9

Rhythmic patterns cue word order*

MOHINISH SHUKLA AND MARINA NESPOR

9.1 Introduction

Human language can be construed as a system of rules that governs the organization of the words of a language to convey meaning. In most languages, the basic ordering of words in an overt linear sequence is determined by fundamental syntactic rules of organization (Baker 2001). The basic word order shows language-specific variations. For example, while in some languages, like English or Greek, direct objects (O) follow verbs (V), in others, like Hindi or Turkish, objects precede the verbs (Greenberg 1963; Comrie 1981). It has been observed that the basic VO or OV order of a language captures major generalizations about the ordering of other words and phrases in that language (Greenberg 1963; Dryer 1992).

However, the variation among OV and VO languages—in total about 86% of the languages of the world (Dryer 2005)[1]—presents a problem for acquisition. Given a pair of words X-Y, how is the young learner to determine whether this represents a VO or an OV case, before the knowledge of words? In this chapter, we consider the hypothesis that the different word orders are manifested as different rhythmic patterns in spoken languages.

Overt speech is more than just a linear sequence of words. It is organized into prosodic domains that link the fundamental units of spoken language to multi-word utterances (Selkirk 1984; Nespor and Vogel 1986 [2008]). In 9.2, we outline our proposal that relates a specific domain of the prosodic

* For interesting discussions on many issues related to the topic of the present paper, we thank Ricardo Bion, Ansgar Endress, Judit Gervain, Jacques Mehler, and Marcela Peña. The research described in this paper was funded by the ESF Eurocores OMLL grant, the Italian National Grants (COFIN) 2003–7, and the James S. McDonnell Foundation.

[1] Dryer (2005) reports ~14% of languages as lacking a dominant word order. The rest include all possible orders of S,V, and O. However, SOV and SVO are the most frequent of all languages, accounting for ~76% of his sample.

hierarchy, the phonological phrase, to the acquisition of word order. This proposal is based on the observation that different physical manifestations of phonological phrase prominence trigger either trochaic or iambic grouping—the so-called Iambic–Trochaic Law (ITL).

In 9.3, the phonological phrase is discussed in greater detail. We review linguistic evidence for the phonological phrase. We discuss (a) how prominence in phonological phrases reflects word order, and (b) acoustic cues that mark phonological phrases in languages with differing word orders. In 9.4, we review psycholinguistic evidence for the phonological phrase in speech processing by adults and infants. In 9.5, we discuss the acquisition of grammar by infants. We show how infants might acquire word order through the perception of language-specific prominence patterns of phonological phrases. Further, we suggest that general perceptual capacities, coupled with some elementary distributional computations, can account for the remarkable achievements of pre-linguistic infants in acquiring the grammar of their language. Concluding remarks are presented in 9.6.

9.2 The phonological phrase organization reflects word order

Word order is one of the first syntactic properties infants acquire, as evidenced by the fact that they do not make mistakes when they start uttering their first two-word sentences and by the fact that, long before then, they react differently to good and bad word order (Brown 1973; Meisel 1992; Clahsen and Eisenbeiss 1993). How can we account for this remarkable achievement? We propose that infants have the possibility of acquiring word order prelexically, that is, before the acquisition of the lexicon. Based on Nespor et al. (2008), in this chapter we propose that they do so on the basis of the rhythmic patterns of the language of exposure.

