Developmental Science

Developmental Science 19:2 (2016), pp 235–250 DOI: 10.1111/desc.12300

PAPER

# Preschoolers use phrasal prosody online to constrain syntactic analysis

Alex de Carvalho,[1,2] Isabelle Dautriche[1,2] and Anne Christophe[1,2]

*1. Laboratoire de Sciences Cognitives et Psycholinguistique (ENS, EHESS, CNRS), Département d'Études Cognitives (École Normale Supérieure – PSL Research University), Paris, France*
*2. Maternité Port-Royal, AP-HP, Faculté de Médecine Paris Descartes, France*

## Abstract

*Two experiments were conducted to investigate whether young children are able to take into account phrasal prosody when computing the syntactic structure of a sentence. Pairs of French noun/verb homophones were selected to create locally ambiguous sentences (*[la petite **ferme**] [est très jolie] *'the small **farm** is very nice' vs.* [la petite] [**ferme** la fenêtre] *'the little girl **closes** the window' – brackets indicate prosodic boundaries). Although these sentences start with the same three words,* ferme *is a noun (*farm*) in the former but a verb (*to close*) in the latter case. The only difference between these sentence beginnings is the prosodic structure, that reflects the syntactic structure (with a prosodic boundary just before the critical word when it is a verb, and just after it when it is a noun). Crucially, all words following the homophone were masked, such that prosodic cues were the only disambiguating information. Children successfully exploited prosodic information to assign the appropriate syntactic category to the target word, in both an oral completion task (4.5-year-olds, Experiment 1) and in a preferential looking paradigm with an eye-tracker (3.5-year-olds and 4.5-year-olds, Experiment 2). These results show that both groups of children exploit the position of a word within the prosodic structure when computing its syntactic category. In other words, even younger children of 3.5 years old exploit phrasal prosody online to constrain their syntactic analysis. This ability to exploit phrasal prosody to compute syntactic structure may help children parse sentences containing unknown words, and facilitate the acquisition of word meanings.*

## Research highlights

- In two experiments, 3.5- to 5-year-old children used phrasal prosody to disambiguate locally ambiguous sentences using noun/verb homophones.
- The effect of prosody was observed from the ambiguous word onset, indicating that children integrate prosody online.
- This is the first study to report that children under 4 years of age use phrasal prosody for syntactic analysis.
- This suggests that phrasal prosody as a cue to syntactic analysis would be available early on in development.

## Introduction

Parsing sentences into meaningful phrases and clauses is an essential step both in language comprehension and in acquisition. While the syntactic structure of sentences is not directly accessible from the input, it is often correlated with other features of the signal that are perceptually available. One such feature is phrasal prosody, the rhythm and melody of speech, that naturally structures utterances into phrases whose boundaries are aligned with syntactic constituent boundaries (e.g. Nespor & Vogel, 1986).

Past studies have shown that adults rapidly integrate phrasal prosody information when computing the syntactic structure of sentences (Kjelgaard & Speer, 1999; Millotte, René, Wales & Christophe, 2008; Millotte, Wales & Christophe, 2007; Snedeker & Yuan, 2008; Snedeker & Trueswell, 2003; Weber, Grice & Crocker, 2006). For example, Millotte *et al.* (2008) constructed locally ambiguous sentences in *French* using pairs of homophones that can be either an adjective or a verb. When the ambiguous word was a verb, there was a prosodic phrase boundary preceding it (e.g. [*Le petit*

Address for correspondence: Alex de Carvalho, Laboratoire de Sciences Cognitives et Psycholinguistique, École Normale Supérieure, 29 rue d'Ulm, PJ, 75005 Paris, France; e-mail: x.de.carvalho@gmail.com

*chien*] [***mord** la laisse*] [*qui le retient*] / [The little dog][**bites** the leash] [that holds it back], where prosodic boundaries are signaled by brackets) and following it when it was an adjective (i.e. [*Le petit chien **mort***] [*sera enterré demain*] / [The little **dead** dog] [will be buried tomorrow]). In a word detection task, adults detected adjectives faster and more accurately when listening to adjective sentences than when listening to verb sentences, and vice versa for verbs. Crucially, they could do so even before they heard the disambiguating content that followed the ambiguous word, showing that prosody was integrated on the fly to constrain syntactic analysis.

The idea that phrasal prosody could be used to guide the interpretation of sentences even in the absence of relevant lexical information has fostered a great interest in the language acquisition literature. Because phrasal prosody is easily recoverable from the speech signal itself, even in the absence of prior linguistic knowledge, it has been proposed that a prosodic analysis of the speech signal might inform early syntactic acquisition and processing (the prosodic bootstrapping hypothesis; Morgan & Demuth, 1996; Morgan, 1986), in conjunction with highly frequent elements such as function words (see e.g. Gervain, Nespor, Mazuka, Horie & Mehler, 2008; Gervain & Werker, 2013; Shi, 2014). Many experimental studies have shown that infants are sensitive to prosodic information from very early on. For example, infants exploit phrasal prosody to identify their mother tongue from birth onwards (e.g. Mehler, Jusczyk, Lamsertz, Halsted, Bertoncini *et al.*, 1988; Nazzi, Bertoncini & Mehler, 1998), they are sensitive to the coherence of prosodic constituents (at 4 months, for intonational phrases; Hirsh-Pasek, Kemler Nelson, Jusczyk, Cassidy, Druss *et al.*, 1987; from 6 months on, for smaller prosodic units; Gerken, Jusczyk & Mandel, 1994; Soderstrom, Seidl, Nelson & Jusczyk, 2003), they show better memory for whole prosodic units than for chunks that span prosodic boundaries (Mandel, Jusczyk & Nelson, 1994; Nazzi, Iakimova, Bertoncini & Alcantara, 2006) and they use prosodic boundaries to constrain lexical access by 10 months of age (Gout, Christophe & Morgan, 2004; Johnson, 2008; Millotte, Margules, Dutat, Bernal & Christophe, 2010).

However, despite the large literature showing the extensive experience that infants have with prosody, as far as we can tell no study has provided direct evidence that toddlers are able to use prosodic boundaries not only to facilitate memory or lexical access, but also to constrain syntactic computations. Given the interest the prosodic bootstrapping hypothesis has received, it may seem surprising that nobody has yet attempted such a demonstration. One potential reason might be that investigating the role of prosody in early syntactic processing is methodologically challenging: it requires presenting infants with sentences that contain a syntactic ambiguity (either local or global), and such sentences are difficult to come up with, especially given infants' reduced lexicon.

Given this methodological difficulty, researchers have instead examined preschoolers' ability to exploit prosody to recover the syntactic structure of ambiguous sentences (Choi & Mazuka, 2003; Choi & Trueswell, 2010; Snedeker & Trueswell, 2003; Snedeker & Yuan, 2008; Vogel & Raimy, 2002) based on the rationale that if toddlers are able to use phrasal prosody to break into syntax then prosody should still serve as a parsing cue in preschoolers. Surprisingly, although preschoolers have had extensive experience with prosody, and despite young infants' efficiency in processing phrasal prosody, most of these studies have failed to observe an effect of prosody on syntactic ambiguity resolution (Choi & Mazuka, 2003; Snedeker & Trueswell, 2003; Vogel & Raimy, 2002). A notable exception is the study conducted by Snedeker and Yuan (2008) showing that English-learning 5-year-olds successfully exploit prosody to interpret globally ambiguous sentences such as 'Could you tap the frog with the feather?', where the prepositional phrase 'with the feather' can be interpreted either as a modifier of the noun or as an instrument, depending on the prosodic structure. Sentences with an instrument interpretation were structured with a prosodic break after the first noun phrase (i.e. [could you tap the frog] [with the feather]) while sentences with a modifier interpretation had a prosodic break after the verb (i.e. [could you tap] [the frog with the feather]). However, these disambiguating prosodic breaks are not part of the normal prosodic structure of these sentences; rather, they can be intentionally added by the speaker when she is aware of the ambiguity (the default prosodic structure is [could you tap] [the frog] [with the feather] for both readings; Snedeker & Yuan, 2008). It is therefore difficult to infer from these studies whether or not younger children do exploit phrasal prosody in their processing of everyday non-ambiguous sentences.

