

English-speaking preschoolers can use phrasal prosody for syntactic parsing

Alex de Carvalho,^{1,a)} Jeffrey Lidz,² Lyn Tieu,¹ Tonia Bleam,²
and Anne Christophe¹

¹Laboratoire de Sciences Cognitives et Psycholinguistique (ENS, EHESS, CNRS),
Département d'Études Cognitives, École normale supérieure - PSL Research University,
29 rue d'Ulm, 75005 Paris, France

²Department of Linguistics, University of Maryland, College Park, Maryland 20742, USA
x.de.carvalho@gmail.com, lyn.tieu@gmail.com, anne.christophe@ens.fr
jlidz@umd.edu, tbleam@umd.edu

Abstract: This study tested American preschoolers' ability to use phrasal prosody to constrain their syntactic analysis of locally ambiguous sentences containing noun/verb homophones (e.g., [The baby flies] [hide in the shadows] vs [The baby] [flies his kite], brackets indicate prosodic boundaries). The words following the homophone were masked, such that prosodic cues were the only disambiguating information. In an oral completion task, 4- to 5-year-olds successfully exploited the sentence's prosodic structure to assign the appropriate syntactic category to the target word, mirroring previous results in French (but challenging previous English-language results) and providing cross-linguistic evidence for the role of phrasal prosody in children's syntactic analysis.

© 2016 Acoustical Society of America
[DDO]

Date Received: January 29, 2016 Date Accepted: May 3, 2016

1. Introduction

According to the prosodic bootstrapping hypothesis,¹ phrasal prosody (the rhythm and melody of speech) may provide a useful source of information for parsing the speech stream into words and phrases. This hypothesis rests on the observation that across languages, sentences have a prosodic structure (i.e., the nested hierarchy of prosodic units) whose boundaries align with syntactic constituent boundaries.² Salient prosodically conditioned acoustic information (i.e., suprasegmental cues), such as phrase-final lengthening, pitch variations, and pauses, may therefore allow listeners to identify prosodic boundaries, and use this information to identify boundaries between some syntactic constituents. This correspondence between prosodic and syntactic structure should facilitate on-line sentence processing in adults, and may even allow young listeners to identify syntactic constituents before they have acquired an extensive vocabulary.

Previous studies have found that adults indeed rapidly integrate suprasegmental cues to recover the syntactic structure of sentences.³⁻⁶ Developmental studies, however, have found little or no effect of prosody on syntactic ambiguity resolution in English-⁷⁻⁹ and Korean-speaking children.¹⁰ This is surprising, given the extensive literature on infants' ability to perceive boundaries between prosodic constituents from 6 months of age,¹¹ and their use of prosodic boundaries to find word boundaries before their first birthday.^{12,13} This literature would suggest that although young children have early access to phrasal prosody and can exploit it for lexical access, they apparently do not use it to constrain syntactic analysis — contra the prosodic bootstrapping hypothesis.¹

A more recent study, however, demonstrates a strong impact of phrasal prosody on children's syntactic analysis.¹⁴ In this study, French 3- to 6-year-old children were presented with sentences containing local ambiguities arising from the presence of noun/verb homophones. For example, “ferme” is a verb in the sentence: [*la petite*] [*ferme le coffre à jouets*] “[the little girl] [**closes** the toy box],” but is a noun in the sentence: [*la petite ferme*] [*lui plait beaucoup*] “[the little **farm**] [pleases him a lot],” where brackets indicate prosodic units, which reflect the syntactic structure of the sentences. Children presented with the beginning of such ambiguous sentences (e.g., “la petite ferme”) were able to associate the target word with a noun or a verb meaning depending on the prosodic structure in which the critical word was contained.

The question that arises is why these results differ from previous findings from English-speaking children. One possible explanation lies in the syntactic structure used.

^{a)} Author to whom correspondence should be addressed.

