Phrasal prosody constrains syntactic analysis in toddlers

Phrasal prosody constrains syntactic analysis in toddlers

Alex de Carvalho[a,b*], Isabelle Dautriche[a,b,c], Isabelle Lin[d] and Anne Christophe[a,b]

a. Laboratoire de Sciences Cognitives et Psycholinguistique, DEC-ENS / EHESS / CNRS, Ecole normale supérieure – PSL Research University, Paris, France

b. Maternité Port-Royal, AP-HP, Faculté de Médecine Paris Descartes, France

c. School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, United Kingdom

d. Department of Linguistics, University of California, Los Angeles - USA

* Corresponding author: correspondence concerning this article should be addressed to Alex de Carvalho, Laboratoire de Sciences Cognitives et Psycholinguistique, Ecole normale supérieure, 29 rue d'Ulm – P.J, 75005 Paris – France. E-mail: x.de.carvalho@gmail.com

Research highlights

. Toddlers use phrasal prosody to constrain syntactic analysis

. From 20 months on, they disambiguate noun/verb homophones

. They rapidly integrate phrasal prosody during sentence processing

. This suggests that phrasal prosody cues syntactic structure for young children

. An early access to syntactic structure may facilitate the learning of word meanings

TODDLERS USE PROSODY TO CONSTRAIN SYNTACTIC ANALYSIS

Abstract

This study examined whether phrasal prosody can impact toddlers' syntactic analysis. French noun-verb homophones were used to create locally ambiguous test sentences (e.g., using the homophone as a noun: [*le bébé* ***souris***] [*a bien mangé*] - [the baby **mouse**] [ate well] or using it as a verb: [*le bébé*] [***sourit*** *à sa maman*] - [the baby] [**smiles** to his mother], where brackets indicate prosodic phrase boundaries). Although both sentences start with the same words (le-bebe-/suʁi/), they can be disambiguated by the prosodic boundary that either directly precedes the critical word /suʁi/ when it is a verb, or directly follows it when it is a noun. Across two experiments using an intermodal preferential looking procedure, 28-month-olds (Exp. 1 and 2) and 20-month-olds (Exp. 2) listened to the beginnings of these test sentences while watching two images displayed side-by-side on a TV-screen: one associated with the noun interpretation of the ambiguous word (e.g., a mouse) and the other with the verb interpretation (e.g., a baby smiling). The results show that upon hearing the first words of these sentences, toddlers were able to correctly exploit prosodic information to access the syntactic structure of sentences, which in turn helped them to determine the syntactic category of the ambiguous word and to correctly identify its intended meaning: participants switched their eye-gaze toward the correct image based on the prosodic condition in which they heard the ambiguous target word. This provides evidence that during the first steps of language acquisition, toddlers are already able to exploit the prosodic structure of sentences to recover their syntactic structure and predict the syntactic category of upcoming words, an ability which would be extremely useful to discover the meaning of novel words.

Key-words: phrasal prosody; language acquisition, infants speech perception, syntactic ambiguity resolution; parsing; sentence processing; eye movements.

Phrasal prosody constrains syntactic analysis in toddlers

1. Introduction

Learning word meanings can be a very complex task for toddlers during language acquisition. In their daily life, toddlers need to extract word forms from the speech stream and associate them with possible meanings in their environment. But what kind of information can children use when they need to identify the meaning of a novel word? The *syntactic bootstrapping hypothesis* (Gleitman, 1990; Landau & Gleitman, 1985; see also Fisher, Hall, Rakowitz, & Gleitman, 1994; Fisher, 1996) proposes that having access to the syntactic structure of sentences can help children to discover the meaning of novel words. According to this hypothesis, syntax can serve as a "zoom lens" to help learners figure out which part of the world is being talked about, and hence to identify candidate meanings for novel words. In other words, the range of syntactic environments in which a given word occurs can be informative about its meaning (see Gillette, Gleitman, Gleitman, & Lederer, 1999).

In the simplest case to illustrate this idea, it has been shown that around the age of two, children are able to learn that a novel word such as "*larp*" refers to an action, when listening to sentences in which it appears as a verb, as in "*He is larping that*"; but when exposed to sentences like "*This is a larp*" in which "*larp*" appears in a noun position, they learn that "*larp*" refers to an object (e.g., Bernal, Lidz, Millotte, & Christophe, 2007; Waxman, Lidz, Braun, & Lavin, 2009). This suggests that children exploit the syntactic frames in which novel words occur to infer their possible referent. Going further, it has been shown that toddlers can also learn that a novel verb such as "*blicking*" refers to a causal action between two participants when listening to transitive sentences such as "*She is blicking the baby*", but they do not make the same inference when listening to intransitive sentences such as "*She is blicking*" (Yuan & Fisher, 2009; Yuan, Fisher, & Snedeker, 2012). In Ferguson, Graf and Waxman (2014), 19-month-olds exposed to sentences like "*The dax is crying*" were

able to infer that "*dax*" referred to an animate entity (i.e., a novel animal), because it appeared in the subject position of a verb that requires an animate agent; but when exposed to sentences like "*The dax is right here*", they did not show any preference for the animate entity at test. Taken together, these studies show the important role played by syntactic structure to assist language acquisition: at an age when toddlers do not have an extensive vocabulary yet, the syntactic structure of sentences helps them to discover the meaning of novel words. The question that arises is how toddlers manage to access the syntactic structure of sentences before acquiring an extensive vocabulary.

A potential cue that has triggered a great deal of interest is phrasal prosody: the rhythm and melody of speech. Across the world's languages, the prosodic organization of speech is such that every prosodic phrase boundary is always aligned with a syntactic constituent boundary (Nespor & Vogel, 1986; Shattuck-Hufnagel & Turk, 1996), although the reverse is not true, since many syntactic boundaries are not marked prosodically. Crucially, however, prosodic information such as phrase-final lengthening, pitch contour variations and pauses between prosodic units may allow young listeners to find the boundaries between some of the syntactic constituents of a sentence, even in the absence of a very extensive vocabulary (Christophe, Millotte, Bernal, & Lidz, 2008; Morgan & Demuth, 1996; Morgan, 1986). This ability to exploit phrasal prosody to identify syntactic constituent boundaries, in addition to the perception of function words (Hallé, Durand, & de Boysson-Bardies, 2008; Höhle, Weissenborn, Kiefer, Schulz, & Schmitz, 2004; Höhle & Weissenborn, 2003; Shafer, Shucard, Shucard, & Gerken, 1998; Shi, Werker, & Cutler, 2006; Shi & Melançon, 2010), has been proposed to be potentially important for infants to bootstrap their way into syntactic acquisition, because phrasal prosody would allow them to identify some of the syntactic constituents in a sentence, while function words would allow them to determine the syntactic nature of these constituents (Christophe, Millotte, Bernal, & Lidz, 2008; Shi, 2014).

Supporting this hypothesis, several studies have shown that the perception of prosodic boundaries can indeed help adults and preschoolers to constrain their syntactic analysis and resolve syntactic ambiguities (in English: de Carvalho, Lidz, Tieu, Bleam, & Christophe, 2016; Kjelgaard & Speer, 1999; Snedeker & Yuan, 2008 and in French: de Carvalho, Dautriche, & Christophe, 2016; Millotte, René, Wales, & Christophe, 2008; Millotte, Wales, & Christophe, 2007). However, little is known about young children who are still in the process of acquiring the words of their language: can they exploit the prosodic structure of sentences as a cue to access their syntactic structure? Such an ability would be extremely important during the first steps of syntactic acquisition, since accessing the syntactic structure of sentences may allow children to determine the syntactic category of unknown words and therefore constrain their meaning.

A long series of studies shows that infants develop an impressive expertise with prosody from their first days of life. Newborns are able to exploit rhythmic information to discriminate between languages (Mehler et al., 1988; Nazzi, Bertoncini, & Mehler, 1998); from 4.5 months onwards, infants are sensitive to the coherence of prosodic constituents (Gerken, Jusczyk, & Mandel, 1994; Hirsh-Pasek et al., 1987; Jusczyk, Hohne, & Mandel, 1995; Männel & Friederici, 2009; Soderstrom, Seidl, Nelson, & Jusczyk, 2003), they show better recognition and memory for segments that correspond to whole prosodic units than for those which span prosodic boundaries (Mandel, Jusczyk, & Nelson, 1994; Nazzi, Iakimova, Bertoncini, Frédonie, & Alcantara, 2006) and they can rely on prosodic cues to segment the speech stream into words and constrain their lexical access (Gout, Christophe, & Morgan, 2004; Johnson, 2008; Millotte et al., 2010; Shukla, White, & Aslin, 2011). All of these findings, together with the reliable relationship between prosodic and syntactic structures, suggest that toddlers might be able to use phrasal prosody, not only to facilitate memory and lexical access, but also to constrain their syntactic analysis (see Christophe et al., 2008;

Hawthorne & Gerken, 2014; Massicotte-Laforge & Shi, 2015; Morgan & Demuth, 1996; Morgan, 1986; Shi, 2014).

In the experiments that follow, we directly examined whether toddlers, who are still in the process of learning the syntax and the lexicon of their language, exploit phrasal prosody to constrain their syntactic analysis.

2. Experiment 1

We tested toddlers' ability to use phrasal prosody as a cue to access the syntactic structure of sentences and to constrain their interpretation of an ambiguous word. Pairs of French noun-verb homophones were used to create locally ambiguous sentences. For instance, the word-form "/suʁi/" was used as a noun in: [*Le bébé*$_{\text{ADJ}}$ ***souris***$_{\text{NOUN}}$][*a bien mangé*] 'The baby$_{\text{ADJ}}$ **mouse**$_{\text{NOUN}}$ ate well' (hereafter the noun prosody condition), and it was used as a verb in: [*Le bébé*$_{\text{NOUN}}$][***sourit***$_{\text{VERB}}$ *à sa maman*] 'The baby$_{\text{NOUN}}$ **smiles**$_{\text{VERB}}$ to his mom' (hereafter the verb prosody condition) – brackets indicate prosodic boundaries. Although these two sentences start with the same three words (e.g., le-bébé-/suʁi/), they are disambiguated by their prosodic structures, reflecting their different syntactic structures. When the ambiguous word was used as a verb, there was a prosodic boundary just before it, corresponding to the boundary between the subject noun phrase and the verb phrase (i.e., [*Le bébé*] [***sourit***.. - [The baby] [smiles…). However, when the homophone was used as a noun, the prosodic boundary appeared just after it, because in this case all three words belonged to a single prosodic unit, corresponding to the subject noun phrase (e.g., [*Le bébé* ***souris***] … - [The baby mouse] …)[1]. Crucially, all words following the homophone were masked with babble noise, such that prosodic cues were the only disambiguating information.