The experiments that follow explore whether preschoolers exploit phrasal prosody to guide their syntactic analysis of sentences when the prosodic cues to syntactic structure are systematic and present in natural speech. Our interest in this question is twofold: First, showing a robust effect of naturally occurring prosody in preschoolers would clarify the mixed results that were previously obtained with rare and non-systematic prosodic cues. Second, although studying online sentence processing in preschoolers cannot directly inform the prosodic bootstrapping hypothesis, finding an effect of prosody on syntactic processing in preschoolers would leave open the possibility that phrasal prosody could be used at a younger age, a hypothesis that

was previously neglected following preschoolers' failure to exploit prosody.

More specifically, we tested children on locally ambiguous sentences which differ in their default prosodic structure, so that the disambiguating prosodic information is naturally produced by naive speakers – whether the sentence is ambiguous or not (Millotte *et al*., 2007). As in Millotte *et al*. (2008), pairs of homophones belonging to different syntactic categories (here, noun and verb) were used to create locally ambiguous sentences such as the following:

1 [*la petite*$_{\text{ADJ}}$ ***ferme***$_{\text{NOUN}}$] [*est très jolie*]
[the small$_{\text{ADJ}}$ **farm**$_{\text{NOUN}}$] [is very nice] (noun prosody)
2 [*la petite*$_{NOUN}$][1] [***ferme***$_{VERB}$ *la fenêtre*]
[the little one$_{\text{NOUN}}$] [**closes**$_{\text{VERB}}$ the window] (verb prosody)

Although both sentences start with the same three words, which have the same pronunciation (i.e., /lapətitfɛʁm/), they are disambiguated by their prosodic structure. That is, when the critical word *ferme* is a noun, it is part of the first prosodic phrase, and it is immediately followed by a prosodic boundary (see example 1). By contrast, when *ferme* is a verb, it is part of the second prosodic phrase, immediately preceded by a prosodic boundary (see example 2). Thus, in both sentences, when the ambiguous word is being processed, only the prosodic structure may allow listeners to determine its syntactic category.

In two experiments, an oral completion task (Experiment 1) and an intermodal preferential looking task (Experiment 2), we investigated whether 3.5- and 4.5-year-old children are able to take into account the position of a word within the prosodic structure when computing its syntactic category (noun vs. verb).

## Experiment 1: Oral completion task

In this experiment, participants listened to the beginnings of sentences that were cut just after the end of the ambiguous word (i.e. after *ferme* in the examples above). Sentences were produced naturally, but all words following the homophone were replaced by an acoustic mask made with babble noise. As a result, only the prosodic structure of the beginning of the sentence could be used to decide whether the target word was a noun or a verb.

[1] In French, the adjective petite can be used as a noun (i.e. *la petite*, meaning *the little 'girl'*, where the pronoun (one) is omitted). Many other adjectives allow for a similar use (e.g. *le grand* / *la grande* – the big boy / the big girl).

In this task, children were asked to complete the sentences in any way they liked. The nature of their completion allowed us to determine whether they interpreted the ambiguous word as a noun or as a verb. For example, if a child heard the sentence beginning '*la petite ferme*…' (either 'the small farm…' or 'the little girl closes…' depending on its prosody), an answer such as '…is very nice' (containing a verb and its complement) suggested that the target word was processed as a noun (part of the subject noun phrase): we called 'noun completions' all completions where the critical word was unambiguously a noun. By contrast, an answer such as '…the door' suggested that the child had interpreted the target word as a verb, and we called these answers 'verb completions'. If children exploit prosodic information to constrain their syntactic analysis, we would expect to observe more noun completions for sentences uttered with a noun prosody and more verb completions for sentences uttered with a verb prosody.

### *Method*

#### Participants

Sixteen 4- to 5-year-old monolingual French-speaking children (4;3 to 5;3, $M_{\text{age}}$ = 4;9, nine boys) were tested in a public preschool in Paris. Their parents signed an informed consent form. An additional three children were tested, but were not included in the final analysis because they failed to complete all training sentences prior to the test phase.

#### Materials

Eight pairs of experimental sentences were created from eight pairs of noun-verb homophones in French. Most of these words were likely to be known by 3-year-old children according to the McArthur database for French (Kern, Langue, Zesiger & Bovet, 2010; Kern, 2007). For each pair of homophones, we created two sentences: one with the ambiguous word used as a noun (hereafter the noun prosody condition, e.g. [*La*$_{\text{DET}}$ *petite*$_{\text{ADJ}}$ ***ferme***$_{\text{NOUN}}$] [*lui plait beaucoup*]) and a second one with the ambiguous word used as a verb (hereafter the verb prosody condition, e.g. [*La*$_{\text{DET}}$ *petite*$_{\text{NOUN}}$ ] [***ferme***$_{\text{VERB}}$ *le coffre à jouets*]; see the Appendix for a complete list of test sentences). All sentences were recorded in a soundproof booth by a female French speaker (the last author) who was aware of the purpose of the study and used child-directed speech. The sentences were recorded in pairs, each with a noun or verb prosodic structure. Note that the prosodic differences between the two types of sentences are naturally produced by naïve adults even

when they are unaware of the syntactic ambiguity of the target words (Millotte *et al.*, 2007) and are consistent with theoretical descriptions of the relationship between prosody and syntax (Nespor & Vogel, 1986). Nonetheless, we assessed the differences between conditions by conducting acoustic analyses (duration and pitch) on the segments around the critical region using Praat.

The analysis of duration (Figure 1) revealed a significant phrase-final lengthening, as expected from the literature (Delais-Roussarie, 1995; Jun & Fougeron, 2002; Millotte *et al.*, 2008, 2007; Shattuck-Hufnagel & Turk, 1996). We analyzed the prosodic boundaries marked in the figure by black vertical lines: just before the ambiguous word in the verb prosody condition and just after it in the noun prosody condition. The rhyme of the syllable immediately preceding the prosodic phrase boundary in the verb condition (e.g. /it/ in Figure 1) was lengthened by 98% compared to the noun condition ($M_{\text{verb}}$ = 403 ms, $SD_{\text{verb}}$ = 50.4 vs. $M_{\text{noun}}$ = 204 ms, $SD_{\text{noun}}$ = 22.01; $t(7) = -3.85$, $p < .01$), and the rhyme of the syllable immediately preceding the prosodic phrase boundary in the noun condition (e.g. /ɛrm/ in Fig. 1) was lengthened by 35% compared to the verb condition ($M_{\text{noun}}$ = 427 ms, $SD_{\text{noun}}$ = 50.6 vs. $M_{\text{verb}}$ = 317 ms, $SD_{\text{verb}}$ = 34.9; $t(7) = 3.77$, $p < .01$). In addition, following Fougeron and Keating (1997), we also analyzed phrase-initial strengthening:[2] the onset of the target word in the verb condition (phrase-initial) was lengthened by 70% compared to the noun condition (phrase-medial; $M_{\text{verb}}$ = 205 ms, $SD_{\text{verb}}$ = 16.2 vs. $M_{\text{noun}}$ = 121 ms, $SD_{\text{noun}}$ = 9.2; $t(7) = -5.02$, $p < .01$). Pitch analyses[3] compared the maximum F0 of the first vowel of the target word with the last vowel of the preceding word (e.g. /i/ from /pətit/ and /ɛ/ from /fɛrm/) in both prosodic conditions. These vowels were on each side of the prosodic boundary in the verb condition and belonged to the same prosodic unit in the noun condition. This analysis revealed a significant difference between conditions, consistent with the literature describing French as having a tendency for a rising pitch contour towards the end of prosodic units (+50 Hz in the noun condition versus −35 Hz in the verb condition, $t(14) = 18.04$, $p < .01$) (Di Cristo, 2000; Welby, 2003, 2006). In the noun condition, this surfaced as a rising pitch pattern between the last syllable of the adjective (e.g. /i/ from 'petite') and the noun (e.g. /ɛ/ from 'ferme' when both syllables were at the end of prosodic unit, +50 Hz). In the verb condition, this resulted in a falling contour between the noun 'petite' and the verb 'ferme' (the vowels then spanning the prosodic boundary). In addition, no pauses were observed between any of the words in both prosody conditions. Thus to differentiate between the noun and the verb prosodic structures, children had to be able to correctly interpret the prosodic structure of the sentences and could not have relied on a simpler strategy such as exploiting pauses to recognize the boundaries between syntactic constituents.