In the French study, the default prosodic structure directly reflects the syntactic structure: when the prosodic boundary falls before the critical word, the latter can only be interpreted as a verb; when the prosodic boundary falls after, the word is interpreted as a noun. In contrast, the English experiments used sentences such as “Can you touch the frog with the feather?,” in which the prepositional phrase “with the feather” can be interpreted either as an instrument of the verb “touch” or as a modifier of the noun “frog.”⁶⁻⁸ Crucially, the default prosodic structure is the same for the two readings, i.e., [Can you touch] [the frog] [with the feather]. Speakers who are aware of the ambiguity can intentionally disambiguate by exaggerating the relevant prosodic break, i.e., “[Can you touch the frog] [with the feather]?” for the instrument interpretation, vs “[Can you touch] [the frog with the feather]?” for the modifier interpretation.⁶ Snedeker & Trueswell (2001) found that children failed to use prosody in interpreting such sentences. Subsequent experiments,^{6,8} which controlled for lexical and perseveration biases, again revealed that children presented with both instrument and modifier sentences failed to use prosody to disambiguate (although they succeeded when presented with only one kind of sentence). These authors argued that children’s failure might be due to the fact that the disambiguating prosodic breaks are not part of the normal prosodic structure of these sentences, arising only when the speaker is consciously trying to disambiguate. Children may have difficulties using this kind of prosodic information because they lack experience with such optional prosodic structures.

Alternatively, the discrepancy in the French and English results could arise from differences between the two languages. In English, suprasegmental cues are used to mark word stress, as well as focus (e.g., “JOHN ate the apple”), while French has no word-level stress and uses non-prosodic devices such as fronting to mark focus (e.g., *C’est Jean qui a mangé la pomme* “It is John who ate the apple”). Suprasegmental cues might thus be more ambiguous in English than in French, where they are used mainly to cue phrasal prosodic structure. Thus, although prosodic units are marked by suprasegmental cues in both French and English, they might be easier to perceive in French than in English.

In this paper, we propose to disentangle these two alternative explanations. If American children tested on the same kind of structure as in French fail to disambiguate using suprasegmental information, we will conclude that the discrepancy in previous results is indeed a matter of language specificity. The transparency with which suprasegmental information reflects prosodic structure may vary from one language to the next, such that it can reliably be used to constrain syntactic analysis only in a subset of the world’s languages. In languages where this is not the case, suprasegmental information may be useful for purposes other than identifying syntactic constituency. However, if American preschoolers, like their French counterparts, exploit phrasal prosody to constrain their interpretations of sentences, we will have cross-linguistic evidence of a role for phrasal prosody in syntactic analysis.

2. Experiment

To test whether American preschoolers are able to use phrasal prosody to constrain syntactic analysis, we used homophones belonging to different syntactic categories (noun and verb) to create pairs of sentences containing local syntactic ambiguities (e.g., [The baby flies] [hide in the shadows] vs [The baby] [flies his kite all day long], where brackets indicate prosodic boundaries). Crucially, all of the words following the homophone were acoustically masked with babble noise; only the prosodic structure from the beginnings of the sentences could be used to decide if the target word was a noun or a verb. Preschoolers were given an oral completion task, where they heard the beginnings of these locally ambiguous sentences (e.g., “The baby flies”) and had to complete the sentences however they wanted. Since they had no access to lexical disambiguating information, any difference in responses between the noun and verb sentence beginnings could only be due to suprasegmental differences. If children exploited suprasegmental information to constrain their syntactic analysis, they should give more noun completions after hearing the beginning of noun sentences and more verb completions after hearing the beginning of verb sentences.

3. Method

3.1 Participants

Sixteen 4- to 5-year-old monolingual English-speaking children (4;3 to 5;2, $M_{age} = 4;8$, five boys) were tested in a preschool in the Maryland area or in the Project on Children’s Language Learning Babylab at the University of Maryland. Parents signed an informed consent form. An additional five children were tested but excluded from

analysis because they failed to complete the training sentences prior to the test phase ($n = 3$) or because they were distracted during the experiment ($n = 2$).

3.2 Materials

From eight English noun-verb homophones, eight pairs of experimental sentences were created. Each pair consisted of a sentence in which the ambiguous word was used as a noun (hereafter the noun sentence condition, e.g., [The baby **flies**] [hide in the shadows]) and a sentence in which the ambiguous word was used as a verb (hereafter the verb sentence condition, e.g., [The baby] [**flies** his kite all day long]); see the [Appendix](#) for a complete list of test sentences. Nouns and verbs had similar average log frequencies (1.89 for nouns and 1.77 for verbs, $t(7) < 1$). Utterances in the noun sentence condition contained a phonological phrase boundary after the target word, while utterances in the verb sentence condition had the phrase boundary before the target word. A female native English speaker recorded the sentences in child-directed speech.

