# Fossil markers of language development: phonological 'deafnesses' in adult speech processing[*]

*Emmanuel Dupoux*[†] *and Sharon Peperkamp*[†‡]

## Abstract

The sound pattern of the language(s) we have heard as infants affects the way in which we perceive linguistic sounds as adults. Typically, some foreign sounds are very difficult to perceive accurately, even after extensive training. For instance, native speakers of French have troubles distinguishing foreign words that differ only in the position of main stress, French being a language in which stress is not contrastive.

In this paper, we propose to explore the perception of foreign sounds cross-linguistically in order to understand the processes that govern early language acquisition. Specifically, we propose to test the hypothesis that early language acquisition begins by using only regularities that infants can observe in the surface speech stream (Bottom-Up Bootstrapping), and compare it with the hypothesis that they use all possible sources of information, including, for instance, word boundaries (Interactive Bootstrapping).

We set up a research paradigm using the stress system, since it allows to test the various options at hand within a single test procedure. We distinguish four types of regular stress systems the acquisition of which requires different sources of information. We show that the two hypotheses make contrastive predictions as to the pattern of stress perception of adults in these four types of languages. We conclude that cross-linguistic research of adults speech perception, when coupled with detailed linguistic analysis, can be brought to bear on important issues of language acquisition.

Index: phonological 'deafness', language acquisition, phonological bootstrapping, stress perception, cross-linguistic studies

## 1. Introduction

Recent research has demonstrated that language acquisition begins at a very early age and proceeds at an amazingly fast pace. During the first year of life, infants find out important parts of the phonological structure of their maternal language and begin to recognize and store frequent words. One crucial question for models of language acquisition concerns the mechanisms underlying these developments. In particular, the question arises as to what type of learning algorithm infants use to extract phonological and lexical information. The more specific question we are interested in is whether learning the words of one's language is independent from learning the phonology, or whether the two processes are interleaved and interdependent.

In this paper, we claim that cross-linguistic studies of the end-state in adults can shed light on these developmental questions. We rest our claim on the finding that early exposure to a language has a lasting impact on speech processing routines in adults. That is, listeners use a processing apparatus specifically tuned to their maternal language. Consequently, they have a lot of difficulty in dealing with sound structures that are alien to the language they heard as infants. They

[*] October, 1999. Reprinted by the authors from B. Laks and J. Durand (Eds) (to appear). *Cognitive Phonology*. Oxford University Press. The names of the authors are in alphabetic order. Research for this paper was supported by a Fyssen grant to the second author. We would like to thank Leticia Maria Sicuro Correa, Anne Christophe, François Dell, Jacques Mehler, and Franck Ramus for comments and discussion.

[†] Laboratoire de Sciences Cognitives et Psycholinguistique, Ecole des Hautes Etudes en Sciences Sociales, 54 Bd Raspail, 75006 Paris, France. Email: dupoux@lscp.ehess.fr. Web: http://www.ehess.fr/centres/lscp/persons/dupoux

[‡] Département des Sciences du Langage, Université Paris 8, 2 Rue de la Liberte, 93535 Saint Denis. Email: sharon@lscp.ehess.fr

display what we call phonological ‘deafnesses’; that is, they have troubles discriminating phonological contrasts that are not used in their native language. Moreover, the phonological ‘deafnesses’ are robust, in that - analogously to patterns of foreign accent in production - they are resistant to learning a second language, and even to specific training. We hypothesize that phonological ‘deafnesses’ originate in the acquisition during the first few years of life. We therefore propose to look at perception in adults cross-linguistically, in order to gain insight into the acquisition processes that they went through during these first years. Crucially, we propose to test the predicted ‘deafnesses’ in various languages according to different theoretical options regarding early language acquisition.

The outline of this paper is as follows. In section 2, we expose some background data regarding, on the one hand, phonological processing and acquisition, and, on the other hand, lexical processing and acquisition. In section 3, we present two hypotheses concerning the relationship between phonological and lexical acquisition, i.e. Bottom-up Bootstrapping and Interactive Bootstrapping. In section 4, we propose to test these hypotheses by means of a cross-linguistic study on the perception of stress by speakers of languages that differ crucially in their stress rules. That is, we consider four types of regular stress systems the acquisition of which requires different sources of information, and we discuss the predictions following from the two hypotheses regarding the pattern of stress perception of adults. Finally, a summary and some concluding remarks are presented in section 5.

# 2. Experimental data

## 2.1 Phonological processing and acquisition

During the first year of life, infants acquire many phonological properties of their native language and lose their sensitivity to phonological contrasts that are not pertinent. In this section, we review data concerning language-specific phonological processing in adults and its relation to phonological acquisition. In particular, we deal with the segmental inventory, phonotactics, and suprasegmentals, respectively.

### 2.1.1 Segments

It has long been known that speech perception is influenced by phonological properties of the maternal language (Sapir 1921; Polivanov 1974). Much experimental evidence has been gathered concerning this influence. For instance, Goto (1971) has documented that Japanese listeners map American [l] and [r] onto their own, single, [R] category, and, as a result, have a lot of difficulties in discriminating between them. Similarly, the contrast between the retroflex and dental stops [ʈ] – [t̪] is very difficult for the English, but not for the Hindi speaker (Werker & Tees 1984b). This contrast is, in fact, phonemic in Hindi, whereas neither of the stop consonants involved occurs in English; rather, English uses the alveolar stop [t]. Note, however, that not all foreign contrasts are difficult (Polka 1991; Best, McRoberts, & Sithole 1988). For instance, Best *et al*. found that English subjects do not have difficulties in discriminating different Zulu clicks from one another.

The Perceptual Assimilation Model (Best 1994) aims at explaining these differences. In particular, it states that a foreign sound is assimilated to an existing segment in the native language if the acoustic characteristics of both sounds are close enough. In this case, listeners are only able to judge whether the foreign sound is a good or a bad exemplar of the segment to which it has assimilated, but they do not have access to its detailed phonetic characteristics. Consequently, two distinct foreign sounds that are assimilated to the same segment in the native language and that are equally bad exemplars of this segment will be very difficult to discriminate. Such is the case for Japanese speakers with the American [l] - [r] contrast, both [l] and [r] being assimilated to Japanese [R], as well as for the English speakers with the Hindi [ʈ] – [t̪] contrast, both [ʈ] and [t̪] being assimilated to English [t]. By contrast, foreign sounds that are too distant from all of the native segments will not be assimilated at all and listeners will be able to make fine phonetic discriminations between them. This is illustrated by the successful discrimination of different Zulu clicks by native speakers of English.

The effects of the maternal inventory on adult perception are very robust, and resist to the acquisition of a second language. Thus, even early and highly fluent bilinguals have difficulties with non-native vowel contrasts, as shown in Pallier, Bosch & Sebastián-Gallés (1997). In this study, subjects are Spanish-Catalan bilinguals who have begun to acquire their second language between four and six years of age, and have used it extensively thereafter. The two languages differ as to the number of vowel phonemes; Spanish has a five-vowel system, while Catalan uses two more vowels. Crucially, subjects whose native language is Spanish are shown to use exclusively the more restricted Spanish vowel system for the purposes of speech perception and spoken word recognition.

