Developmental Science 6:2 (2003), pp 211–220

PAPER

Prosodic structure and syntactic acquisition: the case of the head-direction parameter

Anne Christophe,[1] Marina Nespor,[2] Maria Teresa Guasti[3] and Brit Van Ooyen[1]

1. *Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS-ENS-CNRS, Paris, France*
2. *Department of Humanities, University of Ferrara, Italy*
3. *Department of Psychology, University of Milano-Bicocca, Italy*

Abstract

We propose that infants may learn about the relative order of heads and complements in their language before they know many words, on the basis of prosodic information (relative prominence within phonological phrases). We present experimental evidence that 6–12-week-old infants can discriminate two languages that differ in their head direction and its prosodic correlate, but have otherwise similar phonological properties (i.e. French and Turkish). This result supports the hypothesis that infants may use this kind of prosodic information to bootstrap their acquisition of word order.

Introduction

Word order is one of the language-specific properties that has to be acquired by young language learners. For example, while in French the typical order for a sentence is 'J'ai écrit le livre' ('I wrote the book'), in Turkish it is reversed: 'KitabI yazdim' (the book-acc wrote-1sg, '(I) the book wrote'). This paper discusses the problem of how infants may learn this property of their mother tongue. We propose a *prosodic bootstrapping* hypothesis, whereby infants rely on the prosodic characteristics of their mother tongue to infer some of its syntactic properties (Gleitman & Wanner, 1982; Mazuka, 1996; Morgan & Demuth, 1996).

In order to account for the universal properties of phrasal structure, linguists have proposed the existence of one basic configuration shared by all human languages: the X-bar structure (Chomsky, 1981). This structure features a head, X, its complements and specifiers. X can be one of several syntactic categories such as N (noun), V (verb) or P (preposition). Languages divide into two categories: either complements follow their head, as in English and French, or vice versa, as in Turkish and Japanese. This variation in word order across languages is captured by the head-direction parameter which takes one of two values, 'head-complement' or 'complement-head'. In some languages, such as German or Dutch, the relative order of head and complements differs according to the phrase type. Even in such languages, however, the great majority of phrases have one order, on the basis of which the relative order of main and subordinate clauses is also defined. We would like to note, in addition, that the learning problem does not depend on the linguistic formalism or on the specific theory chosen to account for surface word order variation across languages. Infants have to eventually acquire the surface word order of their mother tongue (in order to correctly interpret sentences and produce them). Thus, even in a theory in which the underlying word order is supposed to be universally Subject Verb Object (Kayne, 1994), the fact that surface word order varies across languages has to be accounted for through language-specific movement operations: in this case, children have to learn about these movement operations, a problem which is equivalent to learning about underlying word order when it is assumed to vary across languages.

How do babies learn about the (surface) word order of their language? Until fairly recently the implicit assumption in the literature has been that babies have to parse at least one sentence correctly (e.g. 'close the door!', see Radford, 1990). However, as pointed out by Mazuka (1996), 'parsing a few sentences' is far from easy, and requires a lot of sophisticated processing: infants have to be able to segment the sentence into

Address for correspondence: Anne Christophe, LSCP, MSH, 54 Bd Raspail, 75270 Paris Cédex 6, France; e-mail: anne@lscp.ehess.fr

words, to understand the meaning of each of the individual words, know their syntactic category, and to extract the meaning of the sentence from the situation. None of these problems is trivial to solve, especially since they have to be solved in the absence of knowledge about language-specific syntactic properties. For instance, finding out the meaning of individual words solely from observing real-life situations is extremely difficult, as has been convincingly shown by Lila Gleitman and her colleagues (Gillette, Gleitman, Gleitman & Lederer, 1999; Gleitman, 1990; Gleitman & Gleitman, 1997). The main reason for this is that in any given situation many different sentences can be uttered (which makes it very hard to guess which particular meaning is intended). Especially for verbs, Gleitman showed that learning was significantly improved when the set of syntactic structures in which a verb may appear was available (for instance, 'thought' verbs can take phrasal complements or no complement at all; communication verbs can typically take three NP complements, corresponding to subject, object and recipient; and these correspondences hold universally). By contrast, knowing which nouns co-occurred with the target verb triggered a much smaller improvement in learning. It thus seems that those elements which are needed to acquire language-specific syntactic structures are rather hard to acquire in the absence of language-specific syntactic structure, which leads to a vicious circle, or *bootstrapping problem*. Which comes first? In this paper, we suggest a way in which word order may be acquired using prosodic information, which would be available early in the learning process and would not depend on the prior learning of any word.

