

de Carvalho, A., He, A. X., Lidz, J., & Christophe, A. (to appear). Prosody and function words cue the acquisition of word meanings in 18-month-old infants. *Psychological Science*

Prosody and function words cue the acquisition of word meanings in 18-month-old infants

Alex de Carvalho^{a,b,c*}, Angela Xiaoxue He^d, Jeffrey Lidz^e and Anne Christophe^{a,b}

x.de.carvalho@gmail.com; angelaxe.axh@gmail.com; jlidz@umd.edu; anne.christophe@ens.fr

a. Laboratoire de Sciences Cognitives et Psycholinguistique (ENS, EHESS, CNRS) - Département d'Études Cognitives, École normale supérieure – PSL University, 29 rue d'Ulm – 75005 Paris, France.

b. Maternité Port-Royal, AP-HP, Université Paris Descartes - 53 Boulevard de l'Observatoire – 75014 Paris, France.

c. Department of Psychology – University of Pennsylvania. 425 S University Ave, Philadelphia, PA 19104, USA.

d. Department of Speech, Language & Hearing Sciences – Boston University. 635 Commonwealth Ave. Boston, MA 02215 MA, USA.

e. Department of Linguistics - University of Maryland, 1401 Marie Mount Hall, College Park – MD 20742, USA.

*Corresponding Author: Alex de Carvalho, Department of Psychology, University of Pennsylvania, 425 South University Avenue (Office 252), Philadelphia, PA 19104, USA.
Email: x.de.carvalho@gmail.com

Abstract

Language acquisition presents a formidable task for infants, in which word learning is a crucial yet challenging step. Syntax (the rules for combining words into sentences) has been robustly shown to be a cue to word meaning. But how can infants access syntactic information when they are still acquiring the meanings of words? We investigate the contribution of two cues that may help infants break into the syntax and give a boost to their lexical acquisition: phrasal prosody (speech melody), and function words, both of which are accessible early in life and correlate with syntactic structure in the world's languages. We show that 18-month-old infants use prosody and function words to recover sentences' syntactic structure, which in turn constrains the possible meaning of novel words: participants ($n = 48$ in each of two experiments) interpreted a novel word as referring either to an object or an action, given its position within the prosodic-syntactic structure of sentences.

Key-words: phrasal prosody; function words; language acquisition; syntactic acquisition; lexical development

Prosody and function words cue the acquisition of word meanings in 18-month-old infants

Humans acquiring language face the challenging task of learning the meanings of words: they have to map the sounds of each word to a possible meaning. Given that in fluent speech, words are not separated from one another by clear acoustic markers (such as a silent pause), and given that for each spoken sentence the world offers a wide array of possible referential intentions, how do babies manage to achieve this sound-to-meaning mapping? A central problem in language acquisition is to determine what sources of information infants can exploit to go from sound to meaning.

Syntactic structure, which governs the organization of words into sentences, has been proposed to be a universal and reliable source of information that children may exploit to discover the meaning of words (e.g., Gleitman, 1990). For instance, by two years of age, children infer that a novel word, such as *dax*, refers to an action when it occupies a verb position in a sentence, as in *He is daxing that*, or to an object when it occupies a noun position, as in *This is a dax* (e.g., Waxman, Lidz, Braun, & Lavin, 2009). These findings demonstrate that the syntactic structure in which the words occur is an important source of information for children: they exploit the syntactic environment of a word to determine its syntactic category (e.g., a noun or a verb) and use the syntactic category to restrict the kind of meaning the novel word can have (e.g., verbs refer to actions/events).

This ability to exploit and learn from syntactic structures so early, although impressive, seems rather counterintuitive. Given that syntactic structure defines the relationships between words in a sentence, and allows listeners to compute the meaning of a sentence from the meaning of the individual words that compose it, one would expect that infants would first need to learn the words and their meanings, to then be able to learn how to organize words into sentences. We are thus faced with a chicken-and-egg problem: children seem to need words to learn syntax, and to need syntax to learn words. How can infants avoid this circularity? Here we experimentally tested whether 18-month-olds can compute the

syntactic structure of a sentence, by relying on phrasal prosody and function words, two sources of information that are available early during language acquisition and convey information about syntactic structure.

Phrasal prosody is the rhythm and melody of speech: when we speak, words are not pronounced one after the other in a monotone way, rather they are grouped together into intonational units (i.e., prosodic phrases). For example, the sentence “*The little cat is running fast*” tends to be pronounced as [the little cat] [is running fast], where brackets represent prosodic units. In all the world’s languages, the boundaries between prosodic units always coincide with syntactic boundaries¹ (e.g., Shattuck-Hufnagel & Turk, 1996). In our example, this boundary appears between the noun phrase [“the little cat”], and the verb phrase [“is running fast”]. Infants are sensitive to phrasal prosody from birth (e.g., Mehler et al., 1988) and they perceive prosodic cues marking the boundaries between groups of words a few months afterwards (Shukla, White, & Aslin, 2011; Soderstrom, Seidl, Nelson, & Jusczyk, 2003). Thus, if they can pay attention to salient prosodically-conditioned acoustic information (e.g., phrase-final lengthening, pauses, pitch contour discontinuity) correlating with syntactic constituent boundaries, they might be able to not only identify potential subdivisions in fluent speech, but also to infer the location of some syntactic boundaries.

Function words and morphemes are elements that serve certain grammatical functions (e.g., articles, auxiliaries, pronouns...). They are acquired within the first year of life, because they are highly frequent (much more than content words: nouns, verbs, adverbs), and possess perceptual and distributional characteristics that distinguish them from content words (e.g., Shi, Morgan, & Allopenna, 1998). Because functional elements tend to consistently co-occur with content words from specific word classes (e.g. determiners such as “the” or “a” typically co-occur with nouns, while pronouns like “she” and “they” tend to co-occur with verbs),

¹ The reverse is not true, because not all syntactic boundaries are marked prosodically; for instance in [he eats], the syntactic boundary between the subject and the verb phrase is unmarked prosodically.

infants could use statistical/distributional information in their input to learn about function words and to identify which words or sets of words co-occur with words from specific categories (e.g., Mintz, 2003). Validating this hypothesis, previous studies showed that between 12 and 24 months old, infants can use function words to categorize content words (e.g., "the blick" vs. "I blick", Cauvet et al., 2014; He & Lidz, 2017; Shi & Melançon, 2010). However, in real-life not all content words are preceded by function words (e.g., in: "The baby flies...", flies can be either a noun or a verb). In such cases, infants would need to integrate additional information into their distributional analysis, and take into account syntactic constituents to constrain their parsing (i.e., fly is a noun in: [The baby flies]_{NP}..., but a verb in: [The baby]_{NP}[flies...his kite]_{VP}). Thus function words/morphemes and phrasal prosody, together, may allow young infants to build at least a rudimentary representation of the syntactic structure of sentences (Christophe et al. 2016; 2008). Supporting this hypothesis, computational work demonstrates that models relying on phrasal prosody, function words and a minimal semantic knowledge successfully predict the syntactic category of prosodic-syntactic units and unknown words (Christodoulopoulos, Roth, & Fisher, 2016; Gutman, Dautriche, Crabbé, & Christophe, 2015).

The situation is thus as follows: phrasal prosody and function words are jointly predictive of syntactic structure in natural languages and young infants are sensitive to each of these sources of information. What has never been investigated is whether infants can jointly use phrasal prosody and function words to access the syntactic structure of sentences and constrain their acquisition of word meanings. This ability would be crucial for language acquisition because it would allow infants to break free of the chicken-and-egg problem, since phrasal prosody and function words are acquired well before infants know many words.

In the current study we experimentally tested this hypothesis, investigating whether 18-month-old French-learning infants are able to exploit function words (Experiment 1) and phrasal prosody together with function words (Experiment 2) to constrain the acquisition of

nouns and verbs.

Experiment 1: Function words constrain the acquisition of word meanings

This experiment tested whether 18-month-olds are able to infer that a novel word such as *bamoule* refers to an object when listening to sentences such as *It is a bamoule*, and to an action when listening to sentences such as *She is bamouling*.