Finally, electrophysiological studies suggest that effects of the phoneme inventory occur very early during on-line speech processing. Thus, Näätänen *et al*. (1997) presented repeated sequences of the same vowel, followed occasionally by a change in vowel. They found that the Mismatch Negativity (an electrophysiological component that correlates with detection of a change in stimulation) is modulated as a function of whether the vowel is part or not of the phoneme inventory of the language. That is, 100 to 240 ms. after the vowel onset, the perceptual system is already influenced by the vowel inventory of the native language.

All in all, these data suggest that listeners use a set of language-specific phoneme categories during speech perception. This raises the question as to how and when infants acquire their native segmental inventory. Concerning the age at which infants tune to the inventory of their language, Kuhl *et al*. (1992) have shown that American and Swedish 6-month-old infants react specifically to vowels that are prototypic in their maternal language. Furthermore, Polka & Werker (1994) found that at the same age, infants begin to lose their sensitivity for non-native vowel contrasts. That is, 6-month-old English-acquiring infants fail to discriminate between the German lax vowels /ʏ/ and /ʊ/ as well as between their tense counterparts /y/ and /u/; the first, front, vowel in each pair is not part of the English inventory. Between 10 and 12 months, infants similarly lose the ability to discriminate non-native consonantal contrasts (Werker & Tees 1984a). Regarding the way in which infants acquire their segmental inventory, Kuhl *et al*. (1997) found that mothers addressing their infants produce acoustically more extreme vowels than they do when addressing adults, resulting in a 'stretching' of the vowel space. This shows that language input to infants provides well-specified information about the phonemes, and suggests that a purely statistical extraction algorithm could establish prototypes for the sounds of the native language.

### 2.1.2 Phonotactics

Language phonologies differ in properties other than the inventory of phonemes. Notably, they differ in the way in which phonemes can co-occur in words. These *phonotactic* properties appear to influence speech processing routines, in that non-words with a high phonotactic probability are processed faster and more accurately than non-words with a low phonotactic probability (Brown & Hildum 1956; Vitevitch *et al*. 1997; Gathercole *et al*. 1999). Furthermore, phonotactics have been shown to bias the perception of individual segments. For instance, Massaro & Cohen (1983) found that synthetic stimuli that are ambiguous between /r/ and /l/ tend to be perceived as /r/ when preceded by /t/ and as /l/ when preceded by /s/. The interpretation given by Massaro and Cohen is that perception is biased towards segments that yield the legal clusters /tr/ and /sl/, rather than the illegal clusters /tl/ and /sr/ (see also Pitt 1998). Similarily, Hallé *et al*. (1998) found that illegal onset clusters in French are perceived as legal ones. In particular, illegal /dl/ and /tl/ are perceived as legal /gl/ and /kl/, respectively.

It has even been reported that in some contexts, illegal sequences of phonemes yield perception of illusory segments. For instance, Japanese syllables cannot have complex onsets (except for consonant-glide onsets) and cannot have codas (except for nasal consonants and the first half of geminates). Dupoux *et al*. (1999) found that Japanese subjects report the presence of an epenthetic vowel [u] between consonants in non-words like [ebzo]. They also found that Japanese subjects have problems discriminating between, for instance, [ebzo] and [ebuzo]. This was found even in subjects who were quite proficient in French, a language which authorizes both coda consonants and complex onsets. Moreover, in an electrophysiological study, Dehaene-Lambertz, Dupoux & Gout (in press) found that the effect for phonotactics arises as early as that of the phoneme inventory investigated by Näätänen *et al*. (1997) and reported above. These results, then, suggest that phonotactics play a role so important as to create the illusory perception of segments.

As to infants' sensitivity to phonotactic properties, there is evidence that it equally arises during the first year of life. For instance, Friederici & Wessels (1993) showed that 9-month-old Dutch infants prefer to listen to phonotactically legal words rather than to illegal ones. Similarly, Jusczyk, Luce & Charles-Luce (1994) found that 9-month-old American infants, when listening to monosyllabic non-words, prefer those with a high-probability phonotactic pattern rather than those with a low-probability phonotactic pattern. Jusczyk *et al*. (1993) reported, furthermore, that 9-month-old American infants listen longer to unfamiliar English words than to Dutch words. The latter contain segments and sequences that are illegal in English, suggesting again that infants of this age are sensitive to the phonotactics of their language. This is corroborated by the finding that no differences are found when the stimuli are low-pass filtered, hence did not contain any segmental information. Note, however, that we do not yet know whether at the same age, the presentation of illegal clusters yields the perception of illusory segments as documented for adults.

### 2.1.3 Suprasegmentals

Finally, languages differ in suprasegmentals and in particular, in the use of word-internal prosody. Two examples might illustrate this. First, consider stress in Spanish and French. In Spanish, stress falls on one of the word's last three syllables (Navarro Tomas 1965), and there are minimal pairs of words that differ only as far as the location of stress is concerned, for instance *bébe* '(s/he) drinks' - *bebé* 'baby'. In French, by contrast, stress does not carry lexical information. Rather, it predictably falls on the word's final vowel (Schane 1968; Dell 1973). Thus, speakers of Spanish have to process and represent stress to identify the lexical item(s) intended by the speaker. Speakers of French, by contrast, do not need to process stress, at least not in the same way.[1] Dupoux *et al*. (1997) found that French subjects are 'deaf' to stress. That is, French subjects – as opposed to Spanish subjects - exhibit great difficulties in discriminating non-words that differ only in the location of stress. Another example of word-internal prosody concerns the use of vowel length. In Japanese, but not in French, vowel length is contrastive. Thus, we find minimal pairs in Japanese such as [to] 'door' and [too] 'tower'. In French, by contrast, vowel length is not used to make lexical distinctions. Accordingly, Dupoux *et al*. (1999) found that French, but not Japanese listeners have great difficulties in distinguishing between non-words that differ only in vowel length.

Little is known about the acquisition of suprasegmentals in young infants. However, there is evidence that newborns are already sensitive to some global suprasegmental properties. In particular, it appears that on the basis of these properties, they distinguish between their native language and a foreign language, as well as between different foreign languages. Thus, Mehler *et al*. (1988, 1996) found that French infants discriminate both between French and Russian utterances and between English and Italian utterances. The stimuli were low-pass filtered, indicating that discrimination can be achieved on the basis of suprasegmental information only. Similarly, Moon, Cooper & Fifer (1993) found that English-acquiring newborns prefer to listen to English rather than to Spanish sentences, while Spanish-acquiring newborns show the reverse preference pattern (see also Nazzi, Bertoncini & Mehler 1998).

Furthermore, Jusczyk *et al*. (1993) showed that 6-month-old American infants prefer to listen to English rather than to Norwegian words, while they fail to show a preference for English words as opposed to Dutch words. These results also hold when the stimuli are low-pass filtered, suggesting that infants are sensitive to suprasegmental properties that are typical of their native language; English and Dutch indeed share many of these properties, whereas Norwegian suprasegmentals differ substantially from those of English. Along the same lines, Jusczyk, Cutler, & Redantz (1993) found that 9-month-old American infants prefer to listen to disyllables with the metrical pattern which is predominant in English, i.e. with stress on the first rather than on the second syllable. However, this experiment has not been carried out with a language that shows the reverse metrical pattern. It could, therefore, be the case that the obtained preference of American infants stems from a universal bias, rather than being related to the predominant metrical pattern of disyllables in the native language.

[1] Rather, stress may be used as a cue to word segmentation (Trubetzkoj 1939; Rietveld 1980).

