Article

Current Biology

Word Frequency Is a Cue to Lexical Category for 8-Month-Old Infants

Highlights

- Eight-month-old infants use word frequency to categorize functors and content words
- They allow substitutions for content words, processing them as an open class
- But they do not accept novel functors, suggesting those constitute a closed class
- They recognize the relative order of the two classes in their native language

Authors

Caterina Marino, Carline Bernard, Judit Gervain

Correspondence

judit.gervain@parisdescartes.fr

In Brief

In six behavioral experiments, Marino et al. demonstrate that 8-month-old infants process frequent words as belonging to closed classes resembling grammatical functors and infrequent words as belonging to open classes like content words, and they map the relative order of these categories to the basic word order of their native language.

Marino et al., 2020, Current Biology *30*, 1–7
April 20, 2020
https://doi.org/10.1016/j.cub.2020.01.070

CellPress

Please cite this article in press as: Marino et al., Word Frequency Is a Cue to Lexical Category for 8-Month-Old Infants, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.01.070

Word Frequency Is a Cue to Lexical Category for 8-Month-Old Infants

Caterina Marino,[1,2] Carline Bernard,[1,2] and Judit Gervain[1,2,3,*]
[1]Integrative Neuroscience and Cognition Center (INCC UMR8002), Université Paris Descartes, 45 rue des Saints-Pères, 75006 Paris, France
[2]Integrative Neuroscience and Cognition Center (INCC UMR8002), CNRS, 45 rue des Saints-Pères, 75006 Paris, France
[3]Lead Contact
*Correspondence: judit.gervain@parisdescartes.fr
https://doi.org/10.1016/j.cub.2020.01.070

SUMMARY

The linguistic distinction between function words (functors) (e.g., *the*, *he*, *that*, *on*…), signaling grammatical structure, and content words (e.g., *house*, *blue*, *carry*…), carrying meaning, is universal across the languages of the world. These two lexical categories also differ in their phonological makeup (functors being shorter and more minimal) and frequency of occurrence (individual functors being much more frequent than most content words). The frequency-based discrimination of the two categories could constitute a powerful mechanism for infants to acquire the basic building blocks of language. As functors constitute closed classes and content words come in open classes, we examined whether 8-month-old monolingual infants relied on word frequency to categorize and track functors and content words. In six artificial grammar-learning experiments, we have found that infants process frequent words as belonging to closed classes, and infrequent words as belonging to open classes, and they map the relative order of these categories following the basic word order of their native language. These findings provide the earliest evidence that infants use word frequency as a cue to lexical categories and combine them to build rudimentary representations of grammar.

INTRODUCTION

A universal feature of human language is the division of labor between function words (functors) such as *the*, *he*, *on*, etc., marking grammatical structure, and content words like *rainbow*, *write*, *beautiful*, etc., carrying lexical meaning. The ability to identify functors is a crucial first step for infants on their way to their native grammar. Functors are highly frequent, while content words occur less frequently. It has, therefore, been hypothesized that infants rely on frequency of occurrence as a particularly useful cue to establish these basic lexical categories [1–5]. However, direct evidence is still lacking. It is thus unknown how and at what age infants first use these basic building blocks to parse the input into grammatically relevant patterns. Here, we show that French monolingual 8-month-olds rely on word frequency and sequential position to categorize and track functors and content words, processing frequent words as a closed-class category (functors) and infrequent words as open classes (content words). This suggests that infants start acquiring basic knowledge of grammar well before they produce speech or have a sizeable lexicon.

The majority of natural languages are configured in one of two ways: functors either precede or follow the content words with which they form syntactic units. In English or French, functors appear at the beginning of phrases, e.g., *in London*, *à Paris* (in Paris), whereas in Japanese or Basque they occur at the end, e.g., *Tokyo ni* (Tokyo to), *garren atzean* (flame behind), respectively. It has been proposed that learners use functors in the input as anchors, encoding the position of other words in relation to them [6–8]. Indeed, adult speakers prefer the relative order of functors and content words in an artificial grammar that is coherent with the basic word order of their native language [9, 10]. Similarly, 8-month-old infants exposed to languages with opposite word orders, e.g., functor-initial Italian and functor-final Japanese, show opposite preferences for word order in an artificial grammar task. Italian infants prefer sequences starting with a frequent word, while Japanese infants prefer sequences starting with an infrequent word, mirroring their native word orders [5]. As the difference in frequency between functors and content words is universal across languages, the frequency-based bootstrapping of lexical categories could provide a powerful tool for young infants to break into language.

To test this hypothesis, we used a distinctive feature of functors and content words as a diagnostic tool. Content words form open classes: new words are added to the lexicon every day (e.g., *iPad*, *Brexit*, etc.), whereas functors constitute closed classes into which new items cannot be added without a major language change. If infants' linguistic representations are such that frequent words come in closed classes and infrequent ones in open classes, this establishes sensitivity to one of the differential features of functors and content words. Together with infants' already documented knowledge of other distinctive features of the two classes, such as their different phonological forms [11] and functional differences [5, 12, 13], our study will provide convergent evidence for infants' ability to represent these two universal lexical classes early in language development.

To this effect, we used the Head-turn Preference Paradigm (HPP, Figure 1) to familiarize six groups of infants with an artificial language (Figure 2) in which frequent words (*F*), mimicking

Please cite this article in press as: Marino et al., Word Frequency Is a Cue to Lexical Category for 8-Month-Old Infants, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.01.070

