedures

[2] The mora is a metrical unit that can either be syllabic (V or CV) or subsyllabic (second half of a long vowel, post-vocalic nasal consonant, or post-vocal portion of a geminate consonant).

[3] We draw here a distinction between rhythm at the lexical level (which we call metrical information, or stress pattern) and rhythm at the sentence level (for which we reserve the term rhythm, and rhythmic classes). However, as we point out below, these two levels are not entirely dissociated.

rests on an early sensitivity to prosody, and more precisely here, linguistic rhythm specified at a non-segmented level. Note that this proposal meets other recent proposals regarding infants' early sensitivity to prosodic information, and the importance of prosody in early language acquisition (Fernald and Kuhl, 1987; Jusczyk, 1997; Jusczyk et al., 1993a,b; Jusczyk and Thompson, 1978; Karzon and Nicholas, 1989; Kuhl and Miller, 1982; Morse, 1972; Nazzi et al., 1998a,b; Spring and Dale, 1977).

More specifically, our proposal is that infants' sensitivity to rhythm at the utterance/supra-segmental level will allow them to specify the type of rhythm of their native language, and develop the procedure appropriate to its segmentation. Hence, the emergence of the metrical segmentation procedures is not based on infants learning that words in their native language have a specific stress pattern (which requires knowledge of at least some words), but on infants specifying the type of rhythm of their native language. This acquisition scenario works because there is a relation between rhythm at the sentence level and meter at the lexical level, so that the metrical units segmented by the infants will more or less correspond to words. We further hypothesize that infants will discover the correspondence between metrical units and lexical units at some point between the ages of 7.5 months (onset of speech segmentation) and 10 to 12 months (acquisition of the first words).

To validate the above (prosodic bootstrapping) proposal, three points need to be validated. First, we need to demonstrate that there are acoustic correlates to the rhythmic classes. Such evidence has been reported in several recent investigations (Arvaniti, 1994; Den Os, 1988; Fant et al., 1991; Nazzi, 1997; Shafer et al., 1999; Ramus et al., 1999). Second, we need to determine that infants are sensitive to rhythmic class information, and, third, that this sensitivity plays a role in the acquisition of the segmentation procedures. The rest of this paper is a review of the research bearing on infants' sensitivity to and acquisition of rhythmic class information, which has studied infants' ability to discriminate between languages to explore how they perceive linguistic rhythm and acquire the rhythmic properties of their native language.

2. Language discrimination at birth

Studies exploring newborns' ability to discriminate between languages first contrasted their to-be-native language to a foreign language. Some of the experiments in Mehler et al. (1988) explored the ability of French newborns to discriminate utterances drawn from languages of different rhythmic classes. Those infants discriminated utterances in French from others in Russian, but did not seem to discriminate English from Italian. This led to the conclusion that early language discrimination was based on the recognition of the native language (see also Moon et al., 1993), rather than on rhythmic class discrimination. However, the latter conclusion was undermined by a subsequent reanalysis of these data by Mehler and Christophe (1995), in which the data for the different subgroups were merged. This new analysis showed that the newborns did discriminate English and Italian, thus raising the prospect that newborns might be able to discriminate between two foreign languages. Moreover, because the languages presented in these experiments always belonged to different rhythmic classes (a stress-based language contrasted with a syllable-based one), these data are consistent with the hypothesis that infants are relying on rhythmic differences when discriminating languages.

In the following, we present the results of two complementary studies that investigated more precisely the basis of the language discriminations observed at birth. The first study used a crosslinguistic approach in which different sets of languages were presented, allowing for the systematic variation of their rhythmic distance (in terms of rhythmic class membership). The second study used an acoustic approach, which consisted in varying systematically the acoustic and linguistic cues present in the stimuli.

2.1. The role of rhythm—a crosslinguistic approach

To clarify the issue of whether young infants discriminate because they recognize their native language or because they extract the rhythmic characteristics of utterances and use these characteristics to sort utterances (languages) into a limited number of rhythmic classes, Nazzi et al. (1998a) conducted a new study with newborns from French-speaking families with sets of sentences in different foreign languages. Four different speakers of each language recorded 10 different sentences each (hence a total of 40 different sentences per language). Moreover, because this study was about the role of rhythmic/prosodic information in discrimination, the stimuli were low-pass filtered with a cutoff frequency of 400 Hz, a transformation of the signal that removes most segmental/lexical information while preserving most prosodic information.

