reprinted by the authors from:

Carlos Gussenhoven and Natasha Warner, eds. 2002. *Laboratory Phonology 7*. Berlin: Mouton de Gruyter, 203-240.

A typological study of stress 'deafness'

Sharon Peperkamp
Laboratoire de Sciences Cognitives et Psycholinguistique (EHESS/CNRS), 54 B[d] Raspail, 75006 Paris, France.
Département de Sciences du Langage, Université de Paris VIII, 2 Rue de la Liberté, 93535 Saint Denis, France.
E-mail: sharon@lscp.ehess.fr
Web page: http://www.lscp.net/persons/peperkamp

Emmanuel Dupoux
Laboratoire de Sciences Cognitives et Psycholinguistique (EHESS/CNRS), 54 B[d] Raspail, 75006 Paris, France.
E-mail: dupoux@lscp.ehess.fr
Web page: http://www.lscp.net/persons/dupoux

Abstract

Previous research has shown that native speakers of French, as opposed to those of Spanish, exhibit stress 'deafness', i.e. have difficulties distinguishing stress contrasts. In French, stress is non-contrastive, while in Spanish, stress is used to make lexical distinctions. We examine three other languages with non-contrastive stress, Finnish, Hungarian and Polish. In two experiments with a short-term memory sequence repetition task, we find that speakers of Finnish and Hungarian are like French speakers (i.e. exhibit stress 'deafness'), but not those of Polish. We interpret these findings in the light of an acquisition framework, that states that infants decide whether or not to keep stress in their phonological representation during the first two years of life, based on information extractable from utterance edges. In particular, we argue that Polish infants, unlike French, Finnish and Hungarian ones, cannot extract the stress regularity of their language on the basis of what they have already learned. As a consequence, they keep stress in their phonological representation, and as adults, they do not have difficulties in distinguishing stress contrasts.

1. Introduction

It has long been known that speech perception is influenced by phonological properties of the listener's native language (Sapir 1921; Polivanov 1974). Much experimental evidence has been gathered concerning this influence, suggesting that listeners use a set of language-specific phoneme categories during speech perception. For instance, Goto (1971) has documented that Japanese listeners map American [l] and [r] onto their own, single, [ɾ] 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].

The influence of suprasegmental properties of the native language has also been investigated. Research in this area has concentrated on the perception of tone (Kiriloff 1969; Bluhme & Burr 1971; Gandour 1983; Lee & Nusbaum 1993; Wang, Spence & Sereno 1999), but the perception of stress was recently investigated as well. In particular, Dupoux, Pallier, Sebastián-Gallés & Mehler (1997) found that French subjects exhibit great difficulties in discriminating non-words that differ only in the location of stress. In French, stress does not carry lexical information, but predictably falls on the word's final vowel. Speakers of French, then, do not need to process stress to identify lexical items; given its fixed position, stress may instead be used as a cue to word segmentation (Rietveld 1980).

The term 'deafness' is meant to designate the effect of listeners having difficulties in discriminating non-words that form a minimal pair in terms of certain non-native phonological contrasts, be it segmental or suprasegmental. We put the term 'deafness' between quotes, since listeners do not completely fail to perceive these contrasts (see Dupoux, Peperkamp & Sebastián-Gallés (2001) for extensive discussion on this issue). Segmental 'deafnesses' have been shown to arise early during language development. In fact, at 6 months, infants begin to lose their sensitivity for non-native vowel contrasts (Polka & Werker 1994), while between 10 and 12 months, they lose the ability to discriminate non-native consonantal contrasts (Werker & Tees 1984a). Suprasegmental 'deafnesses', by contrast, have not yet been attested experimentally in infants. However, it has been shown that 6- and 9-month-old infants are sensitive to suprasegmental properties of words in their native language (Jusczyk, Cutler & Redanz 1993; Jusczyk, Friederici, Wessels, Svenkerud & Jusczyk 1993). We therefore assume that suprasegmental 'deafnesses' similarly arise during infancy.

In this paper, we propose to extend the study of stress 'deafness' initiated by Dupoux *et al*. (1997). Two questions appear to be pertinent. The first one concerns the linguistic parameters that govern the presence of stress 'deafness'. Dupoux *et al*. (1997) suggest that the stress 'deafness' in French speakers is due to the fact that French has non-contrastive stress. This hypothesis, however, needs to be tested with speakers of other languages with non-contrastive stress. French might indeed be a special case, in that - as is sometimes suggested - it has no word stress at all (Grammont 1965); stress 'deafness', then, might be restricted to French only. Alternatively, as we will argue here, languages with non-contrastive stress might vary on other dimensions that are relevant for stress 'deafness'; accordingly, not all of these languages necessarily yield stress 'deafness'. The second question concerns the point during language development at which stress 'deafness' arises in infants. Dupoux & Peperkamp (2002) proposed a framework that allows us to explore these two questions simultaneously. In this framework, the perception of stress is assessed cross-linguistically in adults and inferences are drawn concerning the early acquisition of non-contrastive stress. In this paper, we will take up this framework and present experimental data concerning the perception of stress by adult speakers of several languages.

The outline of this paper is as follows. In section 2, we present the theoretical framework, state our assumptions, and outline different predictions regarding the perception of stress in four classes of languages with non-contrastive stress. We exemplify our typology with languages belonging to each one of the four classes in section 3. In section 4, we recall the results of Dupoux, Peperkamp & Sebastián-Gallés (2001), attesting stress 'deafness' in French speakers with a new paradigm, and report on two new experiments with the same paradigm. These latter experiments assess the perception of stress by speakers of Finnish, Hungarian, and Polish. Section 5 contains our conclusions.

2. Theoretical framework

2.1. Stress 'deafness' as a window onto early language acquisition

It is largely agreed upon that words are stored in a mental lexicon. As to the phonological representation of words in this lexicon, two opposite views have been advanced in the literature. On the one hand, Klatt (1980), Marslen-Wilson & Warren (1994), and Goldinger (1998) argue for a universal acoustically or phonetically based representation of words. On the other hand, Church (1987), Frazier (1987), and Mehler, Dupoux & Segui (1990) propose that words are represented in an abstract phonological format that is tuned to the properties of the maternal language. Mehler *et al*. (1990) argue that from the onset of lexical acquisition, infants store words in this language-specific phonological format. They claim that having learned beforehand a language-specific phonological representation helps lexical acquisition. That is, 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. By contrast, knowledge of what constitutes a lexical entry should facilitate the subsequent task of finding word meanings.

In Dupoux & Peperkamp (2002), we endorsed the hypothesis by Mehler *et al*. (1990) and further proposed that tuning of the phonological representation occurs during the first two years of life, before full mastery of the language. Such tuning, we argued, is based on an analysis of distributional regularities of the phonetic stream, rather than on a contrastive analysis involving minimal pairs. Given that infants have a limited knowledge of their language during the tuning of the phonological representation, they may include certain non-contrastive variation simply since they fail to observe its non-contrastiveness. Furthermore, we proposed that once tuned, the phonological representation of words becomes fixed and is relatively unaffected by later acquisitions in either the same or a different language (for empirical support, see Goto

1971; Best, McRoberts & Sithole 1988; Pallier, Bosch & Sebastián-Gallés 1997; Dupoux *et al.* 1997; Dupoux, Kakehi, Hirose, Pallier & Mehler 1999). This, then, allows us to gain insight into the nature and content of the phonological representation by conducting experiments in adults.

In the present study, we concentrate on the perception of stress, and focus on languages in which stress is signaled only by suprasegmental properties, i.e. duration, pitch, and energy. Regarding the representation of stress, the infants' problem is to decide, on the basis of a limited amount of information, whether stress is contrastive or not in their language, and, consequently, whether it should be encoded in the phonological representation or not. We dub this binary option the Stress Parameter and propose that it is set within the first two years of life. In its default setting, stress is encoded in the phonological representation. We distinguish three cases regarding the setting of the Stress Parameter during acquisition. Suppose a language with non-contrastive stress. If by the time the Stress Parameter is set infants can observe the stress regularity, they will set the Stress Parameter such that stress is not encoded. If, by contrast, they fail to deduce that stress is non-contrastive, they will redundantly keep stress in the phonological representation. Finally, for languages with contrastive stress, infants will observe no stress regularity and hence correctly keep stress in the phonological representation.