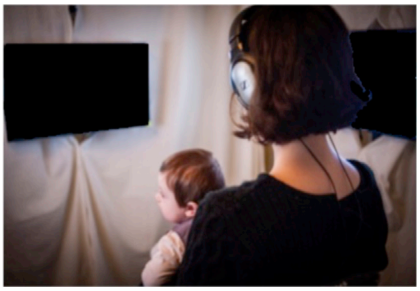

Figure 1. Experimental Setup

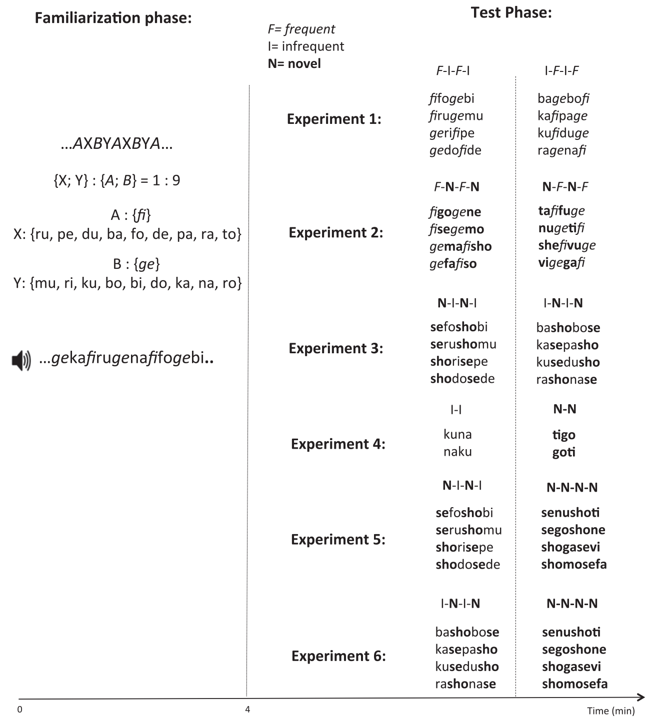

Figure 2. Artificial Grammar Task Used in Experiments 1–6

functors, and infrequent words (I) corresponding to content words strictly alternated (….*geka*fi*mu*ge*na*fi*fo*ge*bi…), replicating the paradigm used in [5]. As phase information was masked by ramping the beginning (15 s) and end (15 s) of the stream in amplitude, the structure of the stream was ambiguous between a frequent-word-initial and a frequent-word-final parse. All six groups were familiarized with the same stream but differed in the test items they received. We measured infants' looking times to the test items in all six experiments.

RESULTS

Eight-Month-Old Monolinguals Map the Relative Order of Frequent and Infrequent Words to the Basic Word Order of Their Native Language

We tested the baseline group (Experiment 1, n = 21) on test sequences taken from the familiarization stream. Half of them started with a frequent word (*F*-I-*F*-I: e.g., *fi*fo*ge*bi) and the other half with an infrequent word (I-*F*-I-*F*: e.g., ba*ge*bo*fi*). French being a functor-initial language, we predicted that infants would show a preference for the frequent-word-initial (*F*-I-*F*-I) sequences, similarly to the Italian infants in previous studies [5]. Crucially, both *F*-I-*F*-I and I-*F*-I-*F* sequences appeared in the stream and were thus equally likely parses. Infants' predicted *F*-I-*F*-I preference must derive from their knowledge of the native functor-initial word order. This experiment was identical to the one conducted in [5], where opposite word order preferences for Italian and Japanese infants had been found.

Infants in Experiment 1 demonstrated the predicted frequent-word-initial preference (Figure 3), showing longer looking times to these items than to frequent-word-final items (*F*-I-*F*-I looking times: M = 7.96 s; CI [confidence interval] = 6.76 to 9.16; I-*F*-I-*F*: looking times M = 6.08 s; 95% CI of mean = 5.0 s to 7.14 s; t (20) = 4.41; p = 0.0003; d = 0.77; power (1-β) = 0.91), corresponding to French word order. This result thus establishes that 8-month-old French-learning infants do indeed represent the basic word order of their native language in terms of the relative order of frequent and infrequent words.

To further validate that this finding is consistent with previous evidence using the same paradigm [5], we directly compared our results with those collected from Japanese infants (n = 20; *F*-I-*F*-I looking times: M = 5.68 s; 95% CI of mean = 4.52 s to 6.84 s; I-*F*-I-*F* looking times: M = 6.98 s; 95% CI of mean = 5.83 s to 8.13 s; t (19) = 2.157; p = 0.04; data from [5]). We ran an ANOVA with language (French versus Japanese) as a between-subject factor and word order (frequent-initial versus frequent-final) as a within-subject factor. Similarly to the Italian-Japanese comparison reported in [5], we did not find a main effect of test item (F (1,39) = 0.646, p = 0.43) or of language (F (1,39) = 1.026, p = 0.32), but crucially, we obtained a significant language X test item interaction (F (1,39) = 18.941, p = 0.04), as French infants looked longer to the frequent-word-initial items and Japanese infants to the frequent-word-final items. This confirms that French and Japanese infants exhibit opposite word order preferences, and that this preference is related to the native language of the participants, as both Japanese and French infants were tested with exactly the same material.

Eight-Month-Old Monolinguals Rely on Word Frequency to Categorize Frequent Words as Belonging to Closed Classes

For a second group (Experiment 2, n = 21), we replaced the infrequent words in the test items with novel ones (*F*-**N**-*F*-**N**: e.g., *fi***go***ge***ne** versus **N**-*F*-**N**-*F*: e.g., **ta***fi***fu***ge*). If infants process infrequent words as content words belonging to open classes, they should maintain their frequent-word-initial preference, as the frequent "functors" providing the structural skeleton of the sequences remain in place. The novel words in the test items could be categorized as "infrequent," i.e., content words, despite the fact that no frequency information is available about them, due to their position with respect to the unchanged frequent words.

