Word frequency and prosody bootstrap basic word order in prelexical infants

Judit Gervain¹2

¹DPSS, University of Padua, Padua, Italy
²INCC, CNRS & Université de Paris, Paris, France

Abstract

Languages systematically vary in their basic word order, which infants need to learn as they acquire their native language. Here I present evidence for the prosodic bootstrapping of word order. Specifically, two behavioral experiments and one brain imaging study word are reviewed supporting the hypothesis that word frequency and phrasal prosody serve as powerful cues to help infants bootstrap the basic lexical categories of functors and content words and guide infants about the relative order of these two categories in their native language. The acoustic realization of prosodic prominence in phonological phrases correlates with basic word order, as functor-initial languages typically rely on phrase-final lengthening, while functor-final language on phrase-initial pitch and/or intensity rise (Nespor et al. 2008). The first study shows that 8-month-old infants can use this acoustic cue to determine the word order of an artificial language. A second study shows that infants expect this prosodic information to be aligned with word frequency, i.e. frequent words to be prosodically non-prominent, as are natural language functors. A near-infrared spectroscopy imaging study suggests that sensitivity to the acoustic realization of prosodic prominence and the resulting rhythmic (iambic/trochaic) grouping derives from babies’ prenatal experience with speech.

Index Terms: language acquisition, prosodic bootstrapping, word order, infants

1. Introduction

Languages systematically vary in their basic word order. In many languages like English or Italian, functors typically precede content words, e.g. English: in the house; Italian: nella casa. In other languages, like Turkish or Basque, functors typically follow content words, e.g. Basque: garren azean flame behind ‘behind the flame’; Turkish: evde house.in ‘in the house’. To be able to understand and produce multiword utterances, infants thus need to learn which order characterizes their mother tongue. This is a crucial early step in acquiring the native language, and because the relative order of function and content words correlates with other word order phenomena (e.g. the order of verbs and objects [1]), discovering functor-directionality may be a useful first cue to bootstrap other aspects of grammar.

The paper provides empirical evidence for the hypothesis that phrasal prosody, in conjunction with word frequency, is a powerful cue that pre-verbal infants as young as 8 months of age rely on to bootstrap the relative order of functors and content words in their native language.

The prosodic bootstrapping of word order relies on the fact that in natural languages, correlations exist between perceptually available properties of words such as their frequency or phonological realization and their grammatical function. Specifically, the distinction between content words and function words is universal in natural languages [2], [3]. Content words such as nouns (chocolate, rainbow etc), verbs (run, feel etc), or adjectives (pretty etc) carry lexical meaning, while functors such as prepositions (on, in etc.) or pronouns (you, they etc.) indicate grammatical relations. The two categories differ not only in their functions, but also in their frequency and phonological properties. Individual function words are thus typically more frequent than individual content words [4], [5]. Also, functors tend to be phonologically minimal, while content words are typically heavy [6]. The exact features in which they differ vary across languages, but functors are typically unstressed, have simpler syllable structure, reduced vowels, shorter duration etc. Content words, by contrast, receive lexical (and phrasal) stress, have a greater number of and more complex syllables etc. The first tenet of the bootstrapping hypothesis is that these two perceptually available features help infants identify at least some of the functors and some of the content words heard in the input.

Furthermore, functor-directionality, i.e. whether a language is functor-initial or functor-final, correlates with the location and realization of phrase-level prosodic prominence [7], [8]. As prominence is carried by the content words, it occurs phrase-finally in functor-initial languages, but phrase-initially in functor-final languages. Importantly, phrasal prominence is realized through increased duration in functor-initial languages (English: in Rayne), but through increased pitch or intensity in functor-final languages (Japanese: “Tokyo ni Tokyo in “in Tokyo”). Thus the acoustic cues indicating prosodic prominence correlate with functor directionality. It is important to note that the three acoustic cues, duration, pitch and intensity, are all modulated in most languages. However, the one that is contrastive and thus indicates prosodic prominence at the phrasal level varies as a function of word order, i.e. functor-directionality. This has been empirically demonstrated in a variety of languages, such as French, Turkish, English, Korean, Japanese, Farsi etc. [7], [8]. The second tenet of the prosodic bootstrapping of word order argues that infants are sensitive to this prosodic cue, and can use it to select between functor-initial and functor-final word order.

2. Prosodic bootstrapping of functor directionality

Two behavioral and one brain imaging (near-infrared spectroscopy, NIRS) study will be presented as empirical support for the prosodic bootstrapping hypothesis.

The behavioral studies are based on the same paradigm, an artificial grammar learning study, which shows that infants as
young as 8 months of age use word frequency to establish the categories of functors and content words, as well as their relative order [5]. In the two behavioral studies presented below, prosody is added to the artificial grammar.