In this study, the rhythmic distance between the languages presented was systematically varied, in that the languages belonged either to two different rhythmic classes (stress-based British English and mora-based Japanese) or to the same class (stress-based British English and Dutch). If infants' discriminations are based on a recognition of the native language, they should fail to discriminate in both experiments as all the languages presented are foreign languages. If infants discriminate languages based on their rhythmic types, then they should discriminate English and Japanese, but fail to discriminate English and Dutch. Nazzi et al. (1998a) also conducted a language 'categorization' experiment, in which infants were familiarized with utterances from two languages, and then tested on whether they could discriminate them from utterances from two other languages. The rationale of this experiment was to show that one could determine whether or not infants would perceive a language change by using the same four languages but varying the way they were paired (i.e., according or orthogonally to the rhythmic classification).

Nazzi et al. (1998a) results supported the hypothesis that language discrimination is based on rhythm and rhythmic classes. French newborns discriminated the two foreign languages from different rhythmic classes (English and Japanese), but did not discriminate the two languages from the stress-based class (English and Dutch). These results are consistent with Mehler and Christophe's (1995) reanalysis of the Mehler et al. (1988) experiment on English/Italian discrimination by French newborns. In the third experiment, discrimination was found when the languages were paired according to the rhythmic classes (i.e. stress-based British English and Dutch vs. syllable-based Italian and Spanish), but not when the pairing was made across the rhythmic classes (i.e. British English and Italian vs. Dutch and Spanish). Hence, newborns discriminated the languages only when there was a rhythmic basis for doing so. Moreover, it appeared that newborns were not sensitive to the prosodic differences existing between languages from the same rhythmic class or at least that the intra-class differences were less important than the intra-class similarities.

2.2. The role of rhythm—an acoustic approach

The more parsimonious interpretation of the above studies is that infants discriminated between different types of rhythm. However, there are many acoustic differences between languages, and one needs to be sure that infants did not attend to some other non-rhythmic speech cue which would define a partition of the languages used in the previous studies equivalent to the rhythmic classes. The way the above studies minimized this risk was by presenting infants with stimuli low-pass filtered at a cutoff frequency of 400 Hz. Hence, Mehler et al. (1988) found that newborns could discriminate between French and Russian utterances both when the stimuli were natural, and when they were low-pass filtered (but not when they were presented backward, a transformation that destroys the temporal organization of speech). Moreover, all the experiments in Nazzi et al. (1998a) used low-pass filtered stimuli. It is therefore unlikely, for instance, that phonetic or phonotactic differences might have provided the basis for the observed discriminations. Yet, because the lowest 400 Hz of speech contain at least one cue other than rhythm, that is intonation (i.e. the melodic contour carried by the fundamental frequency or $F0$), it would be useful to confirm the above findings with studies using stimuli in which the acoustic cues are even more stringently controlled for.

This led Ramus and Mehler (1999) to look for alternative ways to control speech cues. Their strategy was to resynthesize speech from real utterances, which allowed them to control the parameters of the synthesized speech stimuli. They used a diphone synthesizer, MBROLA (Dutoit et al., 1996), whose input consists of a string of phonemes together with their respective durations and corresponding $F0$ contours. A number of manipulations of those parameters were devised. [4] The *saltanaj* manipulation consisted in replacing each phoneme by just one exemplar of the same manner of articulation, without modifying its duration: /s/ for fricatives, /a/ for vowels, /l/ for liquids, /t/ for plosives, /n/ for nasals and /j/ for glides. When applied to two languages using the same diphone database, this effectively removes all phonetic differences and first-order phonotactic differences, while still preserving the overall prosody. A more radical manipulation, *sasasa*, replaced all consonants by /s/ and all vowels by /a/, thereby only preserving the alternation of consonants and vowels, rhythm and intonation. In parallel, it was also possible to degrade the $F0$ contour, by making it constant or replacing it by another contour.

Using this methodology, Ramus et al. (2000) showed that French newborns were able to discriminate between stress-based Dutch and mora-based Japanese when the utterances had been resynthesized using the saltanaj manipulation. Given that both rhythm and intonation were available in these stimuli, this result paralleled the discrimination of low-pass filtered English and Japanese (Nazzi et al., 1998a). In another experiment, the original $F0$ contours of the sentences were replaced with the same artificial contours for both languages, thereby eliminating any differences in intonation. The results again showed evidence of a discrimination by the newborns, although the effect was much weaker (Ramus, in press). This suggests that, for the two languages considered, intonational differences cannot by themselves explain newborns ability to discriminate between them, hence supporting a discrimination based on rhythm. However, this does not imply that intonation plays no role at all. Indeed, rhythm and intonation are not entirely independent as, for example, phrase-final lengthening is also often marked by a decrease in pitch. It is therefore conceivable that this co-variation between rhythm and intonation enhances discrimination abilities.