One might want to argue that infants simply attend to one-word utterances in order to deduce whether stress is contrastive. However, infant-directed speech does not necessarily contain many one-word utterances (Aslin, Woodward, LaMendola & Bever 1996; Van de Weijer 1999), and it is unclear how infants could distinguish between one-word and multi-word utterances (Christophe, Dupoux, Bertoncini & Mehler 1994). Alternatively, we propose that in order to set the Stress Parameter, infants rely on cues concerning the distribution of stresses at utterance boundaries.[1] Indeed, if word stress is regular, then this regularity will be present at either the beginning or the end of utterances, depending on whether stress is assigned at the left or the right edge of the word, respectively. For instance, in French, all utterances end with a stressed syllable. Assuming that the location of main stress in the last word of the utterance does not differ from that of other words, infants can deduce that stress is always word-final and hence need not be encoded. In Spanish, by contrast, stress is largely unpredictable and falls on one of the word's last three syllables (Navarro Tomás 1965). Hence, utterances neither begin nor end consistently with a main stressed syllable. Neither utterance edge thus presents a regular surface stress pattern, and infants therefore decide to keep stress in the phonological representation. Finally, in section 2.3 we will see cases in which infants might fail to deduce that stress is non-contrastive; this concerns languages in which the stress regularity is harder to extract from one of the utterance edges than in French. Before going in detail into the predictions of our framework, though, we will spell out various assumptions that underlie our proposal.

2.2. Background assumptions

First of all, we assume that utterance edges are easily attended to by young infants. This is an uncontroversial claim, since utterance edges are typically signaled by pauses in the discourse and/or by universal prosodic markers such as final lengthening. Experimental evidence that infants can segment speech into utterances is provided by Hirsh-Pasek, Kemler-Nelson, Jusczyk, Wright Cassidy, Druss & Kenndey (1987).

Stress can be instantiated by a variety of phonetic cues, i.e. duration, pitch, and energy (Lehiste 1970). Our second assumption is that infants can perceive word stress categorically by the time they come to fix the Stress Parameter. This is a relatively strong assumption, since there is no one-to-one correspondence between the abstract notion of linguistic stress and the three phonetic cues. On the one hand, the relative weighting of these three cues in the realization of stress is language-specific. For instance, in languages with contrastive vowel length, duration is used to a lesser extent than pitch and energy; and in languages with lexical tone, pitch is avoided as a stress cue (Hayes 1995). We assume that by the time they set the Stress Parameter, infants have acquired a sensitivity to the way in which word stress is realized in their language. In particular, we assume that they know whether their native language is a tone language, a pitch accent language, or a stress language, and whether or not it uses contrastive vowel length. An indication that this is a realistic assumption is that newborns rapidly become attuned to some global suprasegmental and rhythmic properties of their native language (Mehler, Jusczyk, Lambertz, Halsted, Bertoncini & Amiel-Tison 1988; Moon, Cooper & Fifer 1993; Mehler, Bertoncini, Dupoux & Pallier 1996; Nazzi, Bertoncini & Mehler 1998; Ramus, Nespor & Mehler 1999). On the other hand, in addition to signaling word stress, duration and pitch typically serve various grammatical functions. For instance, they are used to mark phrasal boundaries, focused constituents, and interrogatives. We assume that by the time they set the Stress Parameter, infants know the language-specific cues that signal the boundaries of prosodic constituents such as phonological utterances, intonational phrases, and phonological phrases (Nespor & Vogel 1986). This assumption is supported by the fact that during the first year of life, infants develop a sensitivity to increasingly smaller prosodic units (Hirsh-Pasek *et al*. 1987; Gerken, Jusczyk & Mandel 1994). Similarly, we assume that infants know the language-specific intonation contours that are associated with interrogatives, focused constituents, etc.

Third, we assume that the Stress Parameter is set at roughly the same age in all languages. This follows from a more general assumption that languages are equally learnable. Along the same lines, we assume that the tuning of the phonological representation is finished at a certain age, regardless of the language under consideration. This is important, since, as we will show below, languages differ considerably as to when the Stress Parameter can be set correctly in principle.

Fourth, we assume that infants acquire aspects of their maternal language in basically three steps. First, they acquire the segmental inventory between 6 and 12 months (Kuhl, Williams, Lacerda, Stevens & Lindblom 1992; Polka & Werker 1994; Werker & Tees 1984a). Second, they acquire the inventory of function words between 10 and 12 months (Shady 1996). Third, starting at around 10 months, they begin to acquire a lexicon of content words (Benedict 1979; Hallé & Boysson-Bardies 1994). Neither the function nor the content words are linked to a meaning at this stage. That is, they are stored in a recognition lexicon, containing word forms only. Regarding the difference between function words and content words, the former share a number of acoustic, phonological, and distributional properties that set them apart from the latter. In particular, they are mostly high frequent, unstressed, monosyllables that typically occur at phrasal edges, and have a simple syllable structure, a short duration, and a low relative amplitude. This might explain why infants can strip function words off the speech signal before they can segment content words out of the speech signal (Shi, Morgan & Allopenna 1998). At 12 months, the content word lexicon is still very small (around 40 words), but it grows rapidly; at 16 months, for instance, it contains around 100 words (Benedict 1979).

Our final assumption concerns the linguistic input and states that stress is phonologically transparent at utterance edges. By this we mean that the distribution of word stress is not obscured at utterance edges, in that initial or final words do not have a deviant stress pattern. In particular, contrastive stress should not be neutralized at the beginning or end of some phrasal constituent; such unattested neutralization would give rise to a regular pattern at one of the utterance edges, thus inducing infants to incorrectly conclude that stress is not contrastive (Peperkamp 2000).[2]

2.3. Hypotheses and predictions

Our central hypothesis is that before having acquired the entire lexicon, infants use stress patterns at utterance edges to infer whether stress is contrastive or not and, hence, set the Stress Parameter. However, the stress regularity of languages with predictable (i.e. non-contrastive) stress is not always as easy to extract from one of the utterance edges as it is in French, the language introduced in section 2.1. For instance, in Hungarian, stress falls on the word-initial syllable. Due to the presence of utterance-initial unstressed function words, though, not all utterances begin with a stressed syllable. In Dupoux & Peperkamp (2002), we established a typology of languages with non-contrastive stress. This typology distinguishes four classes of languages with a phonological stress rule, corresponding to four types of information that are needed to correctly set the Stress Parameter. In languages of Class I, the stress rule can be acquired on the basis of a universal phonetic representation only; in languages of Class II, it can be acquired once language-specific phonological information has been extracted; in languages of Class III and IV, it can be acquired only after all function words and all content words, respectively, can be segmented out of the speech stream.

Given our assumption about the time course of language acquisition in section 2.2, the stress rules of languages belonging to Class I are the first that can be acquired, followed by those of languages in Class II, III, and IV, respectively. This yields the following prediction regarding the cross-linguistic pattern of attested stress 'deafness': 'deafness' in speakers of a language of Class N implies 'deafness' in speakers of all languages belonging to the same or to a lower Class. To see why, suppose a language with a purely phonological stress rule belonging to Class N, the adult speakers of which are stress 'deaf'; we take this as an indication of the absence of stress in the phonological representation of words. Hence, before the Stress Parameter gets set, infants acquiring this language have correctly inferred that stress is non-contrastive and needs not be encoded. In other words, infants have access to the type of information necessary to deduce the stress rule at hand. But then they have also access to the types of information necessary to compute stress rules of other languages belonging to the same or to a lower class. Therefore, infants acquiring any language of Class N or less can deduce that stress is not contrastive before the setting of the Stress Parameter, and, as a consequence, adult speakers of such a language should be stress 'deaf' as well.