Method

The studies reported in this paper, including the entire method, analysis and criteria for exclusion of participants were pre-registered on the OSF (Open Science Framework) database before running the experiments (the formal preregistration can be accessed with the following link: https://osf.io/9n4u3/?view_only=ea4ac33f605240b2a89000924bb531ea). The materials, collected data, and data analysis are freely available to readers through the following link: https://osf.io/u2xct/?view_only=f61af293ef524e1cbb550a45341148f7

Participants. Forty-eight French-learning 18-month-olds participated in the study, 24 in each experimental group (mean age = 18.1 months, range = 17.7 to 18.6 months; $SD = 0.2$; 27 girls). An additional thirty-two infants came to the lab, but were not included in the final sample for one of the following reasons: because of fussiness not allowing them to finish the experiment ($n = 13$); because they failed to meet the pre-set habituation criterion ($n = 6$); because of parental interference ($n = 3$); technical problem ($n = 1$); or because they cried during the experiment ($n = 9$). Parents signed an informed consent form. This research was approved by the local ethics committee.

Parents were asked to complete a French adaptation of the long version of the MacArthur CDI (Kern & Gayraud, 2010) within 1 week of participation, as a measure of infants' receptive and productive vocabularies. MCDIs were obtained for 34/48 infants. Reported comprehension vocabulary ranged from 10 to 502 words ($M = 184$, $SD = 98.6$) and reported production ranged from 4 to 128 words ($M = 24$, $SD = 29.1$). No correlation was found

between infants' vocabulary size and performance in this experiment (see supplementary material in the OSF).

At the beginning of this project, we had planned (and pre-registered) to test 16 participants in experimental group (32 total), based on the number of participants tested in previous studies using the same design (e.g., He & Lidz, 2017; Werker et al., 1998). However, recent studies investigating the effects of sample size and statistical power in infant looking time research (e.g., Oakes, 2017) suggest that infant studies with sample sizes inferior to 24 participants per cell can be underpowered and result in false positive and false negative results. We thus ran a power analysis based on the effect size observed in He & Lidz (2017; Cohen's $d = 0.802315$), which resulted in 25.38 participants in each condition to reach a significance level $< .05$ and a power of 80%, with a two-sided alternative. Since counterbalancing requires a number of participants divisible by eight, we decided to test a final sample of 24 infants in each condition (closest to 25.38).

Design. A Habituation-Switch paradigm (e.g., Werker et al., 1998; He & Lidz, 2017; cf.: Figure 1) was used to habituate infants with two video stimuli showing a penguin doing two different actions (one-participant agentive actions - e.g., spinning, cartwheeling), one in each video. During the presentation of one of the videos (e.g., a penguin spinning), infants heard sentences using a novel word as a noun (e.g., “*Regarde! C’est une bamoule!*” – “Look! It’s a bamoule!”, where *bamoule* is naming an object, the penguin), and during the presentation of the other video (e.g., a penguin cartwheeling), they heard sentences presenting another novel word as a verb (e.g., “*Regarde! Elle doripe!*” – “Look! She is doriping!”, where *doripe* is naming an action, cartwheeling). Note that the syntactic category of the novel words and the associations with the videos were counterbalanced across participants. Thus half of the participants had *bamoule* as a noun and *doripe* as a verb, and half had the reverse. Half had spinning as the verb meaning, and half had cartwheeling as the verb meaning. This

habituation phase gave the infants the opportunity to guess a possible meaning for each of the two novel words: if they exploit the linguistic context provided by the function words as adults would, then they should infer that the novel word employed as a noun (e.g., *C'est une bamoule* – ‘It’s a bamoule’) refers to the penguin (the only object present in the video), and that the novel word employed as a verb (e.g., *Elle doripe* – ‘She is doriping’) refers to the action that the penguin was doing (e.g., spinning or cartwheeling, counterbalanced across participants). When infants reached a pre-defined habituation criterion (three consecutive trials for which the average looking time was less than 65% of the average looking time for the most-attended three consecutive trials), the habituation phase stopped and the test phase started immediately.

At test, all infants were presented with a fixed number of 2 trials in which they were exposed to a switch of the audio tracks of the videos (as illustrated in Figure 1), such that half of the infants saw a Noun-Switch condition (e.g, hearing noun sentences with *bamoule*, while watching the penguin cartwheeling) and half saw a Verb-Switch condition (e.g., hearing sentences with the verb *doripe*, while watching the penguin spinning). Given that the noun sentences refer to an object (i.e., a penguin, present in both videos), while the verb sentences refer to an action (i.e., either “spinning” or “cartwheeling”), if children correctly used the linguistic context to infer the meanings of the novel noun and the novel verb, they should be more surprised (look more toward the videos) in the Verb-Switch condition than in the Noun-Switch condition. Indeed, changing the action in the Verb-Switch condition violates the inference constructed about the verb meaning: “cartwheeling” and “spinning” are different actions. However, the Noun-Switch condition does not violate the inference constructed about the noun meaning, because infants can still see a penguin in the video (and the fact that the action has changed is irrelevant). In other words, if children correctly exploit the linguistic context of the novel words to infer their meaning, we expect an asymmetrical pattern of results: more dishabituation in the Verb-Switch than in the Noun-Switch condition. Such an

asymmetry between conditions, if we observe it, can only be due to the asymmetry in the linguistic contexts in which the novel words are presented (verb context vs noun context).

Indeed, if infants attempted to associate the novel words to some aspect of the video without taking into account the linguistic context, then the results should come out symmetrical (either dishabituation in both conditions, or no dishabituation in both conditions).