The VO or OV pattern in various languages reflects a general tendency to place heads either before or after their complements, and to place main clauses either before or after subordinate clauses. Thus, the VO or OV word order is related to the direction of branching: right-branching in head-complement (or VO) languages and left-branching in complement-head (or OV) languages (see Dryer 1992). In 9.3 we will see that the prosodic level that signals word order is the phonological phrase (PhPh). The location of prominence within a phonological phrase has been proposed to depend on whether it is mapped from a head-complement or complement-head language: rightmost in the former, leftmost in the latter (Nespor and Vogel 1982, 1986/2008). It has further been proposed that in a language with both orders of heads and complements, the location of

prominence varies in a similar manner, according to the order within individual phrases (Nespor, Guasti, and Christophe 1996). In addition, the physical manifestation of prominence differs in the two cases: it is manifested mainly through duration in iambic PhPhs and mainly through pitch and intensity in trochaic PhPhs (Nespor et al. 2008). Since in a head-complement pair of words with unmarked prosody the complement bears phrasal stress (Nespor and Vogel 1982, 1986/2008; Cinque 1993; Féry and Herbst 2004), infants might be forced to identify the complement, and thus its order with respect to the head, because of a mechanism of general perception.

A uniform rhythmic pattern throughout the language will lead the infant to conclude that its language of exposure has a uniform order of heads and complements, and thus of subordinate clauses with respect to main clauses. The head-complement parameter could thus be set with a general perception mechanism.[2] Infants exposed to languages in which the order varies, a problem for parameter setting, will also identify word order, realizing that it is mixed. In these languages, there is usually a main pattern and some deviations from it. The same mechanism would thus allow the acquisition of word order both in a regular and in a mixed language, accounting for the fact that there is no delay in the acquisition of the latter with respect to the former.

9.3 The phonological phrase

Connected speech can be analyzed as a hierarchical organization of constituents ranging from the syllable and its constituent parts until the utterance (Selkirk 1984; Nespor and Vogel 1986/2008). The phonological phrase is one of the phrasal constituents of the phonological, or prosodic, hierarchy, crucial both in signaling syntactic constituency and in providing cues to word order.

As is the case for the other constituents of the prosodic hierarchy, PhPhs are signaled both by phenomena that apply throughout their domain and by edge phenomena. These different cues are used both in the processing of speech by adults, for example to disambiguate certain sentences with an identical sequence of words but different syntactic structures (Nespor and Vogel 1986 [2008]; Christophe et al. 2004, Millotte et al. 2007; Millotte et al. 2008), and in language acquisition by infants (Gout et al. 2004).

[2] In this chapter we use the term 'parameter setting' in its most general sense, as a shorthand to mean the acquisition of a certain language-specific property, like the relative order of heads and complements.

The domain of the phonological phrase extends from one edge of a syntactic phrase until, and including, its head. There are thus two possibilities; which one is chosen depends on syntactic structure. If the structure is right-branching, the phonological phrase starts from the left edge of a syntactic phrase (X") and it ends at the right edge of its head (X). If the structure is left-branching, the phonological phrase starts from the right edge of a syntactic phrase and it ends at the left edge of its head. These two possibilities are illustrated in (1a) and (1b), where the underlined parts indicate the domain of the phonological phrase in right-branching (or head-complement) and left-branching (or complement-head) structures, respectively.

(1) a. $X"[\underline{\ldots[\ ]}_X \ldots\ldots]$
b. $[\ldots\ldots X\underline{[\ ]\ldots}]_{X"}$

That is, the domain over which the phonological phrase extends depends on syntactic structure, in that it extends on opposite parts of the head, and thus varies depending on word order. The two possibilities are illustrated in (2) and (3), on the basis of adpositional phrases in Italian and Turkish, respectively. Notice that, though adpositions are heads syntactically, they behave as non-heads, as do all closed-class items, in phonology (Nespor and Vogel 1986/2008).

(2) a. per Luca
'for Luca'
b. accanto al giardino
'next to the garden'
c. dietro alla porta
'behind the door'

(3) a. tren ile
train—with
'by train'
b. Aynura göre
Aynur—according to
'according to Aynur'
c. dakika gibi
minutes—like
'like minutes'

In addition to this basic PhPh domain, hypothesized to be universal, there is a possibility of restructuring: if non-branching, the first complement or

modifier on the recursive side of the head can be restructured into the same PhPh of the head (Nespor and Vogel 1986/2008). The level at which the non-branching condition holds appears to present some cross-linguistic variation: it is the level of the word, e.g. in Italian, but the level of the clitic group, e.g. in English (Hayes 1989), as exemplified in (4) and (5), respectively.