Available empirical evidence suggests that the head-direction parameter is set very early. Children do not seem to make mistakes on word order when they start combining words (Bloom, 1970; Brown, 1973; Pinker, 1994). Earlier than that, one-word stage children (age 17 months) appear to use knowledge about word order in a comprehension task: Hirsh-Pasek and Golinkoff (1996) had babies watch two video screens while listening to spoken sentences. They found that babies looked longer at the scene which corresponded to the meaning of the spoken sentence (e.g. 'Cookie Monster is tickling Big Bird') than to the one which did not match the spoken sentence (e.g. a scene showing Big Bird tickling Cookie Monster).

To summarize the argument, the task of acquiring a language would be facilitated if babies knew about the relative order of heads and complements at a time when they start learning about the meanings of words, before they start combining words themselves. In this paper, we wish to examine how babies may learn about the word order of their native language before they are able to parse sentences, and even before they know words. Following Mazuka (1996), we propose that babies may use prosodic information to this end. While Mazuka argued that the relevant prosodic cues are those that distinguish main from subordinate clauses, we suggest that the relevant prosodic information is included in small chunks of utterances. Finally, we substantiate our prosodic bootstrapping hypothesis with empirical evidence obtained with infants.

The prosodic bootstrapping hypothesis

Our hypothesis rests on the fact that the head-direction parameter has a prosodic correlate: within phonological phrases, prominence systematically falls on the right for head-initial languages such as English, French or Greek, and on the left for head-final languages such as Turkish, Japanese or Bengali (see Nespor & Vogel, 1986).[1] Perceiving prominence within phonological phrases may thus allow babies to decide the relative order of heads and complements in their native language (Christophe, Guasti, Nespor, Dupoux & van Ooyen, 1997; Nespor, Guasti & Christophe, 1996). For the more complex cases, as e.g. German or Dutch, the additional assumption must be made that infants are able to find out about the prominence pattern of the great majority of phonological phrases, and that phrases exhibiting the minority pattern (which always occur at the right edge of an intonational contour) are processed later in the learning process (see Nespor *et al.*, 1996, for a fuller discussion of this point).

For this hypothesis to be plausible, one has to show that infants are able to perceive boundaries between phonological phrases and prominence within phonological phrases.[2] Before turning to the presentation of an experiment designed to test this hypothesis, we wish to briefly review the relevant evidence from the literature, which suggests that: (1) phonological phrase boundaries are signaled in the speech stream, (2) adults perceive the corresponding acoustic information and are able to exploit it to find phonological phrase boundaries (therefore it is perceivable by the human speech perception

[1] This correlation is widely accepted in the literature; see among others Hayes (1989), Hayes and Lahiri (1991), Frota (2000).

[2] Note that not every single phonological phrase bears the head-direction information: whenever there is only one word in a phonological phrase, there is no way of deciding whether prominence is initial or final. Phonological phrases which contain only one word do not constitute the majority, however. It is in fact rather hard to build up sentences which contain only one-word phonological phrases (such as: '[yesterday] [Jonathan] [arrived]'). Most utterances are therefore likely to contain information about prominence within phonological phrases.

system) and (3) infants from the age of 9 months also perceive phonological phrase boundaries. Regarding point 1, Wightman, Shattuck-Hufnagel, Ostendorf and Price (1992) observed significant lengthening in the vicinity of prosodic phrase boundaries (see also Gussenhoven & Rietveld, 1992, among others). In addition, Morgan (1996) fed a number of prosodic measurements of individual phonemes (duration, pitch and energy) into a neural network and showed that this network was quite good at finding phonological phrase boundaries in the absence of explicit coaching, by generalizing from the signals found at utterance boundaries, which are clearly audible since they are also marked by pauses. This suggests that phonological phrase boundaries are reliably marked in the speech signal. Regarding point 2, de Pijper and Sanderman (1994) showed that Dutch adults can perceive phonological phrase boundaries in their language, even in the absence of lexical information. They had adults rate the relative depth of all boundaries between the words of sentences, and found similar ratings whether subjects could understand the sentences or not (speech was rendered unintelligible through a manipulation involving resynthesis with fixed formant values, which preserved only the prosody of the original sentences). Finally, with respect to point 3, Gerken, Jusczyk and Mandel (1994) showed that American 9-month-olds reacted to the disruption of phonological phrases in their language (see also Jusczyk, Kemler Nelson, Hirsh-Pasek, Kennedy, Woodward, Piwoz *et al.*, 1992). In addition, Gout, Christophe and Morgan (submitted) observed that American 10- and 13-month-old infants were able to rely on phonological phrase boundaries in order to constrain lexical access. Namely, when trained to turned their heads for the target word 'paper', infants responded to sentences which did contain the word 'paper', but not to sentences which contained both syllables of this word, separated by a phonological phrase boundary, as in 'the butler with the least *pay* # *per*formed the best'.