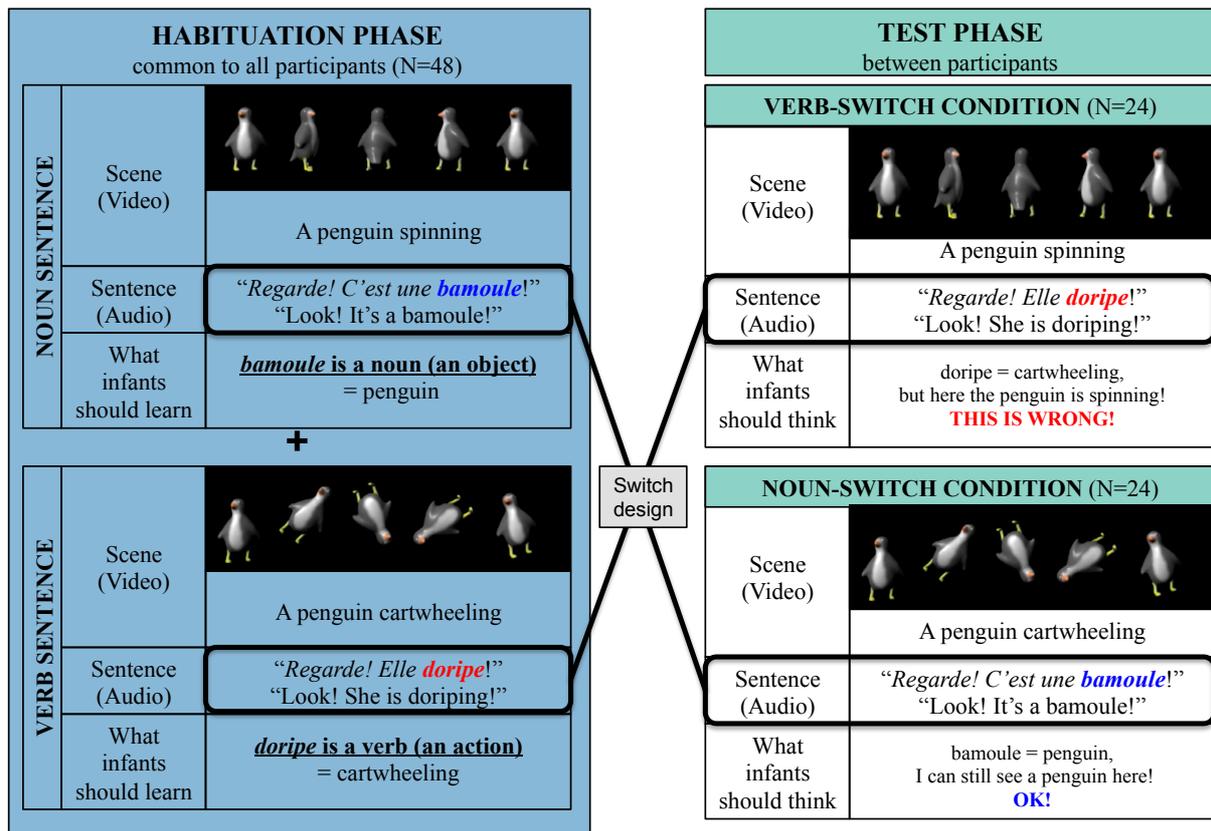


Figure 1: Experimental Design: Habituation-Switch paradigm, used in Experiments 1 and 2. All the infants were habituated with the same two video stimuli showing a penguin doing two different actions (e.g., spinning, cartwheeling, one in each video), while listening to noun sentences in which a novel word was used as a noun, and to verb sentences, in which another novel word was used as a verb. Then, as soon as infants reached a pre-defined habituation criterion, they were exposed to a test phase with a switch between the sentences and the videos, such that half of the children saw a Noun-Switch condition and half saw a Verb-Switch condition. Given that the noun refers to an object – the penguin – while the verb refers to an action, if children correctly use function words (Experiment 1) and phrasal prosody (Experiment 2) to infer the meanings of the novel noun and the novel verb during the habituation phase, at test, they should be more surprised (and look more toward the videos) in the Verb-Switch condition (because the action changed, which is problematic for the verb interpretation) than in the Noun-Switch condition (where they can still see a penguin in the video, and the fact that the action has changed is irrelevant). Note that the syntactic category of the novel words and the associations with the videos were counterbalanced across participants. Thus half of the children tested had *bamoule* as a noun and *doripe* as a verb, and half had the reverse. Half had *spinning* as the verb meaning, and half had *cartwheeling* as the verb meaning.

Material. Two novel words in French (*bamoule*; *doripe*) were used as target words. For each novel word, two sentences were created: one sentence using it as noun (e.g., “Oh regarde! C’est une bamoule! Tu la vois la bamoule?” – “Oh look! It’s a bamoule! Do you see the

bamoule?”) and another one using it as a verb (e.g., “*Oh regarde! Elle bamoule! Tu la vois qui bamoule?*” – “Oh look! She is bamouling! Do you see her bamouling?”). In order to create the audio tracks of the videos, each sentence was repeated twelve times, resulting in a 37-second-long passage for each target word in each condition; each repetition was introduced by an audio prompt (e.g., “Oh”; “Wow”; “Hey”). All the passages had exactly the same audio prompts. The assignment of target words to syntactic categories was counterbalanced across participants, such that half had the target word *bamoule* as a noun and *doripe* as a verb, and half had the reverse. All the stimuli were recorded by a female native speaker of French (last author) who recorded the sentences in child-directed register.

The presentations of these passages were paired with two video stimuli showing a penguin doing two different actions (e.g., spinning, cartwheeling), one in each video. These videos had exactly the same duration as the audio tracks (37 seconds).

Additionally, a silent video of a butterfly perched on a leaf was used as an attention-getter to recapture infant’s attention when they looked away for more than 2 seconds.

Apparatus and procedure. Infants were tested individually in a sound-attenuated double-walled booth, they sat on a parent’s lap, facing a 27-inch TV-screen positioned 70cm away from them. A camera positioned on the top of the TV-screen was connected with an LCD monitor placed outside the cabin: the experimenter observed the infant’s eye fixation to the screen during the experiment and coded their looking behavior online (pressing a button on a keyboard when the infant was looking at the screen, and releasing it as soon as the infant looked elsewhere). Parents wore headphones and listened to masking music during the entire experiment. The presentation of the stimuli and the online coding were controlled and recorded through the Habit program, version 1.0 (Cohen, Atkinson, & Chaput, 2004).

The experiment was composed of two phases: habituation and test. The procedure started by displaying the attention-getter on the screen (i.e., the silent video of a butterfly

perched on a leaf). Once the child looked toward the screen, the experimenter, who was outside the cabin monitoring the child's looking behavior through a video camera, initiated the first trial. The experimenter was blind about the type of trial (Habituation vs Test). During each trial, the experimenter pressed a button when the toddler looked toward the screen, and released it when the toddler looked away. If the toddler reoriented toward the screen within 2 seconds, the trial continued to play, but the time spent looking away was subtracted from their looking time. Each trial lasted until the child looked away for more than two seconds, or until the maximum length of the trial was reached (i.e., 37 seconds).

During the habituation phase, the videos were presented repeatedly one after the other, for as much time as the child wanted to look at the TV-screen, for a minimum of four trials and a maximum of twelve trials, depending on how fast the child reached the pre-defined habituation criterion. This criterion was reached when an infant's average looking time during any block of 3 consecutive trials dropped to less than 65% of the average looking time for the most-attended block (i.e. the 3-trial block that had the longest total looking time). Habituation trials were randomized by blocks of two to avoid the same action-sentence pair occurring more than twice in a row, and to ensure that the number of trials for each action-sentence pair was as balanced as possible, independently of the duration of the habituation phase. Once infants reached the habituation criterion, the habituation phase stopped and the test phase started immediately.

At test, all infants were presented with a fixed number of 2 trials in which the audio sentences of the videos presented during the habituation phase were switched, such that participants assigned to the Noun-Switch condition saw two trials in which the noun sentence was presented together with the video previously associated with the verb sentence, while participants assigned to Verb-Switch condition saw two trials in which the verb sentence was presented together with the video previously associated with the noun sentence. Half of the participants were assigned to the Noun-Switch condition and half to the Verb-Switch

condition.

Data processing and analyses. Data analyses and graphics were performed with R software version 3.2.2 (R Team, 2015). We used the average looking time of the last two trials of the habituation phase, and of the two test trials, and we compared the increase in looking time from habituation to test in the two experimental conditions (Noun-Switch vs Verb-Switch). If infants are able to exploit function words to access the syntactic structure of the sentences in which the novel words occurred, and if they can use this information to infer the syntactic category of the novel words and constrain their possible meaning, we expect a greater increase in looking time from habituation to test in the Verb-Switch condition than in the Noun-Switch condition. To test this, we performed an ANOVA on log-transformed mean looking times as the dependent measure, with participants as the random factor, Condition (Noun-Switch vs Verb-Switch) as a between-participant factor, and Phase (Habituation vs Test) as a within-participant factor. The expected effect should appear as a significant interaction between Condition and Phase. Note that looking times were log-transformed before running the ANOVA, because the data did not follow a normal distribution, which is a necessary condition to conduct an ANOVA.

Results

The results of Experiment 1 are shown in Figure 2.