2.3. Summary

The above studies provide converging evidence that newborns discriminate languages based on their rhythmic properties at the utterance level. Indeed, it was found that newborns can discriminate foreign languages if and only if they belong to different rhythmic classes, and even when the confounding factor of intonation is eliminated. This supports the proposal that infants are, very early on, sensitive to rhythm, and that their sensitivity allows them to separate languages into classes of languages sharing similar global rhythmic properties. These studies then suggest that sensitivity to rhythm, rather than recognition of the native language, supports language discrimination at birth. Hence, this suggests that the preference for the native language found in some studies (Mehler et al., 1988; Moon et al., 1993) would be better qualified as a preference for the native language rhythmic class.

Finally, the above studies were motivated by our proposal that infants' acquisition of the metrical segmentation procedure appropriate to the processing of (the rhythmic class of) their native language might rely on an early sensitivity to rhythm. Our findings that young infants are sensitive to the same rhythmic properties that define the different language classes that emerged from adult research provides an element of validation/feasibility to that proposal. In the following, we present further research on older infants' discrimination of languages, in which we sought to determine whether there is a link between the evolution of language discrimination and the acquisition of the metrical unit and segmentation procedure of the native language.

3. Acquisition of the rhythmic properties of the native language

Following the pattern of language discrimination found in their newborn study, Nazzi et al.

[4] Samples of the stimuli can be heard on http://www.ehess.fr/lscp/persons/ramus/resynth/ecoute.htm.

(1998a) proposed a more precise prosodic bootstrapping account of the acquisition of the metrical procedures found to be used by adults. According to their proposal, called the rhythmic class acquisition hypothesis, infants' initial sensitivity to rhythmic classes would allow them to specify the common rhythmic properties of their native rhythmic class, from which they would develop its associated metrical segmentation procedure. It would also enable them to perform more detailed analyses of the rhythmic properties of languages within the native class, leading to improvements in discriminating languages within that class. Accordingly, by a few months of age, infants should discriminate any two languages from the native rhythmic class, but should not discriminate languages belonging to the same foreign rhythmic class.

Nazzi et al. (2000) later suggested two other possible acquisition patterns. One of these acquisition patterns, called the native language acquisition hypothesis, is that newborns' sensitivity to rhythmic classes might lead to the acquisition of the rhythmic properties of their native language (whether or not they are shared with other languages of its class) rather than the common properties of the native language class as a whole. This pattern of acquisition is more similar to the one observed for the acquisition of the phonetic and phonotactic properties of the native language (Jusczyk et al., 1993b; Werker and Tees, 1984). This hypothesis predicts that the ability to discriminate languages evolves during the first months with the emergence of the ability to discriminate the native language from foreign languages belonging to the native rhythmic class. However, infants should not discriminate two foreign languages belonging to the same rhythmic class, be they native or foreign.

Finally, Nazzi et al. (2000) suggested that infants may undergo a general maturation process between birth and 4 months that allows them to make more subtle language discriminations. This maturation hypothesis holds that, at some point in development, infants should discriminate any two languages, provided that they differ sufficiently at the acoustic level. Furthermore, infants should discriminate languages both within the native language class and within foreign classes.

Some studies have tested infants' language discriminations a few months after birth. A well-established result is that 2- to 4-month-old infants can discriminate their native language from a foreign language (Bahrick and Pickens, 1988; Dehaene-Lambertz and Houston, 1998; Mehler et al., 1988). But the data are more difficult to interpret when it comes to language pairs involving only foreign languages, or languages from the same rhythmic class. Indeed, using a habituation/dishabituation paradigm, Christophe and Morton (1998) tested the ability of 2- to 3-month-olds English infants to discriminate between three pairs of languages. Two of the pairs involved foreign languages belonging to different rhythmic classes: syllable-based French vs. mora-based Japanese, and stress-based Dutch vs. Japanese. The last pair involved the native language, stress-based English, and a foreign language from the same rhythmic class, Dutch. The rationale for testing Dutch in those two last experiments was to determine whether 2- to 3-month-olds considered it as their native language or not. The results for French and Japanese showed no discrimination, a result not predicted by any of the three acquisition hypotheses presented above (but see below for a methodological account of that absence of discrimination). The results for the other two experiments showed close but non-significant discrimination, which suggests that these infants might be at a transitional age, some of them starting to distinguish their native language (e.g., English) from other languages of the same rhythmic class (e.g., Dutch). But even if this interpretation is correct, it still does not allow us to select between the acquisition hypotheses we presented, given that crucial language pairs, involving foreign languages from either the native rhythmic class or a foreign class, were not tested.