(4) a. $[\text{mangia}]_{PP}\ [\text{frutta}]_{PP} \rightarrow [\text{mangia frutta}]_{PP}$
'(s/he) eats fruit'
$[\text{mangia}]_{PP}\ [\text{la frutta}]_{PP} \rightarrow {}^{*}[\text{mangia la frutta}]_{PP}$
'(s/he) eats the fruit'

b. $[\text{scrive}]_{PP}\ [\text{poesie}]_{PP} \rightarrow [\text{scrive poesie}]_{PP}$
'(s/he) writes poems'
$[\text{scrive}]_{PP}\ [\text{delle poesie}]_{PP} \rightarrow {}^{*}[\text{scrive delle poesie}]_{PP}$
'(s/he) writes the poems'

(5) a. $[\text{eat}]_{PP}\ [\text{fruit}]_{PP} \rightarrow [\text{eat fruit}]_{PP}$
$[\text{eat}]_{PP}\ [\text{the fruit}]_{PP} \rightarrow [\text{eat the fruit}]_{PP}$

b. $[\text{he wrote}]_{PP}\ [\text{poems}]_{PP} \rightarrow [\text{he wrote poems}]_{PP}$
$[\text{he wrote}]_{PP}\ [\text{the poems}]_{PP} \rightarrow [\text{he wrote the poems}]_{PP}$

9.3.1 *The relation of prominence in phonological phrases to word order*

Different phonological phenomena, that apply either throughout the constituent or at (one of) its edges, signal that the elements that constitute a phonological phrase have a certain level of cohesion. In addition to segmental phenomena, prominence marks one of the PhPh edges. As in all constituents of the phonological hierarchy, relative prominence in a PhPh is assigned to its daughter constituents: the constituent located at one of the edges is marked as strong and all the other constituents are marked as weak (Liberman and Prince 1977).

The location of the strong element of the phonological phrase—as well as the domain over which it extends—depends on the relative order of heads and complements (Nespor and Vogel 1986/2008). In head–complement structures the strongest element within a PhPh is rightmost, while in complement–head structures it is leftmost. Thus the element of a PhPh that bears the main prominence is either the rightmost or the leftmost depending on the recursive side of a given language: in right recursive languages, e.g., Greek, French, or Arabic, the strongest element is at the right edge; in left recursive languages, e.g., Turkish, Basque, or Japanese, the strongest element is at the left edge.

Some languages, like German or Dutch, do not have a uniform rhythmic pattern—either iambic or trochaic. While mainly iambic, these two languages

have some trochaic phonological phrases. These different rhythmic patterns reflect word order in all cases. While both in German and in Dutch, in most phrases, heads precede their complements, in certain cases, verb phrases have the order complement–verb and adpositional phrases are postpositional, as exemplified in (6) and (7), respectively, on the basis of Dutch.

(6) a. Ik koop land
'I buy land'

b. Paul weet dat ik land koop
'Paul knows that I buy land'

(7) a. op de trap
'on the stairs (state)'

b. de trap op
'on the stairs (movement)'

These different orders of heads and complements do not reflect two values of the head—complement parameter: there are different syntactic reasons that determine the two orders (Koster 1975; Haider and Prinzhorn 1986, among others). Nevertheless, different rhythms mark the specific surface word orders.

9.3.2 *Acoustic differences mark different prominence patterns*

We saw, above, that different word orders are reflected in different prominence patterns. In this section, we will see that the different prominence patterns are accompanied by different acoustic cues.

Several authors have examined the physical, acoustic correlates of the various prosodic domains, which include intonation patterns, pausing, phrase-final lengthening and constituent-initial articulatory strengthening (e.g., Wightman et al. 1992; Fisher and Tokura 1996; Keating et al. 2003).