To examine whether 28-month-olds exploit phrasal prosody to constrain their syntactic analysis, an intermodal preferential looking task with an eye-tracker was designed. Toddlers listened to the beginnings of these ambiguous sentences while watching two images displayed

[1] Note that in French there is no difference in pronunciation between "*souris*" and "*sourit*", the final 's' and 't' are not pronounced, and both words are pronounced as /suʁi/.

side-by-side on a TV screen: one associated with the noun interpretation of the ambiguous target word (e.g., a mouse) and the other one with the verb interpretation (e.g., a baby smiling). Their looking behavior was measured with an eye-tracker. If toddlers are able to take into account the prosodic structure of these sentences when conducting their syntactic computations, we expect them to look more often toward the noun picture when listening to sentences in the *noun prosody condition* than to sentences in the *verb prosody condition.*

2.1. Method

The stimuli, data and analyses of the experiments reported in this paper are accessible to readers on the OSF (Open Science Framework) database through the following link: https://osf.io/744pq/?view_only=c50cd5300feb4832ad58d3566dd041ee

2.1.1. Participants

Forty toddlers, from 27.6 (months.days) to 28.28, with a mean of 27.26 (*SD*=0.5, 19 girls) participated in this experiment. An additional four children participated in the study but were not included in the final analysis because of fussiness during the experiment resulting in more than 50% (4 out of 8) unusable test trials with missing eye-tracking data. All participants were monolingual native French speakers. Parents signed an informed consent form. This research was approved by the local ethics committee.

2.1.2. Materials

Eight pairs of French noun-verb homophones likely to be known to young children (Kern, 2007; Veneziano & Parisse, 2010, 2011) were selected to create eight pairs of experimental sentences. For each pair of homophones, two sentences were created: one using the ambiguous word as a noun (the noun prosody condition, e.g. [*Le bébé*$_{\text{ADJ}}$ ***souris***$_{\text{NOUN}}$][*a bien mangé*] – [The baby$_{\text{ADJ}}$ **mouse**$_{\text{NOUN}}$] [ate well]) and a second one using the ambiguous word as a verb (the verb prosody condition, e.g., [*Le bébé*$_{\text{NOUN}}$] [***sourit***$_{\text{VERB}}$ *à sa maman*] – [The

baby$_{NOUN}$] [**smiles**$_{VERB}$ to his mom]; see Appendix 1 for a complete list of test sentences). Sentences uttered in the noun prosody condition had a prosodic boundary after the ambiguous target word and sentences uttered in the verb prosody condition had a prosodic boundary before the target word, consistent with theoretical descriptions of the relationship between prosodic and syntactic boundaries (e.g., Jun, 2005; Nespor & Vogel, 1986). A female French native speaker (the last author) recorded all the sentences in a child-friendly register. Note that the prosodic boundaries associated with each prosodic condition were found to be naturally produced by naïve adult native speakers, even when they were not aware of the syntactic ambiguity (Millotte et al., 2007). To estimate toddlers' knowledge of the ambiguous words, the parents of the participants in this experiment filled a short questionnaire. Overall, most toddlers understood most of the words used in this study (mean number of words comprehended: 13.8 out of 16; range: 10-16).

In addition to the experimental sentences, six filler sentences were created using target words that were unambiguously either a noun or a verb (e.g., noun: *chat* 'cat' in the sentence: [*Le petit **chat***] [*est très mignon*] The little **cat** is very cute vs. verb: *lave* 'to wash' in the sentence: [*La vieille*] [***lave** sa jupe*] The old lady **washes** her skirt).

To ensure that prosodic cues would be the only information available to participants to determine whether the ambiguous word was a noun or a verb[2], each test and each filler sentence was cut at the offset of the target word, and its end replaced by 1000ms of babble

[2] To control for the possibility that sub-phonemic cues might allow listeners to distinguish between the noun/verb homophones (as suggested by a reviewer), we conducted a control experiment in which adults (n=12) listened to the ambiguous words spliced out from the test sentences used in Experiment 1 and had to judge whether the word was a noun or a verb, in a two-alternative forced-choice task where the alternatives were visually presented (e.g. *la souris* - the mouse vs *elle sourit* - she smiles). Participants were at chance, with 53% noun answers in the noun prosody condition and 50% in the verb prosody condition (β= -0.26; z = -0.43; p = 0.66). The details of this control experiment can be found on the OSF database: https://osf.io/744pq/?view_only=c50cd5300feb4832ad58d3566dd041ee.

noise obtained by superimposing the end of all filler sentences. Thus, there was no lexical disambiguating information following the ambiguous word[3].

There were 16 test sentences, 8 in the verb prosody condition and 8 in the noun prosody condition. Each participant was presented with only one member of each pair. Two counterbalanced lists of stimuli were used, each list containing four sentences in the noun prosody condition and four sentences in the verb prosody condition, plus four filler sentences (two of them having an unambiguous noun as a target and the other two having an unambiguous verb as a target). The order of sentences within each list was randomized, with the constraint that there were no more than two test sentences in a row and no more than two items from the same syntactic category in a row. To create the intermodal preferential looking task, for each sentence beginning (e.g., le-bébé-/suʁi/), two images were created, one depicting the noun interpretation of the ambiguous word (e.g., a mouse) and another one depicting its verbal interpretation (e.g., a baby smiling). For filler sentences, one image corresponded to the target word and the other was unrelated but represented a word from the opposite syntactic category. For instance, if a given filler target was a noun then the other image depicted an action. In total, 28 images (16 for the test sentences and 12 for the filler sentences) were created. An artist (the third author) provided line drawings of approximately equal size and complexity depicting each of these images. The experimental materials, both sentences and images, were the same as those used in de Carvalho, Dautriche and Christophe (2016) with preschoolers.

2.1.3. Acoustic analyses

In order to assess prosodic differences between the two conditions, acoustic measurements (duration and pitch) were conducted on the sentence beginnings (see Fig. 1).

[3] Additionally, to ensure that no co-articulatory cues would differentiate sentences across conditions, in all test sentences, the word following the target word always started with the same segment (e.g. noun prosody condition: *le bébé souris*$_N$ **a** bien mangé and verb prosody condition: *le bébé*$_N$ *sourit*$_V$ **à** sa maman, both words starting with the same vowel /**a**/).

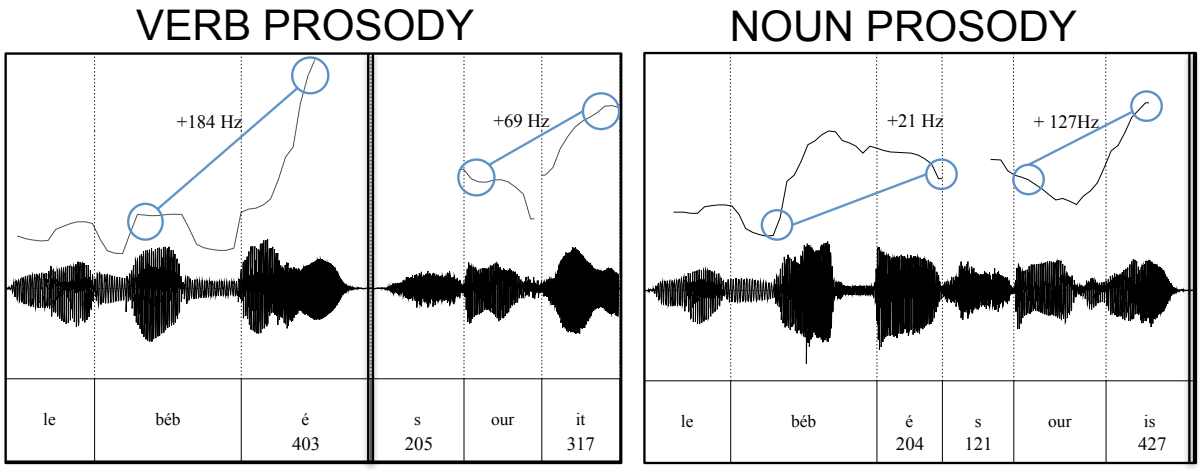


Figure 1: Mean duration of the different segments, and pitch contours in the ambiguous region. Prosodic boundaries are represented with thick black lines. Blue circles delimit the areas where pitch analyses were performed, subtracting the pitch value at the beginning from the pitch value at the end of the words around the prosodic boundaries. Note that while waveforms and pitch curves in the figure correspond to the experimental sentences for the homophone "/suʁi/", the values for duration and pitch correspond to mean values across all stimuli.

The analysis of duration revealed a significant pre-boundary lengthening, as expected from the literature (Cooper & Paccia-Cooper, 1980; Delais-Roussarie, 1995; Jun & Fougeron, 2002; Millotte et al., 2008, 2007; Nespor & Vogel, 1986; Shattuck-Hufnagel & Turk, 1996; Soderstrom, Blossom, Foygel, & Morgan, 2008): the rhyme of the word placed just before the prosodic phrase boundary (marked in Fig1 by thick black lines) in the verb condition (e.g., last vowel -/e/ from *bebe*) was lengthened by 98% compared to this same rhyme in the noun condition (403 vs 204 ms, see Table 1), and the rhyme of the word placed just before the prosodic phrase boundary in the noun condition (e.g., -/i/ from /suʁi/) was lengthened by 35% compared to this same rhyme in the verb condition (427 vs 317 ms). Additionally, we also observed a phrase-initial consonant strengthening (see Fougeron & Keating, 1997): the onset of the target word in the verb condition (205ms, phrase-initial position) was lengthened by 70% compared to the noun condition (121ms, phrase-medial position).

