Available online at www.sciencedirect.com
ScienceDirect
Cognitive Psychology 57 (2008) 56–74
Cognitive Psychology
www.elsevier.com/locate/cogpsych

Bootstrapping word order in prelexical infants: A Japanese–Italian cross-linguistic study ☆

Judit Gervain [a,*], Marina Nespor [b], Reiko Mazuka [c,d], Ryota Horie [c], Jacques Mehler [a]

[a] *SISSA—Scuola Internazionale Superiore di Studi Avanzati Cognitive Neuroscience Sector, Via Beirut 2–4, 34014 Trieste, Italy*
[b] *University of Milano Bicocca, Milan, Italy*
[c] *Brain Science Institute, RIKEN, Tokyo, Japan*
[d] *Duke University, Durham, USA*

Accepted 7 December 2007
Available online 1 February 2008

Abstract

Learning word order is one of the earliest feats infants accomplish during language acquisition [Brown, R. (1973). *A first language: The early stages*, Cambridge, MA: Harvard University Press.]. Two theories have been proposed to account for this fact. Constructivist/lexicalist theories [Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74*(*3*), 209–253.] argue that word order is learned separately for each lexical item or construction. Generativist theories [Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.], on the other hand, claim that word order is an abstract and general property, determined from the input independently of individual words. Here, we show that eight-month-old Japanese and Italian infants have opposite order preferences in an artificial grammar experiment, mirroring the opposite word orders

☆ This work was supported in the frame of the CALACEI NEST PATHFINDER initiative 'What in means to be human', by a program grant from the J. S. McDonnell Foundation and by the PRIN 2005 'General and specific mechanisms in the acquisition of recursivity'. The Japan corpus and infant experiments were funded by RIKEN Brain Science Institute. Judit Gervain acknowledges the financial support of the Central European Initiative (CEI) Research Fellowship. We thank Luca Bonatti for his comments on an earlier version of the manuscript, Alessio Isaja and Luca Filippin for technical support, Jeanne Moussu and Marijana Sjekloća for their help in testing the Italian infants, Kyoko Shirasawa-Potapov, Mihoko Hasegawa, Mahoko Yamaguchi, and Mits Ota for their assistance in testing the Japanese infants, and Ken'ya Nishikawa for providing information about the Japanese infant-directed corpus.
* Corresponding author. Fax: +39 040 3787615.
E-mail address: gervain@sissa.it (J. Gervain).

0010-0285/$ - see front matter
doi:10.1016/j.cogpsych.2007.12.001

of their respective native languages. This suggests that infants possess some representation of word order prelexically, arguing for the generativist view. We propose a frequency-based bootstrapping mechanism to account for our results, arguing that infants might build this representation by tracking the order of functors and content words, identified through their different frequency distributions. We investigate frequency and word order patterns in infant-directed Japanese and Italian corpora to support this claim.

Keywords: Language acquisition; Bootstrapping; Word order; Japanese; Italian; Prelexical infants; Infant-directed speech; Corpus; Headturn preference procedure

1. Introduction: Word order in language acquisition

Learning word order is one of the major tasks during early language acquisition. Indeed, young learners exhibit some knowledge of basic word order from their earliest multiword utterances (Brown, 1973; Guasti, 2002). In the present paper, we investigate what cues and learning mechanisms infants might use to acquire this structural property of their mother tongue.

There exist at least two main theoretical stances about how word order is acquired. The constructivist (or lexicalist) view (Chang, Lieven, & Tomasello (in press); Tomasello, 2000) holds that word order is learnt from frequently encountered examples in the input. According to this view, young learners are sensitive to co-occurrence statistics in the language they hear, and their early competence contains semi-abstract constructions derived from this statistical information. For instance, from frequent occurrences of *Can you see...?, Can you go...?, Can you eat...?*, the infant might construct the semi-general frame *Can you X...?*, where X is a placeholder for possible substitutions, in this case, for certain verbs. Thus, this view claims that young learners have no general and fully abstract representations of syntactic structure, including word order. Rather, their knowledge is linked to specific lexical items or frames. This view, then, implies that learning word order proceeds together with or after acquiring an initial lexicon, but not before.

The generativist view (Chomsky, 1995; Guasti, 2002), on the other hand, argues that language acquisition relies on abstract prewired structural representations. Some of these hold universally true for all languages (‘principles’) and thus need not be learned. Others (‘parameters’) specify choices between possible structures, and languages vary as to which of the specified options they implement (Rizzi, 1986). Language acquisition, then, amounts to setting the parameters to the value that characterizes the target language, using overtly available cues in the input. For instance, the Head–Complement parameter formalizes whether languages choose to place the Head of a syntactic phrase first, and its Complement second; or the other way round. Japanese and Turkish, for instance, are Complement–Head languages. This entails that they have Object–Verb (OV) order (1a), postpositions (1b), complementizers that follow their subordinate clause (1c), just to mention a few phrase types. Head–Complement languages, like English and Italian, on the other hand, have VO order (2a), prepositions (2b), and complementizers that precede the subordinate clause (2c).

(1) Japanese
  a. Taroo ga tegami o kaita. Taroo.nom letter.acc wrote 'Tagoo wrote a letter.'
  b. Kobe ni, Kobe to 'to Kobe'
  c. Mary ga [John ga hon o yon da to] omottei ru Mary.nom John.nom book.acc read.past that think.pres 'Mary thinks [that John read a book].'

(2) Italian
  a. mangiare una melaeat.inf an apple 'to eat an apple'
  b. sul tavolo on-the table 'on the table'
  c. Credo che piova believe.1sg that rain.3sg.subjunctive 'I think it is raining.'

To learn word order, the infant has to bootstrap the correct setting of the Head–Complement parameter (and other word order parameters) on the basis of surface cues found in the input, e.g. prosodic patterns (Mazuka, 1996; Nespor et al., under review). Word order is thus acquired independently of the lexicon.[1] Therefore, there is no logically necessary chronological ordering between them.

In the light of the above, the litmus test to decide between the two theories is to investigate whether infants have any general knowledge of word order prior to learning their words. By comparing Japanese and Italian eight-month-olds, we tested whether they already show signs of language-specific word order preferences. If such language-related differences exist at this prelexical age, then word order is necessarily learned independently of lexical items, arguing for a generativist view. In general, beyond this concrete theoretical debate, our study aims at exploring the acquisition of word order in the pre-linguistic stage. Investigating this early period might contribute to identifying the very first surface cues that get syntax started. Furthermore, it can provide evidence about the schedule of syntax acquisition by determining at what point in development infant populations learning different languages start to diverge and show language-specific structural knowledge for the first time.

If we find evidence for a prelexical word order representation, it requires an explanation in terms of the surface cues that trigger it. In particular, we need to identify (i) a reliable surface cue that even prelexical infants can detect, (ii) which correlates well with the targeted abstract structural property. We propose that the relative position of functors (grammatical elements like *the, of, his* etc.) and content words (elements carrying lexical meaning, like *dog, eat, good* etc.) provides learners with precisely such a cue. Indeed, the distinction of functors and content words is a universal property of all languages in the world (Chomsky, 1995), so it can be used to bootstrap word order in any language. Therefore, we need to find a surface cue that reliably distinguishes the two categories. We propose that one such possible cue is frequency. Moreover, we need to show that the relative order of functors and content words correlates well with the general basic word order in a given language.