Finally, it is currently unknown at what age infants begin to exhibit the type of ‘deafness’ to foreign prosodic contrasts that has been found in adults.

## 2.2 Lexical processing and acquisition

During the first year of life, infants not only acquire many phonological properties of their language, they also begin to build a lexicon. In this section, we review data regarding lexical processing and acquisition.

It has been argued that in adult speech processing, function words and content words are processed differently. For instance, Friederici (1985) reported that in a word-monitoring experiment, responses to function words were faster than responses to content words. This suggests that function words and content words are not stored together. In fact, given that the set of function words is extremely limited, search procedures within this set are faster than that within the set of content words. Similarly, Neville, Mills & Lawson (1992) found that the brain elicits qualitatively different electrophysiological responses to function words and content words, respectively.

There is recent experimental evidence that the distinction between function words and content words is acquired early in life. In particular, Shady, Jusczyk & Gerken (1988) (cf. Jusczyk, in press) found that 10½-month-old American infants listen shorter to passages in which function words are replaced by non-words having the same phonological properties. By contrast, they do not listen shorter if content words are replaced by non-words having the same phonological properties. This suggests that at this age, infants not only make a distinction between function words and content words, but also recognize the actual function words of English. By contrast, they do not know the semantics of these words, as evidenced by a follow-up experiment. That is, infants listened equally long to normal passages and to passages in which the function words were exchanged among each other, leading to ungrammatical sentences.

Gerken (1996) and Morgan, Shi & Allopenna (1996) proposed that infants acquire function words before content words and that they do so on the basis of phonological cues. Specifically, function words typically share phonological properties that set them apart from content words. In English, for instance, function words are characterized by having a short duration and low relative amplitude, a simple syllable structure, and centralized vowels. Moreover, they tend to occur utterance-initially. Shi (1995) showed that taken together, these cues are sufficient for a self-organizing neural network to classify words as function words or content words with an accuracy of 85-90 percent. Shi, Morgan & Allopenna (1998) obtained similar results with two unrelated languages, i.e. Mandarin Chinese and Turkish, suggesting that function words can universally be set apart from content words on the basis of acoustic, phonological and distributional cues only.

As to the beginning of the compilation of a lexicon of content words, Jusczyk & Aslin (1995) showed that it could lie as early as 7½ months of age. That is, infants of this age listen longer to passages containing a word to which they are habituated than to passages that do not contain such a word. The same results are obtained if infants are habituated to passages containing several instances of certain words and tested on these words in isolation. Thus, infants listen longer to words that are contained in the passages they heard previously than to words that are not contained in the passages. Moreover, words appeared to be stored in a detailed phonetic representation. For instance, when trained on *cup*, infants show no recognition of *tup*, which differs only as far as the place of articulation of the first segment is concerned.

Benedict (1979) reported - on the basis of comprehension tests as well as observational data in mothers’ diary notes - that English-learning infants of 10 months comprehend around 10 words; this figure grows to around 40 at 12 month, and to 100 or more at 16 months. Recent experimental work is consistent with this report. For instance, Hallé & Boysson-Bardies (1994) found that at 10 months of age, French infants prefer to listen to a list of 12 familiar rather than to a list of 12 unfamiliar words. More surprisingly, Mandel, Jusczyk & Pisoni (1995) reported that 4½-month-old infants recognize their own name. Specifically, infants were shown to prefer to listen to their own name rather than to names with the same number of syllables and the same stress pattern. Moreover, such preference disappears if the initial phoneme of the infants name is changed (Nazzi & Jusczyk, personal communication).

Hence, recognition of the proper name of the infant seems to be based on precise segmental information rather than on global prosodic properties.

### 2.3 Open questions

As it is apparent in the above review, both phonological acquisition and lexical acquisition begin very early in life and develop rapidly. However, many questions regarding the mechanisms that are responsible for such speedy acquisition remain open and the proposed learning mechanisms contain paradoxical circularities. On the one hand, the acquisition of certain phonological properties, such as phonotactics or the typical prosodic shape of words, seems to require the prior acquisition of a lexicon of a reasonable size in order to extract some stable statistics. On the other hand, lexical acquisition itself seems to require some prior phonological knowledge, such as phoneme categories and language-specific word boundary cues. This raises questions regarding the relative time course of phonological and lexical acquisition, as well as the potential interactions between the two processes. In the next section, we examine the various theoretical alternatives.

## 3. Setting up the research framework

### 3.1 Assumptions and hypotheses

Before discussing different theoretical pathways in early language acquisition, let us first describe three underlying assumptions (cf. Mehler, Dupoux & Segui 1990). First, infants have an innate *universal phonetic representation*. This representation encodes speech sounds and keeps all phonetic distinctions that can in principle be used contrastively, while reducing the importance of non-linguistic variables such as talker voice, length of vocal tract, and background noise. Second, infants similarly acquire a *lexicon* of word forms during the first years of life. This lexicon distinguishes between content words and function words. Third, during the first years of life, infants acquire a language-specific *prelexical representation.* This representation is intermediate between the universal phonetic representation and the lexicon. It is discrete and encodes only a small subset of those segmental and suprasegmental distinctions that are available at the universal phonetic level. It thus specifies the format under which the lexical items are stored. Once acquired, the prelexical representation is assimilatory; that is, foreign sound patterns are assimilated to the closest native sound pattern.[2] This holds not only for the segmental inventory, as in the Perceptual Assimilation Model of Best (1994), but also for all other aspects of phonological structure, such as phonotactics and suprasegmentals. For instance, foreign words that contain illegal syllabic structure will be regularized through the assimilation to existing syllables in the native language (Dupoux *et al*. 1999). Finally, the prelexical representation is crystallized in the adult; that is, it remains stable even after extensive exposure to a second language. All these properties result in patterns of phonological 'deafness' for certain non-native contrasts, i.e. those contrasts involving sounds that are assimilated to a single sound pattern in the prelexical representation.

Within this framework, the acquisition problem can be stated as follows: At birth, the language-specific prelexical representation and lexicons of content words and function words are unavailable to the infant, while a few years later, the acquisition of these components has been completed; how, then, has this been accomplished? The paradox is that each one of these components of the processing system seems to require the prior acquisition of the other one before it can itself be acquired. On the one hand, lexical acquisition seems to require a language-specific prelexical representation. In fact, a given word can surface in a near infinity of phonetic forms that - if the lexicon were constructed on the basis of a universal phonetic representation - would all be mapped onto separate lexical entries. On the other hand, in order to acquire a prelexical representation, infants seem to need minimal pairs in order to decide which variation is pertinent; that is, they need to have a lexicon. This bootstrapping problem is one of the most puzzling questions in early language development. In the following, we discuss two theoretical possibilities to solve the puzzle. The first one we dub *Bottom-Up*

---

[2] If the foreign sound is too far away from any native sound (as in the case of Zulu clicks for English speakers), it is not assimilated but perceived on a purely phonetic basis (see Best *et al*. 1988).