Table 1

Duration analyses – Mean duration in ms (standard error of the mean)			
	Noun Prosody [le bébé suʁi]	**Verb Prosody** [le bébé][suʁi	**Analysis (2-tailed t-tests)**
Rhyme - word preceding Target (e.g., -/e/ from "*bébé*")	204 (22)	403 (50.4)	*t* (7) = -3.85, *p* <.01**
Onset - Target word (e.g., /s/ from "/suʁi/"	121 (9.2)	205 (16.2)	*t* (7) = -5.02, *p* <.01**
Rhyme – Target word (e.g., -/i/ from "/suʁi/")	427 (50.6)	317 (34.9)	*t* (7) = 3.77, *p* <.01**

Pitch analyses – Mean pitch change, in Hz, from the beginning to the end of the target words (standard error of the mean).			
Dependent variable	**Noun Prosody** [le bébé suʁi]	**Verb Prosody** [le bébé][suʁi	**Analysis (2-tailed t-tests)**
Word preceding Target (e.g., last pitch value at the last vowel from "bébé" minus first pitch value from the first vowel of "bébé")	21 (20.4)	184 (38.1)	*t* (7) = -5.29, *p* <.01**
Target word (e.g., last pitch value of "-i" from "souri" minus first pitch value of "/u/" from "/suri/")	127 (23.5)	69 (25.8)	*t* (7) = 4.47, *p* <.01**

The analysis of pitch contours also revealed significant differences between conditions, consistent with the literature describing French as having a tendency for a rising pitch contour towards the end of prosodic units (Di Cristo, 2000; Welby, 2003, 2006). A greater pitch rise was observed on the target word in the noun prosody condition (+127Hz) compared to the verb prosody condition (+69Hz). This difference is due to the fact that in the noun prosody condition the target word was in a phrase-final position, while in the verb prosody condition it was placed at the beginning of a phrase. For the same reasons, the word preceding the target word (e.g., "bébé") had a greater rise in pitch in the verb prosody condition (+184Hz) than in the noun prosody condition (+21Hz). All of these differences were significant (see Table1).

2.1.4. Apparatus and Procedure

Toddlers were tested individually in a sound-attenuated double-walled booth. They were sitting on their parent's lap, facing a 42-in screen positioned 70cm away from them. Toddlers' eye movements during the experiment were recorded by an eye-tracker (Eyelink-1000) placed below the screen (operating in remote mode). Parents wore opaque glasses and were asked not to interact with their children during the experiment. The experimenter remained outside the booth during the test and used a 5-point calibration procedure to calibrate the eye-tracker.

In order to introduce toddlers to the task, the experiment started by a practice block in which they were presented with two filler sentences (one having an unambiguous noun as a target and the other an unambiguous verb). Right after that, toddlers started the test block, composed of eight ambiguous test sentences and four filler sentences.

Each trial started with an inspection period to provide toddlers enough time to inspect each of the images individually on each side of the screen. For instance, one image was presented on the left (or right) side of the screen for three seconds, accompanied by a neutral audio prompt (e.g. 'Hey look!'), then the other image was presented on the opposite side of the screen for another 3s (with another neutral audio prompt). Five hundred milliseconds later, both images were presented side-by-side on the screen for 3s, without any acoustic stimulus. Then these images disappeared and a colorful fixation target appeared in the middle of the screen. Once participants looked at this fixation point for at least 500ms, the two images reappeared on the screen at the same time as the auditory test sentence was played. The time course of each trial is illustrated in Figure 2.

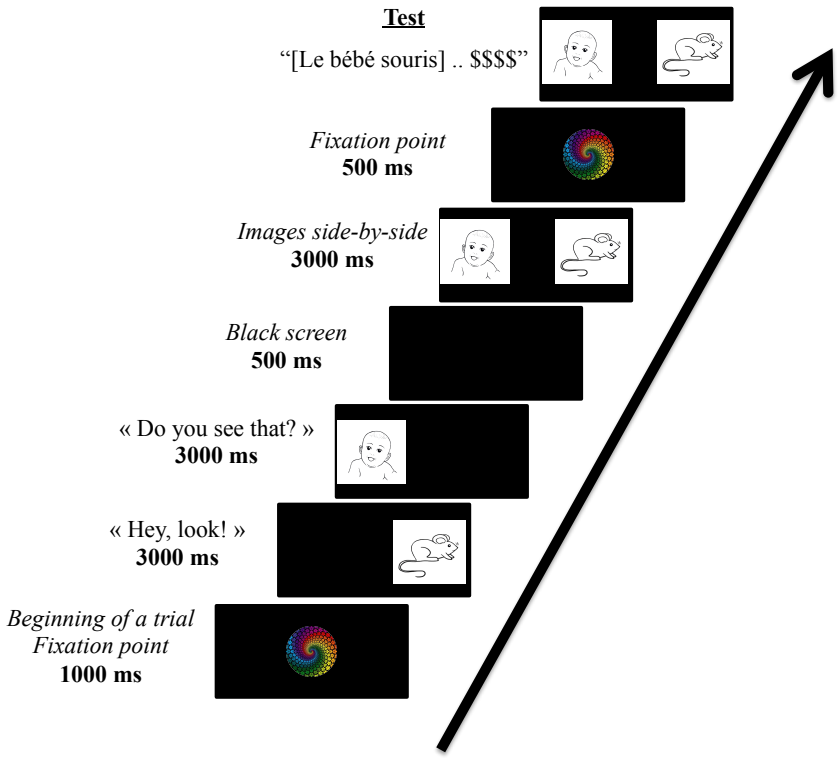


Figure 2: Time-course of a trial

2.1.5. Data processing and analysis

Toddlers' eye-gaze towards the images was recorded by an Eyelink-1000 while they listened to the test sentences, with a time-sample collected every 2ms. Before statistical analysis, the data was down-sampled by a factor of 10, by averaging the data from 10 adjacent samples, so that the final sampling rate was one sample every 20ms. Thirty-nine trials out of 320 were removed from the statistical analysis (17 in the noun condition and 22 in the verb condition), because more than 25% of the data frames were missing between the onset of the ambiguous word and the end of the audio stimuli. The eye-gaze analysis uses the proportion of fixations toward the noun image as a dependent variable, because fixations to noun vs. verb image in this task are complementary (apart from the time spent looking away). To find the time-window(s) which exhibited a significant difference between conditions, a cluster-based permutation analysis was conducted (as in Dautriche, Swingley, & Christophe, 2015; de Carvalho et al., 2016; Hahn, Snedeker, & Rabagliati, in press, Von Holzen & Mani, 2012; see Maris & Oostenveld, 2007, for a formal presentation of the analysis itself). This analysis

allows us to test for the effect of Condition without inflating the rate of Type I error. It proceeds in two phases. First, for each time point, a paired two-tailed *t*-test testing for the effect of Condition (noun prosody vs. verb prosody) is conducted (on the proportion of looks toward the noun picture). Adjacent time points with a *t*-value greater than some predefined threshold (here, $t = 1.5$)[4] are grouped together into a cluster. The size of the cluster is defined as the sum of the *t* values at each time point within the cluster. Second, to obtain the probability of observing a cluster of that size by chance, we conducted 1000 simulations where we randomly shuffled the conditions (noun prosody, verb prosody) for each trial. For each simulation, we calculated the size of the biggest cluster identified with the same procedure that was applied to the real data. A cluster of adjacent time points from the real data shows a significant effect of condition if its size is greater than the size of the largest cluster found in 95% of the simulations (ensuring a p-value of .05). This analysis was conducted on a time-window extending from -700 ms before the onset of the ambiguous word until 2000 ms after the onset of the ambiguous word. Plots of eye-gaze data were performed with the ggplot2 package (Wickham, 2009).

2.2. Results

Figure 3 shows the average proportion of looks toward the noun image in the noun prosody condition (red curve) and in the verb prosody condition (blue curve), time-locked to the beginning of the ambiguous word onset. This reflects toddlers' online interpretation of sentences as the linguistic input unfolds (e.g., Trueswell, 2008; Trueswell & Gleitman, 2007; Trueswell, Sekerina, Hill, & Logrip, 1999).

Visual inspection of the data reveals that toddlers tended to look more toward the verb image at the beginning of the sentences. For instance, at the onset of the target word, vertical black line, both curves are at 40% looks toward the noun picture, perhaps revealing a simple

[4] The same threshold was used in de Carvalho, Dautriche and Christophe (2016). Note that the value of the threshold does not affect the rate of false alarms of the test, since the significance of the cluster is estimated through the permutation procedure.

preference for looking at human beings in the verb picture. Crucially, however, when listening to the beginning of a noun sentence, toddlers increased their looks toward the noun picture, from around the offset of the ambiguous word, thus switching their eye-gaze toward the correct image after hearing the relevant prosodic information. In contrast, when they were listening to the beginning of a verb sentence, toddlers increased their looks toward the verb picture.