The experimental evidence just presented makes it reasonable to assume that infants may have access to phonological phrase boundary information sometime during the first year of life. However, just knowing about phonological phrase boundaries is not enough, and one would like to know whether infants can perceive where prominence falls within a phonological phrase. Experimental evidence on this aspect is lacking; however, there is some evidence that young infants are sensitive to the relative prominence of syllables, within strings containing two or three syllables. Sansavini, Bertoncini and Giovanelli (1997) showed that newborn Italian infants reacted to a change from a list of bisyllabic items accented on the first syllable (such as MAma), to a list of items accented on the last syllable (such as maMA, see also Jusczyk & Thompson, 1978; Kuhl, 1983; Morse, 1972). Even though one cannot infer from these findings that infants are able to discover the most prominent syllable within a phonological phrase in continuous speech (in order to decide whether it falls more towards the left or towards the right end of the phonological phrase) these results do provide evidence for a crucial prerequisite, namely the ability to perceive acoustic prominence (at least within short and repetitive strings of syllables).

Prominence within phonological phrases as a cue to the relative order of heads and complements: experimental evidence

The results reviewed above suggest that infants may be able to perceive prominence within phonological phrases. However, this evidence is not conclusive, and therefore a direct empirical test is needed. Such a test could take a number of forms: for instance, one could study the acoustic correlates of prominence within phonological phrases and examine whether infants can perceive these acoustic correlates, using carefully controlled synthetic speech. Alternatively, one could use normal, continuous speech, and test whether infants can perceive the difference between naturally produced sentences from languages that differ in head direction. We selected the second possibility.

To draw conclusions on whether infants are sensitive to the relative prominence in phonological phrases, a pair of languages must be selected that differ in head direction but have otherwise similar prosodic structures. We selected French and Turkish for our experimental test, because they are matched on all factors which are thought to influence language prosody (namely, word stress, syllabic structure and vowel reduction): both languages have word-final stress, a similar syllabic structure (with between-words resyllabification), and no vowel reduction. But while French is head-initial, Turkish is head-final: the phonological phrase prominence is thus final in French and initial in Turkish. We constructed pairs of sentences that were perfectly matched as to the number of syllables, the position of word boundaries, word stresses, phonological phrase and intonational phrase boundaries. Thus the sentences in the two languages did not differ in any respect other than the prominence within the phonological phrases (e.g. they did not differ in number of syllables per word or phonological phrase, etc.). An example of a pair of such sentences is displayed in Figure 1. Square brackets show the position of phonological phrase boundaries, prominence within

An example of matched French and Turkish sentences

Original pair of sentences:

French:

```
                   '                  '     phonological phrase prominence
      '            '         '        '     word stress
[Le grand orang-outang]  [était énervé]
```

Turkish:

```
        '                    '              phonological phrase prominence
        '        '           '       '      word stress
    [Yeni kitabımı]     [almak istiyor]
```

Resynthesized pair of sentences:

French:

```
               '                '           phonological phrase prominence
     '         '        '       '           word stress
[leplem pelemepem]  [epe pemelse]
```

Turkish:

```
      '                 '                   phonological phrase prominence
      '        '        '       '           word stress
  [jeme pepepeme]  [elmep espejel]
```

Figure 1 *An example of a pair of matched French and Turkish sentences: identical number of syllables, same word and phonological phrase (square brackets) boundaries, same word stresses (see first level of stresses). Only the phonological phrase stresses (second level of stresses) differentiate the two sentences. Below the same sentences are represented when resynthesized. (The French sentence means: 'The big orang-outang was nervous'; the Turkish sentence means: '(S/he) wants to buy my new book').*

words is marked with one stress and prominence within phonological phrases with two stresses. It may be seen, in these examples, that of the three phonological properties indicated, only the phonological phrase prominence distinguishes the two sentences, in that it is final in French and initial in Turkish.

It is crucial to our hypothesis that infants can discriminate between French and Turkish sentences on the basis of their prosody and not on the basis of other differences, such as the phonemic repertoire. To control for this, the sentences were first pronounced naturally by native speakers of French and Turkish; then, in order to eliminate phonetic content while preserving prosodic structure, the sentences were resynthesized through diphone synthesis using an algorithm developed at IPO in Eindhoven, The Netherlands.[3] After resynthesis, all sentences were pronounced by the same voice (which recorded the diphones), and they all contained the same Dutch phonemes. All vowels were mapped to schwa, and consonants were mapped by manner of articulation (stops to /p/, fricatives to /s/, nasals to /m/, liquids to /l/ and semi-vowels to /j/; see Figure 1).

As a first step, we presented the resulting sentences to French adults in a categorization task, in order to check that the prosodic difference between French and Turkish sentences is perceivable by the human processing system (if adults can perceive this prosodic difference, it becomes more plausible that infants may do so, too). Participants were five native speakers of French; the modifications applied to the sentences were explained to them, and they had to learn to categorize the resynthesized French and Turkish sentences. In the training phase they heard half the sentences (20 in each language) and the computer gave them feedback as to the correctness of their response (the training phase was repeated up to three times unless participants reached a level of more than 75% correct during the first or second

[3] Many thanks to René Collier and Jan Roelof de Pijper for allowing us to use facilities at IPO, Eindhoven, The Netherlands.