The basic paradigm (Figure 1) consists of a continuous speech stream in which frequent and infrequent words strictly alternate, mimicking functors and content words, respectively, in natural language. The beginning and end of the stream are ramped in amplitude, suppressing cues about the first and last words. This manipulation renders the stream structurally ambiguous between a parse that is frequent word initial and a parse that is frequent word final. Infants listen to this stream for 4 minutes during familiarization. In the test phase, infants are then presented with 4-syllable-long sequences that begin or end with a frequent element. Importantly, test items are always taken from the familiarization stream, so both frequent-word-initial and frequent-word-final test items are familiar. If infants show a preference for one word order over the other, this preference derives from knowledge the infants bring to the task.

<table>
<thead>
<tr>
<th>A) Lexicon</th>
</tr>
</thead>
<tbody>
<tr>
<td>X: (ru, pe, du, ba, lo, de, pa, ra, lo)</td>
</tr>
<tr>
<td>Y: (mu, ri, ku, bi, do, ka, na, ro)</td>
</tr>
<tr>
<td>B) Familiarization</td>
</tr>
<tr>
<td>frequent-initial</td>
</tr>
<tr>
<td>…AXBYAXBYAXYA…</td>
</tr>
<tr>
<td>…getoflubegoflucogafmogagefo…</td>
</tr>
<tr>
<td>frequent-final</td>
</tr>
<tr>
<td>…getoflubegoflucogafmogagefo…</td>
</tr>
<tr>
<td>OR</td>
</tr>
</tbody>
</table>

C) Test Items

<table>
<thead>
<tr>
<th>frequent-initial</th>
<th>frequent-final</th>
</tr>
</thead>
<tbody>
<tr>
<td>filogebei</td>
<td>bagebofi</td>
</tr>
<tr>
<td>firugenu</td>
<td>kafipage</td>
</tr>
<tr>
<td>gedoflida</td>
<td>kufiduge</td>
</tr>
<tr>
<td>gerifipoe</td>
<td>ragenafi</td>
</tr>
</tbody>
</table>

![Figure 1: The basic artificial grammar paradigm from [5]](image)

This task was used to show that by 7-8 months of age, infants track the frequency distribution of the elements in the linguistic input, and prefer to listen to test items that follow the word order of their native language. Specifically, Japanese infants, exposed to a functor-final language, prefer items with the frequent-final word order, while English, French and Italian babies (all functor-initial languages) prefer the frequent-initial word order [5], [8], [9]. This paradigm thus shows that across multiple different languages, infants rely on word frequency to bootstrap word order.

However, this may not always be sufficient. For instance an English-Japanese bilingual infant encounters both frequent word initial and frequent word final patterns in their input, from English and Japanese, respectively. How can they learn to distinguish the two? Phrasal prosody provides a powerful cue. As shown above, the position and acoustic realization of prominence in phonological phrases correlates with word order.

### 2.1. Infants can use prosody to bootstrap word order

Can infants use this prosodic correlate to bootstrap word order? This was tested in a behavioral study with bilingual infants exposed to both a functor-initial and a functor-final language [8]. The functor-initial language was English, the functor-final language was one of Japanese, Korean, Hindi or Farsi.

Two groups of bilinguals were exposed to a prosodically enriched variant of the artificial grammar. One group heard the familiarization stream with a durational contrast typical of functor-initial languages. Thus the infrequent words, mimicking content words, were longer (288 msec) and thus more prominent, while the frequent words were shorter in duration (240 msec). For both words, pitch was constant at 200 Hz. The other group heard the stream with the typical frequent-final pitch contrast. In this case, pitch was higher (224 Hz) on the prominent infrequent words and lower (200 Hz) on the non-prominent frequent words, both with a constant duration of 240 msec. Both groups were tested on the same test items as in the original paradigm, with no modulation of prosody, i.e. of duration or pitch. Unlike in the original paradigm, where the familiarization was ambiguous about word order, here it provided a clear cue. If, as the prosodic bootstrapping hypothesis predicts, infants are sensitive to phrasal prosodic prominence as a cue, they should show a word order preference in the test.

This is exactly what was observed (Figure 2). Bilinguals exposed to the durational contrast prosody (n = 12) showed a preference for frequent word initial items, while bilinguals (n = 12) exposed to the pitch contrast prosody looked longer to frequent word final items. This provides evidence that bilinguals can rely on prosody to disambiguate between the two word orders present in the linguistic input.

![Figure 2: Looking times observed for the prosodically enriched conditions in [8] (OV: frequent word final, pitch contrast prosody; VO: frequent word initial, durational contrast prosody; light grey: looking times to frequent word final items; dark grey: looking times to frequent word initial items)](image)

### 2.2. Infants expect prosody and frequency to be aligned at the word level

One question that arises is whether infants process prosody and word frequency in a unified manner, forming a combined representation or separately, on two independent tiers. In other words, is it enough for them to hear a certain type of prosody to bootstrap word order independently of the frequency information or do they actually expect the two cues to be aligned, i.e. do they expect frequent words to be non-prominent and infrequent ones to be the ones carrying prominence, as is the case in natural languages with functors and content words, respectively?