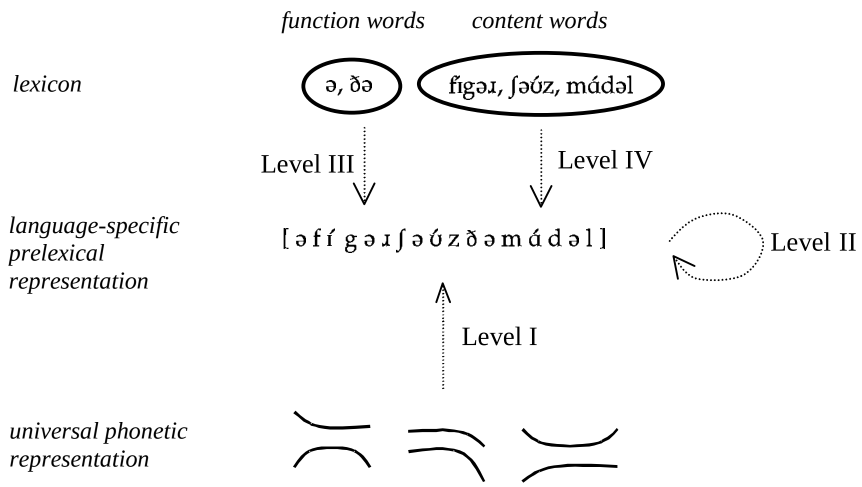


Figure 1: compilation of prelexical representation in early language acquisition

*Bootstrapping*, the second one *Interactive Bootstrapping*.

In Bottom-Up Bootstrapping, acquisition begins on the basis of the acoustic signal only, and proceeds sequentially; once a component is acquired, the next component comes into play, using only information available at lower levels (cf. Mehler, Dupoux & Segui 1990; Christophe & Dupoux 1996; Mehler *et al*. 1996). In our framework, this would mean that the prelexical representation would be acquired first, exclusively on the basis of information that is available in the universal phonetic representation. This representation would then be used to extract words. Given that only partial linguistic information would be available for the compilation of the prelexical representation, it would not necessarily be fully optimized for the native language. In particular, in the absence of lexical information, evidence based on the presence or absence of minimal pairs would not be taken into account. Allophones, then, would generally be encoded separately.[3] In other words, the prelexical representation would contain some redundancy.

In Interactive Bootstrapping, by contrast, lexical and phonological acquisition begin simultaneously and interact with one another. In our framework, this would mean that part of the phonological representation would be acquired bottom-up through inspection of the phonetic representation, and part would be acquired on the basis of lexical information. Similarly, the first words would be acquired from the universal phonetic representation, while they would be recoded more and more abstractly as more phonology becomes available. Thus, in Interactive Bootstrapping, each component could in principle be optimized for all language-specific properties, all relevant sources of information being available. As to the prelexical representation, its acquisition would continue until all redundancy would have been removed. In other words, all and only those distinctions that are used contrastively would be encoded.

### 3.2 The acquisition of the prelexical representation: a typology

In the remaining part of this paper, we compare Bottom-up Bootstrapping and Interactive Bootstrapping by focussing on the acquisition of the prelexical representation. The question we investigate concerns the types of information that allow infants to compile the prelexical representation and decide whether a given phonological distinction is to be kept or not in this representation. The prelexical representation is shown in Figure 1; the square brackets represent the utterance boundaries. The four types of information that could be taken into account during the compilation of the prelexical representation are numbered in increasing order of complexity. First of all, Type I information can be extracted from the universal phonetic representation. Type II information can be extracted from the prelexical representation itself by making reference to already acquired phonological properties. Type III information

[3] In Peperkamp & Dupoux (submitted), we argue that some allophones can be detected on the basis of the phonetic representation only.

regards the distribution and shape of function words. Finally, Type IV information concerns the location of content word boundaries in the utterance.

Bottom-up Bootstrapping and Interactive Bootstrapping make different predictions as to the types of information that can affect the compilation of the prelexical representation. That is, according to Bottom-up Bootstrapping, only low-level information, i.e. Type I and Type II, can affect prelexical acquisition; according to Interactive Bootstrapping, by contrast, high-level information, i.e. Type III and IV, might also come into play. 'Deafnesses' originating from Type I and Type II information, then, would support both Bottom-up Bootstrapping and Interactive Bootstrapping, while 'deafnesses' originating from Type III and Type IV information would provide evidence in favor of Interactive Bootstrapping only.

We can now classify phonological generalizations according to the type of information that is required in order for them to be observed. Thus, a phonological generalization of Type 1 is a generalization that can be observed on the basis of Type 1 information, and so forth. We exemplify one by one the four types of generalizations and determine the corresponding deafnesses that are or might be attested. First, regarding Type I, certain generalizations relevant to the prelexical representation are directly observable on the basis of the universal phonetic representation. For instance, there is some evidence that the distribution of vowels around language-specific prototypes is observable at an acoustic level (Kuhl *et al.* 1997). Consequently, one observes an early prelexical acquisition of prototypic vowels in the language (Kuhl *et al.* 1992), yielding language-specific 'deafnesses' to contrasts of vowels that are phonetically different but map onto the same prototype (e.g. Pallier, Bosch & Sebastián-Gallés 1997).

Second, Type II generalizations require the prior acquisition of other aspects of the phonology of the language. For instance, German has both voiced and unvoiced obstruents, but syllable-finally only the latter can occur (Wiese 1996). In order for German infants to observe this regularity, they should have access to German syllable structure. Thus, once syllable structure is encoded in the prelexical representation, the regularity concerning the absence of syllable-final voiced obstruents can be made. Provided that the latter observation feeds back on the compilation of the prelexical representation, the voicing feature will not be encoded in the prelexical representation for syllable-final consonants. German adults, then, are predicted to exhibit a 'deafness' to syllable-final voicing contrasts.

Similarly, in Italian, vowel length is allophonic: vowels are lengthened if and only if they occur in an open stressed syllable (Vogel 1982). In order to find this regularity, infants need to have acquired both the distinction between open and closed syllables and the distinction between stressed and unstressed syllables in their language. Once such information is contained in the prelexical representation, the higher-order observation (that is the correlation of syllable structure, stress, and vowel length) becomes available. Again, provided that this higher order observation can feed back on the compilation of the prelexical representation, vowel length is predicted to be removed from the prelexical representation. As a consequence, Italian-speaking adults should have difficulties distinguishing between short and long vowels; that is, they should exhibit a 'deafness' to vowel length contrasts.

Third, Type III generalizations can be observed only once the distinction between content words and function words is made. For instance, in Dutch, words can begin with any vowel except schwa. The prohibition on word-initial schwa, though, does not hold for function words, witness examples such as *een* [ən] 'a' and *het* [ət] 'it' (Booij 1995). Once infants can extract function words out of the speech stream, they can observe that the remaining strings do not begin with schwa. If this regularity is taken into account during the compilation of the prelexical representation, we might expect that in adult speech perception, foreign words with an initial schwa are misperceived and assimilated to a legal word-initial vowel.

Fourth, in order for Type IV generalizations to be observed, the boundaries of content word have to be available. For instance, in Northern varieties of Standard Italian, the contrast between [s] and [z] is allophonic; that is, within words, [z] surfaces before voiced consonants and in intervocalic position, while [s] occurs everywhere else (Camilli 1965; Nespor & Vogel 1986). In intervocalic position within phrases, then, both [z] and [s] occur, depending on the presence or absence of a word boundary within the sequence. This is illustrated in (1).

| | | | |
|---|---|---|---|
| (1) | a. | bu[z]ecca | 'bovine tripes' |
| | b. | bu[s] ecologic | 'ecological bus' |

Therefore, infants need to segment utterances into separate words in order to find the generalization concerning the distribution of [s] and [z].[4] If word boundaries or lexical knowledge can influence the compilation of the prelexical representation, one might expect that intervocalic [z] will be recoded as underlying /s/. Italian-speaking adults, then, should exhibit 'deafness' to the contrast between intervocalic [s] and [z].