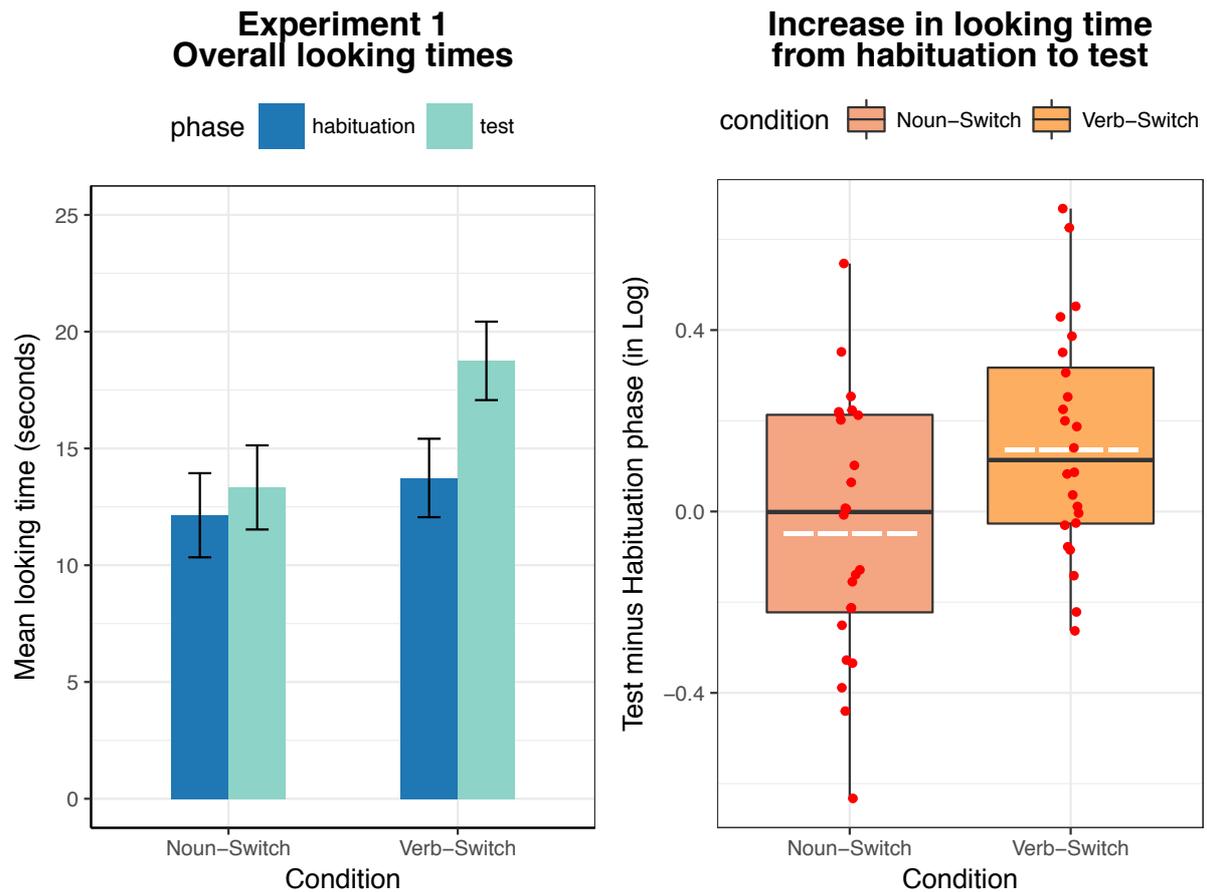


Figure 2: On the left side: Mean looking time in seconds toward the videos during the last two trials of the habituation phase (in blue) and during the two trials of the test phase (in green) for children assigned to the Noun-Switch Condition (N=24) and to the Verb-Switch Condition (N=24). Error bars represent the standard error of the mean of the differences between Test minus Habituation. On the right side: Boxplot of the increase in log-transformed mean looking times from habituation to test in each group. Red dots represent the average for each participant in each group. White dashed lines represent the means of the distributions.

Infants looking times increased more between habituation and test in the Verb-Switch condition than in the Noun-Switch condition: An ANOVA on log-transformed mean looking times revealed a significant interaction between Condition (Noun-Switch vs Verb-Switch) and Phase (Habituation vs Test), $F(1,46) = 5.65, p = .022; d = 0.665$, which confirms that relative to the habituation phase, at test children looked more toward the videos in the Verb-Switch condition than in the Noun-Switch condition. This is consistent with the interpretation that during the habituation phase, infants inferred that the novel verb referred to the action (e.g., cartwheeling), while the novel noun referred to the object (the penguin). Thus, at test, infants in the Verb-Switch condition were surprised when watching the penguin performing the other action (e.g., spinning), while listening to sentences containing the verb that they had associated to another action (e.g., cartwheeling) during the habituation phase. For instance,

since the penguin was not cartwheeling but spinning, there was a discrepancy between the original meaning they had inferred for this word, and the current situation. Note that this increase in looking time can be explained either by the fact that infants thought that the target word was not used correctly at test (i.e., the speaker was making a mistake and using the word incorrectly), or by the fact that infants realized that they had to broaden the meaning that they had initially inferred for this word (e.g. instead of meaning cartwheeling specifically, *doripe* might refer to a broader class of movements, perhaps involving rotation – which is present in both actions). In contrast, infants tested in the Noun-Switch condition, who listened to noun sentences with *bamoule* while watching the penguin cartwheeling, did not show surprise during test: indeed the meaning they had inferred during habituation (that *bamoule* means penguin), was perfectly consistent with the test video they were watching, since there was still a penguin on the screen.

Consistently with the predictions from distributional learning theories mentioned in the introduction, these results suggest infants exploited the information carried by function words to constrain their interpretation of novel content-word meanings. Given that both groups were exposed to exactly the same videos and sentences during habituation, the only way to explain the asymmetry observed at test is that infants paid attention to the syntactic context instantiated by function words to correctly assign a syntactic category to the novel words and constrain their meanings. When we switched the audio tracks of the videos, this caused a violation of the inference that infants constructed about the verb meaning (i.e., “cartwheeling” and “spinning” are different actions), but not the inference about the noun meaning (i.e., it is still the same penguin in both videos).

Discussion

Experiment 1 shows that 18-month-olds use function words to compute the syntactic category of novel words and to constrain their probable meanings. However, because not all

content words are directly preceded by function words, this strategy may not always be sufficient. Since the prosodic structure of an utterance correlates with its syntactic structure, listeners could exploit prosodic boundaries together with function words to constrain syntactic analysis (Christophe et al., 2016, 2008; Morgan & Demuth, 1996), a hypothesis tested in Experiment 2.

Experiment 2: Phrasal prosody and function words constrain the acquisition of word meanings

This experiment investigated whether infants take into account the position of a word within the prosodic structure of a sentence when computing its syntactic category (noun vs. verb).

Method

Participants. Forty-eight French-learning 18-month-olds participated in the study, 24 in each experimental group (mean age = 18.2 months, range = 17.8 to 18.8 months; $SD = 0.2$; 24 girls). An additional twenty-three infants came to the lab, but were not included in the sample: because of fussiness not allowing them to finish the experiment ($n = 11$); because they failed to meet the pre-set habituation criterion ($n = 4$); because of parental interference ($n = 1$); technical problem ($n = 2$); or because they cried during the experiment ($n = 5$). Parents signed an informed consent form. This research was approved by the local ethics committee. Parents were asked to complete a French adaptation of the long version of the MacArthur CDI (Kern & Gayraud, 2010) within 1 week of participation as a measure of each infant's receptive and productive vocabularies. MCDIs were obtained for 35/48 infants. Reported comprehension vocabulary ranged from 13 to 684 words ($M = 213$, $SD = 125.3$) and reported production ranged from 1 to 225 words ($M = 33$, $SD = 43.3$). No correlation was found between infants' vocabulary size and performance.

Design. We used exactly the same paradigm used in Experiment 1, but now we investigated whether, in addition to function words, infants are able to rely on the relationship between the prosodic and syntactic structure of sentences to guide their syntactic interpretation and constrain the acquisition of word meanings. Instead of having a function word (e.g., article or pronoun) immediately preceding the to-be-learned word and unambiguously cueing its syntactic category, children had to take into account the prosodic structure in which a novel word appeared to compute its syntactic category. For instance, the novel word *bamoule* was presented as a noun in the sentence: [*Regarde la petite bamoule*]! – [Look at the little bamoule]! and the novel word *doripe* was presented as a verb in the sentence: [*Regarde*]! [*la petite*] [*doripe*]! – [Look]! [The little one] [is doriping]! (square brackets indicate prosodic phrase boundaries). As an illustration, note that we can find similar examples in English, with sentences such as [Do you see the baby flies?], where ‘flies’ is a noun, naming the insect, versus [Do you see?] [the baby] [flies!], where ‘flies’ is a verb, naming the action that the baby is doing.