In recent work, it has been proposed that the physical manifestation of prominence within phonological phrases differs according to whether the strong element is in initial or in final position. Specifically, it has been hypothesized, in agreement with the Iambic–Trochaic Law, that it is realized more through pitch and intensity if initial and more through duration if final (Nespor et al. 2008). That initial prominence is mainly marked by intensity and final prominence by duration has been shown to be the case for music (Bolton 1894; Woodrow 1951; Cooper and Meyer 1960) as well as for stress at the foot level (Hayes 1995). The different physical manifestations of prominence depending on location thus appear to be, more generally, derived from acoustic perception. The Iambic–Trochaic law states that if the sounds in a sequence alternate in degrees of intensity, being identical in all other respects,

humans perceive them as a sequence of binary trochaic (i.e. strong first) groups. If the sounds of a sequence alternate in duration, they are instead perceived as a sequence of binary iambic (i.e. strong last) groups. This is graphically represented in (8a) and (8b), respectively.

(8) a. H=High intensity, l = low intensity
... H l H l H l H l H l ... →
... [H l] [H l] [H l] [H l] [H l] ...

b. L = Long duration, s = short duration
... L s L s L s L s L s ... →
... [s L] [s L] [s L] [s L] ...

Nespor et al. (2008) examined the acoustic characteristics of phonological phrases with differing word orders both crosslinguistically and within the same language. To look at crosslinguistic variation, Turkish phonological phrases were used to exemplify initial PhPh prominence and French PhPhs to exemplify final prominence. These specific languages were chosen because they both have word-final stress and a similar syllabic structure, so that it was possible to find quite well-matched phrases that crucially differ in the location of PhPh stress. The specific material used consisted of phonological phrases in which the words occupying the position of phrasal prominence were identical in the two languages: either French words borrowed into Turkish, like *cognac*, or Turkish words borrowed into French, like *kilim*. These authors demonstrated that, at the level of the phonological phrase, higher pitch and intensity characterize the initial position, while increased duration characterizes the final position.

In the same study, confirmation of the hypothesis that different acoustic correlates characterize either initial or final PhPh prominence came from German, a language in which, as we said, both orders are found. In PhPhs consisting of two words, either N(object)–V or V–N(object), phrasal prominence falls on the complement independent of its location. In addition, it was found that the prominence on the complement N was characterized more by intensity and pitch when preverbal and more by lengthening when postverbal (Nespor et al. 2008).

The conclusion may thus be drawn that specific (acoustic) types of phrasal stress signal whether a language is head–complement or complement–head, and, within one language, the order of head and complement for each phrase.

9.4 Detecting and using phonological phrases

It has been well established that prosody plays a key role in understanding spoken sentences (e.g., Lehiste 1973; Nespor and Vogel 1983; Warren 1996;

Cutler et al. 1997; Frazier et al. 2006). While in the previous section, we looked at linguistic reasons to motivate phonological phrases and the acoustic cues that mark phonological phrases, in this section we review psycholinguistic evidence that both infants and adults can both detect and utilize phonological phrases in speech processing.

9.4.1 *Adult data*

In adults, in general, prosodic effects have been observed in two kinds of studies: (a) disambiguating possible syntactic (and semantic) parses of a spoken sentence and (b) identifying word boundaries in speech.

Since larger prosodic groupings like the intonational phrase (IP) are more salient in speech, several researchers have examined the effects of IP-breaks in disambiguating garden-path sentences like *The workers considered the last offer from the management was a real insult*, where *the last offer from the management* might be initially construed as the object of the verb *consider* instead of a sentence complement *[that] the last offer* (Marslen-Wilson et al. 1992; see also, e.g., Nespor and Vogel 1983, Clifton et al. 2006, and also Snedeker and Yuan 2008, for data from 4–6 year olds). In addition, evidence has also been found that, under some conditions, smaller prosodic phrases, corresponding to the phonological phrase, can also be used to disambiguate syntax (example in (9) below; see also Kjelgaard and Speer 1999; Clifton et al. 2006).