The inventory presented above does not exhaust the type of 'deafnesses' that could theoretically arise. We might envision cases where only knowledge of syntactic categories or morphological decomposition can inform the prelexical level that some distinction is irrelevant and hence can be removed from the representation. The predictions are the same: if such high-level information is available to infants while they are compiling the prelexical phonological representation, then a 'deafness' to this type of distinction should be observed in adults.

In the examples given above, different phenomena (neutralization, allophonic variation, phonotactics) were involved. Hence, testing these four types of predicted 'deafness' would involve using different materials and experimental paradigms. In the remaining part of this paper, we exemplify the four-way distinction with the stress system, which allows to construct a single experimental test that can be applied cross-linguistically.

## 4. A cross-linguistic test case: stress 'deafness'

One domain in which we find limited capacities to perceive phonological contrasts that are not pertinent in the native language is that of stress. For instance, recall from section 2.1.3 that French listeners have great difficulties perceiving the difference between non-words that are distinguished only with regard to the position of stress, such as *vásuma* - *vasúma* - *vasumá* (Dupoux *et al*. 1997). In contrast, Spanish subjects have no problem in making this distinction.

In our framework, the stress 'deafness' in French adults could arise as a consequence of the way infants acquire stress. In French, stress is predictable in that main stress always falls on the word's final syllable (Schane 1968; Dell 1973), a generalization that can be inferred by infants on the basis of the universal phonetic representation. Indeed, since – as we will show in detail in section 4.1 – all utterances end with a stressed syllable, infants can deduce that stress is not contrastive by paying attention to the ends of utterances only. Therefore, they will not encode stress in the prelexical representation, and as adults they will exhibit a 'deafness' to stress contrasts. In Spanish, by contrast, stress falls on one of the word's last three syllables. Although there are certain restrictions on stress placement, the existence of minimal pairs such as *bebé* 'baby' – *bébe* 'drinks' shows that there is no way to reliably predict the stress location in all cases. Consequently, stress will be encoded in the prelexical representation.

French and Spanish hence represent two extreme cases: one in which stress is Type I predictable, and one in which it is unpredictable. In this section, we will explore languages that display stress regularities of the four types defined in section 3. The patterns of stress perception by adult speakers of these languages will help to determine the types of information that are taken into account in the compilation of the prelexical representation during infancy, and hence to draw conclusions regarding the validity of the two bootstrapping hypotheses.

Before going into this, however, we should define what is meant exactly by studying the perception of stress cross-linguistically. We will examine stress as it is realized in the subjects' native language, and hence test the perception of the acoustic cues that are typical of stress in the language under consideration. Crucially, we make sure that we do not manipulate variables that could be perceived as something else than stress in that language. The acoustic correlates of stress are loudness, pitch and duration (Lehiste 1970), but not every language uses all three cues to realize

[4] We abstract away from problems posed by prefixes for the distribution of [s] and [z] (see, e.g., Nespor & Vogel 1986; Peperkamp 1997)

stress. For instance, in languages with contrastive vowel length, duration is avoided as a correlate of stress (Hayes 1995). Therefore, when processing a foreign language with duration as phonetic correlate of stress, speakers of languages with contrastive length might map stressed vowels onto long vowels and unstressed vowels onto short vowels. Thus, they can assimilate stress to length, and consequently, stress ‘deafness’ will not be observed. Similarly, tone languages typically avoid pitch as a correlate of stress. Thus, when processing a foreign language with pitch as phonetic correlate of stress, speakers of tone languages might map stressed vowels onto high tone vowels and unstressed vowels onto low tone vowels. In other words, they can assimilate stress to tone, and again, stress ‘deafness’ will not be observed. For convenience’s sake, we will continue to use the term stress ‘deafness’ in referring to ‘deafness’ to the phonetic correlate(s) of stress as present in the subject’s native language. For an experimental paradigm that enables testing the perception of stress in adult speakers, see Peperkamp, Dupoux & Sebastián-Gallés (1999).

We will now turn to languages with different stress rules and spell out our predictions concerning the perception of stress by native speakers of these languages. The languages are classified according to the type of regularity presented by their stress rules.

## 4.1 Type I

First, let us look at languages in which the stress rule corresponds to a Type I generalization. French is a case at hand, the stress rule (‘stress the final vowel’) making reference only to phonetic notions. Notice that even clitics, which are typically unstressed elements, attract stress if they are phrase-final.[5] This is illustrated in (2).

(2) a. coupez [kupé] ‘cut$_{\text{IMP-PL}}$’
b. coupez-les [kupelé] ‘cut$_{\text{IMP-PL}}$-them’
c. coupez-vous-en [kupevuzã́] ‘cut$_{\text{IMP-PL}}$-yourself$_{\text{PL-DAT}}$-of them’

[5] Exceptions are *je* ‘I’ and *ce* ‘it’. These clitics have schwa as their vowel, which is deleted in phrase-final position, as in *suis-je* [sɥíʒ] ‘am I’ and *est-ce* [ɛ́s] ‘is it’. In other words, these two clitics do not undermine the surface observability of stress falling on the final vowel either.

Moreover, although eurhythmic principles induce destressing rules in French, the final and strongest stress of an utterance is never reduced (Dell 1984). Hence, neither the occurrence of enclitics nor destressing interferes with the phonetic observability of the stress rule. Therefore, all utterances in French have stress on their final syllable.[6] Infants, then, can infer that stress is not contrastive, hence, needs not be encoded in the prelexical representation.

Pitta-Pitta (Blake 1979; Hayes 1995) is another example of a language with a stress rule that should give rise to Type I ‘deafness’. This Australian language has fixed word-initial stress (3), and there are no proclitics.

(3) a. kárru ‘boy’
b. mílilpu ‘eyebrow’
c. yápariri ‘young man’

Pitta-Pitta has no monosyllabic words; stress clash, therefore, never occurs. Thus, all utterances have stress on the first syllable. Given that stress is predictable and observable on the basis of the universal phonetic representation, both the Bottom-up Bootstrapping hypothesis and the Interactive Bootstrapping hypothesis predict that stress is not encoded in the prelexical representation and that speakers of Pitta-Pitta exhibit stress ‘deafness’.

## 4.2 Type II

An example of a stress system involving a Type II regularity is presented by Fijian (Schütz 1985; Dixon 1988; Hayes 1995). In this Austronesian language, word stress falls on the final syllable if it is heavy; otherwise stress is penultimate. The language has only two syllable types, (C)VV and (C)V, where the former is heavy and the latter is

[6] In southern varieties of French, words can end in an unstressed schwa. In these varieties, stress thus falls on the word’s last full vowel. The acquisition of the stress regularity, then, is more complex, in that infants first have to acquire the difference between full vowels and schwa. Consequently, the stress ‘deafness’ is of the second rather than of the first type.

Note also that, alternatively, French has been characterized as having phrasal stress, with stress falling on phrase-final syllables (Grammont 1965). The question as to whether French has word stress or phrasal light. Suffixes are within the stress domain,[7] and there are no enclitics. Examples from Dixon's (1988) description of the Boumaa dialect are given in (4); syllable boundaries are indicated by dots.