The preference for the frequent-initial items was indeed maintained in Experiment 2 (Figure 3; F-**N**-*F*-**N** looking times: M = 7.08 s; 95% CI of mean = 5.82 s to 8.34 s; **N**-*F*-**N**-*F* looking times: M = 5.98 s; CI = 4.61 to 7.34; t (20) = 2.592, p = 0.017; d = 0.381; power (1-β) = 0.381). This finding shows that infants

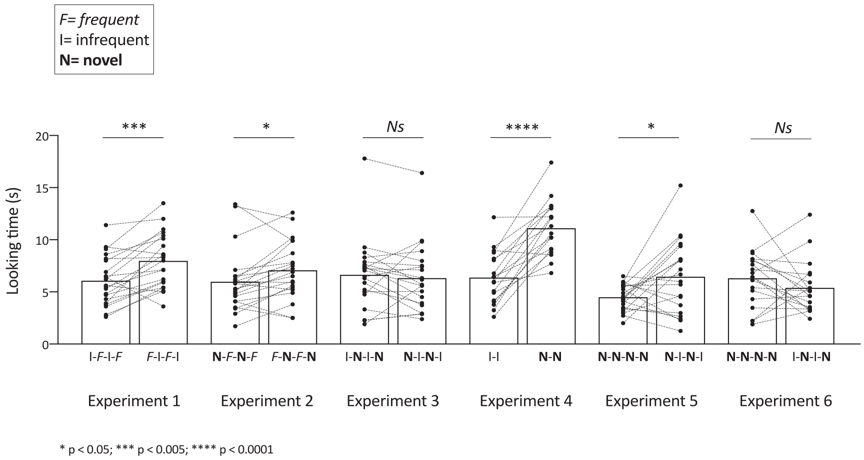

Figure 3. Results of Experiments 1–6

Looking times for Experiments 1–6. The x axis shows the different experimental groups. The y axis shows looking time in seconds. Bars represent group means; connected dots represent individual participants' looking times in the two experimental conditions.

process the infrequent category as an open class, which is a key feature of content word categories in natural language.

For a third group of infants (Experiment 3, n = 21), the frequent words were replaced with novel ones (**N**-I-**N**-I: e.g., **se**fo**sho**bi versus I-**N**-I-**N**: e.g., ba**sho**bo**se**). We expected this to disrupt infants' preference, as they could no longer use the frequent words as structural anchors. Importantly, it is still possible to categorize the novel words as "frequent" due to their sequential position, i.e., their position with respect to the infrequent words, and thus establish the preferred word order. However, if infants categorize infrequent words as content words, they may be less likely to rely on them for structural information, and when given the choice between the two possible word orders, they might fail to show a preference. Note that in Experiments 2 and 3, it is the positional and not the frequency information that allows generalization to the novel items in the test phase, and the two experiments are similar in this regard. In other words, if the lack of frequency information about novel items in the test sequences impacts infants' preference, then we should see the same pattern of results in Experiments 2 and 3, whereas if positional information is used, possibly only derived from frequent items, then different patterns may obtain for Experiments 2 and 3, which is what we predict. Importantly, in both Experiments 2 and 3, if a preference is found, it provides insight about what information in the test items is sufficient to drive a word order preference and what information is not necessary. If items from a category are replaced by novel ones, and a preference is still found, then items of the other category are sufficient to cue infants about word order, and the identity of the items of the other category is not necessary, which we take as an indication that this category is an open class.

We found that in Experiment 3, infants no longer showed a word order preference (**N**-I-**N**-I looking time: M = 6.32 s; 95% CI of mean = 4.85 s to 7.79 s; I-**N**-I-**N** looking time: M = 6.63 s; CI = 5.11 to 8.16; t (20) = 0.677, p = 0.506, Ns.; d = 0.125; power (1-β) = 0.09), suggesting that they do not accept novel frequent words, and they likely process this category as a closed class.

To establish that infants behaved differently in Experiment 3 than in Experiment 1 or Experiment 2, we conducted a linear mixed effects model. We generated the most complex model, which included the fixed factors word order (frequent-initial versus frequent-final) as a within-subject factor and experiment (1 versus 2 versus 3) as a between-subject factor, as well as their interaction. Subject was added as a random factor (σ^2 = 4.86; SD = 2.19) and word order had a random intercept. We then compared this model with simpler models removing terms step-by-step. The model that best fit the data contained the interaction of the two fixed factors (word order X experiment) and the random factor Subject. An ANOVA was then calculated for this model, using the function ANOVA in R. The ANOVA revealed a significant main effect of word order (χ^2 (1) = 11.897; p = 0.0006) and a significant word order X experiment interaction (χ^2 (1) = 7.014; p = 0.030), but no significant main effect of experiment (χ^2 (1) = 0.648; p = 0.723). The results are presented in Table 1.

Importantly, the results of Experiment 2 and 3 have two interrelated implications. The positive result of Experiment 2 suggests that the presence of the frequent words is sufficient to maintain the frequent-initial structure. Additionally, it also means that infants' preference is not disrupted by the novel infrequent words; i.e., they categorize infrequent words as belonging to open lexical classes, or at least they accept their variation without changing word order representations. Similarly, the null preference in Experiment 3 suggests that replacing frequent items is disruptive (i.e., no variation is accepted in the category, so it is a closed class), and relatedly, that infants do not compute word order on the basis of infrequent items. On the basis of Experiments 2 and 3 alone, it is difficult to assess exactly how much weight each of these factors carries in the observed performance. Infants may simply ignore the novel items or may be unable to process or remember them, especially in Experiment 3, where a null preference was found. The three control experiments presented below address this question.

Eight-Month-Old Monolinguals Succeed in Remembering the Infrequent Items from the Familiarization

We presented a fourth group of infants (Experiment 4, n = 18) with pairs of infrequent words from the familiarization stream contrasted with pairs of novel words (e.g., I-I: kuna versus **N**-**N**: **tigo**). This allowed us to make sure that any difference found between Experiments 2 and 3 were not simply due to infants failing to remember the infrequent words in the test items of Experiment 3. The purpose of Experiment 4 was thus to serve as a control against such memory effects and not to establish word order preference. It, therefore, differed from the previous three experiments in the structure of its test items as well as in the predictions. If infants could remember the infrequent words

Table 1. Summary of the Linear Mixed Effects Model Comparing Experiments 1–3

	β	Stnd Error	DF	T
Word order 1	2.09	0.61	403.3	3.45
Experiment 2	−0.55	0.80	60.4	−0.68
Experiment 3	−0.57	0.80	59.8	−0.71
Word order 1 X Experiment 2	−1.04	0.86	403.3	−1.21
Word order 1 X Experiment 3	−2.26	0.85	403.3	−2.64

from the stream, then we predicted that they would show a preference for the novel words in Experiment 4; i.e., a decreased interest in the familiar infrequent words. We predicted this novelty preference based on previous evidence and theoretical proposals suggesting that task or stimulus complexity influences when infants make the "familiarity-to-novelty" shift [14].