Nazzi et al. (2000) set out to explore this issue beyond 2 to 3 months of age, by studying the discrimination abilities of English-learning American 5-month-olds. Because testing 5-month-olds required adapting the headturn preference procedure to provide a discrimination measure, infants were first tested on language contrasts involving languages from different rhythmic classes (for which discrimination is expected according to all three considered developmental hypotheses). In that paradigm, infants were first familiarized with passages from one language, then tested with four new passages of the familiarization language and four passages of a new language, presented in random order.

The between-class language pairs first tested were: stress-based British English (hence a variant of the infants' native language) vs. mora-based Japanese, and two languages from different non-native language rhythmic classes, syllable-based Italian vs. mora-based Japanese (a rhythmic contrast not tested previously). Infants discriminated both pairs, establishing that 5-month-olds can discriminate languages from different rhythmic classes, even when both languages are foreign, and extending the evidence of between-class discrimination to the syllable- vs. mora-based contrast. These results are consistent with the newborn data (Nazzi et al., 1998a), but not with the absence of discrimination between French and Japanese by English 2- to 3-month-olds (Christophe and Morton, 1998). However, a way to reconcile the 2- to 3-month-old data is given by an experiment reported in the discussion of Experiment 2 of Nazzi et al. (2000), which used a habituation/dishabituation paradigm similar to the one used in the studies with the younger infants, and failed to show the American 5-month-olds' ability to discriminate British English and Japanese that emerges with the other test paradigm. Hence, Nazzi et al. (2000) suggested that the habituation/dishabituation paradigm might be well adapted to the testing of newborns but not that of 5-month-olds, the 2-month-olds falling in between these two extremes.

The four other experiments in Nazzi et al. (2000) tested pairs of languages belonging to the same class. Pairs were varied according to whether or not they belonged to the native rhythmic class, and whether or not the native language was presented. The results suggest a critical role of both rhythmic class and native language. Indeed, when the languages belonged to a rhythmic class different from that of the native language (i.e., syllable-based Italian and Spanish), 5-month-olds failed to discriminate. But when the languages belonged to the native language class, some discrimination ability was found. However, the results appeared to be predicted by whether or not the native language was presented. Indeed, the infants discriminated British English from Dutch, and their native dialect (American English) from another dialect (British English), but failed to discriminate German from Dutch.

These results underline the evolution of language discrimination abilities between birth and 5 months, and allow a specification of the acquisition mechanisms involved. Indeed, this evolution cannot be the product of a general maturation of processing abilities, given the fact that 5-month-olds discriminated the subtle distinction between two variants of their native language (American and British English), while failing to discriminate more distinctive language pairs (Italian/Spanish; Dutch/German). Moreover, the failure to discriminate between Dutch and German is incompatible with the rhythmic class acquisition hypothesis according to which infants should discriminate between all languages from the native rhythmic class, given that development consists of the acquisition of the rhythmic properties shared by the different languages of that class. The best account of the present results is the native language acquisition hypothesis which states that infants learn the specific rhythmic features of their native language, rather than those of the rhythmic class as a whole. Accordingly, languages are discriminated when one is more similar to the native patterns than the other (for American infants: American vs. British English; British English vs. Dutch), but not when they are equally dissimilar (for American infants: Dutch vs. German; Italian vs. Spanish).

Finally, it emerges that language discrimination at 5 months is based on prosodic information, even when the languages presented share global rhythmic properties. This is shown here by the fact that infants discriminated the same sentences of English spoken in two different variants (British and American), which supports the hypothesis that their discriminations were based on prosodic rather than phonotactic (and to a lesser degree phonetic) information. This is further supported by the fact that within-class discrimination can be obtained with low-pass filtered sentences (i.e., stimuli with degraded phonetic/phonotactic information, but preserved prosody) in 4-month-olds (Bosch and Sebastián-Gallés, 1997), or that American infants cannot discriminate American words from Dutch words (i.e., stimuli differing phonetically and phonotactically, but lacking rich prosodic information) before 9 months of age (Jusczyk et al., 1993b).