There are thus four theoretical possibilities as far as the cross-linguistic perception of stress is concerned, depending on the moment during language development at which the Stress Parameter gets set. First, the Stress Parameter could be set at a late point in development, that is, after much of the lexicon of word forms has been acquired. This possibility corresponds to the idea that the phonological representation encodes all and only those features that are used contrastively in the lexicon (cf. Dupoux *et al*. 1997). We refer to this hypothesis as the Lexical Parameter Setting hypothesis. It predicts that stress 'deafness' should be attested in speakers of languages belonging to any of the four classes, since in none of these classes is stress used contrastively. Second, the Stress Parameter could be set after the acquisition of all other phonological properties of the language as well as of the set of function words, but before the acquisition of a full word form lexicon. This predicts that only languages belonging to Class I-III should yield a 'deafness'. Third, the Stress Parameter could be set after most of the phonology of the language has been acquired, but prior to the acquisition of the function words. This predicts that 'deafness' should be restricted to languages belonging to Class I and II. Finally, the Stress Parameter could be set on the basis of phonetic information only, in which case only languages belonging to Class I should yield a 'deafness'. We globally refer to the last three hypotheses as Non-lexical Parameter Setting hypotheses. They predict the existence of languages with non-contrastive stress whose speakers nonetheless encode stress in the phonological representation.

Table I summarizes the predictions from the four hypotheses, where each column represents a cross-linguistic possibility ('+' stands for the presence of stress 'deafness' and '-' for its absence); the absence of stress 'deafness' in languages with predictable stress is shaded.

Language Class	Lexical Parameter Setting (Dupoux *et al*. 1997)	Non-lexical Parameter Setting		
		Phonetics, phonology, and function words available	Phonetics and phonology available	Phonetics only available
Class I	+	+	+	+
Class II	+	+	+	–
Class III	+	+	–	–
Class IV	+	–	–	–
Contrastive stress	–	–	–	–

Table I

Four alternative hypotheses regarding the presence (+) or absence (-) of stress ‘deafness’ in speakers of languages belonging to Class I-IV or having contrastive stress

Our aim, then, is to experimentally assess the perception of stress by adult speakers of languages belonging to each one of the four classes. If we find that speakers of languages belonging to any of the four classes exhibit stress ‘deafness’, as represented in the first column of Table I, then the Lexical Parameter Setting hypothesis is corroborated. By contrast, if we find one of the three remaining cross-linguistic patterns in Table I, then the corresponding Non-lexical Parameter Setting hypothesis is corroborated. Finally, if we find a cross-linguistic pattern that is not included in Table I, then the whole theoretical framework of Dupoux & Peperkamp (2002) has to be revised.

Before turning to the experiments, we will discuss examples of languages belonging to the four different classes of our typology, and spell out the predictions according to the various Non-lexical Parameter Setting hypotheses. (Recall that according to the Lexical Parameter Setting hypothesis, all languages in the typology should yield stress ‘deafness’.)

3. Exemplification of the typology

3.1. Class I

Examples of languages belonging to Class I are French and Finnish. First, in French, stress falls on the word’s final vowel (Schane 1968). Function words, which are typically unstressed elements, attract stress if they are phrase-final.[3] This is illustrated in (1).

(1) 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’

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 phrase-final function words nor destressing interferes with the observability of the stress rule at utterance endings. That is, all utterances in French have stress on their final vowel.[4] Infants, then, can infer that stress is not contrastive by focussing on universal phonetic cues at utterance endings. In fact, the utterance’s final vowel can be singled out on the basis of the acoustic signal only.

Second, in Finnish, stress is word-initial (Karlsson 1999). This is illustrated in (2).

(2)	a.	usko	[úsko]	‘belief’
	b.	uskollisuus	[úskolːisuːs]	‘fidelity; loyalty’

There are no unstressed function words that appear phrase-initially. Moreover, monosyllabic content words are not destressed when they are involved in a stress clash.[5] All utterances therefore have stress on the first vowel, and infants can extract the stress rule by focussing their attention on universal phonetic cues at utterance beginnings.

Given that the French and Finnish stress rules can be deduced without having access to any language-specific cues, French- and Finnish-acquiring infants can deduce the stress rule of their language before the Stress Parameter needs to be set. All three Non-lexical Parameter Setting hypotheses therefore predict that adult speakers of these languages exhibit stress ‘deafness’.

3.2. Class II

An example of a language belonging to Class II is 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 light. There are no phrase-final unstressed function words. Examples from Dixon’s (1988) description of the Boumaa dialect are given in (3); syllable boundaries are indicated by dots.

(3)	a.	kám.ba	kam.bá.ta	‘climb - climb it’
	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 syllable structure and the distinction between light and heavy syllables are available to infants at the time they set the Stress Parameter. Thus, among the Non-lexical Parameter Setting hypotheses, only the last one predicts that speakers of Fijian are not stress ‘deaf’; according to this hypothesis, the Stress Parameter is set before the acquisition of both other phonological properties and the set of function words. The remaining hypotheses predict that speakers of Fijian are stress ‘deaf’.

3.3. Class III

In languages belonging to Class III, stress is predictable and observable *modulo* the occurrence of function words. In fact, contrary to the situation in French described in section 3.1, function words are typically unstressed, regardless of their position in the utterance. Consider, for instance, the case of Hungarian. In this language, stress falls on word-initial syllables, and function words are systematically unstressed (Vago 1980). This is illustrated in (4).

(4)	a.	emberek	[émberek]	‘men’
	b.	az emberek	[azémberek]	‘the men’

In Hungarian, then, utterances that begin with a function word 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, since utterance-initial words are never destressed (Vogel 1988).

In order for 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. If infants have acquired the set of function words prior to setting the Stress Parameter, they should thus exhibit stress ‘deafness’ as adults. Among the Non-lexical Parameter Setting hypotheses, then, only the first one predicts that speakers of Hungarian are stress ‘deaf’; according to this hypothesis, in fact, the set of function words is available by the time the Stress Parameter gets set. The remaining hypotheses predict that speakers of Hungarian are not stress ‘deaf’.

3.4. Class IV

Languages in Class IV have a stress rule that is observable only if the boundaries of content words are available. Consider, for instance, Polish, a language in which word stress is on the penultimate syllable (Comrie 1967; Hayes 1995). Some examples are given in (5).

(5)	a.	gázet - gazéta - gazetámi	‘newspaper $_{\text{GEN-PL; NOM-SG; INST-PL}}$’
	b.	jézyk – jezýka - jezykámi	‘language $_{\text{NOM-SG; GEN-SG; INST-PL}}$’

Polish has many monosyllabic content words. If a monosyllabic word is followed by another monosyllabic word or by a disyllabic word, a stress clash arises, which is resolved by destressing of the first word (Rubach & Booij 1985). Destressing therefore does not interfere with the generalization concerning surface stress patterns at utterance edges, which can be formulated as follows: in utterances that end in a monosyllabic word, the final syllable is stressed, whereas in all other utterances, the penultimate syllable is stressed.[6] In order to extract the rule regarding penultimate stress, infants should have access to content word boundaries. Stress ‘deafness’ in adults, then, depends upon the availability of full word segmentation by the time the Stress Parameter gets set. Hence, none of the Non-lexical Parameter Setting hypothesis predicts a stress ‘deafness’ in speakers of Polish. Rather, such a ‘deafness’ is predicted only by the Lexical Parameter Setting hypothesis.

4. Experiments assessing the perception of stress

In the remaining part of this paper, we report on experiments assessing the perception of stress by adult native speakers of French and Finnish (Class I), Hungarian (Class III), and Polish (Class IV).[7] As a control language, we use Spanish, which does not belong to any of the four classes since it has contrastive stress. We predict that native speakers of Spanish will not exhibit stress ‘deafness’.

Before turning to the experiments, however, we should define what is meant exactly by studying the perception of stress cross-linguistically. Recall from section 2.2 that the acoustic correlates of stress, i.e. duration, pitch and energy, are not used to an equal extent in all languages to realize stress. In this study, we manipulate all three stress cues, in order to create a maximally perceptible contrast that contains valid stress cues in all languages under consideration. Using these very different stimuli ensures that whenever a stress contrast is not perceived in a given language, the ‘deafness’ effect is not due to the confusability of the stimuli. One caveat is in order, though: it is important to make sure that we do not manipulate variables that could be perceived as something other than stress in a given language. For instance, in languages with contrastive vowel length, such as Finnish and Hungarian, duration is avoided as a correlate of stress (Hayes 1995). Therefore, when processing a foreign language with duration as a 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. This is why, given that our sample of languages include Finnish and Hungarian, we take special care to insure that the durational cues of our stimuli do not yield the perception of a lexical vowel length contrast.

