The role of phrasal prosody and function words in the acquisition of word meanings

Alex Lopa de Carvalho

THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres
PSL Research University

Préparée à l'École Normale Supérieure

The role of phrasal prosody and function words in the acquisition of word meanings

Le rôle de la prosodie phrasale et des mots grammaticaux dans l'acquisition du sens des mots

Ecole doctorale n°158

ED3C – École Doctorale Cerveau Cognition Comportement

Spécialité SCIENCES COGNITIVES

COMPOSITION DU JURY :

Mme SEBASTIAN-GALLES Nuria
Universitat Pompeu Fabra, Rapporteur

Mme D'IMPERIO Mariapaola
Aix-Marseille Université, Rapporteur

M. RIZZI Luigi
Université de Genève, Examinateur
Membre du jury

M. NAZZI Thierry
Université Paris Descartes, Président du jury

Mme CHRISTOPHE Anne
École Normale Supérieure,
Directrice de thèse

Soutenue par Alex Sander
LOPA DE CARVALHO
le 15 septembre 2017

Dirigée par **Anne CHRISTOPHE**

ECOLE NORMALE SUPERIEURE
UNIVERSITÉ DE RECHERCHE PARIS SCIENCES ET LETTRES
PSL RESEARCH UNIVERSITY
ED3C – ECOLE DOCTORALE CERVEAU COGNITION COMPORTEMENT
DEPARTEMENT D'ETUDES COGNITIVES
Laboratoire de Sciences Cognitives et Psycholinguistique
UMR 8554 CNRS-ENS-EHESS

The role of phrasal prosody and function words in the acquisition of word meanings

Par Alex Sander LOPA DE CARVALHO

Thèse de doctorat de Sciences Cognitives

Dirigée par Anne CHRISTOPHE

Présentée et soutenue publiquement le 15 septembre 2017

Devant un jury composé de:

Anne CHRISTOPHE, directrice de la thèse	- Directrice de Recherche, CNRS, École normale supérieure
Núria SEBASTIAN-GALLES, rapporteur	- Professeur, Universitat Pompeu Fabra, Barcelona
Mariapaola D'IMPERIO, rapporteur	- Professeur, Aix-Marseille Université
Luigi RIZZI, examinateur	- Professeur, Université de Genève, Suisse
Thierry NAZZI, président du Jury	- Directeur de Recherche, CNRS, Université Paris Descartes

Paris
1er Semestre
Année Universitaire 2017-2018

Abstract

Previous research demonstrates that having access to the syntactic structure of sentences helps children to discover the meaning of novel words. This implies that infants need to get access to aspects of syntactic structure before they know many words. Since in all the world's languages the prosodic structure of a sentence correlates with its syntactic structure, and since function words/morphemes are useful to determine the syntactic category of words, infants might use phrasal prosody and function words to bootstrap their way into lexical and syntactic acquisition. In this thesis, I empirically investigated the role of phrasal prosody and function words to constrain syntactic analysis in young children (PART 1) and whether infants exploit this information to learn the meanings of novel words (PART 2).

In part 1, I constructed minimal pairs of sentences in French and in English, testing whether children exploit the relationship between syntactic and prosodic structures to drive their interpretation of noun-verb homophones. I demonstrated that preschoolers use phrasal prosody online to constrain their syntactic analysis. When listening to French sentences such as [*La petite* ***ferme***][… – [The little **farm**][…, children interpreted *ferme* as a noun, but in sentences such as [*La petite*][***ferme***…] – [The little girl][**closes**…, they interpreted *ferme* as a verb (Chapter 3). This ability was also attested in English-learning preschoolers who listened to sentences such as "The baby flies…": they used prosodic information to decide whether "flies" was a noun or a verb (Chapter 4). Importantly, in further studies I demonstrated that even infants around 20-months use phrasal prosody to recover syntactic structures and to predict the syntactic category of upcoming words (Chapter 5), an ability which would be extremely useful to discover the meaning of unknown words.

This is what I tested in part 2: whether the syntactic information obtained from phrasal prosody and function words could allow infants to constrain their acquisition of word meanings. A first series of studies relied on right-dislocated sentences containing a novel verb in French: [*il*$_i$ *dase*], [*le bébé*$_i$] - 'he$_i$ is *dasing*, the baby$_i$' (meaning 'the baby is dasing') which is minimally different from the transitive sentence [*il dase le bébé*] – [he is *dasing* the baby]. 28-month-olds were shown to exploit prosodic information to constrain their interpretation of the novel verb meaning (Chapter 6). In a second series of studies, I investigated whether phrasal prosody and function words constrain the acquisition of nouns and verbs. I used sentences like "*Regarde la petite bamoule*", which can be produced either as [*Regarde la petite bamoule!*] - Look at the little

bamoule!, where "*bamoule*" is a noun, or as [*Regarde*], [*la petite*] [*bamoule!*] - Look, the little (one) is bamoul*ing*, where *bamoule* is a verb. 18-month-olds correctly parsed such sentences and attributed a noun or verb meaning to the critical word depending on its position within the syntactic-prosodic structure of the sentences (Chapter 7).

Taken together, these studies show that infants exploit function words and the prosodic structure of an utterance to recover the sentences' syntactic structure, which in turn constrains the possible meaning of novel words. This powerful mechanism might be extremely useful for infants to construct a first-pass syntactic structure of spoken sentences even before they know the meanings of many words. Although prosodic information and functional elements can surface differently across languages, our studies suggest that this information may represent a universal and extremely useful tool for infants to access syntactic information through a surface analysis of the speech stream, and to bootstrap their way into language acquisition.

Résumé

Des études précédentes démontrent qu'avoir accès à la structure syntaxique des phrases aide les enfants à découvrir le sens des mots nouveaux. Cela implique que les enfants doivent avoir accès à certains aspects de la structure syntaxique avant même de connaître beaucoup de mots. Étant donné que dans toutes les langues du monde la structure prosodique d'une phrase corrèle avec sa structure syntaxique, et que par ailleurs les mots et morphèmes grammaticaux sont utiles pour déterminer la catégorie syntaxique des mots, il se pourrait que les enfants utilisent la prosodie et les mots grammaticaux pour initialiser leur acquisition lexicale et syntaxique. Dans cette thèse, j'ai étudié le rôle de la prosodie phrasale et des mots grammaticaux pour guider l'analyse syntaxique chez les enfants (PARTIE 1) et la possibilité que les jeunes enfants exploitent cette information pour apprendre le sens des mots nouveaux (PARTIE 2).

Dans la partie 1, j'ai construit des paires minimales de phrases en français et en anglais afin de tester si les enfants exploitent la relation entre les structures prosodique et syntaxique pour guider leur interprétation des homophones noms-verbes. J'ai démontré que les enfants entre 3 et 5 ans utilisent la prosodie phrasale en temps réel pour guider leur analyse syntaxique. En écoutant des phrases telles que [La petite ferme][.., les enfants interprètent *ferme* comme un nom, mais pour les phrases telles que [La petite][ferme...], ils interprètent *ferme* comme un verbe (Chapitre 3). Cette capacité a également été observée chez les enfants américains: en écoutant des phrases telles que «*The baby flies…* », ils utilisent la prosodie des phrases pour décider si *flies* est un nom ou un verbe (Chapitre 4). Par la suite, j'ai démontré que même les enfants d'environ 20 mois utilisent la prosodie des phrases pour récupérer leur structure syntaxique et pour en déduire la catégorie syntaxique des mots (Chapitre 5), une capacité qui serait extrêmement utile pour découvrir le sens des mots inconnus.

C'est cette hypothèse que j'ai testé dans la partie 2, à savoir si l'information syntaxique obtenue à partir de la prosodie phrasale et des mots grammaticaux permet aux enfants d'apprendre le sens des mots. Une première série d'études s'appuie sur des phrases disloquées à droite contenant un verbe nouveau en français: [il$_i$ dase], [le bébé$_i$] qui est minimalement différente de la phrase transitive [il dase le bébé]. Mes résultats montrent que les enfants de 28 mois exploitent les informations prosodiques de ces phrases pour contraindre leur interprétation du sens du nouveau verbe (Chapitre 6). Dans une deuxième série d'études, j'ai étudié si la prosodie et les mots grammaticaux guident l'acquisition de noms et de verbes. J'ai utilisé des phrases comme «Regarde la petite bamoule» qui peuvent être produites soit comme [Regarde la petite bamoule!], où

«bamoule» est un nom, ou [Regarde], [la petite] [bamoule!], où bamoule est un verbe. Les enfants de 18 mois ont correctement analysé ces phrases et ont attribué une interprétation de nom ou de verbe au mot bamoule selon sa position dans la structure prosodique-syntaxique des phrases (Chapitre 7).

Ensemble, ces études montrent que les jeunes enfants exploitent les mots grammaticaux et la structure prosodique des phrases pour inférer la structure syntaxique et contraindre ainsi l'interprétation possible du sens des mots. Ce mécanisme peut permettre aux enfants de construire une représentation initiale de la structure syntaxique des phrases, avant même de connaître la signification des mots. Bien que les informations prosodiques et les mots grammaticaux puissent prendre des formes différentes selon les langues, nos études suggèrent que cette information pourrait représenter un outil universel et qui permettrait aux enfants d'accéder à certaines informations syntaxiques des phrases qu'ils entendent, et d'initialiser l'acquisition du langage.

Dedication

To my grandmother, Anna, who from the earliest years of my infancy taught me the true meaning of Love and continues to be an example that it has no limits.

À ma chère Mamie, Anna, celle qui, depuis mes premières années de vie m'a appris le sens du mot Amour et qui continue à être l'exemple que celui-ci n'a pas de limites.

À minha querida avó Anna que desde os meus primeiros anos de vida me ensinou o verdadeiro significado da palavra Amor e continua a ser um exemplo de que ele nao tem limites.

Acknowledgments

They say it takes a village to raise a child, and the same goes for a thesis. I have to say that all the work accomplished during this thesis would not have been possible without the contribution of numerous individuals who have been on my side throughout all these years. As such, I want to take the time to acknowledge and thank all those whom I have met along this road such as: my colleagues, my professors, my collaborators, my friends, my family, all the children who participated in my studies and of course their caregivers, the preschools who so kindly welcomed our team each year, and so many other individuals who contributed directly or indirectly for many aspects of this work.

I am afraid that these words will never fully have the power to express the degree of my gratitude for you all. Nonetheless, I want to say once more how grateful I am for all those whom have give me aid or lent a kind ear throughout this process of research and discovery. Despite my global sense of gratitude for you all, one must all the same thank a certain number of individuals personally. I would like to underline that this does not mean that I appreciate any less those whom I did not name personally, but to the contrary I appreciate to the fullest of my ability all those who have been lights to me in this storm of research.

Firstly, I would like to thank all the researchers who accepted to be part of this thesis committee: **Nuria Sebastian-Galles**, **Mariapaola D'Imperio**, **Luigi Rizzi** and **Thierry Nazzi**. I am truly honored by the time and effort you have all taken to read and give me feedback in regards to this thesis. I am looking forward to having you all here in Paris for my defense and thus have the opportunity to discuss and review in person all your thoughts and commentary on my work.

I would also equally like to thank the members of my thesis advisory board: **Judit Gervain** and **John Trueswell**. Thank you so much for all your feedbacks on my work during the last three years, and most importantly thanks for the support that you have given me all through this process.

Finally, the thanks of all thanks has to go to an amazing person whom has been an inextricable part to this whole process: **Anne Christophe**. Without you, I can easily say that I would neither be who or where I am today. You have been a guiding light through the thick and thin of it all! You are an example to be followed, and I am so truly grateful to have had the opportunity to work and get to know you as my thesis advisor. Anne, I learned so much with you!! I am completely aware of how lucky I am to have had the honour to be your student. I am sure these words will not be enough to express all what I feel, but I would like to let you know that I am truly thankful for everything you did for me. During one of the talks that I gave, someone said "You are so productive! How does Anne feed you?". I can easily answer this question! Tu me nourris moi et tous ceux qui t'entourent avec ta passion pour ce que tu fais, ta joie de vivre, ta bonne humeur, ton optimisme, tes réponses à toutes nos questions, ton désir le plus sincère que tout se passe toujours bien, ton soutien inconditionnel qui nous permet d'oser, de nous lancer dans nos hypothèses, et surtout, tu nous donnes de la confiance. Merci pour tout**,** Anne!

Je tiens également à remercier celle qui a été beaucoup plus qu'une collaboratrice dans mes études, mais aussi ma grande sœur dans la recherche: **Isabelle Dautriche**. Isa, tu as été là dès le début! Tu m'as même aidé à me préparer pour le concours d'entrée à l'ENS, avant que je ne commence le Cogmaster ! Je ne serais sans doute pas là aujourd'hui si je n'avais pas pu compter avec ton aide inconditionnelle, ton soutien, et ta patience pour m'apprendre des fondamentaux comme les statistiques et la programmation qui peuvent être très compliquées pour quelqu'un qui vient d'une filière complètement littéraire comme moi. Tu as été un exemple à suivre ! Je t'admire beaucoup et tes encouragements et ta détermination m'ont fait croire que « c'est possible ! » et que l'on doit toujours foncer pour attendre nos objectifs. Je te suis très reconnaissant pour tout ce que tu as fait et ce que tu continues à faire pour m'aider.

I owe a very personal thanks to **Jeffrey Lidz** who hosted me for a month and a half in his lab to allow me to conduct the experiment with English-speaking preschoolers. It is thanks to this brief but intense exchange that I was able to produce two of the studies presented in this thesis. Jeff, my time with you and all the amazing team of the linguistic department of Maryland allowed me to learn a lot and inundated me with your joie de vivre and most importantly with all your energizing discussions about linguistics! I will never forget this experience! I also need to thank **Colin Phillips**, **Juliana Gerard**, **Tonia Bleam**, **Tara Mease**, **Mina Hirzel**, and **Tim Dawson** who significantly contributed to the great time that I had in Maryland and who helped me a lot with different aspects of the research that I conducted there.

Many thanks to all my old and new collaborators: **Isabelle Dautriche**, **Anne Reboul, Angela He**, **Lyn Tieu**, **Sandra Waxman**, **Jeffrey Lidz**, **Tatjana Nazir**, **Jean-Baptiste Van der Henst**, **Anne Chelyus**, **Isabelle Lin**, **Naomi Havron**, **Mireille Babineau,** with whom I had/have the great opportunity to work. It is and it has been a pleasure, and you all have taught me a lot. Thank you!

I also need to thank the group of indivuals whom willingly agreed to be my guinea pigs for the day during the filming of so many stimuli videos that I used in part of the studies conducted during this thesis: **Adriana Guevara Rukoz**, **Gerda Ana Melnik**, **Alexander Martin**, **Lucas Tiphine**, **Marie Palu**, **Gwendal Kerdavid**, **Laia Fibla**, **Lorna Le Stanc**, **Margaux Romman-Monnier**, **Cécile Crimon**, **Christina Bergmann**, **Benjamin Rebouillat**, **Aurélien Jutel**,

Maddalena Napolitani, **Julia Carbajal**, **Matthieu Koroma**, **Adeline Lucchesi**, **Gabriel Synnaeve**, **Bogdan Ludusan**, **Perrine Brusini,** and many others that were not kept in the final sample because of fussiness! I am afraid the festival at Cannes did not grant you all the Césars which you all so merited. However, I want you all to know the videos of you will be watched by many children to come and you are all my stars.

I would additionally like to thank the following people, for their support in various ways before and during my PhD:

Anne Reboul, je pense que sans ton soutien, je ne serais pas arrivé jusqu'à là non plus. Merci de m'avoir encouragé à partir, et de m'avoir soutenu dans mon souhait d'étudier l'acquisition du langage. Je me souviens toujours de ta phrase « on ne vit qu'une vie ! ».

Anne-Caroline Fiévet, je n'aurais sans doute pas testé autant d'enfants dans mes expériences sans ton aide et sans le travaille magnifique que tu développes en tant que Babylab manager. Le succès des études avec les bébés au LSCP est en grande partie dû à toi et l'efficacité de ton travail! Merci pour tout ton soutien, pour tous les samedis pendant lesquels nous avons testé ensemble et pour toute ton aide décisive pour réaliser mes études dans les écoles maternelles. T'as été une de mes plus grandes alliées pendant toutes ces années !

Radhia Achheb, Vireack Ul, Michel Dutat, Isabelle Brunet, Clémentine Eyraud, Luca Filippin, Auguste Filippi, Isabelle Mistral, Amine Dali, Salou Kanoute, Ina Valintelyte, vous êtes nos anges gardiens, et je suis très reconnaissant de tous vos efforts et de tout ce que vous avez fait pour moi et mes recherches pendant toutes ces années. Merci !

Alejandrina Cristia, tu as été là dès le début aussi ! Tu as contribué significativement à différents aspects de ma formation et de mon développement professionnel. Merci pour toutes nos discussions, pour tes conseilles et ce que tu m'as appris sur la recherche, mais aussi sur la vie ! Ton feedback attentif et toujours précis a beaucoup apporté aux études que j'ai pu réaliser pendant cette thèse. Je te remercie!

Sharon Peperkamp, Emmanuel Chemla, **Franck Ramus**, **Emmanuel Dupoux**, **Jérôme Sackur** et **Sid Kouider**, merci de tout votre soutien, votre aide, et de tous vos conseils pendant toutes ces années !

Alexandre Martin, merci beaucoup d'avoir été littéralement à mes côtés pendant toutes ces années. Nos échanges pendant le master et ensuite pendant la thèse m'ont beaucoup apporté ! On s'est toujours aidé et soutenu dans les bons et « moins bons » moments et je crains que tu me manqueras beaucoup lors que je ne te verrai plus à côté de moi au bureau.

Axel Barrault, avoir eu l'occasion d'encadrer tes stages m'a beaucoup apporté et je te remercie également de ta contribution pour tester avec moi tous les enfants de notre étude sur la négation.

Benjamin Spector, merci pour tout ce que tu fais afin de t'assurer qu'on puisse avoir les meilleures conditions pour apprendre, pour enseigner, pour se développer professionnellement, pour développer nos études et surtout pour concrétiser nos rêves.

Aurélien Jutel, **Chase Crook**, **Sonia Leyglene**, **Marcela Lopez**, **Thiago Queiroz**, **Isabelle Lecompte, Claire Maurin, Charlotte Van Den Driessche, Christophe Marco, Sylvain Guyonnet, Maddalena Napolitani**, **Marco Doudin**, **Emelyn Lih**, **Aline Waltzing**, **Igor Rolemberg**, **Eckhard Wallis**, **Anastasia de La Fage**, **Julien Calvani**, **Ruben Trevino**, **Mirela Botaro**, **Philipp J**, **Melody Courson**, **Emmanuel Ahr**, **Robin Millman**, **Claire Trebitsch**, **Guy et Lucette Bordron** (**Papi et Mamie**) vous êtes plus que mes amis, vous êtes ma famille en France et sans votre soutien, votre amitié, vos mots d'encouragement, je ne serais sans doute pas arrivé jusqu'à là. Je vous aime et je vous remercie de tout mon cœur.

Lucas Tiphine, merci beaucoup pour tout ton soutien pendant ces années. Tu as sans doute été mon coach #1 ! Je sais à quel point je suis chanceux d'avoir une amitié comme la tienne. Tu as été toujours là pour moi dans les évènements les plus importants, les plus décisifs, et si j'ai réussi, c'est sans doute grâce à ton soutien aussi.

Laurent Lemercier, **Olivier Granit**, **Martine Fragoas**, **Isabelle Hesling,** merci pour tout ce que vous avez fait pour moi lors de mon année à Bordeaux. Vous avez été à la base de tous ces projets et cette réussite je tiens également à la partager avec vous.

Aniela França, my very first research advisor! você foi à primeira orientadora que eu tive! Foi você quem me apresentou o fantástico mundo da pesquisa em aquisição da linguagem e sem você eu também não teria chegado até aqui. Obrigado por ter me mostrado o caminho à seguir e ter me apoiado à partir.

Thaïs Cristófaro Silva, **Celina Mello**, **Marcus Maia**, **Aleria Lage**, **Thiago Motta, Juliana Gomes, Daniela Cid, Isabella Sampaio, Leonardo Marcotulio, Simone Betoni,** obrigado por todos os conselhos, todas as palavras de incentivo e todo o apoio que vocês me forneceram em diversas etapas dessa longa trajetória!

Luiz Carlos Balga, obrigado pelo seu apoio e incentivo nos momentos mais importantes. Obrigado pela oportunidade de ter trabalhado com você e obrigado por tudo o que você me ensinou.

Pedro Paulo Catharina, muito obrigado por tudo o que você fez para que eu pudesse chegar até aqui.

Eloisa Lima, obrigado pelo seu apoio sempre incondicional! Ter tido a oportunidade de trabalhar com você durante o tempo que eu ainda cursava a Universidade foi sem duvida decisivo para que eu pudesse chegar até aqui hoje. Obrigado por tudo o que você me permitiu aprender junto com você e por tudo que você e toda a equipe do DICE me ensinaram sobre crianças.

Katia Teonia, Magna Magistra, você foi e é uma fonte de inspiração como professora, e como amiga você se tornou um alicerce muito importante da minha estrutura, da minha base. Obrigado por estar ao meu lado durante todos esses anos. Obrigado pelo seu apoio incondicional, por se manter presente mesmo quando temos um oceano entre nós. Seu carinho e seu apoio foram essenciais nessa trajetória.

A minha **família** e em especial**, minha avó e mãe Anna, meu pai, tia Sylvia, minha mãe Sandra, Flavinha, minha irmã Mariana, tio Germano**, muito obrigado por tudo! Cada um de vocês sabe o quanto contribuiu para que eu chegasse até aqui.

Merci!

Contents

Part 2

Chapter 1: Introduction

Human language acquisition intrigues anyone who has already observed a child uttering her first sentences. How is it possible that between the age of two and three years, toddlers have yet to master the use of a fork, but they already know how to utter grammatical sentences in their native language? This ability is all the more impressive given that they start from scratch, and have to guess the meaning of the words that are used around them. Discovering the words composing their language is one of the most challenging tasks faced by language learners. This ability requires children to find a way to map the sounds of each word they hear to a possible meaning in their environment. However, given that in fluent speech there are no pauses between two consecutive words and given that for each spoken sentence the world offers a wide array of possible referential intentions (e.g., Quine, 1960), how do babies manage to achieve this sound-to-meaning mapping?

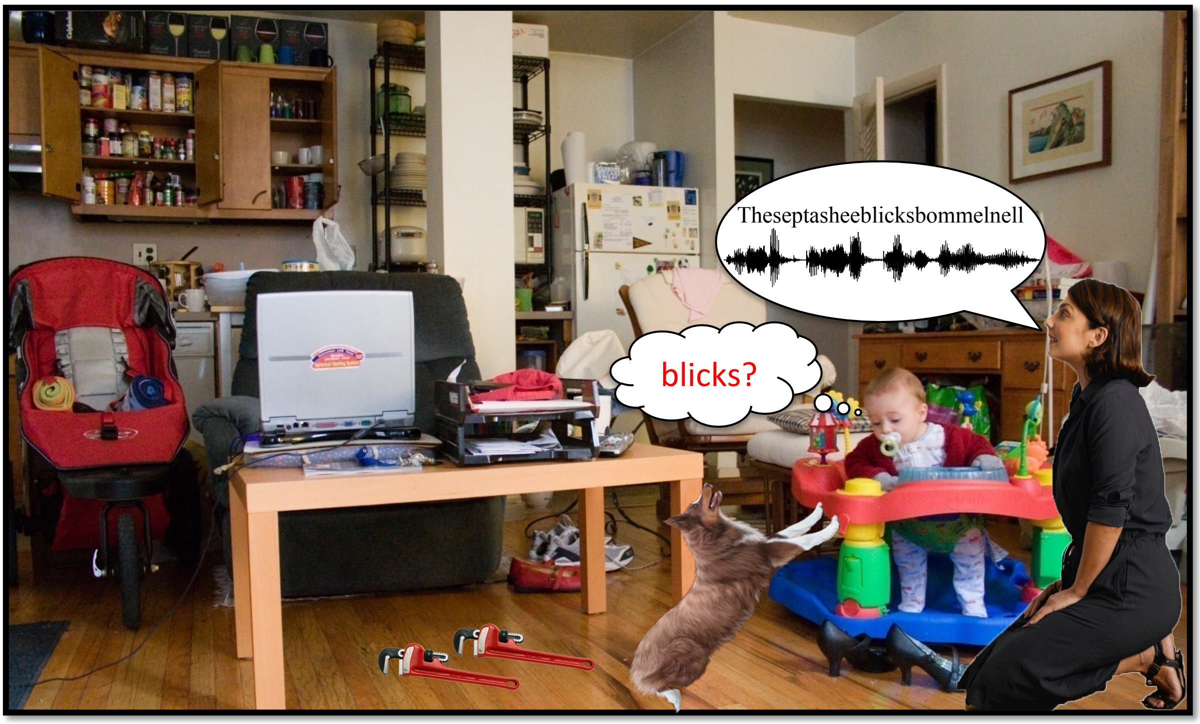


Figure 1: Example of an environment in which a child learns language. Figure adapted from Medina, Snedeker, Trueswell and Gleitman (2011). If we consider that in this sceneario that baby succeed to extract a novel word such as "blicks" from the fluent speech her mother is uttering, what would be the possible meaning/referent for this word in this context?

In order to understand the complexity of the word learning task, consider for instance the situation represented in Figure 1. The mother utters a simple sentence containing words that the child does not know yet. Given that when a sentence is uttered, the words are not separated from one another by small pauses, as the spaces that we put between words when we write, how could this child segment the speech that her mother is uttering into words? Note that there would be several possibilities to segment this spoken input (e.g., the septasheeblicks bommel nell; the septashee blicks bommel nell; the septa, she blicks bommel nell; the septa sheeblicks bommel nell; etc..), but what sources of information can children exploit in order to learn the best way to segment the fluent speech into words in their language?

Once infants become able to extract the word-forms from the continuous speech, they will have to associate meanings to these words. How can infants link words with possible meanings in their environment? Consider for instance, that the child in Figure 1 assumed that "blicks" is a word. What would be the possible meaning/referent for this word in the scenario illustrated in Figure 1? *Blicks* could be anything present in the figure (i.e., the red objects below the table, one of the toys that the child is using, something that is inside the wardrobes, etc.); *Blicks* could be a word used to describe the kind of movement that someone is doing (i.e., maybe the dog "blicks"); and *Blicks* could refer to something that is not even present in the scenario, which would make the learning of word meanings even more challenging for young children. A central problem to understand how humans acquire language is to determine what sources of information infants can exploit to go from the sounds of words to their meanings. In this thesis, I investigated some of the mechanisms allowing young children to accomplish this impressive task and some sources of information they can use to learn the meanings of words in their native language.

A classical hypothesis about language acquisition suggests that the syntactic structure, which governs the organization of words into sentences, could be an important source of information that children may exploit to discover the meaning of words (i.e., the *syntactic bootstrapping hypothesis* - e.g., Fisher, Hall, Rakowitz, & Gleitman, 1994; Gillette, Gleitman, Gleitman, & Lederer, 1999; Gleitman, 1990; Landau & Gleitman, 1985). In the simplest case to illustrate this idea, this hypothesis considers that if in the scenario presented in Figure 1, a child hear the sentence "Do you see the blicks?" she will probably infer that the novel word "blicks" refers to a kind of object, because it occupies the position of a noun in the sentence. She would therefore focus her attention on the objects present in her environment trying to identify which one could potentially represent the meaning of "blicks". However, if the sentence was "Do you see? She blicks!", the child would infer that "blicks" refers to an action, because it occupies a verb position in this sentence. She would therefore focus her attention on the actions happening in her environment in order to understand what "blicks" means. In other words, this hypothesis suggests that children would be able to exploit the syntactic context in which a novel word appears to determine its syntactic category (e.g., a noun or a verb) and use the syntactic category of the words to restrict the kind of meaning the novel word can have (e.g., verbs tends to refer to actions while nouns tends to refer to objects).

This hypothesis seems highly counterintuitive. Given that the syntactic structure defines the relationships between words in a sentence (e.g., determining who is the subject, the verb, or the object) and allows listeners to compute the meaning of a sentence from the meaning of the individual words that compose it (e.g., who is the agent or the patient of an action), one would expect that infants would first need to learn the words and their meanings, to then be able to learn how to organize words into sentences and therefore learn about syntax itself (e.g., Radford, 1990; Tomasello, 2000a, 2000b). Following this idea, one would expect that the child would first need to learn what “to eat”, “boy”, and “cake” means, to only then be able to know that in their language they can say “the boy eats the cake”, but not “the eats cake the boy”. We are therefore faced with a chicken-and-egg problem for language acquisition: infants would need words to learn syntax, but they would need syntax to learn words. How can infants avoid this circularity?

1.1. The syntactic bootstrapping hypothesis for language acquisition

The syntactic bootstrapping hypothesis proposes that having access to the syntactic structure of sentences can help children to discover the meaning of novel words because syntax could serve as a “zoom lens” to help learners figure out which part of the world is being talked about, and therefore to identify possible candidate meanings for the words that they do not know yet. In other words, this framework proposes that children can constrain their hypotheses about novel word meanings based on the kind of syntactic structures in which the words occur.

This idea is based on the fact that there is a close relationship between the meaning of words and the kind of syntactic structures in which they can appear. Note that we can say “She eats” or “The balls”, but not “She balls” or “The eats”. In the case of verbs more precisely, this relationship is expressed by the fact that the syntactic structure in which a verb appear is consistently related to the semantic properties of the verb. For instance, verbs describing actions that involves *moving/transferring an object* from one place to another (e.g., to give) may occur in structures taking three NPs as argument (the agent, the object, and the destination; e.g., as in the sentence "Alex gives his dissertation to the jury"). However, verbs describing actions that involves *mental states* will typically occur in structures taking a subordinate clause as argument (e.g., "Alex thinks that the baby is smart"). Thus, the syntactic structures in which the words occur reveal aspects of the words' meaning, and children may use this information to discover the possible meaning of words that they do not know yet.