3.1. Summary

So far, the developmental pattern across the above studies illustrates an early evolution of infants' processing abilities, resulting in selectively finer language discrimination performance between birth and 5 months of age. A complete description of how language discrimination abilities develop involves a combination of two components. An innate sensitivity to rhythmic types, independent of any knowledge of the rhythmic properties of the native language, explains abilities to discriminate languages from different rhythmic classes from birth onwards. However, infants' growing knowledge of their native language rhythmic properties seems to be the key to their ability to discriminate languages from within the native language rhythmic class by 5 months of age. In the following, we evaluate the consequences of these results on our proposal regarding the implication of rhythmic sensitivity in the acquisition of the metric segmentation procedure appropriate to the processing of the native language.

4. Conclusion and future directions

In this paper, we have presented a developmental proposal for the acquisition of the metrical speech segmentation procedures used in adulthood. Given the bootstrapping issue that these procedures, which operate at the lexical level, need to be learnt at a level independent of the word, we have proposed that these procedures could develop from an early sensitivity to and acquisition of the rhythmic properties of speech at the utterance/suprasegmental level. Based on data from the adult and phonetic literature, we have proposed that infants' sensitivity to differences in global rhythms (rhythms that seem to separate the languages into rhythmic classes) could be the basis for the later development of the metrical segmentation procedure appropriate to their native language (Mehler et al., 1996). More specifically, the proposal is that, based on early sensitivity to rhythmic classes, infants will specify the common rhythmic properties of their native rhythmic class (and its metrical unit), from which they will develop its associated metrical segmentation procedure (Nazzi et al., 1998a, 2000).

We have then presented studies testing this proposal. These studies explored infants' sensitivity to rhythmic classes, from birth to 5 months, through the study of their ability to discriminate utterances from different languages. Overall, there is converging evidence that newborns are sensitive to rhythm and to the way languages fall into rhythmic classes (Nazzi et al., 1998a,b; Ramus et al., 2000), a sensitivity which further appears to be independent from any knowledge of the native language. Studies with older infants (up to 5 months of age) suggested that infants rapidly learn the prosodic properties of their native language, rather than those of their native language class as a whole. This is attested by the fact that, contrary to younger infants, 5-month-olds were found to discriminate between their native language and a foreign language of the same rhythmic class, but still could not discriminate two foreign languages from their native class (Nazzi et al., 2000). At first sight, these data go against our acquisition proposal. However, it could be that the infants tested by Nazzi et al. (2000) were too young, and had not yet learnt the global properties of their native language class. In order to test this hypothesis, older infants are starting to be tested (Nazzi, Jusczyk, and Johnson, work in progress). This new study tests infants around the age of 7.5 months, which is the age at which English-learning American infants were found to show the first signs of having learnt the predominant lexical stress pattern/metric unit of their native language (Jusczyk et al., 1993a), and of having developed its appropriate metrical segmentation procedure (Jusczyk et al., 1999b). Data showing that, by that age, infants have started discriminating foreign languages from their native language class (while still failing at discriminating foreign languages from a non-native class) would support our acquisition proposal.

Acknowledgements

We would like to thank the co-authors of the different papers cited here, for their crucial contributions to our theoretical and experimental work. We would also like to thank the babies that participated in the studies and their families, and, of course, Bert Schouten for organizing and inviting us to the Nature of Speech Perception workshop where both authors presented their work. FR was supported by the Délégation Générale pour l'Armement and a Marie Curie fellowship from the European Commission.

References

Abercrombie, D., 1967. Elements of General Phonetics. University of Edinburgh Press, Edinburgh.

Arvaniti, A., 1994. Acoustic features of Greek rhythmic structure. Journal of Phonetics 22, 239–268.

Bahrick, L.E., Pickens, J.N., 1988. Classification of bimodal English and Spanish language passages by infants. Infant Behavior and Development 11, 277–296.

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

Christophe, A., Morton, J., 1998. Is Dutch native English? Linguistic analysis by 2-month-olds. Developmental Science 1, 215–219.

Cutler, A., Mehler, J., 1993. The periodicity bias. Journal of Phonetics 21, 103–108.

Cutler, A., Norris, D.G., 1988. The role of strong syllables in segmentation for lexical access. Journal of Experimental Psychology: Human Perception and Performance 14, 113–121.

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

Cutler, A., Mehler, J., Norris, D., Segui, J., 1992. The monolingual nature of speech segmentation by bilinguals. Cognitive Psychology 24, 381–410.

Dehaene-Lambertz, G., Houston, D., 1998. Faster orientation latencies toward native language in two-month-old infants. Language and Speech 41, 21–43.