4.1. General method

Dupoux, Peperkamp & Sebastián-Gallés (2001) present a novel paradigm for assessing the perception of stress, based on a short term memory task. In this paradigm, the recall performance of a stress contrast is compared with that of a control phonemic contrast, across different levels of memory load. The experiment is divided into two parts. In each part, subjects are required to learn two CVCV non-words that are a minimal pair differing only in one phonological dimension, i.e. place of articulation of the second consonant or location of stress. In each part, subjects are taught to associate the two non-words to the keys [1] and [2], respectively, of a computer keyboard. After some training with an identification task, subjects listen to longer and longer random sequences of the two non-words, which they are required to recall and transcribe as sequences of [1] and [2]. Within the sequences, the non-words are randomly instantiated by one of 6 acoustically different tokens. The length of the sequences varies from 2 to 6, and is augmented by one after 8 trials. Hence, each part of the experiment contains 40 (5 x 8) trials. The segmental contrast in the first part is phonemic in all languages under consideration and hence equally easy for all subjects; this contrast is thus used to establish baseline performance.

In order to diminish the likelihood that subjects use explicit recoding strategies, the stimuli are short and the tokens in the sequences are separated from one another by a very short interval, i.e. 80 ms. Moreover, in order to prevent subjects from using echoic memory, every sequence is followed by the word ‘OK’. On average, the experiment lasts about 20 minutes. Responses that are a 100% correct transcription of the input sequence are coded as correct; all other responses are coded as incorrect, regardless of the number of tokens within the sequence that are transcribed incorrectly. Among the incorrect responses, those that are a 100% incorrect transcription - i.e. with each token of the sequence labeled incorrectly - are coded as reversals. Subjects with more reversals than correct responses in either the phoneme or the stress condition are rejected, the high percentage of reversals suggesting that they might have confused the number key associated to the first item with the one associated to the second item.

In Dupoux *et al*. (2001), we used the novel paradigm to test the perception of stress by speakers of French and Spanish. In a previous experiment, using a different paradigm, Dupoux *et al*. (1997) found that speakers of French, but not those of Spanish, exhibit stress ‘deafness’, i.e. they had much more difficulties with the stress contrast than with the phonemic contrast. These results were confirmed with the novel paradigm. That is, in Experiment 3 of Dupoux *et al*. (2001), we compared the phonemic contrast [kúpi - kúti] with the stress contrast [mípa - mipá], and found that the French subjects made significantly more errors with the latter ($F(1,11) = 71.0$; $p < .0001$), whereas the Spanish subjects showed a nonsignificant trend in the other direction ($F(1,11) = 3.7$; $.1 > p > .05$). The interaction between the factors language and contrast was highly significant ($F(1,22) = 70.3$; $p < .0001$).

4.2. Experiment 1: Finnish

In the first of the two experiments to be presented in this paper, we used the same materials and the same paradigm as Dupoux *et al*. (2001) to test 12 native speakers of Finnish, a language with contrastive vowel length. As in French and Spanish, all the items are composed of existing phonemes in Finnish that appear in a combination in accordance with the phonotactics of the language. Hence, apart from the location of stress, they are possible but non-existing words.

The recordings, used previously for the French and Spanish subjects, were made by a female trained phonetician whose native language is Dutch. The mean durations of the tokens used for the phonemic and the stress contrast were 439 ms, and 513 ms, respectively. In order to introduce more phonetic variation among the tokens, we manipulated the global pitch with a waveform editor. Specifically, the pitch contours of the 6 tokens of each item were multiplied with the values 95, 97, 99, 101, 103, and 105, respectively (see Dupoux *et al*. (2001) for discussion). Figure 1 displays the pitch contours of the 6 tokens of each item used for the stress contrast before the global pitch modifications were realized, as well as the mean duration and intensity of their segments. As can be seen, the pitch contours of the [mípa] tokens and those of the [mipá] tokens are highly dissimilar. Indeed, all tokens of [mípa] show a drop in pitch between the first and second syllable, the magnitude of which is between 7 and 9 semitones, whereas all tokens of [mipá] show an increase in pitch between the first and second syllable of a magnitude between 4 and 5 semitones. There are also significant differences in the mean duration of segments in stressed versus unstressed syllables. This is seen most prominently in the vowels. Indeed, across the two sets of tokens, stressed vowels are on average 21 ms longer than unstressed vowels, a numerically small but significant difference given the small between-token variance ($F(1,5) = 43.2$; $p <.001$). Finally, stressed vowels are on average 3.7 dB louder than unstressed vowels, again, a numerically small but significant difference ($F(1,5) = 78.2$; $p < .001$).[8]

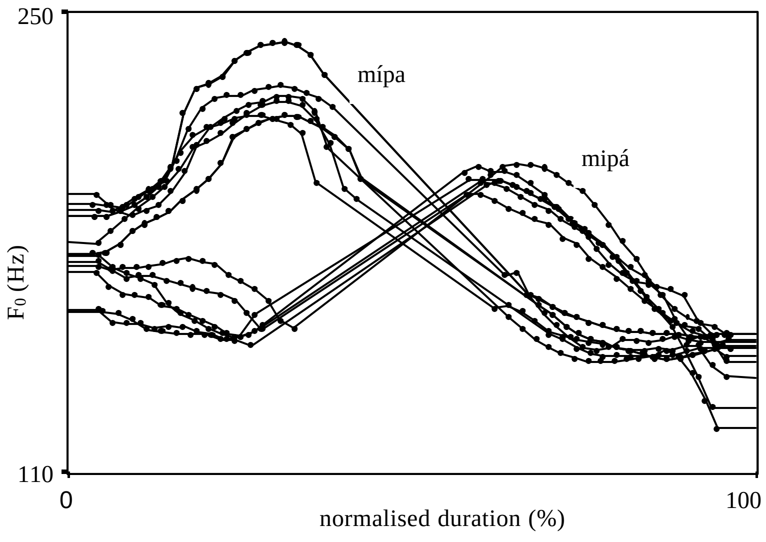
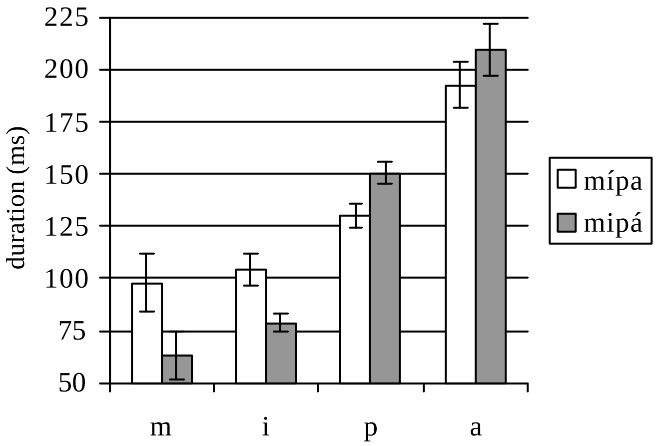
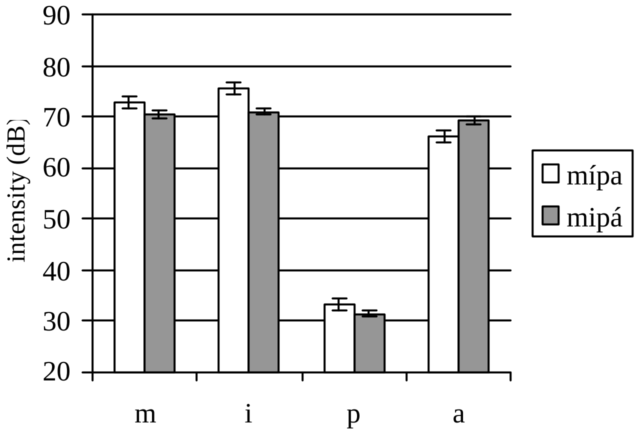


Figure 1: Acoustic measurements of 6 tokens [mípa] and 6 tokens [mipá] used in Experiment 1

a. F_0 as a function of normalised duration



b. Duration of segments



c. Intensity of segments

In an informal pilot with a native speaker of Finnish, we found that the stressed vowels in [mípa] and [mipá] were not perceived as long vowels. We thus decided to run the experiment with the 12 Finnish subjects using the same recordings of the four non-words. The results were that analogously to the French subjects in Dupoux *et al*., the Finnish subjects made significantly more errors with the stress contrast than with the phonemic contrast ($F(1,11) = 15.9$; $p < .003$). Comparing the Finnish subjects with the Spanish subjects of Dupoux *et al*. (2001) in a post-hoc analysis of variance, we found a significant interaction between language and contrast ($F(1,22) = 19.5$; $p < .0001$). This interaction was due to the fact that there was an effect of contrast for the Finnish but not for the Spanish subjects. In a comparison of the Finnish subjects with the French subjects of Dupoux *et al*., the interaction between language and contrast was marginally significant ($F(1,22) = 3.6$; $p < .074$). The effect of contrast for the Finnish subjects was indeed smaller than that for the French subjects.