Infants indeed showed a preference for the novel items (novel items looking time (**N**-**N**): M = 11.1 s; CI = 9.8 to 12.41; familiar items looking time (I-I): M = 6.38 s; 95% CI of mean = 5.13 s to 7.61 s; t (17) = 7.084, $p < 0.0001$; d = 1.842; power (1-β) = 1). This novelty preference is an indication that infants were more familiar with the infrequent words that appeared in the stream than with the novel items. Their differential looking behavior in Experiments 2 and 3 cannot thus be attributed to memory limitations in encoding the words of the stream. However, it is also compatible with an alternative claim, namely that infants fail to show a preference, because they fail to process the infrequent items.

Eight-Month-Old Monolinguals Encode the Position of Infrequent Items

Experiment 4 shows that infants remember the infrequent words, at least when not presented in a structural sequence. But can they process them when the infrequent words are inserted into the alternating test item sequences? If yes, do infants have knowledge about their position? If yes, then the null preference in Experiment 3 cannot be attributed to infants' failure to process infrequent items. In Experiments 5 and 6, we addressed these questions. The aim was to determine whether the null preference is obtained because infants ignore or fail to process infrequent words altogether, or whether they process and have some knowledge about them but do not use them to select word order, possibly because they treat them as content words and hence uninformative about structure. In Experiment 5 (n = 20), we presented infants with test items in which frequent words were replaced by novel syllables in the initial position (**N**-I-**N**-I, identical to Experiment 3) and contrasted them with items in which both frequent and infrequent words were replaced by novel ones (**N**-**N**-**N**-**N**: e.g., **se**fo**sho**bi versus **senushoti**). These latter items carry neither frequency, nor positional information; i.e., no cues at all. If infants under these conditions show a preference for the items in which infrequent words are in the native-like final position, then that would suggest that even if they do not readily rely on infrequent words as structural anchors when making a word order choice as in Experiment 3, they may nevertheless have implicit knowledge about their sequential position.

As predicted, infants showed a preference for the items in which infrequent words appeared in their native-like final position over all-novel sequences (**N**-I-**N**-I looking time: M = 6.45 s; CI = 4.8 to 8.1; **N**-**N**-**N**-**N** looking time M = 4.49 s; 95% CI of mean = 3.89 s to 5.08 s; t (19) = 2.55, p = 0.019; d = 0.738; power (1- β) = 0.88).

In Experiment 6 (n = 19), we presented a sixth group of infants with test items in which frequent words were replaced by novel syllables in the final position (I-**N**-I-**N**, as in Experiment 3) as well as with the all-novel test items also presented in Experiment 5 (**N**-**N**-**N**-**N**: e.g., ka**se**pa**sho** versus **senushoti)**. Infants' preference in Experiment 5 may reflect, as we expect, that they have some knowledge about the position of infrequent items, although they don't use them as structural anchors. Alternatively, they may prefer **N**-I-**N**-I items over all-novel ones, because the former are more familiar: they exhibit an alternating pattern like the familiarization and contain familiar infrequent items. Experiment 6 allows us to tease these two alternatives apart. If familiarity played a role in Experiment 5, a preference for the I-**N**-I-**N** items is to be expected in Experiment 6, as well, despite the non-native-like word order, as these items also contain familiar infrequent words and show alternation. If, however, a genuine sensitivity to the native word order governed infants' performance in Experiment 5, then we predict a null preference in Experiment 6, as the infrequent items are no longer in the final position characteristic of French.

In Experiment 6, infants did not indeed show a preference (I-**N**-I-**N** looking time: M = 5.38 s; 95% CI of mean = 4.18 to 6.59; **N**-**N**-**N**-**N** looking time M = 6.31 s; 95% CI of mean = 4.99 s to 7.62 s; t (18) = 1.27, p = 0.218; d = 0.354; power (1- β) = 0.31).

Taken together, Experiments 5 and 6 show evidence that infants do not ignore infrequent items and can recognized the position they typically occupy in their native languages.

Since Experiments 5 and 6 serve as controls for the predicted null result in Experiment 3, it is informative to compare these experiments. However, it is not possible to include these three experiments into a single linear mixed effects model as we did for Experiments 1–3, because Experiments 3, 5, and 6 do not have the same test item types. Experiment 3 contrasts frequent-initial and frequent-final items, whereas Experiments 5 and 6 both have a test item type that carries no word order information; i.e., the all-novel **N**-**N**-**N**-**N** sequences. We thus compared looking times to the test items that were identical in the different experiments pairwise, using unpaired samples t tests, with Bonferroni correction for multiple comparisons (α = 0.05/3 = 0.0167). We found no difference in looking times to **N**-I-**N**-I items, shared between Experiments 3 **(**M = 6.32 s) and 5 (M = 6.45 s; t (38.28) = 0.123; p = 0.9 Ns.; d = 0.03; power (1-β) = 0.07). Similarly, there was no difference in looking times to the I-**N**-I-**N** items between Experiments 3 (M = 6.63 s) and 6 (M = 5.38 s; t (36.72) = 1.353; p = 0.184, Ns.; d = 0.42; power (1-β) = 0.132). By contrast, we found a significant difference in the looking times for the all-novel items across Experiments 5 (M = 4.49 s) and 6 (M = 6.31 s; t (25.13) = 2.652; p = 0.013; d = 0.86; power (1-β) = 0.574).

Why did infants show a preference for **N**-I-**N**-I items when those were contrasted with all-novel items, but not when they were contrasted with I-**N**-I-**N** items? We argue that this is because the infants' task was more complex in the latter case

(Experiment 3), where they had to choose between two possible word orders, than in the former (Experiment 5), in which they simply needed to choose between the native word order and a sequence with no identifiable order.

Taken together, the results of Experiments 4, 5, and 6 shed light on the relative contributions of the two factors playing a role in the results of Experiments 2 and 3; i.e., infants' ability to compute word order on the basis of available information and to maintain word order despite novel information. In Experiment 2, where a positive preference was observed, infants necessarily needed to process infrequent words as replaceable; i.e., an open category. Additionally, infants could use the frequent words, which were maintained, as anchors to word order. In Experiment 3 we obtained a null result, which, albeit predicted, can in itself be due to several factors: infants not accepting novel items in the frequent word position, infants not relying on infrequent words for structure, or alternatively, because infants fail to process the infrequent words in the test items.