The premises of the syntactic bootstrapping hypothesis is that the ability to extract such information from the syntactic structures of sentences can be rapidly generalized by young children during language acquisition. This abstract representation would provide children a preliminary format to support their acquisition of the meanings of other novel words that will appear in similar syntactic structures in their language (e.g., Fisher et al., 1994; Gillette et al., 1999; Gleitman, 1990; Landau & Gleitman, 1985). For instance, when we hear a sentence such as "the gorp is pilking the blicket", although we don't know the meaning of these words, we can still infer that "the gorp" is doing something (i.e., pilking) to "the blicket", and thus that gorp might be the agent of the action and blicket its patient. The question that arises is whether infants can indeed extract such abstract representation of sentences during language acquistion.

Decades of empirical research demonstrated that around the age of two, children seem to have already acquired important knowledge about syntax in their native language and that they can indeed use this information to constrain their syntactic analysis and therefore to constrain their interpretation of word meanings (e.g., Arunachalam & Waxman, 2010; Bernal, Dehaene-Lambertz, Millotte, & Christophe, 2010; Bernal, Lidz, Millotte, & Christophe, 2007; Brusini et al., 2017; Brusini, Dehaene-Lambertz, Dutat, Goffinet, & Christophe, 2016; Ferguson, Graf, & Waxman, 2014; He & Lidz, 2017; Naigles, 1990; Waxman, Lidz, Braun, & Lavin, 2009; Yuan & Fisher, 2009; Yuan, Fisher, & Snedeker, 2012). For instance, studies measuring event-related potentials (ERPs) in infants as young as 18 months demonstrated that they are able to compute syntactic structures online when listening to spoken sentences: ungrammatical sentences such as "**Je la poire*" – "I pear it" elicited a late positivity (resembling a P600) that was not observed for grammatical sentences such as "Je la mange" – "I eat it" (Bernal et al., 2010; Brusini et al., 2017). This suggests that during the online processing of these sentences, infants were sensitive to the fact that a syntactic frame such as "Je la…" predicts a verb but not a noun. Consistently with these findings, a recent study showed that 18-month-olds are able to learn that a novel word such as "pratch" refers to an action, when listening to sentences in which it appears as a verb, as in "Look! It is pratching!"; but when exposed to sentences like "Look! It is a pratch!" in which "pratch" appears in a noun position, they learn that "pratch" refers to an object (He & Lidz, 2017). These results support the idea that children can indeed exploit the syntactic frames in which novel words occur to determine the syntactic category of upcoming words. Infants interpreted the words as nouns or as verbs depending on the syntactic context in which the words appeared: words appearing after a pronoun or an auxiliary were interpreted as verbs while words appearing after an article or a determiner were interpreted as nouns. Moreover, they used this information to constrain their interpretation of the possible meaning of these novel words, mapping nouns to objects and verbs to actions.

In addition, the "zoom lens" provided by syntactic information can also reveal other aspects of the meaning of words. Studies demonstrated for instance that toddlers can also learn that a novel verb such as "blicking" refers to a causal action between two participants when listening to transitive sentences such as "She is blicking the baby", but they do not make the same inference when listening to intransitive sentences such as "She is blicking" (Arunachalam, Escovar, Hansen, & Waxman, 2013; Arunachalam & Waxman, 2010; Yuan & Fisher, 2009; Yuan et al., 2012). Going further, 19-month-olds exposed to sentences like "The dax is crying" were able to infer that "dax" referred to an animate entity (i.e., a novel animal), because it appeared in the subject position of a verb that requires an animate agent (i.e., the verb "crying" requires as a subject someone able to cry); but when exposed to sentences like "The dax is right here", they did not show any preference for the animate entity in comparison to an inanimate object at test, when they were asked to find "where is the dax" (Ferguson et al., 2014).

Taken together, all these studies show the important role played by the syntactic structure of sentences to bootstrap language acquisition in young children. We observe that at an age when toddlers do not have an extensive vocabulary yet, the syntactic structure of sentences provide them with an abstract representation of sentences in their language and they can use this information to discover the meaning of novel words. What remains unknown however is how children manage to access the syntactic structure of sentences so early during their development given that before two years old they do not know much about the meaning of words in their language yet.

1.2. How children access the syntactic structure of sentences before knowing the meaning of words: the role of phrasal prosody and function words

A rich literature suggests that a surface analysis of the speech signal itself could allow children to learn some aspects of the syntactic structure of their native language (i.e., the *prosodic bootstrapping hypothesis*, Morgan, 1986; Morgan & Demuth, 1996) even in absence of knowledge about the meaning of words. More specifically, two sources of information could allow children to extract syntactic information from the spoken input: phrasal prosody and function words (Christophe, Dautriche, de Carvalho, & Brusini, 2016; Christophe, Millotte, Bernal, & Lidz, 2008; Morgan & Demuth, 1996; Shi, 2014).

Phrasal prosody is the rhythm and melody of speech, it reflects the organization of sentences into prosodic constituents above the word level. Phrasal prosody in spoken languages is conveyed by variations in pitch, duration, and energy, of speech sounds over the course of an utterance. The importance of phrasal prosody for syntax comes from the fact that in all the languages of the world, when we speak, the words tend to be grouped together into intonational units (i.e., prosodic phrases) and the boundaries between these prosodic units (at the phrasal level: phonological and intonational phrases, as defined by Nespor & Vogel, 1986) always coincide with the boundaries between syntactic constituents in a sentence (e.g., Nespor & Vogel, 1986; Shattuck-Hufnagel & Turk, 1996).

To illustrate the relationship between prosodic and syntactic structures, consider for instance what happens when we utter a sentence such as "The little rabbits are jumping high". This sentence tends to be pronounced as [the little rabbits] [are jumping high], in which the words are organized into two group of words (brackets represent the prosodic units). Note that between these two group of words there is a boundary between two prosodic units that coincide with the syntactic boundary between the noun phrase, containing the words "the little rabbits", and the verb phrase containing the words "are jumping high". Although it is not the case that all syntactic constituent boundaries are marked by a prosodic boundary[1], whenever a

---

[1] Note that we could have said [the little rabbits are jumping high] into one single unit.

[2] The reverse is not true, since not all syntactic boundaries are marked with a prosodic boundary; for instance in short sentences such as [The boy eats], or [he eats], the syntactic boundary between the subject and the verb

prosodic boundary is perceived in the speech stream, it always coincides with a syntactic constituent boundary (Nespor & Vogel, 1986). As a result, salient prosodically-conditioned acoustic information such as phrase-final lengthening, pitch variations, and pauses, may allow listeners to identify prosodic boundaries, which in turn might be useful to segment the continuous speech stream into relevant units such as words and syntactic constituents, and to infer the position of some syntactic boundaries in their language (those coinciding with prosodic boundaries; Christophe et al., 2016, 2008; Morgan, 1986; Morgan & Demuth, 1996; Shi, 2014).

Function words and morphemes (i.e., highly frequent functional elements, such as articles, auxiliaries, pronouns, conjugation endings, etc.) are also salient to young children, because they are much more frequent than content words (nouns, verbs, adverbs; e.g., Gervain, Nespor, Mazuka, Horie, & Mehler, 2008; Kučera & Francis, 1967) and have perceptible characteristics that distinguish them from content words (Shi, Morgan, & Allopenna, 1998). Because functional elements carry almost exclusively morpho-syntactic information, they may allow infants to determine the syntactic nature of the constituents delimited by prosodic boundaries. In a language such as English for example, if infants indentify a prosodic unit contatining the words “the blicks” and if they know that articles typically precede nouns, they might infer that blicks is a noun. However, if infants identify a prosodic unit containing the words “She blicks” and if they know that pronouns typically precede verbs, they might infer that blicks is a verb. Thus, infants may exploit the information provided by function words to categorize the constituents delimited by phrasal prosody as noun phrases or verb phrases for instance (e.g., Gutman, Dautriche, Crabbé, & Christophe, 2015).

Taken together, this litterature suggests that phrasal prosody and function words could work as anchors to help infants access syntactic information from the speech stream, during the first steps of language acquisition (Christophe et al., 2016, 2008; Shi, 2014).

1.3. The syntactic skeleton hypothesis

To better understand the role that would be played by function words and phrasal prosody to provide syntactic information from the speech, consider for instance a situation in which, as young language learners, we do not know the meaning of words yet, but we have access to phrasal prosody and function words. When listening to a sentence such as **[*the*** feeze blickets**]** **[*are*** dax***ing*]**, thanks to the information provided by phrasal prosody, we could rapidly infer that there are two groups of words in this sentence **[**xXXx**]** **[**xXx**]**. Of course, the prosodic boundaries *per se*, do not tell us anything about the syntactic nature of these constituents. However, since we know the function words occurring inside each of these constituents, we can use this information to determine the syntactic nature of these prosodic units. For instance, by knowing that a noun is more likely to be preceeded by an article while a verb is more likely to be preceeded by a pronoun or an auxiliary, we could construct a first-pass syntactic structure of this sentence in the form **[*the*** XXx**]**$_{NP}$ **[*are*** X***ing*]**$_{VP}$, in which we consider that the first unit contains a noun because it is starting by the article "the", while the second unit, which follows a noun phrase, starts with the auxiliary "are" and contains the verbal morpheme "-ing" is much more likely to contain a verb. Thus, phrasal prosody would serve as a cue to delimitate units, while function words and morphemes would serve as a cue to determine the syntactic nature of these units delimitated by phrasal prosody.

This rudimentary syntactic representation, called a *syntactic skeleton* by Christophe and colleagues (Christophe et al., 2016, 2008) may be available to young children even without knowing the meaning of the content words making up the sentences: note that in our example the unknown content words were represented simply as syllables in the form of Xs, and we were still able to construct our syntactic analysis based on the information provided by the prosodic units and the function words. This mechanism may be all that children have to support the acquisition of word meanings during the first steps of language acquisition. Nevertheless, this mechanism could be sufficient to bootstrap language acquisition, if from the abstract representation provided by the syntactic skeleton of sentences, children could infer that words that occupy a noun position in the sentences typically refer to objects in their language while words that occupy verb positions typically refer to actions (e.g., Gleitman, 1990; Landau & Gleitman, 1985; Waxman & Lidz, 2006; Waxman et al., 2009; Waxman & Leddon, 2011).

This powerful mechanism would be extremely useful for infants to build their way into lexical and syntactic acquisition and to construct a first-pass syntactic structure of spoken sentences even before they know the meanings of many words. Recent computational work demonstrate that models relying on phrasal prosody and function words successfully predict the syntactic category of unknown words (e.g., Gutman et al., 2015). However, the plausibility of this mechanism for language acquisition needs to be supported by experimental evidence showing that children can indeed exploit phrasal prosody and function words during sentence processing and that they can use this information to constrain the acquisition of word meanings. This is exactly what I investigated in the current thesis. While decades of research suggest that young children exploit the syntactic structure in which words occur to discover the meaning of words, in the current thesis I investigated how children can access such abstract representation of syntactic structures without knowing the meanings of words in the first place. I investigated whether phrasal prosody and function words could work as anchors to help infants access syntactic information from the speech stream, during the first steps of language acquisition.

Given that phrasal prosody and function words are acquired well before infants know many words, the ability to jointly use this information to access the syntactic structure of sentences and to constrain the acquisition of word meanings would be crucial for infants to break free of the chicken-and-egg problem presented above.

The plan of this introductory chapter is as follows. I will provide the state of the art of what we know about the role played by phrasal prosody and function words during the first steps of language acquisition. First, I will present studies showing that infants are sensitive to phrasal prosody and function words from birth onwards. Second, I will review empirical evidence showing that infants can use phrasal prosody and function words to segment the fluent speech into words and that function words can allow children to determine the syntactic category of content words. Finally, I will present experimental evidence suggesting that infants can use phrasal prosody and function words together, for instance to infer the typical word order in their native language.

Chapter 2: State of the Art

2.1. Infants' early sensitivity to prosodic cues

The prosody of a language can be described at different levels (i.e., syllables, words, phrases, and whole utterances) and impacts linguistic interpretations in various ways. For example, prosodic information can be used to convey lexical meaning at the word level through variations in stress pattern or lexical tone (e.g. 'beBE' and 'BEbe' are two different words in Spanish, differing only in their stress patterns). Prosody is also used to mark whether a sentence is declarative or interrogative (e.g., Zhou, Crain, & Zhan, 2012; Zhou, Su, Crain, Gao, & Zhan, 2012) and it conveys useful discourse information such as information structure (e.g. focus, new vs. old information; Hirschberg & Pierrehumbert, 1986; Pierrehumbert & Hirschberg, 1990). Prosody can also be used to change the interpretation of an utterance (e.g. irony, disbelief, etc.), and it can even reflect the emotional state of the speaker (e.g., Armstrong, Andreu, & Esteve-gibert, 2016; Jun, 2005, 2014; Ladd, 2008). As for phrasal prosody which is the cue in which we will focus our studies in this thesis, the prosodic structure of an utterance can be described hierarchically, in the sense that utterances contain one or more intonational phrases, intonational phrases contain one or more phonological phrases, and phonological phrases contain one or more prosodic words, which in turn contain syllables (e.g., Nespor and Vogel, 1986).

Considerable work has shown that infants have extensive experience with prosody from their first days of life (Christophe, Mehler, & Sebastián-Gallés, 2001; Decasper & Spence, 1986; Mehler et al., 1988; Nazzi, Bertoncini, & Mehler, 1998; Shi, Werker, & Morgan, 1999). An important study conducted by Mehler and colleagues (1988) showed that four-day-old infants are already sensitive to prosodic information when listening to sentences. Testing French newborns with French and Russian, and American infants with English and Italian, the authors observed that infants showed a preference for listening to their native language over a foreign one. Given that a few days after birth, infants have received very little postnatal language input, it is possible that prenatal exposure plays a role in this early preference for native speech (e.g., Decasper & Spence, 1986). Indeed, some prosodic characteristics from the speech stream, such as rhythm, stress and intonation, pass through the skin and uterus to the fetus (Decasper and Spence, 1986). To test whether newborns' preference arises from an early knowledge of the prosodic characteristics of their native language, Mehler et al. (1988) tested another group of French and American infants using low-pass filtered speech samples, where the prosodic information remained intact while phonetic information was stripped away. Their results showed that French newborns preferred to listen to low-pass filtered French speech over filtered Russian speech and American infants showed the same pattern for filtered English vs. filtered Italian speech. These results suggest that from birth onwards, infants are already sensitive to prosodic information, which helps them to distinguish their native language from a foreign language.

In a subsequent study, Nazzi, Bertoncini and Mehler (1998), used low-pass filtered sentences in foreign languages to test infants' sensitivity to prosody to distinguish between non-native languages. They showed that French newborns were able to discriminate foreign languages based on their prosodic patterns. For example, French newborns were able to discriminate stress-timed English from mora-timed Japanese, but failed to discriminate between English and Dutch because these languages share the same rhythmic properties (both stress-timed languages). Taken together, these results show not only that infants can use prosodic cues to discriminate their native language from a foreign one but also that they can use prosody to categorize languages based on their rhythmic and intonational properties.

Subsequent studies provided further evidence about the role played by prosody during the first steps of language acquisition, showing in particular that infants use it to segment the continuous speech stream into chunks. One of the earliest infant studies on continuous speech perception showed that 7-month-old infants are able to rely on prosodic information to recognize clauses in the speech stream: Hirsh-Pasek et al. (1987) showed that 7-to-10-month-olds are sensitive to the coherence of intonational phrases in the speech stream. Inserting a one-second pause either at clause boundaries or at within-clause locations in the speech stream, the authors observed that infants prefer to listen to speech containing pauses at clause boundaries (i.e., "*Cinderella lived in a great big house*-PAUSE-*but it was sort of dark*-PAUSE-...") than within clauses (i.e., "*Cinderella lived in a great big house, but it was*-PAUSE-*sort of dark because she had*-PAUSE-…"; Hirsh-Pasek et al., 1987; see also Männel & Friederici, 2009 for similar results at 5 months measured by the EEG technique in infants).

Other studies showed that infants are sensitive to smaller prosodic units, phonological phrases, that may correspond to noun phrases or parts of verb phrases in a sentence, from about the age of 6 months (e.g., Gerken, Jusczyk, & Mandel, 1994; Johnson & Seidl, 2008; Soderstrom, Seidl, Nelson, & Jusczyk, 2003). Moreover, infants show better memory for units from the speech stream that correspond to whole prosodic units than for chunks of speech that span prosodic boundaries (Mandel, Jusczyk, & Nelson, 1994; Nazzi, Iakimova, Bertoncini, Frédonie, & Alcantara, 2006). These studies show that infants are not only sensitive to prosodic grouping information, but that they recognize the prosodic well-formedness of speech chunks, and find it harder to process chunks of speech that are not well-formed prosodically.

It is important to note that some of the cues that mark prosodic units are the same in all the world's languages, especially for larger prosodic units such as intonational phrases. These tend to be followed by a silent pause, systematically exhibit final lengthening, and often show marked pitch excursions (e.g., a decline, or a rise for questions; e.g., Jun, 2005, 2014; Ludusan et al., 2016; Ludusan & Dupoux, 2016; Nespor & Vogel, 1986; Shattuck-Hufnagel & Turk, 1996). Note that these kind of boundary cues are also found in other domains such as musical phrases (e.g., Jusczyk & Krumhansl, 1993; Knosche et al., 2005; Patel, Gibson, Ratner, Besson, & Holcomb, 1998). However, other cues may vary between languages, so that infants need to learn some language-specific properties before they can exploit those cues efficiently, which may explain why sensitivity to smaller prosodic units such as phonological phrases arises later during development. For instance, the pitch contours typical of smaller prosodic units such as phonological phrases vary between languages, which means that children need experience with their native language before they can exploit these cues as reliable boundary markers in their language. Experimental work suggests that while infants around 6 months (even across different languages) tend to rely on strong and universal prosodic markers, such as pauses, older infants (around 8-to-10 months) can make use of more subtle prosodic cues such as pitch contours and lengthening even in the absence of pauses (Johnson & Seidl, 2008; Seidl, 2007; Wellmann, Holzgrefe, Truckenbrodt, Wartenburger, & Höhle, 2012).

Taken together, the studies reviewed in this section show that prosody is an important source of information for young infants: newborns discriminate between languages on the basis of their rhythmic properties and during the first year of life infants become able to exploit the cues that mark prosodic boundaries between groups of words in a sentence. This information may promote early language learning in a variety of ways, such as to constrain the segmentation of fluent speech into words and to recognize groups of words that correlate with syntactic constituents in a sentence.

2.2. Infants' early sensitivity to function words

As stated earlier in the introduction, function words and morphemes are highly frequent functional elements that appear almost systematically in every sentence that we produce (i.e., note that it would be hard for us to utter an entire sentence without any determiner or verbal morpheme in it). Moreover, those functors elements have perceptible acoustic characteristics that distinguish them from content words (Shi, Morgan, & Allopenna, 1998). Taking these characteristics of function words into account, several studies investigated whether this information would be salient for infants in child-directed speech. In favor of this idea, distributional, phonological and acoustic analyses of the child spoken input reveal that function words are highly frequent relative to content words (Gervain, Nespor, Mazuka, Horie, & Mehler, 2008; Kučera & Francis, 1967), function words typically occur at the edges of syntactic phrases, and tend to be short and unstressed relative to content words (Shi, Morgan, & Allopenna, 1998). The question that arises is, do infants exploit this information to distinguish between content and function words in their language?

The perceptual characteristics that differentiate function words from content words have been shown to be perceived by infants, from their first days of life (Shi & Werker, 2001; Shi et al., 1999). Newborns who were habituated to listen to a list of content words, showed surprise when they were then exposed to a list of function words during the test, and vice-versa for when they were habituated with a list of function words and then tested with a list of content words. Crucially, another group of newborns did not show any surprise when they listened to the same type of list of words between habituation and test, which suggests that from the first days of life infants are already able to discriminate between the acoustical properties of function versus content words (Shi et al., 1999).

Before their first birthday, around 11 months, infants have already acquired sensitivity to frequent function words in their native language and they can even recognize mispronounciations in function words (Hallé, Durand, & de Boysson-Bardies, 2008; Höhle & Weissenborn, 2003; Höhle, Weissenborn, Kiefer, Schulz, & Schmitz, 2004; Kedar, Casasola, Lust, & Parmet, 2017; Shafer, Shucard, Shucard, & Gerken, 1998; Shi, Werker, & Cutler, 2006). For instance, in Kedar et al., (2017), in a preferential looking task, 12-month-olds were exposed to both grammatical sentences using the determiner "the" (i.e., "Can you see **the** ball?") and ungrammatical conditions in which "the" was replaced by another English function word or omitted (e.g., "Can you see **by** ball?"). The results showed that infants oriented faster to a target image (and had more correct looks to target words) following grammatical sentences than ungrammatical sentences.

Function words tend to consistently co-occur with content words from a specific word class: determiners such as “the” or “a” typically co-occur with nouns, while pronouns such as “she” and “they” tend to co-occur with verbs (e.g., Mintz, 2003). Several studies suggested that the acquisition of function words during the first year of life might allow infants to anticipate the lexical category of a subsequent content word (Christophe et al., 2016, 2008; Hochmann, 2013; Hochmann, Endress, & Mehler, 2010; Shi & Melançon, 2010; Shi, 2014). For instance, infants might infer that a novel word appearing after a determiner might be a noun (e.g., the blicks), while a novel word appearing after a pronoun might be interpreted as a verb (e.g., she blicks). Moreover, if infants use the morpho-syntactic information carried by function words to anticipate the lexical category of a subsequent content word, they could also use this information to infer the meaning of novel words (mapping nouns to objetcs and verbs to actions, e.g., Christophe et al., 2016; Gleitman, 1990; Shi, 2014).

Supporting this hypothesis, studies demonstrated that around 14 months, infants exposed to a novel content word following a determiner, such as “**the** *blicket*”, infer that this novel word can also follow other determiners as in “**a** *blicket*”, but it cannot follow pronouns as in “**she** *blicket*” (Höhle et al., 2004; Shi & Melançon, 2010). Very recently, He and Lidz (2017) showed that 18-month-olds were able to infer that a novel word such as “pratch” refers to an action, when listening to sentences in which “pratch” occurs in a verb position in a sentence, as in “Look! It ***is*** pratch***ing***!”; but when exposed to sentences like “Look! It is ***a*** pratch!” in which “pratch” appeared right after the determiner “**a**”, infants learned that “pratch” refered to an object. Consistently with these results, other studies have pointed out that by 18 months of age, lexical access is speeded and more accurate when familiar nouns and verbs are preceded by a function word from an appropriate category than when they are preceded by a function word from an inappopriate category (i.e. determiners and pronouns preceding words such as “ball” or “eats”, versus the reverse; Cauvet et al., 2014; Kedar, Casasola, & Lust, 2006; Zangl & Fernald, 2007).

Taken together, these studies suggest that function words are acquired within the first year of life: by 12 months, infants have already acquired some form of sensitivity to determiners and pronouns in sentence processing, and they use this information when computing the syntactic category of content words.

2.3. How do infants extract words from the fluent speech?

Discovering the words composing their language is also one of the challenging tasks faced by infants during language acquistion. During their development, infants have to extract the word-forms from the continuous speech stream, and associate the extracted word-forms with a possible meaning. However, given that in fluent speech there are no pauses between two consecutive words, how can infants segment the speech stream into words? Several studies have shown that infants can use a variety of cues to discover word boundaries; although none of them is sufficient on its own, together they may allow infants to discover many words in their input (see He & Arunachalam, 2017, for a review). Word segmentation cues that have been proposed and studied include phonotactic constraints (e.g., Jusczyk, 1997), frequently occuring words such as words heard in isolation or the infants own names (e.g., Bortfeld, Morgan, Golinkoff, & Rathbun, 2005; Lew-Williams, Pelucchi, & Saffran, 2011), the statistical structure of the input (e.g., transitional probabilities; e.g., Saffran, Aslin, & Newport, 1996), word-level prosodic information such as typical word stress pattern (e.g., Cutler & Butterfield, 1992; Echols, 1993; Echols & Newport, 1992; Jusczyk, Houston, & Newsome, 1999), function words that appears very frequently in the speech stream (Gervain et al., 2008; Haryu & Kajikawa, 2016; Johnson, Christophe, Demuth, & Dupoux, 2014; Shi & Lepage, 2008), as well as phrasal prosody (e.g., Gout, Christophe, & Morgan, 2004; Johnson, 2008; Shukla, Nespor, & Mehler, 2007; Shukla, White, & Aslin, 2011). Since in the current thesis we are interested in the role played by phrasal prosody and function words to bootstrap the acquisition of word meanings, in this section I will turn our attention towards the role played by phrasal prosody and function words to extract word forms from the fluent speech.

2.3.1. The role of function words on word segmentation

Given that function words often appear before or after a content word and that infants are sensitive to this information very early during their development, several studies suggested that the acquisition of function words could also allow infants to rely on functors to segment the speech stream into words. For instance, by simply noticing that "the" is a function word in their language, infants listening to a chunk of words such as "theblicks" could infer that "blicks" is a word on its own (Christophe et al., 2016, 2008; Hochmann, 2013; Hochmann, Endress, & Mehler, 2010; Shi & Melançon, 2010; Shi, 2014).

Supporting this hypothesis, empirical research demonstrated that infants from 8 months of age can use frequent functors to segment potential word forms from continuous speech in English (Shi, Cutler, Werker, & Cruickshank, 2006), in French (Shi & Lepage, 2008), in Dutch (Höhle & Weissenborn, 2003) and even in Japanese (Haryu & Kajikawa, 2016). In Shi and Lepage (2008) for instance, French-learning infants were familiarized to two utterance types: an unknown noun preceded by a frequent function word in French (e.g., **des** *preuves* – proofs), and another unknown noun preceded by a prosodically matched nonsense function word (e.g., **kes** *sangles* – **kes** straps). After familiarization, infants' segmentation of these two nouns was assessed in a test phase where they were presented to the nouns (*preuves*, *sangles*) in isolation. The results showed that infants who were exposed to "des preuves" during the familiarization phase, recognized the word "preuves" as a single word during the test. However infants familiarized with "kes sangles" during familiarization did not recognize the word "sangles" at test, because they may have encoded "kessangles" as one single word. These results suggest that infants were able to use familiar function words (as opposed to nonsense functors such as *kes*) to extract unknown nouns from fluent speech.

In a follow-up experiment, Shi and Lepage (2008) tested whether similar results would be observed, if function words that are not "familiar" for infants (e.g., *vos* – your) were used to preceed the unknown nouns (preuves, sangles). The results showed that the "infrequent" function word "vos" in French, did not facilitate word segmentation in infants, which suggests that the frequency of function words is a crucial factor. The conclusions that we can draw from these results is that frequent function words can help infants to segment words early on, since these function words provide infants a cue to extract word forms from continous speech before their first birthday.

Taken together, the studies presented in this section (together with the section about infants' early sensitivity to function words) demonstrate that function words are acquired within the first year of life and are used for categorization between 12 and 18 months. Additionally, these studies suggest that early on in life, function words can help infants to segment the speech stream into words and these function words can trigger infants' expectations about the word class of upcoming words in the speech.

2.3.2. The role of phrasal prosody on word segmentation

As we demonstrated in section 2.1, several studies have shown that infants are able to perceive prosodic phrase boundaries and exploit them to find the boundaries between groups of words from 6 months onwards (Gout et al., 2004; Johnson, 2008; Millotte et al., 2010; Shukla et al., 2011). Since prosodic units such as phonological phrases are constructed by grouping words together, whenever a prosodic boundary is perceived, it has to correspond to a word boundary (Nespor & Vogel, 1986). Thus, sensitivity to phrasal prosody might also provide cues to speech segmentation and, therefore, constrain lexical search in infants. Of course, since most prosodic units contain more than one word, many word boundaries will not be marked by phrasal prosody – word segmentation within phonological phrases will have to rely on some other cues mentioned above. As an illustration of the impact of prosodic boundaries on word segmentation, the sentences below (1a and 1b) both contain the two syllables *pay* and *per*, however, only the first one contains the word *paper* (1a):

(1) a. [The college] [with the biggest *paper* forms] [is best]
b. [The butler] [with the highest *pay*] [*per*forms the most]

In the second sentence (1b), we observe that the prosodic boundary between the syllable 'pay' and the syllable 'per' should block lexical access to the word paper: indeed, it has been shown with adult listeners that prosodic boundaries constrain lexical access (Christophe, Peperkamp, Pallier, Block, & Mehler, 2004; Endress & Hauser, 2010; Warner, Otake, & Arai, 2010). To test infants' ability to use phrasal prosody to segment the speech stream into words, Gout, Christophe, and Morgan (2004) used the above sentences (1) in a conditioned head-turn procedure. In a first session, they trained American 10-month-olds to turn their heads toward a puppet whenever they heard the word *paper*, for instance. Then, during a test phase infants were exposed to full sentences, such as (1a) and (1b). Their results showed that infants trained to respond to the word *paper* turned their head more often when listening to paper-sentences (1a) than to pay#per-sentences (1b). In contrast, infants trained to respond to the target word *pay* turned equally often for both types of sentences (since the syllable pay was present in both sentences, the target word *pay* might have been noticed in both sentences). This result shows that 10-month-old American infants can use phrasal prosody to segment the speech stream into words and, therefore, to constrain their lexical access.

Further studies showed similar results with French 16-month-olds (Millotte et al., 2010) and other experiments in English confirmed these results, showing that 12-month-olds can use phrasal prosody to constrain lexical access within strings of nonsense syllables differing in their prosodic structure (Johnson, 2008). Moreover, Shukla, White, and Aslin (2011) showed that 6-month-olds were able to better associate a visual referent to a novel word aligned with a prosodic phrase boundary, than to a novel word that straddled a prosodic boundary. In other words, infants were able to associate a novel word (e.g., AB) appearing at the edge of a prosodic boundary (i.e., as in the sentence "[xAB] [yz]") with an object moving on the screen. But they did not learn the same association when exposed to that novel word straddling two prosodic units (e.g., [xA][Byz]). Thus in the test phase, when listening to the novel word [AB], infants who were exposed to "[xAB] [yz]" sentences look more towards the object moving than infants who were exposed to "[xA] [Byz]" sentences.

Taken together, the studies reviewed in this section show that infants are sensitive to the fact that words are aligned with prosodic phrase boundaries, and exploit this information to facilitate word learning. These results highlight the importance of phrasal prosody for segmenting the speech stream into words and to constrain lexical access and its acquisition.