A summary of the results is shown in Figure 2.

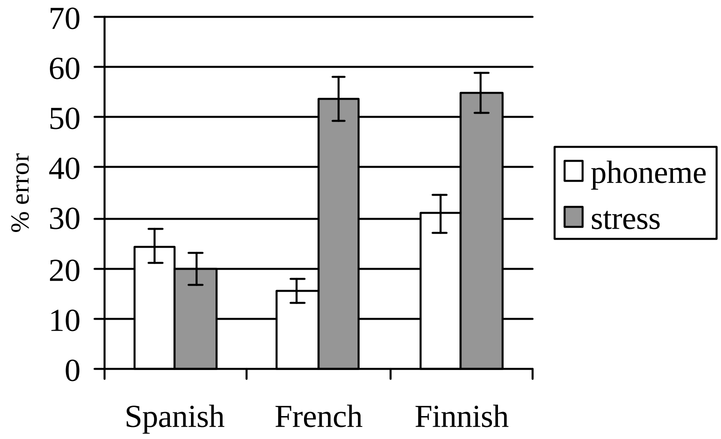


Figure 2: Percent recall error as a function of contrast for 12 Spanish, 12 French, and 12 Finnish subjects

The Finnish results replicate the results with the French subjects in Dupoux *et al*. (2001), thus corroborating the hypothesis that Class I languages yield stress 'deafness'. This is an interesting result in itself, since it shows that stress 'deafness' is not limited to a single language, and, moreover, that it is independent of the position of word stress in the language. Recall, in fact, that stress is word-final in French and word-initial in Finnish. If we are correct in positing that infants acquire the stress regularity of their language by focusing on utterance edges, these results are evidence that infants can focus equally well on utterance endings and on utterance beginnings.[9]

To sum up, native speakers of French and Finnish, as opposed to native speakers of Spanish, exhibit stress 'deafness'. The 'deafness' effect appears to be somewhat larger in the French than in the Finnish subjects; this difference, however, is not significant. We need to test more subjects to verify whether the interaction between language and contrast reaches significance. If this turns out to be the case, then a possible explanation might be related to the fact that in Finnish but not in French, vowel length is contrastive. That is, despite our precautions, some Finnish subjects might encode stressed vowels as long vowels, and hence rely on the lexically distinctive property of vowel length to do the task.

4.3. Experiment 2: Hungarian and Polish

In our second experiment, we tested the perception of stress by native speakers of Hungarian and Polish, languages belonging to Class III and Class IV, respectively. Given that [kúti] is a real word in Hungarian, we could not use the pair [kúpi - kúti] of the previous experiments. We therefore chose new items, which are possible but non-existing words not only in Hungarian and Polish, but also in Finnish, Fijian, and our control language Spanish, allowing us to use the new materials in future experiments.

For the phonemic contrast, we used the pair [níka - níta]. As in the previous experiment, all segments making up the items are phonemic in the languages at hand, and appear in a phonotactically legal combination. There are two reasons why the pair [níka - níta] is comparable to the pair [kúpi - kúti] used in Experiment 1 as far as perceptual difficulty is concerned. First, the consonants involved in the two contrasts, i.e. [p], [t], and [k], are all unvoiced stops; the place of articulation of the consonant present in both pairs, i.e. [t], lies between that of the consonants with which it forms a contrast, i.e. [p] and [k], respectively. Moreover, the consonantal contrasts [p-t] and [t-k] are known to be equally difficult to perceive (Miller & Nicely 1955). Second, the consonants are embedded within comparable vocalic contexts, consisting of two different cardinal vowels each that differ on either the front/back dimension ([kúpi - kúti]) or on the height dimension ([níka - níta]). Specifically, in both pairs, one of the consonants shares its place of articulation with the following vowel and the other one shares its place of articulation with the preceding vowel. Thus, in [kúpi – kúti], [p] and the preceding [u] are both labial, while [t] the following [i] are both coronal; and in [níka - níta], [t] and the preceding [i] are both coronal, while [k] and the following [a] are both dorsal. Therefore, there is no reason to believe that coarticulation will make one of the contrasts more difficult than the other.

As to the stress contrast, we constructed 10 minimal pairs, in order to show that the 'deafness' effect is not limited to a single non-word. These pairs are shown in (6).

(6)	[kánu - kanú]	[númi - numí]
	[míku - mikú]	[támi - tamí]
	[nátu - natú]	[támu - tamú]
	[nímu - nimú]	[tímu - timú]
	[núma - numá]	[túka - tuká]

All items were recorded 6 times by the same speaker that recorded the materials for the experiments reported in the previous section. The mean duration of the tokens for the phonemic contrast was 444 ms. The mean duration of the tokens for the stress contrasts was 425 ms. In this experiment, we wanted to reduce the importance of duration as a stress cue, in order to prevent Hungarian subjects from relying on the lexical difference between short and long vowels. The speaker therefore took special care to reduce the durational differences with respect to that in the materials used in Experiment 1. As in Experiment 1, we introduced more phonetic variation among the tokens, by multiplying the pitch contours of the 6 tokens of each item with the values 95, 97, 99, 101, 103, and 105, respectively.

We carried out several acoustical analyses of the stress tokens. Due to the large number of tokens, we restricted these analyses to only one token per item. For each of the 20 items, we randomly chose one of the 6 tokens for the analyses. Figure 3 displays the pitch contours of these tokens before the global pitch modifications were realized, as well as the mean duration and the intensity of their segments.

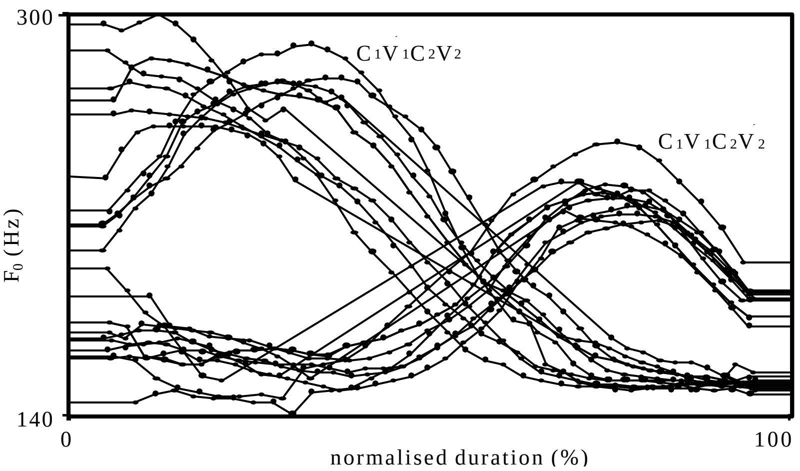
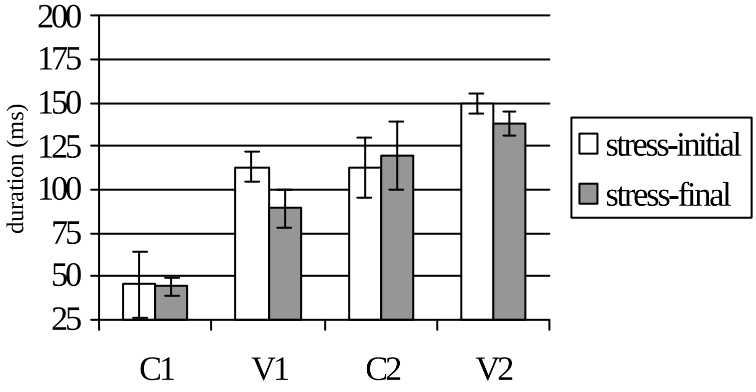
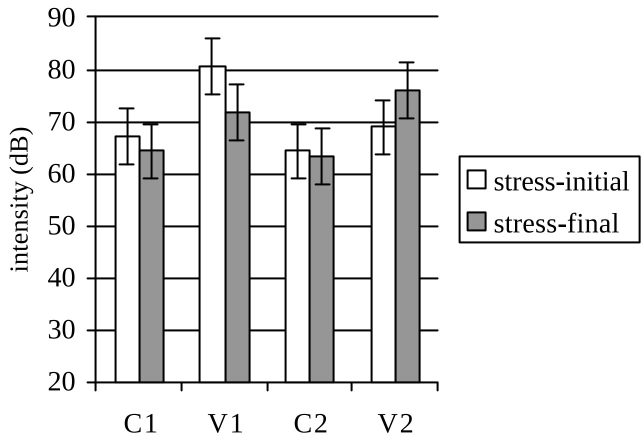