To disentangle these possibilities, we first demonstrated that infants remember the infrequent items from the familiarization (Experiment 4), arguing against the memory deficit alternative. Second, the fact that infants show a preference for test items in which infrequent words occupy their native-like position and combine with novel ones in this canonical order (**N**-I-**N**-I items, Experiment 5), but not when this canonical order is violated (I-**N**-I-**N** items, Experiment 6), indicates that they are able to associate positional information with infrequent words and apply their native word order representation to them correctly. Consequently, the lack of preference for these same **N**-I-**N**-I items in Experiment 3 cannot be explained by infants' inability to encode infrequent items.

Altogether, the most coherent interpretation of this set of results is that infants rely on frequent words to compute word order and they process these as not replaceable; i.e., on the basis of positional information alone, it is not possible to add new items into this category. Moreover, even if we cannot establish with certainty that infants categorize infrequent words as belonging to open classes, they do not ignore them altogether and have some implicit knowledge about their sequential position.

DISCUSSION

In six artificial grammar-learning experiments, we have found that infants implicitly categorize frequent words as anchors for word order and process them as belonging to closed classes. Furthermore, they may possibly process infrequent words as belonging to open classes, and they map the relative position of these categories onto the order of functors and content words in their native language.

Functors and content words universally differ in a set of properties: their function, their frequency of occurrence, their sequential position, their phonological make-up, and the closed-versus open-class nature of the lexical categories to which they belong. Previous research has shown that infants are sensitive to the functional difference, expecting infrequent words, but not frequent ones, to have a referent [12, 13]. They can discriminate functors and content words on the basis of their phonological differences at birth [11] and know their relative positions in the native language at 8 months [5, 15, 16]. By establishing infants' sensitivity to the open-class versus closed-class nature of the two categories, the current study aims to add the last piece to the puzzle, demonstrating infants' sensitivity to the only hitherto unexplored distinctive feature of lexical categories. Taken together with these previous findings, our results provide the earliest evidence that at 8 months, infants already use word frequency as a potential cue to lexical categories, which they combine functionally to build rudimentary representations of word order. This implies that the acquisition of early grammatical knowledge begins in parallel with the development of native phonology and the lexicon.

This early acquisition can readily be accounted for within a bootstrapping framework. Such theories argue that learners are able to extract abstract, structural, and hence directly unobservable properties of the target language from perceptually available cues in the input that correlate with the underlying structure [17]. Our results suggest that the differential frequency distribution of functors and content words is a potential cue, which infants can use to bootstrap the basic word order of their native language.

The ability to categorize words on the basis of frequency has also been found in non-human animals; e.g., rats [18]. This suggests that frequency-based categorization is a general mechanism shared across species. Crucially, however, rats always show a preference for frequent-initial or familiar-initial sequences independently of the structure of the familiarization or test sequences. Contrarily to infants, we therefore hypothesize that rats would exhibit a preference for the infrequent initial items in Experiments 3 and 6 (I-**N**-I-**N**), since they show a preference for sequences with a familiar initial syllable irrespective of the 4-syllabic structure—a prediction future research will need to address. Importantly, this behavior is very different from that of infants, who prefer the order that matches the word order found in their native language; i.e., frequent-initial for French and Italian infants, but frequent-final for Japanese infants. This implies that frequency-based lexical categorization interacts with language experience and feeds into linguistic representations, constituting a valuable bootstrapping strategy [5, 16].

In apparent contradiction with our hypothesis, functors are generally produced later than content words. Crucially, however, while infants produce functors relatively late, in perception they are sensitive to them much earlier. This may be due to the different phonological make-up of the two classes. Functors are universally phonologically more reduced than content words; e.g., they may not carry stress, have a simpler syllable structure than content words, often contain reduced vowels, etc. [19]. The specific features in which they differ from content words vary across languages, but they are always more minimal. In French, for instance, grammatical functors cliticize onto their content words (e.g., *je t'aime* /ʒ(ə)**'t**ɛm(ə)/ I you.acc love "I love you"); they are typically shorter and have simpler syllable structure than content words. In perception, infants can already discriminate lists of functors and content words on the basis of their different phonological properties at birth [11], and only a few months later prefer to listen to content words [20, 21]. Between 11 and 13 months, infants start discriminating between existing functors of their native language and their close mispronunciations, and can use them to segment the associated content words [22, 23]. Furthermore, 17-month-olds are more likely to associate infrequent words than frequent ones with a possible referent, suggesting that they expect infrequent words, i.e., content words, to have semantic content [12, 13]. At 24 months, an age when infants do not yet produce many functors, they nevertheless understand better sentences in which functors are in place as compared to sentences in which functors are omitted or scrambled [24, 25]. Generally, infants produce functors later than content words due to production constraints attributable to their phonological minimality (e.g., tendency to omit unstressed rather than stressed word units) [26, 27]. The fact that infants do not produce functors early is thus not incompatible with our result that at 8 months they can already track them in the input and use them as structural anchors.

Importantly, since at this age infants do not yet have a sizeable lexicon, this knowledge is most likely not item based. Rather, bootstrapping appears to be a learning mechanism independent of vocabulary learning. It is important to note in this regard, however, that our experimental paradigm itself is not constrained between the morphological (word) and syntactic (phrase) levels. Nevertheless, the results of our previous studies, especially [5], suggest that the task taps into syntactic or phrase-level processing, as Japanese infants show infrequent-initial word order preferences, while Italian infants show the opposite preference. This difference is clearly related to the word order differences of Japanese and Italian infants' native language. Additional evidence comes from another study using the same paradigm [16], in which phrasal level prosody was added to the familiarization stream, effectively guiding bilinguals learning an object-verb and a verb-object language toward the word order whose prosody was added. Thus, in two of our previous studies, which both used the same paradigm as our current experiments, we have found independent evidence for phrase-level processing. Additionally, in natural languages, word order phenomena at the syntactic and the morphological levels are correlated. Typological studies [28, 29] as well as corpus analyses [e.g., 5, 16] show that there is a correlation between the position of free functors (i.e., syntax) and bound functors (i.e., morphology). Object-verb languages such as Japanese are predominantly suffixing while verb-object languages tend toward prefixing. Thus, their functor-directionality is the same at the syntactic and morphological levels. Since both free and bound functors are highly frequent, our hypothesis that infants process frequency as a cue to membership in the closed class of functors is compatible with both a syntactic and a morphological level of processing.