In what follows, we will evaluate each of these assumptions in a cross-linguistic context, comparing Japanese and Italian. In Sections 2 and 3, we lay the groundwork by examining

[1] In certain formulations of generativist acquisition theories (e.g. Pinker, 1984), infants are assumed to know some vocabulary before the acquisition of syntax begins. However, even in these accounts, the knowledge of the lexicon is a background assumption, rather than a logical necessity.

the properties of infant-directed Japanese and Italian in order to provide evidence that functors are indeed distinguishable from content words by their higher frequency of occurrence, and that the opposite word orders of the two languages are reflected in the different sentential positions of the two categories. In Section 4, we report our main experimental finding showing that eight-month-old Japanese and Italian infants exhibit opposite word order preferences in an artificial grammar learning situation, mirroring the different orders of functors and content words found in their native languages. In Section 5, we discuss the results and elaborate our hypothesis further by clarifying some theoretical issues. In Section 6, we summarize our findings and outline their relevance for the broader framework of language acquisition.

2. Experiment 1: Distinguishing functors and content words

A major functional distinction in language is the division of labor between functors, which signal syntactic and morphological structure, and content words, which carry lexical meaning. Functors are organized into closed classes with a few elements (e.g. articles: *a, an, the*; personal pronouns: *I, you, he, she, it, we, you, they*; prepositions: *of, on, up* etc.), and do not tolerate the introduction of new items unless considerable grammatical reorganization takes place. Content words belong to large, open classes (e.g. nouns: *car, dog, baby* etc., verbs: *eat, kiss* etc., adjectives: *beautiful, good* etc.), into which new elements are routinely inserted every day (*xerox, podcast* etc.). A good illustration of this division of labor comes from Lewis Carroll's Jabberwocky poem: '*Twas bril lig, and the slithy toves/ Did gyre and gimble in the wabe*...' The content words are replaced by novel tokens, while the grammatical structure is maintained by leaving functors intact. Although languages differ with respect to the actual grammaticalization of content words and functors, the divide between functors and content words has been shown to be universal (Abney, 1987; Fukui, 1986).

This functional distinction is accompanied by a series of surface differences in the phonological realizations and frequency distributions of the two categories. It has long been observed that functors tend to be shorter and more reduced than content items (Nespor & Vogel, 1986; Selkirk, 1984). Recently, Morgan, Shi, and Allopenna (1996) conducted a systematic cross-linguistic study to explore the perceptual differences between functors and content words, comparing infant-directed English and Mandarin Chinese on a series of phonological and acoustic measures (e.g. number of syllables, syllable complexity, diphthongization, vowel duration, amplitude etc.). They found that functors were more reduced than content words on all of these measures. Thus, they concluded that functors are perceptually minimal, while content words are not.

Indeed, in later work, Shi, Werker, and Morgan (1999) showed that even newborns are able to distinguish the phonological cues correlated with the two categories. At 6 months of age, infants start to show a preference for content words (Shi & Werker, 2001), but by 13 months, they are also able to represent functors in some phonological detail (Shi, Cutler, Werker, & Cruickshank, 2006a; Shi, Werker, & Cutler, 2006b).

Phonological minimality, however, is implemented by different phonological properties in different languages. In English, for instance, function words have reduced vowels (*the* [ðə], *of* [əv] etc.), and start with consonants that rarely appear at the beginning of content words (*th*-, *wh*-), while in Hungarian, there is no vowel reduction, and the set of consonants used in functors (*m*-, *h*-) frequently occur in content words, too. Rather, in Hungarian, functors differ from content words in having fewer syllables. Therefore, while the minimality of functors is universal, their differentiating features are language-specific, and thus require some familiarity with the language.

In contrast, it has been suggested that, given their grammatical role, functors universally have a much higher token frequency than content words. This was confirmed for English in several corpus studies. For instance, Cutler and Carter (1987) and Cutler (1993) report that functors made up 59% of the word tokens of their corpus, while they constitute only about 1% of all the word types, i.e. the lexical entries of English. Moreover, it has been observed that in English, there is little overlap in the frequency distributions of functors and content items. In Kucera & Francis's classical study (1967), the 50 most frequent lexical items were found to be function words. Therefore, if universally valid, frequency alone might be a sufficient heuristic predictor of category membership. Below, we test this assumption on our Japanese and Italian infant-directed corpora.

2.1. Methods

2.1.1. Corpora

For Japanese, we made use of the corpus of infant-directed Japanese collected at the Laboratory of Language Development, Brain Science Institute, RIKEN (Mazuka, Igarashi, & Nishikawa, 2006). For our purposes, we extracted 22 mothers' utterances addressed to their infants during free play or directed story-telling (using specific story books), but we excluded their conversations with adults, e.g. the experimenter. Our corpus thus comprises 14,958 utterances, made up of 47,071 word tokens[2], falling into 5205 word types[3] Utterances were phonologically transcribed.

For Italian, we used the adult utterances of the Italian language subcorpora (Antelmi, n.d.; Antinucci & Parisi, 1973; Cipriani et al., 1989; Tonelli, n.d.; Volterra, 1976; Volterra, 1984) of the CHILDES database (MacWhinney, 2000). These subcorpora contain recordings of free adult–infant interactions in a variety of conditions (dyadic and group contexts, home as well as institutional environments, typically and atypically developing children etc.). The corpus comprises 51,489 utterances, made up of 233,137 word tokens, falling into 9538 word types. Utterances were phonologically transcribed.

2.1.2. Measures

We calculated the frequency ranks (Zipf, 1935) of the word types, and counted the number of functors and content words among the 100 most frequent words. We also computed 'overall coverage', i.e. what percentage of the corpora is covered by the most frequent functors and content words.

[2] We use the token/type distinction as customary in computational linguistics. Word tokens are individual occurrences of word types, i.e. a given word form. For example, the sentence "Rage, rage against the dying of the light" (Dylan Thomas: *Do not Go Gentle into That Good Night*) is made up of 8 word tokens, falling into 6 word types. Word types "rage" and "the" have 2 tokens each, the other types have one token.

[3] In accordance with our purposes, as well as Japanese orthographic tradition, grammatical particles and markers are counted as separate words.

2.2. Results

Fig. 1 illustrates the frequency distributions of the 100 most frequent words in the two languages. In Japanese, this list contained 57 functors and 43 content words. In Italian, the list had 67 functors and 33 content words. Importantly, as expected, in the highest frequency range, the distributions of the two categories are non-overlapping. Indeed, the most frequent Japanese content word is 21st in the rank (*Hora?* 'See?'), followed by only four other content words in the first third of the distribution (*So.* 'I see', *I!* 'Good! Great!', *mi* 'to see', *ko* 'child'). The most frequent Italian content word (*Guarda!* 'Look!') is 14th in rank, and it is followed by only two more content words in the first third of the distribution (*Fa!* 'Do!' and *mamma* 'Mum'). Note, in addition, that these content words are not used in a genuinely referential way. Rather, they function as phatic or discursive elements, e.g. forms of address (*mamma*, *ko*) and interjections (*Guarda!, So.* etc.).