2.4. *Infants can exploit prosody and function words together to discover the word order in their native language*

Prosody and function words have also been proposed to help infants discover the word order of their native language. In fact, many languages differ with respect to the canonical word order they use (e.g, the position of the verb and its object in a sentence), or more generally, the respective position of heads and complements. Either complements tend to follow their heads, or they tend to precede them. This organization, in turn, can impact the order of function words with respect to content words in a sentence (Dryer, 1992; Gervain et al., 2008; Kučera & Francis, 1967). For example, in VO (Verb-Object) languages such as English and French, function words typically appear before content words and at the beginning of phrases (e.g., *Le bateau* 'The boat'; *de Paris* 'from Paris' – head-initial languages), while in OV languages like Turkish and Japanese, function words tend to appear after content words and at the end of phrases (e.g., *Tokyo kara* 'Tokyo from' – head-final languages; Dryer, 1992). In addition, the head-direction of a language also determines which element will be more prosodically prominent within phonological phrases: the first one in head-final languages (typically marked with higher pitch), and the last one in head-initial languages (typically marked with lengthening; Nespor et al., 2008). Thus, infants could use both the prosodic cues and the relative position of frequent (function words) and infrequent (content words) elements in a sentence to infer the basic word order of their native language (Bernard & Gervain, 2012; Christophe, Nespor, Guasti, & Van Ooyen, 2003; Gervain et al., 2008; Gervain & Werker, 2013; Höhle, Weissenborn, Schmitz, & Ischebeck, 2001).

Consistently with this hypothesis, Christophe, Nespor, Guasti, and Van Ooyen (2003) showed that 2-month-old infants were able to distinguish between two languages that have very similar phonology, but differ in their head-direction: French (head-initial) versus Turkish (head-final), suggesting that this kind of prosodic information might be used by young listeners to obtain information about the word order in their native language. Additionally, Gervain and colleagues (2008) showed that 8-month-olds were sensitive to the typical position in which frequent and infrequent elements appear in their native language in order to infer the position of function and content words (e.g., Italian: VO, frequent-infrequent; or Japanese: OV, infrequent-frequent). Thus, in an artificial grammar experiment, when exposed to an unsegmented string of syllables in which some syllables were highly frequent (i.e., playing the role of function words) and others infrequent (i.e. playing the role of content words), infants segmented this continuous signal in such a way that the position of the frequent elements respected the typical order of function and content words in their native language (Italian infants prefered to have the frequent elements in initial position, and Japanese infants in final position).

Interestingly however, in the case of bilingual infants acquiring both VO and OV languages, frequency alone does not provide enough information about word order since both frequent-final and frequent-initial phrases occur in their input. Gervain and colleagues have proposed that prosodic information could cue word order in this case. For instance, Gervain and Werker (2013) showed that bilingual 7-month-olds acquiring simultaneously an OV and a VO language, exploit prosodic information to determine the relative order of frequent and infrequent elements in an unsegmented string of syllables. When familiarized with strings of syllables consistent with an OV prosodic pattern (with high-low-high-low pitch alternations), infants preferred to listen to chunks of syllables with the frequent elements at the end; in contrast, when familiarized with strings of syllables consistent with a VO prosodic pattern (short-long-short-long), they showed the reverse preference. These studies suggest that prosody, together with frequency information (of function words), impacts word order acquisition in infants. These results suggest that young infants are able to exploit these sources of information together during sentence processing.

2.5. Summary: Are infants able to jointly use phrasal prosody and function words to access the syntactic structure of sentences and to constrain their acquisition of word meanings?

All the empirical studies presented above show that phrasal prosody and function words are jointly predictive of syntactic structure in natural languages and that infants are sensitive to each of these sources of information early during their development. We saw that infants acquire function words before their first birthday and that they can use function words to predict the grammatical category of an upcoming content word. Additionally, phrasal prosody and function words have an important role in infant speech perception, because they allow infants to discriminate between languages, infer aspects of the word order of their language, and segment the speech stream into words and clauses. What has never been investigated however is whether infants are able to jointly use phrasal prosody and function words to access the syntactic structure of sentences and to constrain their acquisition of word meanings.

In this thesis, I experimentally tested whether young children can use phrasal prosody and function words to access the syntactic structure of sentences and to contrain their syntactic analysis (Part 1) and whether this information in turn allows infants to constrain the acquisition of word meanings (Part 2).

Part 1

Can infants use phrasal prosody and function words to constrain their syntactic analysis?

As stated in the introduction, since the prosodic structure of an utterance partially reflects the syntactic structure of a sentence (Nespor and Vogel, 1986), the *prosodic bootstrapping hypothesis* proposes that phrasal prosody might help infants to discover the syntactic structure of sentences (e.g., Morgan and Demuth, 1996; Morgan, 1986). The relationship between prosodic structure and syntactic structure is such that prosodic boundaries are aligned with syntactic constituent boundaries. Thus, prosodic information such as phrase-final lengthening, pitch variation and pauses could help listeners to identify prosodic boundaries, and therefore to find some syntactic constituent boundaries (Christophe et al., 2016, 2008; Morgan, 1986; Morgan & Demuth, 1996). Taking this hypothesis into account, prosodic information might facilitate on-line sentence processing in adults, and might provide a way for infants to identify some of the syntactic constituents of an utterance even before they have acquired an extensive vocabulary.

In support to this hypothesis, many studies have shown that adults integrate phrasal prosody online to recover the syntactic structure of sentences (Kjelgaard & Speer, 1999; Michelas & D'Imperio, 2015; Millotte, René, Wales, & Christophe, 2008; Millotte, Wales, & Christophe, 2007; Snedeker & Trueswell, 2003; Weber, Grice, & Crocker, 2006). Given the extensive literature I reviewed in chapters 1 and 2 showing that infants have an early access to phrasal prosody, and are able to exploit it for lexical segmentation, one would naturally expect that infants might also be able to use phrasal prosody, as adults do, to constrain their syntactic analysis. However several studies investigating whether preschoolers can exploit phrasal prosody to constrain their syntactic analysis have found that children have difficulties using prosody for syntactic ambiguity resolution in English (Snedeker & Trueswell, 2001, 2003; Vogel & Raimy, 2002) and in Korean (Choi & Mazuka, 2003). Most of these studies in English used sentences with a prepositional phrase attachment ambiguity, 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". In such sentences, the default prosodic structure however is the same for the two possible interpretations (i.e. [Can you touch] [the frog] [with the feather] – Snedeker & Yuan, 2008), but speakers who are aware of the ambiguity can intentionally disambiguate by exaggerating one of the prosodic breaks, in order to favor one interpretation over the other, for example: "[Can you touch the frog] [with the feather]?" for the instrument interpretation versus "[Can you touch] [the frog with the feather]?" for the modifier interpretation (Snedeker & Trueswell, 2003).

Snedeker and Trueswell (2001) initially found that 4-to-7-year-olds failed to use this kind of prosodic information when interpreting sentences with PP-attachment ambiguities. In a subsequent experiment however, which controlled for children's perseveration biases, Snedeker and Yuan (2008) observed that preschoolers succeeded in this task, when they were presented with only one kind of sentences: either modifier-only, or instrument-only. However, when children were presented with both instrument and modifier sentences across the experiment, as in the previous study, they failed to use prosody to constrain their syntactic interpretations (Snedeker and Yuan, 2008). My interpretation of these findings by the time I started this thesis was that children's difficulty using prosody in this kind of sentences might be due to the fact that the disambiguating prosodic breaks they needed to use were not part of the normal prosodic structure of these sentences. Instead, these cues are only produced when the speaker is consciously trying to disambiguate (as established in Snedeker and Trueswell, 2003). Thus, children may have had difficulties using this kind of optional prosodic information. It is therefore difficult to infer from these studies whether or not younger children do exploit phrasal prosody in their processing of everyday sentences to constrain their syntactic analyses.

To avoid this problem of optional prosodic disambiguation when testing children's ability to exploit phrasal prosody to constrain their syntactic analyses, in the first part of this thesis I exploited the case of locally ambiguous sentences featuring noun/verb homophones, in which the default prosodic structure differed between conditions. For example, the word "watch" is a verb in the sentence: [Mommies] [**watch** TV every night], but it is a noun in the sentence: [Mommy's **watch**] [ticks very noisily]. Here, brackets indicate prosodic units, which reflect the syntactic structure of each sentence. Crucially, in both cases, there is a prosodic break (marked by phrase-final lengthening and pitch change) between the subject Noun Phrase and the Verb Phrase: this break falls after the critical word when it is used as a noun, and before it when it is used as a verb. Note that the prosodic boundary between the noun phrase and the verb phrase in these examples is part of the normal prosodic structure of sentences and was found to be naturally produced by naive speakers who were unaware of the local ambiguity (Millotte et al., 2007). This kind of prosodic boundaries is present even in non-ambiguous sentences, for instance between the noun phrase and the verb phrase in the sentence "[the little rabbit] [eats a lot of food]".

In **Chapter 3**, I tested whether 3-to-5-year-old children can use phrasal prosody online to constrain their syntactic analysis of sentences and recover the meaning of a noun-verb homophone in French. For instance, if they listen to a sentence such as "[*La petite*

ferme]… – [The little **farm**]… , they should interpret “ferme” as a noun, but in a sentence such as [*La petite*] [***ferme***… – [The little girl] [**closes**…, they should interpret “ferme” as a verb (de Carvalho, Dautriche, & Christophe, 2016, *DevSci*). In **Chapter 4**, I investigated whether this ability would also be attested with English-speaking preschoolers when listening to sentences such as “the baby flies” in which the prosodic structure can help them to decide whether “flies” is a noun or a verb (de Carvalho et al., 2016 *JASA*) and finally in **Chapter 5**, I investigated whether the relationship between prosodic and syntactic structures could be exploited even at a younger age, during the first steps of syntactic and lexical acquisition. I tested whether infants around 20 months of age, who are still in the process of learning their language, can use phrasal prosody and function words to recover the syntactic structure of sentences and to predict the syntactic category of words, an ability which would be extremely useful to discover the meaning of words (de Carvalho et al., 2017 *Cognition*).

Chapter 3

Preschoolers use phrasal prosody online to constrain syntactic analysis

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

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_{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*$_{DET}$ *petite*$_{ADJ}$ ***ferme***$_{NOUN}$] [*lui plait beaucoup*]) and a second one with the ambiguous word used as a verb (hereafter the verb prosody condition, e.g. [*La*$_{DET}$ *petite*$_{NOUN}$] [***ferme***$_{VERB}$ *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

[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).

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_{verb} = 403 ms, SD_{verb} = 50.4 vs. M_{noun} = 204 ms, SD_{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_{noun} = 427 ms, SD_{noun} = 50.6 vs. M_{verb} = 317 ms, SD_{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_{verb} = 205 ms, SD_{verb} = 16.2 vs. M_{noun} = 121 ms, SD_{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*$_N$ ***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).

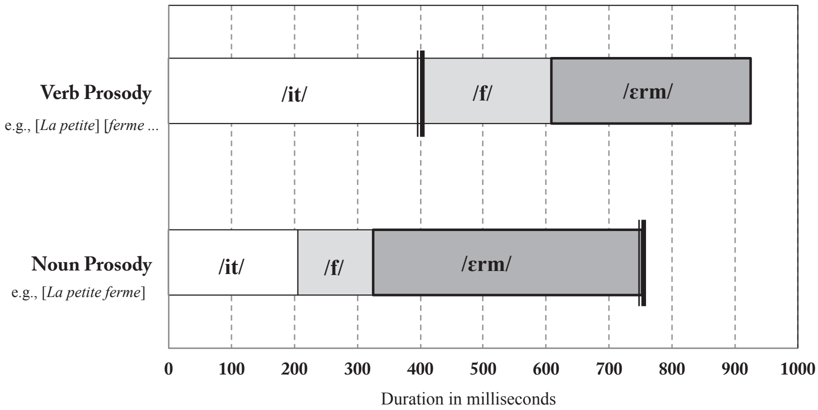


Figure 1 *Mean duration (in ms) of the different segments around the prosodic boundaries for both conditions: noun and verb prosody (phonological phrase boundaries are represented by thick black lines). Note that to illustrate, we use the segments for the experimental sentences of the item /fɛrm/, but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

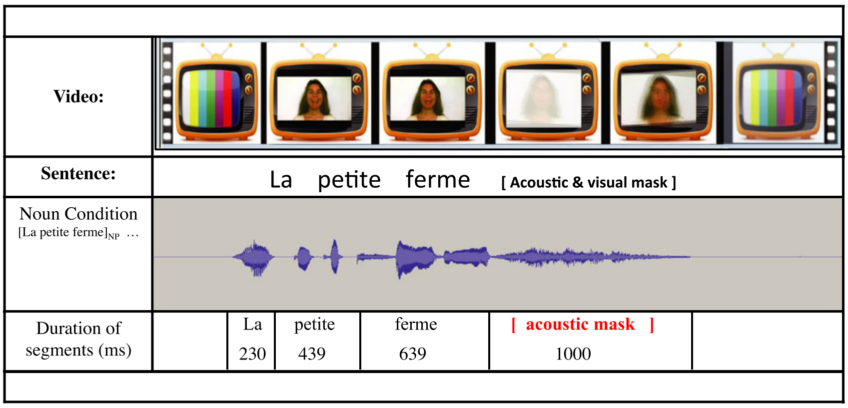


Figure 2 *Example of a test sentence used in the completion task (Experiment 1) together with its waveform and the duration of each of the components.*

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

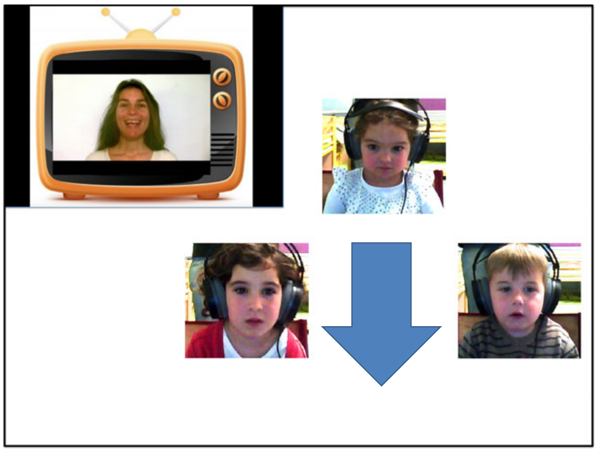


Figure 3 *Example of the scenario used in the completion task (Experiment 1) for each trial: first the blue arrow turned and selected which child would play. Then, children saw the video in the upper left corner of the screen as illustrated here. Finally, they completed the sentence they had heard in 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 interceptβ_0, reflecting the baseline probability of giving a noun answer, and a slope estimateβ_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. β_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 fromβ_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 slopeβ_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:

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

where e_{is} is the normally distributed error for the observation. β 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 χ^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 (β = 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,

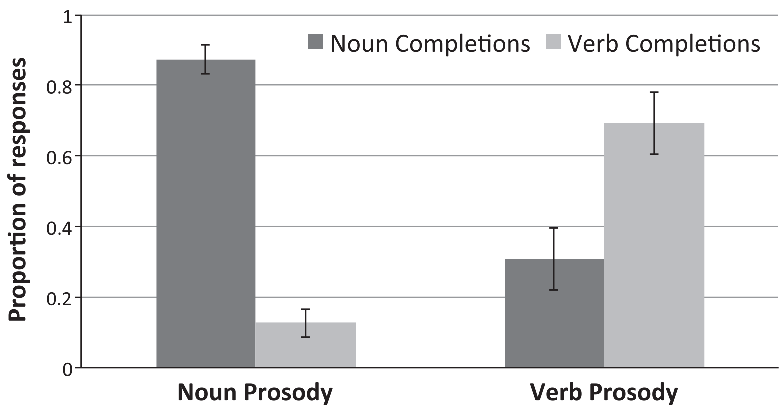


Figure 4 *Proportion of noun and verb completions for each prosody condition. Error bars represent the standard error of the mean.*

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

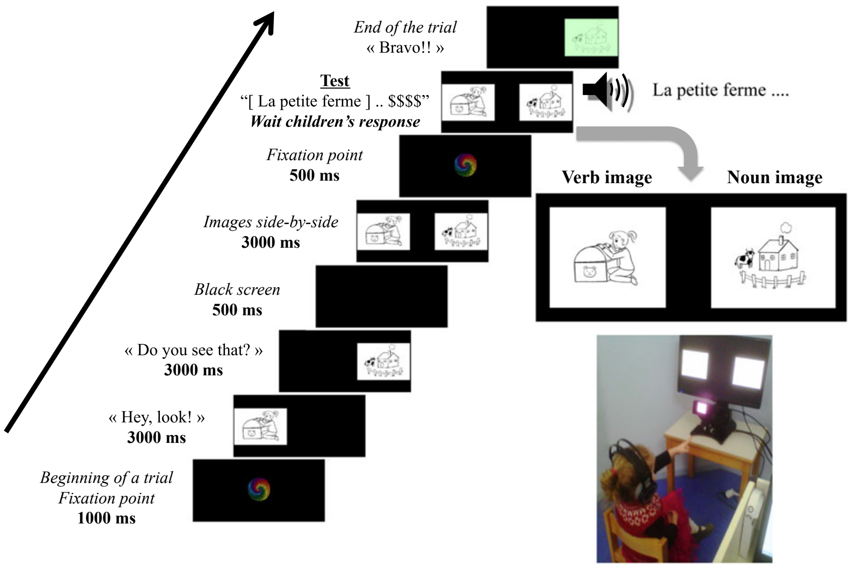


Figure 5 *Time course of a trial in Experiment 2. Each trial started with a fixation point in the middle of the screen for 1000 ms. Then, each image was presented alone for 3 seconds on the left or the right side of the screen with an audio prompt. Then, after a 500-ms black screen, both images were 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)

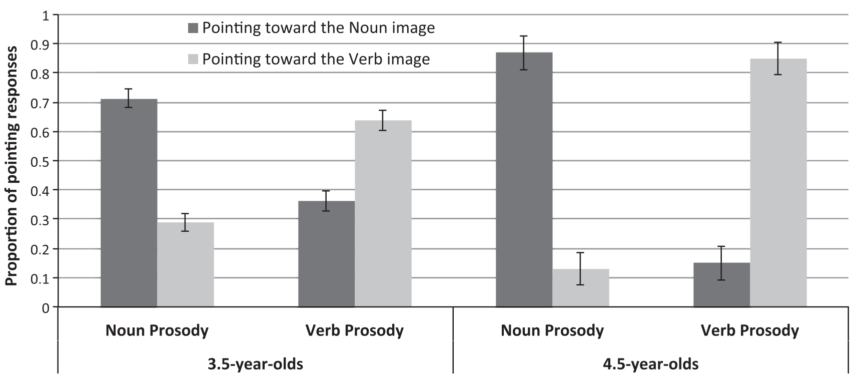


Figure 6 *Proportion of pointing responses toward the noun image and the verb image after listening to the target word, broken down by prosody condition, for each group of participants. Error bars represent the standard error of the mean.*

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

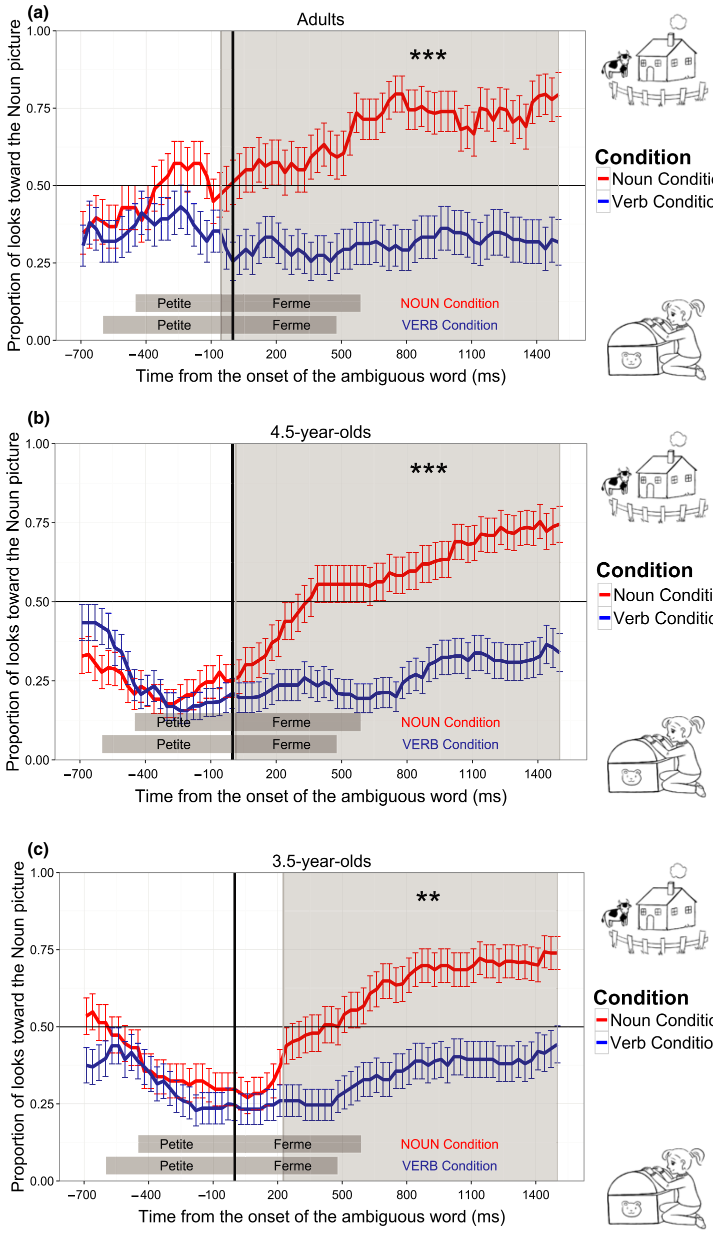


Figure 7 *Proportion of looks toward the noun image, time-locked to the onset of the ambiguous word (vertical black line) for (a) Adults, (b) 4.5-year-olds, and (c) 3.5-year-olds, for the noun prosody condition (red curve) and the verb prosody condition (blue curve). Error bars represent the standard 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*$_N$] [*blicks*$_V$ *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 Pair of ambiguous words	Syntactic category	Target	Full sentence recorded
fermer × la ferme *to close × the farm*	Verb	Ferme	La petite **ferme** le coffre à jouets *The little one **closes** the toy box*
	Noun		La petite **ferme** lui plait beaucoup *The small **farm** pleases him a lot*
lire × le lit *to read × the bed*	Verb	Lit	Le grand **lit** souvent des histoires à son petit frère *The big one often **reads** stories to his younger brother*
	Noun		Le grand **lit** sera pour les parents *The big **bed** will be for the parents*
marcher × la marche *to walk × the step (of a staircase)*	Verb	Marche	La grande **marche** lentement toute la journée *The big one **walks** slowly all day long*
	Noun		La grande **marche** la fait tomber *The big **stair** makes her fall*
moucher × la mouche *to blow somebody's nose × the fly*	Verb	Mouche	La maman **mouche** le bébé malade *The mother **blows** the sick baby's nose*
	Noun		La maman **mouche** laisse son bébé tout seul *The mummy **fly** leaves her baby alone*
porter × la porte *to carry × the door*	Verb	Porte	La vieille **porte** sa montre à réparer *The old lady **carries** her watch to be repaired*
	Noun		La vieille **porte** sera réparée demain *The old **door** will be repaired tomorrow*
montrer × la montre *to show × the watch*	Verb	Montre	La grande **montre** ses jouets à son frère *The big one **shows** her toys to her brother*
	Noun		La grande **montre** sera réparée demain *The big **watch** will be repaired tomorrow*
sourire × la souris *to smile × the mouse*	Verb	[suri]	Le bébé **sourit** à sa maman *The baby **smiles** to his mom*
	Noun		Le bébé **souris** a bien mangé *The baby **mouse** ate well*
pêcher × les pêches *to fish × the peaches*	Verb	[pɛʃ]	Les grosses **pêchent** mon poisson préféré pour le dîner *The fat ones **fish** my favorite fish for dinner*
	Noun		Les grosses **pêches** me font très envie *The big **peaches** tempt me a lot*

Chapter 4

English-speaking preschoolers can use phrasal prosody for syntactic parsing

de Carvalho *et al.*: JASA Express Letters [http://dx.doi.org/10.1121/1.4954385] Published Online 23 June 2016

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 Christophe[1]
[1]*Laboratoire 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*
x.de.carvalho@gmail.com, lyn.tieu@gmail.com, anne.christophe@ens.fr
[2]*Department of Linguistics, University of Maryland, College Park, Maryland 20742, USA*
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.

[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

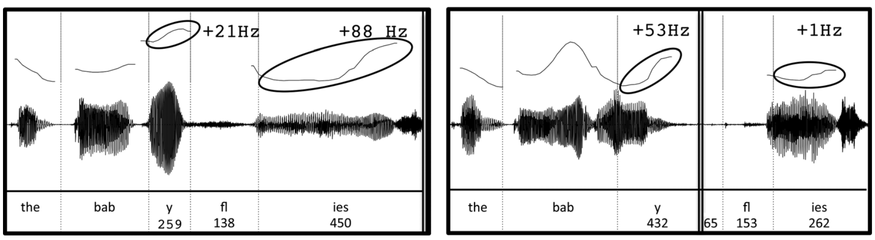


Fig. 1. Mean duration of the different segments, and pitch contours in the ambiguous region. Prosodic boundaries are represented with thick black lines. Ellipses delimit the areas where pitch analyses were performed, subtracting pitch at the beginning of the rime to pitch at the end of the rime, to determine 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	*t* 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 - *mean differences in semitones in italics*.

Dependent variable	Noun sentence	Verb sentence	Difference	*t* test (2-tailed)
Rime: word preceding Target (e.g., y from "baby")	21 (20.8) *1.17*	53 (23.3) *4.03*	−31 (33.3) *−2.86*	$t(7) = -0.94$; $p = 0.37$
Rime: Target word (e.g., ies from "flies")	88 (17.4) *6.56*	1 (10.1) *0.12*	87 (20.7) *6.45*	$t(7) = 4.20$; $p < 0.01$**

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 offline 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 in "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 verb[2] 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 preschoolers[14] 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

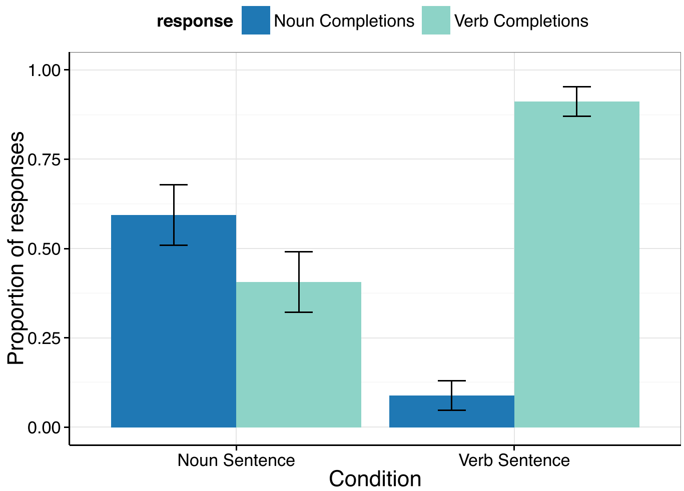


Fig. 2. (Color online) Proportion of noun and verb completions for each condition. Error bars represent the standard error of the mean.

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 IIe-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

[1]The full model was thus as follows: 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.

[2]The 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.

[1]J. Morgan, *From Simple Input to Complex Grammar* (MIT Press, Cambridge, MA, 1986).
[2]M. Nespor and I. Vogel, *Prosodic Phonology* (Foris, Dordrecht, the Netherlands, 1986).
[3]M. 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).
[4]S. Millotte, R. Wales, and A. Christophe, "Phrasal prosody disambiguates syntax," Lang. Cogn. Process. **22**, 898–909 (2007).
[5]S. 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).
[6]J. Snedeker and J. Trueswell, "Using prosody to avoid ambiguity: Effects of speaker awareness and referential context," J. Mem. Lang. **48**, 103–130 (2003).
[7]J. 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).
[8]J. Snedeker and S. Yuan, "Effects of prosodic and lexical constraints on parsing in young children (and adults)," J. Mem. Lang. **58**, 574–608 (2008).
[9]I. Vogel and E. Raimy, "The acquisition of compound vs. phrasal stress: The role of prosodic constituents," J. Child Lang. **29**, 225–250 (2002).
[10]Y. Choi and R. Mazuka, "Young children's use of prosody in sentence parsing," J. Psycholinguist. Res. **32**, 197–217 (2003).
[11]M. 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).
[12]A. Gout, A. Christophe, and J. L. Morgan, "Phonological phrase boundaries constrain lexical access II. Infant data," J. Mem. Lang. **51**, 548–567 (2004).
[13]E. K. Johnson, "Infants use prosodically conditioned acoustic-phonetic cues to extract words from speech," J. Acoust. Soc. Am. **123**, EL144–EL148 (2008).
[14]A. de Carvalho, I. Dautriche, and A. Christophe, "Preschoolers use phrasal prosody online to constrain syntactic analysis," Dev. Sci. **19**, 235–250 (2016).
[15]W. E. Cooper and J. Paccia-Cooper, *Syntax and Speech* (Harvard University Press, Boston, MA, 1980).
[16]M. 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).
[17]A. W. F. Huggins, "Just noticeable differences for segment duration in natural speech," J. Acoust. Soc. Am. **51**, 1270–1278 (1972).
[18]T. F. Jaeger, "Categorical data analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models," J. Mem. Lang. **59**, 434–446 (2008).
[19]L. Gleitman, "The structural sources of verb meanings," Lang. Acquis. **1**, 3–55 (1990).
[20]A. 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).
[21]S. 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).
[22]A. 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).