In addition to experimental sentences, we created 11 filler sentences featuring target words that were unambiguously either a noun or a verb (e.g. [*Le bébé* ***oiseau***] [*mange beaucoup*] 'the baby **bird** eats a lot'; [*La maîtresse*] [***parle*** *aux enfants*] 'the teacher **talks** to the children').

In order to make the experiment child-friendly, all stimuli were videotaped recordings of the female speaker. Each sentence was cut after the target word and 1000 ms of babble noise, created by superimposing the end of all filler sentences, was added. This babble noise was identical across test sentences. To create an analogous effect in the visual domain, the video of the speaker lost contrast, became blurred, and trembled, starting right at the offset of the target word (making lip-reading fully impossible, see Figure 2). This manipulation gave credit to the story that 'the television didn't work properly', and ensured that participants could only rely on prosodic information to interpret sentences, since the disambiguating information following the ambiguous word was not available (no acoustic or visual information was available after the end of the target word).

To ensure that there were no co-articulatory differences between words of the same homophone pair across conditions, the word following the target word always started with the same consonant (e.g. noun prosody condition: *la petite ferme*$_{\text{N}}$ ***l****ui plait beaucoup* and verb prosody condition: *la petite ferme*$_{V}$ ***l****e coffre à jouets*, both words start with an /l/).

An example of a trial outline is depicted in Figure 2.

In total, we created 16 test videos from the eight pairs of homophones; eight in the verb condition and eight in the noun condition. We created two lists of stimuli, so that each member of a given sentence pair appeared in a different list. Each list contained four sentences with the noun prosody and four sentences with the verb prosody,

[2] According to Fougeron and Keating (1997), the onset of words located at the beginning of a prosodic unit should be lengthened relative to when they are located in the middle or at the end of a prosodic unit. Thus, the onset of the ambiguous word in the verb prosody condition (e.g. /f/ for 'ferme', where 'ferme' is phrase-initial) should be longer than the onset of this same word in the noun prosody condition (where it is phrase-medial).

[3] Intonation in French is characterized by a sequence of rising pitch movements demarcating phonological phrase boundaries (Jun & Fougeron, 2002) and the final full syllable of a word at the end of a prosodic unit typically bears a rise in fundamental frequency (Vaissière & Michaud, 2006) together with longer duration and possibly a higher intensity (Di Cristo, 1998; Jun & Fougeron, 2002).

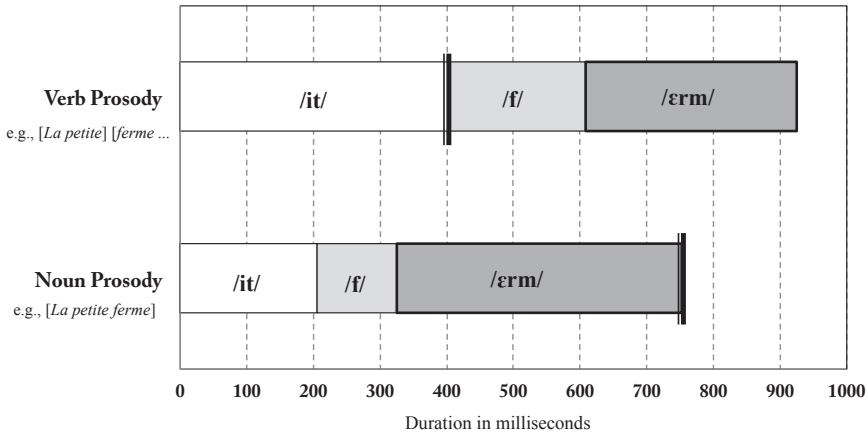


**Figure 1** *Mean duration (in ms) of the different segments around the prosodic boundaries for both conditions: noun and verb prosody (phonological phrase boundaries are represented by thick black lines). Note that to illustrate, we use the segments for the experimental sentences of the item /fɛrm/, but the numbers correspond to mean values across all test sentences.*

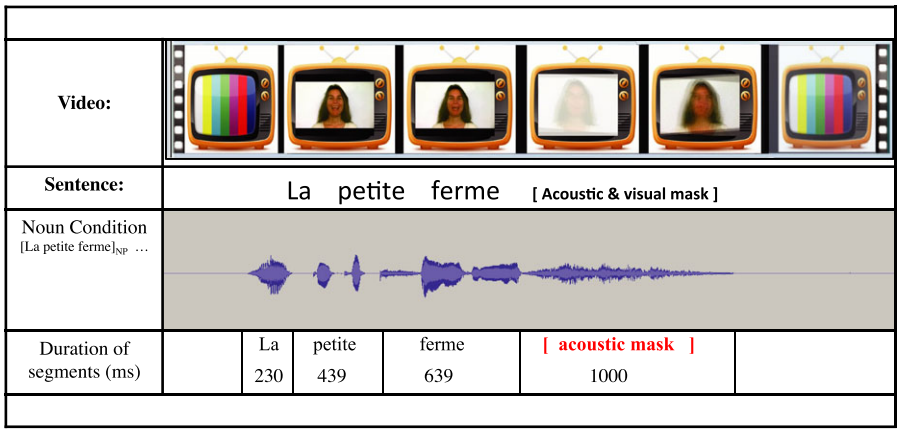


**Figure 2** *Example of a test sentence used in the completion task (Experiment 1) together with its waveform and the duration of each of the components.*

plus four filler sentences. Each participant listened to only one list. Half of the participants were assigned to each list and the order of sentences within each list was randomized with the constraint that no more than two test sentences could appear one after the other.

## Procedure

Children were tested individually in a quiet room in their preschool. During the experiment, children sat in front of a computer and wore headphones to listen to the stimuli. A game-like task was used to elicit children's completions of the test stimuli. At the beginning of the task, the experimenter told the child that he or she would listen to a woman on a television screen. However, because the television was broken, the child could not hear the end of the story and would have to guess what the woman might have said. To motivate children to give an answer to all sentences, the experimenter told them that they were in competition with other children and that the one who gave the most story completions would win the game. A screenshot of the screen viewed by children is shown in Figure 3.

As depicted in Figure 3, for each trial an arrow rotated in the middle of the screen and selected one of the children to complete a sentence. If the arrow pointed downward, it was the participant's turn to answer. The virtual children were chosen only to answer filler sentences. All test sentences had to be completed by the participant. When a virtual child was selected to respond, a pre-recorded sentence was played; these sentences were previously recorded from children of the

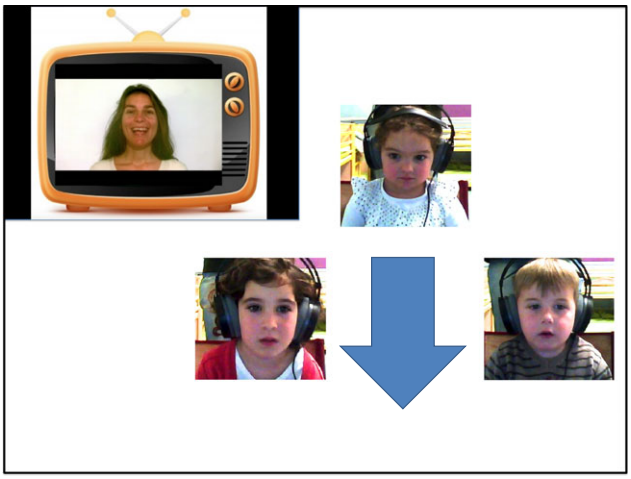

**Figure 3** *Example of the scenario used in the completion task (Experiment 1) for each trial: first the blue arrow turned and selected which child would play. Then, children saw the video in the upper left corner of the screen as illustrated here. Finally, they completed the sentence they had heard in the video.*

same age as our participants. When the arrow selected the participant, the experimenter asked her to pay attention to the video that was coming up and to complete the sentence in any way she wanted to. When the arrow pointed toward a virtual child's picture, the experimenter interacted with this virtual child in the same way he did with the actual participant, providing encouragement to respond. All participants listened to the same virtual children's answers for all the filler sentences.

Participants started the experiment with a practice block. In this block, children were presented exclusively with filler sentences. The virtual children answered the first two completion trials of this block in order to introduce the participant to the task. Then, starting from the third completion trial, the arrow chose the child participant. All children completed between two and seven of these filler sentences and as soon as they had given two correct answers, the test session started.