We also calculated the cumulative token frequencies or 'overall coverage' of the first 100 most frequent words, i.e. how much of the input they account for. In Japanese, the 100 most frequent words altogether make up 47.45%, i.e. almost half of the corpus. Of this, functors make up 79.73% (corresponding to 37.83% of the whole corpus), content words 20.27% (corresponding to 9.62% of the whole corpus). Functors in the highest frequency range, i.e. the first third, where the distribution of the two categories is almost completely non-overlapping, account for 31.69% of the corpus. In Italian, the 100 most frequent words together account for 61.27% of all word tokens, i.e. almost two thirds of the whole corpus. Of this, functors make up 83.13% (corresponding to 50.93% of the

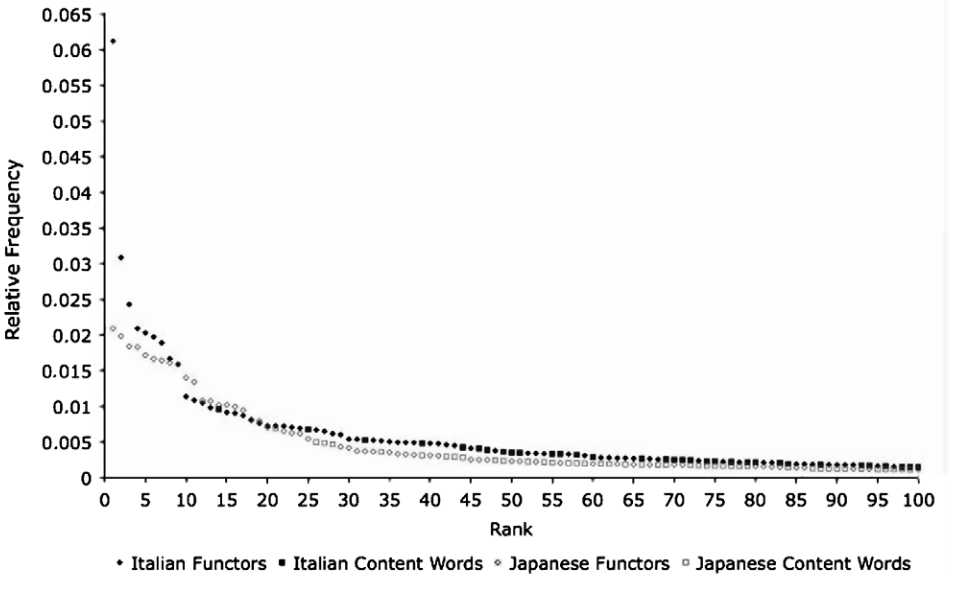


Fig. 1. Histogram showing the frequency distributions of the 100 most frequent words in the Japanese and Italian infant-directed corpora used in Experiment 1. Light grey, empty markers represent Japanese functors (diamonds) and content words (squares). Black, filled markers represent Italian functors (diamonds) and content words (squares). The x-axis corresponds to the rank of a word in the frequency list. The y-axis shows relative frequencies (=absolute frequency/number of word tokens). Absolute frequencies could not be used, because the two corpora are not of the same size.

whole corpus), content words cover 16.87% (corresponding to 10.33% of the whole corpus). Functors in the highest frequency range account for 39.26% of the corpus.

2.3. Discussion

From the above, it is clear that the frequency distributions of functors and content words show different patterns. As expected, individual functors occur more frequently than individual content words. Indeed, the thirty most frequent words are almost exclusively functors both in Italian and in Japanese, and these few functors account for about one third of the entire input that infants are exposed to. Therefore, frequency is a useful heuristic predictor of category membership.

3. Experiment 2: Information about word order in the input

The second premise of our hypothesis is that the input contains information about the order of functors and content words in a form that even young learners can have access to, and that this order correlates with the general word order pattern of the language. We investigate the two corpora to evaluate this premise.

It has long been noted (Morgan et al., 1996) that functors tend to appear at edges of syntactic units. However, languages systematically differ in whether functors come at the left or the right edges of phrases. For example, Japanese, Basque or Turkish have postpositions, whereas English, Italian or French use prepositions. Importantly, it has been extensively documented in language typology (Dryer, 1992; Greenberg, 1963; Mehler, Sebastian Gallés, & Nespor, 2004) that the relative order of functors and content words correlates with a series of other word order phenomena, such as the basic word order of verbs and their objects, the order of complementizers and subordinate clauses, or prepositions vs. postpositions, as illustrated before. It is precisely this empirical observation that is formally captured by the word order parameters of generative grammar.

Nevertheless, since young learners do not know where the boundaries of syntactic units lie within utterances,—this is precisely what needs to be learnt—this general information cannot be used to bootstrap structure. However, there is a special type of syntactic boundary that is available even to infants, namely utterance boundaries (Aslin, Woodward, Lamendola, & Bever, 1996). Therefore, we looked at the occurrences of functors and content words at utterance boundaries. Since Japanese is an OV language, we expected frequent words to appear phrase-, and thus utterance-finally, whereas in Italian, which is a VO language, frequent words were assumed to occur phrase-, and thus utterance-initially. If Japanese and Italian indeed exhibit the opposite patterns, then we can conclude that the order of frequent and infrequent words at utterance boundaries is a useful cue to bootstrap basic word order.

In the light of Experiment 1, functors were operationally defined as frequent words, content words were defined as infrequent words.

3.1. Methods

3.1.1. Corpora

The same corpora as in Experiment 1 served as the basis for this experiment. However, since one-word utterances are not informative about word order, we discarded these, and extracted only the multiword utterances. With this manipulation, we obtained a corpus of 9889 utterances in Japanese and 42 955 utterances in Italian.

3.1.2. Measures

We used the multiword utterances of the corpora to calculate how often frequent and infrequent words appear at initial and final positions at utterance boundaries. Frequent and infrequent words (FW and IW) were defined as having a (relative) frequency of occurrence higher and lower, respectively, than four predefined thresholds: $T_1 = 0.01$, $T_2 = 0.005$, $T_3 = 0.0025$ and $T_4 = 0.001$. T_1 defines 12 words as frequent in Italian, and 20 in Japanese, roughly corresponding to the highest frequency range where only functors appear. T_2 defines 36 words as frequent in Italian, and 34 in Japanese, still corresponding to the frequency ranges where there is little overlap between the distributions of the two categories. T_3 defines 72 words as frequent in Italian, and 63 in Japanese. Finally, T_4 defines 133 words as frequent in Italian, and 144 in Japanese. All other words in the corpora were categorized as infrequent. For further descriptive statistics about the four thresholds, see Table 1. No further thresholds were used, because the words categorized as frequent by T_4 already cover about two thirds of the corpora. Further decreasing the threshold would have rendered the frequent/infrequent distinction meaningless, as almost all words would be categorized as frequent.

Using the frequent and infrequent categories as defined by the four thresholds, we calculated the percentages of the different possible word orders at the boundaries of multiword utterances. We obtained these measures in the following way. We identified the first and the last two words of all utterances, that is two-word 'phrases' at the left and right utterance boundaries. If the 'phrase' had a [FW IW] order, it was counted as 'frequent-initial'. If it had an [IW FW] structure, it was counted as 'frequent-final'. 'Phrases' where both words were of the same category, i.e. [FW FW] or [IW IW] did not enter into the counts[4], as they were not informative about the relative order of frequent and infrequent words. Since the two corpora were not of equal size, the counts were transformed into percentages.