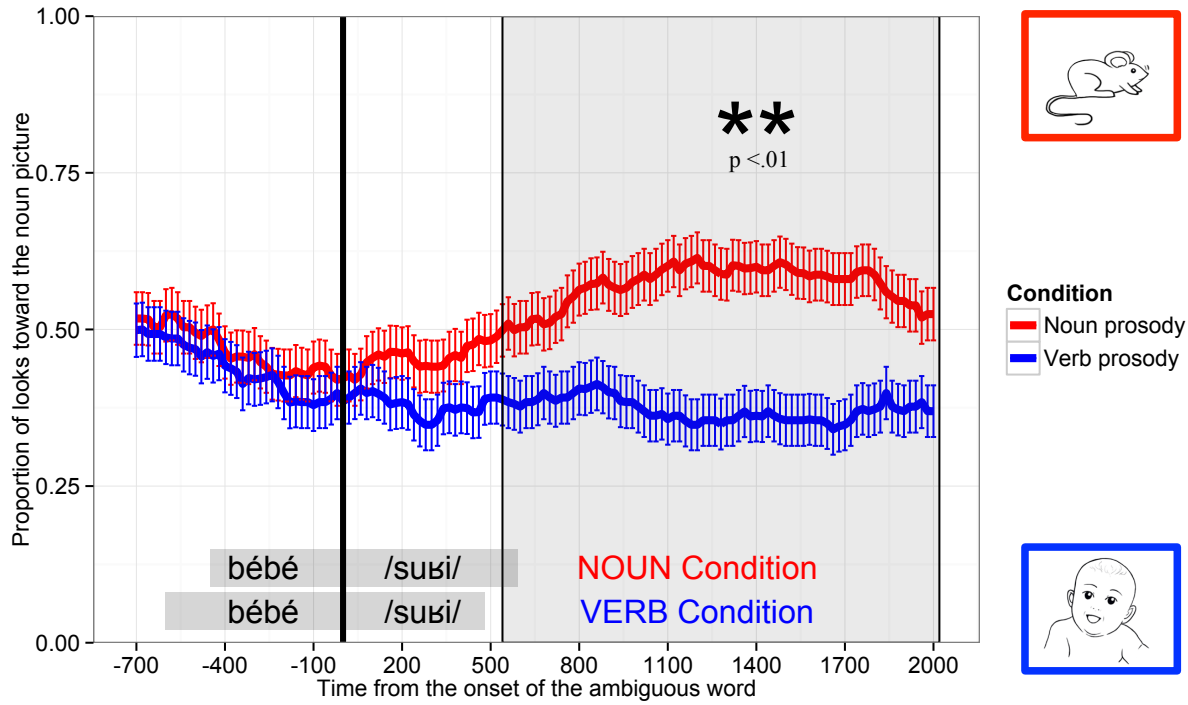


Figure 3: Proportion of looks toward the noun image, time-locked to the onset of the ambiguous word (thick vertical line), for the noun prosody condition (red curve) and the verb prosody condition (blue curve). Error bars represent the standard error of the mean. A nonparametric cluster-based permutation test revealed a significant difference between the noun prosody and the verb prosody conditions, starting around the offset of the ambiguous target word (grey time-window; from 540 ms after the beginning of the critical word, '**' $p < .01$).

The cluster-based analysis found a significant time-window where the proportion of looks toward the noun picture was significantly different in the noun condition compared to the verb condition, from 540 ms after the beginning of the critical word until the end of the trial at 2000 ms ($p < .01$). This shows that 28-month-olds were able to exploit prosodic information to recover the syntactic structures of sentences and use this syntactic structure to

compute the syntactic category of the homophones and therefore constrain their interpretation of the ambiguous target word.

2.3. Discussion

The experiment reported here tested whether 28-month-olds exploit phrasal prosody online to access the syntactic structure of sentences and constrain their syntactic analysis. In an intermodal preferential looking task, toddlers were able to exploit the position of an ambiguous word within the prosodic structure of sentences to compute its syntactic category. They interpreted the ambiguous word as a noun (and looked more toward the noun picture) when it was embedded in a sentence from the noun prosody condition, and as a verb when it was embedded in a sentence from the verb prosody condition.

The time course of toddlers' eye-gaze suggests that they integrate prosodic information online during sentence parsing. Although children were initially biased to look toward the verb image, soon after they heard the critical word in the noun prosody condition, they switched their eye-gaze toward the noun image, while they increased their looks toward the verb image when hearing the critical word in the verb prosody condition. This behavior was reflected by a strong effect of prosodic condition, starting 540ms after the target word onset and remaining stable until the end of the trials. Considering that it takes toddlers 300 to 500 ms to orient their eye-gaze toward pictures of familiar objects when listening to simple sentences such as "Where is the ball?" (e.g., Ferguson, Graf, & Waxman, 2014; Fernald, Zangl, Portillo, & Marchman, 2008; Swingley & Aslin, 2000), it is impressive that they took only slightly longer in the present experiment (around 540ms), even though the target word was a homophone, and corresponded to a verb half the time (since action pictures are more complex than pictures of concrete objects).

This effect mirrors previous results obtained with adults and preschoolers in French (de Carvalho et al., 2016), although in the present study toddlers seem to be slower than 4-

year-old children, by about 300ms. This difference could be due to the fact that preschoolers (and adults) exploit prosodic information more efficiently than toddlers, or simply to the fact that toddlers have less attentional skills than their older counterparts (leading to noisier behavior). Although we cannot disentangle between these interpretations, the main result is that, just like adults and preschoolers, toddlers who are still in the process of learning the syntax of their language, can use phrasal prosody online to access the syntactic structure of sentences and constrain their syntactic analysis. Given that prosodic phrase boundaries are perceived and exploited by infants from six months onwards (Gerken et al., 1994; Gout et al., 2004; Shukla et al., 2011; Soderstrom, 2003), it is possible that even younger toddlers might be able to use phrasal prosody as cue to recover the syntactic structure of sentences.

In order to investigate this question, Experiment 2 aims to directly test whether 20-month-old infants are able to use prosodic structure to access the syntactic structure of sentences and constrain their syntactic analysis. A pre-test of Experiment 1 with a small group of 18-month-olds (n=20) revealed that this task was not appropriate for testing this age group, for the following reasons: a) The task seemed to be too long for them, they became fussy before the end of the experiment, and tended not to finish the task; b) some toddlers were afraid of the babble noise masking the end of sentences and started crying during the experiment; c) the duration of each trial seemed to be too short for 18-month-olds, not leaving them enough time to choose the correct image. In the current experiment, trials ended one second after the offset of the target words (i.e., the duration of the babble noise mask), while younger infants may have needed more time to process the sentences and to switch their eye-gaze toward the correct image. Supporting this idea, previous eye-tracking studies with 19- and 21-month-olds have shown that it can take them between 1 and 4 seconds after target word offset to look toward a noun or a verb referent (Arunachalam, Escovar, Hansen, & Waxman, 2013; Ferguson et al., 2014). Thus, in Experiment 2 we adapted the experimental procedure to test younger toddler's ability to use phrasal prosody to constrain syntactic

analysis.

3. Experiment 2

To adapt the experimental design to 20-month-olds, several changes were implemented. The experiment was shortened by half by using only four of the previous eight pairs of noun-verb homophones. To avoid using the babble noise mask, only homophones for which the verb could be used in an intransitive structure were used (either intransitive verbs, or verbs that accepted omission of their complement). Finally, to give infants more time to process the sentences, each ambiguous sentence was repeated twice.

These changes led us to create minimal pairs of globally ambiguous sentences, such as *'Regarde le bébé* /suʁi/*'*, which can be produced either as [*Regarde le*$_{\text{DET}}$ *bébé*$_{\text{ADJ}}$ /suʁi/$_{\text{NOUN}}$!] - Look at the$_{\text{DET}}$ baby$_{\text{ADJ}}$ mouse $_{\text{NOUN}}$!, where /suʁi/ is a noun, or as [*Regarde*], [*le*$_{\text{DET}}$ *bébé*$_{\text{NOUN}}$] [/suʁi/$_{\text{VERB}}$!] - Look! The$_{\text{DET}}$ baby$_{\text{NOUN}}$ smiles $_{\text{VERB}}$!, where /suʁi/ is a verb (brackets indicate prosodic boundaries). As in Experiment 1, both sentences are composed of exactly the same words, and can be disambiguated by their prosodic structures, which reflect the different syntactic structures. If 20-month-olds exploit phrasal prosody to constrain their syntactic analysis, we expect them to look more toward the noun picture when listening to sentences in the noun prosody condition, than when listening to sentences in the verb prosody condition. In order to directly compare the performance of the 20-month-olds and the 28-month-olds, we tested two groups of toddlers in this experiment: the younger group of 20-month-olds, and a new group of 28-month-olds, in which we expected to replicate the same effect found in Experiment 1.

3.1. Method

3.1.1. Participants

Sixty-four toddlers participated in this experiment. They were all monolingual native French speakers and were divided into two age groups (with 32 toddlers in each age group): the 20-month-old group, ranging in age from 19.0 (months.days) to 21.3, with a mean age of

19.19 (*SD*=0.6; 14 girls) and the 28-month-old group, ranging in age from 26.19 to 28.27, with a mean of 27.20 (SD=0.6; 18 girls). Within each age group, half of the participants heard the test sentences in the noun prosody condition and half heard them in the verb prosody condition. An additional twenty-six children completed the experiment (eleven 28-m.o and fifteen 20-m.o) but they were not included in the final sample because of fussiness during the experiment resulting in more than 50% of trials with missing eye-tracking data (n = 19), because they cried (n = 4), or because of technical problems (n = 3). Parents signed an informed consent form. This research was approved by the local ethics committee.

3.1.2. Material

Four pairs of French noun-verb homophones were used to create eight experimental sentences, four using the target word as a noun (e.g. [*Regarde le*$_{\text{DET}}$ *bébé*$_{\text{ADJ}}$ /suʁi/$_{\text{NOUN}}$!] [*Tu vois le*$_{\text{DET}}$ *bébé*$_{\text{ADJ}}$ /suʁi/$_{\text{NOUN}}$?] - Look at the$_{\text{DET}}$ baby$_{\text{ADJ}}$ mouse $_{\text{NOUN}}$! Do you see the$_{\text{DET}}$ baby$_{\text{ADJ}}$ mouse $_{\text{NOUN}}$?) and four using the ambiguous word as a verb (e.g., [*Regarde*], [*le*$_{\text{DET}}$ *bébé*$_{\text{NOUN}}$] [/suʁi/$_{\text{VERB}}$!] [*Tu vois*?] [*le*$_{\text{DET}}$ *bébé*$_{\text{NOUN}}$] [/suʁi/$_{\text{VERB}}$!] - Look! The$_{\text{DET}}$ baby$_{\text{NOUN}}$ smiles $_{\text{VERB}}$! Do you see? The$_{\text{DET}}$ baby$_{\text{NOUN}}$ smiles$_{\text{VERB}}$!; see the Appendix 2 for a complete list of test sentences). In each trial, the target word was repeated twice, to give infants more time to process the sentences. As in Experiment 1, sentences uttered in the verb prosody condition had a prosodic boundary before the target word (i.e., corresponding to the boundary between the noun and the verb phrases), while in sentences uttered in the noun prosody condition all the words were grouped together into one single prosodic unit. The same speaker as in Experiment 1 recorded all the sentences using a child-directed register. An example of each kind of sentence is depicted in Figure 4. As in Experiment 1, parents' reports suggest that most of the participants understood the majority of the words (mean number of words comprehended: 7.75 out of 8; range: 7-8; for the 28-month-olds; and 6.34 out of 8; range: 3-8; for the 20-month-olds).