The test session was composed of eight test sentences and four filler sentences. Half of the test sentences were in the noun prosody condition and half in the verb prosody condition. All filler sentences were completed by the virtual children, and all test sentences were completed by our participants. Using filler sentences in this task allowed us to justify the 'competition game' proposed to children (since these sentences were completed by the virtual children), and in addition it minimized the risk that participants could become aware of the presence of ambiguous words in the experiment.

## Data analysis

To examine children's use of prosody to disambiguate ambiguous noun/verb homophones, their answers were coded as *noun answers* when they gave a completion using the target word as a noun (e.g. '. . . is very nice'), or as *verb answers* when they used the target word as a verb (e.g. '. . . the window'). Children's responses were coded offline by two independent coders who each listened to all the recordings of children's answers, without knowing which of the sentence beginnings had been heard. Agreement between coders was 100%. Seventeen out of the 128 responses were excluded from our analysis (11 from the verb prosody condition) because the child did not answer ($n$ = 7), or because the answer was consistent with both interpretations of the target word ($n$ = 10). For example, for a sentence with the target word '*marche*', ambiguous between the noun 'step' (from a staircase) and the verb 'to walk', a response such as 'on the floor' was considered to be ambiguous between both interpretations (because the child could have meant either 'the large step on the floor' or 'the tall girl walks on the floor' – the prosody of the child's utterance was not taken into account when coding the answers).

Because noun and verb responses in this task were complementary, we chose the occurrence of a noun answer (0 or 1) as our dependent measure. Since we analyzed categorical responses we modeled them using logit models (following Jaeger, 2008). We ran mixed model analyses using R 2.15 and the lme4 package (v 1.0; Bates & Sakar, 2007). Each response $R_{is}$ for item $i$ and subject $s$ is modeled via an intercept$\beta_0$, reflecting the baseline probability of giving a noun answer, and a slope estimate$\beta_1$ of the predictor variable Condition $C_i$ (Noun prosody or Verb prosody depending on the item $i$), reflecting the likelihood of occurrence of $R_{is}$ with the predictor $C_i$. $\beta_1$ thus reflects the increase in the probability of noun responses in the noun condition relative to the verb condition. Since we used the maximal random effect structure (as suggested by Barr, Levy, Scheepers & Tily, 2013), we also included by-subject and by-item intercepts ($S_{0s}$ and $I_{0i}$ allowing the baseline to vary from a fixed amount from$\beta_0$ for each subject $s$ and each item $i$) and slopes ($S_{1s}$ and $I_{1i}$, respectively, allowing each subject and item to deviate from the population slope$\beta_1$ in their sensitivity to the condition factor). We assumed no effect of trial order or list presentation beyond the effect of items. The resulting equation for the model is the following:

$$\text{Logit}(\text{P}(R_{is}=1)) = \beta_0 + S_{0s} + I_{0i} + (\beta_1 + S_{1s} + I_{1i})C_i + e_{is}$$

where $e_{is}$ is the normally distributed error for the observation. $\beta$ estimates are given in log-odds (the

space in which the logit models are fitted). To compute the increase in absolute probability of giving a noun answer across different levels of $C_i$ (the prosodic condition: verb vs. noun), we can calculate: P($R_{is}$ = 1; $C_i$ = Noun condition) − P($R_{is}$ = 1; $C_i$ = Verb condition) by taking the inverse logit of the right-hand side of the previous equation using the estimates β given by the model.

We computed two tests of significance: the Wald's Z statistic, testing whether the estimates are significantly different from 0, as well as a $\chi^2$ test over the change in likelihood between two mixed models that both had the maximal random structure (as recommended by Barr *et al.*, 2013) but differed in the presence or the absence of the considered predictor ($C_i$ factor). Since the results are similar for the two tests, we report the Z statistic only. The categorical predictor Condition $C_i$ was coded as 0 for the verb prosody and 1 for the noun prosody. Hence the intercept corresponds to the probability of giving a noun response when children are in the verb prosody condition, while the slope corresponds to the increase in the probability of giving a noun response in the noun prosody condition relative to the verb prosody condition.

### *Results*

Figure 4 presents the average proportion of noun and verb answers for each prosody condition.

Children gave more noun answers in the noun prosody condition than in the verb prosody condition. This was reflected in our mixed model analysis by a main effect of the predictor Condition ($\beta$ = 3.83; $z$ = 5.29; $p$ < .001), corresponding to an increase of 0.73 in the probability of giving a noun response in the noun condition relative to the verb condition.

### *Discussion*

In an oral completion task, 4.5-year-olds assigned different syntactic categories to an ambiguous word depending on its position within the prosodic structure of the sentence. Upon hearing '*la petite ferme*' where the word '*ferme*' is ambiguous between a noun and a verb, they gave more noun completions (e.g. 'is really nice') in the noun condition ([*la petite ferme*]$_{NP}$ / the small farm) than in the verb condition ([*la petite*]$_{NP}$ [*ferme*]$_{VP}$ / the little 'one' closes) even though the only disambiguating information between the two sentence beginnings was phrasal prosody.

These results mirror previous results with adults (Millotte *et al.*, 2007, 2008) and show that 4.5-year-olds are able to use the prosodic structure of a sentence to solve local syntactic ambiguities. Yet, while children's interpretation of sentences is influenced by the prosodic structure of the sentence, it is unclear *when* the prosodic information is integrated during the parsing process. Since children were free to take as much time as they wanted to complete the test sentences, the prosodic information might be integrated relatively late during the parsing process in this task. To investigate whether children integrate prosodic information online, we conducted a second experiment using a paradigm tapping into the time course of sentence interpretation.

## Experiment 2: Intermodal preferential looking task

To investigate whether children use prosody during online sentence processing and its syntactic analysis, we conducted a second experiment using the same audio stimuli as in Experiment 1. However this time,

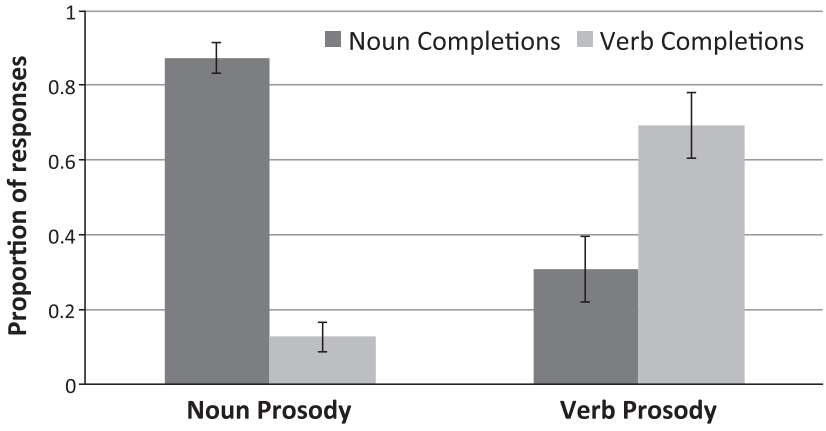


**Figure 4** *Proportion of noun and verb completions for each prosody condition. Error bars represent the standard error of the mean.*

the beginnings of the ambiguous sentences (e.g. '*la petite ferme* . . .') were paired with two images displayed side-by-side on a screen. One of these images was associated with the noun interpretation of the ambiguous word (e.g. a farm) and the other one with the verb interpretation (e.g. a little girl closing something). Children were asked to point toward the image that represents, in their opinion, the correct interpretation of the sentence they just heard. During this task, both the time course of children's eye-gaze and their pointing responses toward the images were recorded.

To perform well in Experiment 1, children's lexicon had to be quite advanced. Not only did they have to understand the meaning of all ambiguous words, they also had to complete the sentences in their own words. Experiment 2, in contrast, is less demanding, in that no explicit production was required. For this reason, we were able to test a second group of children of 3.5 years of age. If children exploit prosodic information online during sentence processing, we expect them to choose the image representing the noun interpretation more often when they listen to the beginning of noun sentences than when they listen to the beginning of verb sentences. We also expect them to switch their eye-gaze towards the correct image as soon as they start processing the prosodic information.