To evaluate the results statistically, we divided both corpora into 10 equal-sized subcorpora, calculated the percentages for the individual subcorpora using all four thresholds, and conducted ANOVAs over these datasets. We expected to find an interaction between languages and word orders, as an indication of opposite word orders in Japanese and Italian.

3.2. Results

Fig. 2A–D presents the percentages of frequent-initial and frequent-final utterances in the two languages using the four different thresholds. As expected, Japanese and Italian show the opposite patterns, Japanese having more frequent-final utterances, Italian more frequent-initial ones. Numerically, T_1 identifies 47% of the multiword utterances as frequent-final and 27% as frequent-initial in Japanese, 25% as frequent-final and 54% as frequent-initial in Italian. T_2 identifies 55% as frequent-final, and 31% as frequent-initial in

[4] Of course, their number is easily calculable by subtracting the sum of the frequent-initial and frequent-final utterances from the total number of multiword utterances.

Table 1
Some quantitative properties of the category ‘frequent word’ in Japanese and Italian, as defined by the four different relative frequency thresholds used in Experiment 2

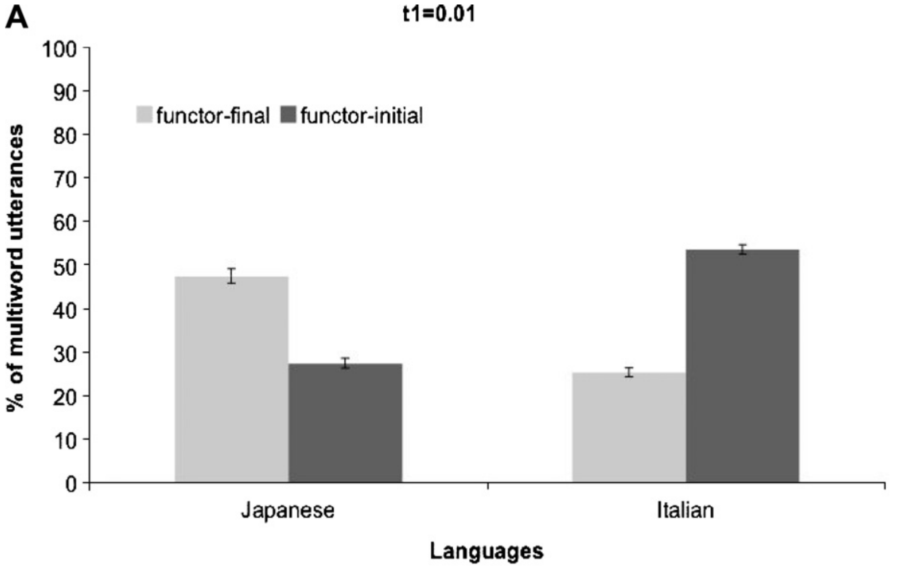
Relative frequency threshold	Number of word types in frequent word category		Absolute frequency of word type immediately above threshold		Percentage of word tokens covered by words types in frequent word category	
	Japanese	Italian	Japanese	Italian	Japanese (%)	Italian (%)
$T_1 = 0.01$	20	12	471	2457	29	26
$T_2 = 0.005$	34	36	236	1173	38	43
$T_3 = 0.0025$	63	72	119	595	48	56
$T_4 = 0.0001$	144	133	48	235	61	65

Japanese, and 26% as frequent-final and 64% as frequent-initial in Italian. T_3 identifies 54% as frequent-final and 31% as frequent-initial in Japanese, 26% as frequent-final and 66% as frequent-initial in Italian. T_4 identifies 46% as frequent-final and 29% as frequent-initial in Japanese, 24% as frequent-final and 62% as frequent-initial in Italian.

We carried out an ANOVA with factors Language (Japanese/Italian) and Order (frequent-initial/frequent-final) for each threshold using the percentages of frequent-initial and frequent-final ‘phrases’ in the 10 subcorpora as the dependent measure. For T_1, we obtained no main effect of Language. But there was a significant main effect of Order ($F(1,39) = 37.822$, $p < .001$), indicating that there were more frequent-initial phrases than frequent-final ones. Crucially, there was a significant interaction Language X Order ($F(1,39) = 1311.3$, $p < .0001$) due to the opposite order patterns attested in the two languages. For T_2, the ANOVA showed no main effect of Language, but a significant main effect of Order ($F(1,39) = 59.560$, $p < .0001$), once again reflecting the fact that there were more frequent-initial phrases overall in the two languages than frequent-final ones. Just as before, we also obtained a significant interaction Language X Order ($F(1,39) = 1161.6$, $p < .0001$), indicating that Japanese had more frequent-final phrases, while Italian had more frequent-initial ones. Unlike in the previous two cases, the ANOVA for T_3 revealed a significant main effect of Language ($F(1,39) = 42.118$, $p < .001$), indicating that this threshold filtered in more sentences in the Italian corpus than in the Japanese one. In addition, as before, we also found a significant main effect of Order ($F(1,39) = 135.60$, $p < .0001$), as well as a significant Language X Order interaction ($F(1,39) = 1709.0$, $\mathrm{p} < .00001$), indicating opposite orders in the two language. Using T_4, a similar pattern was obtained, with a significant main effect of Language ($F(1,39) = 72.158$, $p < .0001$) and Order ($F(1,39) = 178.74$, $p < .0001$), and a significant interaction between the two factors ($F(1,40) = 1213.3$, $p < .0001$).

3.3. Discussion

These results confirm the prediction that the relative order of frequent and infrequent words is the opposite in Italian and Japanese, as expected on the basis of the theoretical linguistic characterization of word order in these languages. Italian has more frequent-initial phrases at utterance boundaries than frequent-final ones, while Japanese has more of the latter type. This observation is true in both languages irrespectively of how FWs are defined, i.e. what frequency threshold was used. In addition to this relative difference



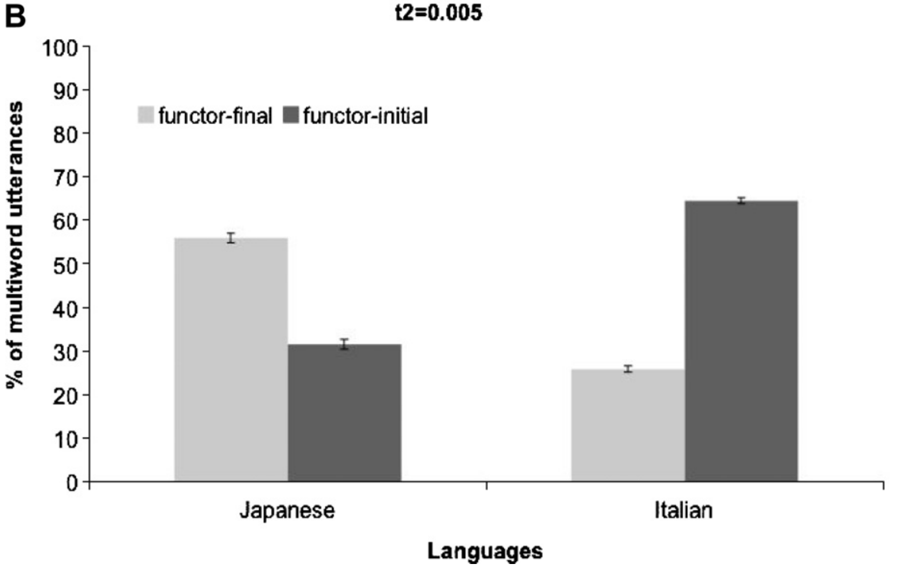


Fig. 2. The percentage of frequent-initial and frequent-final phrases at utterance boundaries in the Japanese and Italian infant-directed corpora, using four different relative frequency thresholds to define frequent words. A–D show the results at the four different thresholds. Light grey bars represent frequent-final phrases. Dark grey bars represent frequent-initial ones. The *y*-axis corresponds to the percentage of multiword utterances. Errors bars show standard errors of the means.