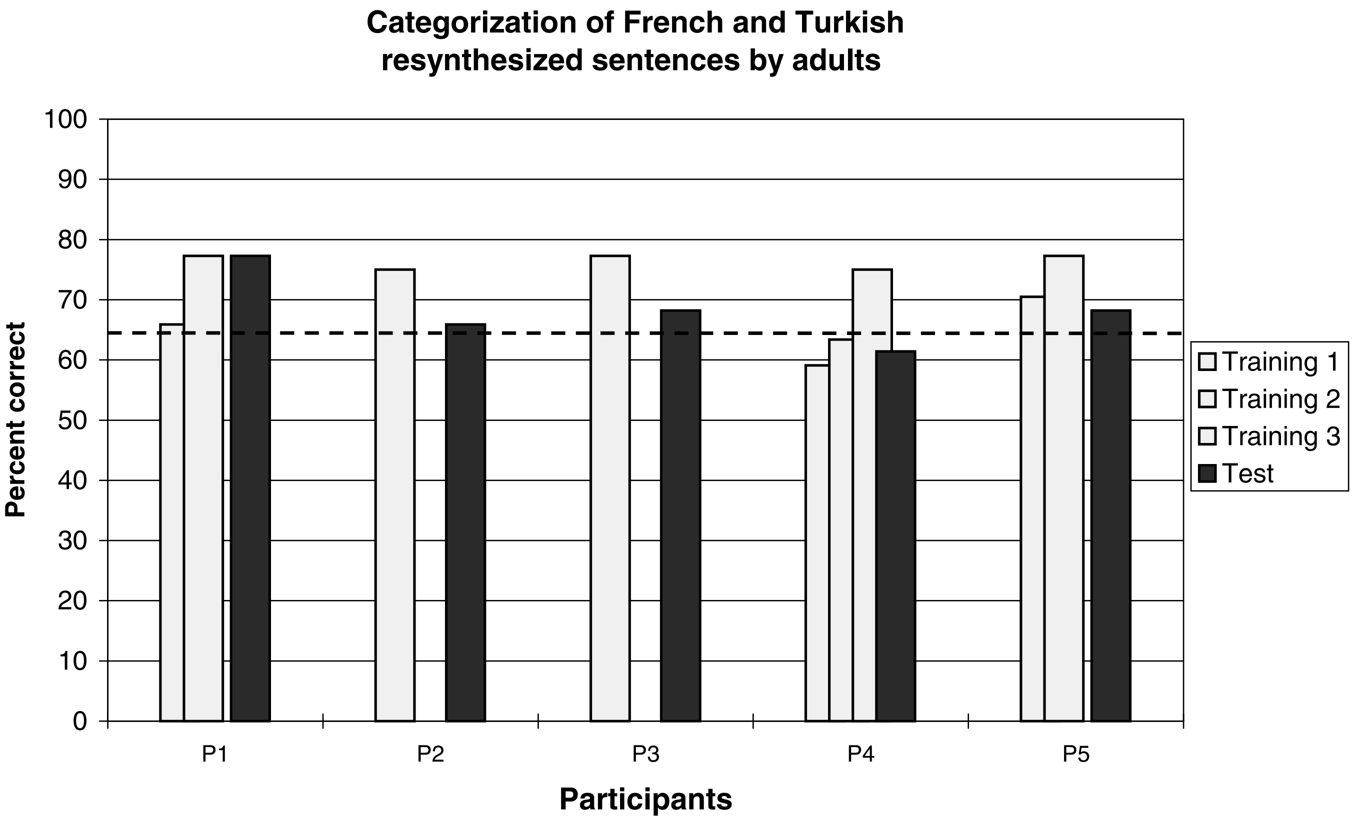

Figure 2 *Results of a categorization experiment with 5 French adults listening to resynthesized French and Turkish sentences. Each group of bars represents the percentage of correct responses from one participant; light bars represent responses in the Training phases with feedback (repeated up to three times, until each participant reached 75% correct), and dark bars represent responses in the Test phase (with sentences never heard before and without feedback). The dotted line represents better-than-chance performance (65% correct on one phase of the experiment).*

training phase, in which case they were immediately switched to the test phase). In the test phase, they had to generalize what they had learned to the second half of the sentences. We observed that all subjects but one performed better than chance (65% correct) during the test phase (see Figure 2). These results show that adult native speakers of French can learn to categorize French and Turkish sentences on the basis of the limited information preserved in the resynthesized sentences (prosodic information). Therefore we can conclude that there is some perceivable difference between the resynthesized French and Turkish sentences. Of primary interest to us is to investigate whether this difference is available to infants.