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

The duration analysis revealed a significant pre-boundary lengthening, consistent with previous literature^{2,15,16}: the rime of the word just before the phrase boundary in the noun condition (e.g., -ies from “flies”) was lengthened by 72% compared to the same rime in the verb condition (450 vs 262 ms); the rime of the word just before the phrase boundary in the verb condition (e.g., -y from “baby”) was lengthened by 67% compared to the same rime in the noun condition (432 vs 259 ms, see Table 1). Note that these duration differences are well above the just-noticeable difference for segment duration in speech, which is evaluated to be around 15% to 25%.¹⁷

The observation of pitch contours in both prosodic conditions revealed that most often, the subject noun phrase exhibited a low-high-low-high pitch contour (see Fig. 1). In the noun prosody condition, this pattern spread over all of the words that made up the sentence beginning, including the critical ambiguous word (e.g., “the baby flies”). In the verb prosody condition, this pattern was restricted to the first words of the sentence (e.g., “the baby”), while the verb, belonging to the next prosodic phrase, typically exhibited a flat low pitch. To quantify these impressions, we computed the variation in pitch over the rime of the critical word (e.g., -ies in “flies”), and the rime of the preceding word (e.g., -y in “baby”). Consistent with the above-described pattern, the rime of the critical word (e.g., -ies in “flies”) showed a rising pitch pattern in the noun prosody condition when it was phrase-final (+88 Hz), but not in the verb prosody condition when it was phrase-initial (+1 Hz); this difference was significant (see Table 1). The rime of the word preceding the target word (e.g., -y in “baby”) showed a rise in both conditions, corresponding to the phrase-medial rise in the noun prosody condition (+21 Hz), and to the phrase-final rise in the verb prosody condition (+53 Hz); this difference was not significant.

In addition to the target sentences, eight filler sentences were created, containing unambiguous sentence beginnings (e.g., [The baby **mouse**] [eats cheese all the time] or [Mommies] [**like** to have a kiss from their babies]).

In order to make the experiment child-friendly, the speaker was videotaped. Each sentence was cut off at the offset of the target word and its end was replaced by 1200 ms of babble noise, which was obtained by superimposing the ends of all of the filler sentences. The visual stimuli were also masked by having the image tremble and then fade away, starting from the end of the target word. Since the ends of the

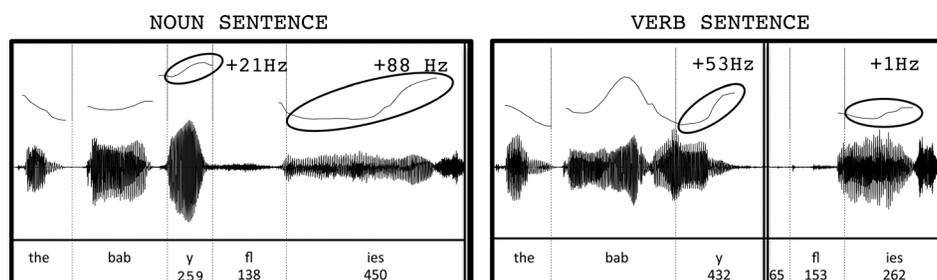


Fig. 1. Mean duration of the different segments, and pitch contours in the ambiguous region. Prosodic boundaries are represented with thick black lines. Ellipses delimit the areas where pitch analyses were performed, subtracting pitch at the beginning of the rime to pitch at the end of the rime, to determine if the pitch contour was rising or falling (also reported in semitones in Table 1). Note that while the waveforms and pitch curves in the figure correspond to the experimental sentences for the target word “flies,” the values for duration and pitch correspond to mean values across all stimuli.

Table 1. Duration and pitch analysis for the stimuli. Mean duration (in ms) and pitch (in Hz) for the segments around the prosodic boundaries for both noun and verb sentence conditions.