between word orders, the absolute numbers of word order types are also informative. Indeed, in Italian, utterances starting or finishing with a frequent-initial phrase constitute the absolute majority of all utterances at any of the four thresholds. In Japanese, utterances with frequent-final phrases outnumber other utterances at all four thresholds, and

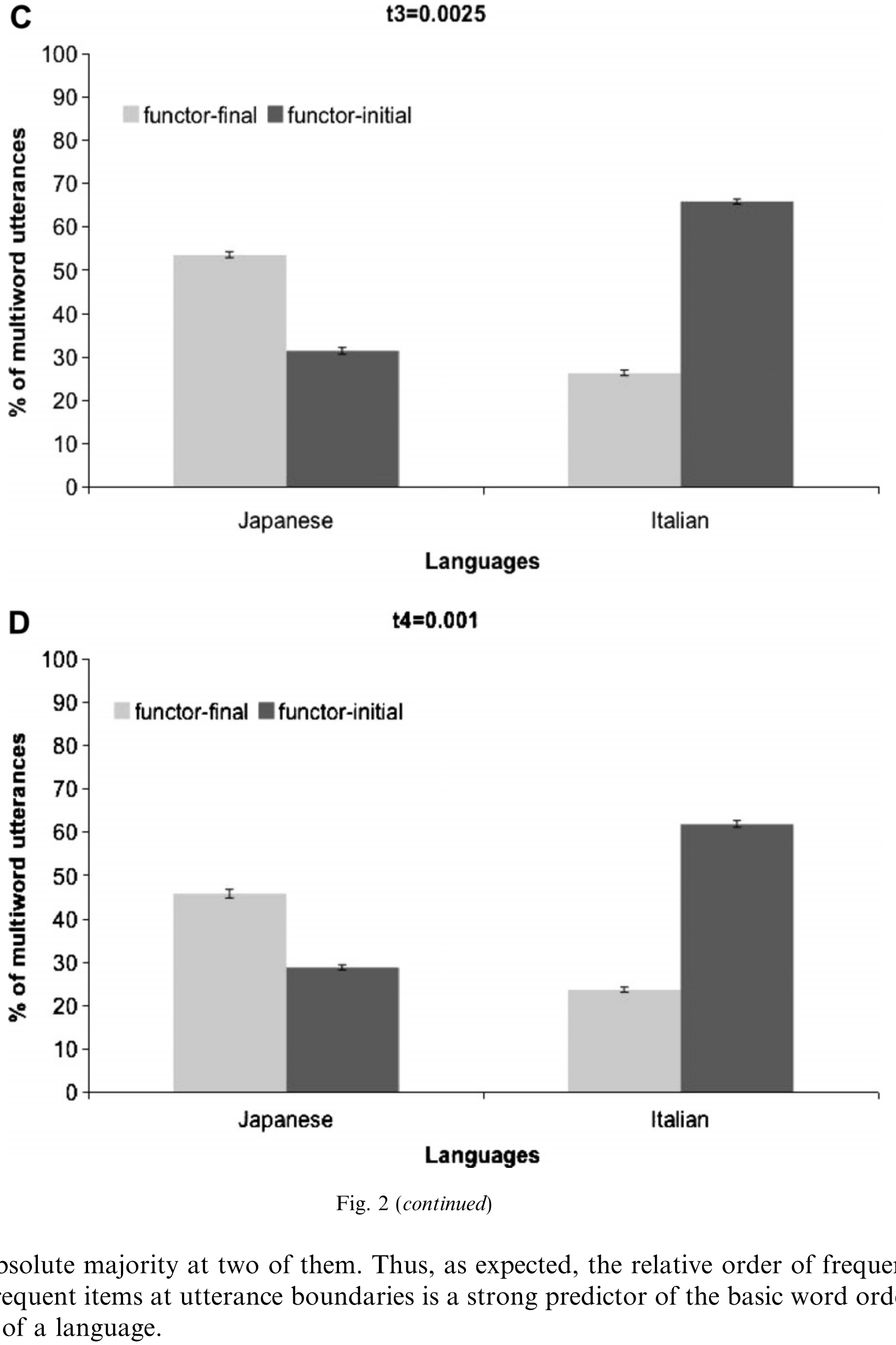


Fig. 2 (*continued*)

reach absolute majority at two of them. Thus, as expected, the relative order of frequent and infrequent items at utterance boundaries is a strong predictor of the basic word order pattern of a language.

Interestingly, in addition to the above finding, two more results were obtained: (i) overall, there were more frequent-initial utterances than frequent-final ones (at all thresholds), and (ii) at thresholds T_3 and T_4, more sentences were identified for Italian than for Japanese. Our assumption is that these results are at least partly attributable to another word order property, namely the relative order of determiners and nouns (formally, the Specifier–Head parameter). The order is [Det(erminer) N(oun)] in both languages (Japanese: *kono hon* 'his book'; Italian: *la tavola* 'the table'). However, Japanese has fewer determiners than Italian. Importantly, it lacks articles altogether, and only has demonstratives, numeral classifiers etc. Since determiners are functors, thus frequent words, while nouns are less frequent content words, the [Det N] pattern, common to both languages, increases the overall amount of frequent-initial utterances. Additionally, given that in Italian, there are more determiners than in Japanese, the lower thresholds identified more utterances in the former language than in the latter. This is does not happen for the two higher thresholds, because the most frequent functors are not determiners, and thus follow the Head–Complement, rather than the Specifier–Head pattern. In addition, other factors, such as the more varied nature of the Italian corpus, might also contribute to the presence of the additional effects. Whatever the definitive explanation may be, these effects are much smaller than the effect of the opposite word orders (cf. the statistical results above), thus they do not blur the strong interaction between language and word order that we are focusing on here.

In sum, our corpus data shows that frequency information in the input can be used as a heuristic predictor of category membership. This information, in turn, can be used to extract basic word order from phrases at utterance boundaries. Therefore, cues are available in the input to help infants' bootstrap word order even prelexically. As a next step, we need to show that infants are sensitive to this information at a prelexical age.

4. Experiment 3: Opposite word order preferences in Japanese and Italian infants

In order to investigate whether young learners have a prelexical representation of the distribution of functors and content words in their native language, we tested eight-month-old Japanese and Italian infants. As predicted by the information contained in the input, we expected to find opposite order preferences in the two populations. We tested this prediction in an artificial grammar learning experiment using the headturn preference paradigm. We familiarized both populations with the same artificial language, which allowed for two, symmetrically opposite organizations in terms of word order. Then we tested whether the two groups preferred test items instantiating opposite orders.