Figure 3: Acoustic measurements of 10 items $C_1\acute{V}_1C_2V_2$ (one token per item) and 10 items $C_1V_1C_2\acute{V}_2$ (one token per item) used in Experiment 2

a. F_0 as a function of normalised duration



b. Duration of segments



c. Intensity of segments

The durational difference between stressed and unstressed vowels is 6.5 ms ($F(1,9) = 5.7$; $p < .04$), which is smaller than the corresponding difference in the stimuli in Experiment 1. Duration, then, is a less important stress cue than in the previous experiments, as desired. As to the pitch contours of the stress-initial tokens and those of the stress-final tokens, they are again highly dissimilar, and even more so than in Experiment 1. Indeed, all stress-first tokens show a drop in pitch between the first and the second syllable, the magnitude of which is between 9 and 11 semitones; all stress-second tokens show an increase in pitch between the first and the second syllable, the magnitude of which is between 6 and 8 semitones. Stressed vowels are on average 8 dB louder than unstressed vowels ($F(1,9) = 116.5$; $p < .001$), again a greater difference than in the previous experiment.[10] Hence, in this new recording, the reduction of duration as a stress cue appears to be compensated by an increase of pitch and energy as stress cues.

As to the experimental paradigm, we made one change with respect to Experiment 1. That is, due to the fact that the experiment requires a lot of concentration and is quite long (on average 20 minutes), we shortened it by taking out the sequences of length 3 and 5. Hence, we used only sequences of length 2, 4, and 6. An analysis of the results obtained in Dupoux *et al*. (2001) shows that using only these sequence lengths does not impair the power of the paradigm.

We tested 10 native speakers of Hungarian and 10 native speakers of Polish. Each subject was tested on the phonemic contrast and on one of the stress contrasts. One Hungarian and one Polish subject were replaced, due to too many reversals among their responses with the stress contrast and with the phonemic contrast, respectively. We found that the Hungarian subjects exhibited stress 'deafness'; that is, they made significantly more errors with the stress contrast than with the phonemic contrast ($F(1,9) = 37,4$; $p < .0001$). The Polish subjects also made more errors with the stress contrast, but the effect was only marginally significant ($F(1,9) = 4.5$; $p < .056$). The interaction between language and contrast was also marginally significant ($F(1,18) = 3.4$; $p < .084$).

The results are summarized in Figure 4.

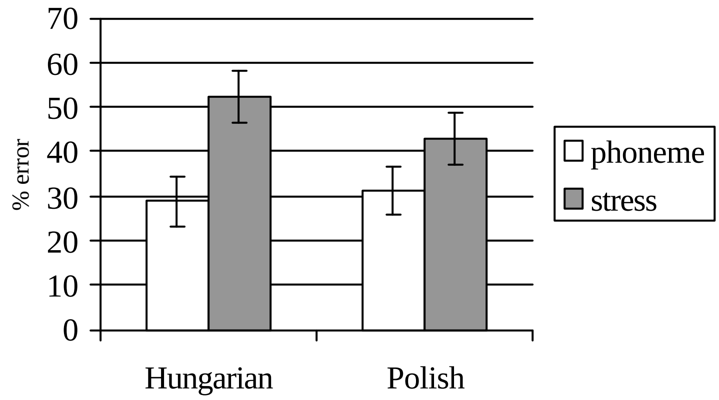


Figure 4: Percent recall error as a function of contrast for 10 Hungarian and 10 Polish subjects

These results show that Hungarian subjects are stress 'deaf'. Polish subjects, by contrast, are not stress 'deaf'; indeed, they did not make significantly more errors with the stress contrast than with the phoneme contrast. In Polish, vowel length is not contrastive but is used as a stress cue. Consequently, for some Polish subjects the materials in set 2 might have been ambiguous as far as stress is concerned, the vowel length difference between stress-initial and stress-final items not being sufficiently large. This could explain why we found a tendency towards stress 'deafness' in the Polish subjects.

Before closing this section, some additional remarks concerning Polish are in order. It is well-known that the Polish stress rule has some lexical exceptions, which are borrowings of mainly Greek and Latin origin, with either antepenultimate (7a,b) or final stress (7c).

(7)	a.	muzyka	[múzika]	'music'
	b.	uniwersytet	[univɛ́rsitɛt]	'university'
	c.	menu	[menú]	'menu'

We assume that the existence of several lexical exceptions does not interfere with the classification of Polish as having a regular phonological stress rule. In particular, due to their small number as well as their low frequency of occurrence, infants are likely not to take lexical exceptions into account while deducing the stress regularity of their language, if they are exposed to them at all.

We take the presence of lexical exceptions as an indication that speakers of Polish cannot be stress 'deaf'. Indeed, if they were, they would not perceive lexical exceptions as being deviant and hence they would not be able to produce them with the exceptional stress pattern. Consequently, foreign loans should be completely regularized to the native stress pattern.[11] This is indeed what happens in, for instance, French, which does not have any lexical exceptions and whose speakers typically do not recall where stress falls in foreign words. Based on a large sample of languages, Peperkamp (submitted) argues that languages with non-contrastive stress present lexical exceptions if and only if they belong to Class IV, suggesting that speakers of languages belonging to Class I-III but not those of languages belonging to Class IV exhibit stress 'deafness'.

4.4 General discussion

Using a paradigm that we initially set up to study the perception of stress in French and Spanish, we presented experiments in three new languages that exemplify different levels of our linguistic typology. In order to get a clearer idea of the overall pattern of results, we need a direct comparison of the results obtained with the French, Finnish and Spanish subjects on the one hand and those obtained with the Hungarian and Polish subjects on the other hand. Before conducting such a comparison, however, an important caveat has to be raised. Indeed, in these two sets of languages, we used different sets of stimuli for both the phonemic and the stress contrast. Given our constraint on the non-word status of the items as well as the differences in the use of vowel length in the various languages, it is extremely hard to construct materials that can be used in all five languages. The materials of set 1 contained stressed vowels that were longer than those of the materials in set 2. Speakers of languages without contrastive vowel length use duration as a stress cue and are thus best tested with set 1. By contrast, speakers of languages with contrastive vowel length are best tested with set 2, since the materials in set 1 might induce them to perceive a lexical vowel length contrast. In our experiments, Finnish and Polish were tested with materials from the non-optimal set. That is, Finnish subjects, who use vowel length contrastively, were tested with set 1, while Polish speakers, who do not use vowel length contrastively, were tested with set 2. Stress 'deafness', therefore, might be underestimated in Finnish subjects and overestimated in Polish subjects.