More recently, Millotte and colleagues (Millotte et al. 2007; Millotte et al. 2008) have provided direct evidence that phonological phrase prosody can directly influence online syntactic analyses. For example, consider the phrases (9a) and (9b) (from Millotte et al. 2008; phonological phrases are marked):

(9) a. [[Le petit chien]$_{PP}$ [*mord* la laisse]$_{PP}$ [qui le retient]PP]$_{IP}$
The little dog bites the leash that holds it back.

b. [[Le petit chien *mort*]$_{PP}$ [sera enterré demain]$_{PP}$]$_{IP}$
The little dead dog will be buried tomorrow.

The word *mord/mort* (pronounced identically in the two cases, /mɔʀ/) is initially ambiguous as to it being an adjective or a verb. In a sentence-completion task, the authors found that when the prosody indicated a grouping as in (9a), participants were more likely to infer a verb continuation. Conversely, they were more likely to infer an adjective continuation for (9b).

Using the same sets of stimuli as in (9), Millotte et al. (2008) also found that the detection of a category-labeled word (/mɔʀ/ as noun or adjective) was influenced by prosody: if /mɔʀ/ came at the end of a phonological phrase, as

in (9b), it was more quickly detected if the participant was instructed to respond to *mort*, rather than to *mord*.

Thus, the metalinguistic ability of adult participants to detect words in fluent speech is conditioned by the prosodic phrasing of the sentence. However, other studies have shown that even the online, implicit segmentation of fluent speech into words is influenced by prosody (e.g., Salverda et al. 2003; Christophe et al 2004).

In particular, Christophe et al (2004) asked French participants to monitor target words, e.g. *chat* /**ʃa**/ (cat) that occurred either in a locally ambiguous context (e.g., chat grincheax /**ʃagʀɛ̃ʃø**/, where the chagrin /**ʃagʀɛ̃**/ is a French word, or an unambiguous context (e.g., chat drogué /**ʃadʀoge**/; there's no French word beginning with /**ʃad**/). These authors found delayed lexical access when a local lexical ambiguity was included in one phonological phrase. In contrast when a phonological phrase boundary intervened between the two words of the lexical competitor, no delay was observed. From this study the conclusion was drawn that lexical access is constrained by phonological phrases.

9.4.2 *Infant data*

Although infants have been shown to be sensitive to prosodic groupings in speech (e.g., Hirsch-Pasek et al. 1987; Kemler et al. 1989; Jusczyk et al. 1992; Morgan 1994; Nazzi et al. 2000), their limited linguistic prowess makes it difficult to empirically investigate how they might use prosody to resolve syntactic ambiguities.

Nevertheless, newborns can discriminate bisyllables (e.g. *latí*) that come from within (e.g., ge*latí*na) or across (e.g., gorí*la tí*sico) prosodic phrases, suggesting that they are sensitive to and can utilize the acoustic cues that distinguish the two types of bisyllables (Christophe et al. 2001; Christophe et al. 2004).

Further, Soderstrom et al. (2003) showed that 6- and 9-month-old infants, when familiarized with a word sequence (e.g., *people b(u)y the (w)hole*), preferred passages where such sequences were well-formed phrases (. . . *people by the hole*. . .), as compared to passages where a (phonological) phrase boundary interrupted the sequence (. . . *people # buy the whole*. . .)