Den Os, E., 1988. Rhythm and Tempo of Dutch and Italian: A Contrastive Study. Drukkerij Elinkwijk BV, Utrecht, Holland.

Dutoit, T., Pagel, V., Pierret, N., Bataille, F., van der Vrecken, O., 1996. The MBROLA project: towards a set of high-quality speech synthesizers free of use for non-commercial purposes. In: ICSLP'96, Philadelphia.

Echols, C.H., Crowhurst, M.J., Childers, J.B., 1997. 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.

Friederici, A.D., Wessels, J.M., 1993. Phonotactic knowledge and its use in infant speech perception. Perception and Psychophysics 54, 287–295.

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

Jusczyk, P.W., 1997. The Discovery of Spoken Language. MIT Press, Cambridge, MA.

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., Thompson, E.J., 1978. Perception of a phonetic contrast in multisyllabic utterances by two-month-old infants. Perception and Psychophysics 23, 105–109.

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., Luce, P.A., Charles-Luce, J., 1994. Infants' sensitivity to phonotactic patterns in the native language. Journal of Memory and Language 33, 630–645.

Jusczyk, P.W., Hohne, E.A., Baumann, A., 1999a. Infants' sensitivity to allophonic cues for word segmentation. Perception and Psychophysics 61, 1465–1476.

Jusczyk, P.W., Houston, D., Newsome, M., 1999b. The beginnings of word segmentation in English-learning infants. Cognitive Psychology, 39.

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

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.T.L., Coolen, R., Houston, D., Cutler, A., 1998. Using the headturning technique to explore cross-linguistic performance differences. Advances in Infancy Research 12, 205–220.

Mattys, S.L., Jusczyk, P.W., 2001. Phonotactic cues for segmentation of fluent speech by infants. Cognition 78, 91–121.

Mattys, S.L., Jusczyk, P.W., Luce, P.A., Morgan, J.L., 2001. Phonotactic and prosodic effects on word segmentation in infants. Cognitive Psychology 38, 465–494.
Mehler, J., Christophe, A., 1995. Maturation and learning of language in the first year of life. In: Gazzaniga, M.S. (Ed.), The Cognitive Neurosciences. Bradford Books, Cambridge, MA, pp. 943–954.
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., 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.
Mehler, J., Dupoux, E., Nazzi, T., Dehaene-Lambertz, G., 1996. Coping with linguistic diversity: the infant's viewpoint. In: Morgan, J.L., Demuth, K. (Eds.), Signal to Syntax. Erlbaum, Mahwah, NJ, pp. 101–116.
Moon, C., Panneton-Cooper, R., Fifer, W.P., 1993. Two-day-olds prefer their native language. Infant Behavior and Development 16, 495–500.
Morais, J., Content, A., Cary, L., Mehler, J., Segui, J., 1989. Syllabic segmentation and literacy. Language and Cognitive Processes 4, 57–67.
Morgan, J.L., Saffran, J.R., 1995. Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. Child Development 66, 911–936.
Morse, P.A., 1972. The discrimination of speech and nonspeech stimuli in early infancy. Journal of Experimental Child Psychology 13, 477–492.
Nazzi, T., 1997. Du rythme dans l'acquisition et le traitement de la parole. Ph.D., Ecole des Hautes Etudes et Sciences Sociales, Paris, unpublished.
Nazzi, T., Bertoncini, J., Mehler, J., 1998a. 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., Floccia, C., Bertoncini, J., 1998b. Discrimination of pitch contours by neonates. Infant Behavior and Development 21, 779–784.
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.
Otake, T., Hatano, G., Cutler, A., Mehler, J., 1993. Mora or syllable? Speech segmentation in Japanese. Journal of Memory and Language 32, 258–278.
Pike, K., 1945. The Intonation of American English. University of Michigan Press, Ann Arbor, MI.
Ramus, F., (in press). Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues. Annual Review of Language Acquisition, 2.
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.
Saffran, J.R., Aslin, R.N., Newport, E.M., 1996. Statistical learning by 8-month-old infants. Science 274, 1926–1928.
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.
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.
Spring, D.R., Dale, P.S., 1977. Discrimination of linguistic stress in early infancy. Journal of Speech and Hearing Research 20, 224–232.
Turk, A.E., Jusczyk, P.W., Gerken, L.A., 1995. Do English-learning infants use syllable weight to determine stress? Language and Speech 38, 143–158.
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.
Werker, J.F., Tees, R.C., 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. Infant Behavior and Development 7, 49–63.