Lastly, even if this set of experiments is not fully conclusive about how infants process infrequent items, the results are suggestive, showing that at 8 months, infants have acquired some basic knowledge about their sequential position and readily accept their replacement.

In conclusion, our findings suggest that 8-month-old infants are sensitive to frequency and positional information in the language input. Moreover, they use this information to build rudimentary representations of grammar. The distinction between function and content words is a universal feature of human language, which infants need to learn early in development. Frequency-based bootstrapping, the mechanism we have uncovered here, provides an account of how they might achieve this early and efficiently.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Participants
- METHOD DETAILS
  - Procedure
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND CODE AVAILABILITY

ACKNOWLEDGMENTS

This work was supported by the "PredictAble" Marie Skłodowska-Curie Actions (MSCA) Innovative Training Network (ITN) (SEP-210134423) for understanding and predicting developmental language abilities and disorders in multilingual Europe (ERC Consolidator grant "BabyRhythm" 773202) to J.G.

AUTHOR CONTRIBUTIONS

C.M., C.B., and J.G. developed the study concept. All authors contributed to the study design. Testing, data collection, analysis, and interpretation were performed by C.M. and C.B. under the supervision of J.G. C.M., C.B., and J.G. wrote the manuscript. J.G. secured funding. All authors approved the final version of the manuscript for submission.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 5, 2019
Revised: November 23, 2019
Accepted: January 22, 2020
Published: March 12, 2020

REFERENCES

1. Braine, M.D. (1966). Learning the positions of words relative to a marker element. J. Exp. Psychol. *72*, 532–540.
2. Green, T.R.G. (1979). The necessity of syntax markers: Two experiments with artificial languages. J. Verbal Learn. Verbal Behav. *18*, 481–496.
3. Morgan, J.M., and Newport, E.L. (1981). The role of constituent structure in the induction of an artificial language. J. Verbal Learn. Verbal Behav. *20*, 67–85.
4. Mori, K., and Moeser, S.D. (1983). The role of syntax markers and semantic referents in learning an artificial language. J. Verbal Learn. Verbal Behav. *22*, 701–718.
5. Gervain, J., Nespor, M., Mazuka, R., Horie, R., and Mehler, J. (2008). Bootstrapping word order in prelexical infants: a Japanese-Italian cross-linguistic study. Cognit. Psychol. *57*, 56–74.
6. Valian, V., and Levitt, A. (1996). Prosody and Adults' Learning of Syntactic Structure. J. Mem. Lang. *35*, 497–516.
7. Valian, V., and Coulson, S. (1988). Anchor points in language learning: The role of marker frequency. J. Mem. Lang. *27*, 71–86.
8. Morgan, J.L., Meier, R.P., and Newport, E.L. (1987). Structural packaging in the input to language learning: contributions of prosodic and morphological marking of phrases to the acquisition of language. Cognit. Psychol. *19*, 498–550.
9. Gervain, J., Sebastián-Gallés, N., Díaz, B., Laka, I., Mazuka, R., Yamane, N., Nespor, M., and Mehler, J. (2013). Word frequency cues word order in

adults: cross-linguistic evidence. Front. Psychol. *4*, 689. https://doi.org/10.3389/fpsyg.2013.00689.

10. De la Cruz-Pavía, I., Elordieta, G., Sebastián-Gallés, N., and Laka, I. (2015). On the role of frequency-based cues in the segmentation strategies of adult OV- VO bilinguals. Int. J. Biling. Educ. Biling. *18*, 225–241.
11. Shi, R., Werker, J.F., and Morgan, J.L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. Cognition *72*, B11–B21.
12. Hochmann, J.R., Endress, A.D., and Mehler, J. (2010). Word frequency as a cue for identifying function words in infancy. Cognition *115*, 444–457.
13. Hochmann, J.R. (2013). Word frequency, function words and the second gavagai problem. Cognition *128*, 13–25.
14. Hunter, M.A., and Ames, E.W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. Advances in infancy research *5*, 69–95.
15. Bernard, C., and Gervain, J. (2012). Prosodic cues to word order: what level of representation? Front. Psychol. *3*, 451. https://doi.org/10.3389/fpsyg.2012.00451.
16. Gervain, J., and Werker, J.F. (2013). Prosody cues word order in 7-month-old bilingual infants. Nat. Commun. *4*, 1490. https://doi.org/10.1038/ncomms2430.
17. Morgan, J.L., and Demuth, K. (1996). Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition. *Hillsdale* (NJ: Lawrence Erlbaum Associates, Inc), pp. 41–53.
18. Toro, J.M., Nespor, M., and Gervain, J. (2016). Frequency-based organization of speech sequences in a nonhuman animal. Cognition *146*, 1–7.
19. Morgan, J.L., Shi, R., and Allopenna, P. (1996). Perceptual bases of rudimentary grammatical categories: toward a broader conceptualization of bootstrapping. In Signal to Syntax, J.L. Morgan, and K. Demuth, eds. (Hillsdale, NJ: Lawrence Erlbaum Associates), pp. 263–283.
20. Shi, R., and Werker, J.F. (2001). Six-month-old infants' preference for lexical words. Psychol. Sci. *12*, 70–75.
21. Shi, R., and Werker, J.F. (2003). The basis of preference for lexical words in 6-month-old infants. Dev. Sci. *6*, 484–488.
22. Shafer, V.L., Shucard, D.W., Shucard, J.L., and Gerken, L. (1998). An electrophysiological study of infants' sensitivity to the sound patterns of English speech. J. Speech Lang. Hear. Res. *41*, 874–886.
23. Shi, R., Werker, J., and Cutler, A. (2006). Recognition and representation of function words in English-learning infants. Infancy *10*, 187–198.
24. Gerken, L., Landau, B., and Remez, R.E. (1990). Function morphemes in young children's speech perception and production. Dev. Psychol. *26*, 204–216.
25. Shipley, E.F., Smith, C.S., and Gleitman, L.R. (1969). A study in the acquisition of language: Free responses to commands. Language *45*, 322–342.
26. Demuth, K. (1994). On the 'underspecification' of functional categories in early grammars. In Syntactic Theory and First Language Acquisition: Cross-Linguistic Perspectives, B. Lust, M. Suñer, and J. Whitman, eds. (Hillsdale, N.J.: Lawrence Erlbaum Associates), pp. 119–134.
27. Gerken, L., and McIntosh, B.J. (1993). Interplay of function morphemes and prosody in early language. Dev. Psychol. *29*, 448–457.
28. Greenberg, J.H. (1963). Universals of Human Language (Stanford University Press).
29. Dryer, M.S. (1992). The Greenbergian word order correlations. Language *68*, 81–138.
30. Dutoit, T. (1997). An Introduction to Text-to-Speech Synthesis (Kluwer Academic Publishers). https://doi.org/10.1007/978-94-011-5730-8.
31. Nelson, D.G., Jusczyk, P., Mandel, D., Myers, J., Turk, A., and Gerken, L. (1995). The head-turn preference procedure for testing auditory perception. Infant Behav. Dev. *18*, 111–116.
32. Saffran, J.R., Johnson, E.K., Aslin, R.N., and Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. Cognition *70*, 27–52.
33. Faul, F., Erdfelder, E., Buchner, A., and Lang, A.G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav. Res. Methods *41*, 1149–1160.
34. Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software *67*, 1–48.
35. Fox, J., and Weisberg, S. (2011). Multivariate linear models in R. In An R Companion to Applied Regression (Los Angeles: Thousand Oaks).