In addition to experimental sentences, two filler sentences contained a non-ambiguous target word at the end of the sentence (one noun sentence "[*Regarde le petit **chat**!*] [*Tu vois le petit **chat**?*] – Look at the little **cat**! Do you see the little **cat**?) and one verb sentence "[*Regarde*]! [*la petite*] [***dort**!*] [*Tu vois?*] [*la petite*][***dort**!*] – Look! The little girl is **sleeping**! Do you see? The little girl is **sleeping**). These two filler sentences were used at the beginning of the experiment to familiarize toddlers with the task.

To make the experiment as simple as possible for young toddlers, each participant was presented either with sentences in the noun prosody condition, or with sentences in the verb prosody condition, in a between-participants design. Half of the participants listened to four sentences in the noun prosody condition and the other half listened to four sentences in the verb prosody condition, for a total of 6 trials (2 filler trials followed by 4 test trials). Test sentences were presented in random order.

For each homophone used in the experiment, two images were created, one depicting the noun interpretation of the homophone and the other depicting the verb interpretation. For the two filler items used, one image corresponded to the target word and the other was unrelated but represented a word from the opposite syntactic category. In total, 12 images were created: 8 for the test sentences and 4 for the filler sentences. These images were drawn by the same person as in Experiment 1, and were colored in order to make the experiment more interesting for young children.

3.1.3. Acoustic analyses

In order to assess prosodic differences between the two prosodic conditions, acoustic measurements (duration and pitch) were conducted on the test sentences. The analysis of duration revealed a significant pre-boundary lengthening, as expected from the literature: the rhyme of the word preceding the target word (e.g., last vowel -/e/ from *bébé*) in the verb condition (where it was placed just before the prosodic phrase boundary) was lengthened by

211% compared this same segment in the noun condition (where it was placed in the middle of a prosodic unit; 395 vs 127 ms, see Table 2). A silent pause of 232 ms preceding the target word (i.e., between "bébé" and /suʁi/) was observed in the verb condition, while there was no pause between these words in the noun condition. Additionally, a phrase-initial strengthening was observed: the onset of the target word in the verb condition (205ms, phrase-initial position) was lengthened by 88% compared to the noun condition (109ms, phrase-medial position). The rhyme of the target words (e.g., -/i/ from /suʁi/) were utterance-final in both conditions (contrary to Experiment 1); it was lengthened by 49% in the verb condition relative to the noun condition (480 vs 383 ms), possibly because the verb was alone in its prosodic unit.

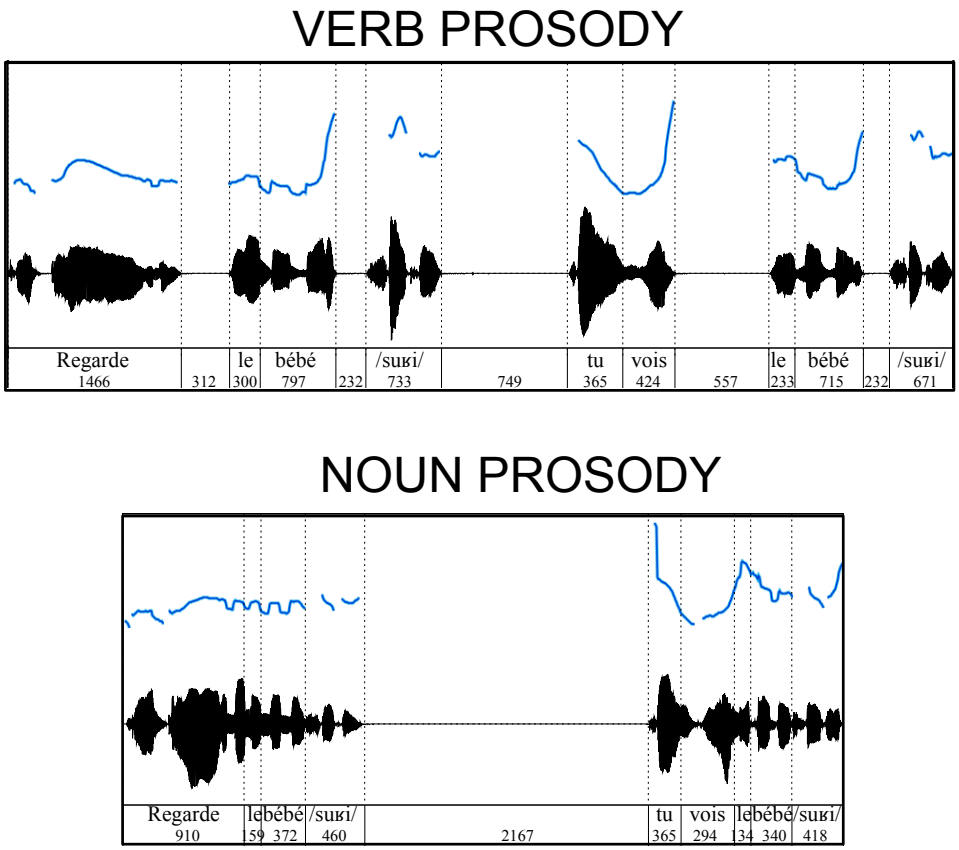


Figure 4: Representation of the ambiguous sentences, with, from bottom to top, the mean duration (in ms) of the different segments, the transcription, the waveform, and the pitch contour. Note that while the waveforms and the pitch contours correspond to the experimental sentences of the item /suʁi/, the duration of the segments correspond to the mean values observed across all stimuli.

Table 2

Duration analyses – Mean duration in ms (standard error)			
	Noun Prosody [le bébé suʁi]	**Verb Prosody** [le bébé][suʁi]	**Analysis (2-tailed t-tests)**
Rhyme - word preceding Target (e.g., e from "bébé")	127 (14.5)	395 (69.2)	*t* (7) = -4.59, *p* <.01**
Pause - before Target (e.g., between "bébé" and "/suʁi/")	0 (0)	232 (52.2)	*t* (7) = -4.44, *p* <.01**
Onset - Target word (e.g., s from "/suʁi/")	109 (11.7)	205 (16.1)	*t* (7) = -6.44, *p* <.01**
Rhyme – Target word (e.g., i from "/suʁi/")	323 (61.8)	480 (116)	*t* (7) = -2.76, *p* = .03*

Pitch analyses – Mean pitch change, in Hz, from the beginning to the end of the target words (standard error of the mean).			
Dependent variable	**Noun Prosody** [le bébé suʁi]	**Verb Prosody** [le bébé][suʁi]	**Analysis (2-tailed t-tests)**
Word preceding Target (e.g., last pitch value at the last vowel from "bébé" minus first pitch value from the first vowel of "bébé")	-53 (16.4)	185 (28.1)	*t* (7) = -14.32, *p* <.01**
Target word (e.g., last pitch value of "-i" from "souri" minus first pitch value of "/u/" from "/suʁi/")	118 (46.9)	29 (30.7)	*t* (7) = 1.94, *p* = .09

The analysis of pitch contours in both prosodic conditions revealed a significant difference between conditions (see Table 2), consistent with the literature describing French as having a tendency for a rising pitch contour towards the end of prosodic units. The word preceding the target word (e.g., bébé) exhibited a greater rising pitch pattern in the verb prosody condition (+185Hz; because of its position at the end of a prosodic unit), than in the noun prosody condition (-53Hz; when it was placed in the middle of a prosodic unit). Given that in both conditions, the target word was placed in the end of a prosodic unit, no particular hypothesis was made regarding their differences in pitch. The target word in the noun prosody condition (e.g., /suʁi/) seemed to exhibit a greater rising pitch pattern in the noun prosody condition (+118Hz) than in the verb prosody condition (+29Hz), but this difference was not

significant.

3.1.3. Apparatus and procedure

The procedure was similar to that of Experiment 1 (although it took place in a different sound-attenuated booth, from IAC Acoustics). Toddlers sat on their parent's lap about 70cm away from a 27-in television screen and as before, their movements were recorded by an eye-tracker (Eyelink-1000) placed below the screen. The caregivers wore headphones and listened to masking music during the experiment.

As in Experiment 1, the experiment began by presenting toddlers with two filler trials (one asking them to look toward a familiar noun (i.e., *chat* - 'cat') and another one asking them to look toward a familiar action (i.e., *dormir* – 'to sleep'). The test block was composed of four ambiguous test sentences (repeated twice for each item). No filler sentences were used into the test block.

As in Experiment 1, each trial started with an inspection period, to provide infants enough time to inspect each of the images individually, on each side of the TV-screen. However, because younger children may benefit from having more time to inspect the images, the inspection period for each image was increased from 3s in Experiment 1 to 5s in the current experiment. Thus, each image was first presented alone for 5 seconds on the left or the right side of the TV-screen and a neutral audio prompt was played at the same time (e.g. 'Hey look! Do you see that?'). Both images were then presented together on the screen, without any acoustic stimulus, during five seconds. Then the images disappeared and a colorful fixation point appeared in the middle of the screen. Once participants looked at the fixation point for 500 ms, the trial started: the two images were presented side-by-side on the screen at the same time that infants listened to the audio sentences and their eye-gaze was recorded, for a total duration of 9s.