Chapter 5

Phrasal prosody constrains syntactic analysis in toddlers

Cognition 163 (2017) 67–79

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Original Articles

Phrasal prosody constrains syntactic analysis in toddlers

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

[a] *Laboratoire de Sciences Cognitives et Psycholinguistique, DEC-ENS/EHESS/CNRS, Ecole normale supérieure – PSL Research University, Paris, France*
[b] *Maternité Port-Royal, AP-HP, Faculté de Médecine Paris Descartes, France*
[c] *School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, United Kingdom*
[d] *Department of Linguistics, University of California, Los Angeles, USA*

ARTICLE INFO

Article history:
Received 16 June 2016
Revised 7 December 2016
Accepted 28 February 2017
Available online 10 March 2017

Keywords:
Phrasal prosody
Language acquisition
Infants speech perception
Syntactic ambiguity resolution
Parsing
Sentence processing
Eye movements

ABSTRACT

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

1. Introduction

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

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

* Corresponding author at: Laboratoire de Sciences Cognitives et Psycholinguistique, Ecole normale supérieure, 29 rue d'Ulm – P.J, 75005 Paris, France.
E-mail address: x.de.carvalho@gmail.com (A. de Carvalho).

http://dx.doi.org/10.1016/j.cognition.2017.02.018
0010-0277/

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

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

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

A long series of studies shows that infants develop an impressive expertise with prosody from their first days of life. Newborns are able to exploit rhythmic information to discriminate between languages (Mehler et al., 1988; Nazzi, Bertoncini, & Mehler, 1998); from 4.5 months onwards, infants are sensitive to the coherence of prosodic constituents (Gerken, Jusczyk, & Mandel, 1994; Hirsh-Pasek et al., 1987; Jusczyk, Hohne, & Mandel, 1995; Männel & Friederici, 2009; Soderstrom, Seidl, Nelson, & Jusczyk, 2003), they show better recognition and memory for segments that correspond to whole prosodic units than for those which span prosodic boundaries (Mandel, Jusczyk, & Nelson, 1994; Nazzi, Iakimova, Bertoncini, Frédonie, & Alcantara, 2006) and they can rely on prosodic cues to segment the speech stream into words and constrain their lexical access (Gout, Christophe, & Morgan, 2004; Johnson, 2008; Millotte et al., 2010; Shukla, White, & Aslin, 2011). All of these findings, together with the reliable relationship between prosodic and syntactic structures, suggest that toddlers might be able to use phrasal prosody, not only to facilitate memory and lexical access, but also to constrain their syntactic analysis (see Christophe et al., 2008; Hawthorne & Gerken, 2014; Massicotte-Laforge & Shi, 2015; Morgan & Demuth, 1996; Morgan, 1986; Shi, 2014).

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

2. Experiment 1

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

To examine whether 28-month-olds exploit phrasal prosody to constrain their syntactic analysis, an intermodal preferential looking task with an eye-tracker was designed. Toddlers listened to the beginnings of these ambiguous sentences while watching two images displayed side-by-side on a TV screen: one associated with the noun interpretation of the ambiguous target word (e.g., a mouse) and the other one with the verb interpretation (e.g., a baby smiling). Their looking behavior was measured with an eye-tracker. If toddlers are able to take into account the prosodic structure of these sentences when conducting their syntactic computations, we expect them to look more often toward the noun picture when listening to sentences in the *noun prosody condition* than to sentences in the *verb prosody condition.*

2.1. Method

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

2.1.1. Participants

Forty toddlers, from 27.6 (months.days) to 28.28, with a mean of 27.26 (*SD* = 0.5, 19 girls) participated in this experiment. An additional four children participated in the study but were not included in the final analysis because of fussiness during the

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

experiment resulting in more than 50% (4 out of 8) unusable test trials with missing eye-tracking data. All participants were monolingual native French speakers. Parents signed an informed consent form. This research was approved by the local ethics committee.

2.1.2. Materials

Eight pairs of French noun-verb homophones likely to be known to young children (Kern, 2007; Veneziano & Parisse, 2010, 2011) were selected to create eight pairs of experimental sentences. For each pair of homophones, two sentences were created: one using the ambiguous word as a noun (the noun prosody condition, e.g. [*Le bébé*$_{ADJ}$ ***souris***$_{NOUN}$] [*a bien mangé*] – [The baby$_{ADJ}$ **mouse**$_{NOUN}$] [ate well]) and a second one using the ambiguous word as a verb (the verb prosody condition, e.g., [*Le bébé*$_{NOUN}$] [***sourit***$_{VERB}$ *à sa maman*] – [The baby$_{NOUN}$] [**smiles**$_{VERB}$ to his mom]; see Appendix A for a complete list of test sentences). Sentences uttered in the noun prosody condition had a prosodic boundary after the ambiguous target word and sentences uttered in the verb prosody condition had a prosodic boundary before the target word, consistent with theoretical descriptions of the relationship between prosodic and syntactic boundaries (e.g., Jun, 2005; Nespor & Vogel, 1986). A female French native speaker (the last author) recorded all the sentences in a child-friendly register. Note that the prosodic boundaries associated with each prosodic condition were found to be naturally produced by naïve adult native speakers, even when they were not aware of the syntactic ambiguity (Millotte et al., 2007). To estimate toddlers' knowledge of the ambiguous words, the parents of the participants in this experiment filled a short questionnaire. Overall, most toddlers understood most of the words used in this study (mean number of words comprehended: 13.8 out of 16; range: 10–16).

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

To ensure that prosodic cues would be the only information available to participants to determine whether the ambiguous word was a noun or a verb,[2] each test and each filler sentence was cut at the offset of the target word, and its end replaced by 1000 ms of babble noise obtained by superimposing the end of all filler sentences. Thus, there was no lexical disambiguating information following the ambiguous word.[3]

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

2.1.3. Acoustic analyses

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

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

The analysis of pitch contours also revealed significant differences between conditions, consistent with the literature describing French as having a tendency for a rising pitch contour towards the end of prosodic units (Di Cristo, 2000; Welby, 2003; Welby, 2006). A greater pitch rise was observed on the target word in the noun prosody condition (+127 Hz) compared to the verb prosody condition (+69 Hz). This difference is due to the fact that in the noun prosody condition the target word was in a phrase-final position, while in the verb prosody condition it was placed at the beginning of a phrase. For the same reasons, the word preceding the target word (e.g., "bébé") had a greater rise in pitch in the verb prosody condition (+184 Hz) than in the noun prosody condition (+21 Hz). All of these differences were significant (see Table 1).

2.1.4. Apparatus and procedure

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

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

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

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

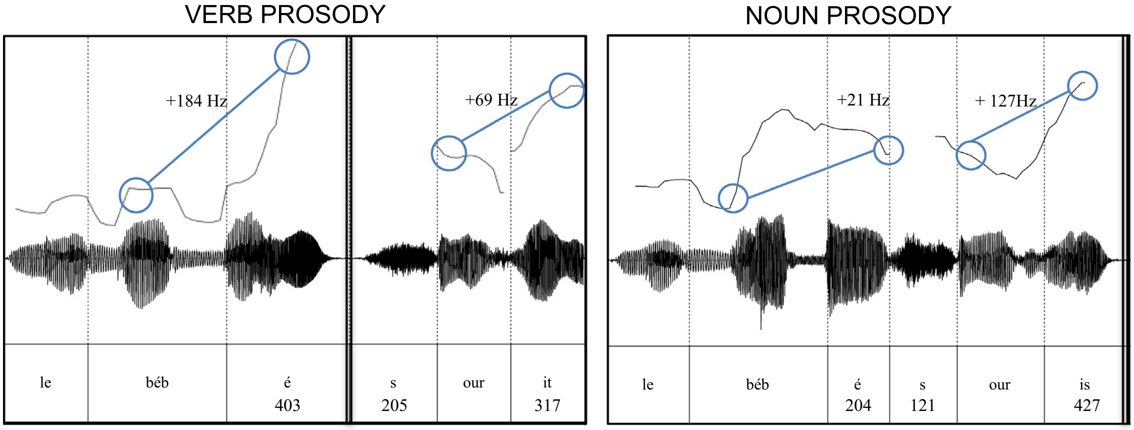


Fig. 1. Mean duration of the different segments, and pitch contours in the ambiguous region. Prosodic boundaries are represented with thick black lines. Blue circles delimit the areas where pitch analyses were performed, subtracting the pitch value at the beginning from the pitch value at the end of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

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

Duration analyses – mean duration in ms (standard error of the mean)			
Dependent variable	**Noun prosody** [le bébé suʁi]	**Verb prosody** [le bébé][suʁi	**Analysis (2-tailed t-tests)**
Rhyme - word preceding Target (e.g., -/e/ from "*bébé*")	204 (22)	403 (50.4)	t (7) = −3.85, $p < .01^{**}$
Onset - Target word (e.g., /s/ from "/suʁi/"	121 (9.2)	205 (16.2)	t (7) = −5.02, $p < .01^{**}$
Rhyme – Target word (e.g., -/i/ from "/suʁi/")	427 (50.6)	317 (34.9)	t (7) = 3.77, $p < .01^{**}$
Pitch analyses – mean pitch change, in Hz, from the beginning to the end of the target words (standard error of the mean).			
Dependent variable	**Noun prosody** [le bébé suʁi]	**Verb prosody** [le bébé][suʁi	**Analysis (2-tailed t-tests)**
Word preceding Target (e.g., last pitch value at the last vowel from "bébé" minus first pitch value from the first vowel of "bébé")	21 (20.4)	184 (38.1)	t (7) = −5.29, $p < .01^{**}$
Target word (e.g., last pitch value of "-i" from "souri" minus first pitch value of "/u/" from "/suri/")	127 (23.5)	69 (25.8)	t (7) = 4.47, $p < .01^{**}$

started the test block, composed of eight ambiguous test sentences and four filler sentences.

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

2.1.5. Data processing and analysis

Toddlers' eye-gaze towards the images was recorded by an Eyelink-1000 while they listened to the test sentences, with a time-sample collected every 2 ms. Before statistical analysis, the data was down-sampled by a factor of 10, by averaging the data from 10 adjacent samples, so that the final sampling rate was one sample every 20 ms. Thirty-nine trials out of 320 were removed from the statistical analysis (17 in the noun condition and 22 in the verb condition), because more than 25% of the data frames were missing between the onset of the test sentences and the end of the ambiguous word. The eye-gaze analysis uses the proportion of fixations toward the noun image as a dependent variable, because fixations to noun vs. verb image in this task are complementary (apart from the time spent looking away). To find the time-window(s) which exhibited a significant difference between conditions, a cluster-based permutation analysis was conducted (as in Dautriche, Swingley, & Christophe, 2015; de Carvalho et al., 2016; Hahn, Snedeker, & Rabagliati, 2015; Von Holzen & Mani, 2012; see Maris & Oostenveld, 2007, for a formal presentation of the analysis itself). This analysis allows us to test for the effect of Condition without inflating the rate of Type I error. It proceeds in two phases. First, for each time point, a paired two-tailed t-test testing for the effect of Condition (noun prosody vs. verb prosody) is conducted (on the proportion of looks toward the noun picture). Adjacent time points with a t-value greater than some predefined threshold (here, $t = 1.5$)[4] are grouped together into a cluster. The size of the cluster is defined as the sum of the t values at each time point within the cluster. Second, to obtain the probability of observing a cluster of that size by chance, we conducted 1000 simulations where we randomly shuffled the conditions (noun prosody, verb prosody) for each trial. For each simulation, we calculated the size of the biggest cluster identified with the same procedure that was applied to the real data. A cluster of adjacent time points from the real data shows a significant

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

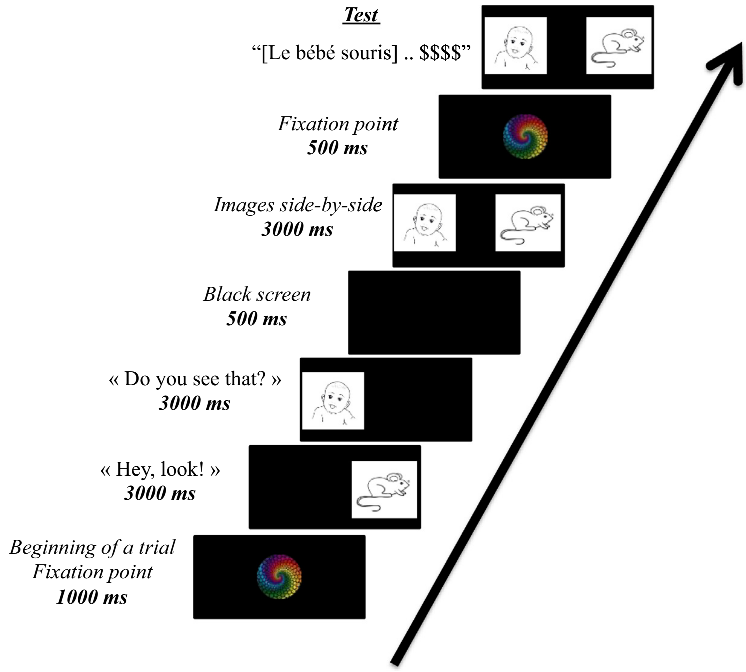


Fig. 2. Time-course of a trial.

effect of condition if its size is greater than the size of the largest cluster found in 95% of the simulations (ensuring a p-value of .05). This analysis was conducted on a time-window extending from −700 ms before the onset of the ambiguous word until 2000 ms after the onset of the ambiguous word. Plots of eye-gaze data were performed with the ggplot2 package (Wickham, 2009).

2.2. Results

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

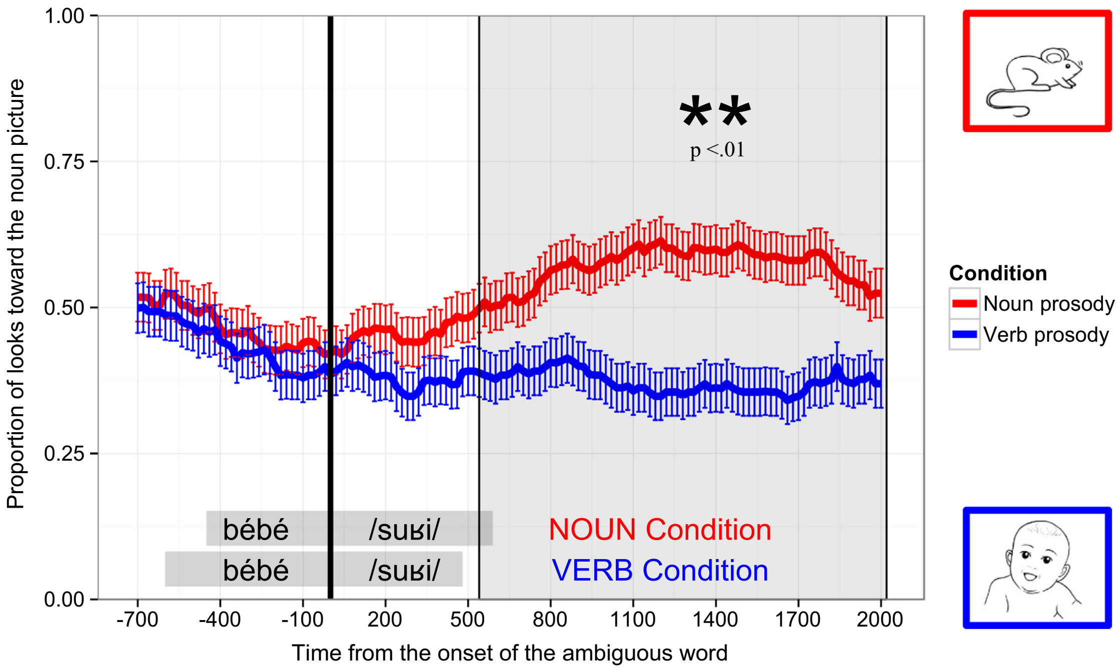


Fig. 3. Proportion of looks toward the noun image, time-locked to the onset of the ambiguous word (thick vertical line), for the noun prosody condition (red curve) and the verb prosody condition (blue curve). Error bars represent the standard error of the mean. A nonparametric cluster-based permutation test revealed a significant difference between the noun prosody and the verb prosody conditions, starting around the offset of the ambiguous target word (gray time-window; from 540 ms after the beginning of the critical word, '**'$p < 0.01$).

2008; Trueswell & Gleitman, 2007; Trueswell, Sekerina, Hill, & Logrip, 1999).

Visual inspection of the data reveals that toddlers tended to look more toward the verb image at the beginning of the sentences. For instance, at the onset of the target word, vertical black line, both curves are at 40% looks toward the noun picture, perhaps revealing a simple preference for looking at human beings in the verb picture. Crucially, however, when listening to the beginning of a noun sentence, toddlers increased their looks toward the noun picture, from around the offset of the ambiguous word, thus switching their eye-gaze toward the correct image after hearing the relevant prosodic information. In contrast, when they were listening to the beginning of a verb sentence, toddlers increased their looks toward the verb picture.

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

2.3. Discussion

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

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

This effect mirrors previous results obtained with adults and preschoolers in French (de Carvalho et al., 2016), although in the present study toddlers seem to be slower than 4-year-old children, by about 300 ms. This difference could be due to the fact that preschoolers (and adults) exploit prosodic information more efficiently than toddlers, or simply to the fact that toddlers have less attentional skills than their older counterparts (leading to noisier behavior). Although we cannot disentangle between these interpretations, the main result is that, just like adults and preschoolers, toddlers who are still in the process of learning the syntax of their language, can use phrasal prosody online to access the syntactic structure of sentences and constrain their syntactic analysis. Given that prosodic phrase boundaries are perceived and exploited by infants from six months onwards (Gerken et al., 1994; Gout et al., 2004; Shukla et al., 2011; Soderstrom et al., 2003), it is possible that even younger toddlers might be able to use phrasal prosody as cue to recover the syntactic structure of sentences.

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

3. Experiment 2

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

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

3.1. Method

3.1.1. Participants

Sixty-four toddlers participated in this experiment. They were all monolingual native French speakers and were divided into two age groups (with 32 toddlers in each age group): the 20-month-old group, ranging in age from 19.0 (months.days) to 21.3, with a mean age of 19.19 ($SD = 0.6$; 14 girls) and the 28-month-old group, ranging in age from 26.19 to 28.27, with a mean of 27.20 (SD = 0.6; 18 girls). Within each age group, half of the participants heard the test sentences in the noun prosody condition and half heard them in the verb prosody condition. An additional twenty-six children completed the experiment (eleven 28-m.o and fifteen 20-m.o) but they were not included in the final sample because of fussiness during the experiment resulting in more than 50% of trials with missing eye-tracking data (n = 19), because they cried (n = 4), or because of technical problems (n = 3). Parents signed an informed consent form. This research was approved by the local ethics committee.

3.1.2. Material

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

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

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

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

3.1.3. Acoustic analyses

In order to assess prosodic differences between the two prosodic conditions, acoustic measurements (duration and pitch) were conducted on the test sentences. The analysis of duration revealed a significant pre-boundary lengthening, as expected from the literature: the rhyme of the word preceding the target word (e.g., last vowel -/e/ from *bébé*) in the verb condition (where it was placed just before the prosodic phrase boundary) was lengthened by 211% compared this same segment in the noun condition (where it was placed in the middle of a prosodic unit; 395 vs 127 ms, see Table 2). A silent pause of 232 ms preceding the target word (i.e., between "bébé" and /suʁi/) was observed in the verb condition, while there was no pause between these words in the noun condition. Additionally, a phrase-initial strengthening was observed: the onset of the target word in the verb condition (205 ms, phrase-initial position) was lengthened by 88% compared to the noun condition (109 ms, phrase-medial position). The rhyme of the target words (e.g., -/i/ from /suʁi/) were utterance-final in both conditions (contrary to Experiment 1); it was lengthened by 49% in the verb condition relative to the noun condition (480 vs 383 ms), possibly because the verb was alone in its prosodic unit.

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

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

Duration analyses – mean duration in ms (standard error)

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

Pitch analyses – Mean pitch change, in Hz, from the beginning to the end of the target words (standard error of the mean).

Dependent variable	**Noun prosody** [le bébé suʁi]	**Verb prosody** [le bébé][suʁi]	**Analysis (2-tailed t-tests)**
Word preceding Target (e.g., last pitch value at the last vowel from "bébé" minus first pitch value from the first vowel of "bébé")	−53 (16.4)	185 (28.1)	t (7) = −14.32, p < .01**
Target word (e.g., last pitch value of "-i" from "souri" minus first pitch value of "/u/" from "/suʁi/")	118 (46.9)	29 (30.7)	t (7) = 1.94, p = .09

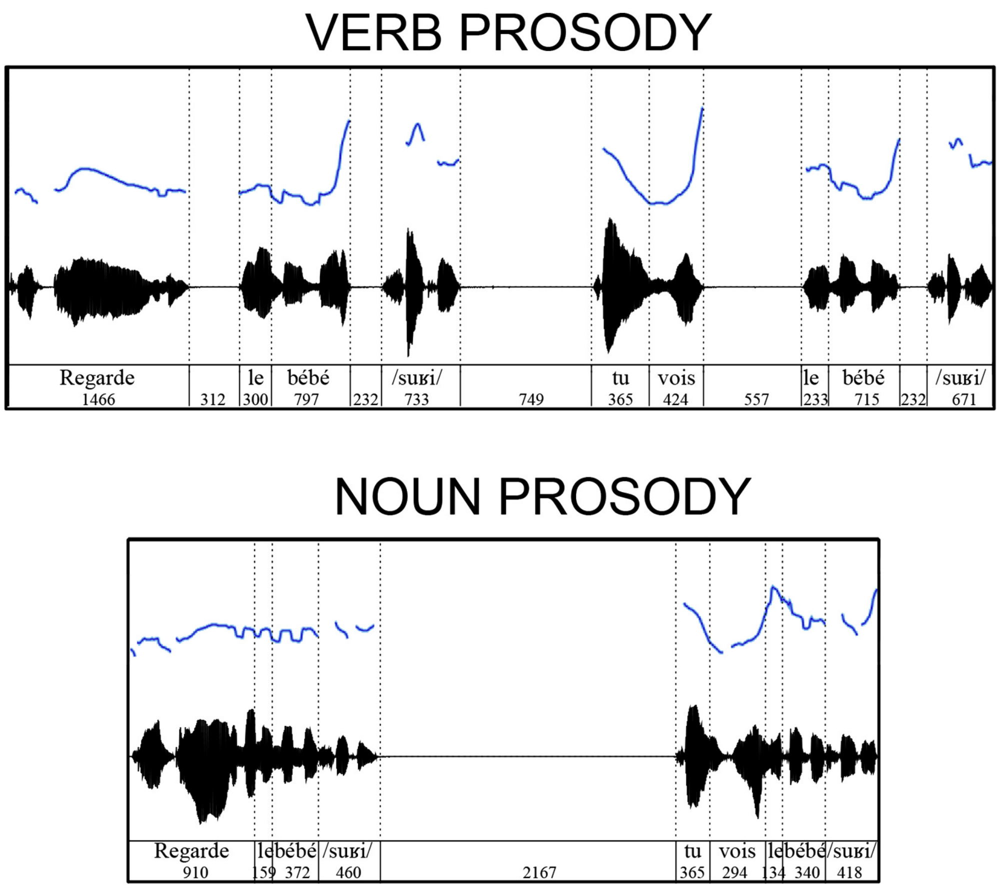


Fig. 4. Representation of the ambiguous sentences, with, from bottom to top, the mean duration (in ms) of the different segments, the transcription, the waveform, and the pitch contour. Note that while the waveforms and the pitch contours correspond to the experimental sentences of the item /suʁi/, the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

3.1.4. Apparatus and procedure

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

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

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

3.1.5. Data processing and analysis

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

3.2. Results

Fig. 5 shows the proportion of looks toward the noun image for toddlers in the noun prosody condition (red curve) and in the verb prosody condition (blue curve), time-locked to the beginning of the first onset of the ambiguous word, for the 20-month-old group (A) and for the 28-month-old group (B).

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

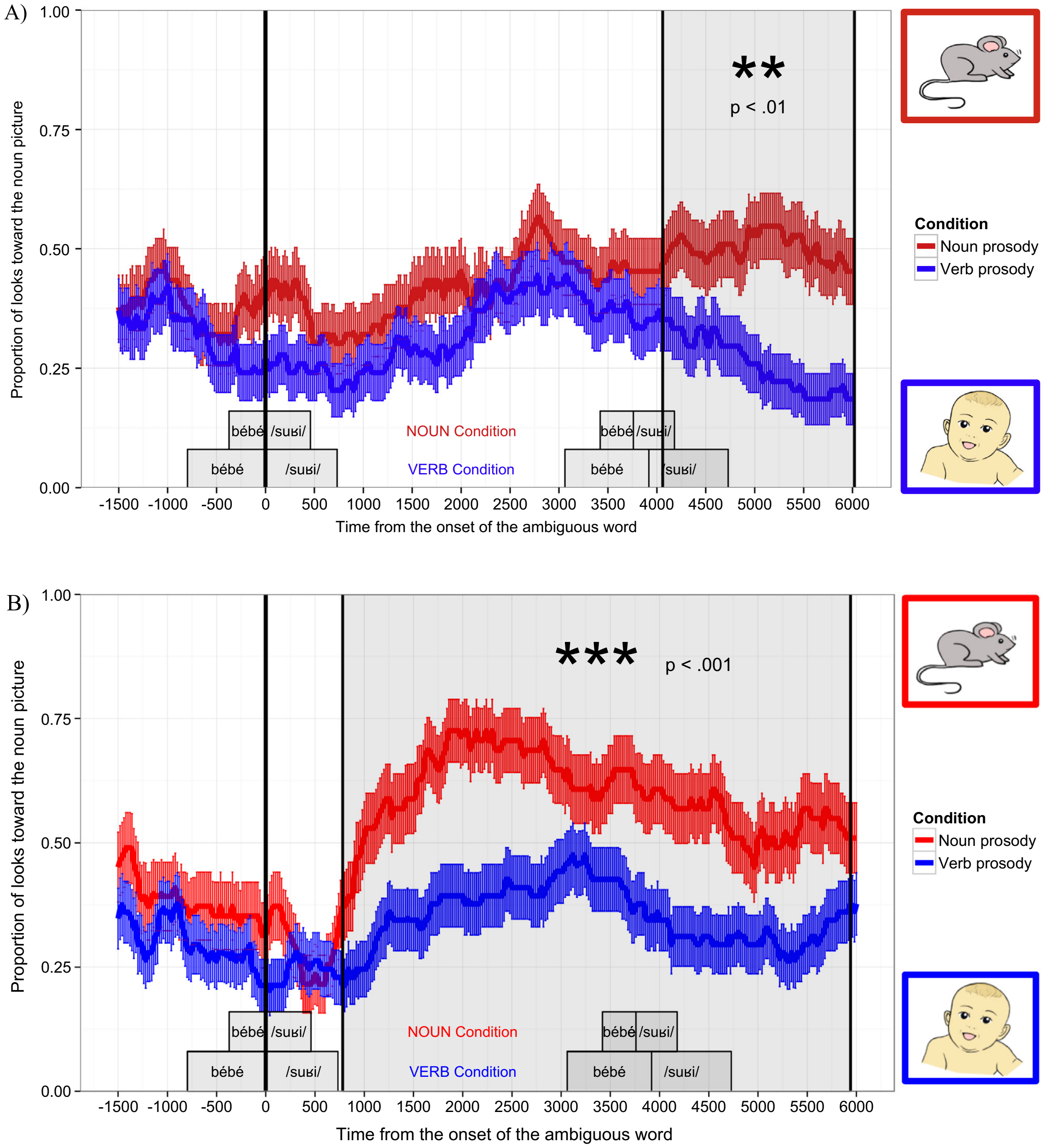


Fig. 5. Proportion of looks toward the noun image, time-locked to the onset of the ambiguous word (vertical black line) for 20-month-olds (A), and 28-month-olds (B), for children in the noun prosody condition (red curve) and in the verb prosody condition (blue curve). Error bars represent the standard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

able to exploit prosodic information to guide their interpretation of the ambiguous target word.

The cluster-based analysis found a significant time-window where the proportion of looks toward the noun picture was significantly different from children in the noun prosody condition compared to children in the verb prosody condition, for both age groups: 28-month-olds (from 780 ms after the onset of the first repetition of the critical word; $p < 0.001$), and 20-month-olds (from 4060 ms after the onset of the first critical word or about 300 ms after the onset of the second critical word; $p < 0.01$). These results show that, despite their speed difference, both 20- and 28-month-olds looked more towards the noun picture in the noun prosody condition than in the verb prosody condition.

3.3. Discussion

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

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

4. General discussion

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

To succeed in our experiments, toddlers may have used phrasal prosody and function words together to constrain their syntactic analysis. This hypothesis is based on the fact that while the perception of prosodic boundaries in our experiments allowed toddlers to group words into syntactic constituents, and informed them about the location of syntactic boundaries, the prosodic boundaries *per se* do not directly provide the syntactic labels of constituents (e.g. noun phrase, verb phrase). To interpret the homophone as a noun or a verb, toddlers may have used the additional information carried by function words[5], together with the prosodic structure of sentences. For instance, in Experiment 1, when participants heard a sentence such as [*le bébé*] [/*suʁi*/ ..., the prosodic boundary before the target word signaled the presence of two prosodic units. Given that the first unit (e.g., [*le bébé*]) started with an article (e.g., *le* – the), this unit could be identified as a noun phrase (e.g., [Le_{DET} $bébé_{NOUN}$]$_{NP}$ - [The_{DET} $baby_{NOUN}$]$_{NP}$). Having identified the first unit as a full noun phrase, toddlers might expect it to be followed by a verb phrase, which allows them to rapidly identify the ambiguous word (e.g., /suʁi/) as a verb. In the noun prosody condition in contrast, given that all three words appeared together into one single prosodic unit starting with an article (e.g., [*le bébé* /suʁi/]), this information led toddlers to interpret the entire constituent as a noun phrase, which entailed that /suʁi/ had to be interpreted as a noun. Similarly, in Experiment 2, the presence of a prosodic boundary just before the ambiguous word triggered a verb interpretation, while the ambiguous word was identified as a noun when it belonged to the same prosodic unit as the first three words ([*Regarde le bébé souris*], 'look at the baby mouse'). It is important to note that the use of prosodic information to constrain syntactic analysis is not limited to the kind of syntactic ambiguity resolution featured in our experiments. The relationship between prosodic and syntactic structures is present in all sentences, whether or not they contain ambiguous words. For instance, in a sentence such as [The little cat] [jumps really high], listeners can perceive the prosodic boundary between the subject noun phrase and the verb phrase, as in many sentences that children hear in their everyday lives. In other words, although sentences containing homophones are useful to test listeners' abilities to rely on phrasal prosody to recover syntactic structure, listeners can learn the relationship between prosodic and syntactic structures from unambiguous everyday sentences.