4.1. Method

4.1.1. Participants

The Japanese group consisted of 20 eight-month-old infants (mean age: 235 days, age range: 201–254 days; 9 females, 11 males). They were born to monolingual Japanese families, and had no record of neurological or auditory impairment. An additional 11 babies were tested, but not included in the analysis for the following reasons: failure to complete the experiment due to crying (3), fussiness (7), and experimenter error (1).

The Italian group consisted of 20 eight-month-old infants (mean age: 234 days, age range: 214–256 days; 10 females, 10 males). They were born to monolingual Italian families, and had no record of neurological or auditory impairment. An additional 10 babies were tested, but not included in the analysis for the following reasons: failure to complete the experiment due to crying (4), fussiness (4), experimenter error (1), and technical error (1).

A parent of each infant gave informed consent prior to participation. The study was approved by the Ethics Committee of RIKEN (where the Japanese infants were tested) and the Ethics Committee of SISSA (where the Italian infants were tested).

4.1.2. Material

We constructed an artificial grammar that allowed us to have a structurally ambiguous speech stream to be used for familiarization. We repeatedly concatenated a four-syllable-long basic unit: AXBY, where A and B represent constant syllables, while X and Y come from two categories containing 9 syllable tokens each (Fig. 3A). Thus, we obtained an alternating sequence of frequent (A & B) and infrequent (X & Y) syllables, mimicking functors and content words, respectively. By ramping the amplitude of the initial and final 15 s of this familiarization stream, phase information was suppressed. This resulted in an ambiguous underlying structure, since the basic unit could be identified either as having a frequent-initial (AXBY) or a frequent-final (XBYA) order (Fig. 3B).

The familiarization stream was synthesized using the fr4 female diphone database of MBROLA (Dutoit, 1997), with a monotonous pitch of 200 Hz and a constant phoneme duration of 120 s. The four-syllabic basic unit was repeated 243 times (each possible _X_Y syllable combination was used 3 times), resulting in a 3 min 53 s long familiarization stream.

The test items were eight four-syllabic 'sentences' of the language (Fig. 3C), four instantiating the frequent-initial order (AXBY), the other four the frequent-final one (XBYA). The infrequent X & Y syllables making up the test items were chosen in such a way that

(A) Lexicon	**A: {fi}**	
	X: {ru, pe, du, ba, fo, de, pa, ra, to}	
	B: {ge}	
	Y: {mu, ri, ku, bo, bi, do, ka, na, ro}	
(B) Familiarization	**...AXBYAXBYAXBYA...**	
	...gefofibugedefikogepafimoge...	
frequent-initial	**...gefofibugedefikogepafimoge...**	
	OR	
frequent-final	**...gefofibugedefikogepafimoge...**	
(C) Test Items	***frequent-initial***	***frequent-final***
	fifogebi	**bagebofi**
	firugemu	**kafipage**
	gedofide	**kufiduge**
	gerifipe	**ragenafi**

Fig. 3. The material used in Experiment 2. (A) The syllables used in the different categories. (B) The familiarization stream. The first row indicates the underlying structure. The second row is a chuck of the familiarization stream. The third and fourth rows show the two possible 'word orders' of the stream. Different shades of gray and underlining separate the constituent units. (C) The test items listed according to 'word order'.

the transitional probabilities (TP) between all syllable pairs within test items be zero or very low both in Japanese and Italian, as measured in our respective corpora of infant-directed speech used in Experiments 1 and 2. In an ANOVA with factors Language (Japanese/Italian) X Order (frequent-initial/frequent-final), using as dependent measure the TPs between the syllable pairs contained in the test items, we found no significant main effect (Language: $F(1,44) = 0.13$, $p = .72$; Order: $F(1,44) = 0.54$, $p = .46$) or interaction ($F(1,44) = 1.39$, $p = .24$). Thus, there was no bias in the test items from the TPs of the native language.

A test trial consisted of 15 repetitions of the same test item, separated by 500 ms pauses. The order and side of presentation of the test trials was randomized and counter-balanced across subjects in such a way that at most two consecutive trials could be of the same order type (frequent-initial/frequent-final).

4.1.3. Procedure

We used the headturn preference paradigm as described in Saffran, Johnson, Aslin, and Newport (1999) to test infants' word order preferences. Infants were tested individually while sitting on a parent's lap in a dimly lit, sound-attenuated cubicle. Parents were listening to masking music and were wearing dark sunglasses throughout the experiment to avoid all parental influence on infants' behavior. Infants first listened to the almost 4-min-long familiarization stream, while they watched attention-getter lights at the two sides or the center of the testing cubicle. The blinking of the lights was contingent upon the infants' looking behavior, but there was no systematic relation between the lights and the sounds. During the experiment, an experimenter, blind to the stimuli and seated outside the testing cubicle, monitored infants' looking behavior and controlled the lights and the stimuli. Infants were videotaped during the experiment for the subsequent off-line coding of their looking behavior.

Immediately after familiarization, infants were tested for their word order preference in eight test trials. Each trial started with the blinking of the central light to attract infants' attention. Once infants attended to the central light, one of the side lights started blinking and the central light was extinguished. When infants stably fixated on the blinking side light (defined as a 30° head turn towards the light), the associated test item started playing from a loudspeaker on the corresponding side. The sound file continued until the end (22 s) or until infants looked away for more than 2 s. After this, a new trial began.

Japanese infants were tested at the Laboratory of Language Development of the Brain Science Institute of RIKEN, Tokyo, Japan. In this laboratory, the testing cubicle had real lamps mounted on its walls. Italian infants were tested at the Language, Cognition and Development Laboratory of SISSA, Trieste, Italy. Here, the attention-getter lights were implemented as movies of blinking lamps displayed on flat screens attached to the walls of the cubicle.

4.2. Results

Infants' looking times were coded and measured off-line. They were averaged across all trials of the same type (frequent-initial/frequent-final) for both groups (Fig. 4). An ANOVA with factors Language (Japanese/Italian) as a between subject variable and

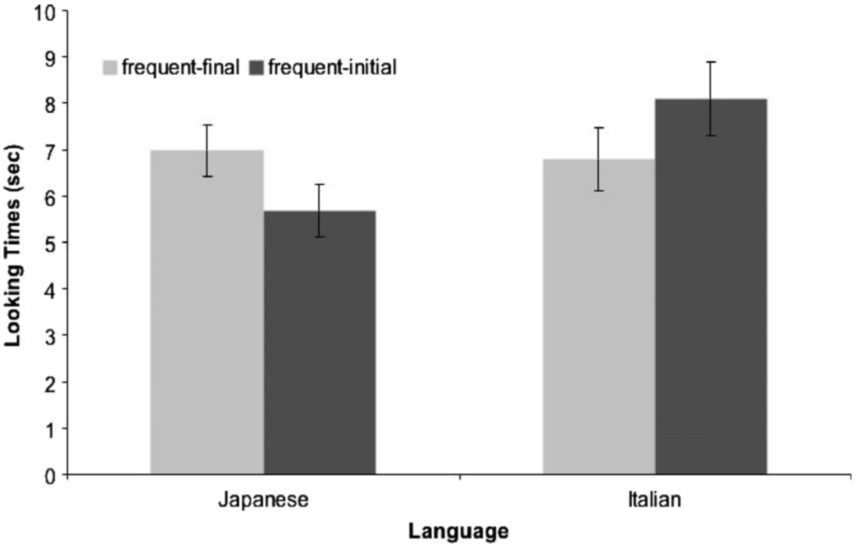


Fig. 4. The average looking times of the two groups. The *y*-axis indicates time in seconds. Light grey bars represent average looking times to the frequent-final test items. Dark grey bars represent average looking times for the frequent-initial test items. Error bars show standard errors of the mean.