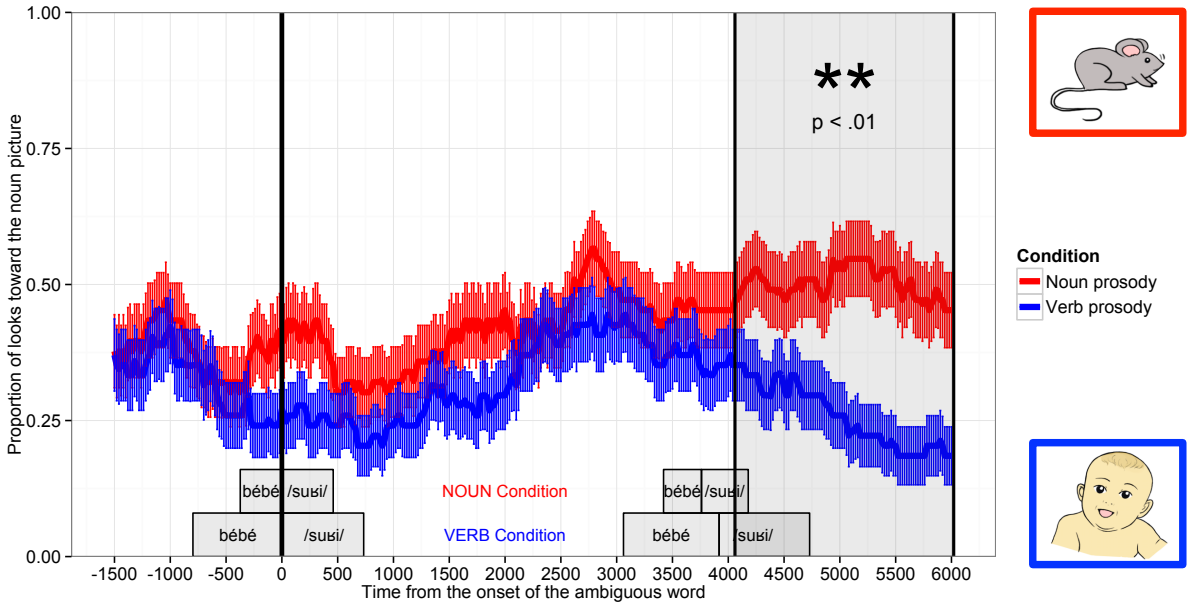
3.1.4. Data processing and analysis

Data processing and analysis followed the same criteria as in Experiment 1. This analysis was conducted on a time-window extending from -1500 ms before the onset of the ambiguous word until 6000 ms after the onset of the ambiguous word (i.e., the end of the trial). Thirty-four trials out of 256 were removed from the statistical analysis because more than 25% of the data frames between the first onset of the ambiguous word and the end of the trial were missing (21 in the noun condition and 13 in the verb condition).

3.2. Results

Figure 4 shows the proportion of looks toward the noun image for toddlers in the noun prosody condition (red curve) and in the verb prosody condition (blue curve), time-locked to the beginning of the first onset of the ambiguous word, for the 20-month-old group (A) and for the 28-month-old group (B).

A)



B)

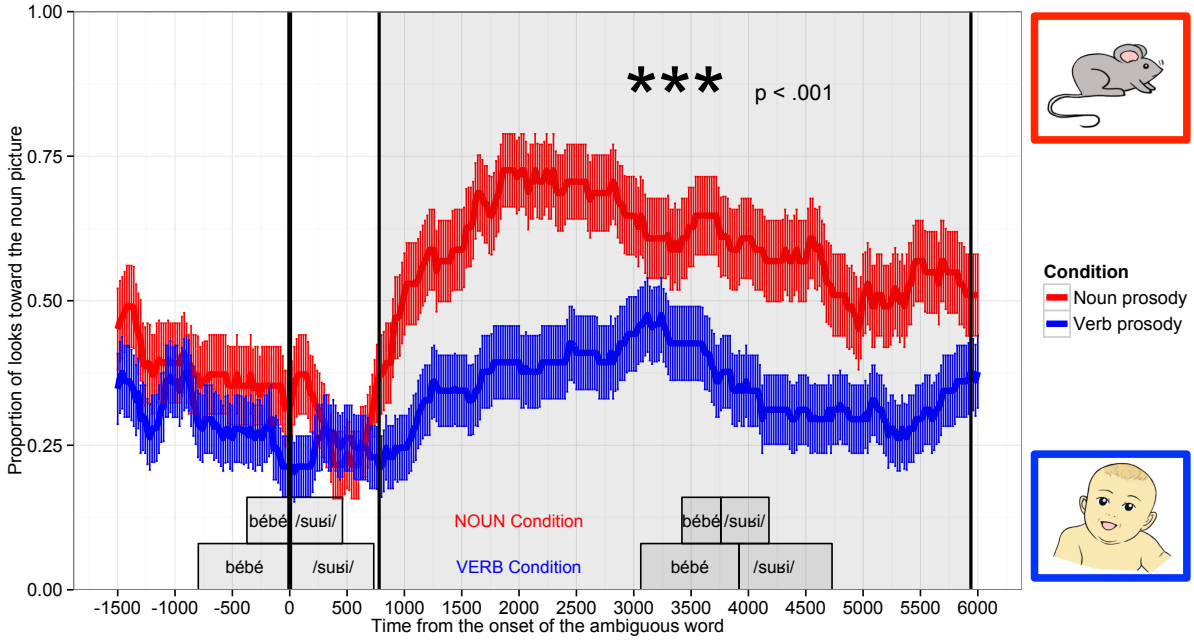


Figure 5: Proportion of looks toward the noun image, time-locked to the onset of the ambiguous word (vertical black line) for 20-month-olds (A), and 28-month-olds (B), for children in the noun prosody condition (red curve) and in the verb prosody condition (blue curve). Error bars represent the standard error of the mean. As in Experiment 1, participants initially looked more toward the verb image, but both age groups switched to the noun image in the noun prosody condition. The cluster-based permutation test revealed significant differences between the noun prosody and the verb prosody conditions (dark grey window) starting slightly after the offset of the first ambiguous target word for the 28-month-olds (about 780ms after onset of the critical word); and during the second repetition of the ambiguous word for the 20-month-olds.

Just as in Experiment 1, visual inspection of the data shows that both groups of toddlers tended to look more toward the verb image at the beginning of the trials. However, toddlers in the noun prosody condition increased their looks toward the noun image, starting slightly after the offset of the first critical word for 28-month-olds, and around the second repetition of the critical word for 20-month-olds. This suggests that while the 28-month-olds were faster than the 20-month-olds in this task, both groups were able to exploit prosodic information to guide their interpretation of the ambiguous target word.

The cluster-based analysis found a significant time-window where the proportion of looks toward the noun picture was significantly different from children in the noun prosody condition compared to children in the verb prosody condition, for both age groups: 28-month-olds (from 780 ms after the onset of the first repetition of the critical word; $p < .001$), and 20-

month-olds (from 4060 ms after the onset of the first critical word or about 300ms after the onset of the second critical word; $p < .01$). These results show that, despite their speed difference, both 20- and 28-month-olds looked more towards the noun picture in the noun prosody condition than in the verb prosody condition.

3.3. Discussion

The results obtained here provide direct evidence that from 20 months on, children exploit prosodic information to access the syntactic structure of sentences, and use this syntactic structure to identify the syntactic category of an ambiguous word (noun/verb homophone). In an intermodal preferential looking task, when listening to minimal pairs of sentences such as *Regarde le bébé* /suʁi/, which can be produced either as [*Regarde le bébé* /suʁi/!] – 'Look at the baby mouse!', where '/suʁi/' is a noun, or as [*Regarde*], [*le bébé*] [/suʁi/!] – 'Look, the baby smiles!', where '/suʁi/' is a verb, 20- and 28-month-olds correctly interpreted the ambiguous word as either a noun or a verb, depending on the prosodic structure of the sentence they were listening to.

Although both age groups switched their eye-gaze toward the correct image, 20-month-olds appeared to be much slower than 28-month-olds. For 28-month-olds, the two conditions diverged right after the first repetition of the ambiguous word, while for 20-month-olds this happened during the second repetition of the target word. This difference in processing speed across age groups may be due to differences in attentional skills between the two ages, and/or to the fact that the younger children knew the homophones less well. In any case, these results show that 20-month-olds can use phrasal prosody to access the syntactic structure of sentences and that they use this information to recover the intended meaning of a homophone.

4. General Discussion

The results reported in this paper show that from 20 months on, toddlers are able to exploit phrasal prosody to access the syntactic structure of sentences, which in turn allows them to identify the syntactic category of an ambiguous word and access its meaning. In a preferential looking task, both 28-month-olds (Experiment 1 and 2) and 20-month-olds (Experiment 2) were able to correctly assign a grammatical category to an ambiguous word (noun vs. verb) depending only on its position within the prosodic structure of sentences. When presented with ambiguous sentences that were phonemically identical but syntactically and prosodically distinct, toddlers were able to exploit the prosodic structure of sentences to infer their syntactic structures, and use this information to decide whether an ambiguous target word was a noun or a verb. They interpreted the ambiguous target word as a noun when it was embedded in a noun sentence and as a verb when it was embedded in a verb sentence, even though the only cue to syntactic structure came from phrasal prosody. This study is the first to report that children under two years old exploit phrasal prosody to recover the syntactic structure of sentences, and use this syntactic structure to compute the syntactic category of an ambiguous word and to access its meaning.

To succeed in our experiments, toddlers may have used phrasal prosody and function words together to constrain their syntactic analysis. This hypothesis is based on the fact that while the perception of prosodic boundaries in our experiments allowed toddlers 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 homophone as a noun or a verb, toddlers may have used the additional information carried by function words[5], together with

[5] Function words have already been shown to be used by 18-month-olds to categorize neighbouring content words (e.g., Cauvet et al., 2014; He & Lidz, 2014; Höhle et al., 2004; Shi & Melançon, 2010).

the prosodic structure of sentences. For instance, in Experiment 1, when participants heard a sentence such as [*le bébé*] [/*suʁi*/ …, the prosodic boundary before the target word signaled the presence of two prosodic units. Given that the first unit (e.g., [*le bébé*]) started with an article (e.g., *le* – the), this unit could be identified as a noun phrase (e.g., [Le_{DET} $bébé_{NOUN}$]$_{NP}$ - [The_{DET} $baby_{NOUN}$]$_{NP}$). Having identified the first unit as a full noun phrase, toddlers might expect it to be followed by a verb phrase, which allows them to rapidly identify the ambiguous word (e.g., /suʁi/) as a verb. In the noun prosody condition in contrast, given that all three words appeared together into one single prosodic unit starting with an article (e.g., [*le bébé* /suʁi/]), this information led toddlers to interpret the entire constituent as a noun phrase, which entailed that /suʁi/ had to be interpreted as a noun. Similarly, in Experiment 2, the presence of a prosodic boundary just before the ambiguous word triggered a verb interpretation, while the ambiguous word was identified as a noun when it belonged to the same prosodic unit as the first three words ([*Regarde le bébé souris*], 'look at the baby mouse'). It is important to note that the use of prosodic information to constrain syntactic analysis is not limited to the kind of syntactic ambiguity resolution featured in our experiments. The relationship between prosodic and syntactic structures is present in all sentences, whether or not they contain ambiguous words. For instance, in a sentence such as [The little cat] [jumps really high], listeners can perceive the prosodic boundary between the subject noun phrase and the verb phrase, as in many sentences that children hear in their everyday lives. In other words, although sentences containing homophones are useful to test listeners' abilities to rely on phrasal prosody to recover syntactic structure, listeners can learn the relationship between prosodic and syntactic structures from unambiguous everyday sentences.