In order to test infants' sensitivity to the difference between French and Turkish sentence prosody, we used the modified version of the non-nutritive sucking procedure described by Hesketh, Christophe and Dehaene-Lambertz (1997). In this procedure, the presentation of sentences is made contingent upon the infants' high-amplitude sucking. Each baby was tested on two changes in stimulation, one experimental (change of language, from Turkish to French or vice versa) and one control (from one set of sentences to another within the same language). It is assumed that discrimination takes place whenever infants increase their sucking (and therefore the number of sentences triggered) more for the experimental change than for the control change.

Method

Material

The stimuli consisted of 40 matched pairs of French and Turkish sentences such as those shown in Figure 1. They were between 10 and 23 syllables long. They were recorded by one male native speaker of French and one male native speaker of Turkish whose voices were matched in pitch (range 100 Hz to 210 Hz, mean 150 Hz, for both speakers), since pitch is the only property of the voice that is preserved in the resynthesis process. Speakers were naïve as to the aim of the experiment and

Table 1 *Frequency of occurrence of different syllabic types and consonantal types in the French and Turkish sentences from our corpus*

Syllabic types:	CV	CVC	CCV	V	VC	CCVC	Other syllabic types
French	52%	23%	9%	6%	4%	4%	<1%
Turkish	57%	29%	3%	4%	5%	2%	<1%

Consonantal types:	liquids	stops	nasals	fricatives	semi-vowels
French	26%	31%	16%	21%	6%
Turkish	17%	37%	20%	19%	7%

were instructed to speak as naturally as possible.[4] Each sentence was digitized at 16 kHz and stored in the computer. Sentences were hand-labeled so that the beginning of each phoneme was marked. For each sentence, the original string of phonemes was translated using the grid mentioned above (vowels to schwa, consonants by manner of articulation); the resulting string of phonemes was used to diphone-synthesize a new sentence; the duration and pitch of each phoneme from the original sentence was copied onto the synthetic sentences by means of the PSOLA algorithm (Mouline & Charpentier, 1990). As a result, the synthetic sentences replicated the prosody of the original sentences, but contained only Dutch phonemes. This procedure was performed through a software program developed at IPO, Eindhoven. The sentences were split into two blocks matched for syllabic length (mean 15.5 syllables, range 10 to 23 syllables) and duration (mean 2.7 seconds, range 1.6 to 4.1 seconds).

French and Turkish are matched on a number of prosodic characteristics; in addition, we constructed pairs of sentences which were matched for the position of word, phonological phrase and intonational phrase boundaries (so that characteristics like the mean number of syllables per word, phonological phrase or intonational phrase were identical between French and Turkish, in our corpus). The only aspects which were not exactly matched between sentences of a pair were the syllabic structure and phonemic content. For these aspects, we computed statistics on our corpus (see Table 1). French and Turkish have similar syllabic repertoires, including V, CV, CCV, VC, CVC, CCVC, VCC, CVCC, CCVCC, for both languages, and some rare CCC onsets in French only. We observed that the frequency with which these syllabic types occurred was also similar in French and Turkish in our corpus; CV and CVC accounted for more than 75% of syllables in both languages. Regarding consonantal types (by manner of articulation, since this is the only information which was preserved in the resynthesis process), we also observed similar distributions, the only small difference being slightly more liquids and less stops in French relative to Turkish.

Apparatus and procedure

Subjects were tested individually at the Maternité Port-Royal, Paris. They were seated in a relax chair placed in a sound-proof room and offered a standard (heat sterilized) pacifier. One experimenter, out of view of the baby and deaf to the stimuli (listening to masking noise over sound-proof headphones[5]), checked that the pacifier stayed in the baby's mouth throughout the experiment. A second experimenter monitored the experiment on the computer which recorded the pressure of the infant's sucks via an analogue-digital card (Data Translation), detected the sucking responses on the basis of speed of increase and decrease in sucking and an amplitude threshold and delivered the sentences through an OROS sound board according to the reinforcement schedule (see below). The computer also saved both the moment and amplitude of each suck as well as the stimuli triggered by the sucks. Hesketh *et al.* (1997) reported that the number of sentences triggered was a cleaner measure than the number of sucking responses. Only this measure will be reported here.