Since in this experiment both sentences are composed of the same words and functional elements in the same order (*regarde-la-petite-bamoule/doripe*), an analysis in terms of which words precede *bamoule* or *doripe* is not sufficient to determine its syntactic category (since they are the same in both conditions): rather, the syntactic difference between these two sentences is reflected in their different prosodic structures. When *doripe* is a verb, there is a prosodic boundary preceding it (i.e., the boundary between the noun phrase and the verb phrase) while when *bamoule* is a noun, it is embedded in a single prosodic unit together with the other words of the sentence, corresponding to the verb ‘look’ and the following noun phrase. If infants are able to use the information provided by the prosodic structure of a sentence to access its syntactic structure as adults and preschoolers do (de Carvalho, Dautriche, & Christophe, 2016; de Carvalho, Lidz, Tieu, Blean, & Christophe, 2016;

Millotte, Wales, & Christophe, 2007; Snedeker & Yuan, 2008), and if they can use this information to constrain the meaning of novel words, during the habituation phase they should infer that the novel word used as a noun refers to the penguin (the only object present in the video), and that the novel word used as a verb refers to the action that the penguin is doing (e.g., spinning or cartwheeling, counterbalanced across participants). Thus, as in Experiment 1, during the test phase, we expect infants to look more toward the video (being more surprised) in the Verb-Switch condition than in the Noun-Switch condition.

Material. The same two novel words in French (*bamoule*; *doripe*) were used as target words to create minimal pairs of sentences that differed only in their prosodic structures. Thus, for each novel word two sentences were created, one presenting the target word in a noun position within the prosodic-syntactic structure (Noun sentence: [*Regarde la petite bamoule*]! [*Tu vois la petite bamoule*]? – [Look at the little bamoule]! [Do you see the little bamoule]?, and another one presenting the novel word in a verb position (Verb sentence: [*Regarde*]! [*la petite*] [*bamoule*]! [*Tu vois*]? [*la petite*] [*bamoule*]! - [Look]! [The little one] [is bamouling]! [Do you see]? [The little one] [is bamouling]! (square brackets indicate prosodic boundaries). An example of each kind of sentence is depicted in Figure 3.

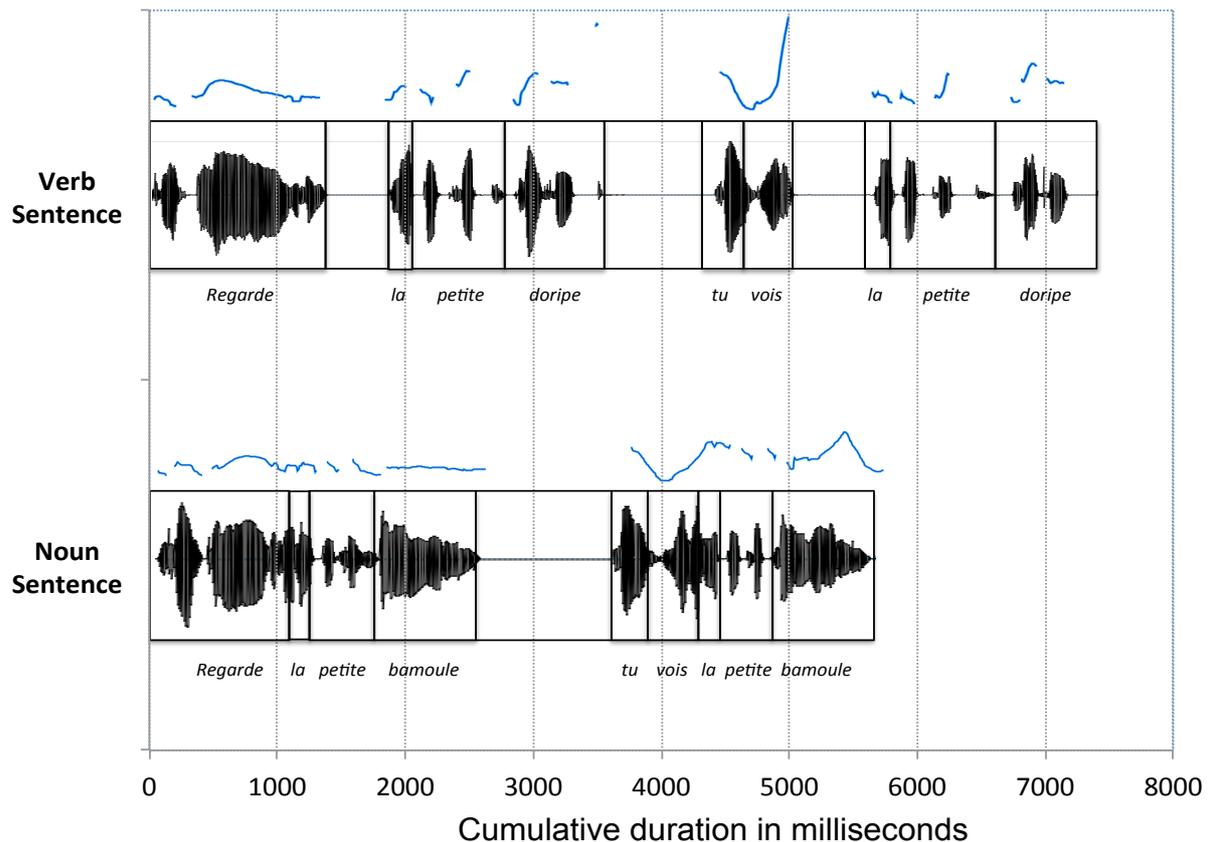


Figure 3: An example of the minimal pair of sentences created for Experiment 2, with the Verb sentence on top and the Noun sentence at the bottom. The x-axis represents the time-course, for each sentence. From bottom to top, we present the transcription, the waveform, and the pitch contour (blue curve) of each sentence. Both sentences were composed of exactly the same words (*la-petite-bamoule/doripe*) but differed only in their prosodic structure, which reflected their different syntactic structures. In Noun sentences all the critical words were grouped together into a single prosodic unit [*Regarde la petite bamoule*]! - [Look at the little bamoule]!; In contrast, in Verb sentences these words were spread into three different prosodic units [*Regarde*]! [*la petite*] [*bamoule*]! - [Look]! [The little one] [is bamouling]!. The syntactic category of the novel words and the associations with the videos were counterbalanced across participants. Thus half of the participants had *bamoule* as a noun and *doripe* as a verb, and half had the reverse. Half had spinning as the verb meaning, and half had cartwheeling as the verb meaning.

Sentences uttered in the verb condition had a phonological phrase boundary before the target word (i.e., corresponding to the boundary between the noun and the verb phrase); in contrast, in sentences uttered in the noun condition all the words were grouped together into a single prosodic unit; these prosodic structures are consistent with theoretical descriptions of the relationship between prosodic and syntactic boundaries (e.g., Nespor & Vogel, 1986).

In order to create the audio tracks, each sentence was repeated twelve times, resulting in a 50-second-long passage for each target word in each condition; each repetition was introduced by an audio prompt (e.g., “Oh”; “Wow”; “Hey”). As in Experiment 1, the assignment of a syntactic category to the two novel words was counterbalanced across

participants. All the stimuli were recorded by the same speaker as Experiment 1, in child-directed register. The audio tracks were paired with the same video stimuli used in Experiment 1 (but with a duration of 50 seconds). The same silent video of a butterfly perched on a leaf was used as attention-getter.

Acoustic analyses

To assess prosodic differences between the two conditions, acoustic measurements (duration and pitch) were conducted on each of the 8 exemplars of the test sentences (*bamoule* as a noun and *bamoule* as a verb, *doripe* as a noun and *doripe* as a verb, with the novel word repeated twice in each case, see Figure 3). As expected from the literature (e.g., Jun & Fougeron, 2002), the analysis of duration revealed a significant pre-boundary lengthening: the rhyme of the word preceding the target word (e.g., “ite” from *petite*) in the verb condition (where it was placed just before the prosodic phrase boundary) was lengthened by 211% compared to this same segment in the noun condition (where it was placed in the middle of a prosodic unit; 381 vs 130 ms, see Table 1). A silent pause of 62 ms preceding the target word (i.e., between “petite” and “bamoule”) was observed in the verb condition, while there was no pause between these words in the noun condition.