Duration analysis—Mean duration in ms (standard error of the mean)				
Dependent variable	Noun sentence	Verb sentence	Difference	<i>t</i> test (2-tailed)
Rime: word preceding Target (e.g., y from “baby”)	259 (22.9)	432 (35.4)	−173 (22.1)	$t(7) = -7.85; p < 0.001^{**}$
Pause: before Target (e.g., between “baby” and “flies”)	0 (0)	65 (18.1)	−65 (18.1)	$t(7) = -3.59; p < 0.01^{**}$
Onset: Target word (e.g., fl from “flies”)	138 (12.2)	153 (13.7)	−15 (10.1)	$t(7) = -1.48; p = 0.18$
Rime: Target word (e.g., ies from “flies”)	450 (23.6)	262 (17.7)	188 (15.3)	$t(7) = 12.32; p < 0.001^{**}$
Pitch analysis—Mean pitch contour in Hz (standard error of the mean) computed as the difference in pitch between the beginning and the end of the rimes around the prosodic boundaries - <i>mean differences in semitones in italics.</i>				
Dependent variable	Noun sentence	Verb sentence	Difference	<i>t</i> test (2-tailed)
Rime: word preceding Target (e.g., y from “baby”)	21 (20.8)	53 (23.3)	−31 (33.3)	$t(7) = -0.94; p = 0.37$
	<i>1.17</i>	<i>4.03</i>	<i>−2.86</i>	
Rime: Target word (e.g., ies from “flies”)	88 (17.4)	1 (10.1)	87 (20.7)	$t(7) = 4.20; p < 0.01^{**}$
	<i>6.56</i>	<i>0.12</i>	<i>6.45</i>	

sentences were both acoustically and visually masked, the only disambiguating information available to participants was prosodic in nature.

The 8 pairs of sentences gave rise to 16 target audiovisual stimuli; 8 in the verb sentence condition and 8 in the noun sentence condition. Each participant saw only one member of each pair. Two counterbalanced lists of stimuli were used, each containing four noun targets, four verb targets, and four unambiguous fillers (two nouns and two verbs). The order of sentences within each list was randomized, with the constraint that there be no more than three target sentences in a row and no more than two consecutive test items from the same syntactic category.

3.3 Procedure

Children sat in front of a computer and listened to the stimuli through headphones. The experiment was presented as a game in which the participants were told that they were “competing” with children from another school. They saw a picture of three children on the screen, which created the illusion that they were communicating by Skype. The child was told that in this game she was going to listen to a woman on a television screen. However, because the television was “broken,” the end of the sentences could not be heard, and she would have to guess what the woman might have said. To motivate children to answer all items, they were told that the child who gave the most completions would win the game.

On each trial, an arrow rotated in the middle of the screen and selected one of the children to complete the sentence. Whenever the arrow pointed downward, it was the participant’s turn to answer. The virtual children on the screen were selected only on unambiguous filler trials, while the participant answered only the target sentences containing the ambiguous noun-verb homophone. When a virtual child was selected to respond, a pre-recorded sentence was played; these “answers” were previously recorded by children of the same age as our participants.

The experiment started with a practice block to familiarize children with the task. In this block, children were presented only with filler sentences (e.g., “The giant castle...”). The first two trials of this block were completed by the virtual children, so as to introduce the participant to the task. From the third trial on, the arrow started selecting our participants and as soon as they correctly completed two filler trials, the test session started.

3.4 Data analysis

We coded children’s responses as noun answers when they gave a completion consistent with the noun interpretation of the target word (e.g., for “the baby flies...,” a

completion such as "...drink milk"), and as verb answers when the completion was consistent with the verb interpretation (e.g., "...away"). Children's responses were coded off-line by two independent coders who each listened to all of the recordings of children's answers, blinded to the condition in which the sentences had been presented. Agreement between coders was 100%. Ten out of the 128 responses were excluded from analysis because the child did not provide a completion compatible with the sentence beginning ($n=6$), e.g., for the sentence "the ladies ring" one child did not use the target word at all and instead said: "The ladies went to a farm," or because the answer was consistent with either interpretation of the target word ($n=4$). For example, the continuation "The little girls paint.... and the little girls wanted to paint" was considered to be ambiguous between the two interpretations, as the child could have interpreted "paint" as a noun or a verb in the first utterance, before using it as a verb in the continuation. We did not take into account the prosody of the child's utterances when coding the answers.