The experiment started with a short baseline during which the baby sucked on the pacifier but heard no sentences (about 30 s). Next, the first phase of the experiment began, during which infants heard sentences in either Turkish or French (depending on the condition) contingently upon their high-amplitude (HA) sucks. During baseline and the first minute of the experiment, the experimenter adjusted the sensitivity of the sucking detection apparatus to ensure a correct rate of HA sucks. After a short shaping phase, three HA sucks were required to trigger each sentence (such that there was less than one second between consecutive sucks). There was an inter-stimulus interval (ISI) of at least 600 ms between consecutive sentences. For each phase of the experiment, the order of presentation of the sentences was quasi-random and different for each baby (i.e. all sentences were presented once in random order, then

[4] Note that the stimuli are adult-directed speech: given the highly constrained nature of the stimuli, it would have been impossible to obtain a set of well-matched sentences with naturally produced infant-directed speech. However, many successful language discrimination experiments were performed with adult-directed speech (e.g. Nazzi, Bertoncini & Mehler, 1998), and there is therefore no reason to think that very young infants fail to pay attention to the prosody of adult-directed speech.

[5] The masking noise was made using several superimposed streams of continuous speech ('babble noise'). It sounds like cocktail party noise, and is very efficient in masking speech stimuli.

reshuffled and presented once more, and so on to the end of the phase).

Each infant received two changes in stimulation, one experimental (language) change, and one control change (change of sentences within the same language). Half the infants received the experimental change first and the control change second and the other half received the control change first and the experimental change second. In addition, the order of presentation of languages was counterbalanced across subjects. This yielded the following four conditions: 1. French, Turkish, Turkish; 2. Turkish, French, French; 3. French, French, Turkish; 4. Turkish, Turkish, French.

A switch in stimulation occurred as soon as a predefined habituation criterion had been met. For two consecutive minutes the infant's HA sucking rate had to be less than 80% of the maximum number of HA sucks per minute, counting from the beginning of the experiment but excluding the very first minute of stimulation. In addition, the first switch of stimulation could not occur until the baby had shown a sufficient level of activity (at least 20 HA sucks in at least one minute). Finally, at least one sentence had to be triggered during the last minute before a switch in stimulation, to avoid having a silent period between two phases. Each phase of the experiment lasted at least 5 full minutes.

Subjects

Sixteen infants aged between 6 and 12 weeks (mean age: 62.5 days) successfully completed the experiment (4 in each subcondition). In addition, 23 infants were tested but their data were excluded from the final analysis for the following reasons: completed one switch only (6); cried (13); sucked at low amplitude or not at all (2); fell asleep (2). Subjects were randomly assigned to one of the four conditions prior to testing. All infants had a monolingual French background.

Results

The number of sentences triggered per minute was analyzed. To assess the effect of the experimental manipulation, we compared the increase in response rate for the Experimental switch (i.e. change in language) to the increase in response rate for the Control switch (i.e. change of sentences within the same language). This comparison was within-subject. Two dishabituation scores were computed for each baby: the Experimental dishabituation score was the mean response rate for two minutes after the Experimental switch minus the mean response rate for two minutes before this switch; the Control dishabituation score was the same measure for the Control switch. Figure 3 shows mean Experimental and Control dishabituation scores.

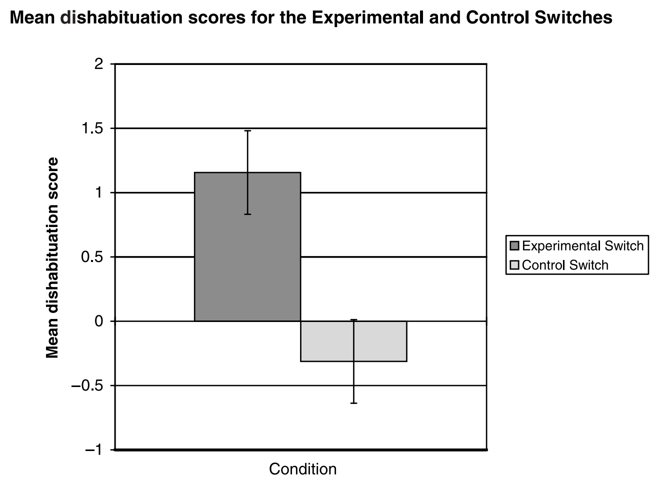

Figure 3 *Results of a non-nutritive sucking experiment with 16 2-month-old French infants. Each infant listens to two switches of stimulation: one Experimental (language change) and one Control (change of sentences within the same language). For each switch of stimulation, a dishabituation score is computed: the difference between the mean number of sentences triggered 2 minutes after and 2 minutes before the switch. The graph represents mean Experimental and Control dishabituation scores (each error bar represents one standard error of the mean above and below the mean).*

An ANOVA was performed, using as dependent measure the Experimental and Control dishabituation scores. The main Experimental factor was within-subject (Experimental vs. Control switch) and there were also two between-subject counterbalancing factors, Order (Experimental switch first, versus Control switch first) and Language (French first vs. Turkish first). The analysis showed a significant effect of the Experimental factor ($F(1, 12) = 7.4$, $p < .02$), indicating that infants increased their response rate more in the Experimental than in the Control switch. The counterbalancing factors showed no significant main effect, nor did they interact with the Experimental factor. In addition, a non-parametric test was also performed on the same data, since the number of infants is relatively small: a discrimination index was computed for each infant, as the difference between the Experimental and the Control dishabituation scores; a non-parametric Wilcoxon signed rank test showed that the median of the discrimination indices was significantly above 0 ($Z = 2.2$, $p < .03$), indicating that most infants dishabituated to the Language switch more than to the Control switch.