## *Method*

### Participants

Forty children participated in this experiment. All were monolingual native French speakers. Children fell into one of two age groups: either the 3.5-year-old group (3;4 to 4;0, $M_{age}$ = 3;7, $n$ = 20) or the 4.5-year-old group (4;3 to 5;10, $M_{age}$ = 4;8, $n$ = 20). Children were tested in a public preschool in Paris and their parents signed an informed consent form. An additional five children participated in the study but were not included in the final analysis because they were exposed to languages other than French at home ($n$ = 3), or because of fussiness during the experiment resulting in more than 50% (out of eight) of unusable test trials with missing eye-tracking data ($n$ = 2).

In addition, 14 adults, native speakers of French, participated in the same test, to provide us with a baseline.

### Material

We used the same eight pairs of ambiguous test sentences and the 11 unambiguous filler sentences recorded for Experiment 1, extracted from the videos. Sentences were played while children were presented with two images displayed side-by-side on the screen. For filler sentences, one image corresponded to the target word and the other was unrelated but represented a word from the opposite syntactic category. Thus, if the filler target word was a noun then the other image depicted an action. For each pair of noun-verb ambiguous sentences, one image represented the noun meaning and the other one the verb meaning. A total of 38 images (16 for test sentences and 22 for filler sentences) were created. These images were drawn by a designer and they were line drawings of approximately equal size and complexity.

### Procedure

Children were tested individually in a silent room in their own preschool. During the experiment, participants were seated approximately 60 cm away from a 19″ computer screen displaying the visual stimuli. As in Experiment 1, children wore headphones to listen to the audio stimuli. Children were told that they were going to play a game in which they would have to find the image belonging to the sentence they would listen to.

As in Experiment 1, each participant started the experiment with a practice session consisting of filler sentences in which the target word was unambiguous. The practice session consisted of at least four filler sentences. As soon as participants gave two correct pointing responses, the experimenter started the test session. The test session was composed of 12 trials: eight test sentences and four filler sentences, half with verb prosody and half with noun prosody counterbalanced between participants. We used the same two lists of stimuli as in Experiment 1 so that each child heard only one sentence from each noun-verb pair.

Each trial started with an inspection period to provide the children with sufficient time to inspect the pair of images displayed on the screen. Each image was first presented alone for 3 seconds on the left or the right side of the screen and a neutral audio prompt was played at the same time (e.g. 'Oooh look!'). Both images were then simultaneously presented on the screen, 17 cm apart from one another, without any acoustic stimulus for 3 seconds. Then these images disappeared and a colorful fixation target appeared in the middle of the screen. Once participants fixated the central fixation point, the two images reappeared on the screen and the auditory sentence was played. Following auditory sentence presentation, participants had to choose which image matched the sentence they heard. After children gave their response, the experimenter, who was standing behind the child but could not hear what the child heard, selected the image the child pointed to and the

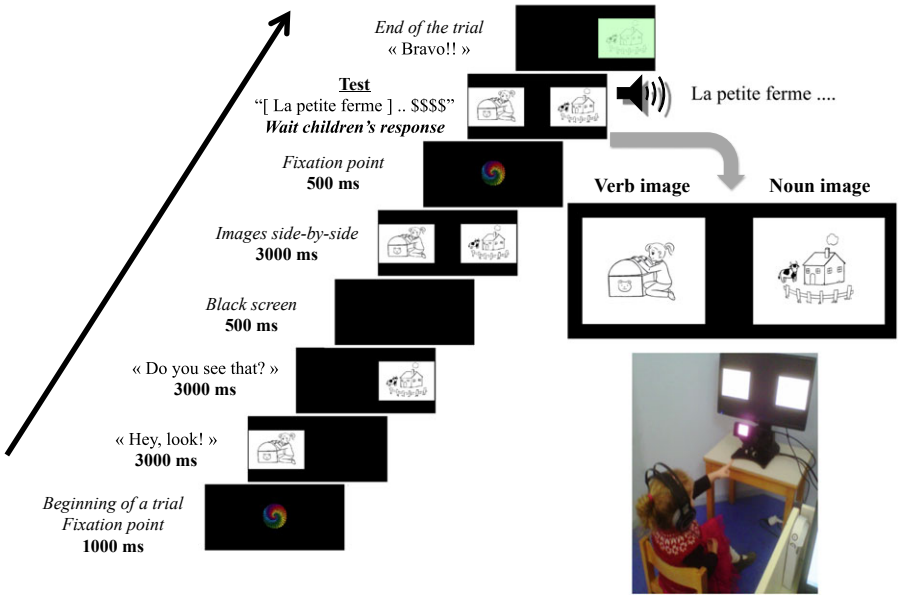


**Figure 5** *Time course of a trial in Experiment 2. Each trial started with a fixation point in the middle of the screen for 1000 ms. Then, each image was presented alone for 3 seconds on the left or the right side of the screen with an audio prompt. Then, after a 500-ms black screen, both images were presented simultaneously side-by-side without any audio materials for 3 seconds. The fixation point reappeared and as soon as participants fixated this fixation target, the test period started. The test sentence started playing immediately once the images appeared on the screen. Finally, participants had to point to the image which they thought corresponded to the sentence they heard. The selected picture then started blinking in green and participants heard a clapping sound.*

selected picture started blinking in green. At that point, the child also heard a clapping sound, regardless of whether the response was correct. The time course of each trial is described in Figure 5.

## Data processing

Participants' eye-gazes were recorded using an Eyelink 1000 while they listened to each test sentence and until they pointed toward one of the two images. Seventeen trials out of 320 (nine in the noun condition and six in the verb condition) were removed from the statistical analysis because more than 25% of the data frames between the onset of the ambiguous word and the end of the audio stimuli were missing. Note that these trials were still included in the pointing responses. Children pointed for every trial because the experimenter prompted them to do so.

## Data analysis

As in Experiment 1, we conducted a mixed model analysis for the pointing data (see data processing section). For the eye-gaze data, we analyzed for each age group the proportion of fixations toward the noun image (since fixations to noun vs. verb image are almost complementary, apart from the time spent looking away), and conducted a cluster-based permutation analysis (Maris & Oostenveld, 2007) to find a time window where a significant effect of condition was observed. This analysis allows us to test for the effect of Condition on each time point without inflating the rate of Type I error. For each time point we conducted a paired two-tailed $t$-test on the proportion of looks toward the noun picture between the noun and the verb prosody condition. Adjacent time points with a $t$-value greater than some predefined threshold ($t = 1.5$)[4] were grouped together into a cluster. The statistic for the cluster was defined as the sum of the $t$ statistics of each time point within the cluster. 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 computed

[4] The value of the threshold does not affect the rate of false alarms of the test. In our case, we chose a rather small threshold to detect subtle differences of timing between the three age groups.

the statistic 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 statistic is greater than the statistic of the largest cluster found in 95% of the simulations (ensuring a *p*-value of .05). This analysis was conducted from −700 ms before the onset of the ambiguous word until 1500 ms after the end of the ambiguous word. Note that in 41 trials (six for 3.5-year-olds, 19 for 4.5-year-olds and 16 for adults), participants gave their answer before 1500 ms (no more than 200 ms before). For the analysis to work properly, we extended the participant's final data point until the end of the trial. Plots of eye-gaze data have been created using the ggplot2 package (Wickham, 2009).

### *Results*

We report two analyses looking at (1) the pointing responses, reflecting children's final interpretation of the target word and (2) the time course of children's and adults' eye-gaze, reflecting their online interpretation of sentences as the linguistic input unfolds.

#### Pointing task

Figure 6 presents the average proportion of pointing responses toward the noun and the verb images for each condition (noun prosody or verb prosody) for both groups of children.