The statistical analysis of children's performance were assessed by analyzing the occurrence of a noun answer (0 or 1) in each condition.¹⁴ We modeled their answers using a logit mixed-effects model.¹⁸ The model included the categorical factor condition (noun \times verb) as well as a random intercept and random slope for condition for both subject and item.¹

4. Results

The average proportions of noun and verb² answers for each condition are presented in Fig. 2. Children gave more noun answers in the noun sentence condition than in the verb sentence condition. This was reflected in our mixed model analysis by a main effect of condition ($\beta = 3.91$; $z = 2.88$; $p < 0.01$), corresponding to an increase of 0.63 in the probability of giving a noun response in the noun condition relative to the verb condition.

5. Discussion

In this experiment, English-speaking 4.5-year-olds were able to assign different syntactic categories to an ambiguous word, depending only on the word's position within the prosodic structure of the sentence. In an oral completion task, upon hearing the beginning of locally ambiguous sentences like: "the baby flies," preschoolers gave more noun completions in the noun sentence condition than in the verb condition. Given that the two sentence beginnings differed only in prosodic structure, this shows children were able to exploit phrasal prosody to constrain their syntactic analysis, correctly assigning syntactic categories to the ambiguous words. The results mirror the strong prosodic effect obtained with French preschoolers¹⁴ and adults,⁵ and confirm that American preschoolers can use phrasal prosody to constrain their syntactic analysis.

The previously reported discrepancies between English and French are thus not due to specific properties of these languages, but rather to a difference in the syntactic structures that were tested, specifically the reliability with which the prosodic structure reflected the syntactic structure. The English sentences used in previous studies (e.g., [can you touch] [the frog] [with the feather]) were such that the two readings shared the same default prosodic structure.⁸ In contrast, our sentence beginnings had

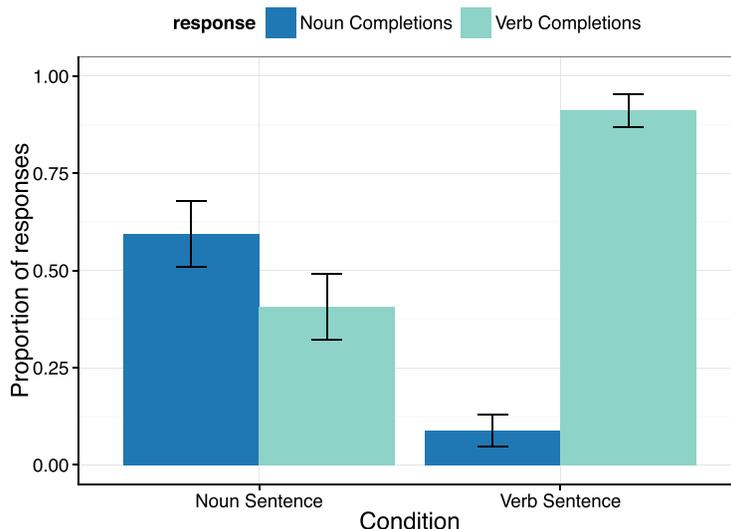


Fig. 2. (Color online) Proportion of noun and verb completions for each condition. Error bars represent the standard error of the mean.

different default prosodic structures, with the prosodic boundary falling either before or after the critical word. Because the prosodic boundary between the subject noun phrase and the verb phrase is present in many sentences that children hear everyday, including unambiguous sentences (e.g., [The little boy] [runs really fast]), children can rely almost systematically on the phrasal prosody to recover aspects of the syntactic structure. This may explain our participants' remarkable ability to integrate prosodic information in their computation of syntactic structure.