These results show that French infants aged between 6 and 12 weeks hear a difference between French and

Turkish sentences that were resynthesized in order to remove phonemic information. Recall, in fact, that all phonemes from the original French and Turkish sentences were replaced by Dutch phonemes. We thus know that infants did not react to a difference in phonemes. We carefully selected French and Turkish for our experimental test, because they are matched on all factors which are thought to influence language prosody (namely, word stress, vowel reduction, syllabic structure). In addition, the sentences in French and Turkish were matched on word, phonological phrase and intonational phrase boundaries, so that characteristics like the number of syllables per word or per phonological phrase were identical in our French and Turkish sentences. Even the frequencies of syllabic and phonemic types (which were not exactly matched in our sentence pairs) were similar in our corpus. The most likely interpretation of the present experiment is therefore that infants discriminated between sentences in French and in Turkish on the basis of the prosodic difference linked to prominence within phonological phrases. This in turn lends plausibility to the hypothesis that infants may rely on prosodic information in order to decide whether their language is head-initial or head-final. This prosodic information is available to young infants before the end of the first year of life, and might therefore be used in the acquisition of the syntax of their native language.

Conclusion

In this paper, we presented a prosodic bootstrapping hypothesis for setting the head-direction parameter: if infants can perceive prominence within phonological phrases, they may decide whether their maternal language is head-final or head-initial on the basis of prosodic information only. We presented an experiment that suggests that infants may be sensitive to precisely that prosodic information before the end of the first year of life: 2- to 3-month-old infants were able to distinguish between sentences in French (head-initial) and Turkish (head-final), even though these sentences had been resynthesized to take phonemic information away, leaving solely prosodic information.

The experiment presented in this paper shows that a first prediction of our hypothesis was borne out: we have observed that both infants and adults can distinguish between sentences from French and Turkish, presumably on the basis of some prosodic difference resulting from the fact that French is head-initial while Turkish is head-final. This experimental result makes our hypothesis more plausible: it could indeed be the case that infants decide about the relative order of heads and complements in their native language on the basis of its prosody, sometime during the first year of life. However, the present result is not a proof that infants actually decide about their language's word order by listening to its prosody.

To further confirm the plausibility of our hypothesis we can carry out a detailed study of the acoustic correlates of prominence within phonological phrases across languages. Indeed, for our proposal to be valid, we have to assume that prominence is detectable by infants from a relatively uninformed analysis of whole sentences. This would be impossible if, for instance, prominence was marked in a different way in each language. Instead, we have to assume that across the languages of the world, there exist only a limited number of ways to mark prominence. Once features which typically mark prominence are established, then one may conduct some experimental studies to check whether infants are sensitive to them. In fact, one study of the acoustic correlates of prominence is in progress, and results so far suggest that prominence may be marked mainly through raised pitch in head-final languages while it would be marked mainly through increased duration in head-initial languages (Avesani, Donati & Nespor, in preparation; Pierrehumbert, personal communication).

Finally, the idea that one part of syntax is available through prosodic information contained in the signal may become more plausible if it does not constitute an isolated case. The question we want to ask is thus whether only the relative order of heads and complements is signaled through prosody, or other main syntactic properties of languages could also be so signaled. Nespor (2002) has put forward another such hypothesis, namely, that the location of the main prominence within intonational phrases (IP) may signal other word order properties of languages. It is proposed that the location of IP main prominence may give some information about the relative freedom of occurrence of phrases within a sentence. In languages with free word order, IP main prominence is always final, while prominence in other locations indicates that the language has strict word order. The specific location of this prominence (e.g. at IP left edges or in the middle of IP) would indicate in which portion of the sentence different orders are allowed. A relatively free word order may in turn give some information about the obligatoriness of specific complements (see also Donati & Nespor, in press). While the location of prominence within the phonological phrase may cue the relative order of the constituents of a syntactic phrase, the location of prominence within the intonational phrase may cue the relative order of syntactic phrases within a sentence.

An important aspect of the present prosodic bootstrapping proposal is that the head-direction parameter could be set prelexically, without any knowledge of the lexicon. This implies that as soon as infants comprehend some words (around the age of 12 months, see Oviatt, 1980, or even before) they may start working on the meaning of whole sentences. That is, knowing the basic word order of their language of exposure and the meaning of some words may allow them to understand the meaning of at least some sentences and to start working on more complex ones.