Overall, the ability to use phrasal prosody and function words together helps infants to generate a first parse of the syntactic structure of sentences, and allows them to calculate the syntactic category of an ambiguous word. Note that toddlers seem not to be bothered by the noun-verb homophony, in these cases, because the critical words occur in disambiguating contexts (contrary to what has been proposed in the literature, e.g. Conwell & Morgan, 2012). We suspect that cross-category homophones such as these will most often appear in disambiguating contexts, and therefore not hinder children's language acquisition (see Dautriche, Fibla, & Christophe, 2015; Dautriche, 2015; Dautriche et al., 2015, for a fuller discussion of this aspect).

The ability to assign a syntactic category to a word according to its context may be extremely important during language acquisition, especially when children do not yet know the meanings of many words. Indeed, children may exploit the fact that an unknown word occurs in a noun context to infer that it probably refers to an object, while words occurring in verb contexts probably refer to actions (e.g., Gillette et al., 1999; Gleitman, 1990). For instance, He and Lidz (2017) showed that 18-month-olds (but not 14-month-olds) were able to infer that a novel word such as 'doke' referred to an object when listening to sentences such as "Look, it's a doke!", and that a novel word such as 'pratch' referred to an action when listening to sentences such as "Look! It's pratching!". However, not all content words are immediately preceded or followed by a disambiguating function word or morpheme as in "a doke" or "is doking" (e.g., in: "*The giant bears...*", *bears* can be either a noun or a verb). In such cases, a more sophisticated analysis in terms of syntactic constituents, signaled by prosodic boundaries, might be extremely informative for infants. For example, in a sentence like "[Do you see the baby blicks]?", infants might be able to infer that "blick" is a noun, referring to an object; but in a sentence such as: "[Do you see]? [The baby] [blicks]!" they might be able to infer that "blick" is a verb, referring to an action. Note that this hypothesis is rather plausible, since to correctly interpret the novel word "blick" as a noun or a verb in this situation, infants would need to exploit exactly the same kind of information they were shown to use in the present experiments.

Other recent findings support the importance of phrasal prosody for syntactic computations in toddlers, showing that prosody

[5] Function words have already been shown to be used by 18-month-olds to categorize neighbouring content words (e.g., Cauvet et al., 2014; He & Lidz, 2017; Höhle et al., 2004; Shi & Melançon, 2010).

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

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

To sum up, we provided evidence that from 20 months old, toddlers readily exploit the prosodic structure of an utterance to constrain its syntactic analysis, and access the meaning of an ambiguous target word. We showed that toddlers use phrasal prosody to segment the continuous speech stream into prosodic units, use them to infer the presence of syntactic constituent boundaries, and exploit function words and syntactic boundaries to assign a syntactic category to ambiguous words and recover their meanings. Given that at this age, toddlers are still in the process of learning their lexicon, this ability to assign a syntactic category to words depending on their context may help infants to constrain the acquisition of word meanings. These findings suggest that phrasal prosody plays an important role in language acquisition, since it provides access to a first-pass syntactic structure of sentences which may help infants to bootstrap language acquisition.

Acknowledgements

This research was supported by a PhD fellowship from the École normale supérieure to Alex de Carvalho, and by a an ESRC Future Research Leaders award (ES/N017404/1) to Isabelle Dautriche. It was also supported by grants from the Région Ile-de-France, Fondation de France, LabEx IEC (ANR-10-LABX-0087), IdEx PSL (ANR-10-IDEX-0001-02), as well as the ANR 'Apprentissages' (ANR-13-APPR-0012). The authors thank Anne-Caroline Fiévet, Luce Legros and Mélanie Rouelle for their assistance with recruitment of participants and data collection as well as all the parents and babies who participated in this study.

Appendix A. Experimental sentences of Experiment 1

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

Test sentences			
Pair of ambiguous word	Syntactic category	Target	Full sentence before acoustical mask
Fermer x la ferme	Verb	Ferme	La petite **ferme** le coffre à jouets *The small one **closes** the toy box*
to close x the farm	Noun		La petite **ferme** lui plait beaucoup *The small **farm** pleases him a lot*
Lire x le lit	Verb	Lit	Le grand **lit** souvent des histoires à son petit frère *The big one often **reads** stories to his younger brother*
to read x the bed	Noun		Le grand **lit** sera pour les parents *The big **bed** will be for the parents*
Marcher x la marche	Verb	Marche	La grande **marche** lentement toute la journée *The big one **walks** slowly all day long*
to walk x the stairs	Noun		La grande **marche** la fait tomber *The big **stair** makes her fall*
Moucher x la mouche	Verb	Mouche	La maman **mouche** le bébé malade *The mother **blows the nose of** the sick baby*
to nose x the fly	Noun		La maman **mouche** laisse son bébé tout seul *The mother **fly** leaves her baby alone*
Porter x la porte	Verb	Porte	La vieille **porte** sa montre à réparer *The old lady **carries** her watch to be repaired*
to carry x the door	Noun		La vieille **porte** sera réparée demain *The old **door** will be repaired tomorrow*
Montrer x la montre	Verb	Montre	La grande **montre** ses jouets à son frère *The big one **shows** her toys to her brother*
to show x the watch	Noun		La grande **montre** sera réparée demain *The big **watch** will be repaired tomorrow*

(*continued on next page*)

Appendix A (*continued*)

Test sentences			
Pair of ambiguous word	Syntactic category	Target	Full sentence before acoustical mask
Sourire x la souris	Verb	[suri]	Le bébé **sourit** à sa maman *The baby* ***smiles*** *to his mom*
to smile x the mice	Noun		Le bébé **souris** a bien mangé *The baby* ***mouse*** *ate well*
Pêcher x les pêches	Verb	[pɛʃ]	Les grosses **pêchent** mon poisson préféré pour le dîner *The fat ones* ***fish*** *my favorite fish for dinner*
to fish x the peaches	Noun		Les grosses **pêches** me font très envie *The big* ***peaches*** *tempt me a lot*

Appendix B. Experimental sentences of Experiment 2

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

References

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

Bernal, S., Lidz, J., Millotte, S., & Christophe, A. (2007). Syntax constrains the acquisition of verb meaning. *Language Learning and Development, 3*(4), 325–341. http://dx.doi.org/10.1080/15475440701542609.

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

Cauvet, E., Limissuri, R., Millotte, S., Skoruppa, K., Cabrol, D., & Christophe, A. (2014). Function words constrain on-line recognition of verbs and nouns in French 18-month-olds. *Language Learning and Development, 10*(1), 1–18. http://dx.doi.org/10.1080/15475441.2012.757970.

Christodoulopoulos, C., Roth, D., & Fisher, C. (2016). An incremental model of syntactic bootstrapping. In *Proceedings of the seventh workshop on cognitive aspects of computational language learning*.

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

Conwell, E., & Morgan, J. L. (2012). Is it a noun or is it a verb? resolving the ambicategoricality problem. *Language Learning and Development, 8*(2), 87–112. http://dx.doi.org/10.1080/15475441.2011.580236.

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

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

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

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

de Carvalho, A., Dautriche, I., & Christophe, A. (2016a). Preschoolers use phrasal prosody online to constrain syntactic analysis. *Developmental Science, 19*(2), 235–250. http://dx.doi.org/10.1111/desc.12300.

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

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

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

Ferguson, B., Graf, E., & Waxman, S. R. (2014). Infants use known verbs to learn novel nouns: evidence from 15- and 19-month-olds. *Cognition, 131*(1), 139–146. http://dx.doi.org/10.1016/j.cognition.2013.12.014.

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

Fisher, C. (1996). Structural limits on verb mapping: the role of analogy in children's interpretations of sentences. *Cognitive Psychology, 31*(1), 41–81. http://dx.doi.org/10.1006/cogp.1996.0012.

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

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

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *The 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.
Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition, 73*(2), 135–176.
Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1*, 3–55.
Gout, A., Christophe, A., & Morgan, J. L. (2004). Phonological phrase boundaries constrain lexical access II. Infant data. *Journal of Memory and Language, 51*(4), 548–567. http://dx.doi.org/10.1016/j.jml.2004.07.002.
Gutman, A., Dautriche, I., Crabbé, B., & Christophe, A. (2015). Bootstrapping the syntactic bootstrapper: probabilistic labeling of prosodic phrases. *Language Acquisition, 22*(3), 285–309. http://dx.doi.org/10.1080/10489223.2014.971956.
Hahn, N., Snedeker, J., & Rabagliati, H. (2015). Rapid Linguistic Ambiguity Resolution in Young Children with Autism Spectrum Disorder: Eye Tracking Evidence for the Limits of Weak Central Coherence. *Autism Research, 8*(6), 717–726. http://dx.doi.org/10.1002/aur.1487.
Hallé, P. A., Durand, C., & de Boysson-Bardies, B. (2008). Do 11-month-old French infants process articles? *Language and Speech, 51*(Pt 1–2), 23–44.
Hawthorne, K., & Gerken, L. (2014). From pauses to clauses: prosody facilitates learning of syntactic constituency. *Cognition, 133*(2), 420–428. http://dx.doi.org/10.1016/j.cognition.2014.07.013.
Hawthorne, K., Rudat, L., & Gerken, L. (2016). Prosody and the acquisition of hierarchical structure in toddlers and adults. *Infancy, 1–22*. http://dx.doi.org/10.1111/infa.12130.
He, A. X., & Lidz, J. (2017). Verb learning in 14- and 18-month-old English-learning infants. *Language Learning and Development*. 5441 (March).
Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Cassidy, K. W., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition, 26*, 269–286.
Höhle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science, 6*(2), 122–127. http://dx.doi.org/10.1111/1467-7687.00261.
Höhle, B., Weissenborn, J., Kiefer, D., Schulz, A., & Schmitz, M. (2004). Functional elements in infants' speech processing: the role of determiners in the syntactic categorization of lexical elements. *Infancy, 5*(3), 341–353. http://dx.doi.org/10.1207/s15327078in0503_5.
Johnson, E. K. (2008). Infants use prosodically conditioned acoustic-phonetic cues to extract words from speech. *The Journal of the Acoustical Society of America, 123*(6). http://dx.doi.org/10.1121/1.2908407. EL144-EL148.
Jun, S. (2005). Prosodic typology. In S. Jun (Ed.). *Prosodic typology: the phonology of intonation and phrasing* (pp. 430–458). Oxford University Press.
Jun, S., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus, 14*(1), 147–172. http://dx.doi.org/10.1515/prbs.2002.002.
Jusczyk, P. W., Hohne, E., & Mandel, D. (1995). Picking up regularities in the sound structure of the native language. In W. Strange (Ed.), *Speech perception and linguistic experience: theoretical and methodological issues in cross-language speech research* (pp. 91–119). Timonium, MD: Yorck Press.
Kern, S. (2007). Lexicon development in French-speaking infants. *First Language, 27*(3), 227–250. http://dx.doi.org/10.1177/0142723706075789.
Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language, 40*(2), 153–194. http://dx.doi.org/10.1006/jmla.1998.2620.
Landau, B., & Gleitman, L. (1985). *Language and experience: evidence from the blind child*. Cambridge, MA: Harvard University Press.
Mandel, D. R., Jusczyk, P. W., & Nelson, D. G. (1994). Does sentential prosody help infants organize and remember speech information? *Cognition, 53*(2), 155–180.
Männel, C., & Friederici, A. D. (2009). Pauses and intonational phrasing: ERP studies in 5-month-old German infants and adults. *Journal of Cognitive Neuroscience, 21*(10), 1988–2006. http://dx.doi.org/10.1162/jocn.2009.21221.
Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190. http://dx.doi.org/10.1016/j.jneumeth.2007.03.024.
Massicotte-Laforge, S., & Shi, R. (2015). The role of prosody in infants' early syntactic analysis and grammatical categorization. *Journal of Acoustical Society of America, 138*(October), 441–446. http://dx.doi.org/10.1121/1.4934551.
Mehler, J., Jusczyk, P., Lamsertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29*, 143–178.
Millotte, S., René, A., Wales, R., & Christophe, A. (2008). Phonological phrase boundaries constrain the online syntactic analysis of spoken sentences. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 34*(4), 874–885. http://dx.doi.org/10.1037/0278-7393.34.4.874.
Millotte, S., Wales, R., & Christophe, A. (2007). Phrasal prosody disambiguates syntax. *Language and Cognitive Processes, 22*(6), 898–909. http://dx.doi.org/10.1080/01690960701205286.
Millotte, S., Margules, J., Bernal, S., Dutat, M., & Christophe, A. (2010). Phrasal prosody constrains word segmentation in French 16-month-olds. *Journal of Portuguese Linguistics, 9–2*(10–1), 67–86.
Morgan, J. (1986). *From simple input to complex grammar*. Cambridge MA: MIT Press.
Morgan, J., & Demuth, K. (1996). Signal to syntax: Bootstrapping from speech to grammar in early acquisition (Eribaum). Mahwah, NJ.
Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology. Human Perception and Performance, 24*(3), 756–766.
Nazzi, T., Iakimova, G., Bertoncini, J., Frédonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: emerging evidence for crosslinguistic differences. *Journal of Memory and Language, 54*, 283–299.
Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris.
Shafer, V., Shucard, D., Shucard, J., & Gerken, L. (1998). An electrophysiological study of infants' sensitivity to the sound patterns of English speech. *Journal of Speech, Language, and Hearing Research, 41*, 874–886.
Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research, 25*(2), 193–247.
Shi, R. (2014). Functional morphemes and early language acquisition. *Child Development Perspectives, 8*(1), 6–11. http://dx.doi.org/10.1111/cdep.12052.
Shi, R., & Melançon, A. (2010). Syntactic Categorization in French-Learning Infants. *Infancy, 15*(5), 517–533. http://dx.doi.org/10.1111/j.1532-7078.2009.00022.x.
Shi, R., Werker, J., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy, 10*(2), 187–198.
Shukla, M., White, K. S., & Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences, 108*(15), 6038–6043. http://dx.doi.org/10.1073/pnas.1017617108.
Snedeker, J., & Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults)⁎. *Journal of Memory and Language, 58*, 574–608. http://dx.doi.org/10.1016/j.jml.2007.08.001.
Soderstrom, M., Blossom, M., Foygel, R., & Morgan, J. (2008). Acoustical cues and grammatical units in speech to two preverbal infants. *Journal of Child Language, 35*(4), 869–902. http://dx.doi.org/10.1017/S0305000908008763.
Soderstrom, M., Seidl, A., Nelson, D. G., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: evidence from prelinguistic infants. *Journal of Memory and Language, 49*(2), 249–267. http://dx.doi.org/10.1016/S0749-596X(03)00024-X.
Swingley, D., & Aslin, R. (2000). Spoken word recognition and lexical representation in very young children. *Cognition, 76*, 147–166.
Trueswell, J. C. (2008). Using eye movements as a developmental measure within psycholinguistics. In I. Sekerina, E. Fernandez, & H. Clahsen (Eds.), *Developmental psycholinguistics: on-line methods in children's language processing* (pp. 73–96). Amsterdam: John Benjamins.
Trueswell, J. C., & Gleitman, L. R. (2007). Learning to parse and its implications for language acquisition. In M. G. Gaskell (Ed.), *Oxford handbook of psycholinguistics* (pp. 1–39). Oxford: Oxford University Press.
Trueswell, J., Sekerina, I., Hill, N., & Logrip, M. (1999). The kindergarten-path effect: studying on-line sentence processing in young children. *Cognition, 73*(2), 89–134.
Veneziano, E., & Parisse, C. (2010). The early comprehension of noun-verb distinction in French: an experimental method. In *XVIIth biennial international conference on infant studies*. Baltimore, Maryland, United States.
Veneziano, E., & Parisse, C. (2011). Retrieving the meaning of words from syntactic cues: a comprehension study of 2 to 4 yrs old French-speaking children. In *International conference on the study of child language*, IASCL 2011. Montreal, Canada.
Von Holzen, K., & Mani, N. (2012). Language nonselective lexical access in bilingual toddlers. *Journal of Experimental Child Psychology, 113*(4), 569–586. http://dx.doi.org/10.1016/j.jecp.2012.08.001.
Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology, 59*(1), 67–95. http://dx.doi.org/10.1016/j.cogpsych.2009.02.001.
Welby, P. (2003). *The slaying of lady mondegreen: being a study French of tonal association and alignment and their role in speech segmentation*. Columbus: Ohio State University.
Welby, P. (2006). French intonational structure: evidence from tonal alignment. *Journal of Phonetics, 34*(3), 343–371. http://dx.doi.org/10.1016/j.wocn.2005.09.001.
Wickham, H. (2009). ggplot2. In R. Gentleman, K. Hornik, & G. Parmigiani (Eds.), New York, NY: Springer, New York. doi: http://dx.doi.org/10.1007/978-0-387-98141-3.
Yuan, S., & Fisher, C. (2009). "Really? she blicked the baby?": Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science, 20*, 619–626.
Yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the nouns: simple structural cues to verb meaning. *Child Development, 83*(4), 1382–1399.

Summary and Discussion

Chapters 3, 4 and 5 together show that from 20 months on, children are able to exploit phrasal prosody and function words to constrain their syntactic analysis of ambiguous words in French and in English. Children were able to use phrasal prosody to recover the syntactic structure of sentences and function words to predict the syntactic category of an ambiguous word, which in turn allowed them to access its possible meaning. In Chapter 3 and 5 we saw that when children listened to a sentence such as "[Tu vois la petite marche?] – Do you see the little stair?, they interpreted "marche" as a noun, but in a sentence such as [Tu vois?] [la petite] [marche]! – Do you see? The little girl walks!, they interpreted "marche" as a verb (de Carvalho et al., 2016 *DevSci*; 2017 *Cognition*). Importantly, in Chapter 4, I demonstrated that this ability is also present in English with sentences such as "Do you see the baby flies" in which preschoolers interpreted "flies" as a noun or as a verb depending on the prosodic information they heard (de Carvalho et al., 2016 *JASA*). These results show that young children can use prosodic boundaries to infer the presence of a syntactic constituent boundary, and in turn use that information to figure out the syntactic category of an ambiguous word.

The discrepancy between these recent studies in French (Chapter 3 and 5) and the previous literature showing children's failure to use prosody to constrain their syntactic analysis, rests on the relationship between prosodic and syntactic structures that was used in previous studies and on the reliability of the prosodic information used. As we discussed in Chapter 4, 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, which shows that the link between prosodic and syntactic structure was not sufficiently systematic in the structures that were tested. This may explain why English-learning preschoolers had difficulties using prosody to disambiguate between the two possible interpretations. In contrast, the sentences used here have different default prosodic structures, with the prosodic boundary falling either before or after the critical ambiguous word. Thus we believe that the relationship between prosodic and syntactic structures was more systematically marked in our studies than in the previous studies. In this case, we observed that children were just as sensitive to prosody as one might expect.

Note that the disambiguating prosodic information that children used to succeed in our tasks (i.e., the prosodic boundary between the noun phrase and the verb phrase) is part of the normal prosodic structure of sentences, and it is present even in non-ambiguous sentences (e.g. [the little frog] [eats a lot of food]). This may explain our participants' remarkable ability to integrate prosodic information in their computation of syntactic structure.

We used familiar homophones as a test case, to assess the role of phrasal prosody and function words in children's syntactic analysis. However, we do not think that prosody is only used by children when they encounter an ambiguous word or sentence. Although we used ambiguous sentences to test listeners' abilities to rely on phrasal prosody to recover syntactic structure, we believe that infants can learn the relationship between prosodic and syntactic structures from unambiguous everyday sentences, and that they can potentially use this information to parse sentences that contain unknown words.

Taken together, chapters 3, 4 and 5 provide evidence that during the first steps of language acquisition, toddlers are already able to exploit the prosodic structure of sentences to recover their syntactic structure and that they exploit function words to predict the syntactic category of upcoming words, an ability which would be extremely useful to assign a syntactic category to novel words, and constrain their possible meaning. This hypothesis will be investigated in the second part of this thesis (Part 2).

Part 2

Can infants use phrasal prosody and function words to constrain the acquisition of word meanings?

Part 2: Can infants use phrasal prosody and function words to learn word meanings?

The ability to exploit information from the speech signal to access the syntactic structure of sentences, and generate predictions regarding the syntactic category of upcoming words, can be extremely useful for infants during language acquisition, especially when they do not yet know the meanings of many words. As proposed by the syntactic bootstrapping hypothesis (originally proposed in Gleitman, 1990; Landau & Gleitman, 1985), having access to the syntactic structure of sentences can help infants to discover the meaning of novel words. For instance, children may exploit the fact that an unknown word occuring in a noun context might refer to an object, while words occurring in verb contexts probably refer to actions (e.g., Gillette et al., 1999; Gleitman, 1990).

In a very recent study, He and Lidz (2017) showed that 18-month-olds (but not 14-month-olds) listening to a sentence such as "Look, it is a doke!" were able to infer that the novel word 'doke' referred to an object (i.e., a penguin), but when they were listening to sentences like "Look! It is pratching!", infants inferred that 'pratching' referred to an action (i.e., a spinning action). In this case, the critical word, 'doke' or 'pratch', was preceded and/or followed by disambiguating function morphemes (a, it's… -ing). However, in children everyday lives, not all content words are immediately preceded by function words, as shown in Part 1 with noun/verb homophones (e.g., flies can be either a noun or a verb in the sentence: "The baby flies..."). In such cases, a representation in terms of syntactic constituents can be very useful. For example, we showed that preschoolers interpreted "flies" as a noun when listening to [The baby flies]$_{NP}$…, but they interpreted it as a verb in [The baby]$_{NP}$[flies..]$_{VP}$ (de Carvalho, Lidz, Tieu, Bleam, & Christophe, 2016).

The question that arises is whether infants exploit prosodic boundaries together with function words to constrain their syntactic analysis of sentences containing novel words: If so, can they use this information to determine the syntactic nature of these novel words and therefore constrain their meaning? For example, in a sentence like "[Do you see the baby blicks]?", infants might be able to infer that "blick" is a noun, referring to an object; but in a sentence such as: "[Do you see]? [The baby] [blicks]!" they might be able to infer that "blick" is a verb, referring to an action. Note that this hypothesis is rather plausible, since to correctly interpret the novel word "blick" as a noun or a verb in this situation, infants would need to exploit exactly the same kind of information they were shown to use in the experiments presented in the first part of the current thesis (see.: Part 1).

Part 2: Can infants use phrasal prosody and function words to learn word meanings?

In the second part of this thesis I will experimentally investigate the plausibility of the *prosodic bootstrapping* and the *syntactic skeleton hypothesis*, testing whether infants are able to jointly use phrasal prosody and function words to access the syntactic structure of sentences and to constrain their acquisition of novel word meanings. To address these questions, I will explore two situations: first I will focus on the acquisition of verb meanings by asking whether young children are be able to use phrasal prosody to recover the syntactic structures of right-dislocated sentences containing a novel verb in French (e.g., "Il_i dase, le $garçon_i$ – meaning the boy is dasing; Chapter 6). Second, I will investigate the importance of phrasal prosody and function words to constrain the acquisition of nouns and verbs: a crucial syntactic distinction present in all languages (Chapters 7 and 8).

More precisely, in **Chapter 6**, I tested whether 28-month-olds were able to exploit phrasal prosody to recover the syntactic structure of sentences and to learn the meaning of a novel verb presented in right-dislocated sentences. I relied on the studies showing that toddlers learn that a novel verb such as "blicking" refers to a causal action between two participants when listening to transitive sentences such as "She is blicking the baby", but they do not make the same inference when listening to intransitive sentences such as "She is blicking" (e.g., Yuan & Fisher, 2009). In this chapter I tested whether toddlers would be able to constrain their interpretation of a novel verb meaning (whether it is referring to a two-participants action or to a one-participant action) based only on the prosodic information reflecting the syntactic structure of sentences. Toddlers might interpret the novel verb "dase" as refering to a two-participants action when listening to simple transitive sentences such as "[*Il dase le garçon*]" - He is **dasing** the boy' (meaning: someone is dasing the boy). However, if toddlers are able to exploit the prosodic information in right-dislocated sentences (i.e., a prosodic boundary between the novel verb and the dislocated unit), when listening to a sentence such as "[*Il_i dase*] [*le $garçon_i$*]" - He_i is **dasing,** the boy_i' (meaning: the boy is dasing), they should interpret the novel verb as being intransitive (i.e., not referring to a two-participants action, as they did in previous studies when listening to a simple intransitive sentences such as "He is dasing"; de Carvalho, Dautriche & Christophe, in prep).

In **Chapter 7**, I investigated the importance of phrasal prosody and function words to constrain the acquisition of nouns and verbs. Specifically, I tested whether at 18 months of age (an age at which the number of words children know is still relatively small), infants can rely on prosodic information present in the speech signal to structure spoken input into syntactic constituents, and whether in conjunction with function words infants could use this information to constrain their interpretation of novel word meanings. In a first experiment I tested whether infants were able to infer that a novel word such as *bamoule* in French refers to an object when listening to sentences such as "*C'est une bamoule*" - It is a bamoule, and to an action when listening to sentences such as "Elle *bamoule*" - She is bamouling. In a second experiment I used French sentences like '*Regarde la petite bamoule*', which can be produced either as one intonational unit, [*Regarde la petite bamoule*]! – [Look at the little bamoule]!, where bamoule is a noun, or as three intonational units, [*Regarde*]! [*la petite*] [*bamoule*]! – [Look]! [the little one] [is bamouling], where bamoule is a verb. I tested whether infants could correctly infer that bamoule refered to an object when it appeared in a noun position, and to an event (action) when it appeared in a verb position in the prosodic-syntactic structure of the sentences they heard (de Carvalho, He, Lidz & Christophe, submitted).

In **Chapter 8**, I tested the role of a special function word (i.e., the negative functors: ne+pas in French, meaning "not" in English) during the interpretation of novel word meanings. Negation is, in some extent, a function word with highly abstract meaning, since it has the power to completely change the meaning of a sentence. Contrary to content words such as nouns, verbs, and adjectives, negative function words do not refer to anything in the world. The meaning of negation comes from a syntactic computation that change the meaning of the propositions in which the negative elements appear. Studying how infants interpret the negative function words is important, since there is not a language in the world without negation. To investigate whether infants can interpret this kind of function words, I went a step further from the results obtained in Chapter 7, and tested whether infants would be able to correctly interpret negative sentences such as "*Regarde! Ce n'est pas une bamoule*" – Look! It is not a bamoule, and "*Regarde elle ne bamoule pas*" – Look! She is not bamouling (de Carvalho, Barrault & Christophe, submitted).

Chapter 6

Toddlers can exploit visual and syntactic cues to flexibly adapt their interpretation of novel verb meanings

Toddlers can exploit visual and syntactic cues to flexibly adapt their interpretation of novel verb meanings

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

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

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

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

Author Note

* Joint first authorship

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

Research highlights

- Toddlers integrate visual and syntactic information to recover from parsing bias
- They revise the default interpretation assigned to non-canonical sentences
- They use phrasal prosody to constrain the interpretation of novel verb meanings
- At 28 months, toddlers flexibly interpret novel verbs following different sources of information
- This suggests that toddlers can adjust their prior syntactic expectations depending on the context

Abstract

Previous studies show that because linguistic communication is often noisy and uncertain, the ability to successfully interpret sentences requires listeners to integrate their prior expectations about likely utterances (world-knowledge, linguistic regularities, etc.,) with the information they extract from the input (auditory, visual, etc), and to weigh the importance assigned to certain linguistic cues depending on the context. Here we test whether toddlers learning their language engage in a similar process while interpreting novel verbs. We relied on the work of Dautriche et al., (2014) showing that although 28-month-olds are able to correctly interpret familiar verbs in right-dislocated sentences such as [*Il$_i$ mange*] [*le lapin$_i$*] - 'He$_i$ is eating, the bunny$_i$' (meaning 'the bunny is eating'), when they found a novel verb in the same kind of structure (e.g., [*Il$_i$ dase*] [*le garçon$_i$*] - 'He$_i$ is dasing, the boy$_i$'; meaning 'the boy is dasing'), they relied on a parsing heuristic based on the canonical word order and the number of NPs in the sentence (e.g, NPagent+Verb+NPpatient), and considered each NP as a distinct argument of the verb, irrespective of phrasal prosody indicating that the final (right-dislocated) NP was co-referring to the first NP-agent. Toddlers' failure to integrate prosodic cues when interpreting novel (but not familiar) verbs was not a failure to use prosody *per se* but rather a reflection of the strategy that children use to interpret N+V+N structures in the presence of uncertainty about the meaning of the verb. In our study, we took advantage of this situation to test whether, across two experiments, toddlers would be able to integrate the information provided by different syntactic structures (in Experiment 1) or by the visual context illustrating the semantic representation of the possible meanings for the novel verb (in Experiment 2) to switch from the strategy based on the number of NPs in the sentence to the intransitive interpretation. The results showed that toddlers were able to adjust their interpretation of right-dislocated sentences, and correctly reached the intransitive interpretation provided by the prosody, thanks to the additional information extracted from the syntactic or the visual context. Similarly to adults, toddlers seem to be able to reach the most probable interpretation by weighing the plausibility of different information sources during language processing. This ability might be extremely important during language acquisition, since it would allow toddlers to evaluate the plausibility of different levels of linguistic information, to flexibly adjust their interpretations and constrain their acquisition of word meanings.

Key-words: language acquisition; syntactic bootstrapping; language processing; verb learning; parsing bias; noisy channel; phrasal prosody; eye movements; toddlers.

Toddlers can exploit visual and syntactic cues to flexibly adapt their interpretation of novel verb meanings

Introduction

Several studies suggest that the syntactic structure of sentences can guide young listeners in the discovery of word meanings: the syntactic bootstrapping hypothesis for language acquisition (Fisher, 1996; Gillette, Gleitman, Gleitman, & Lederer, 1999; Gleitman, 1990; Landau & Gleitman, 1985). This hypothesis has been validated experimentally and studies show for instance that at 18 months, infants can learn that a novel word such as "*doke*" refers to an action, when listening to sentences in which this word occupies a verb position, as in "*Look! It is doking*"; or that this novel word refers to an object when listening to sentences in which it appears as a noun, as in "*Look! It is a doke*" (He & Lidz, 2017). Similarly, 19-month-olds exposed to sentences such as "*The dax is crying*" infer that "*dax*" refers to an animate entity (i.e., a novel animal), because *dax* appeared in the subject position of a verb requiring an animate agent (i.e., someone able to cry); but when exposed to sentences like "*The dax is right here*", toddlers did not show any preference for the animate entity (in comparison to an inanimate object) at test (Ferguson, Graf, & Waxman, 2014). Syntactic bootstrapping works because a partial analysis of sentence structure helps children to construct a first parse of sentences and to build an abstract representation that provides useful linguistic generalizations to infer the meanings of novel words (e.g., Fisher, 1996; Lidz, Gleitman, & Gleitman, 2003).