In addition, Gout et al. (2004) showed that 10 to 12.5-month-old American infants use phonological phrase boundaries to constrain online lexical access. Thus, when trained to turn their heads to isolated bisyllabic English words like 'paper', these infants were more likely to turn towards sentences in which 'paper' did not straddle a phonological phrase (e.g., [The scandalous **paper**]$_{PP}$

[sways him]$_{PP}$ [to tell the truth]$_{PP}$) than to sentences in which 'paper' straddled a phonological phrase (e.g., [The outstanding **pay**]$_{PP}$ [**per**suades him]$_{PP}$ [to go to France]$_{PP}$). Similar results were found for French 16-month-old infants (Millotte 2005).

These observations underline the importance of phonological phrases not just within linguistic theory (previous section), but also in understanding the online processing of speech both in adults and in infants. With the phonological phrase motivated on both linguistic and psycholinguistic grounds, we can ask how these can be used in acquiring word order. Indeed, we can ask more generally, how grammar acquisition can be aided by perceptual mechanisms.

9.5 Perceptual mechanisms in grammar acquisition

One of the basic requirements of any theoretical proposal of grammar is that the structures proposed be acquirable by the infant. There is however a gap between the knowledge acquired and the richness of the input available to the infant, the so-called Plato's problem (Chomsky 1984). Any theory of first language acquisition must account, among other things, for the fact that some of the basic grammatical properties appear to be acquired at the prelexical stage. However, there are not many proposals as to the precise learning mechanisms that are responsible for the early acquisition of the major syntactic properties.

The problem of grammar acquisition is eased if infants have access to, and can utilize, cues in the input that robustly mark specific syntactic properties of the language. A first step would be to show that the grammatical rules to be acquired have distinct, perceptually accessible correlates in fluent speech. Therefore, understanding basic human perceptual capacities can contribute to understanding both the nature of the language competence and how it is acquired.

With this in mind, we can now understand how phonological phrases can be used to acquire word order.

9.5.1 *Using phonological phrases to acquire word order*

Given both inter-linguistic (Turkish and French) and intra-linguistic (German) evidence for different physical correlates of phonological phrase prominence depending on the relative order of head and complement, it is feasible that this may be exploited in first language acquisition.

Experimental evidence shows that 6–12-week-old infants discriminate French from Turkish exclusively on the basis of phonological phrase stress

(Christophe et al. 2003). The material for that experiment consisted of French and Turkish utterances containing two phonological phrases. The utterances were well matched as to the number of syllables and were resynthesized so that discrimination between the two languages could not be based on phonemic or phonetic segmental properties.

The results suggested that infants react to the difference between (10a) and (10b).

(10) a. . . . #wsws . . .
b. . . . #swsw . . .

Notice, however, that if there are no boundaries (#), it is impossible to know whether in a sequence . . . swswswsws . . . the strong node is initial or final in its phrase. Thus the signal must contain some cues to the location of the constituent edges.

The different manifestation of prominence based on the Iambic–Trochaic Law may be one such cue and thus account for infants' ability to discriminate the two languages. That is, even in the absence of knowledge of phrase boundaries, the acoustic nature of prominence—whether more intense and higher in pitch or not—can indicate whether the strong node is initial or final.

That is, to an infant, an utterance may appear as a series of weak syllables punctuated by perceptually prominent, strong syllables. Depending on the acoustic properties of such prominent syllables, the infant can infer the word-order properties of the language.

Not only are there such differing acoustic cues to the two kinds of prominence in phonological phrases that differ in word order, but it has also been shown that infants are sensitive to such cues. Several authors have shown that infants can discriminate relative pitch levels and can memorize the pitch and temporal characteristics of auditory material (e.g., Trehub 2003). Indeed, Krumhansl and Jusczyk (1990) and Jusczyk and Krumhansl (1993) found evidence that 4.5-month-old infants organize tone sequences such that tonal phrases start with a high pitch and end with a long duration.

Note that all this data is still correlational. That is, although existing evidence shows that (a) specific acoustic patterns mark specific word orders at the phonological phrase level, (b) that infants are sensitive to such acoustic cues, and (c) infants can parse phonological phrases in speech, it has not yet been established if infants indeed use such acoustic cues to determine word order.