Overall, the ability to use phrasal prosody and function words together helps infants to generate a first parse of the syntactic structure of sentences, and allows them to calculate the syntactic category of an ambiguous word. Note that toddlers seem not to be bothered by the

noun-verb homophony, in these cases, because the critical words occur in disambiguating contexts (contrary to what has been proposed in the literature, e.g. Conwell & Morgan, 2012). We suspect that cross-category homophones such as these will most often appear in disambiguating contexts, and therefore not hinder children's language acquisition (see Dautriche, Fibla, & Christophe, 2015; Dautriche, 2015; Dautriche et al., 2015, for a fuller discussion of this aspect).

The ability to assign a syntactic category to a word according to its context may be extremely important during language acquisition, especially when children do not yet know the meanings of many words. Indeed, children may exploit the fact that an unknown word occurs in a noun context to infer that it probably refers to an object, while words occurring in verb contexts probably refer to actions (e.g., Gillette et al., 1999; Gleitman, 1990). For instance, He and Lidz (2014) showed that 18-month-olds (but not 14-month-olds) were able to infer that a novel word such as 'doke' referred to an object when listening to sentences such as "Look, it's a doke!", and that a novel word such as 'pratch' referred to an action when listening to sentences such as "Look! It's pratching!". However, not all content words are immediately preceded or followed by a disambiguating function word or morpheme as in "a doke" or "is doking" (e.g., in: "*The giant bears...*", *bears* can be either a noun or a verb). In such cases, a more sophisticated analysis in terms of syntactic constituents, signaled by prosodic boundaries, might be extremely informative for infants. For example, in a sentence like "[Do you see the baby blicks]?", infants might be able to infer that "blick" is a noun, referring to an object; but in a sentence such as: "[Do you see]? [The baby] [blicks]!" they might be able to infer that "blick" is a verb, referring to an action. Note that this hypothesis is rather plausible, since to correctly interpret the novel word "blick" as a noun or a verb in this situation, infants would need to exploit exactly the same kind of information they were shown to use in the present experiments.

Other recent findings support the importance of phrasal prosody for syntactic computations in toddlers, showing that prosody facilitates learning of syntactic constituency in 19-month-olds (Hawthorne & Gerken, 2014; Hawthorne, Rudat, & Gerken, 2016) and that 20-month-old toddlers use phrasal prosody to identify syntactic constituents (Massicotte-Laforge & Shi, 2015). For instance, 20-month-olds familiarized with jabberwocky sentences such as [*Ton*$_{\text{Det}}$*felli*$_{\text{Adj}}$ *crale*$_{\text{N}}$]$_{\text{NP}}$ [*vur*$_{\text{V}}$ *la*$_{\text{Det}}$ *gosine*$_{\text{N}}$]$_{\text{VP}}$, where the novel word '*crale*' should be considered as a noun, were surprised (listening longer to test trials) when listening to short phrases presenting this novel word as a verb (e.g., "*Tu*$_{\text{Pron}}$ *crales*$_{\text{V}}$" – 'You crale'), but not when the novel word appeared in the expected syntactic context, as a noun "*Le*$_{\text{Det}}$ *crale*$_{\text{N}}$" (Massicotte-Laforge & Shi, 2015). Taken together, these results show that around 20 months, infants are sensitive to the information provided by phrasal prosody and function words when parsing sentences. Our current findings extend these results and show that infants can exploit prosodic structure to identify possible syntactic constituents; this constituent structure helps them to constrain their syntactic analysis and to access the intended meaning of an ambiguous word.

This suggest that at an age where their knowledge of content words is limited, but phrasal prosody and function words are available, infants could rely on phrasal prosody and function words to retrieve a partial syntactic representation of spoken sentences and attribute a noun or a verb meaning to words, depending on their position in the syntactic structure of sentences: a mechanism that might be extremely important during the first stages of language acquisition. Recent computational work supports this idea and shows an excellent performance of models relying on a combination of factors including phrasal prosody, function words and a minimal semantic knowledge, to access the syntactic category of unknown words (Brusini, Amsili, Chemla, & Christophe, 2011; Christodoulopoulos, Roth, & Fisher, 2016; Fisher, 2015; Gutman, Dautriche, Crabbé, & Christophe, 2015).

To sum up, we provided evidence that from 20 months old, toddlers readily exploit the

prosodic structure of an utterance to constrain its syntactic analysis, and access the meaning of an ambiguous target word. We showed that toddlers use phrasal prosody to segment the continuous speech stream into prosodic units, use them to infer the presence of syntactic constituent boundaries, and exploit function words and syntactic boundaries to assign a syntactic category to ambiguous words and recover their meanings. Given that at this age, toddlers are still in the process of learning their lexicon, this ability to assign a syntactic category to words depending on their context may help infants to constrain the acquisition of word meanings. These findings suggest that phrasal prosody plays an important role in language acquisition, since it provides access to a first-pass syntactic structure of sentences which may help infants to bootstrap language acquisition.

Acknowledgements

This research was supported by a PhD fellowship from the École normale supérieure to Alex de Carvalho, and by a postdoctoral fellowship from the Fyssen Foundation to Isabelle Dautriche. It was also supported by grants from the Région IIe-de-France, Fondation de France, LabEx IEC (ANR-10-LABX-0087), IdEx PSL (ANR-10-IDEX-0001-02), as well as the ANR ‘Apprentissages’ (ANR-13-APPR-0012). The authors thank Anne-Caroline Fiévet, Luce Legros and Mélanie Rouelle for their assistance with recruitment of participants and data collection as well as all the parents and babies who participated in this study.

Appendix 1: Experimental Sentences of Experiment 1.

Note that in French several adjectives can be used as nouns. For example, one can say "*le grand*" (literally: *the tall*), meaning *the tall one*, where the pronoun (one) is omitted. The same applies to other adjectives like "*le petit, la petite*" meaning "*the little one*".

Test sentences			
Pair of ambiguous word	**Syntactic category:**	**Target:**	**Full sentence before acoustical mask:**
fermer x la ferme *to close x the farm*	Verb	Ferme	La petite **ferme** le coffre à jouets *The small one **closes** the toy box*
	Noun		La petite **ferme** lui plait beaucoup *The small **farm** pleases him a lot*
lire x le lit *to read x the bed*	Verb	Lit	Le grand **lit** souvent des histoires à son petit frère *The big one often **reads** stories to his younger brother*
	Noun		Le grand **lit** sera pour les parents *The big **bed** will be for the parents*
marcher x la marche *to walk x the stairs*	Verb	Marche	La grande **marche** lentement toute la journée *The big one **walks** slowly all day long*
	Noun		La grande **marche** la fait tomber *The big **stair** makes her fall*
moucher x la mouche *to nose x the fly*	Verb	Mouche	La maman **mouche** le bébé malade *The mother **blows the nose of** the sick baby*
	Noun		La maman **mouche** laisse son bébé tout seul *The mother **fly** leaves her baby alone*
porter x la porte *to carry x the door*	Verb	Porte	La vieille **porte** sa montre à réparer *The old lady **carries** her watch to be repaired*
	Noun		La vieille **porte** sera réparée demain *The old **door** will be repaired tomorrow*
montrer x la montre *to show x the watch*	Verb	Montre	La grande **montre** ses jouets à son frère *The big one **shows** her toys to her brother*
	Noun		La grande **montre** sera réparée demain *The big **watch** will be repaired tomorrow*
sourire x la souris *to smile x the mice*	Verb	[suri]	Le bébé **sourit** à sa maman *The baby **smiles** to his mom*
	Noun		Le bébé **souris** a bien mangé *The baby **mouse** ate well*
pêcher x les pêches *to fish x the peaches*	Verb	[pɛʃ]	Les grosses **pêchent** mon poisson préféré pour le dîner *The fat ones **fish** my favorite fish for dinner*
	Noun		Les grosses **pêches** me font très envie *The big **peaches** tempt me a lot*

Appendix 2: Experimental Sentences of Experiment 2

Test sentences			
Pair of ambiguous word	**Syntactic category:**	**Target:**	**Full sentence before acoustical mask:**
lire x le lit	Verb	Lit	Oh Regarde! Le petit **lit**! Tu vois? Le petit **lit**! *Oh look! The little one **reads**! Do you see? The little one **reads**!*
to read x the bed	Noun		Oh! Regarde le petit **lit**! Tu vois le petit **lit**? *Oh! Look at the small bed! Do you see the small bed?*
marcher x la marche	Verb	Marche	Oh Regarde! La petite **marche**! Tu vois? La petite **marche**! *Oh look! The little one **walks**! Do you see? The little one **walks**!*
to walk x the stairs	Noun		Oh! Regarde la petite **marche**! Tu vois la petite **marche**? *Oh! Look at the small stair! Do you see the small stair?*
porter x la porte	Verb	Porte	Oh Regarde! La petite **porte**! Tu vois? La petite **porte**! *Oh look! The little one **carries**! Do you see? The little one **carries**!*
to carry x the door	Noun		Oh! Regarde la petite **porte**! Tu vois la petite **porte**? *Oh! Look at the little door! Do you see the little door?*
sourire x la souris	Verb	[suri]	Oh Regarde! Le bébé **sourit**! Tu vois? La petite **sourit**! *Oh look! The baby **smiles**! Do you see? The the baby **smiles**!*
to smile x the mice	Noun		Oh! Regarde la petite **porte**! Tu vois la petite **porte**? *Oh! Look at the baby **mouse**! Do you see the baby **mouse**?*

References

Arunachalam, S., Escovar, E., Hansen, M. a, & Waxman, S. R. (2013). Out of sight, but not out of mind: 21-month-olds use syntactic information to learn verbs even in the absence of a corresponding event. *Language and Cognitive Processes*, *28*(4), 417–425.