As can be seen in the figure, children pointed more toward the noun image than toward the verb image when they heard the beginning of test sentences with noun prosody, and vice versa for the test sentences with verb prosody. This was confirmed by our mixed model analysis: we modeled the occurrence of a pointing response toward the noun image with two categorical predictors and their interaction: Condition (Noun prosody, Verb prosody) and Age (3.5-year-olds, 4.5-year-olds). Our final model included by-subject and by-item intercepts and slopes yielding a maximal random effect structure (cf. Barr *et al.*, 2013). For the predictor Age, we coded as −0.5 the 3.5-year-olds and 0.5 the 4.5-year-olds and for the predictor Condition we coded as 0 the verb condition and 1 the noun condition. As a result, the intercept was the proportion of noun answers averaged across the two age groups in the verb condition and the estimate of the predictor Condition could directly be interpreted as a 'main effect' of prosody. This main effect of Condition ($\beta = 2.46$; $z = 5.80$; $p < .001$), which predicts an increase of 0.54 in the probability of pointing to the noun picture in the noun condition compared to the verb condition, was statistically significant. Although there was no significant effect of Age ($p > .6$), nor an interaction between Age and Condition ($p < .15$), inspection of the results suggests that the behavior of the 4.5-year-olds is more stable than that of the 3.5-year-olds. A post-hoc analysis looking at 3.5-year-olds nonetheless revealed a significant effect of Condition ($\beta = 2.08$; $z = 4.62$; $p < .001$) for the younger children, reflecting an increase of 0.46 in the probability of pointing to the noun picture in the noun condition compared to the verb condition. This suggests that both age groups performed well in the task.

#### Temporal analysis of eye movements

Figure 7a–c shows the average proportion of looks toward the noun image in the noun condition (red)

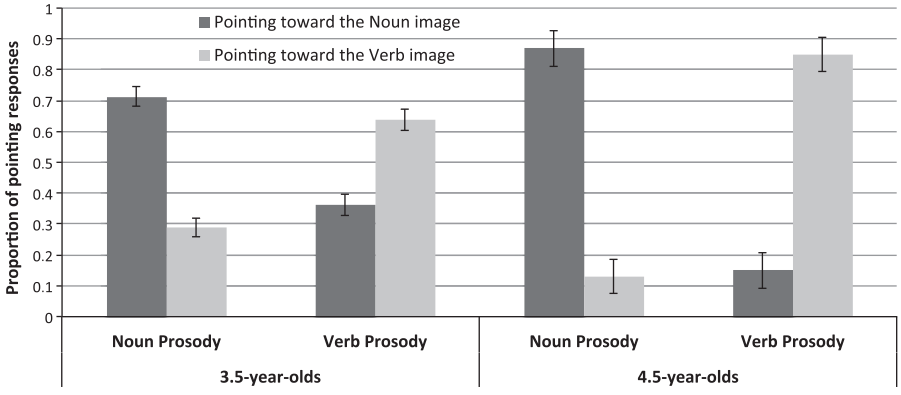


**Figure 6** *Proportion of pointing responses toward the noun image and the verb image after listening to the target word, broken down by prosody condition, for each group of participants. Error bars represent the standard error of the mean.*

and in the verb condition (blue), time-locked to the beginning of the ambiguous word, for the three age groups (i.e. 3.5-year-olds, 4.5-year-olds, and adults).

Visual inspection of the data shows that adults and children look more at the verb image at the onset of the ambiguous word (this was especially pronounced for children as both curves start around the 0.25 level at the beginning of the ambiguous word). This initial gaze is likely to be driven by the interpretation of the adjective (e.g. *La vieille* –'the old'; *La petite* – 'the small one'/'the little one'; *Le bébé* – 'the baby'), which is more likely to describe a human (always pictured in the verb image) than an object (always pictured in the noun image). Crucially, however, participants in all age groups increased their looks toward the noun image more so in the noun condition than in the verb condition, starting at or shortly after the onset of the ambiguous word, depending on the age group.

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 for all three age groups: 3.5-year-olds (from 226 ms after the beginning of the critical word; $p < .01$), 4.5-year-olds (from 14 ms after the beginning of the critical word; $p < .001$) and adults (from 54 ms before the beginning of the critical word; $p < .001$). Thus, adults and 4.5-year-olds were more than 200 ms faster than 3.5-year-olds to switch their gaze toward the noun picture in the noun prosody condition than in the verb prosody condition.

### *Discussion*

In this experiment we tested whether children are able to use prosody online to compute the syntactic category of ambiguous words. The results of the pointing task replicated the findings observed in Experiment 1 for the 4.5-year-olds and extended it to the younger 3.5-year-olds. Children from both age groups correctly interpreted the syntactic category of an ambiguous word based on its position within the prosodic structure of the sentence. Children interpreted the ambiguous word as a noun when it was embedded in a sentence with a noun prosodic structure and as a verb when it was embedded in a sentence with a verb prosodic structure. Moreover, the eye-tracking data reveal that while children initially looked toward the verb image (likely because hearing the adjective led them to turn toward the picture that contained humans), when they heard the beginning of a noun sentence, they appropriately switched their gaze toward the noun image by the end of the ambiguous word. Taking into account the 200–300 ms that are necessary to program an eye movement (Allopenna, Magnuson & Tanenhaus, 1998), this suggests that participants computed the syntactic category of a word before its offset. This pattern of response was observed for all three age groups, although the timing of eye movements was faster and more accurate for adults and 4.5-year-olds, who started to switch their gaze around the onset of the ambiguous word. The slight delay for 3.5-year-olds could be due to one of two reasons (or a combination of both): First, young 3.5-year-olds may be slower at accessing the meaning of words in their lexicon and/or may be slower to integrate prosodic information than 4.5-year-olds and adults. Second, 3.5-year-olds' responses may be more variable as a result of poorer attentional skills. Although our data do not allow us to tease apart these two possibilities, we can conclude that upon hearing the first words of a sentence, both adults and children exploit prosody online to calculate the syntactic category of a word.

One question that remains open is whether this ability is specific to the presence of ambiguity. Because children were presented with side-by-side images – one consistent with the noun interpretation, and the other with the verb interpretation – they might have become aware that the target word had two possible meanings, and might have paid special attention to prosody because the situation was ambiguous. We consider this unlikely for three main reasons: (1) in Experiment 1, children were able to exploit phrasal prosody to constrain their syntactic analysis even though the two interpretations of the ambiguous word were not presented visually. (2) Several studies have shown that when adults are asked to identify unambiguously an object that has a homophonous label (e.g. a baseball bat), they produce the ambiguous label (e.g. 'look at the bat') even when the homophonous object (e.g. an animal bat) is present on the display; in contrast when a second exemplar of the same category is present (e.g. another baseball bat), they disambiguate it with an adjective or a relative clause (e.g. 'look at the red bat') (Ferreira, Slevc & Rogers, 2005; Rabagliati & Snedeker, 2013). This shows that speakers do not spontaneously notice homophones that do not overlap semantically. Although we used a comprehension task rather than a production task (which may make a difference with respect to the processing of ambiguous words), it is worth noting that in our case the semantic distance between the two meanings of the homophone was even larger, since one meaning referred to an object (the noun) and the other to an action (the verb), a feature which should reduce even further the likelihood that subjects will notice the ambiguity. Anecdoctically, none of the adults who took part in this experiment reported being aware of the ambiguity of the test trials. (3) Finally, to minimize the risk that participants could