The structure mapping account (Fisher, 1996) suggests that the origins of syntactic bootstrapping is based on an unlearned bias towards one-to-one mapping between nouns in sentences and participant-roles in events (actions), to constrain some important aspects of verb meanings. In the simplest case to illustrate this idea, consider the situation in which a toddler listens to the sentence "The boy is pushing the girl" while observing the action that the boy is doing on the girl. Noticing that this sentence contains two nouns (i.e., "the boy" and "the girl") will give children the opportunity to map the sentence onto a representation of "pushing" that involves two participants. Additionally, as soon as children know that the most frequent word order in English follows the organization "Subject-Verb-Object", they infer that the first noun (preceding the verb) refers to the agent of the action while the second noun (following the verb) refers to the patient (e.g., Gertner, Fisher, & Eisengart, 2006). Taken together, the set of noun phrases in the sentences may allow children to infer how many participants are involved in the action described by the verb, and which is the role of each participant (agent or patient) in the event (Fisher, 1996, 2002; Naigles, 1990).

Supporting this hypothesis, several studies show that from 19 months of age, toddlers can learn that a novel verb such as "*blicking*" refers to a causal action between two participants, when listening to transitive sentences such as "*She is blicking the baby*", but they do not make the same inference when listening to intransitive sentences such as "*She is blicking*" (Arunachalam & Waxman, 2010; Dautriche et al., 2014; Yuan & Fisher, 2009; Yuan, Fisher, & Snedeker, 2012). At 21 months, toddlers can use the order of nouns in transitive sentences such as "The boy is gorping the girl" to interpret the role of participants in the action. These empirical studies suggest that infants exploit the set of nouns in a sentence to guide their initial interpretation. Validating the strength of this mechanism for language acquisition, studies using computational models to study how children acquire syntax suggest that using these simple representations based on the sets of nouns in a sentence is useful for learning to interpret sentences (Christodoulopoulos, Roth, & Fisher, 2016; Connor, Gertner, Fisher, & Roth, 2008).

Although this ability to map each noun phrase in a sentence as a participant in the action can be extremely useful in distinguishing transitive from intransitive verbs, and sensitivity to the canonical word-order patterns would be an important cue for children to guess the precise role of each participant in the action, these parsing strategies can also provide wrong interpretations. Note that not all the sentences containing two noun phrases (2NPs) are necessarily transitive or imply an event in which the first NP is the agent and the second is the patient of the action. In a sentence such as "John and Mary sleep", if children count each NP as a participant in the action and consider that "John" is the agent and "Mary" the patient, they will consider the verb to be transitive, involving an action between these two participants, which is not true. Thus, an interpretation based on the number of NPs is not always correct (Lidz et al., 2003) and in some situations it can lead children toward a wrong interpretation.

Sensitivity to word order to infer the role of participants in the action can also allow children to make wrong interpretations. For instance, studies have shown that by linking the canonical word-order noun+verb+noun sequence in English with the interpretation agent+action+patient, children mistakenly interpret passive sentences such as "The boy was bumped by the girl", since they experience difficulties to extract the reversed interpretation, and wrongly interpret that the boy was the agent of the action (de Villiers & de Villiers, 1973; Maratsos, 1974; Slobin & Bever, 1982).

In Gertner and Fisher (2012), 21-month-olds provided the same interpretation to a new verb appearing in the sentence "The boy is gorping the girl" or in the sentence "The boy and the girl are gorping!", since in both conditions they interpreted "gorping" as referring to a causative action between two participants. Note that although both sentences contained two noun phrases, "gorping" was transitive in the former example, but intransitive in the latter. Additionally, toddlers exposed to sentences such as "The girl and the boy are gorping" mistakenly used the noun-order in English to assign an agent's role to the first NP in this sentence (i.e., the girl) and interpreted this sentence differently from the sentence "The boy and the girl are gorping!", despite of the fact that both sentences have the same meaning. These results confirm that young children during the first steps of language acquisition have a tendency to interpret each NP in a sentence as a participant in the action and the first of two NP as the agent, even when this interpretation is not correct.

In a recent study however, French 28-months-old were shown to be selective in their application of such parsing heuristics when interpreting sentences (Dautriche, Cristia, Brusini, Yuan, Fisher & Christophe, 2014). The authors used a specific syntactic structure in French: right-dislocated sentences (e.g., "*Il*$_i$ *dort, le lapin*$_i$" – It$_i$ is sleeping, the bunny$_i$), to test whether 28-month-olds could use phrasal prosody to constrain the interpretation of a novel verb appearing in such a structure. The specific prosody of these sentences (in particular an intonational phrase boundary between "sleeping" and "the bunny") should block the interpretation based on the number of NPs, because in this case "the bunny" is co-referring to the pronoun "It", and this sentence is describing an intransitive action with a single participant. Thus, with a correct interpretation of the prosodic information, right-dislocated sentences can contain an intransitive verb between two noun phrases, and describe an action with a single participant. However, insensitivity to prosodic information would make toddlers wrongly interpret right-dislocated sentences as transitives (following the abstract representation N+V+N), and mistakenly interpret "the bunny" as the patient of the action (rather than its agent). The results showed that 28-month-olds correctly interpreted the prosodic information in right-dislocated sentences containing familiar verbs (e.g., "*Il*$_i$ *mange,* *le lapin$_i$*" – He$_i$ is eating, the rabbit$_i$, as meaning 'the rabbit is eating'), but they interpreted right-dislocated sentences containing a novel verb incorrectly (e.g., "*Il$_i$ dase, le garçon$_i$*" – He$_i$ is *dasing*, the boy$_i$). Toddlers exposed to a dialogue phase containing sentences presenting a novel verb in right-dislocated structures, behaved exactly as toddlers who heard the same novel verb in transitive sentences (e.g., "*Il$_i$ dase le garçon$_k$*" – He$_i$ is dasing the boy$_k$): at test, when asked to look towards "*the one who was dasing*", both groups preferred to look towards the video showing a causative action between two-participants, rather than towards a video showing a one-participant action. Importantly, toddlers who were exposed to dialogues in the intransitive condition (e.g., "*Il dase*" - He is dasing) did not show any preference for the two-participants action. These results show that despite the tendency to interpret each noun phrase as a participant in the action, and the first of the two nouns as its agent, toddlers can correctly integrate prosodic information and reach the correct interpretation of the sentences when they contain familiar verbs. However, when a novel verb was involved, toddlers applied the canonical surface-to-meaning mapping to infer the meaning of the novel verb, and apparently ignored the information provided by prosody.

This selectiveness in the use of the canonical surface-to-meaning mapping could be explained by the fact that although right-dislocated sentences are extremely frequent in French, they are nevertheless much less frequent than transitive sentences (see Dautriche et al., 2014). Thus in the presence of uncertainty about the meaning of a verb, toddlers preferred to rely on the interpretation that most often applies for sentences containing 2NPs: the first NP is interpreted as the agent of the action and the second one as the patient, as in transitive sentences.

Taken together, all these studies suggest that the problem faced by children is that sometimes they can have several parsing strategies that can provide different (and opposite) results. In such cases, children would need to evaluate the plausibility of each interpretation, in order to decide which interpretation is correct or incorrect. In other words, children need to find a way to figure out that in some situations the strategy of relying on the number of nouns in the sentence, or the order in which the nouns appear, is not the best option to constrain their interpretation, and that they should instead rely on some other kind of information to constrain their parsing.

The question that arises is how children could manage to deal with conflicting interpretations during sentence processing? This question has been extensively studied with adults and the existing results suggest that when adults deal with conflicting interpretations during sentence processing, they quickly adapt their interpretation to the situation in which they are, and they can estimate the plausibility of each source of information (or interpretation) to decide in which cue they should rely on (Gibson, Bergen, & Piantadosi, 2013; Jaeger, 2010; Levy, 2008). In other words, the *Noisy-channel account* suggests that to understand sentences, adults integrate their prior expectations about likely utterances (world-knowledge, linguistic regularities, etc.; e.g., Trueswell & Kim, 1998) with the information they extract from the input (e.g., auditory, visual, etc.; e.g., Tanenhaus, Spivey, Eberhard, & Sedivy, 1995) to constrain their online sentence processing. However, depending on the level of uncertainty of a given environment (e.g., noise, accents, new talker, etc.), adults can adjust their prior linguistic expectations to weigh the plausibility of different information sources (e.g., Ernst & Banks, 2002; Jacobs, 1999).

In Gibson, Bergen, and Piantadosi (2013) for instance, adults were presented with semantically implausible sentences such as "The mother gave the candle the daughter", which could be interpreted as semantically strange, if the listener interpreted it literally, or they can simply reconstruct the meaning of the sentence, assuming that the speaker meant the more plausible interpretation ("The mother gave the candle ***to*** the daughter") but that for some reason they missed the word "to". Their experiment showed that adults tended to correct these errors more often if they thought that the syntax of the sentences was imperfect (when the experiment contained many small typographical errors involving function words). However, they did not correct these errors if they were led to believe that they were in a silly context (when the experiment contained many other sentences that were semantically implausible). The questions that arise are: are toddlers, who are still in the process of learning their language, able to engage in a similar process when interpreting novel verbs? Can toddlers process language in a similarly flexible way, and assign different weights to different kinds of linguistic cues, depending on the information provided by the context?

To test this, we relied on the work of Dautriche et al. (2014) mentioned above which showed that French 28-month-olds incorrectly expect novel verbs embedded in right-dislocated sentences (e.g., *Il$_i$ VERB, le garçon$_i$* - He$_i$ is VERBing, the boy$_i$', meaning 'the boy is VERBing') to map to a causal action between two participants (assuming that someone else is VERBing the boy), even though the post-verbal prosodic phrase boundary should have blocked this interpretation. Given that toddlers correctly interpret right-dislocated sentences with familiar verbs (e.g., It_i eats, the $bunny_i$), and that studies showed that from 20 months of age, toddlers can use prosodic information to constrain their syntactic analysis (de Carvalho, Dautriche, & Christophe, 2016; de Carvalho, Dautriche, Lin, & Christophe, 2017; de Carvalho, Lidz, Tieu, Bleam, & Christophe, 2016), their failure to integrate prosodic cues when interpreting novel verbs seems to not be a failure to use prosody *per se* but rather a reflection of the strategy that children use to interpret sentences in the presence of uncertainty about the verb meaning.

We assume that when listening to right-dislocated sentences, toddlers can come up with two possible interpretations: an intransitive one (integrating prosodic information), as they did when interpreting familiar verbs, or a transitive one (using the strategy based on the number of NPs in the sentence), as they did when interpreting novel verbs. We suspect that toddlers relied on the prosodic information when interpreting familiar verbs, because they had prior knowledge about the syntactic contexts in which these verbs could appear. Given that many transitive verbs can appear both in transitive and in intransitive frames ("Bob ate an apple"; "Bob ate"), and that such verbs are frequent in child-directed speech (e.g., Scott & Fisher, 2009), it is possible that children knew that a verb such as "eat" could be used both in transitive frames containing two noun phrases (e.g., "the boy is eating an apple") and in intransitive frames containing only one noun phrase (e.g., "the boy is eating") or even in right-dislocated sentences that they may have already heard in their everyday lives. In this situation, the information provided by prosodic information (suggesting that the verb was used in an intransitive way) was easy to integrate since that information was compatible with the linguistic expectations that children had about this verb. However, when listening to right-dislocated sentences containing a novel verb, toddlers had no prior knowledge to base their interpretations on. They repeatedly heard this novel verb in sentences containing 2NPs and they had no prior information on whether or not this novel verb could be used in an intransitive form. In this uncertain situation, although toddlers had the two possible interpretations available (the intransitive one provided by prosody and the transitive one provided by the number of NPs), they preferred to rely on their parsing heuristic and to interpret a structure containing two noun phrases, N+V+N, as agent+action+patient. They assumed that any novel verb appearing in a NP-verb-NP sentence would refer to a causal action where an agent (the first NP) acts on a patient (the second NP).

In the current study we asked whether toddlers would be able to flexibly adjust their reliance on phrasal prosody versus the heuristics based on the number of NP, depending on the additional information that they extract from the input. In experiment 1, we improved children's access to information about the kinds of syntactic structures that the novel verb could enter. We added simple intransitive frames containing a novel verb to the dialogues used in Dautriche et al., (2014) with right-dislocated sentences. Thus, immediately after having heard a right-dislocated sentence, toddlers heard an additional syntactic frame in which the novel verb was used in a simple intransitive sentence (e.g., *Il$_i$ dase, le garçon$_i$. Ah bon, Il dase?* - He$_i$ is dasing, the boy$_i$. Really, he is dasing?). We tested whether showing toddlers that the novel verb could also appear in simple intransitive frames would increase the probability of "dase" being considered as intransitive (i.e., not referring to a causal action between two participants).

In experiment 2, we improved children's access to the semantic representation of events and offered them the possibility to observe both a transitive and an intransitive interpretation for a novel verb, at the same time as they listened to the sentences. Thus, while toddlers heard right-dislocated sentences containing a novel verb, they watched two videos side-by-side on a TV screen: one video showing a person doing a self-generated action (one participant action) and another video showing a person acting on another (a causative action between two participants). We tested whether offering the possibility of observing the intransitive interpretation for a novel verb would impact children's on-line interpretation of right-dislocated sentences and increase their reliance on the prosodic information to interpret the novel verbs as intransitives.

Experiment 1

Experiment 1 tested whether toddlers can adjust their reliance on the prosodic information of right-dislocated sentences containing a novel verb, when provided access to additional information about the syntactic contexts in which the novel verb "*daser*" could appear: namely that it can also appear in intransitive sentences. Following the preferential looking paradigm used in Yuan and Fisher (2009) and in Dautriche et al., (2014), we presented 28-month-olds with dialogues introducing a novel verb "*daser*" in one of four conditions: transitive+intransitive, right-dislocated+intransitive, dislocated-only and intransitive-only (see Figure. 1). In each dialogue condition, toddlers listened to a total of 8 sentences containing a novel verb (i.e., daser). In the transitive+intransitive condition they heard 4 transitive sentences and 4 intransitive sentences in alternation; in the right-dislocated+intransitive condition they heard 4 right-dislocated sentences and 4 intransitive sentences in alternation; in the right-dislocated-only condition they heard 8 right-dislocated sentences and in the intransitive-only condition they heard 8 intransitive sentences. Right after exposure to one of these dialogues, toddlers were then asked to look for "*daser*" while watching two videos displayed side-by-side: a causal action featuring two participants, and a one-participant action.

Given that many transitive verbs can appear both in transitive and intransitive frames, we expected that toddlers exposed to the dialogues in the transitive+intransitive condition should still be able to interpret the novel verb as transitive and referring to a two-participants action at test, even though this verb appeared in transitive and intransitive structures in alternation. If toddlers are flexible in the way they integrate the information provided by the different syntactic frames in which a novel verb appears, in the right-dislocated+intransitive condition they might reach the intransitive interpretation, which would surface as no preference for the two-participants action at test: this might be the case because listening to right-dislocated and intransitive sentences in alternation may increase the probability of *daser* being considered as intransitive, since both syntactic frames indicate that the verb is intransitive.

The intransitive-only and right-dislocated-only conditions were used to provide baseline conditions: a baseline on how toddlers behave when the syntactic context of right-dislocated sentences was not enriched with the intransitive sentences; and how toddlers behave when listening only to simple intransitive sentences. We expect the right-dislocated-only condition to replicate the results observed in Dautriche et al., (2014), and show that toddlers process these sentences as transitive and look more towards the two-participants action during the test. As established in previous studies, toddlers in the intransitive only condition should not show any preference for the two-participants action during the test.

To summarize, we expect that at test, toddlers in the right-dislocated-only condition will behave as toddlers in the transitive+intransitive condition and associate the novel verb to the two-participants causal action. Crucially however, if the presence of intransitive sentences in the right-dislocated+intransitive dialogue increases the plausibility of the intransitive interpretation for right-dislocated sentences, toddlers in this condition should behave as toddlers in the intransitive-only condition and should not show any preference for the causal action between two participants. If so, this experiment will show that toddlers can use the information provided by multiple syntactic contexts to adjust their interpretation of novel verb meanings and to recover from their parsing biases.

Method

The stimuli, data and analyses of the experiments reported in this paper are accessible to readers on the OSF (Open Science Framework) database through the following link: https://osf.io/b5yqp/?view_only=7b34cd9ddba94da28adc1f8958c25fab

Participants. Eighty French 28-month-olds participated in this study (mean age = 28.0 months, range = 26.9 to 30.2; SD =0.7; 40 girls; 20 participants in each condition). All were native French speakers with less than 20% exposure to another language. Twenty toddlers were randomly assigned to each of the four experimental conditions (transitive+intransitive, right-dislocated+intransitive, dislocated-only and intransitive-only). An additional twenty-four children participated in the study but were not included in the final analysis because of fussiness during the experiment (5), distraction during the dialogue phase (6), side bias (2), unusable test trial with missing eye tracking data (8), because they were exposed to other languages than French at home (1), or because of technical problems (2).

Apparatus. Toddlers were tested individually in a sound-attenuated double-walled booth. They were sat on a parent's lap, facing a 42-in television positioned 70cm away from them. Toddlers' eye movements were recorded by an eye-tracker (Eyelink 1000) placed below the screen, and operating in a remote mode with a time-sample collected every 2ms. The caregivers wore opaque glasses and the experimenter stayed outside the cabin during the test.

Materials and Procedure. The stimuli used in this experiment were videos of two women conversing (for the dialogue phases) and videos of people performing actions (for the test phase). The videos of actions were accompanied by sound tracks recorded by a female native French speaker (last author).

The procedure was similar to that of Dautriche et al., (2014) and Yuan & Fisher (2009). The experiment was composed of three blocks: practice, dialogue and test. The experiment began by a practice block to familiarize children with the procedure. In this practice block, toddlers saw two practice items involving familiar verbs, one intransitive (*danser* – to dance, or *marcher* – to walk) and one transitive (*pousser* – to push or *porter* – to carry). These practice trials consisted of two 8s test events in which a synchronized pair of videos was presented side-by-side on the screen along with a sound track that encouraged toddlers to look at one of the videos. For instance, when participants saw the pair of videos showing a girl walking in one video and a girl dancing in the other video, they heard a sentence such as "*Tu la vois qui marche? Regarde celle qui marche!*" - Do you see her walking? Look at the one who is walking!". The target videos were counterbalanced across participants, such that half of the participants were asked to look towards "the girl walking" and the other half towards "the girl dancing". This practice item served to show participants that in this task, one of the two videos matched the soundtracks they heard.

Each trial (practice or test) started with an inspection period to provide participants enough time to inspect each of the videos individually, on each side of the screen. Thus, each video was first presented alone for 5 seconds on the left or the right side of the screen and a neutral audio prompt was played at the same time (e.g., "*Hey, regarde là! Tu as vu ça*?" - Hey, look here! Did you see that?). Next, the two videos disappeared, and a sentence containing the verb was presented during a 5-s black-screen interval (e.g., "*Oh regarde! Elle marche!"* - Hey, look! She is walking!"). Right after that, the two videos appeared side-by-side on the screen for 8s, and at the same time participants heard the test sentences repeating the verb twice (e.g., "*Tu la vois qui marche? Regarde celle qui marche!*" - Do you see her walking? Look at the one who is walking!). Then the videos disappeared and in a new 5-s black-screen interval participants heard another sentence containing the target verb (e.g., She is walking!). Right after that the two videos were played again for 8s while toddlers heard two more repetitions of the target verb (e.g., "*Elle est où, celle qui marche? Regarde celle qui marche!*" - Where is she, the one who is walking? Look at the one who is walking!).

After the practice block, participants started the dialogue block in which they saw the dialogue phase appropriate for their assigned condition. In this block two four-sentence dialogue video clips of 24s separated by a 3s interval were presented in the middle of the screen. Thus, each participant was exposed to eight sentences: half transitive and half intransitive for the transitive+intransitive condition; half right-dislocated and half intransitive for the right-dislocated+intransitive condition; eight right-dislocated sentences in the right-dislocated-only condition, and eight intransitive sentences in the intransitive-only condition (see Figure 1, for the entire list of sentences in each condition). These dialogue videos showed a conversation between two women who uttered sentences using a novel verb "***daser***" in one of the four experimental conditions. They uttered the sentences in child-directed speech. See Figure 1 for sample dialogues.

Dialogue phase (4 conditions)

------------------------------ Transitive+Intransitive ------------------------------

A: *Hey! Il va daser le papa!*	Hey! He will dase the dad!
B: *Ah bon, il va daser ?*	Really, he will dase?
A: *Oui, et en plus ils ont dasé les garçons.*	Yeah! And they dased the boys.
B: *C'est vrai, ils ont dasé!*	That's right, they dased!
--	--
A: *Tu sais quoi? Elles ont dasé les filles!*	Guess what? They dased the girls!
B: *Quoi, vraiment elles ont dasé?*	What, really they dased?
A: *Oui! Et elle va daser la maman!*	Yes! And she will dase the mom!
B: *Waouh, elle va daser!*	Waouh, she will dase!

-------------------------- Right-dislocated+Intransitive --------------------------

A: *Hey! Il va daser , le papa!*	Hey! He$_i$ will dase, the dad$_i$!
B: *Ah bon, il va daser ?*	Really, he will dase?
A: *Oui, et en plus ils ont dasé , les garçons.*	Yeah! And they$_i$ dased, the boys$_i$.
B: *C'est vrai? ils ont dasé!*	That's right, they dased!
--	--
A: *Tu sais quoi? Elles ont dasé , les filles!*	Guess what? They$_i$ dased, the girls$_i$!
B: *Quoi, vraiment elles ont dasé?*	What, really they dased?
A: *Oui! Et elle va daser , la maman!*	Yes! And she$_i$ will dase, the mom$_i$!
B: *Waouh, elle va daser!*	Waouh, she will dase!

------------------------------ Right-dislocated only ------------------------------

A: *Hey! Il va daser , le papa!*	Hey! He$_i$ will dase, the dad$_i$!
B: *Ah bon, il va daser , le papa?*	Really, he$_i$ will dase, the dad$_i$?
A: *Oui, et en plus ils ont dasé , les garçons.*	Yeah! And they$_i$ dased, the boys$_i$.
B: *C'est vrai? ils ont dasé , les garçons!*	That's right, they$_i$ dased, the boys$_i$!
--	--
A: *Tu sais quoi? Elles ont dasé , les filles!*	Guess what? They$_i$ dased, the girls$_i$!
B: *Quoi, vraiment elles ont dasé , les filles?*	What, really they$_i$ dased, the girls$_i$?
A: *Oui! Et elle va daser , la maman!*	Yes! And she$_i$ will dase, the mom$_i$!
B: *Waouh, elle va daser , la maman!*	Waouh, she$_i$ will dase, the mom$_i$!

--------------------------------- Intransitive only ---------------------------------

A: *Hey! Il va daser !*	Hey! He will dase!
B: *Ah bon, il va daser ?*	Really, he will dase?
A: *Oui, et en plus ils ont dasé.*	Yeah! And they dased.
B: *C'est vrai? ils ont dasé!*	That's right, they dased!
--	--
A: *Tu sais quoi? Elles ont dasé !*	Guess what? They dased!
B: *Quoi, vraiment elles ont dasé ?*	What, really they dased?
A: *Oui! Et elle va daser !*	Yes! And she will dase!
B: *Waouh, elle va daser!*	Waouh, she will dase!

Figure 1: Sample of dialogues in Experiment 1 for the four conditions: transitive+intransitive, right-dislocated+intransitive, right-dislocated only and intransitive only. The dialogues were split in two 24-s videos containing four sentences each separated by a 3-s black screen. Transitive+intransitive and right-dislocated+intransitive dialogues were composed exactly of the same words, but differed only in their prosodic structures, reflecting their different syntactic structures.

Three seconds after the end of the dialogue phase, participants started the test block. This block presented participants with two videos illustrating two novel actions. One video showed an action executed by one single participant (a girl making circles with her right arm) and the other video showed a causative action between two participants (a girl swinging another girl's leg), the same test videos were used in Yuan & Fisher (2009) and Dautriche et al., (2014).

The novel action videos were presented following the same procedure as the practice block: inspection period first and then both videos displayed side-by-side on the screen for eight seconds, but for one single test trial. During the test trial, participants heard two sentences featuring the novel verb in an intransitive syntactic structure: "*Tu la vois qui dase? Regarde celle qui dase!*" – Do you see her dasing? Look at the one who is dasing!), and at the same time, toddlers eye-gaze towards the videos were recorded by an eye-tracker. The time-course of the test block presentation is illustrated in Figure 2.

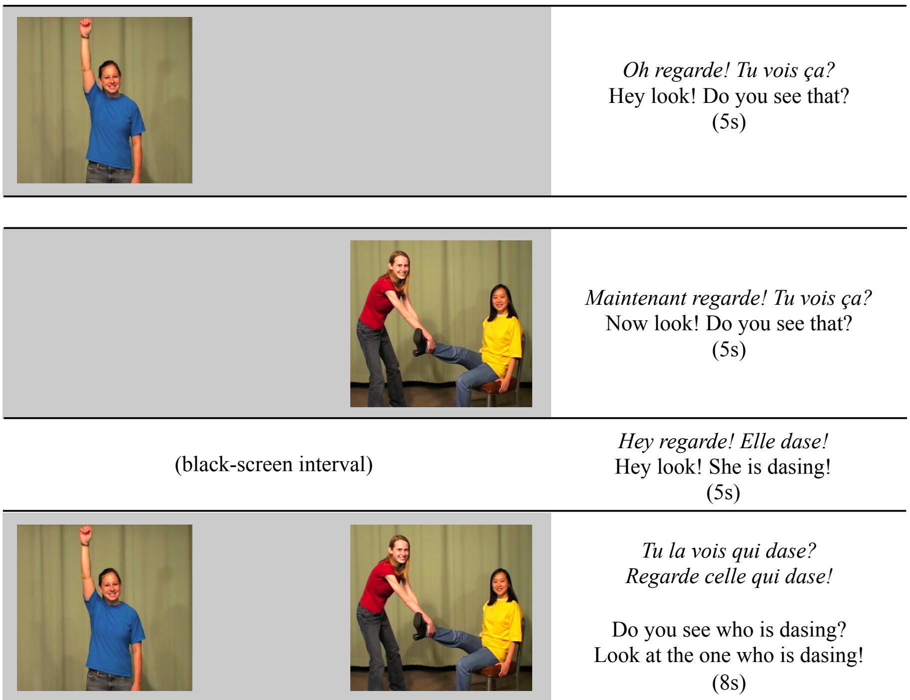


Figure 2: Time-course of the test phase of Experiment 1. After having watched the videos of the dialogue phase, participants were presented with two novel action videos that were first individually presented in a different side of the screen for 5 seconds. Then the two videos were presented simultaneously side-by-side on the screen for 8 seconds and participants were asked to look at the one who was "dasing".

Note that the auditory stimuli in the test trial were identical for all participants. The order of presentation of the practice items and the left/right position of the target videos were counterbalanced across participants, within each of the four conditions.

Acoustic Analysis. Given that the transitive and the dislocated sentences had exactly the same words but differed only with respect to their prosodic structures, reflecting their different syntactic structures, acoustical analysis of the stimuli were conducted. These analyses revealed that there was a clear prosodic boundary between the verb and final DP in the dislocated sentences but not in the transitive sentences. A preboundary syllable lengthening for the last syllable of the novel verb "daser" (i.e., the syllable –"er") in the right-dislocated sentences from the right-dislocated+intransitive condition was observed when compared to the same segment on the transitive sentences from the transitive-intransitive condition (M_{trans} = 199ms vs M_{disloc}= 621ms, t(6) = 5.71, $p < 0.01$**).

Additionally we also observed a pitch drop in the right-dislocated sentences from the right-dislocated-intransitive condition compared to the same pitch values on the same segments of the transitive sentences in transitive-intransitive condition. This analysis compared the max pitch on the last vowel of the novel verb "daser" and on the first vowel of the final NP (e.g., max pitch on the "-e" from "daser" minus the max pitch on the "-e" from "le papa" - the dady; M_{trans} = -6.07Hz vs M_{disloc}= -101.15Hz , t(3) = 13.17, p <0.001***). Note that there were no pauses between the verb and the final NP, which ensures that to differentiate the transitive from the right-dislocated sentences, toddlers had to interpret the prosodic structure of the sentences rather than just ignoring any lexical material occurring after a pause. A pilot experiment with naive adults (n=10), asked participants to listen to each of the sentences used in the dialogues and decide who was performing the action. Participants interpreted transitive and right-dislocated sentences correctly over 94% of the time.

Data processing and analysis. In this experiment, during the test phase all the participants listened to the same sound file (asking them to look towards the one who was dasing). Our prediction is that the dialogues they heard before the test will impact their looking preference towards the two-participants action at test. Given that the looking times toward the two-participants action and toward the one-participant action are almost complementary (except for the away looking time, which was not reliably different between conditions), in our statistical analysis we used the proportion of looking times towards the two-participants action as the dependent variable.

In order to find whether there was a time-window in which children looked towards the two-participants action significantly more than chance (0.5), a non-parametric cluster-based permutation analysis was conducted on the entire duration of the test trial (from 0 to

8000ms). This analysis was run for each of the 4 conditions separately. Note that the same kind of analysis was used in other eye-tracking studies in the literature : Dautriche, Swingley & Christophe, 2015; de Carvalho et al., 2017; Hahn, Snedeker & Rabagliati, 2015; Von Holzen & Mani, 2012; see Maris & Oostenveld, 2007 for a formal presentation of the analysis itself). This analysis has several advantages: it allows us to find a time-window where we observe a significant effect without having to select it arbitrarily. This analysis involves two steps: 1) the identification of time-windows that have a potential effect; 2) the statistical test itself, which quantifies whether these effects are likely to have been generated by chance.