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)*. Montreal (Canada).

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 Seventh Workshop on Cognitive Aspects of Computational Language Learning*.

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

Conwell, E., & Morgan, J. L. (2012). Is It a Noun or Is It a Verb? Resolving the Ambicategoricality Problem. *Language Learning and Development*, *8*(2), 87–112. http://doi.org/10.1080/15475441.2011.580236

Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech*. London, England: Harvard University Press.

Dautriche, I. (2015). *Weaving an ambiguous lexicon*. Unpublished PhD thesis, Universite Paris Descartes, Paris.

Dautriche, I., Fibla, L., & Christophe, A. (2015). Learning homophones: syntactic and semantic contexts matter. In *Paper presented at the 40th Boston University Conference on Language Development*. Boston - MA.

Dautriche, I., Swingley, D., & Christophe, A. (2015). Learning novel phonological neighbors: Syntactic category matters. *Cognition*, *143*, 77–86. http://doi.org/10.1016/j.cognition.2015.06.003

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., Bleam, 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

Delais-Roussarie, E. (1995). *Pour une approche parallèle de la structure prosodique: Etude de l'organisation prosodique et rythmique de la phrase française*. Université de Toulouse - Le Mirail, France.

Di Cristo, A. (2000). Vers une modélisation de l'intonation du français : seconde partie. *French Language Studies*, *9*, 143–179.

Ferguson, B., Graf, E., & Waxman, S. R. (2014). Infants use known verbs to learn novel nouns: evidence from 15- and 19-month-olds. *Cognition*, *131*(1), 139–46. http://doi.org/10.1016/j.cognition.2013.12.014

Fernald, A., Zangl, R., Portillo, A., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. Sekerina, E. Fernandez, & H. Clahsen (Eds.), *Developmental Psycholinguistics: On- line methods in children's language processing* (pp. 97–135). Amsterdam: John Benjamins.

Fisher, C. (1996). Structural Limits on Verb Mapping: The Role of Analogy in Children's Interpretations of Sentences. *Cognitive Psychology*, *31*(1), 41–81. http://doi.org/10.1006/cogp.1996.0012

Fisher, C. (2015). Two mechanisms for syntactic bootstrapping. In *Paper presented at the 40th Boston University Conference on Language Development*. Boston - MA.

Fisher, C., Hall, D. G., Rakowitz, S., & Gleitman, L. (1994a). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. *Lingua*, *92*(C), 333–375. http://doi.org/10.1016/0024-3841(94)90346-8

Fisher, C., Hall, G., Rakowitz, S., & Gleitman, L. (1994b). When it is better to receive than to give: Structural and cognitive factors in acquiring a first vocabulary. *Lingua*, *92*, 333–376.

Fougeron, C., & Keating, P. a. (1997). Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America*, *101*(6), 3728–40.

Gerken, L., Jusczyk, P. W. &, & Mandel, D. R. (1994). When prosody fails to cue syntactic structure 9-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition*, *51*, 237–265.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*(2), 135–176.

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

Gout, A., Christophe, A., & Morgan, J. L. (2004). Phonological phrase boundaries constrain lexical access II. Infant data. *Journal of Memory and Language*, *51*(4), 548–567. http://doi.org/10.1016/j.jml.2004.07.002

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

Hallé, P. a, Durand, C., & de Boysson-Bardies, B. (2008). Do 11-month-old French infants process articles? *Language and Speech*, *51*(Pt 1-2), 23–44.

Hawthorne, K., & Gerken, L. (2014). From pauses to clauses: Prosody facilitates learning of syntactic constituency. *Cognition*, *133*(2), 420–428. http://doi.org/10.1016/j.cognition.2014.07.013

Hawthorne, K., Rudat, L., & Gerken, L. (2016). Prosody and the Acquisition of Hierarchical Structure in Toddlers and Adults. *Infancy*, 1–22. http://doi.org/10.1111/infa.12130

He, A. X., & Lidz, J. (2014). Verb learning in 14- and 18-month-old English-learning infants. In *Paper presented at the 39th Boston University Conference on Language Development*. Boston - MA.

Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Cassidy, K. W., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, *26*, 269–286.

Höhle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science*, *6*(2), 122–127. http://doi.org/10.1111/1467-7687.00261

Höhle, B., Weissenborn, J., Kiefer, D., Schulz, A., & Schmitz, M. (2004). Functional Elements in Infants' Speech Processing: The Role of Determiners in the Syntactic Categorization of Lexical Elements. *Infancy*, *5*(3), 341–353. http://doi.org/10.1207/s15327078in0503_5

Johnson, E. K. (2008). Infants use prosodically conditioned acoustic-phonetic cues to extract words from speech. *The Journal of the Acoustical Society of America*, *123*(6), EL144-EL148. http://doi.org/10.1121/1.2908407

Jun, S. (2005). Prosodic Typology. In S. Jun (Ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing* (pp. 430–458). Oxford University Press.

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

Jusczyk, P. W., Hohne, E., & Mandel, D. (1995). Picking up regularities in the sound structure of the native language. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 91–119). Timonium, MD: Yorck Press.

Kern, S. (2007). Lexicon development in French-speaking infants. *First Language*, *27*(3), 227–250. http://doi.org/10.1177/0142723706075789

Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic Facilitation and Interference in the Resolution of Temporary Syntactic Closure Ambiguity. *Journal of Memory and Language*, *40*(2), 153–194. http://doi.org/10.1006/jmla.1998.2620

Landau, B., & Gleitman, L. (1985). *Language and experience: Evidence from the blind child*. Cambridge, MA: Harvard University Press.

Mandel, D. R., Jusczyk, P. W., & Nelson, D. G. (1994). Does sentential prosody help infants organize and remember speech information? *Cognition*, *53*(2), 155–80.

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

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–90. http://doi.org/10.1016/j.jneumeth.2007.03.024

Massicotte-Laforge, S., & Shi, R. (2015). The role of prosody in infants ' early syntactic analysis and grammatical categorization. *Journal of Acoustical Society of America*, *138*(October), 441–446. http://doi.org/10.1121/1.4934551

Mehler, J., Jusczyk, P., Lamsertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acqusiition in young infants. *Cogntion*, *29*, 143–178.

Millotte, S., René, A., Wales, R., & Christophe, A. (2008). Phonological phrase boundaries constrain the online syntactic analysis of spoken sentences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*(4), 874–885. http://doi.org/10.1037/0278-7393.34.4.874

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

Millotte, Morgan, J., Margules, S., Bernal, S., Dutat, M., & Christophe, A. (2010). Phrasal prosody constrains word segmentaion in French 16-month-olds. *Journal of Portuguese Linguistics*, *9–2*(10–1), 67–86.

Morgan, J. (1986). *From simple input to complex grammar*. Cambridge MA: MIT Press.

Morgan, J., & Demuth, K. (1996). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (Eribaum). Mahwah, NJ.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology. Human Perception and Performance*, *24*(3), 756–766.

Nazzi, T., Iakimova, G., Bertoncini, J., Frédonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French : Emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, *54*, 283–299.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris.

Shattuck-Hufnagel, S., & Turk, a E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, *25*(2), 193–247.

Shi, R. (2014). Functional Morphemes and Early Language Acquisition. *Child Development Perspectives*, *8*(1), 6–11. http://doi.org/10.1111/cdep.12052

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., Werker, J., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy*, *10*(2), 187–198.

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*, 574–608. http://doi.org/10.1016/j.jml.2007.08.001

Soderstrom, M. (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

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*(4), 869–902. http://doi.org/10.1017/S0305000908008763

Soderstrom, M., Seidl, A., 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

Swingley, D., & Aslin, R. . (2000). Spoken word recognition and lexical representation in very youn children. *Cognition*, *76*, 147–166.

Trueswell, J. C. (2008). Using eye movements as a developmental measure within psycholinguistics. In I. Sekerina, E. Fernandez, & H. Clahsen (Eds.), *Developmental Psycholinguistics: On-line methods in children's language processing* (pp. 73–96). Amsterdam: John Benjamins.

Trueswell, J. C., & Gleitman, L. R. (2007). Learning to parse and its implications for language acquisition. In M. G. Gaskell (Ed.), *Oxford Handbook of Psycholinguistics* (pp. 1–39). Oxford: Oxford University Press.

Trueswell, J., Sekerina, I., Hill, N., & Logrip, M. (1999). The kindergarten-path effect: studying on-line sentence processing in young children. *Cognition*, *73*(2), 89–134.

Veneziano, E., & Parisse, C. (2010). The Early Comprehension of Noun-Verb Distinction in French: An Experimental Method. In *XVIIth Biennial International Conference on Infant Studies*. Baltimore, Maryland, United States.

Veneziano, E., & Parisse, C. (2011). Retrieving the meaning of words from syntactic cues: A comprehension study of 2 to 4 yrs old French-speaking children. In *International Conference on the Study of Child Language, IASCL 2011*. Montreal, Canada.

Von Holzen, K., & Mani, N. (2012). Language nonselective lexical access in bilingual toddlers. *Journal of Experimental Child Psychology*, *113*(4), 569–586. http://doi.org/10.1016/j.jecp.2012.08.001

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. (2003). *The slaying of Lady Mondegreen: Being a study French of tonal association and alignment and their role in speech segmentation*. Ohio State University, Columbus.

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

Wickham, H. (2009). *ggplot2*. (R. Gentleman, K. Hornik, & G. Parmigiani, Eds.). New York, NY: Springer New York. http://doi.org/10.1007/978-0-387-98141-3

Yuan, S., & Fisher, C. (2009). "'Really? She Blicked the Baby?'" *Psychological Science*, *20*, 619–626.

Yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the Nouns : Simple Structural Cues to Verb Meaning. *Child Development*, *83*(4), 1382–1399.