(4) a. lú.a lu.á.ca 'vomit – vomit on$_{TR}$'
b. te.ʔe.vúː te.ʔe.vúː.na 'start - start$_{TR}$'
c. pu.lóu pu.lóu.na 'be covered - cover$_{TR}$'

The language permits monosyllabic words provided they are heavy. If a monosyllable is preceded by a word ending in a long vowel or a diphthong, a stress clash arises. We do not know if and how stress clash is resolved in Fijian. Clearly, if the second stress undergoes destressing, utterance-final stress clash configurations disrupt the surface stress pattern. Abstracting away from this potential confound, however, we can formulate the following surface generalization: in utterances that end in a word with a final long vowel or a diphthong, the final syllable is stressed, while in utterances that end in a word with a final short vowel, the penultimate syllable is stressed. Once infants have acquired the distinction between heavy and light syllables, they can observe the stress regularity. Speakers of Fijian, then, are predicted to exhibit stress 'deafness' if this regularity is taken into account during the compilation of the prelexical representation As with the Type I regularities, this follows from both the Bottom-up Bootstrapping hypothesis and the Interactive Bootstrapping hypothesis.

## 4.3 Type III

In many languages, stress is predictable and observable *modulo* the occurrence of clitics. In fact, contrary to the situation in French described in section 4.1, clitics are typically unstressed, regardless of their position in the utterance. Consider, for instance, the case of Hungarian. In this language, stress falls on the word-initial syllable, and clitics are systematically unstressed (Vago 1980). This is illustrated in (5).

(5) a. emberek [émberek] 'men'
b. az emberek [azémberek] 'the men'

stress is irrelevant to our point, given that under both assumptions, all utterances end in a stressed syllable.

[7] Some suffixes and sequences of suffixes, however, form separate stress domains. That is, they behave as independent words phonologically (cf. Dixon 1988).

In Hungarian, then, utterances that begin with a clitic have stress on the second syllable, while all other utterances have stress on the first syllable. Hungarian has monosyllabic content words, but stress clash resolution does not interfere with this surface stress pattern (Vogel 1988).

In order for prelexical infants to discover the stress rule of Hungarian, they should strip off utterance-initial function words and look for generalizations in the remaining string. That is, after removing the initial function word(s), infants can discover that the remaining string of content words always begins with a stressed syllable. Bottom-up Bootstrapping states that such a discovery cannot affect the compilation of the prelexical representation. In fact, according to this hypothesis, only low-level phonetic and phonological information is taken into account during the compilation of the prelexical representation. As a consequence, adult speakers of Hungarian should not exhibit stress 'deafness', even though stress is not contrastive in their language. By contrast, Interactive Bootstrapping states that all possible sources of information – including those pertaining to the distinction between function words and content words - is used in order to determine which contrasts are pertinent in the language. According to this latter hypothesis, stress would not be encoded in the prelexical representation; hence, adult Hungarians should exhibit stress 'deafness'.

## 4.4 Type IV

Certain stress rules are observable only if the boundaries of content words are available. Consider, for instance, Piro, an Arawakan language in which word stress is on the penultimate syllable (Matteson 1965; Hayes 1995). Examples from Matteson (1965) are given in (6).

(6) a. nsó 'genipa'
b. wálo 'rabbit'
c. rutxítxa 'he observes taboo'
d. ʧiyaháta 'he cries'

Given that there are no enclitics, the following generalization emerges. In utterances ending in a monosyllabic word, the final syllable is stressed, whereas in all other utterances, the penultimate syllable is stressed. In order to extract the rule regarding penultimate stress, prelexical infants should have access to content word boundaries.

Therefore, the Interactive Bootstrapping hypothesis but not the Bottom-up Bootstrapping hypothesis predicts that speakers of Piro exhibit stress 'deafness', since only in the former can word boundaries influence the compilation of the prelexical representation.[8]

Another example of a purely phonological stress rule that is observable only if the boundaries of content words are available is presented by Nyawaygi (Dixon 1983; Hayes 1995). In this Australian language, stress is assigned at the left edge of words, as follows. In words beginning with a heavy syllable, stress falls on the first syllable (7a). In words beginning with a light syllable, stress falls on the second syllable if the word contains three or five syllables (7b) and on the first syllable if the word contains two or four syllables (7c). Notice that coda consonants are weightless in Nyawaygi; heavy syllables are syllables containing a long vowel.

(7)a. *heavy initial syllable*:
ŋú:ɡ̡u 'fish'
ɡ̡í:baɽi 'south'
b. *light initial syllable, uneven nb. of syllables*:
bulbíri 'quail'
c. *light initial syllable, even nb. of syllables*:
ɡ̡íŋa 'man'
bíyaɡ̡ala 'water snake'

Stress clash does not occur, since all words are minimally disyllabic and no word has final stress. Furthermore, there are no proclitics. Thus, all utterances whose first word begins with a heavy syllable have initial stress, whereas all other utterances have stress on either the first or the second syllable, depending on the number of syllables in the first word. In order to extract this generalization, infants not only need to be sensitive to syllable weight, but they also need to have access to the number of syllables in the utterance's first word; hence, they must be able to segment the first word out of the utterances. Therefore, as with speakers of Piro, native speakers of Nyawaygi are predicted to exhibit stress 'deafness' according to Interactive Bootstrapping but not according to Bottom-up Bootstrapping.

[8] For the sake of simplicity, we deliberately chose a language without enclitics. Of course, a language with penultimate stress that has enclitics would fall into the same class of our 'deafnesses' typology, since lexical segmentation includes clitic stripping.

### 4.5 A caveat

Before closing our typology of stress 'deafness', we would like to raise the following caveat. Our reasoning is based on the assumption that the stress pattern at utterance edges allows infants to make inferences about the general stress rule of their language. For instance, we postulate that from the observation that all *utterances* end with a stressed syllable, infants will infer that all *words* end with a stressed syllable. This, of course, may be unwarranted. Phonological rules can indeed apply at some phrasal edge only. In that case, a regularity appears at some utterance edge which does not hold, however, for individual words. An example is provided by the tonal system of Slave. In this language, there are two tones, high and low. At the end of intonational phrases, the distinction is neutralized; the final word necessarily surfaces with a low tone (Rice 1987). If infants were to examine the end of utterances only, they would incorrectly conclude that there is no tone in their language.

Yet, stress seems to be special, in that – to the best of our knowledge - this kind of situation does not arise with stress rules. On the contrary, in the cases in which stress is modified at a phrasal edge, the modification seems to enhance the regularity of the stress pattern instead of obscuring it. For instance, recall from section 4.1 that in French, phrase-final enclitics are stressed. This pattern reinforces the utterance-final regularity caused by the word-final stress pattern of French. By contrast, we have found no language in which, analogously to tone in Slave, contrastive stress is neutralized at a phrasal edge. It should be noted that if our model is correct, there cannot be such a language. In fact, if it existed, infants would use the edge regularity to infer that stress is predictable, and hence they would become stress 'deaf'; this, then, would have the effect that stress contrasts would be lost within one generation.