What kinds of data would be required to support our claim? First, we are examining corpora of child-directed speech in different languages to see if the correlation between different types of acoustic cues and different word orders

is indeed reliable in naturalistic input. The stimuli used in Nespor et al. (2008) are a small but carefully constructed set, to best reveal acoustic differences that mark phrases with different word orders. Through the analysis of corpora we will be able to establish the robustness of these findings in speech samples that more closely approximate the input to the infant.

In addition, we can ask experimentally if infants can use prosody in an online task to figure out word order. For example, if we show an unfamiliar transitive action performed by an unfamiliar character, and accompany such a video with a phrase like 'wug pilk', then depending on the prosody, this could be interpreted as either Verb–Object or Object–Verb. Consequently, while the input is ambiguous (two words and two novel referents—an object and an action), we predict that prosody will bias infants towards one interpretation or another. In particular, a trochaic phrasal stress will bias infants towards an Object–Verb reading, while an iambic phrasal stress will bias them towards a Verb–Object interpretation.

9.5.2 *The Iambic–Trochaic Law as a general perception mechanism*

The Iambic–Trochaic Law has been primarily expressed in music and other non-linguistic, acoustic stimuli. However, recent work suggests that this might be a more domain-general perceptual mechanism. Bion et al. (submitted) carried out experiments on sequences of syllables and Peña et al. (in progress) on visual stimuli. The auditory experiments were based on pitch and duration. Intensity was not investigated, since, at least for the acoustic modality, there is ample evidence that intense stimuli are perceived as group-initial (Hay and Diehl 2007). Bion et al. habituated adults with sequences of syllables alternating either in pitch or in duration, and subsequently tested them with pairs of acoustically flat syllables. Participants were found to significantly prefer pairs that in habituation had either higher pitch initially or longer duration finally. That is, the Iambic–Trochaic Law is responsible for segmentation and storage in memory.

9.5.3 *Other perceptual mechanisms for grammar acquisition*

Although we suggest that the Iambic–Trochaic Law can lead the infant to acquiring word order, it is plausible that there are other cues perceptually available in the speech stream that can aid in discovering word order. It is, in fact, desirable that multiple cues lead the infant to converge on the same structure. More generally, perceptual mechanisms might play a role in several aspects of grammar acquisition.

For example, it has been known for several years that infants are capable, at birth, of discriminating their native language from other languages, and also discriminating two foreign languages, when these come from different rhythmic classes (e.g., Moon, Cooper, and Fifer 1993; Nazzi, Bertoncini, and Mehler 1998; Ramus et al. 2000). These early discrimination capacities are well correlated with the amount of vocalic space per utterance (as a percentage of time of the total utterance occupied by vowels, %V) in languages from different rhythmic classes (Ramus, Nespor, and Mehler 1999). Given that differences in %V correspond to the complexity of syllable structures in a language, it has been proposed that a sensitivity to %V might bias infants towards different segmentation strategies (Mehler and Nespor 2003; Nespor et al. 2003).

More recently, it has been proposed that %V also provides an indirect cue to syntax—low %V languages tend to be head–complement, while high %V languages tend to be complement–head (Shukla, Nespor, and Mehler in preparation). Thus this cue, perceived already by newborns, could give a bias as to the order of words in the language of exposure.

Several researchers have explored other aspects of perception that are salient for infants (and adults). It has been proposed that perceptual primitives—specialized mechanisms that render certain kinds of stimuli highly salient—place strong constraints on the kinds of inferences that infants draw from a given stimulus (Endress, Scholl, and Mehler 2005; Endress, Dehaene-Lambertz, and Mehler 2007; Endress and Mehler under review). Such perceptual primitives allow participants to detect both immediate repetitions and configurations at the edges of constituents.