The analysis of pitch contours in both sentence conditions revealed a significant difference between conditions (see Table 1), consistent with the literature describing French as having a tendency for a rising pitch contour towards the end of prosodic units (e.g., Welby, 2006). Thus, the word preceding the target word (i.e., “petite”) exhibited a greater rising pitch pattern in the verb prosody condition (+150Hz; because of its position at the end of a prosodic unit), than in the noun prosody condition (-29Hz; when it was placed in the middle of a prosodic unit). Given that in both conditions, the target word was placed at the end of a prosodic unit, no particular hypothesis was made regarding their differences in pitch and or duration at the final position. The target word in the noun prosody condition (e.g., “bamoule”)

seemed to exhibit a greater rising pitch pattern in the noun prosody condition (+66 Hz) than in the verb prosody condition (+25 Hz), but this difference was not significant.

Table 1.

Acoustic analyses of the stimuli of Experiment 2. Mean duration (in ms) and pitch (in Hz) for the segments around the prosodic boundaries for both noun and verb sentence conditions.

Duration analysis – mean duration in ms (standard error)			
Dependent variable	Noun sentence [la petite <i>bamoule</i>]	Verb sentence [la petite] [<i>bamoule</i>]	Analysis (2-tailed t-tests)
Rhyme - word preceding Target (e.g., ‘ite’ from “petite”)	130 (2.3)	381 (33.9)	$t(3) = -7.35, p = .005^{***}$
Pause - before Target (e.g., between “petite” and “bamoule”)	0 (0)	62 (10.9)	$t(3) = -5.74, p = .010^{**}$
Onset - Target word (e.g., ‘b’ from “bamoule”)	97 (10.8)	127 (5.9)	$t(3) = -1.79, p = .171$
Rhyme – Target word (e.g., ‘oule’ from “bamoule”)	533 (35.8)	443 (8.6)	$t(3) = 2.71, p = .073$

Pitch analysis – Mean pitch change, in Hz, from the beginning to the end of the target words (standard error of the mean).			
Dependent variable	Noun sentence [la petite <i>bamoule</i>]	Verb sentence [la petite] [<i>bamoule</i>]	Analysis (2-tailed t-tests)
Word preceding Target (e.g., last pitch value at the last vowel from “petite” minus first pitch value from the first vowel of “petite”)	-29.2 (3.1)	150 (27.7)	$t(3) = -6.39, p = .008^{***}$
Target word (e.g., last pitch value of ‘u’ from “bamoule” minus first pitch value of ‘a’ from “bamoule”)	66.7 (42.2)	25 (12.7)	$t(3) = 0.77, p = .497$

Note that in previous studies (e.g., Fisher & Tokura, 1996; Soderstrom, Blossom, Foygel, & Morgan, 2008), acoustical analysis of mothers’ spontaneous speech addressed to American English-learning infants show that they naturally produce the kind of acoustical cues associated to prosodic boundaries (both utterance- and internal phrase-level boundaries), such as the one we exploited here, including reliable prosodic cues to grammatical units such as Subject–Verb Phrase boundaries. In French, a production study with adults found that they naturally produce the acoustic cues associated to the prosodic boundaries between noun and verb phrases that we exploited in the current experiment (Millotte et al., 2007). Taken together these studies suggests that the relationship between prosodic and syntactic boundaries in our study may be present in children’s everyday spoken input and could be learnt through exposure to language.

Apparatus and procedure. Same as Experiment 1, the only difference concerns the sentences uttered during the presentation of the videos. During the presentation of one of the videos (e.g., the penguin spinning) infants listened to sentences presenting a novel word in a noun position within the prosodic-syntactic structure (Noun sentence: [*Regarde la petite bamoule*]! – [Look at the little bamoule]!, where *bamoule* is a noun, naming an object, the penguin); during the presentation of the other video (e.g., the penguin cartwheeling), they listened to sentences presenting the novel word in a verb position within the prosodic-syntactic structure (Verb sentence: “[*Regarde*]! [*la petite*] [*doripe*]! - [Look]! [The little one] [is doriping]!, where *doripe* is a verb, naming an event, cartwheeling; square brackets indicate prosodic boundaries). As in Experiment 1, in each trial infants had the opportunity to listen to a maximum of twelve repetitions of the test sentences (for a total duration of 50s, since adding the word “petite” made the sentences longer). Half of the children were tested in the Noun-Switch condition and the other half in the Verb-Switch condition.

Data processing and analyses. Same as Experiment 1.

Results

The results of Experiment 2 are shown in Figure 4.

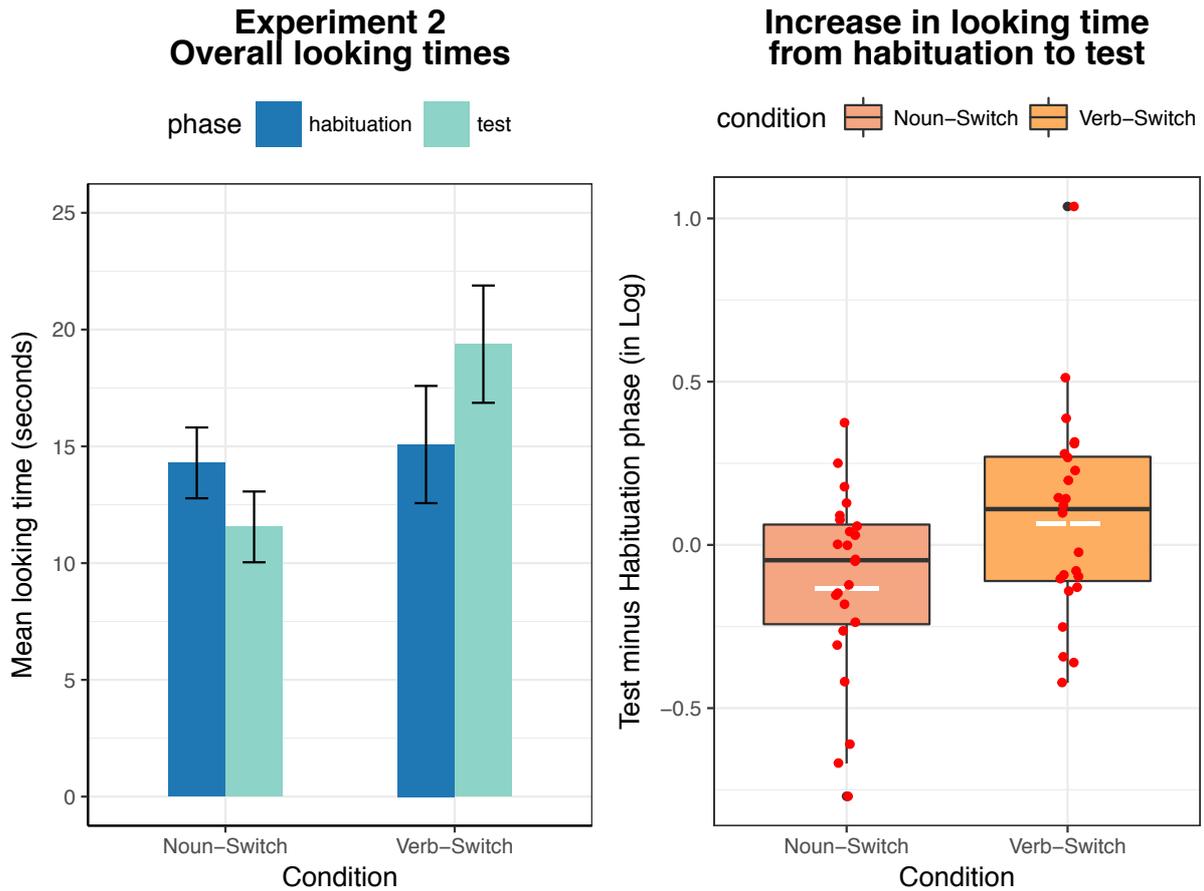


Figure 4: On the left side: Mean looking time in seconds toward the videos during the last two trials of the habituation phase (in blue) and during the two trials of the test phase (in green) for children assigned to the Noun-Switch Condition (N=24) and to the Verb-Switch Condition (N=24). Error bars represent the standard error of the mean of the differences between Test and Habituation. On the right side: Boxplot of the increase in log-transformed mean looking times from habituation to test in each group. Each red dot represents one participant. White dashed lines represent the means of the distributions.