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
PsyScope	http://psy.ck.sissa.it	n/a
PsyCode	http://psy.ck.sissa.it/PsyCode/PsyCode.html	n/a
SPSS	https://www.ibm.com/analytics/spss-statistics-software	n/a
R	https://www.r-project.org	n/a
Prism	https://www.graphpad.com/scientific-software/prism/	n/a

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Judit Gervain (judit.gervain@parisdescartes.fr).

This study did not generate new reagents.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Participants

Parents of all participating infants gave written informed consent prior to participation. All experiments were approved by the ethics boards of the institutions involved (CERES of the Université Paris Descartes) CER-Paris Descartes, approval nr 2016/32.

Experiment 1. Thirty (15 girls) 8-month-old (mean age 8 months and 4 days, range 7.5- 9 months) French infants took part in Experiment 1. Among these 30 infants, nine were not included in the final data analysis, because of technical problems (1), because French was not the main language spoken in the home (2), because they had too short (shorter than 960 ms, the duration of a test item) or too long (longer than 21 ms, the maximal duration of a trial) looking times in more than four trials (4) or for fussiness and crying (2). A final sample of 21 participants were entered into the analysis.

Experiment 2. Thirty-three (16 girls) 8-month-old (mean age 8 months and 20 days, range 8- 9 months) French infants took part in Experiment 2. Among these 33 infants, twelve were not included in the final data analysis, because they had too short or too long looking times (4) or for fussiness and crying (8). A final sample of 21 participants were entered into the analysis.

Experiment 3. Twenty-seven (10 girls) 8-month-old (mean age 8 months and 20 days, (range 8-9 months) French infants took part in Experiment 3. Among these 27 infants, 6 were not included in the final data analysis, because they had too short or too long looking times (3) or for fussiness and crying (3). A final sample of 21 participants was entered in the analysis.

Experiment 4. Thirty (13 girls) 8 month-old (mean age 8 months and 16 days, (range 8-9 months) French infants took part in Experiment 4. Among these 30 infants, five were not included in the final data analysis, because they had too short or too long looking times (1) or for fussiness and crying (4). Additionally, seven infants were excluded from the analysis because of a family risk/history of language impairment. A final sample of 18 participants was entered into the analysis. The reason for this unusual distribution of participants is that the time when Experiment 4 was run coincided with the data collection period for another study, which tested behavioral differences between typical and atypical participants. Some of the at-risk participants, who were originally all recruited for this other study, could not, however, be tested there for different practical reasons and were thus tested in Experiment 4. This data is not reported in the current paper, but will be pursued given the first author's parallel interest in behavioral differences between typical and atypical populations and will be reported, in a later publication, once more data can be collected. Importantly, no participants at risk were present in the final analysis of any of the experiments. Sample sizes were determined on the basis of power/sample size calculations based on [5], as described below in the Quantification and Statistical Analysis.

Experiment 5. Thirty (15 girls) 8-month-old (mean age 8 months and 6 days, range 7.5-9 months) French infants took part in Experiment 5. Among these 30 infants, ten were not included in the final data analysis because of technical problems (1), because they had too short or too long looking times in more than four trials (7) or for fussiness and crying (2). A final sample of 20 participants was entered into the analysis.

Experiment 6. Twenty-five (9 girls) 8-month-old (mean age 8 months and 22 days, (range 8-9 months) French infants took part in Experiment 6. Among these 25 infants, 6 were not included in the final data analysis for fussiness and crying (3) and because they had too short or too long looking times in more than four trials (3). A final sample of 19 participants was entered into the analysis.

METHOD DETAILS

The artificial grammar task was exactly the same as the one used in [5]. During familiarization a 3-min 48 s long speech stream was played with alternating frequent and infrequent words concatenated without pauses. The grammar consisted of a four-syllable-long basic structure (AXBY), where each unit is realized as a consonant-vowel (CV) syllable. In this structure A and B units mimic frequent words (function words), whereas X and Y mimic infrequent words (content word), because the A and B categories have one token each (A: *fi*; B: *ge*), while the X and Y categories contain nine tokens (X: ru, pe, du, ba, fo, de, pa, ra, to; Y: mu, ri, ku, bo, bi, do, ka, na, ro), making individual X and Y tokens nine times less frequent than A and B tokens. This ratio 1:9 ratio has been successfully used in [5] as well as in [15, 16]. The ratio needs to satisfy two opposing constraints: (i) it needs to be large enough to cue the functor/content word distinction (see corpus results in [5] about actual frequencies), while (ii) it needs to be reasonably implementable within the limited syllable repertoire of an artificial grammar paradigm.