Order (frequent-initial/frequent-final) as a within subject variable, using looking times as the dependent measure, yielded no significant main result. Importantly, the interaction Language X Order was significant ($F(1,38) = 8.3301$, $p = .006$), indicating that the two groups showed opposite looking patterns.

In a Scheffe post hoc test, we also compared looking times for the two types of test items in each group. The Japanese group looked significantly longer at the frequent-final items over the frequent-initial ones ($p = .046$), whereas the Japanese group exhibited the opposite pattern ($p = .049$).

4.3. Discussion

These results show that Japanese and Italian infants, who are exposed to languages with opposite word orders, have opposite expectations about the order of frequent and infrequent items in their target language. In other words, they show sensitivity to the frequency distributions and words orders they encounter in the input. This suggests that infants might use the relative order of functors and content words at utterance boundaries to create one of the first rudimentary representations of word order already before they build their lexicon.

5. General discussion

We have proposed a learning mechanism that allows bootstrapping word order before the lexicon is acquired. According to this proposal, infants are sensitive to the frequency distributions of word-like elements in the input, and also track the relative order of at least some of them, for instance, of those that occur at obvious syntactic boundaries, such as the edges of utterances. This mechanism makes use of two universal properties of the linguistic input. The first is that the frequency distributions of functors and content words are sufficiently different to allow the differentiation of the two categories. The second is that the relative order of most functors and content words correlates with other word order phenomena in a language, and can thus be used as a predictor of basic word order. We have shown that these two properties characterize infant-directed corpora in two typologically, genealogically and geographically different languages, Japanese and Italian. The combination of the two properties, that is, the relative order of frequent and infrequent words at utterance boundaries, yields a powerful cue for infants to cue word order.

This frequency-based bootstrapping mechanism raises a number of issues that require further clarification. First, just like other bootstrapping procedures, this is also a heuristic or probabilistic mechanism. It does not identify functors or content words infallibly; neither does it predict the precise word order of each and every phrase of a language. But, as we have shown, the cues are robust enough to inform learners about a general word order property, the order of Verbs and Objects, and their correlates, e.g. adpositions and nouns, complementizers and subordinate clauses, etc.

A second and related question is what level of abstraction this frequency-based representation is encoded at. Does it serve as a trigger to set the abstract word order parameters? Or does it remain a statistical encoding? While our study provides no definitive answer, it can safely be concluded, given the experimental results, that the representation is abstract enough to allow infants to generalize it onto the ‘functors’ and ‘content words’ of an unknown artificial language, in which all the ‘words’ are novel to them. However, more speculatively, we also propose that the frequency-based order representation works in conjunction with other bootstrapping mechanisms such as prosodic bootstrapping (Nespor, Guasti, & Christophe, 1996; Nespor et al., under review) to establish a more fully-fledged representation of word order in the target language. We suggest that the frequency-based mechanism could work as an initial procedure, yielding a general, overall representation of the most dominant word order pattern, characteristic of most phrase types in the target language. Then, this initial representation might be further elaborated by prosodic bootstrapping mechanisms, which assign a precise word order to each phrase type, especially when the word order of a given phrase type is different from the dominant order of the language (e.g. in ‘mixed’ languages, like German and Dutch; for a discussion, see below). Nespor and Vogel (1986) argue that the position of prosodic prominence in phonological phrases correlates with word order, and Nespor et al. (1996) propose that it can thus provide a perceptually available surface cue to it. In Turkish, for example, which is an OV language, the prominence is left-most, i.e. phrase-initial (***kilim*** *için* kilim for ‘for the kilim’), while in the VO language French, prominence is right-most, i.e. phrase-final (*pour châque* ***morale*** ‘for each ethic’). Moreover, this cue is perceptually detectable, since phrase-initial prominence is proposed to be cross-linguistically realized as increased pitch and intensity, while phrase-final prominence is mainly marked by increased duration (Nespor et al., under review).

This is true even when the same phrase type exhibits both word orders. In Dutch, for instance, if a prepositional phrase is pronounced with its canonical preposition–noun order (*op de* ***trap*** ‘up the stairs’), prominence is realized by lengthening the noun. If, on the contrary, the phrase has a non-canonical noun-‘preposition’ order (*de* ***trap*** *op* ‘the stairs up'), motivated by certain pragmatic contexts, prominence is implemented as higher pitch and intensity on the noun (which is now to the left). Nespor et al. (submitted for publication) have recently reported systematic measurements of the acoustic/phonetic correlates of prominence in German, a language that has both VO and OV orders in subordinate clauses. They have found that VO orders are accompanied by right-most lengthening, while OV orders come with increased pitch and intensity at the left edge, confirming that the realization of prominence changes as a function of word order even within a language. This close correlation between prosodic prominence and phrase-level word order is particularly relevant when considering mixed languages, in which different phrase types have different orders or even the same phrase type shows several different orders, such as in Dutch or German above.

Importantly, this mechanism, just like the frequency-based one, allows bootstrapping word order independently of the lexicon, and it also makes use of the edge positions of phrases. Given these representational similarities, it is not implausible that the two mechanisms might complement each other during the acquisition of word order. This hypothesis is in line with other bootstrapping theories that emphasize the importance of convergent cues in language acquisition (Morgan & Demuth, 1996).

A third issue concerns the phonetic form and detail in which the frequency-based representation might be encoded. We tested eight-month-old infants, who are clearly prelexical. From work by Shi et al. (2006a), (2006b), we know that at this age, infants have an underspecified representation of functors, at least for the purposes of segmentation. These authors show that at 8 months of age, English-learning infants are able to use frequent functors, such as *the* (but not less frequent ones, such as *her*) to segment content words. However, their representation is not complete, as they accept both the real English functor *the* ([ðə]) and a similar non-sense functor *kuh* ([kə]) in the segmentation task. Such an underspecified representation, however, does not compromise our hypothesis, since word forms are simply required to be categorized as a functor, their unique identification is not necessary. The fact that eight-month-olds only track the most frequent functors is also consistent with our proposal, since as we have shown, the distributions of functors and content words are maximally distinct precisely in the highest frequency range. As we have observed, a handful of the most frequent functors already cover a large part of the input, and provide reliable information about word order.