Infants looking time increased more between habituation and test in the Verb-Switch condition than in the Noun-Switch condition: An ANOVA on log-transformed mean looking times revealed a significant interaction between Condition and Phase: $F(1,46) = 5.09$, $p = .029$; $d = 0.632$, showing that infants looked longer (were more surprised) when tested in the Verb-switch phase condition than in the Noun-switch condition. This behavior, just like in Experiment 1, suggests that the switch of the actions led to a violation of the inference constructed about the verb meaning, but not about the noun meaning, consequently infants were more surprised when listening to verb sentences than to noun sentences during the test phase.

Looking at the results of Exp. 1 and 2 together, one may note a lesser increase in looking times between habituation and test in Exp. 2 relative to Exp. 1, which surfaces in Exp.

2 as a small decrease between habituation and test in the Noun-switch condition, and a smaller increase than in Exp. 1 between habituation and test in the Verb-switch condition. This might be due to the fact that experimental trials were longer in Exp. 2 than in Exp. 1 (50s vs. 37s), which means that toddlers could reach the habituation criterion with longer average looking times (14.68s ($SD=6.6$) in Exp. 2 vs. 12.94s ($SD=4.86$) in Exp.1), which would leave them more room to go down further at test. This remains speculative, as we do not know if such a difference would replicate (and had not planned for making such a comparison in our pre-registration). Note however that this does not impact our main result, since the expected effect was a significant interaction between Condition and Phase, revealing a greater increase in looking time from habituation to test in the Verb-Switch condition compared to the Noun-Switch condition.

The present result thus shows that at 18 months, infants are already able to use phrasal prosody as a cue to interpret a novel word as either a noun or a verb depending on its position within the prosodic structure of sentences. When listening to minimal pairs of sentences such as *Regarde-la-petite-bamoule/doripe*, which can be produced either as a single prosodic unit, as in [*Regarde la petite bamoule*]! – [Look at the little bamoule]!, where *bamoule* is used as a noun, or as three prosodic units such as [*Regarde*]! [*la petite*] [*doripe*]! – [Look]! [the little one] [is doriping]!, where *doripe* is used as a verb, 18-month-olds correctly interpreted the target word as either a noun or a verb, depending on its position within the prosodic structure of the sentence, and they used this information to constrain the meaning of this novel word.

It is important to note that the noun and verb sentences had exactly the same words (i.e., *Regarde-la-petite-bamoule/doripe*), thus a distributional analysis tracking which functional element was preceding and/or following the target word would not have been sufficient to constrain infants' interpretation in this experiment. This does not suggest however that infants did not exploit function words together with prosodic information to constrain their interpretations. Rather, we believe that prosody and function words were

jointly exploited in this experiment, because while prosodic boundaries allowed infants to group words into syntactic constituents and informed them about the location of syntactic boundaries, the prosodic boundaries per se do not directly provide the syntactic labels of constituents (e.g. noun phrase, verb phrase). To interpret the novel words as a noun or as a verb, infants had to use the information carried by the function words inside the prosodic units to determine the syntactic nature of these constituents delimited by prosody. For instance, when participants heard a sentence such as [*Oh regarde!*] [*la petite*] [*bamoule*], the prosodic boundary before the target word “*bamoule*” signaled the presence of a syntactic constituent boundary. Given that the first prosodic unit (e.g., [*la petite*]) started with an article (e.g., *la* – the), this unit could be identified as a noun phrase (e.g., [*La*_{DET} *petite*_{NOUN}]_{NP} - [The_{DET} little one]_{NP}). Having identified the first unit as a full noun phrase, infants might have expected it to be followed by a verb phrase, which allowed them to interpret *bamoule* as a verb. In the noun condition in contrast, infants interpreted the novel word as a noun because all the three words (*la+petite+bamoule*) appeared together into a single prosodic unit with the well-known verb *regarde* (e.g., [*Regarde la petite bamoule*] – [Look at the little bamoule]).

Importantly, these results show that even in a situation where the information provided by function words alone was not sufficient to compute the syntactic category of the novel words, infants were able to exploit prosodic information to recover the syntactic structure of sentences and, in combination with the other words (*regarde-la-petite...*), infer the syntactic category of novel words and therefore constrain their meanings.

Discussion

These results show that 18-month-olds are able to use phrasal prosody, together with function words, to recover the syntactic structure of sentences and to interpret a novel word as either a noun (referring to an object) or a verb (referring to an action) depending on its position within the prosodic-syntactic structure.

General Discussion

Across two experiments, we demonstrated that 18-month-olds, who are still in the process of learning the syntax and building up the lexicon of their language, can rely on function words and phrasal prosody to access the syntactic structure of sentences and guide their discovery of novel word meanings. In Experiment 1, French infants exploited the functional elements in a sentence to assign a syntactic category to a novel word and to constrain its meaning. In Experiment 2, where the information provided by function words alone was not sufficient to compute the syntactic category of the novel words, infants simultaneously exploited prosodic information together with function words to recover the syntactic structure of sentences. In our study, infants used acoustic/prosodic information to parse spoken sentences into groups of words and identify possible syntactic constituents; they exploited the function words inside these prosodic-syntactic constituents to determine the syntactic nature of these constituents, which in turn allowed them to infer the syntactic category of novel words, and therefore constrain their meanings (associating nouns to objects, and verbs to actions).

Infants' ability to use function words and phrasal prosody to identify noun and verb contexts is truly impressive, and raises the question of *how* infants may have learned which contexts go with nouns and which go with verbs. Some studies suggest that infants could rely on a handful of known words, to act as a *seed* for learning noun and verb categories (Gutman et al. 2015; Brusini, Amsili, Chemla, & Christophe, 2011; Christophe et al., 2016; see also Yarowsky, 1995). This hypothesis relies on two established premises: first, infants know the meaning of a few highly frequent content words (e.g., Bergelson & Swingley, 2012); second, they group concepts into semantic categories (e.g. objects, actions, agents - Carey, 2009). If, in addition, infants expect words from similar conceptual categories to occur in similar syntactic environments (Gleitman, 1990; Pinker, 1984), they could exploit the syntactic

contexts in which the few words they already know appear, in order to infer information about other unknown words that appear in the same contexts. For instance, having observed that the words “bottle” and “teddy bear” label an object kind and often occur after “the”, as in “the bottle”, “the teddy bear”, infants could infer that other words occurring in that syntactic environment will share conceptual properties as well (e.g., “blick” in “the blick” might also refer to an object kind). Taken together, phrasal prosody, function words and a distributional analysis of the contexts in which known words appear may be extremely useful to access the syntactic category of unknown words and allow infants to bootstrap their acquisition of word meanings.

This powerful mechanism might provide infants with a tool to construct a first-pass syntactic structure of spoken sentences, during the first steps of language acquisition. In this sense, the answer we provide to the chicken-and-egg problem of learning word meanings through syntax, and learning syntax through word meanings, is that by providing information about the syntactic structure of sentences, phrasal prosody and function words can work as anchors to help infants access syntactic information. Crucially, since infants are sensitive to phrasal prosody and function words during the first year of life, the ability to jointly exploit phrasal prosody and function words to access syntactic structure may be in place even before infants know many words.

Our results were obtained with French, but we expect that phrasal prosody and function words should support an early access to syntax in many different languages. Although prosodic information and functional elements can surface differently across languages, this information is present in all the world’s languages (Dryer, 1992; Shattuck-Hufnagel & Turk, 1996). Overall, we suggest that phrasal prosody and function words may well represent a universal and extremely useful tool for infants to access syntactic information through a surface analysis of the speech stream, and to bootstrap their way toward successful language acquisition.

Author Contributions

A. de Carvalho, A. He, J. Lidz and A. Christophe designed the study; A. de Carvalho performed research; A. de Carvalho analyzed data; A. de Carvalho wrote the paper, and A. Christophe, J. Lidz and A. He provided critical revisions.