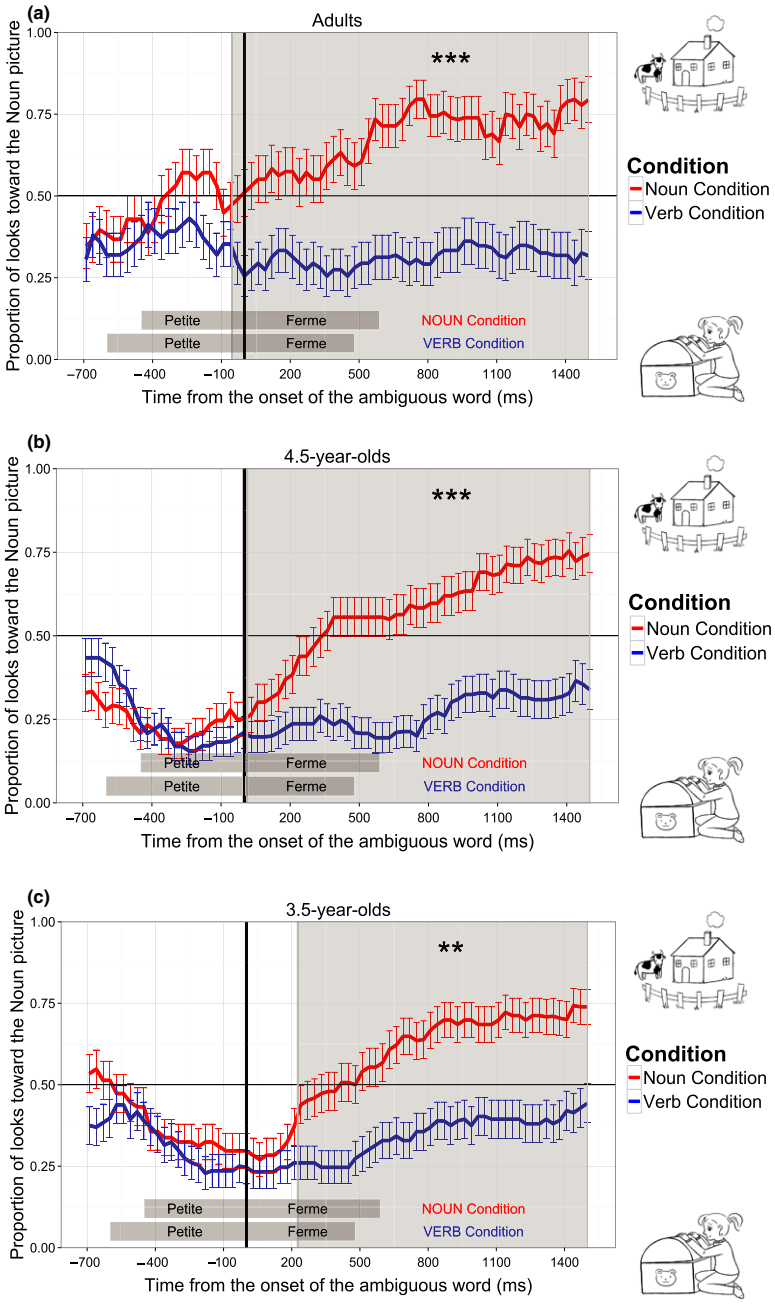


**Figure 7** *Proportion of looks toward the noun image, time-locked to the onset of the ambiguous word (vertical black line) for (a) Adults, (b) 4.5-year-olds, and (c) 3.5-year-olds, for the noun prosody condition (red curve) and the verb prosody condition (blue curve). Error bars represent the standard error of the mean. Participants initially looked more toward the verb image but switched to the noun image in the noun prosody condition. A nonparametric cluster-based permutation test (Maris & Oostenveld, 2007) revealed significant differences between the noun prosody and the verb prosody conditions starting slightly after the onset of the ambiguous target word (dark grey time-window) for all age groups: 3.5-year-olds (from 226ms after the beginning of the critical word, '**'$p < .01$); 4.5-year-olds (from 14ms after the beginning of the critical word, '***' $p < .001$); and Adults (from 54ms before the beginning of the critical word, '***' $p < .001$). Plots of eye-gaze data were created using the ggplot2 package (Wickham, 2009).*

become aware of the ambiguity of test trials, the eight test trials were interleaved with at least six unambiguous filler trials (four during the test block and at least two during the practice block). Such a manipulation should decrease the likelihood that participants would notice the ambiguity.

As a result, we consider it rather unlikely that children in Experiment 2 used prosodic information only because they noticed that the test sentences were ambiguous. Rather, as we discuss next, we propose that children use phrasal prosody to constrain the syntactic analysis of sentences even when they contain non-ambiguous words.

## General discussion

The experiments described in this paper show that by 3.5 years of age, children exploit prosody online to determine the syntactic structure of sentences. In an oral completion task (offline, with 4.5-year-olds) and a preferential looking task (online, with 3.5- and 4.5-year-olds), children were able to correctly assign the grammatical category to an ambiguous word (noun vs. verb) when this ambiguous word was embedded in sentences that began in a phonemically and morphologically identical fashion, but that were syntactically and prosodically distinct. That is, children interpreted the ambiguous target word as a noun when it was embedded in a sentence with a noun prosodic structure and as a verb when it was embedded in a sentence with a verb prosodic structure. Our study is the first to report that children under 4 years of age use phrasal prosody to retrieve the syntactic structure of sentences.

In the introduction we noted that several studies have shown that even older children failed to use prosodic information to interpret ambiguous sentences. At first, this seems at odds with the ease with which children used prosody to guide their interpretation of sentences in our task. One fundamental difference between the present study and previous ones is that in our case the disambiguating prosodic information, namely the phrasal boundary between the noun phrase and the verb phrase, is part of the normal prosodic structure of sentences. Thus, children succeed in the present experiment because the task is so easy to solve: children only need to interpret the prosodic boundary as a syntactic boundary, something that applies to all the sentences they hear daily, whether they contain ambiguous words or not. Because phrasal prosody is found in all languages (e.g. Shattuck-Hufnagel & Turk, 1996), we expect children speaking other languages to succeed equally well, as long as they are presented with sentences for which the default prosodic structure differentiates between the two possible interpretations.

While the detection of prosodic boundaries informed children about the location of syntactic boundaries, prosodic boundaries alone do not directly provide the syntactic label of constituents (e.g. noun phrase, verb phrase). So what enabled children to interpret an ambiguous word as a verb or as a noun depending on the prosodic structure? To derive this interpretation, children likely processed the information carried by function words along with the prosodic information. For example, when participants heard the test sentence [*la petite*] [***ferme*** . . .], the presence of the prosodic boundary before the ambiguous word *ferme* signaled the presence of two prosodic units. The first prosodic unit [***la*** *petite*]$_{NP}$ could furthermore be identified as a noun phrase on the basis of the article. Since the first unit forms a complete noun phrase then children may expect it to be followed by a verb phrase. Thus, upon hearing the beginning of *ferme*, children may expect this word to be a verb or an auxiliary, and quickly identify it to be a verb. In the noun prosody condition, by contrast, the same three words are this time grouped in a single prosodic unit starting with the article ***la*** ([***la*** *petite* ***ferme***] [. . .]), boosting children's interpretation of the constituent as a noun phrase, and of *ferme* as a noun. Prosody would thus be used online to group words into constituents and the function words within the sentence would serve to label them. Using these two sources of information, children could generate a first parse of the sentence, a *syntactic skeleton*, that could help them compute the category of an ambiguous word (Christophe, Millotte, Bernal & Lidz, 2008). Note that children are not bothered by the noun-verb homophony, in this case, because the critical words occur in disambiguating contexts.[5]

In our experiments we used homophones as a test case. However, the ability to generate online predictions regarding the syntactic category of upcoming words would also be very useful to children when perceiving non-ambiguous words, potentially speeding up lexical access. For example, 18-month-old children have been shown to exploit function words to constrain lexical access: They expect a noun after a determiner (Cauvet, Limissuri, Millotte, Skoruppa, Cabrol *et al*., 2014; Kedar, Casasola & Lust, 2006; Van Heugten & Johnson, 2011; Zangl & Fernald, 2007) and a verb after a pronoun (Cauvet *et al*., 2014). For instance, in Cauvet *et al*. (2014), 18-month-olds trained to recognize a target noun ('*la balle*' – 'the ball') were better able to identify it at test when it was preceded by a determiner (a noun context: '*j'aime les* ***balles*** *en mousse*' – I love foam balls) than when it was preceded by a pronoun (a verb context: *'*Pierre, il* ***balle*** *du chocolat*' – *Pierre, he balls some chocolate) and conversely for target verbs. Thus, function words facilitate lexical access to the neighboring content words and constrain online lexical access. Yet, not all content words are immediately preceded by

[5] We suspect that cross-category homophones such as these will most often appear in disambiguating contexts.

function words. In such cases, a more sophisticated analysis in terms of syntactic constituents, signaled by prosodic boundaries, can be very informative and would contribute to fast and efficient lexical access.