The idea that the frequency and the phonological form of functors contribute conjointly to language use and acquisition is not new. Investigating agrammatism in aphasic patients, Kean (1977), (1979 and subsequent work) proposed that it is not so much the syntactic function, but the phonological form of functors that accounts for their omission in agrammatic speech. This proposal prefigured a subsequent body of work emphasizing the prosodic and phonological differences between functors and content words (Morgan & Demuth 1996; Morgan, Meier, & Newport, 1987; Morgan et al., 1996; Shi & Werker, 2001, Shi et al., 1999, 2006a, 2006b), and linking the phonological minimality of the former to their outstandingly high frequency and predictability (Herron & Bates, 1997; Jurafsky, Bell, Gregory, & Raymond, 2000). A seminal paper by Kelly (1988) highlighted another aspect of the contribution of functors to prosodic bootstrapping. The author derived the well-known rhythmic asymmetry between English verbs (weak–strong, i.e. iambic, e.g. *re**cord***) and nouns (strong–weak, i.e. trochaic, e.g. ***rec**ord*) from the different rhythmic contexts they appear in. By analyzing spoken and written English, as well as by eliciting non-word productions from participants, Kelly (1988) showed that the rhythmically weak suffixes that verbs frequently take (e.g. *–ing*) bias them towards an iambic stress pattern, thus conforming to the rhythmic alternation principle (e.g. *sug**gest**ing*, weak–strong–weak). Nouns, on the contrary, typically do not take such weak suffixes in English. Although Kelly's (1988) framework is different from ours, his results can be interpreted as evidence that functors contribute to the prosodic bootstrapping of the two major lexical categories.

6. Conclusion

Word order is an essential and overarching property of languages, which infants learn early on during acquisition. In the above, we have reported a series of experiments that explore a frequency-based mechanism to bootstrap word order in prelexical infants. We have shown, using Japanese and Italian infant-directed corpora, that functors and content words have different frequency distributions, which can serve as a universal predictor of category membership. Furthermore, the relative order of these two categories correlates with other word order phenomena of the language. We have found that prelexical Japanese and Italian infants are able to detect these properties of the input, as shown by their preferential looking behavior, mirroring the dominant word order of their respective native languages. This simple representation of order can serve as a cue to acquire more general word order regularities. Moreover, these findings demonstrate for the first time that infant populations acquiring different native languages start diverging in their structural knowledge as early as 8 months of age.

The division of labor between functors and content words has long been recognized as universal among the languages of the world, and fundamental to their functional design. Our proposal highlights an important aspect of this distinction. We show that, in addition to segmentation, the different frequency distributions and sentential positions of the two categories also contribute to bootstrapping grammatical structure. Therefore, the distinction between functors and content words appears as a design feature of language that plays a crucial role in its learnability.

References

Abney, S. (1987). *The English noun phrase in its sentential aspect. Unpublished doctoral dissertation*. Cambridge, MA: MIT.

Antelmi, D. (n.d.). The Antelmi corpus. CHILDES Database.

Antinucci, F., & Parisi, D. (1973). Early language acquisition: A model and some data. In C. F. D. Slobin (Ed.), *Studies in child language development*. New York: Holt.

Aslin, R., Woodward, J., Lamendola, N., & Bever, T. (1996). Models of word segmentation in fluent maternal speech to infants. In J. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early language acquisition* (pp. 117–134). Mahwah, NJ.: Lawrence Erlbaum Associates.

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

Chang, F., Lieven, E., & Tomasello, M. (in press). Automatic evaluation of syntactic learners in typologically-different languages. *Cognitive Systems Research*.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Cipriani, P., Pfanner, P., Chilosi, A., Cittadoni, L., Ciuti, A., Maccari, A., et al. (1989). *Protocol li diagnostici e terapeutici nel lo sviluppo e nel la patologia del linguaggio*.

Cutler, A. (1993). Phonological cues to open- and closed class words in the processing of spoken sentences. *Journal of Psycholinguistic Research, 22*, 133–142.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language, 2*, 133–142.

Dryer, M. S. (1992). The Greenbergian word order correlations. *Language, 68*, 81–138.

Dutoit, T. (1997). *An introduction to text-to-speech synthesis*. Dordrecht: Kluwer.
Fukui, N. (1986). *A theory of category projection and its applications. Unpublished doctoral dissertation*. Cambridge, MA: MIT.
Greenberg, J. H. (1963). *Universals of Languages*. Cambridge, MA: MIT Press.
Guasti, T. (2002). *Language acquisition*. Cambridge, MA: MIT Press.
Herron, D. T., & Bates, E. A. (1997). Sentential and acoustic factors in the recognition of open- and closed-class words. *Journal of Memory and Language, 37*, 217–239.
Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2000). Probabilistic Relations between Words: Evidence from Reduction in Lexical Production. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.
Kean, M-L. (1977). The linguistic interpretation of aphasic syndromes: Agrammatism in Broca's aphasia, an example. *Cognition, 5*(1), 9–46.
Kean, M-L. (1979). Agrammatism: A phonological deficit? *Cognition, 7*(1), 69–83.
Kelly, M. (1988). Rhythmic alternation and lexical stress differences in English. *Cognition, 30*(2), 107–137.
Kucera, H., & Francis, N. (1967). *A computational analysis of present-day American English*. Providence, RI: Brown University Press.
MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum Associates.
Mazuka, R. (1996). Can a grammatical parameter be set before the first word? Prosodic contributions to early setting of a grammatical parameter. In J. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 313–330). Mahwah, NJ: Lawrence Erlbaum Associates.
Mazuka, R., Igarashi, Y., & Nishikawa, K. (2006). Input for learning Japanese: RIKEN Japanese mother–infant conversation corpus. *The institute of electronics, information and communication engineers technical report, 16*, 11–15.
Mehler, J., Sebastian Gallés, N. v., & Nespor, M. (2004). Biological foundations of language: Language acquisition, cues for parameter setting and the bilingual infant. In M. Gazzaniga (Ed.), *The new cognitive neurosciences III*. Cambridge, MA: MIT Press.
Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning. *Cognitive Psychology, 19*, 498–550.
Morgan, J. L., & Demuth, K. (1996). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.
Morgan, J., Shi, R., & Allopenna, P. (1996). Perceptual bases of rudimentary grammatical categories. In J. Morgan & K. Demuth (Eds.), *Signal to syntax*. Mahwah, NJ: Lawrence Erlbaum Associates.
Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris.
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., Shukla, M., Avesani, C., van de Vijver, R., Schraudolf, H., & Donati, C. (submitted for publication). Different phrasal prominence realizations in VO and OV languages?
Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
Rizzi, L. (1986). Null objects in Italian and the theory of *pro*. *Linguistic Inquiry, 17*, 501–558.
Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tonal sequences by human infants and adults. *Cognition, 70*, 27–52.
Selkirk, E. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
Shi, R., & Werker, J. (2001). Six-month-old infants' preference for lexical words. *Psychological Science, 12*(1), 70–75.
Shi, R., Werker, J., & Morgan, J. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition, 72*(2), B11–B21.
Shi, R., Cutler, A., Werker, J. F., & Cruickshank, M. (2006a). Frequency and form as determinants of functor sensitivity in English acquiring infants. *Journal of the Acoustical Society of America, 119*(6), EL61–EL67.
Shi, R., Werker, J. F., & Cutler, A. (2006b). Recognition and representation of function words in English-learning infants. *Infancy, 10*(2), 187–198.
Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74*(3), 209–253.
Tonelli, L. (n.d.). The Tonelli corpus. CHILDES Database.
Volterra, V. (1976). A few remarks on the use of the past participle in child language. *Journal of Italian Linguistics, 2*, 149–157.
Volterra, V. (1984). Waiting for the birth of a sibling: The verbal fantasies of a two year old boy. In I. Bretherton (Ed.), *Symbolic play*. New York: Academic Press.
Zipf, G. (1935). *The psycho-biology of language*. Boston, MA: Houghton Mifflin.