Acknowledgements

The authors thank Alejandrina Cristia, Sid Kouider and Alexander Martin for their thoughtful comments on previous versions of this manuscript.

Declaration of Conflicting Interests

The authors declare that they had no conflicts of interest.

Funding

This research was supported by a PhD fellowship from the École normale supérieure and by a postdoctoral fellowship from the Fyssen Foundation to Alex de Carvalho. It was also supported by grants from the Région Ile-de-France, Fondation de France, LabEx IEC (ANR-10-LABX-0087), IdEx PSL (ANR-10-IDEX-0001-02), and the ANR ‘Apprentissages’ (ANR-13-APPR-0012).

References

- Bergelson, E., & Swingle, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(9), 3253–8. <http://doi.org/10.1073/pnas.1113380109>
- Bernal, S., Lidz, J., Millotte, S., & Christophe, A. (2007). Syntax Constrains the Acquisition of Verb Meaning. *Language Learning and Development*, *3*(4), 325–341. <http://doi.org/10.1080/15475440701542609>

Brusini, P., Amsili, P., Chemla, E., & Christophe, A. (2011). Learning to categorize nouns and verbs on the basis of a few known examples: A computational model relying on 2-word contexts. In *Paper presented at Society for Research on Child Development Biennial meeting*. Montreal (Canada).

Carey, S. (2009). *The origin of concepts*. Oxford University Press.

Cauvet, E., Limissuri, R., Millotte, S., Skoruppa, K., Cabrol, D., & Christophe, A. (2014). Function Words Constrain On-Line Recognition of Verbs and Nouns in French 18-Month-Olds. *Language Learning and Development*, *10*(1), 1–18.
<http://doi.org/10.1080/15475441.2012.757970>

Christodoulopoulos, C., Roth, D., & Fisher, C. (2016). An incremental model of syntactic bootstrapping. In *Proceedings of the 7th Workshop on Cognitive Aspects of Computational Language Learning* (pp. 38–43). Berlin, Germany.

Christophe, A., Dautriche, I., de Carvalho, A., & Brusini, P. (2016). Bootstrapping the syntactic bootstrapper . In J. Scott & D. Waughtal (Eds.), *Proceedings of the 40th Annual Boston University Conference on Language Development* (pp. 75–88). Somerville, MA: Cascadilla Press.

Christophe, A., Millotte, S., Bernal, S., & Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. *Language and Speech*, *51*(1–2), 61–75.
<http://doi.org/10.1177/00238309080510010501>

Cohen, L., Atkinson, D., & Chaput, H. (2004). Habit X: a new program for obtaining and organizing data in infant perception and cognition studies (version 1.0). Austin: University of Texas.

de Carvalho, A., Dautriche, I., & Christophe, A. (2016). Preschoolers use phrasal prosody

online to constrain syntactic analysis. *Developmental Science*, 19(2), 235–250.

<http://doi.org/10.1111/desc.12300>

de Carvalho, A., Lidz, J., Tieu, L., Blears, T., & Christophe, A. (2016). English-speaking preschoolers can use phrasal prosody for syntactic parsing. *The Journal of the Acoustical Society of America*, 139(6), EL216-EL222. <http://doi.org/10.1121/1.4954385>

Dryer, M. (1992). The greenbergian word order correlations. *Language*, 68(1), 81–138.

Fisher, C., & Tokura, H. (1996). Prosody in speech to infants: Direct and indirect acoustic cues to syntactic structure. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 343–363). Hillsdale, NJ: Erlbaum.

Gervain, J., Nespore, M., Mazuka, R., Horie, R., & Mehler, J. (2008). Bootstrapping word order in prelexical infants: A Japanese–Italian cross-linguistic study. *Cognitive Psychology*, 57(1), 56–74. <http://doi.org/10.1016/j.cogpsych.2007.12.001>

Gleitman, L. (1990). The Structural Sources of Verb Meanings. *Language Acquisition*, 1, 3–55.

Gutman, A., Dautriche, I., Crabbé, B., & Christophe, A. (2015). Bootstrapping the Syntactic Bootstrapper: Probabilistic Labeling of Prosodic Phrases. *Language Acquisition*, 22(3), 285–309. <http://doi.org/10.1080/10489223.2014.971956>

He, A. X., & Lidz, J. (2017). Verb Learning in 14- and 18-Month-Old English-Learning Infants. *Language Learning and Development*, 13(3), 335–356.
<http://doi.org/10.1080/15475441.2017.1285238>

Jun, S., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus*, 14(1), 147–172. <http://doi.org/10.1515/prbs.2002.002>

- Kern, S., & Gayraud, G. (2010). *Inventaire Français du Développement Communicatif (IFDC)*. La Cigale.
- Männel, C., & Friederici, A. D. (2009). Pauses and intonational phrasing: ERP studies in 5-month-old German infants and adults. *Journal of Cognitive Neuroscience*, *21*(10), 1988–2006. <http://doi.org/10.1162/jocn.2009.21221>
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertocini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*(2), 143–178. [http://doi.org/10.1016/0010-0277\(88\)90035-2](http://doi.org/10.1016/0010-0277(88)90035-2)
- Millotte, S., Wales, R., & Christophe, A. (2007). Phrasal prosody disambiguates syntax. *Language and Cognitive Processes*, *22*(6), 898–909. <http://doi.org/10.1080/01690960701205286>
- Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, *90*(1), 91–117. [http://doi.org/10.1016/S0010-0277\(03\)00140-9](http://doi.org/10.1016/S0010-0277(03)00140-9)
- Morgan, J., & Demuth, K. (1996). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (Eribaum). Mahwah, NJ.
- Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris.
- Oakes, L. M. (2017). Sample Size, Statistical Power, and False Conclusions in Infant Looking-Time Research. *Infancy*, *22*(4), 436–469. <http://doi.org/10.1111/infa.12186>
- Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, *25*(2), 193–247.

<http://doi.org/10.1007/BF01708572>

- Shi, R., & Melançon, A. (2010). Syntactic Categorization in French-Learning Infants. *Infancy, 15*(5), 517–533. <http://doi.org/10.1111/j.1532-7078.2009.00022.x>
- Shi, R., Morgan, J. L., & Allopenna, P. D. (1998). Phonological and acoustic bases for earliest grammatical category assignment: a cross-linguistic perspective. *Journal of Child Language, 25*, 169–201. <http://doi.org/10.1017/S0305000997003395>
- Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition, 72*(2), B11–B21. [http://doi.org/10.1016/S0010-0277\(99\)00047-5](http://doi.org/10.1016/S0010-0277(99)00047-5)
- Shukla, M., White, K. S., & Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences, 108*(15), 6038–6043. <http://doi.org/10.1073/pnas.1017617108>
- Snedeker, J., & Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language, 58*(2), 574–608. <http://doi.org/10.1016/j.jml.2007.08.001>
- Soderstrom, M., Blossom, M., Foygel, R., & Morgan, J. (2008). Acoustical cues and grammatical units in speech to two preverbal infants. *Journal of Child Language, 35*(04), 869–902. <http://doi.org/10.1017/S0305000908008763>
- Soderstrom, M., Seidl, A., Kemler Nelson, D. G., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language, 49*(2), 249–267. [http://doi.org/10.1016/S0749-596X\(03\)00024-X](http://doi.org/10.1016/S0749-596X(03)00024-X)
- Team, R. C. (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

- Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology*, 59(1), 67–95. <http://doi.org/10.1016/j.cogpsych.2009.02.001>
- Welby, P. (2006). French intonational structure: Evidence from tonal alignment. *Journal of Phonetics*, 34(3), 343–371. <http://doi.org/10.1016/j.wocn.2005.09.001>
- Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology*, 34(6), 1289–1309. <http://doi.org/10.1037/0012-1649.34.6.1289>
- Yarowsky, D. (1995). Unsupervised word sense disambiguation rivaling supervised methods. In *Proceedings of the 33rd annual meeting on Association for Computational Linguistics* (pp. 189–196). Morristown, NJ, USA: Association for Computational Linguistics. <http://doi.org/10.3115/981658.981684>