To conclude, the absence of languages with contrastive stress that is neutralized at some phrasal edge is quite critical for our model. If such languages exist, we have to revise our proposal, by offering another, more complex, learning principle that allows infants to acquire the stress rule of their language. Note that even in that case, our hierarchy of languages would still be pertinent to describe the relative observability of these stress rules from the infant's viewpoint. For instance, stress in French (Type I) would still be easier to acquire than stress in Nyawaygi (Type IV).

## 5. Summary and conclusion

We have shown that cross-linguistic research of adult speech perception can be brought to bear on issues of language acquisition. In particular, it might allow to solve some of the most puzzling issues in early language development, namely, the paradoxical fact that the acquisition of each processing component seems to require the prior acquisition of the other components.

As a first step to solve the acquisition paradox, we proposed to distinguish two hypotheses, i.e. Bottom-up Bootstrapping and Interactive Bootstrapping. The former states that acquisition proceeds step by step; specifically, the acquisition of each level exclusively relies upon information that is available in the shallower levels of representation. According to the latter hypothesis, in contrast, all levels begin to be acquired independently and there is interaction between levels until a stable state is attained.

We set up our methodology to explore the acquisition of the prelexical representation and focussed on the stress system, since it allows a uniform way of distinguishing the various sources of information that could in principle be taken into account in the compilation of the prelexical representation. We thus proposed to test whether stress is encoded in this representation by speakers of a variety of languages. In particular, we argued that the *presence* of stress 'deafness' is an indication of the *absence* of stress in the prelexical representation. The underlying assumption is that stress should only be encoded in the prelexical representation if it is useful to distinguish lexical items. The infant's problem is to decide, on the basis of a limited amount of information, whether stress should be encoded or not.

We distinguished four classes of languages, corresponding to four types of information that infants could use in order to make this decision. That is, in the first class of languages (Type I), the stress rule can be acquired on the basis of a universal phonetic representation only; in the second class (Type II), it can be acquired once certain language-specific phonological information has been extracted; in the third and the fourth class (Type II and IV, respectively), it can be acquired only after function words and content words, respectively, can be segmented out of the speech stream. We then made contrasting predictions regarding the presence of stress 'deafness' in adults of these four language classes. The reasoning is as follows. Whenever stress 'deafness' is exhibited by speakers of a certain language, we take this as evidence that stress is not encoded prelexically, hence, that the stress regularity is not taken into account during the compilation of the prelexical representation. According to the Bottom-Up Bootstrapping hypothesis, only the universal phonetic and the prelexical representation itself can enter the scene during the compilation of the prelexical representation, hence only 'deafnesses' of Types I and II are predicted. In contrast, the Interactive Bootstrapping hypothesis states that all processing components are allowed to interact till a stable state is attained. Hence, one would expect all four types of 'deafnesses' to be observed. A summary of the languages with their characteristics and the various predictions is given in Table 1.

The empirical investigation that we propose thus allows to test models of early language acquisition. Such empirical investigation is now under way.

As a final note, we would like to mention that although we have focussed on the acquisition of the prelexical representation, our approach can be transposed to other processing components that are involved in early language acquisition, such as word segmentation or lexical access. Provided that these components show clear language-specific properties in adults, a cross-linguistic comparative study can probe what types of information are used by infants to compile such language-specific routines. In other words, the comparative psycholinguistics approach using only adult data is a potentially useful tool that can be added to the panoply of techniques currently used to investigate first language acquisition.

Table 1: Summary of predictions for the four classes of stress systems.

| *Regularity* | *Language* | *Stress Rule* | *Interfering Destressing* | *Clitics* | *Stress 'Deafness'* |
|---|---|---|---|---|---|
| Type I | French | stress final syllable | no | stressed enclitics | attested (Dupoux *et al.* 1997) |
| | Pitta-Pitta | stress initial syllable | not applicable | no proclitics | predicted by both Bottom-up and Interactive Bootstrapping |
| Type II | Fijian | stress heavy penultimate, otherwise antepenultimate syllable | ? | no enclitics | predicted by both Bottom-up and Interactive Bootstrapping |
| Type III | Hungarian | stress first syllable | no | unstressed proclitics | predicted only by Interactive Bootstrapping |
| Type IV | Piro | stress penultimate syllable | ? | no enclitics | predicted only by Interactive Bootstrapping |
| | Nyawaygi | stress initial syllable if heavy or if word contains even number of syllables; otherwise, stress second syllable | not applicable | no enclitics | predicted only by Interactive Bootstrapping |

## References


Benedict, H. 1979 Early lexical development: comprehension and production. *Journal of Child Language* 6. 183-200.

Best, C. 1994 The emergence of native-language phonological influence in infants: A perceptual assimilation model. In J. Goodman and H. Nusbaum, eds. *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*. 167-224. Cambridge, MA: MIT Press.

Best, C., G. McRoberts, and N. Sithole, 1988 Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14. 345-360.

Blake 1979 Pitta-Pitta. In R. Dixon and B. Blake, eds. *Handbook of Australian Languages, Vol. 1*. 182-242. Amsterdam: John Benjamins.

Booij, G. 1995 *The Phonology of Dutch*. Oxford: Clarendon Press.

Brown, R. and D. Hildum 1956 Expectancy and the perception of syllables. *Language* 32. 411-419.

Camilli, A. 1965 *Pronuncia e grafia dell'italiano*. Firenze: Sansoni.

Christophe, A. and E. Dupoux 1996 Bootstrapping lexical acquisition: the role of prosodic structure. *The Linguistic Review* 13. 383-412.

Dell, F. 1973 Les Règles et les Sons. Paris: Hermann.

Dell, F. 1984 L'accentuation des phrases en français. In F. Dell, D. Hirst and J.-R. Vergnaud, eds. *Forme Sonore du Language*. 65-122. Paris: Hermann.

Dixon, R. 1983 Nyawaygi. In R. Dixon and B. Blake, eds. *Handbook of Australian Languages, Vol. 3*. 430-531. Amsterdam: John Benjamins.

Dixon, R. 1988 *A Grammar of Boumaa Fijian*. Chicago: The University of Chicago Press.

Dehaene-Lambertz, G., E. Dupoux and A. Gout (in press) Electrophysiological correlates of phonological processing: a cross-linguistic study. *Journal of Cognitive Neuroscience*.

Dupoux, E., K. Kakehi, Y. Hirose, C. Pallier, and J. Mehler 1999 Epenthetic vowels in Japanese: a perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25. 1-11.

Dupoux, E., C. Pallier, N. Sebastian, and J. Mehler 1997 A destressing 'deafness' in French? *Journal of Memory and Language* 36. 406-421.

Friederici, A. 1985 Levels of processing and vocabulary types: Evidence from on-line comprehension in normals and agrammatics. *Cognition* 19. 133-166.

Friederici, A. and J. Wessels 1993 Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception and Psychophysics* 54. 287-295.

Gathercole, S., C. Frankish, S. Pickering, and S. Peaker 1999 Phonotactic influences on short term memory. *Journal of Experimental Psychology: Learning, Memory and Cognition* 25. 84-95.

Gerken 1996 Phonological and distributional information in syntax acquisition. In J. Morgan and K. Demuth, eds. *Signal to Syntax. Bootstrapping from Speech to Grammar in Early Acquisition*. 411-425. Mahwah, N.J.: LEA.

Goto, H. 1971 Auditory perception by normal Japanese adults of the sounds ‘l’ and ‘r’. *Neuropsychologia* 9. 317-323.