Additionally, this ability to exploit suprasegmental information for syntactic purposes may be extremely important in the early stages of language acquisition, particularly when children do not yet know the meanings of many words. Having access to information that signals syntactic constituent boundaries may help children to identify parts of the syntactic structure of a sentence in which a novel word appears, and use it to constrain its possible meanings.¹⁹ For example, in a sentence like “[Do you see the baby blicks]?”, children might be able to infer that “blick” is a noun, referring to a kind of object; but in a sentence like: “[Do you see]? [The baby] [blicks]!” they may infer that “blick” is a verb, referring to some action in their environment. Very recent studies in French suggest that such a mechanism for language acquisition is plausible: 2-year-olds were shown to exploit suprasegmental information from phrasal prosody to correctly identify noun-verb homophones,²⁰ and 18-month-olds were shown to use this suprasegmental information to interpret novel words as either nouns or verbs, depending on their position within the prosodic-syntactic structure of the sentence.^{21,22}

These recent findings in French, along with our current results in English, lend support to the hypothesis that phrasal prosody cues syntactic structure in early language development, and likely in different languages. Previous difficulties detecting this connection were likely due to the fact that the link between prosodic and syntactic structure was not sufficiently systematic in the structures that were tested. In cases where this relationship is more systematically marked, we observe that children are just as sensitive to prosody as one might expect. These results lend support to the hypothesis that phrasal prosody is an important cue to syntactic structure during language acquisition.

Acknowledgments

This research was supported by a Ph.D. fellowship from the École normale supérieure to A. de Carvalho. It was also supported by grants from the Région Ile-de-France, Fondation de France, LabEx IEC (ANR-10-LABX-0087), IdEx Paris Sciences et Lettres (PSL) (ANR-10-IDEX-0001-02), as well as the ANR “Apprentissages” (ANR-13-APPR-0012). The authors thank Juliana Gerard for her help in preschool testing, Page Piccinini, Tara Mease, Mina Hirzel, Rosa Capetta, Tim Dawson, and the teams from the preschools.

APPENDIX

The experimental sentences are given in Table 2.

Table 2. Experimental sentences.

Test sentences used in the experiment			
Pair of ambiguous words	Syntactic category	Target	Full sentence recorded
A fly × to fly	Noun	flies	The baby flies hide in the shadows
	Verb		The baby flies his kite all day long
A plant × to plant	Noun	plant	The nice kid's plant fell down in the garden
	Verb		The nice kids plant flowers in the garden
A watch × to watch	Noun	watch	Mommy's watch ticks very noisily
	Verb		Mommies watch TV every night
A ring × to ring	Noun	ring	The lady's ring had to be repaired
	Verb		The ladies ring her doorbell every night
Water × to water	Noun	water	The boy's water dripped on the floor
	Verb		The boys water the plants every day
A hand × to hand	Noun	hand	The little girl's hand has a ring on the third finger
	Verb		The little girls hand heavy books to their teacher
Paint × to paint	Noun	paint	The little girl's paint got spilled on the floor
	Verb		The little girls paint go-karts at the track
A swing × to swing	Noun	swing	The little kid's swing fell down in the park
	Verb		The little kids swing frequently at the park