We showed that preschoolers are able to compute online predictions regarding the syntactic category of upcoming content words. Importantly, this opens the possibility that such an ability could also be present at a younger age, and may allow toddlers in the process of learning their lexicon to assign a syntactic category to words they have not yet acquired. For example, if a listener expects a noun in a specific position in a sentence, and hears a novel word such as *blick* in that position, she can infer that *blick* is a noun. Adult studies using jabberwocky sentences where all content words are replaced by invented words, while phrasal prosody and function words are preserved (e.g. [*the moop*$_{\mathbf{N}}$] [*blicks*$_{\mathbf{V}}$ *mabily*]), show that adults readily infer that *moop* is a noun, while *blick* is a verb (Millotte, Wales, Dupoux & Christophe, 2006). Thus, even in the absence of knowledge of any of the content words in the previous sentence, it is possible to retrieve a partial syntactic representation based on phrasal prosody and function words (see Gutman, Dautriche, Crabbé & Christophe, in press, for a computational formalization). This might reflect the situation of 18-month-old toddlers, whose knowledge of content words is limited, but who do have access to phrasal prosody (e.g. Gerken *et al.*, 1994) and use function words for syntactic categorization (e.g. Cauvet *et al.*, 2014; Shi & Melançon, 2010).

Having access to the syntactic category of novel words could help toddlers constrain their acquisition of word meanings, since nouns typically refer to objects while verbs typically refer to actions. More generally, the *syntactic bootstrapping hypothesis* (Gleitman, 1990) proposes that the syntactic structure of sentences constrains the possible meaning of words. For instance, faced with *the moop* ***gorps*** *the dax*, listeners readily infer that ***gorp*** is a causal action involving one agent (*the moop*) and one patient (*the dax*; Gillette, Gleitman, Gleitman & Lederer, 1999). Likewise, 2-year-olds infer that novel verbs embedded in transitive sentences have a causative meaning (e.g. Yuan & Fisher, 2009; Naigles, 1990). Thus, having access to a partial syntactic structure based on prosodic structure and function words may help toddlers constrain the possible meanings of verbs.

In summary, we showed that 3.5- to 4.5-year-olds readily use the prosodic structure of an utterance to constrain its syntactic analysis online and access the meaning of an ambiguous word. Children thus use phrasal prosody to segment the continuous speech stream into prosodic units and exploit function words to assign a syntactic function to these units. Because phrasal prosody is available very early during development (within the first year of life), we expect that such an initial parsing mechanism could be active as early as 18 months, during the first steps of syntactic acquisition.

## Acknowledgements

The project reported in this paper was made possible by an MD and a PhD fellowship from the École Normale Supérieure to Alex de Carvalho, and a PhD fellowship from DGA 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). This work would not have been possible without the support of the CNRS, the PSL Research University and the École Normale Supérieure as well. The authors thank Janet Fodor, Franck Ramus, Emmanuel Dupoux, for valuable comments and insights during the conception and construction of the experiments, Sylvain Charron for technical support, Isabelle Lin who drew all the images used in Experiment 2, and the team from the public preschool where we conducted the experiments with children.

## References

Allopenna, P.D., Magnuson, J.S., & Tanenhaus, M.K. (1998). Tracking the time course of spoken word recognition using eye movements : evidence for continuous mapping models. *Journal of Memory and Language*, **439** (38), 419–439.

Barr, D.J., Levy, R., Scheepers, C., & Tily, H.J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language*, **68** (3), 255–278.

Bates, D., & Sakar, D. (2007). lme4: Linear mixed-effects models using S4 classes. *R package version*, **2**, 6.

Cauvet, E., Limissuri, R., Millotte, S., Skoruppa, K., Cabrol, D. *et al.* (2014). Function words constrain on-line recognition of verbs and nouns in French 18-month-olds. *Language Learning and Development*, **10** (1), 1–18.

Choi, Y., & Mazuka, R. (2003). Young children's use of prosody in sentence parsing. *Journal of Psycholinguistic Research*, **32** (2), 197–217.

Choi, Y., & Trueswell, J.C. (2010). Children's (in)ability to recover from garden paths in a verb-final language: evidence for developing control in sentence processing. *Journal of Experimental Child Psychology*, **106** (1), 41–61.

Christophe, A., Millotte, S., Bernal, S., & Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. *Language & Speech*, **51**, 61–75.

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. (1998). Intonation in French. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 195–218). Cambridge: Cambridge University Press.

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

Ferreira, V.S., Slevc, L.R., & Rogers, E.S. (2005). How do speakers avoid ambiguous linguistic expressions? *Cognition*, **96** (3), 263–284.

Fougeron, C., & Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, **101** (6), 3728–3740.

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.

Gervain, J., Nespor, 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.

Gervain, J., & Werker, J.F. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nature Communications*, **4**, 1490.

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

Gutman, A., Dautriche, I., Crabbé, B., & Christophe, A. (in press). Bootstrapping the syntactic bootstrapper: probabilistic labelling of prosodic phrases. *Language Acquisition*.

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

Jaeger, T.F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, **59** (4), 434–446.

Johnson, E.K. (2008). Infants use prosodically conditioned acoustic-phonetic cues to extract words from speech. *Journal of the Acoustical Society of America*, **123** (6), EL144–148.

Jun, S., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus*, **14** (1), 147–172.

Kedar, Y., Casasola, M., & Lust, B. (2006). Getting there faster: 18- and 24-month-old infants' use of function words to determine reference. *Child Development*, **77** (2), 325–338.

Kern, S. (2007). Lexicon development in French-speaking infants. *First Language*, **27** (3), 227–250.

Kern, S., Langue, J., Zesiger, P., & Bovet, F. (2010). Adaptations françaises des versions courtes des inventaires du développement communicatif de MacArthur-Bates. *ANAE*, **22**, 107–108.

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.

Mandel, D.R., Jusczyk, P.W., & Nelson, D.G. (1994). Does sentential prosody help infants organize and remember speech information? *Cognition*, **53** (2), 155–180.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, **164** (1), 177–190.

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

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

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.

Millotte, S., Wales, R., & Christophe, A. (2007). Phrasal prosody disambiguates syntax. *Language and Cognitive Processes*, **22** (6), 898–909.

Millotte, S., Wales, R., Dupoux, E., & Christophe, A. (2006). Can prosodic cues and function words guide syntactic processing and acquisition? In R. Hoffmann & H. Mixdorff (Eds.), *Speech prosody: 3rd International Conference (Studientexte zur Sprachkommunikation ed., Vol. 40)*. Dresden: TUD Press.

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

Morgan, J., & Demuth, K. (1996). Signal to syntax: an overview. In J. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 1–22). Mahwah, NJ: Erlbaum.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, **17** (2), 357–374.

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., & 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.

Rabagliati, H., & Snedeker, J. (2013). The truth about chickens and bats: ambiguity avoidance distinguishes types of polysemy. *Psychological Science*, **24** (7), 1354–1360.

Shattuck-Hufnagel, S., & Turk, A. (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.

Shi, R., & Melançon, A. (2010). Syntactic categorization in French-learning infants. *Infancy*, **15** (5), 517–533.

Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: effects of speaker awareness and referential context. *Journal of Memory and Language*, **48**, 103–130.

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.

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.

Vaissière, J., & Michaud, A. (2006). Prosodic constituents in French: a data-driven approach. In I. Fónagy, Y. Kawaguchi & T. Moriguchi (Eds.), *Prosody and syntax* (pp. 47–64). Amsterdam: John Benjamins.

Van Heugten, M., & Johnson, E.K. (2011). Gender-marked determiners help Dutch learners' word recognition when gender information itself does not. *Journal of Child Language*, **38** (1), 87–100.

Vogel, I., & Raimy, E. (2002). The acquisition of compound vs. phrasal stress: the role of prosodic constituents. *Journal of Child Language*, **29** (2), 225–250.

Weber, A., Grice, M., & Crocker, M.W. (2006). The role of prosody in the interpretation of structural ambiguities: a study of anticipatory eye movements. *Cognition*, **99** (2), B63–B72.

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

Welby, P. (2006). French intonational structure: evidence from tonal alignment. *Journal of Phonetics*, **34** (3), 343–371.

Wickham, H. (2009). *ggplot2*: Elegant graphics for data analysis. In R. Gentleman, K. Hornik & G. Parmigiani (Eds.), *Use R!*. New York: Springer.

Yuan, S., & Fisher, C. (2009). 'Really? She blicked the baby?' Two-Year-olds learn combinatorial facts about verbs by listening. *Psychological Science*, **20** (5), 619–626.

Zangl, R., & Fernald, A. (2007). Increasing flexibility in children's online processing of grammatical and nonce determiners in fluent speech. *Language & Learning Development*, **3**, 199–231.

Received: 22 July 2014
Accepted: 25 January 2015

## Appendix: Experimental sentences

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