The stream was synthesized using a text-to-speech synthesis software (MBROLA, fr4 French voice) [30] with a pitch of 200Hz (corresponding to the fundamental frequency of female voices) and phoneme duration of 120 ms. The stream thus provided no prosodic cue to its structure. Furthermore, the initial and final 15 s of the stream were ramped in amplitude, suppressing information about the exact beginning and end of the stream. As a result, the structure of the stream was ambiguous between a frequent word initial and a frequent word final parse (e.g., …*ge*fo*fi*bu-*ge*de*fi*ko-*ge*pa*fi*mo-*ge*ku*fi*na… or …*ge*-fo*fi*bu*ge*-de*fi*ko*ge*-pa*fi*mo*ge*-ku*fi*na…).

During the test phase, 8 test items were presented in Experiments 1, 2, 3, 5 & 6. A single test trial consisted of the same test item, repeated 16 times separated by a pause of 500ms (e.g., *fi*fo*ge*bi_ *fi*fo*ge*bi_ *fi*fo*ge*bi_ *fi*fo*ge*bi_ *fi*fo*ge*bi_....), resulting in test trials that lasted maximum 21 s.

For Experiment 4, participants were tested with 4 test items. Two pairs of infrequent words from the familiarization stream (I-I: kuna; naku) were contrasted with two pairs of novel words (**N**-**N**: **tigo**; **goti**). A single test trial consisted of the same test item, repeated 22 times for experiment 4 (e.g., kuna_kuna_kuna_kuna_…), resulting in test trials that lasted maximum 21 s.

The full list of CV syllables and test items for all Experiments are shown in Figure 2.

Procedure

The Headturn Preference Paradigm (HPP) [31, 32] was used to test our hypothesis (Figure 1). This experimental method measures infants' looking behavior to assess preferences for and/or discrimination between at least two different types of auditory stimuli. The experiments were conducted in a quiet testing booth, with three side screens on which visual attention getters (e.g., videos of looming circles imitating blinking lights) were played (one on each side). Below each side screen, loudspeakers were placed for the presentation of the sound stimuli. Infants were seated on a caregiver's lap, sitting on a chair in the middle of the booth. The caregiver listened to masking music in order to avoid influencing the infant's response. Each experimental session consisted of a familiarization phase and a test phase. During the familiarization phase infants listened to the continuous familiarization speech stream, which was played independently of infants' looking behavior. Infants also gained experience with the visual attention getters, which unlike the sounds were presented contingently upon infants' looking behavior (see below). After the end of the familiarization phase, infants immediately went on to the test phase. In the test phase, both the sound and the visual stimulus were contingent upon infants' looking behavior. A typical trial started with the presentation of the central attention getter on the front screen. Once infants reliably fixated on it, the central attention getter was extinguished and one of the side attention getters was turned on (sides were randomized and counterbalanced within and across infants). Once the infant reliably fixated on the blinking side screen, as indicated by a head turn of at least 30° to that side, a sound stimulus started to play from the loudspeaker placed below the corresponding side screen. The trial lasted until the infant turned away for more than a predefined look away criterion (2 s) or until the end of the sound file (21 s). A new trial was then presented. During the study, an experimenter located outside the testing booth and blinded to the stimuli being presented, monitored infants' looking behavior and operated the stimulus presentation software (PsyScope version X B55 run on a Mac OS X, version 10.10.5). Experimental sessions were recorded, and the videos were analyzed offline to measure infants' looking times. For each experiment, we averaged looking times across all trials of the same condition after the offline coding of the videos. One blind coder coded all the videos. Additionally, a second coder coded a set of randomly selected videos, representing 18% of all videos. The correlation between the two coders was $r = 0.85$.

QUANTIFICATION AND STATISTICAL ANALYSIS

Based on effect sizes derived from [5], a study using the same artificial grammar-learning paradigm [5] as the current one, we performed a power calculation to estimate the sample size.

In the previous study, an effect size (Cohen's d) of $d = 0.524$ was obtained. Using this effect size and a power of 0.7, the sample size calculation for a one-tailed, paired sample t test was performed using G*Power [33] and it yielded a required sample size of 19. Therefore, in the five main experimental conditions (Experiments 1, 2, 3, 5 & 6), we aimed for a final sample size (after rejection) of at least 19 infants.

By contrast, as Experiment 4 uses a simple recognition/novelty paradigm and thus it was not directly comparable with the previous study [5], we couldn't rely on previously established effect sizes. We assumed, as it is common in the literature, that a simple recognition/novelty preference has a moderate to large effect size. With a Cohen's *d* of 0.6 corresponding to the conventional value for moderate to large effects and a power of 0.7, the needed sample size was 15 for a one-tailed paired sample t test. We therefore aimed for a final sample size of at least 15 infants.

Looking times were coded offline. One blind coder coded all the videos. Additionally, a second coder coded a set of randomly selected videos, representing 18% of all videos. The correlation between the two coders was r = 0.85. Data from some participants were not included in the final analyses (see details in the Participants section for each experiment). For all four experiments, rejection was performed on the basis of pre-defined criteria, prior to statistical analyses.

Offline coded looking time data were analyzed using the SPSS and R softwares. First, paired samples t tests (with equal variance not assumed) were run to compare the two test items in Experiments 1, 2, 3, 4, 5 and 6 using SPSS. The same software was then used to run the ANOVA between the French data from Experiment 1 and the Japanese data from [5]. The Linear Mixed Effects Model was performed by using the lme4 package in R [34]. *P value*s were then obtained conducting ANOVAs in the package CAR in R [35].

DATA AND CODE AVAILABILITY

The videos of the experiments supporting the current study have not been deposited in a public repository due to European regulations on the protection of personal data, but the offline coded looking times are available from the corresponding author upon reasonable request.