References and links

- ¹The full model was thus as follows: $\text{Logit}(P(R_{is} = 1)) = \beta_0 + S_{0s} + I_{0i} + (\beta_1 + S_{1s} + I_{1i})C_i + e_{is}$ where e_{is} , represents the normally distributed error for the observation, β_0 is the intercept, S_{0s} the intercepts by subjects, I_{0i} the intercepts by items, β_1 the slope for the condition, S_{1s} the slopes by subjects, and I_{1i} the slopes by items. β estimates are given in log-odds.
- ²The pattern of responses shows an overall bias in favor of verb responses. Note that de Carvalho *et al.* (2016) observed an overall bias in favor of noun responses, which disqualifies any sort of general interpretation (e.g., that the prosody of verb sentences is more informative than the prosody of noun sentences). We checked that there was no frequency difference between nouns and verbs using the CELEX database (see Sec. 3.2), and did not find any correlation between the verb-noun frequency difference for each pair of items and the difference in proportion of verb answers in the experiment ($R = 0.16$; $t(6) < 1$). We may wonder whether verb sentences might have been easier to complete than noun sentences in this experiment because noun sentences often contained a genitive (which may have been harder to process for children), but this is purely speculative. Irrespective of the explanation for the verb bias, though, the significant difference between conditions shows that the suprasegmental differences between sentences were exploited by children to constrain their syntactic analysis.
- ¹J. Morgan, *From Simple Input to Complex Grammar* (MIT Press, Cambridge, MA, 1986).
- ²M. Nespors and I. Vogel, *Prosodic Phonology* (Foris, Dordrecht, the Netherlands, 1986).
- ³M. M. Kjelgaard and S. R. Speer, "Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity," *J. Mem. Lang.* **40**, 153–194 (1999).
- ⁴S. Millotte, R. Wales, and A. Christophe, "Phrasal prosody disambiguates syntax," *Lang. Cogn. Process.* **22**, 898–909 (2007).
- ⁵S. Millotte, A. René, R. Wales, and A. Christophe, "Phonological phrase boundaries constrain the online syntactic analysis of spoken sentences," *J. Exp. Psychol. Learn. Mem. Cogn.* **34**, 874–885 (2008).
- ⁶J. Snedeker and J. Trueswell, "Using prosody to avoid ambiguity: Effects of speaker awareness and referential context," *J. Mem. Lang.* **48**, 103–130 (2003).
- ⁷J. Snedeker and J. Trueswell, "Unheeded cues: Prosody and syntactic ambiguity in mother-child communication," Paper presented at the 26th Boston University Conference on Language Development, Boston, MA (2001).
- ⁸J. Snedeker and S. Yuan, "Effects of prosodic and lexical constraints on parsing in young children (and adults)," *J. Mem. Lang.* **58**, 574–608 (2008).
- ⁹I. Vogel and E. Raimy, "The acquisition of compound vs. phrasal stress: The role of prosodic constituents," *J. Child Lang.* **29**, 225–250 (2002).
- ¹⁰Y. Choi and R. Mazuka, "Young children's use of prosody in sentence parsing," *J. Psycholinguist. Res.* **32**, 197–217 (2003).
- ¹¹M. Soderstrom, A. Seidl, D. G. Kemler Nelson, and P. W. Jusczyk, "The prosodic bootstrapping of phrases: Evidence from prelinguistic infants," *J. Mem. Lang.* **49**, 249–267 (2003).
- ¹²A. Gout, A. Christophe, and J. L. Morgan, "Phonological phrase boundaries constrain lexical access II. Infant data," *J. Mem. Lang.* **51**, 548–567 (2004).
- ¹³E. K. Johnson, "Infants use prosodically conditioned acoustic-phonetic cues to extract words from speech," *J. Acoust. Soc. Am.* **123**, EL144–EL148 (2008).
- ¹⁴A. de Carvalho, I. Dautriche, and A. Christophe, "Preschoolers use phrasal prosody online to constrain syntactic analysis," *Dev. Sci.* **19**, 235–250 (2016).
- ¹⁵W. E. Cooper and J. Paccia-Cooper, *Syntax and Speech* (Harvard University Press, Boston, MA, 1980).
- ¹⁶M. Soderstrom, M. Blossom, R. Foygel, and J. Morgan, "Acoustical cues and grammatical units in speech to two preverbal infants," *J. Child Lang.* **35**, 869–902 (2008).
- ¹⁷A. W. F. Huggins, "Just noticeable differences for segment duration in natural speech," *J. Acoust. Soc. Am.* **51**, 1270–1278 (1972).
- ¹⁸T. F. Jaeger, "Categorical data analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models," *J. Mem. Lang.* **59**, 434–446 (2008).
- ¹⁹L. Gleitman, "The structural sources of verb meanings," *Lang. Acquis.* **1**, 3–55 (1990).
- ²⁰A. de Carvalho, I. Dautriche, and A. Christophe, "Phrasal prosody constrains online syntactic analysis in two-year-old children," Paper presented at the 39th Boston University Conference on Language Development, Boston, MA (2014).
- ²¹S. Massicotte-Laforge and R. Shi, "The role of prosody in infants' early syntactic analysis and grammatical categorization," *J. Acoust. Soc. Am.* **138**, 441–446 (2015).
- ²²A. de Carvalho, A. He, J. Lidz, and A. Christophe, "18-month-olds use the relationship between prosodic and syntactic structures to constrain the meaning of novel words," Paper presented at the 40th Boston University Conference on Language Development, Boston, MA (2015).