This being said, in order to compare the different languages, we define a stress 'deafness' index for a population as the mean percentage of errors made with the stress contrast minus the mean percentage of errors made with the phonemic contrast. For each language in our sample, Table II displays the stress 'deafness' index, together with information concerning the class to which the language belongs, the status of vowel length in the language, and the stimuli that were used in the experiment.

language	stress 'deafness' index	Class	vowel length	stimuli
French	38.1	I	non-contrastive	set 1
Finnish	24.0	I	contrastive	set 1
Hungarian	23.7	III	contrastive	set 2
Polish	11.6	IV	non-contrastive	set 2
Spanish	- 4.4	control	non-contrastive	set 1

Table II

Stress 'deafness' index and information regarding class membership, status of vowel length, and stimulus set for five languages tested in Experiment 3 of Dupoux *et al*. (2001) and in Experiments 1 and 2 of the present paper

According to the stress 'deafness' index, speakers of French exhibit the strongest effect, followed by speakers of Finnish, Hungarian, and Polish, respectively. Interestingly, the gradual nature of the 'deafness' effect goes in the direction of our language typology, in that the strongest 'deafness' effect is found in a Class I language, i.e. French, and the weakest 'deafness' effect is found in a Class IV language, i.e. Polish. The intermediate languages, Finnish and Hungarian, belong to Class I and Class III, respectively. It appears, then, that the size of the 'deafness' effect correlates with the ease with which the stress regularity can be acquired by infants. Note that the numerical difference between Finnish and French might be due to the underestimation of 'deafness' in Finnish that we discussed earlier.

We also ran a series of t-tests for the 'deafness' index computed for each individual subject across the different languages. The t-tests are corrected for multiple comparisons, using a Bonferroni corrected p-value of .005. We found that French yielded a significantly higher 'deafness' score than all other languages except Finnish. Similarly, Spanish had a significantly lower score than all other languages except Polish. No other comparison reached significance. Considering French and Spanish as two anchor points, we thus find that Finnish patterns with the former and Polish patterns with the latter; Hungarian is situated in the middle and differs significantly from both anchors.

These results allow us to discard the Lexical Parameter Setting hypothesis, according to which all languages with non-contrastive stress should yield an equal amount of stress 'deafness'. This hypothesis is falsified, given that both Hungarian and Polish yield significantly less stress 'deafness' than French and that Polish does not differ significantly from Spanish. Regarding the Non-lexical Parameter Setting hypothesis, recall that it is actually a grouping of three hypotheses, that vary in the amount of non-lexical information that they state is available for the setting of the Stress Parameter. Given that we did not test speakers of a language belonging to Class II, we cannot distinguish the hypothesis according to which only phonetic information is available from that according to which additional phonological information is available. Both correctly predict that Polish (Class IV) does not differ from the control language Spanish, and that Hungarian (Class III) differs from French (Class I). However, these hypotheses cannot account for the fact that Hungarian also differs from Spanish; according to these hypotheses, in fact, speakers of Hungarian, similarly to those of Spanish, should have no problem with the perception of stress contrasts. The third Non-lexical Parameter Setting hypothesis states that in addition to phonetic and phonological information, the set of function words is available before the Stress Parameter needs to be set. Analogously to the first two, this hypothesis also correctly predicts the pattern of results regarding Polish and fails to account for Hungarian, although for a different reason. That is, under this hypothesis, the difference between Hungarian and Spanish is accounted for, while the difference between Hungarian and French is not.

Due to the intermediate status of Hungarian, our pattern of results is thus intermediate between that predicted by the last non-lexical hypothesis discussed here and that predicted by the first two. Speculatively, we would like to offer two possible explanations for this fact. First, the Stress Parameter might not be binary. It has indeed been proposed that the initial tuning of the phonological representation of words is gradual rather than discrete. For instance, Jusczyk (1993) argues that early exposure to a language results in a deformation of the acoustic space, with relevant dimensions being blown up and irrelevant dimensions being squeezed down. Second, the Stress Parameter might be binary but set according to a statistical criterion, resulting in individual variation. Consequently, in cases where the stress regularity is fairly complex to extract, as in Hungarian, only a certain proportion of infants correctly detects the presence of the regularity and hence becomes stress 'deaf'; the remaining infants would simply fail to detect the regularity and hence retain stress in the phonological representation. We then predict the presence of a bimodal distribution of individual responses, a prediction which can be put to test if a larger number of subjects is used.

5. Conclusion

The present paper is a first step towards a more comprehensive cross-linguistic examination of the perception of suprasegmental information. The results can be accounted for in terms of our acquisition framework; specifically, they support the notion that decisions regarding the format of the phonological representation of words are made before much of a lexicon of word forms is available. Several questions, though, remain open, and future research can be carried out along various directions.

First, we need to replicate our findings in the languages that we have tested so far, and test the generality of our results by independently manipulating duration and pitch of stressed vowels in the stimuli. We also need to test a language belonging to Class II to complete our survey. Furthermore, in order to show that the location of word stress *per se* does not interfere with the presence of stress 'deafness', more cross-linguistic experiments should be carried out with languages that share the location of word stress but belong to different classes of our typology. In particular, we tested two languages with initial stress (i.e. Finnish and Hungarian), but only one language with final stress (i.e. French) and only one language with penultimate stress (i.e. Polish). Given the fact that Polish is the only language in which stress does not fall at one of the word edges, we would especially like to test another language with prefinal stress but belonging to a different class.

Second, it would be interesting to show that the stress 'deafness' effect found with our task truly reflects a property of the phonological representation of words. One way to test this would be to explore with a lexically pertinent task how speakers of a language with stress 'deafness' represent words in a foreign language with contrastive stress that they have learned as a second language. We predict that such speakers fail to build two separate lexical entries for words forming a minimal stress pair.

Third, we might test more directly the hypothesis that infants can set the Stress Parameter by focusing on utterance boundaries. A first step in this direction would be to perform acoustical analysis on natural utterances in the various languages of our sample, in order to determine whether word stress can indeed be observed above and beyond sentence level suprasegmental phenomena. A second step would be to directly probe the age at which infants of languages for which we have found a 'deafness' become less sensitive to stress.

Finally, our framework can be extended to include languages in which the segmental phonology can help to distinguish between stressed and unstressed syllables. For instance, many languages have a rule that reduces certain unstressed vowels; the presence of a reduced vowel, then, is an indication of the absence of stress (Hayes 1995). Cutler (1986) and Cutler and Van Donselaar (2001) found contrastive results in English and Dutch, two languages with vowel reduction. Cutler (1986) used minimal stress pairs in English in which vowel reduction exceptionally does not apply, such as *fórbear – forebéar*. She found that speakers of English do not process stress on-line in order to identify lexical items, and argued that this is due to the fact that English has only a very small number of true minimal stress pairs (that is, pairs of words in which vowel reduction does not apply). Cutler & Van Donselaar (2001) similarly used true minimal stress pairs in Dutch, a language in which the number of such pairs is equally small. They found that speakers of Dutch, contrary to those of English, do use stress on-line in order to identify lexical items. Their interpretation of this finding relies on the observation that English and Dutch differ in the amount of vowel reduction that applies; whereas unstressed syllables in English almost invariably have a reduced vowel, Dutch has much less vowel reduction. Cutler & Van Donselaar, then, raised the hypothesis that a given contrast is not used for lexical access if and only if its information value is relatively low. That is, speakers of English, but not those of Dutch, can rely on the parasitic, segmental, cue of vowel reduction in order to identify stressed syllables. This hypothesis, though, cannot account for the present pattern of results, since in the languages with non-contrastive stress that we tested, the information value of the stress cues is zero.

To conclude, we hope to have convinced the reader that studying models of early language acquisition in a cross-linguistic perspective can open interesting and novel issues in speech perception research. These issues can be tested in adults, and, ultimately, in infants.

Acknowledgments

Research for this paper was funded by grants from the French Ministry of Education, Research, and Technology (Groupement d'Intérêt Scientifique 'Sciences de la Cognition', and Action Concertée Incitative 'Cognitique'), as well as by a grant from the Centre National de la Recherche Scientifique ('Aide à Projet Nouveau'). We are grateful to the Finnish, Hungarian, and Polish cultural institutes in Paris, who provided us with testing facilities. We would like to thank Katherine White for help in running the subjects and Inga Vendelin for practical assistance. Thanks are also due to Anne Christophe, François Dell, Jacques Mehler, Marina Nespor, Christophe Pallier, Franck Ramus, and two anonymous reviewers for comments and discussion.