Grammont 1965 Traité de phonétique. Paris: Delagrave.

Hallé, P. and B. de Boysson-Bardies 1994 Emergence of an early receptive lexicon: infants' recognition of words. *Infant Behavior and Development* 17. 119-129.

Hallé, P., J. Segui, U. Frauenfelder, and C. Meunier 1998 Processing of illegal consonant clusters: a case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance* 24. 592-608.

Hayes, B. 1995 *Metrical Stress Theory*. Chicago: The University of Chicago Press.

Jusczyk, P. 1999 Dividing and conquering linguistic input. In M. Gruber, K. Olson and T. Wysocki, eds. *Proceedings of the Chicago Linguistic Society 34. Volume II: The Panels*. Chicago: Chicago Linguistic Society.

Jusczyk, P. and R. Aslin 1995 Infants’ detection of sound patterns of words in fluent speech. *Cognitive Psychology* 29. 1-23.

Jusczyk, P., A. Cutler, and N. Redantz 1993 Preference for the predominant stress pattern of English words. *Child Development* 64. 675-687.

Jusczyk, P., A. Friederici, J. Wessels, V, Svenkerud, and A. Jusczyk 1993 Infants’ sensitivity to the sound pattern of native language words. *Journal of Memory and Language* 32. 402-420.

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

Kuhl, P., J. Andruski, I. Chistovich, L. Chistovich, E. Kozhevnikova, V. Ryskina, E. Stolyarova, U. Sundberg, and F. Lacerda 1997 Cross-language analysis of phonetic units in language addressed to infants. *Science* 277(5326). 684-6.

Kuhl, P., K. Williams, F. Lacerda, K. Stevens, and B. Lindblom 1992 Linguistic experience alters phonetic perception in infants by six months of age. *Science* 255. 606-608.

Lehiste, I. 1970 *Suprasegmentals*. Cambridge, MA: MIT Press.

Mandel, D., P. Jusczyk, and D. Pisoni 1995 Infants’ recognition of the sound pattern of their own names. *Psychological Science* 6. 314-317.

Massaro, D. and M. Cohen 1983 Phonological context in speech perception. *Perception and Psychophysics* 34. 338-348.

Matteson, E. 1965 *The Piro Arawakan Language*. Berkeley and Los Angeles: University of California Press.

Mehler, J., J. Bertoncini, E. Dupoux, and C. Pallier 1996 The role of suprasegmentals in speech perception and acquisition. In T. Otake and A. Cutler, eds. *Phonological structure and language processing. Cross-linguistic studies*. 145-169. Berlin: Mouton de Gruyter.

Mehler, J., E. Dupoux, and J. Segui 1990 Constraining models of lexical access: The onset of word recognition. In G. Altman, ed. *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. 236-262. Cambridge, MA: MIT Press.

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

Moon, C., R. Cooper, and W. Fifer 1993 Two-day-olds prefer their native language. *Infant behavior and development* 16. 495-500.

Morgan, J., R. Shi, and P. Allopenna 1996 Perceptual bases of rudimentary grammatical categories: Toward a broader conceptualization of bootstrapping. In: J. Morgan and K. Demuth eds. *Signal to Syntax. Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, N.J.: LEA. 263-283.

Näätänen, R., A. Lehtokovski, M. Lennes, M. Cheour, M. Huotilainen, A. Iivonen, M. Vainio, P. Alku, R. Ilmoniemi, A. Luuk, J. Allik, J. Sinkkonen, and K. Alho 1997 Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, 385. 432-434.

Navarro Tomas, T. 1965 *Manual de Pronunciación Española*. Madrid: Consejo Superior de Investigaciones Científicas.

Nazzi, T., J. Bertoncini, and J. Mehler 1998 Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance* 24. 756-66.

Nespor, M. and I. Vogel 1986 *Prosodic Phonology*. Dordrecht: Foris.

Neville, H., D. Mills, and D. Lawson 1992 Fractionating language: Different neural subsystems with different sensitive periods. *Cerebral Cortex* 2. 244-258.

Pallier, C., L. Bosch, and N. Sebastián-Gallés 1997 A limit on behavioral plasticity in speech perception. *Cognition* 64. B9-B17.

Peperkamp, S. 1997 *Prosodic words*. The Hague: Holland Academic Graphics.

Peperkamp, S., and E. Dupoux submitted On infants’ acquisition of phonological alternations. Submitted to: I. Lasser ed. *Proceedings of GALA*. Peter Lang Verlag.

Peperkamp, S., E. Dupoux and N. Sebastián-Gallés 1999 Perception of stress by French, Spanish, and bilingual subjects. *Proceedings of EuroSpeech '99, Vol. 6*. 2683-2686.

Pitt, M. 1998 Phonological processes and the perception of phonotactically illegal consonant clusters. *Perception and Psychophysics* 60. 941-951.

Polivanov, E. 1974 The subjective nature of the perceptions of language sounds. In E. Polivanov. *Selected Works. Articles on General Linguistics*. Compiled by A. Leont'ev. 223-237. The Hague: Mouton

Polka, L. 1991. Cross-language speech perception in adults: Phonemic, phonetic and acoustic contributions. *Journal of the Acoustical Society of America* 89 (6). 2961-2977.

Polka, L. and J. Werker 1994 Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance* 20. 421-435.

Rice, K. 1987 On defining the intonation phrase: evidence from Slave. *Phonology Yearbook* 4. 37-59.

Rietveld, A. 1980 Word boundaries in the French language. *Language and Speech* 23. 289-296.

Sapir, E. 1921 Language. New York: Harcourt Brace Jovanovich.

Shady, M., P. Jusczyk, and L. Gerken 1998 *Infants' sensitivity to function morphemes*. Paper presented at the 23td Annual Boston University Conference on Language Development, Boston, MA.

Schane, S. 1968 French Phonology and Morphology. Cambridge, MA: MIT Press.

Shi, R. 1995 *Perceptual Correlates of Content Words and Function Words in Early Language Input*, unpublished Ph.D. diss., Brown University.

Shi, R., J. Morgan, and P. Allopenna 1998 Phonological and acoustic bases for earliest grammatical category assignment: a cross-linguistic perspective. *Journal of Child Language* 25. 169-201.

Schütz, A. 1985 *The Fijian Language*. Honolulu: University of Hawaii Press.

Trubetzkoy, N. 1939 *Grundzüge der Phonologie*. Travaux du Cercle Linguistique de Prague 7.

Vago, R. 1980 *The Sound Pattern of Hungarian*. Washington, D.C.: Georgetown University Press.

Vitevitch, M., P. Luce, J. Charles-Luce, and D. Kemmerer 1997. Phonotactics and syllable stress: implications for the processing of spoken nonsense words. *Language and Speech* 40. 47-62.

Vogel, I. 1982 *La sillaba come unità fonologica*. Bologna: Zanichelli.

Vogel, I. 1988 Prosodic constituents in Hungarian. In P.M. Bertinetto and M. Loporcaro, eds. *Certamen Phonologicum. Papers from the 1987 Cortona Phonology Meeting*. 231-250. Tornino: Rosenberg & Sellier.

Werker, J. and R. Tees 1984a Cross language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7. 49-63.

Werker, J. and R. Tees 1984b Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America* 75. 1866-1878.

Wiese, R. 1996 *The phonology of German*. Oxford: Oxford University Press.