For example, both adults (Endress et al 2005; Shukla, Nespor, and Mehler 2007) and infants (Seidl and Johnson 2006) have been shown to be better at extracting word-like units or learning rules over syllables at the edges of phrases as compared to their middles. Further, Gervain, Macagno, et al. (2008) have shown that neonates can more quickly process 'words' that contain immediate syllable repetitions.

All these studies are thus aimed at investigating basic perceptual mechanisms that infants bring to the task of language acquisition. These mechanisms can then be invoked in order to build better models of grammar acquisition (see also, e.g., Slobin 1973; Morgan and Demuth 1996; Dupoux and Peperkamp 2002).

9.5.4 *Distributional cues and acquisition*

The proposal that different perception mechanisms contribute to make the identification of word order possible, does not exclude that cues of a different

nature also contribute to signaling word order, independently of the knowledge of words. For example, several researchers have shown that young infants can track distributional properties of syllables in order to extract word-like units from fluent speech (Saffran et al. 1996; Aslin et al. 1998; Peña et al. 2002).

Further, on the basis of an artificial grammar experiment in which frequent items alternate with infrequent ones, it has, in fact, been shown that 7-month-old Japanese and Italian infants have opposite order preferences (Gervain, Nespor, et al. 2008). That is, infants appear to be able to keep track of the most frequent words in the input as well as their location with respect to the less frequent words. Interestingly, the preferred word order corresponds to that of their native language: frequent–infrequent for Italian and infrequent–frequent for Japanese.

Thus signals of different nature might reflect important properties of the syntax of a language. And infants at the prelexical stage may have at their disposal different learning mechanisms based on perception, on the ability to compute distributional regularities and other linguistic biases, that aid them into the acquisition of basic grammatical properties (see, e.g., Peperkamp et al. 2006).

9.6 Conclusions

Although generative theories of grammar provide a substantial understanding of the mature state of the human language faculty, *how* the rules of grammar are acquired has been a much harder question to answer. Since the discovery, on the one hand, that even newborns are very sensitive to the rhythmic properties of language (Mehler et al. 1987, 1988) and, on the other hand, that young, prelexical infants are capable of computing fairly sophisticated distributional properties of speech (e.g., Saffran et al. 1996), the precise mechanisms that are responsible for language acquisition have been at the center of attention. In fact, the notion that such rhythmic and distributional properties might aid in acquiring language—either grammatical or lexical properties—have played a large role in the latest theories of acquisition. In this chapter, we draw attention to another potential source of information that relies on basic perceptual capacities of infants.

Spoken language, that constitutes the primary linguistic evidence for the infant, comprises a set of hierarchically organized phonological constituents that are not in direct correspondence with syntactic units (Selkirk 1984; Nespor and Vogel 1986/2008). One level of the hierarchy, the phonological phrase, differs amongst languages in accordance with the underlying syntax.

In particular, the prominence pattern of phonological phrases reflects the direction of branching in the language—initial prominence for left-branching (or complement–head) languages, and final prominence for right-branching (head–complement) languages (Nespor and Vogel 1986/2008). Infants appear to be sensitive to the different rhythmic patterns at this level of the hierarchy (Christophe et al. 2003). In addition, we find that the acoustic correlates of the two kinds of prominence patterns differ systematically; if prominence is initial, it is marked primarily by higher pitch and intensity, if final, primarily by duration (Nespor et al. 2008).

Indeed, these differing patterns for initial and final elements of a 'phrase' are proposed to be a basic perceptual Gestalt—the Iambic–Trochaic Law. Therefore, innate perceptual capacities might well direct the infant exposed to a given language to the correct word order of that language without any recourse to lexical information. Thus, we may account for the empirical observation that the earliest productions of infants respect the word order of their language.

More generally, we suggest that early perceptual capacities of infants can provide substantial information that aids the infant in acquiring grammar. Such perceptual capacities, coupled with distributional learning, might place strong constraints on the possible grammars, thus substantially easing the problem of language acquisition.