For each time point a t-test comparing the proportion of looks towards the two-participants action to the chance level (0.5) was calculated. Adjacent time points with a t-value greater than some predefined threshold ($t > 2$; on arcsin-transformed data) 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 the conditions were randomly shuffled (chance versus condition of interest). For each simulation, we computed 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 if its statistic is greater than the statistic of the largest cluster found in 95% of the simulations (ensuring a p-value of .05).

Results

Figure 3 shows the proportion of toddlers' looks towards the two-participants action, time-locked to the onset of the target verb "daser" during the test query: "*Tu vois celle qui dase?* - Do you see the one who is dasing?".

In the transitive+intransitive condition, the cluster analysis revealed a significant time window between 1300 ms and 1982 ms from the onset of "daser", in which the proportion of looks toward the two-participants action was significantly different from chance ($p < .01$). For the right-dislocated only condition, a significant time window was found between 1108ms and 1832ms ($p < .02$). However, for the right-dislocated+intransitive condition and for the intransitive-only condition, the proportion of looking times towards the two-participants action did not differ from chance.

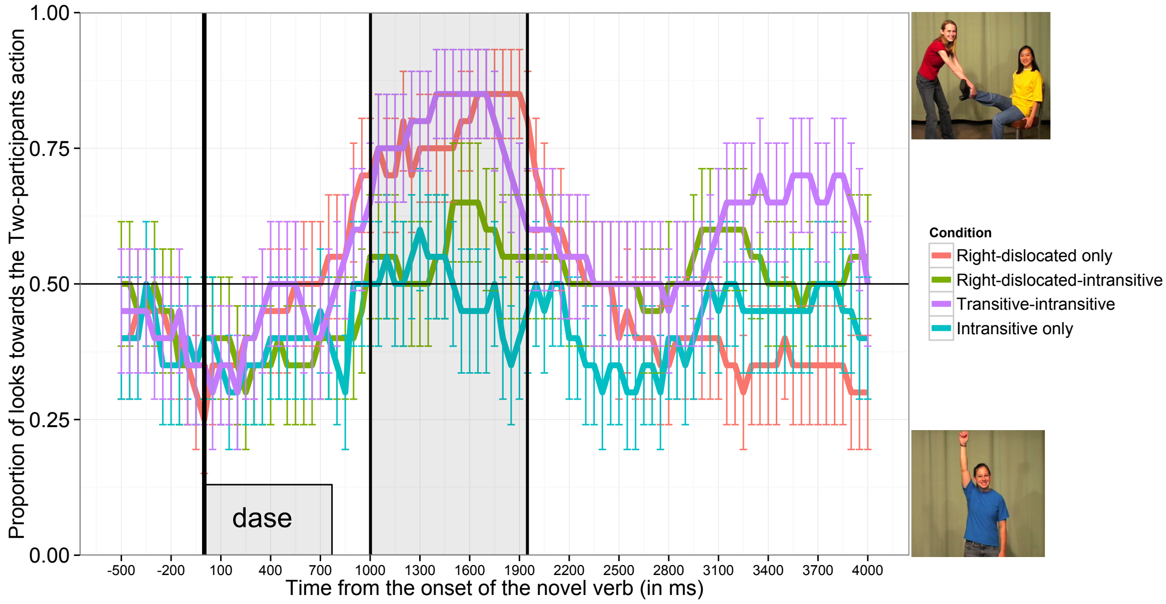


Figure 3: Proportion of looks towards the two-participants action, time-locked to the onset of the novel verb "daser" (vertical black line) for toddlers in the transitive+intransitive condition (purple curve), the right-dislocated+intransitive condition (green curve), the right-dislocated-only condition (magenta curve) and the intransitive-only condition (blue curve). Error bars represent the standard error of the mean. A nonparametric cluster-based permutation test (Maris & Oostenveld, 2007) performed on the whole duration of the test trial (8s) identified a significant time window where the looking curve, averaged across all four conditions, was above chance level ($p < .05$) (time-window marked by vertical bars, 1100-2000ms after verb onset). The same analysis performed on each condition separately revealed that children in the transitive+intransitive (purple) and right-dislocated-only (red) conditions looked significantly more towards the two-participants action (both ps < .05), while children in the right-dislocated+intransitive (green) and in the intransitive-only (blue) condition had no preference for any of the action videos. In addition, a cluster-based permutation test that directly compared conditions pairwise revealed that both the right-dislocated-intransitive and the intransitive-only conditions (green and blue) differed significantly from the transitive+intransitive and dislocated-only conditions (red and purple) (ps < .02).

In order to analyze the differences between conditions without having to select an arbitrary time window, we conducted two additional analyses. First, we averaged all the conditions and conducted an additional cluster-based analysis to identify whether, across all conditions, we could find a time-window in which the proportion of looking times towards the two-participants action was significantly different from chance. This analysis revealed a significant time-window from 1100-2000ms after verb onset ($p < .05$). We used this time-window to conduct a mixed-model analysis of the differences between the 4 conditions (see below). Second, we conducted 6 cluster-based permutation tests that directly compared conditions pairwise. This analysis revealed that both the right-dislocated+intransitive and the intransitive-only conditions (green and blue) differed significantly from the transitive+intransitive and right-dislocated-only conditions (red and purple) (ps < .02).

In order to analyze the differences between conditions in the time-window identified by the analysis above (1100-2000ms, represented with the grey-shaded area in Figure 3), we entered the proportion of looking-times towards the two-participants action in a mixed logit model with a fixed Condition effect with 4 modalities (transitive+intransitive; right-dislocated+intransitive; right-dislocated-only; intransitive-only) and a random Subject effect. We report here the coefficients associated with the fixed effects β and the associated Z statistic (comparing β to 0).

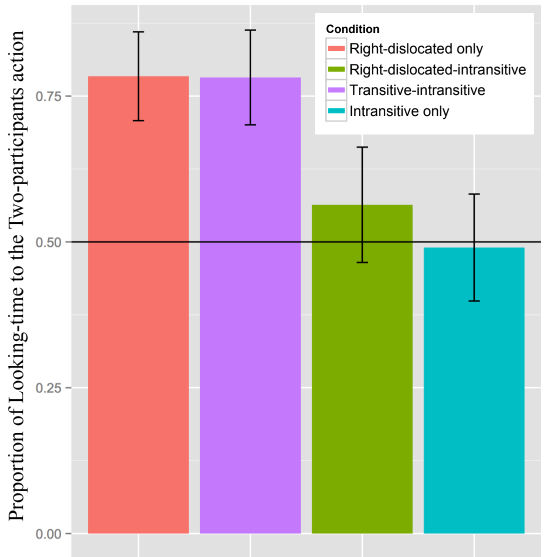


Figure 4: Proportion of looking-times towards the two-participants action in the time-window going from 1100 to 2000 ms after the onset of the novel verb in the test trial, for participants in the right-dislocated-only condition (magenta curve), in the transitive+intransitive condition (purple curve), in the right-dislocated+intransitive condition (green curve) and in the intransitive-only condition (blue curve). This time-window was revealed by the cluster-based permutation analysis (Maris & Oostenveld, 2007) as having a significant preference for the two-participants action when we averaged all the four conditions together. Error bars indicate the standard error of the mean in each group. Toddlers in the right-dislocated only condition behaved as did toddlers in the transitive-intransitive condition: they looked more towards the two-participants action than toddlers in the right-dislocated-intransitive condition and toddlers in the intransitive only condition.

As represented in Figure 4, looking times towards the two-participants action were affected by the dialogue conditions. Toddlers who heard the novel verb "*daser*" in the transitive+intransitive dialogue looked more towards the two-participants action than those who heard the novel verb in the intransitive-only condition ($\beta = 10.20$; $z = 2.4$; $p < 0.05$). To some extent, this result replicates previous studies showing that toddlers use the syntactic structure in which a novel verb appears to infer how many participants are involved in the action (e.g., Arunachalam, Escovar, Hansen, & Waxman, 2011; Arunachalam & Waxman, 2010; Dautriche et al., 2014; Yuan & Fisher, 2009; Yuan, Fisher, & Snedeker, 2012). We observe that despite the fact that children in the transitive+intransitive condition heard transitive and intransitive sentences in alternation, they were still able to interpret the novel verb as referring to a two-participants action at test. Importantly, children in the intransitive-only condition behaved differently from children in the transitive+intransitive condition, since they did not show any preference for the two-participants action.

As expected from Dautriche et al (2014), toddlers in the right-dislocated-only condition behaved similarly to toddlers in the transitive-intransitive condition: both groups had greater looking-times towards the two-participants action, and they did not differ from each other ($\beta = -0.18$; $z = -0.04$); the right-dislocated-only condition differed significantly from the intransitive condition ($\beta = -10.38$; $z = -2.44$, $p < 0.05$). These results replicate the results of Dautriche et al., (2014) and show, once again, that when processing a novel verb that appears only in right-dislocated sentences (i.e., non-canonical structure), toddlers prefer to rely on the canonical surface-to-meaning mapping to infer the meaning of the novel verb rather than to integrate the information provided by prosody.

Crucially however, children in the right-dislocated+intransitive condition looked less to the two-participants action than children in the transitive+intransitive condition ($\beta = 8.97$, $z = 2.10$, $p < 0.05$) or children in the right-dislocated-only condition ($\beta = 9.15$; $z = 2.14$, $p < 0.05$). They behaved as the children in the intransitive-only condition (these two conditions were not significantly different, $\beta = -1.23$; $z = -0.32$). This suggests that contrary to toddlers who heard the novel verb only in right-dislocated sentences, toddlers who heard right-dislocated sentences in alternation with intransitive sentences, did not interpret the novel verb "daser" as referring to a two-participants action. This suggests that enriching the syntactic context of right-dislocated sentences with simple intransitive sentences helped children to correctly interpret right-dislocated sentences as intransitives, in accordance with their prosodic structure.

Discussion

Experiment 1 shows that from 28 months old, French toddlers are able to integrate the information provided by different syntactic structures to correctly build their syntactic interpretations and to constrain their interpretation of a novel verb meaning. Strikingly, toddlers were able to flexibly adjust their reliance on the prosodic information (leading to an intransitive interpretation) compared to the number of NPs in the sentence (leading to a transitive interpretation), depending on the different syntactic contexts in which the novel verb appeared.

While previous research (Dautriche et al., 2014) showed that toddlers incorrectly associated a novel verb embedded into right-dislocated sentences to a causal action between two participants, as did children who listened to transitive sentences, in the current study we show that when the syntactic context is enriched with simple intransitive sentences (in both the transitive and right-dislocated conditions), children became able to correctly interpret right-dislocated sentences containing a novel verb and to assign different interpretations for right-dislocated and transitive sentences. This suggests that by simply observing that the novel verb in the right-dislocated sentences can also appear in intransitive sentences, toddlers were able to boost their reliance on the prosodic information of right-dislocated sentences rather than on the number of NPs in the sentence, to constrain their interpretation of the meaning of the novel verb.

Given that the only difference between the transitive+intransitive and the right-dislocated+intransitive conditions was the prosodic structure of the sentences, this suggests that toddlers can exploit the prosodic structure of a right-dislocated sentence to recover its syntactic structure and to constrain their learning of novel verb meanings.

As expected, children in the right-dislocated-only condition incorrectly associated the novel verb to the causal action, as did children in the transitive-intransitive condition. This replicates the results of Dautriche et al. (2014), and shows that in the absence of any additional information about the syntactic contexts in which the novel verb can appear, toddlers listening to right-dislocated sentences containing a novel verb still prefer to apply the canonical surface-to-meaning mapping rather than to integrate the prosodic information in their interpretations.

Additionally, while previous studies show that toddlers construct different interpretations for novel words presented in transitive sentences versus intransitive sentences, in the current experiment we extend these findings showing that toddlers can still assign a transitive interpretation to a novel verb even when the syntactic context in which this verb appears alternate between transitive and intransitive sentences.

Consistently with these results, Naigles, Bavin and Smith (2005) showed that 22-to-27-month-olds who learned the meaning of a novel verb introduced in transitive sentences, still interpret the verb as referring to a two-participants action, even when they heard this verb in an intransitive structure. This suggests that by the age of 28 months, toddlers may have already noticed that many familiar verbs relating to two participants (e.g., to eat) can enter in an alternating pattern between transitive ("He is eating the cake") and intransitive sentences in which the object is sometimes dropped ("He is eating"). Thus, participants were not prevented from assigning a transitive interpretation to a novel verb, when the syntactic context showed that the verb appeared both in transitive and intransitive sentences.

Taken together, our results show that toddlers are able to do sophisticated inferences about the argument structure of verbs and to adjust their syntactic interpretations following the information provided by different syntactic contexts. Relying on the fact that toddlers interpret novel verbs presented in right-dislocated sentences as transitive, the current study shows that they can change this wrong interpretation, when the syntactic context is enriched with simple intransitive sentences. A question that arises is whether other sources of information could also help children to recover from the noun-counting parsing bias.

In real life, in addition to syntactic information, listeners can also sometimes observe the scenario in which the actions take place and they can use this information to constrain their interpretation of novel verb meanings (e.g., Gillette, Gleitman, Gleitman, & Lederer, 1999; Gleitman, 1990). Thus, it is possible that the visual context in which the actions occur could also influence toddlers' interpretations of novel verb meanings and the reliance on different linguistic cues depending on what they can observe. In the case of right-dislocated sentences, for instance, we wondered whether children would consider the intransitive interpretation of right-dislocated sentences to be more plausible than the transitive interpretation, if at the same time as they hear these sentences, they could observe an action that is compatible with the intransitive interpretation provided by prosody.

In order to investigate these questions, we conducted a second experiment asking whether the observation of the semantic representations of the actions during sentence processing would impact children's online processing of right-dislocated sentences and their interpretations of novel verb meanings.

Experiment 2

In Experiment 2, we tested the hypothesis that offering toddlers the opportunity to observe the semantic representation of an event during the online processing of sentences would help them to avoid the parsing biases attested in previous studies with right-dislocated sentences, and increase their reliance on the prosodic information to interpret the novel verbs as intransitives. Toddlers listened to right-dislocated sentences such as "*Il$_i$ fome, le garçon$_i$*" - He$_i$ is foming, the boy$_i$ (meaning the boy is foming), or transitive sentences such as "*Il$_i$ fome le garçon$_k$*" - He$_i$ is foming the boy$_k$ (meaning someone is foming the boy) while watching two videos side-by-side on a TV-screen: one video showing a person doing a self-generated action (one participant action) and another video showing a person acting on another (a causative action between two participants).

If offering the possibility of observing both the transitive and the intransitive interpretations for a novel verb during sentence processing impacts children's on-line interpretation of the right-dislocated sentences, then children in the right-dislocated condition should look less towards the two-participants action than children in the transitive condition. Additionally, another group of toddlers listened to intransitive sentences such as "*Il dase*" – He is dasing, to provide us a baseline on how children behave in this experiment when they interpret a novel verb as intransitive. Based on previous studies showing that toddlers who listened to transitive dialogues look reliably longer at the two-participant event than those who listened to intransitive dialogues (e.g., Yuan & Fisher, 2009; Yuan et al., 2012), we expected that children in the intransitive condition should look less to the two-participants action than children in the transitive condition.

Method

Participants. Seventy-two 28-month-olds (ranging in age from 26.7 to 29.0, M_{age}=27.7; *SD*= 0.5; 36 girls) participated in this experiment. All were monolingual native French speakers with less than 20% exposure to another language. Twenty-four toddlers were randomly assigned to each of the three experimental conditions (transitive, right-dislocated, and intransitive). An additional twelve toddlers completed the experiment but they were not included in the final sample because of fussiness during the experiment resulting in more than 50% of trials with missing eye-tracking data (n = 9), because they cried during the experiment (n = 2), or because of technical problems (n = 1). Parents signed an informed consent form. This research was approved by the local ethics committee.

Apparatus. The apparatus was similar to that of Experiment 1, although the experiment took place in a different sound-attenuated booth: IAC Acoustics, and with a different screen for stimuli presentation: a 27-in television screen. As before, toddlers sat on their parent's lap about 70 cm away from the television and, their eye-gazes toward the stimuli were recorded by an eye-tracker (Eye-link 1000), placed below the TV-screen and operating in a remote mode with a time-sample collected every 2ms. The caregivers wore headphones and listened to masking music during the experiment. The experimenter stayed outside the cabin during the test.

Materials. Materials consisted of four pairs of color videos showing people performing novel actions. For each pair of videos, one video presented a person doing a self-generated action (one-participant action) and another video presented a person acting on another person (a causative action between two participants). Each pair of videos was used to illustrate the possible interpretations of each of the four novel verbs used in the experiment: “fomer”, “daser”, “raner’, “nuver”, (see Figure 5 for action descriptions).

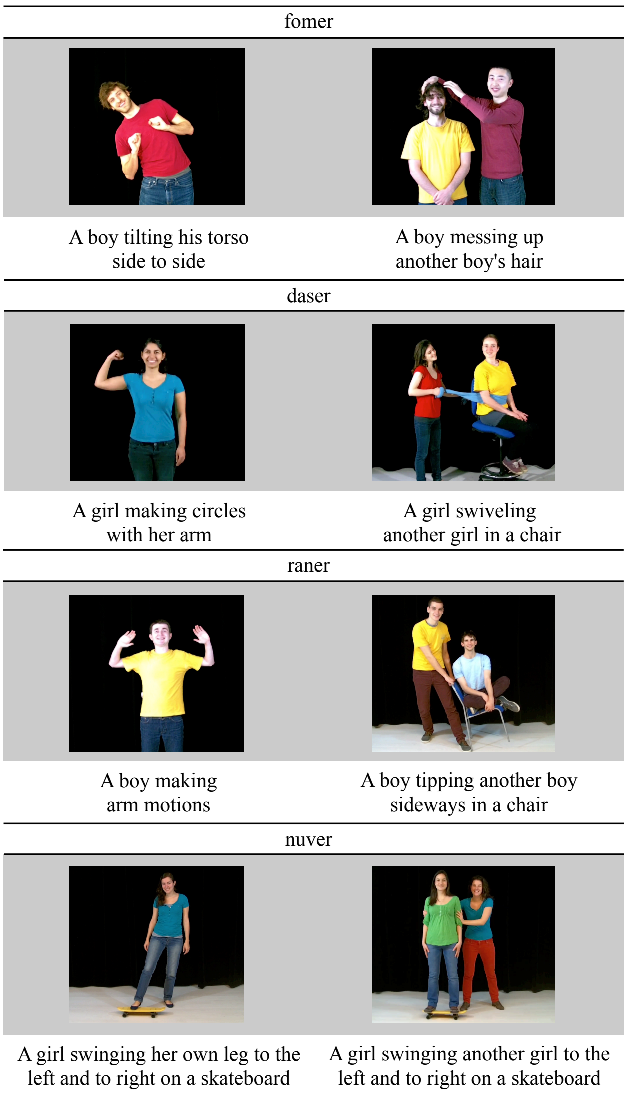


Figure 5: Novel verbs and actions used in Experiment 2

These videos offered toddlers the option of interpreting the postverbal NP (e.g., "le garçon" – the boy) either as the patient of the action (appropriate for the transitive condition, "Il_i fome le $garçon_k$" - He_i is foming the boy_k) which can be observed only in the two-participants action, or as the agent of the action (appropriate for the dislocated condition, "Il_i fome, le $garçon_i$" - He_i is foming, the boy_i) which can be observed in both videos. Additionally, two videos illustrating familiar actions were created. One video presented a familiar intransitive action ("*marcher*" – to walk), and showed a boy walking, and the other video presented a familiar transitive action ("porter" – to carry), and showed a boy carrying another boy. For familiar and novel-verbs actions, all actors on screen had the same gender, so that toddlers could not use the gender of the pronouns to find which action was talked about. Each actor appeared in only one video.

All videos were accompanied by sound tracks recorded by a female native French speaker (last author), who uttered all sentences in child-directed speech. These sound tracks presented the novel verbs in each of the three experimental conditions (transitive, right-dislocated or intransitive). The sound tracks for the transitive condition presented the target verb in transitive structures such as "[*Tu vois ça*?] [*Il_i fome le $garçon_k$*]. [*Regarde*!] [*Il_i fome le $garçon_k$*]" – Do you see that? He_i is foming the boy_k. Look! He_i is foming the boy_k. In the right-dislocated condition the target verb was presented in right-dislocated structures such as "[*Tu vois ça*?] [*Il_i fome*] [*le $garçon_i$*]. [*Regarde*!] [*Il_i fome*] [*le $garçon_i$*]" – Do you see that? He_i is foming, the boy_i. Look! He_i is foming, the boy_i). In the intransitive condition the novel verb was presented in intransitive structures such as "[*Tu vois ça*?] [*Il fome*]. [*Regarde*!] [*Il fome*]" – Do you see that? He is foming. Look! He is foming. Note that for each sound track, for each trial, the target verb was repeated twice.

In order to assess the prosodic differences between the two critical conditions that were composed of exactly the same words but differed only with respect to their prosodic structure, reflecting their different syntactic structures (transitive vs. right-dislocated sentences), acoustical analyses of duration and pitch were conducted on these test sentences. The analyses revealed that as in Experiment 1, there was a clear prosodic boundary between the verb and the final NP in the right-dislocated sentences but not in the transitive sentences. This was attested by a significant preboundary syllable lengthening on the last syllable of the novel verb (e.g., fome) preceding the final NP (e.g., "le garçon" – the boy; M_{trans} = 444ms vs. M_{disloc} = 633 ms, $t(15) = 12.91$, $p < .001$), and an important pitch drop between the max pitch observed on the verb minus the max pitch observed on the first vowel of the final NP in the right-dislocated sentences compared to the same segments on the transitive sentences (M_{trans} = -45Hz vs. M_{disloc} = -151Hz), $t(15) = 5.27$, $p < .001$). As in experiment 1, there was no pause between the verb and the postverbal NP. In a pilot experiment, naive French adults (n=10) were asked to decide who was performing the action in each of these test sentences: they interpreted the transitive and dislocated sentences correctly 92% of the time.

Procedure. The procedure included five items: one practice item involving a familiar intransitive verb ("*marcher*" – to walk) common to all participants, and four novel-verb test items ("fomer", "daser", "raner', "nuver") presented in one of the three experimental conditions, in a between-participants design. Each item included two 8s trials in which a pair of video events was presented together with the sentences.

In order to introduce toddlers to the task, the procedure began with the practice item, involving the familiar intransitive verb "*marche*" – to walk. The target video showed a man walking (one-participant action) and the distractor video showed a different man carrying another man (two-participants action; see Figure 6).

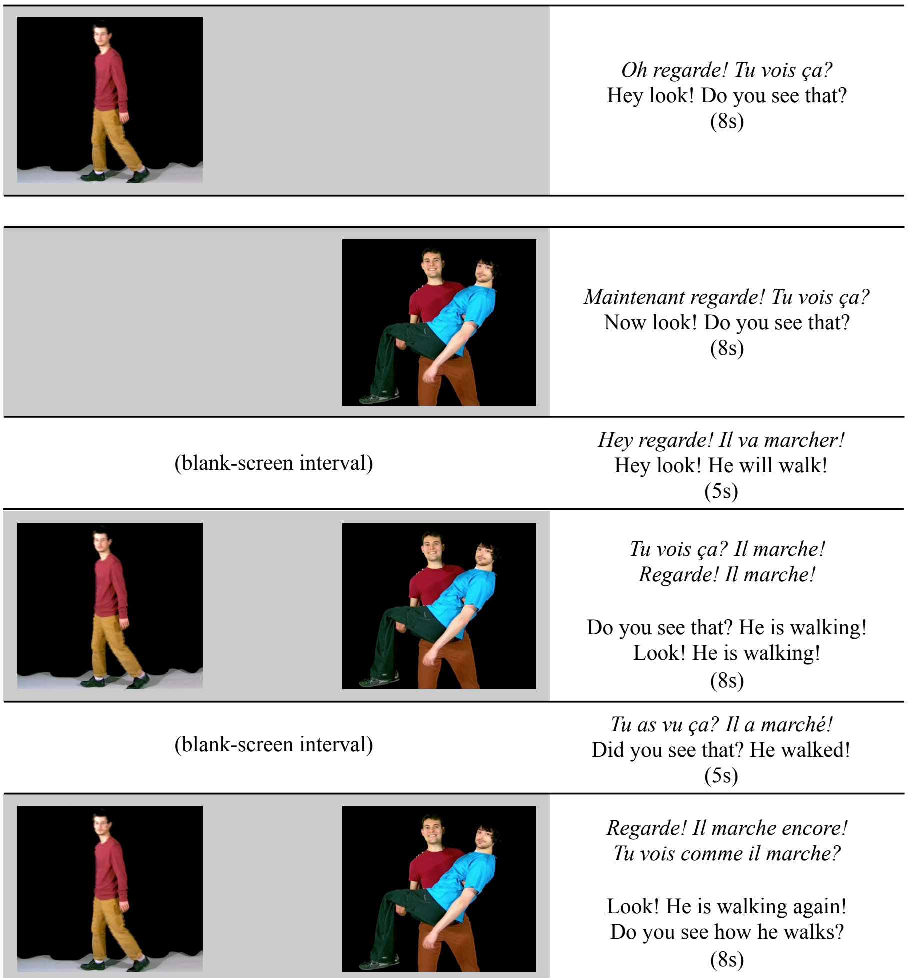


Figure 6: Time-course of the practice familiar-verb item presentation. The novel-verb items were presented in the same way.

In this practice trial, all participants were asked to look towards the one who was walking. This practice trial served to show children that in this experiment, one of the two videos matched the soundtrack they heard. Toddlers' eye-gaze towards the videos was recorded by an eye-tracker during the test trials. The time course of the practice trial is illustrated in Fig 6. The side of the target video (left or right) was counterbalanced across participants.

Each trial started with an inspection period, to provide toddlers enough time to inspect each of the videos individually, on each side of the TV-screen (8s for each video). Five hundred milliseconds later, both videos disappeared, and a sentence containing the verb was presented during a 5-s blank-screen interval (e.g., "Hey, look! He will walk!"). Next, the two videos appeared side-by-side on the screen for 8s, and at the same time participants heard the test sentence repeating the verb twice (e.g., Do you see that? He is walking! Look! He is walking!). This test phase was repeated twice (see Figure 6). Thus in total, for each item, participants heard 6 repetitions of the target word (4 repetitions while watching the two videos side-by-side, 2 in each trial; and two repetitions during the blank-screen interval without no videos on the screen, one in each interval).

The four novel-verb test items were presented exactly in the same way described for the practice item in Figure 6. The side of the test video presentations was counterbalanced within participants, such that for half of the items, a given participant saw the one participant action on the left and the two participants action on the right and for the other half, she had the reverse. The order of presentation of the novel-verb items was random. Depending on condition, toddlers heard the four novel verbs in transitive sentences, or in right-dislocated sentences, or they heard them in intransitive sentences. Figure 7 shows an example of the two test trials for a given novel verb and the kind of sentence that was presented in each condition.

Pair of actions accompanying the novel verb **fomer**

Trial 1

One participant action | Two participants action

Right-dislocated: *Tu vois ça? Il fome, le garçon. Regarde! Il fome, le garçon.*
Do you see that? He is foming, the boy. Look! He is foming, the boy.

Transitive: *Tu vois ça? Il fome le garçon. Regarde! Il fome le garçon.*
Do you see that? He is foming the boy. Look! He is foming the boy.

Intransitive: Tu vois ça? Il fome. Regarde! Il fome.
Do you see that? He is foming. Look! He is foming.

Trial 2

One participant action | Two participants action

Right-dislocated: *Regarde! Il fome encore, le garçon. Tu vois comme il fome, le garçon?*
Look! He is foming again, the boy. Do you see how he fomes, the boy?

Transitive: *Regarde! Il fome encore le garçon. Tu vois comme il fome le garçon?*
Look! He is foming the boy again. Do you see how he fomes the boy?

Intransitive: *Regarde! Il fome encore. Tu vois comme il fome?*
Look! He is foming again. Do you see how he fomes?

Figure 7: Pair of videos and critical sentences used for the novel verb fomer. All the other novel verbs in each condition used the same kind of syntactic frames for each trial.

Data processing and analysis. In this experiment, during the test phase, participants listened to different sound files depending on the condition to which they were assigned. Similarly to Experiment 1, our prediction is that looking times towards the two-participants action would be influenced by the type of sentences that participants heard. To assess these differences, we conducted 3 cluster-based permutation analysis comparing conditions pairwise, to identify time-windows where a significant difference between conditions occurred. Forty-eight trials out of 576 were removed from the statistical analysis because more than 25% of the data frames in the analyzed time-window were missing (20 in the right-dislocated condition, 18 in the transitive condition and 10 in the intransitive condition).

Results

Figure 8 shows the proportion of looks towards the two-participants action for toddlers in the transitive condition (blue curve), the right-dislocated condition (red curve) and in the intransitive condition (green curve), time-locked to the beginning of the test trials.