Acknowledgements

The work reported in this paper was supported by a grant from the European Science Foundation, 'The role of intonation in early language acquisition' (European Neuroscience Programme, Research grant no. 255), as well as by the Human Capital and Mobility programme, the GIS Sciences de la Cognition (no. 99N35/0008) and the ACI Cognitique of the MENRT (Paris). Infants' testing was performed at the Maternité Port-Royal, Paris. We wish to thank René Collier and Jan Roelof de Pijper for providing access to the facilities at IPO, Eindhoven, The Netherlands. We also wish to thank two anonymous reviewers for thoughtful comments on the first version of this manuscript.

References

Avesani, C., Donati, C., & Nespor, M. (in preparation). Acoustic correlates of phonological phrase prominence in Turkish and French.

Bloom, L. (1970). *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Christophe, A., Guasti, M.T., Nespor, M., Dupoux, E., & van Ooyen, B. (1997). Reflections on phonological bootstrapping: its role for lexical and syntactic acquisition. *Language and Cognitive Processes*, **12**, 585–612.

de Pijper, J.R., & Sanderman, A.A. (1994). On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *Journal of the Acoustical Society of America*, **96**, 2037–2047.

Donati, C., & Nespor, M. (in press). From focus to syntax. *Lingua*.

Frota, S. (2000). *Prosody and focus in European Portuguese. Phonological phrasing and intonation*. New York: Garland.

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

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, **73**, 165–176.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, **1**, 3–55.

Gleitman, L., & Gleitman, H. (1997). What is a language made out of? *Lingua*, **100**, 29–55.

Gleitman, L., & Wanner, E. (1982). The state of the state of the art. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 3–48). Cambridge: Cambridge University Press.

Gout, A., Christophe, A., & Morgan, J. (submitted). Phonological phrase boundaries constrain lexical access: II. Infant data.

Gussenhoven, C., & Rietveld, A.C.M. (1992). Intonation contours, prosodic structure and preboundary lengthening. *Journal of Phonetics*, **20**, 282–303.

Hayes, B. (1989). The prosodic hierarchy in meter. In P. Kiparsky & G. Youmans (Eds.), *Rhythm and meter. Phonetics and phonology 1* (pp. 201–260). New York: Academic Press.

Hayes, B., & Lahiri, A. (1991). Bengali intonational phonology. *Natural Language and Linguistic Theory*, **9**, 47–96.

Hesketh, S., Christophe, A., & Dehaene-Lambertz, G. (1997). Non-nutritive sucking and sentence processing. *Infant Behavior and Development*, **20**, 263–269.

Hirsh-Pasek, K., & Golinkoff, R.M. (1996). *The origins of grammar: Evidence from early language comprehension*. Cambridge, MA: MIT Press.

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

Jusczyk, P.W., & Thompson, E.J. (1978). Perception of a phonetic contrast in multisyllabic utterances by two-month-old infants. *Perception & Psychophysics*, **23**, 105–109.

Kayne, R. (1994). *Antisymmetry in syntax*. Cambridge, MA: MIT Press.

Kuhl, P.K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development*, **6**, 263–285.

Mazuka, R. (1996). How can a grammatical parameter be set before the first word? In J.L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 313–330). Mahwah, NJ: Lawrence Erlbaum Associates.

Morgan, J. (1996). Phonological bootstrapping in early acquisition: the importance of starting small versus starting 'sharp'. Paper presented at the conference How to get into language: Approaches to bootstrapping in early language development, Potsdam, Germany.

Morgan, J.L., & Demuth, K. (1996). Signal to syntax: an overview. In J.L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 1–22). Mahwah, NJ: Lawrence Erlbaum Associates.

Morse, P.A. (1972). The discrimination of speech and nonspeech stimuli in early infancy. *Journal of Experimental Child Psychology*, **13**, 477–492.

Mouline, E., & Charpentier, F. (1990). Pitch synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication*, **9**, 453–467.

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. (2002). About parameters, prominence, and bootstrapping. In E. Dupoux (Ed.), *Language, brain and cognitive development: Essays in honor of Jacques Mehler*. Cambridge, MA: MIT Press.

Nespor, M., Guasti, M.T., & Christophe, A. (1996). Selecting word order: the rhythmic activation principle. In U. Kleinhenz (Ed.), *Interfaces in phonology* (pp. 1–26). Berlin: Akademie Verlag.

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

Oviatt, S.L. (1980). The emerging ability to comprehend language: an experimental approach. *Child Development*, **51**, 97–106.

Pinker, S. (1994). *The language instinct*. London: Penguin Books.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax*. Cambridge: Basil Blackwell.

Sansavini, A., Bertoncini, J., & Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental Psychology*, **33**, 3–11.

Wightman, C.W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P.J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, **91**, 1707–1717.

Received: 15 November 2000
Accepted: 30 November 2001