Notes

[1] The idea that infants can acquire phonological regularities of their language by focusing on utterance edges was first explored in Bourgeois (1991).

[2] By contrast, rhythmic readjustments in stress clash configurations potentially interfere with the transparency of stress at utterance edges. We test only languages in which this is not the case (cf. section 3).

[3] Exceptions are *je* 'I' and *ce* 'it'. These words 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'.

[4] In southern varieties of French, utterance-final words can end in an unstressed schwa. In these varieties, the acquisition of the stress regularity is more complex, in that infants first have to acquire the difference between full vowels and schwa. Consequently, these varieties belong to Class II rather than to Class I.

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

[5] This observation is based on judgments of our own recordings of four female native speakers.

[6] In the literary language, there are three enclitics that fall outside the stress domain, i.e. *by*, *śmy*, and *ście*. With these enclitics, then, antepenultimate stress is yielded. Examples are *róbi-łby* 'he would', *robílismy* 'we did', and *robíliscie* 'you$_{PL}$ did'. The latter two, however, can also surface with penultimate stress, and this variant is the most frequent one (Booij & Rubach 1987). We assume that infant-directed speech does not contain sequences of host plus enclitic with antepenultimate stress.

[7] Fijian, the Class II language in our sample, presents two problems. First, the official language of Fiji being English, it is hard to find native speakers who were raised in a monolingual environment. Second, there is much dialectal variation, and it is still unclear whether this variation concerns the stress pattern.

[8] The root mean square of the signal in the portion of speech corresponding to each segment was measured and converted to an intensity value in dB by a logarithmic transform (this was done with the Praat software). The dB-value for each vowel was then entered into a two-way ANOVA with stress and vowel position as main factors.

[9] Aslin, Woodward, LaMendola & Bever (1996) show that when American and Turkish mothers teach their 12-month-old infants new words, they tend to put these words in utterance-final position. This suggests that infants might be more focused on utterance endings that on beginnings, whence the non-triviality of the present finding.

[10] See note 8.

[11] Our subjects informed us that lexical exceptions are explicitly taught as such to school children. Moreover, they are sometimes regularized, either by movement of the stress, or – in the case of antepenultimate stress – by deletion of the vowel in the penultimate syllable. This regularization, however, is not generalized.

References

Aslin, R., Woodward, J., LaMendola, N & Bever, T. (1996) Models of word segmentation in fluent maternal speech to infants. In *Signal to Syntax. Bootstrapping from Speech to Grammar in Early Acquisition* (J. Morgan & K. Demuth, editors), pp. 117-134. Mahwah, N.J.: LEA.

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

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

Bluhme, S. & Burr, R. (1971) An audio-visual display of pitch for teaching Chinese tones. *Studies in Linguistics*, **22**, 51-57.

Booij, G. & Rubach, J. (1987) Postcyclic versus postlexical rules in lexical phonology, *Linguistic Inquiry*, **18**, 1-44.

Bourgeois, T. (1991) *Instantiative Phonology*. Doctoral dissertation, University of Arizona.

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

Church, K. (1987) Phonological parsing and lexical retrieval, *Cognition*, **25**, 53-69.

Comrie, B. (1967) Irregular stress in Polish and Macedonian, *International Review of Slavic Linguistics*, **1**, 227-240.

Cutler, A. (1986) Forbear is a homophone: Lexical prosody does not constrain lexical access, *Language and Speech*, **29**, 201-220.

Cutler, A. & Van Donselaar, W. (2001) Voornaam is not a homophone: Lexical prosody and lexical access in Dutch. Language and Speech, **44**, 171-195.

Dell, F. (1984) L'accentuation dans les phrases en français. In *Forme sonore du language* (F. Dell, D. Hirst & J.-R. Vergnaud, editors), pp. 65-122. Paris: Hermann.

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

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C. & Mehler, J. (1999) Epenthetic vowels in Japanese: a perceptual illusion?, *Journal of Experimental Psychology: Human Perception and Performance*, **25**, 1568-1578.

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

Dupoux, E. & Peperkamp, S. (2002) Fossil markers of language development: phonological 'deafnesses' in adult speech processing. In *Phonetics, Phonology, and Cognition* (J. Durand & B. Laks, editors). Oxford : Oxford University Press, 168-190.

Dupoux, E., Peperkamp, S. & Sebastián-Gallés, N. (2001) A robust paradigm to study stress 'deafness'. *Journal of the Acoustical Society of America*, **110**, 1606-1618.

Frazier, L. (1987) Structure in auditory word recognition, *Cognition*, **25**, 157-187.

Gandour, J. (1983) Tone perception in Far Eastern languages, *Journal of Phonetics*, **11**, 149-175.

Gerken, L.A., Jusczyk, P. & Mandel, D. (1994) When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition*, **51**, 237-265.

Goldinger, S. (1998) Echoes of echoes? An episodic theory of lexical access, *Psychological Review*, **105**, 251-279.

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

Grammont, M. (1965) *Traité de phonétique*. Paris: Delagrave.

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

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

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

Jusczyk, P. (1993) From general to language specific capacities: The WRAPSA model of how speech perception develops, *Journal of Phonetics*, **21**, 3-28.

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

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

Karlsson, F. (1999) *Finnish: An Essential Grammar*. London: Routledge.

Kiriloff, C. (1969) On the auditory perception of tones in Mandarin, *Phonetica*, **20**, 63-67.

Klatt, D. (1980) Speech perception: A model of acoustic-phonetic analysis and lexical access, *Journal of Phonetics*, **7**, 279-312.

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

Lee, L. & Nusbaum, H. (1993) Processing interactions between segmental and suprasegmental information in native speakers of English and Mandarin Chinese, *Perception and Psychophysics*, **53**, 157-165.

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

Marslen-Wilson, W. & Warren, P. (1994) Levels of perceptual representation and process in lexical access: words, phonemes, and features, *Psychological Review*, **101**, 653-675.

Mehler, J., Bertoncini, J., Dupoux, E. & Pallier, C. (1996) The role of suprasegmentals in speech perception and acquisition. In *Phonological Structure and Language Processing. Cross-linguistic Studies* (T. Otake & A. Cutler, editors), pp. 145-169. Berlin: Mouton de Gruyter.

Mehler, J., Dupoux, E. &.Segui, J. (1990) Constraining models of lexical access: The onset of word recognition. In *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives* (G. Altmann, editor), pp. 236-262. Cambridge, MA: MIT Press.

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

Miller, G. & Nicely, P. (1955) An analysis of perceptual confusions among some English consonants, *Journal of the Acoustical Society of America*, **27**, 338-352.

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

Navarro Tomás, T. (1965) *Manual de prononciación española*. Madrid: Consejo Superior de Investigaciones Científicas.

Nazzi, T., Bertoncini, J. & Mehler, J. (1998) Language discrimination by newborns: Towards an understanding of the role of rhythm, *Journal of Experimental Psychology: Human Perception and Performance*, **24**, 1-11.

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

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

Peperkamp, S. (2000) *Two typological gaps in stress systems: arguments from early language acquisition*. Handout of a talk presented at the Fourth Utrecht Biannual Phonology Workshop, Utrecht, The Netherlands.

Peperkamp, S. (submitted) Lexical exceptions in stress systems: Arguments from early language acquisition and adult speech processing.

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), pp. 223-237. The Hague: Mouton.

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

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

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

Rubach, J. & Booij, G. (1985) A grid theory of stress in Polish, *Lingua*, **66**, 281-319.

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

Shady, M. (1996) *Infants' sensitivity to function morphemes*. Doctoral dissertation, State University of New York at Buffalo.

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

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

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

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

Vogel, I. (1988) Prosodic constituents in Hungarian. In *Certamen Phonologicum. Papers from the 1987 Cortona Phonology Meeting* (P.M. Bertinetto & M. Loporcaro, editors), pp. 231-250. Torino: Rosenberg & Sellier.

Wang, Y., Spence, M. & Sereno, J. (1999) Training American listeners to perceive Mandarin tones, *Journal of the Acoustical Society of America*, **106**, 3649-3658.

Weijer, J. van de (1999) *Language Input for Word Discovery*. Doctoral dissertation, University of Nijmegen.

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

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