To answer this question, two artificial grammars were used [10]. One was the same as the durational contrast condition in the previous study, with longer infrequent and shorter frequent words in the familiarization stream. In the
second, frequency and prosody were misaligned: now frequent words were more prominent, i.e. longer, and infrequent ones shorter. This does not happen in natural languages, as function words are phonologically more minimal. However, if infants don’t combine the two cues at the word level and process them independently, then they may still perceive the durational contrast prosody as cueing a frequent word initial order.

French-exposed 8-month-old infants were tested, and as French is a functor-initial language with durational-contrast based prosody, infants were expected to show a frequent word initial preference in the first, aligned condition (n = 20). The question was whether they do so in the misaligned condition (n = 22).

The results (Figure 3) indeed confirmed a strong preference for frequent word initial items in the aligned condition, while no preference was found for the misaligned condition, showing that infants expect prosody and frequency to be coherent at the word level, conforming to the natural language categories of functors and content words.

![Figure 3: Looking times observed for the aligned and misaligned conditions in [10] (light grey: looking times to frequent word final items; dark grey: looking times to frequent word initial items)](image)

### 2.3. Expectations about prosodic grouping are based on prenatal experience

When do infants discover the prosodic patterns of their native language? For prosody to be a useful bootstrapping cue, sensitivity to it needs to emerge early. As prosody is already perceived in utero [11], [12], it may be the case that infants learn about the prosodic groupings typical of their native language even before birth. This hypothesis was tested in a NIRS brain imaging study with newborn infants (age 1-4 days) [13].

Newborns were exposed to one of the three relevant acoustic contrasts, duration, pitch or intensity. Pairs of pure tones were generated to either have a prominence-initia or a prominence-final grouping. Thus in the durational contrast condition, infants listened to sequences of short-long tone pairs (prominence-final, iambic grouping, typical of functor-initial languages) and sequences of long-short (prominence-initial, trochaic grouping). In the pitch contrast condition, infants listened to sequences of high-low (prominence-initial, trochaic, typical of functor-final languages) and low-high (prominence-final, iambic) tone pairs. In the intensity condition, they heard loud-soft (prominence-initial, trochaic, typical of functor-final languages) and soft-loud (prominence-final, iambic) pairs.

A large body of literature on the Iambic-Trochaic Law [7], [14], [15] suggests that adults and under some conditions older infants show a preference for the iambic, prominence-final grouping in the durational condition, as is often found in the phrasal prosody of functor-initial languages, while the trochaic, prominence-initial grouping is preferred for the pitch and intensity cues, as is the case in the phrasal prosody of functor-final languages.

Newborns prenatally exposed to French were tested in all three conditions in three separate groups. A fourth group of infants were tested in the pitch condition. These infants were prenatally exposed to two languages, French as well as another language that relied on a pitch contrast to mark prosodic prominence. In French, prominence in phrasal prosody being marked mainly by a durational contrast, monolingual French infants had prenatal experience with the durational cue, but less so with the pitch and the intensity cue, while the bilingual group had prenatal experience with pitch.

French newborns’ brain responses were higher for the dispreferred trochaic grouping than for the predicted iambic grouping in the durational condition (surprise response), whereas no difference in brain response between the iambic and trochaic groupings were observed for the pitch and intensity conditions. Interestingly, however, for the very same pitch stimuli for which monolingual French-exposed newborns showed no preferential brain responses, bilingual newborns prenatally exposed to a language with pitch-contrast based phrasal prosody showed the predicted differential response, with greater activation to the dispreferred iambic than for the preferred trochaic grouping.

Thus only the two groups who had prenatal experience with the given cue (monolingual French newborns for duration and bilingual newborns for pitch) showed brain activity consistent with a grouping preference. These results suggest that the prosodic patterns of the native language, at least at the phrasal level, may already be acquired prenatally.

### 3. Discussion

Three studies have been presented to support the hypothesis that infants rely on phrasal level prosody to support the acquisition of basic functor – content word order even pre-verbally, as early as 8 months of age.

Developmentally, this bootstrapping mechanism relies on infants prenatal experience with speech, which mainly consists of prosody, given the filtering properties of the intra-uterine environment. Newborns are thus already sensitive to whether their native language has iambic or trochaic prosodic groupings.

Such a sensitivity then lays the foundations for the subsequent acquisition of the basic word order of functors and content words, as the location and physical realization of phrase-level prosodic prominence correlates with word order.

It is important to note that this bootstrapping mechanism relies on the physical / acoustic properties of the speech stream. Thus they do not require abstract, linguistic knowledge on the part of the learner, making it a feasible bootstrapping strategy.
In conjunction with word frequency, another simple and non-language-specific cue, prosody can thus contribute to a very early bootstrapping of a basic property of grammar, even before a sizeable lexicon or linguistic production is in place. Moreover, this cue is useful even in challenging language learning situation such as bilingualism.

4. Conclusions

Prosodic bootstrapping is a powerful perceptual mechanism allowing infants as early as 8 months of age to start acquiring the basic word order of their native language.

5. Acknowledgements

This work was supported by the European Research Council (ERC) Consolidator Grant nr. 773202.

6. References