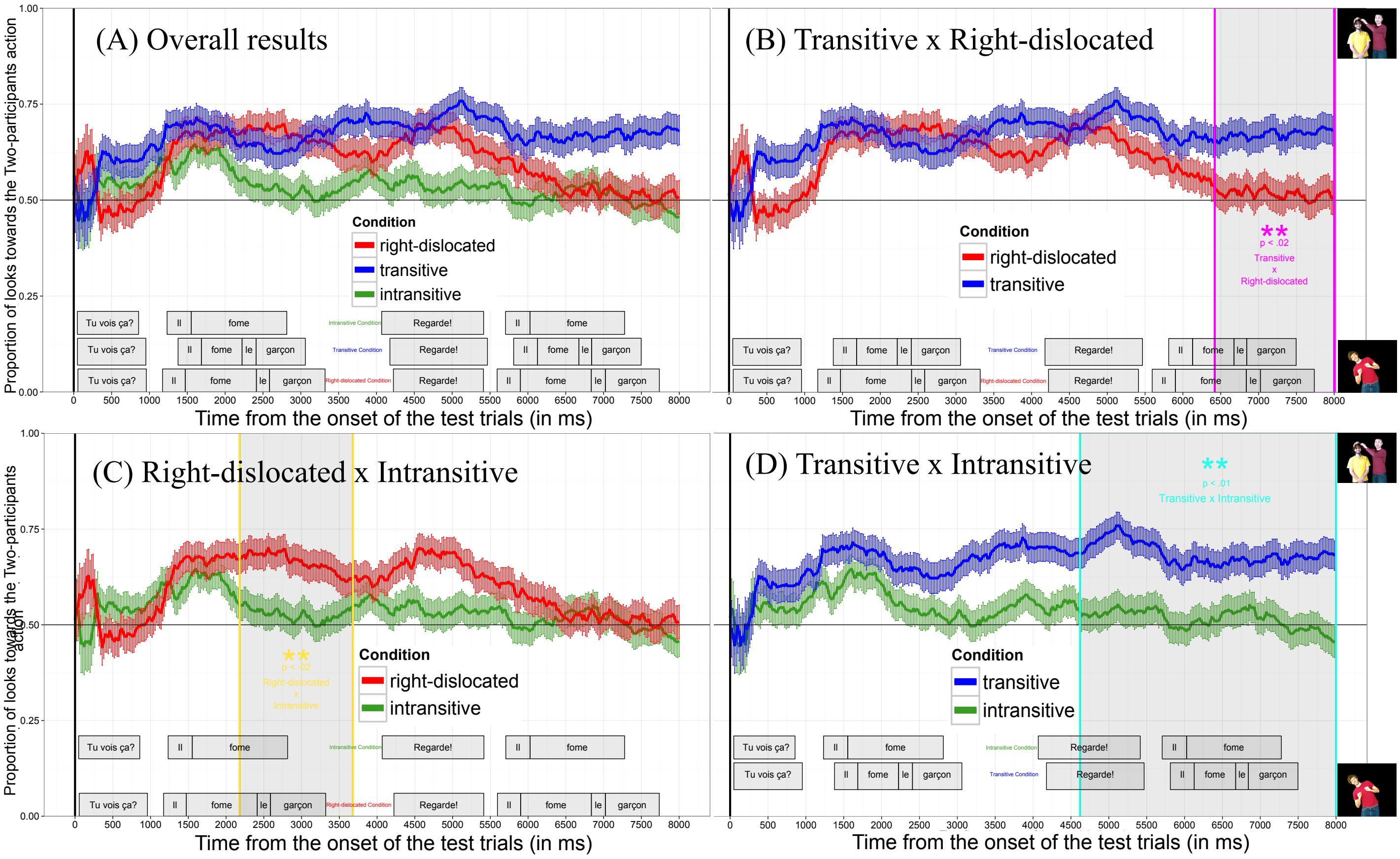


Figure 8: (A) Proportion of looks towards the Two-participants action, time-locked to the onset of the test trials (vertical black line) for toddlers in the transitive condition (blue curve), the right-dislocated condition (red curve) and in the intransitive condition (green curve). Error bars represent the standard error of the mean. (B) The transitive and right-dislocated conditions differed from each other from the second repetition of the novel verbs (about 6420 ms after the onset of the test trials until the end of the trials). (C) The transitive and right-dislocated conditions differed from each other from the first repetition of the novel verbs (from 2180 until 3680 ms after the beginning of the test trials). (D) The transitive and intransitive conditions differed from each other slightly after the offset of the first sentence in the test trials (from 4620 ms after the beginning of the test trials until the end of the trials).

Visual inspection of the data reveals that all groups of toddlers tended to look more toward the two participants action right after the beginning of the trials. Toddlers in the transitive condition showed the greater preference toward the two participants action, and looked more toward the two-participants action during the entire test phase. Toddlers in the intransitive condition (green curve) increased their looks toward the one-participant action as soon as they heard the first repetition of the critical verb, then remained around chance (.05) until the end of the trial (see Fig 8-D). Crucially, we can observe that toddlers in the right-dislocated sentences initially behaved as toddlers in the transitive group and looked more towards the two participants action during the first repetition of the target word. However, from the second repetition, they increased their looks towards the one-participant action and behaved as children in the intransitive condition (see Fig 8-B).

The cluster-based analysis found a significant time-window where the proportion of looks towards the two-participants action was significantly different between toddlers in the transitive condition and toddlers in the right-dislocated condition (Fig 8-B), from 6420 ms until the end of the test-trial ($p < 0.02$) (corresponding to the second repetition of the novel verbs). When we compared the proportion of looks towards the two-participants action between toddlers in the right-dislocated condition and toddlers in the intransitive condition, a significant time-window was found between 2180ms and 3680 ms; $p < 0.02$; which coincides with the first repetition of the novel verbs (Fig 8-C). Finally, the cluster-based analysis found a significant time-window where the proportion of looks towards the two-participants action was significantly different between toddlers in the transitive condition and toddlers in the intransitive condition (Fig 8-D), from 4620 ms until the end of the trial ($p < 0.01$).

Taken together, these results show that toddlers' looking preferences towards the two-participants action was affected by the type of sentence in which they heard the novel verbs. Toddlers who listened to transitive sentences looked longer at the two-participants action than did toddlers who listened to the same novel verbs in right-dislocated sentences, or in intransitive sentences. Importantly, toddlers in the right-dislocated condition were able to integrate the visual aspect of the task to exploit the prosodic information of the right-dislocated sentences and to guide their interpretation of the novel verbs, behaving differently from toddlers in the transitive condition.

Discussion

The results obtained here provide evidence that 28-month-olds can integrate visual information to adapt their interpretations of right-dislocated sentences and to rely more on the information provided by prosody (rather than on the number of NPs in the sentence) to constrain their online interpretation of novel verbs meaning. In the current experiment, toddlers who listened to a novel verb in right-dislocated sentences such as "[*Il*$_i$ *fome*], [*le garçon*$_i$]" (meaning the boy is foming), while watching a one-participant action versus a two-participants action, behaved differently from toddlers who listened to a novel verb in transitive sentences such as "[*Il*$_i$ *fome le garçon*$_k$]" (meaning someone is foming the boy). The results showed that toddlers in the transitive condition looked more to the two-participants action than toddlers in the right-dislocated condition. These results suggest that allowing toddlers to visually inspect a one-participant (intransitive) action while listening to the novel verbs in the right-dislocated sentences helped them to correctly interpret such a structure as intransitive, and to recover from a parsing bias based on the number of NPs in the sentence.

It is important to note however, that although both right-dislocated sentences and intransitive sentences support an intransitive interpretation for the novel verbs, toddlers did not interpret these sentences in the same way. As can be observed in Figure 8-C, during the first repetition of the novel verbs in the test phase, toddlers in the right-dislocated sentences looked more towards the two-participants action than toddlers in the intransitive sentences. This suggests that the presence of the two NPs in the right-dislocated sentences made these sentences harder to interpret than simple intransitive sentences.

General Discussion

Across two experiments, we tested toddlers' ability to flexibly adjust their reliance on different linguistic cues (i.e., on the prosody or on the number of NPs in the sentence), depending on the information provided by the syntactic context (in Experiment 1) or by the visual context (in Experiment 2). Previous studies have established that to understand sentences, adults integrate their prior expectations about likely utterances with the information they extract from the input, but depending on the level of uncertainty of specific cues, in a given environment, they can adjust their prior linguistic expectations to weigh the plausibility of different information sources. Here we provide evidence that 28-month-old toddlers are able to engage in a similar process when interpreting novel verbs. Toddlers were able to flexibly adapt their reliance on the prosodic structure of right-dislocated sentences (leading to an intransitive interpretation) or on the number of NPs in a sentence (leading to a transitive interpretation), depending on the information provided by the syntactic context (in Experiment 1) or by the visual context (in Experiment 2).

In Experiment 1, we added simple intransitive frames containing a novel verb to the dialogues used in Dautriche et al. (2014) with right-dislocated sentences, such that immediately after having heard a right-dislocated sentence, toddlers heard an additional syntactic frame in which the novel verb appeared in a simple intransitive frame. This additional information seems to have increased toddlers' reliance on the information provided by prosody in the right-dislocated sentences. Assuming that when toddlers listen to right-dislocated sentences they can construct both an intransitive interpretation (integrating prosodic information, as they did for familiar verbs) and a transitive interpretation (based on the number of NPs in the sentence, as they did for novel verbs), the presence of the simple intransitive frames containing the novel verb seems to have allowed children to assign more weight to the intransitive interpretation provided by prosody rather than relying on the information provided by the number of NPs in the sentence. The results confirmed this hypothesis, showing that contrary to the transitive+intransitive condition and similarly to the intransitive-only condition, children in the right-dislocated+intransitive condition did not show any preference for the two-participants action during the test.

The benefit of adding the intransitive sentences to make toddlers adjust their interpretation was evident when one compares toddlers in the right-dislocated+intransitive condition with those in the right-dislocated only condition. The results observed for toddlers in the right-dislocated-only condition (i.e., the group without the additional intransitive hints), replicate the results observed in Dautriche et al., (2014) and show that in the absence of any additional information about the novel verb, toddlers rely on the information provided by the NPs in the sentence and mistakenly interpreted the novel verb as transitive. Thus, being exposed to multiple syntactic frames for a single verb allowed children to recover from parsing biases when interpreting non-canonical sentences with the structure N+V+N.

Consistently with our results, other studies also provide evidence that toddlers can use multiple syntactic frames to learn the meaning of a novel verb (e.g., Naigles, 1996; Naigles et al., 2005) and that sometimes they can exploit additional cues on the additional syntactic structures provided to avoid parsing bias (Arunachalam, Escovar, Hansen, & Waxman, 2013). For instance, while previous studies showed that 21-month-olds mistakenly interpret sentences such as "The boy and the girl are gorping" as referring to a causative action between two participants, further research demonstrated that toddlers do not construct the same inference if just after this sentence, they hear another intransitive frame with a pronoun in the place of the conjoined subject (e.g., "The boy and the girl are gorping. Really? They are going to gorp?"; Arunachalam et al., 2013). Note that in the transitive condition, this study presented children with only transitive frames (e.g., "The boy gorped the girl. Really? The boy gorped the girl?"). In this case, pronominalization was the cue that might have helped toddlers to process "The boy and the girl" as a single NP-subject in the intransitive condition. In our study however, note that both the transitive+intransitive and the right-dislocated+intransitive conditions contain the same intransitive syntactic frames and all sentences were composed of exactly the same words. Thus, toddlers' success in our task can not be simply explained by the presence of disambiguating syntactic information, but rather by the rational integration that toddlers were able to do when processing transitive+intransitive sentences in one condition versus right-dislocated+intransitive sentences in the other. Transitive+intransitive sentences gave rise to a final transitive interpretation for the novel verb, while right-dislocated+intransitive sentences gave rise to a final intransitive interpretation.

In Experiment 2, we asked whether toddlers would be able to flexibly adjust their reliance on the prosodic information in comparison to the strategy based on the number of NPs in the sentence, depending on the visual information illustrating the semantic representation of the actions during sentence processing. The results showed that offering the possibility of observing the intransitive interpretation side-by-side with the transitive interpretation, while toddlers listened to right-dislocated sentences, impacted their on-line interpretation. Toddlers listening to a sentence such as "Look! He_i is *foming*, the boy_i", initially preferred to look more towards the two-participants action, as did toddlers in the transitive condition. However, during the experiment, while inspecting the two videos together, toddlers seem to have adapted their parsing strategy following the visual information, and to have accessed the intransitive interpretation for the novel verb in right-dislocated sentences. Thus, they started to look less towards the two-participants action than toddlers in the transitive condition. Importantly, during the second onset of the target words, toddlers in the right-dislocated condition behaved exactly as toddlers in the intransitive condition. This suggests that the integration of visual cues together with the online processing of the sentences, allowed toddlers to adjust their syntactic interpretations of right-dislocated sentences and to recover from parsing bias based on the number of NPs.

Given that in our study both right-dislocated and transitive sentences were composed of exactly the same words, but differed only with respect to the prosodic structure of the sentences, reflecting their different syntactic structures, the only way to interpret the difference between toddlers in the transitive compared to the right-dislocated condition in experiment 2 for instance, is that the latter were able to exploit the prosodic information of the right-dislocated sentences to constrain their online interpretation of the novel verbs.

These results contribute to the hypothesis that phrasal prosody can be an important source of information for children during the early steps of language acquisition and that it provides useful information for syntactic acquisition (Morgan, 1986; Morgan & Demuth, 1996). While recent studies show that from 20 months, toddlers are able to use phrasal prosody to constrain their syntactic analysis and to guide the interpretation of an ambiguous noun-verb homophone (de Carvalho, Dautriche, et al., 2016; de Carvalho et al., 2017; de Carvalho, Lidz, et al., 2016), in the current study we show that toddlers can exploit phrasal prosody to calculate the argument structure of verbs and to constrain the acquisition of novel verb meanings.

Recent studies suggest that because linguistic communication is often noisy and uncertain, the ability to successfully interpret sentences requires listeners to integrate bottom-up cues from speech perception with top-down expectations about what speakers are likely to say, and to weigh the importance attributed to certain linguistic cues depending on the context (Fine, Jaeger, Farmer, & Qian, 2013; Gibson et al., 2013; Jaeger, 2010; Kleinschmidt & Jaeger, 2015). Several studies investigating the noisy-channel account framework demonstrate that adults can indeed flexibly adapt how much they rely on certain kinds of linguistic cues depending on their reliability: they adjust the weight they attribute to phonological, semantic and syntactic information according to experiment-specific statistics (e.g., Fine et al., 2013; Gibson et al., 2013).

In a very recent study with preschoolers, 4-to-5-year-olds were also shown to be able to flexibly adapt to linguistic input, in accordance with the noisy channel account (Yurovsky, Case, & Frank, 2016). In that study, during a familiarization phase, children were exposed either to a 'reliable speaker' who always uttered plausible sentences, such as "my cat has three little kittens", or they were exposed to an 'unreliable speaker' who always uttered implausible sentences such as "my cat has three little hammers". During the test phase, all participants listened to test sentences containing a target word that was really implausible in the situation, but that could be easily re-interpreted as plausible if listeners inferred that one single phoneme was mispronounced. For instance, participants heard "I had carrots and bees for dinner" instead of "I had carrots and peas for dinner". Participants had to choose between two pictures: one depicting the plausible interpretation for the sentence (carrots and peas on a plate) and the other representing an implausible interpretation (carrots and bees on a plate). The results showed that participants exposed to the semantically 'reliable' speaker chose more often the plausible picture (assuming that the speaker said 'peas' instead of 'bees') whereas participants exposed to the semantically unreliable speaker chose more often the implausible picture (with the bees). These recent results suggest that children can deal flexibly with linguistic input and change the importance they attribute to linguistic cues (here phonological representations of words) depending on the context. In our current study we extend these findings providing evidence that even younger children of 28 months are able to engage in a similar process when interpreting the syntactic structure of sentences and inferring the meaning of novel verbs.

Taken together, the experiment reported here suggest that toddlers are able to solve sophisticated inferences about the argument structure of verbs and to adjust their syntactic interpretations following the information provided by different syntactic frames (in Experiment 1) or by the visual context presented during sentence processing (in Experiment 2). Similar to what has been shown with adults, toddlers were able to adjust their prior linguistic expectations to weigh the plausibility of different information sources during language processing. This ability might be extremely important for toddlers during language acquisition because it would allow them to evaluate the plausibility of different levels of linguistic information during parsing and to flexibly adapt their reliance on different linguistic cues to reach different sentences' interpretation and to constrain their acquisition of word meanings.

Acknowledgements

This research was supported by a PhD fellowship from the École normale supérieure to Alex de Carvalho, and by an ESRC Future Research Leaders award (ES/N017404/1) to Isabelle Dautriche. It was also supported by grants from the Région IIe-de-France, Fondation de France, LabEx IEC (ANR-10-LABX-0087), IdEx PSL (ANR-10-IDEX-0001-02), as well as the ANR 'Apprentissages' (ANR-13-APPR-0012). The authors thank all the bright "actors" who participated in the stimuli creation for the videos illustrating actions and Anne-Caroline Fiévet and Laia Fibla, for their assistance with recruitment of participants and data collection as well as all the parents and babies who participated in this study.

References

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

Arunachalam, S., Escovar, E., Hansen, M., & Waxman, S. R. (2011). Verb Learning from Syntax Alone at 21 Months.

Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition*, *114*(3), 442–446. http://doi.org/10.1016/j.cognition.2009.10.015

Christodoulopoulos, C., Roth, D., & Fisher, C. (2016). An incremental model of syntactic bootstrapping. In *Proceedings of the Seventh Workshop on Cognitive Aspects of Computational Language Learning*.

Connor, M., Gertner, Y., Fisher, C., & Roth, D. (2008). Baby {SRL:} Modeling Early Language Acquisition. *Proceedings of the Twelfth Conference on Computational Natural Language Learning, CoNLL 2008, Manchester, UK, August 16-17, 2008*, 81–88.

Dautriche, I., Cristia, A., Brusini, P., Yuan, S., Fisher, C., & Christophe, A. (2014). Toddlers Default to Canonical Surface-to-Meaning Mapping When Learning Verbs. *Child Development*, *85*(3), 1168–1180.

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

de Carvalho, A., Dautriche, I., & Christophe, A. (2016). Preschoolers use phrasal prosody online to constrain syntactic analysis. *Developmental Science*, *19*(2), 235–250. http://doi.org/10.1111/desc.12300 de Carvalho, A., Dautriche, I., Lin, I., & Christophe, A. (2017). Phrasal prosody constrains syntactic analysis in toddlers. *Cognition*, *163*, 67–79. http://doi.org/10.1016/j.cognition.2017.02.018

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

de Villiers, J. G., & de Villiers, P. A. (1973). Development of the use of word order in comprehension. *Journal of Psycholinguistic Research*, *2*(4), 331–41. http://doi.org/10.1007/BF01067055

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion, *415*(January), 429–433.

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

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE*, *8*(10). http://doi.org/10.1371/journal.pone.0077661

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

Fisher, C. (2002). The role of abstract syntactic knowledge in language acquisition: a reply to Tomasello (2000). *Cognition*, *82*(3), 259–78.

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning Words and Rules. *Psychological Science*, *17*(8), 684–692.

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(20), 8051–6. http://doi.org/10.1073/pnas.1216438110

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

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

Hahn, N., Snedeker, J., & Rabagliati, H. (2015). Rapid Linguistic Ambiguity Resolution in Young Children with Autism Spectrum Disorder: Eye Tracking Evidence for the Limits of Weak Central Coherence. *Autism Research*, *8*(6), 717–726. http://doi.org/10.1002/aur.1487

He, A. X., & Lidz, J. (2017). Verb learning in 14- and 18-month-old English-learning infants. *Language Learning and Development*, *5441*(March).

Jacobs, R. A. (1999). Optimal integration of texture and motion cues to depth, *39*, 3621–3629.

Jaeger, F. T. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, *61*(1), 23–62. http://doi.org/10.1016/j.cogpsych.2010.02.002

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception : Recognize the familiar , generalize to the similar , and adapt to the novel. *Psychological Review*, *122*(2), 148–203. http://doi.org/10.1037/a0038695

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

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177. http://doi.org/10.1016/j.cognition.2007.05.006

Lidz, J., Gleitman, H., & Gleitman, L. (2003). Understanding how input matters : verb learning and the footprint of universal grammar. *Cognition*, *87*, 151–178.

Maratsos, M. P. (1974). Children who get worse at understanding the passive: A replication of Bever. *Journal of Psycholinguistic Research*, *3*(1), 65–74. http://doi.org/10.1007/BF01067222

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

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

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, *17*(2), 357–74.

Naigles, L. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition*, *58*(2), 221–251. http://doi.org/10.1016/0010-0277(95)00681-8

Naigles, L., Bavin, E., & Smith, M. (2005). Toddlers recognize verbs in novel situations and sentences. *Developmental Science*, *8*(5), 424–31. http://doi.org/10.1111/j.1467-7687.2005.00431.x

Scott, R. M., & Fisher, C. (2009). 2-year-olds use distributional cues to interpret transitivity-alternating verbs. *Language and Cognitive Processes*, *24*(6), 1–25.

Slobin, D. I., & Bever, T. G. (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition*, *12*(3), 229–265. http://doi.org/10.1016/0010-0277(82)90033-6

Tanenhaus, M. K., Spivey, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of Visual and Linguistic Information in Spoken Language. *Science*, *268*(5217), 1632–1634.

Trueswell, J., & Kim, A. (1998). How to Prune a Garden Path by Nipping It in the Bud : Fast Priming of Verb Argument Structure. *Journal of Memory and Language*, *123*(39), 102–123.

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

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. http://doi.org/10.1111/j.1467-9280.2009.02341.x

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

Yurovsky, D., Case, S., & Frank, M. C. (2016). Preschoolers Flexibly Adapt to Linguistic Input in a Noisy Channel. *Psychological Science*. http://doi.org/10.1177/0956797616668557.

Chapter 7

Prosody cues the acquisition of word meanings in 18-month-old infants

Prosody cues the acquisition of word meanings in 18-month-old infants

Alex de Carvalho[a,b,*], Angela Xiaoxue He[c], Jeffrey Lidz[d] and Anne Christophe[a,b]

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

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

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

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

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

*. To whom correspondence and material requests should be addressed. Email: x.de.carvalho@gmail.com; Address: 29 rue d'Ulm – 75005 Paris, France. Telephone number: +33.1.44.32.29.82

Abstract

Language is a defining feature of being human, and its acquisition presents a formidable task to babies during their development. Extracting words from fluent speech and assigning meanings to them, for instance, is a crucial yet challenging step. Previous research has demonstrated the importance of syntax (the rules for combining words into sentences) as a cue to word meaning. But how can babies access such abstract structure without first knowing the meanings of the words? We investigate the contribution of phrasal prosody (speech melody), which is directly accessible in speech and correlates with syntactic structure in all the world's languages. We show for the first time that 18-month-old infants use prosody to recover sentences' syntactic structure, which in turn constrains the possible meaning of novel words: participants interpreted a novel word as referring either to an object or an action, given its position within the prosodic-syntactic structure of sentences.

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

Prosody cues the acquisition of word meanings in 18-month-old infants

1. Introduction

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

Syntactic structure, which governs the organization of words into sentences, has been proposed to be a universal and reliable source of information that children may exploit to discover the meaning of words (Fisher, et al., 1994; Gillette, et al., 1999; Gleitman, 1990). In the simplest case to illustrate this idea, studies have shown that by two years of age, children can infer that a novel word, such as *dax*, refers to a novel action when it occupies a verb position in a sentence, as in *He is daxing that*, or to a novel object when it occupies a noun position, as in *This is a dax* (e.g., Waxman, Lidz, Braun, & Lavin, 2009). These findings demonstrate that the syntactic structure in which the words occur is an important source of information for children: they can exploit the syntactic environment of a word to determine its syntactic category (e.g., a noun or a verb) and use the syntactic category to restrict the kind of meaning the novel word can have (e.g., verbs refer to actions).

This ability to exploit and learn from syntactic structures so early, although impressive, seems rather counterintuitive. Given that syntactic structure defines the relationships between words in a sentence, and allows listeners to compute the meaning of a sentence from the meaning of the individual words that compose it, one would expect that infants would first need to learn the words and their meanings, to then be able to learn how to organize words into sentences and therefore learn about syntax itself (e.g., Tomasello, 2000a, 2000b). We are thus faced with a chicken-and-egg problem: children seem to need words to learn syntax, but on the other hand they need syntax to learn words. How can infants avoid this circularity? In this paper we experimentally tested whether 18-month-olds, who do not yet know many words, were able to compute the syntactic structure of a sentence, by relying on phrasal prosody and function words, two sources of information that are available early during language acquisition and convey useful information about syntactic structure.

Phrasal prosody is the rhythm and melody of speech: when we speak, words are not pronounced one after the other in a monotone way, but rather they are grouped together into intonational units (i.e., prosodic phrases). For example, a sentence such as "*The little cat is running fast*" tends to be pronounced as [the little cat] [is running fast], where brackets represent prosodic units. This organization into prosodic units is important because it has been shown that in all the world's languages, the boundaries between prosodic units always coincide with syntactic boundaries[2] (e.g., Shattuck-Hufnagel & Turk, 1996). In our example, this boundary appears between the noun phrase, containing the words "the little cat", and the verb phrase containing the words "is running fast". Thus, by paying attention to this kind of prosodic information (e.g., phrase final lengthening, pauses, pitch contour discontinuity) in speech, infants may be able to infer the position of some syntactic boundaries, those that coincide with prosodic boundaries.

Function words and morphemes (i.e., highly frequent functional elements, such as articles, auxiliaries, pronouns, conjugation endings, etc.) are also salient to young children, because they are much more frequent than content words (nouns, verbs, adverbs) and have perceptible characteristics that distinguish them from content words (Shi, Morgan, & Allopenna, 1998). Because functional elements carry almost exclusively morpho-syntactic information, they may allow infants to determine the syntactic nature of the constituents delimited by prosodic boundaries. In a language such as English for example, if infants are able to learn that articles typically precede nouns while pronouns or other full noun phrases typically precede verbs, they may then exploit this information to categorize the constituents delimited by prosody (for instance, a prosodic constituent starting with the article "the" is likely to be a noun phrase: "[the blicks]"). Thus function words/morphemes and phrasal prosody, together, may allow young infants to build at least a rudimentary representation of the syntactic structure of sentences (Christophe, Dautriche, de Carvalho, & Brusini, 2016; Christophe, Millotte, Bernal, & Lidz, 2008). Supporting this hypothesis, computational work demonstrates that models relying on phrasal prosody and function words successfully predict the syntactic category of unknown words (Christodoulopoulos, Roth, & Fisher, 2016; Gutman, Dautriche, Crabbé, & Christophe, 2015).

The situation is thus as follows: on the one hand, phrasal prosody and function words are jointly predictive of syntactic structure in natural languages, and on the other hand, young

---

[2] The reverse is not true, since not all syntactic boundaries are marked with a prosodic boundary; for instance in short sentences such as [The boy eats], or [he eats], the syntactic boundary between the subject and the verb phrase is unmarked prosodically.

infants are sensitive to each of these sources of information: they are sensitive to phrasal prosody from birth (e.g., Mehler et al., 1988) and percieve prosodic cues marking the boundaries between groups of words (Shukla, White, & Aslin, 2011; Soderstrom, Seidl, Nelson, & Jusczyk, 2003). Within the first year of life, infants recognize some function words and they use them to categorize content words between 12 and 24-months old (Cauvet et al., 2014; Haryu & Kajikawa, 2016; He & Lidz, 2017; Shi & Melançon, 2010). What has never been investigated however is whether infants are able to jointly use phrasal prosody and function words to access the syntactic structure of sentences and constrain their acquisition of word meanings. This ability would be crucial for language acquisition because it would allow infants to break free of the chicken-and-egg problem, since phrasal prosody and function words are acquired well before infants know many words. For instance, 18 month-olds still have a relatively small vocabulary, and typically utter one word at a time, but they already have a good knowledge of phrasal prosody and function words in their native language. Having access to syntactic structure at such an early age would allow infants to boost their acquisition of words with this powerful tool. In the current study we experimentally tested this hypothesis investigating whether 18-month-old French-learning infants are able to exploit function words (Experiment 1) and phrasal prosody (Experiment 2) to constrain the acquisition of nouns and verbs: a crucial syntactic distinction present in all languages.

2. Experiment 1: Testing the role of function words to constrain the acquisition of word meanings

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

2.1 Method

The studies reported in this paper, including the entire method, analysis and criteria for exclusion of participants were pre-registered on the OSF (Open Science Framework) database before running the experiments (the preregistration can be accessed with the following link: https://osf.io/u2xct/?view_only=f61af293ef524e1cbb550a45341148f7).

The materials, collected data, and data analysis are freely available to readers following the same link.

2.1.1 Participants. Forty-eight French 18-month-olds participated in the study (mean age = 18.1 months, range = 17.7 to 18.6 months; *SD* = 0.2; 27 girls). An additional thirty-two infants came to the lab, but were not included in the final sample for one of the following reasons: because of fussiness not allowing them to finish the experiment (*n* = 13); because they did not meet the habituation criterion within 12 habituation trials (*n* = 6); because of parental interference (*n* = 3); technical problem (*n* = 1); or because they cried during the experiment (*n* = 9). Parents signed an informed consent form. The number of participants tested was chosen on the basis of a power analysis conducted on the effect size found in previous studies using the same design (He & Lidz, 2017). This research was approved by the local ethics committee.

2.1.2 Design. A Habituation-Switch paradigm (e.g., Werker et al., 1998; cf.: Figure 1) was used to habituate infants with two video stimuli showing a penguin doing two different intransitive actions (e.g., spinning, cartwheeling), one in each video. During the presentation of one of the videos (e.g., a penguin spinning), infants heard sentences using a novel word as a noun (e.g., *"Regarde! C'est une bamoule!"* – "Look! It's a bamoule!", where *bamoule* is naming an object, the penguin), and during the presentation of the other video (e.g., a penguin cartwheeling), they heard sentences presenting another novel word as a verb (e.g., "*Regarde! Elle doripe!*" – "Look! She is doriping!", where *doripe* is naming an action, cartwheeling).

This habituation phase gave the infants the opportunity to guess a possible meaning for each of the two novel words: if they exploit the linguistic context provided by the function words as adults would, then they should infer that the novel word employed as a noun (e.g., *C'est une bamoule* – It's a bamoule) refers to the penguin (the only object present in the video), and that the novel word employed as a verb (e.g., *Elle doripe* – She is doriping) refers to the action that the penguin was doing (e.g., spinning or cartwheeling, counterbalanced across participants). When infants reached a pre-defined habituation criterion (three consecutive trials for which the average looking time was less than 65% of the average looking time for the most-attended three consecutive trials), the habituation phase stopped and the test phase started immediately. At test, all infants were presented with a fixed number of 2 trials in which they were exposed to a switch of the audio tracks of the videos (as illustrated in Figure 1), such that half of the infants saw a Noun-Switch condition (e.g, hearing noun sentences with *bamoule*, while watching the penguin cartwheeling) and half saw a Verb-Switch condition (i.e., hearing sentences with the verb *doripe*, while watching the penguin spinning). Because this kind of switch violates the inference constructed about the verb meaning (i.e., "cartwheeling" and "spinning" are different actions), but not about the noun meaning (i.e., it's always the same penguin in both videos), if infants are able to correctly infer the meanings of the novel words during the habituation phase, we expected that during the test phase, infants would be more surprised, thus look more toward the videos, in the Verb-Switch condition than in the Noun-Switch condition.

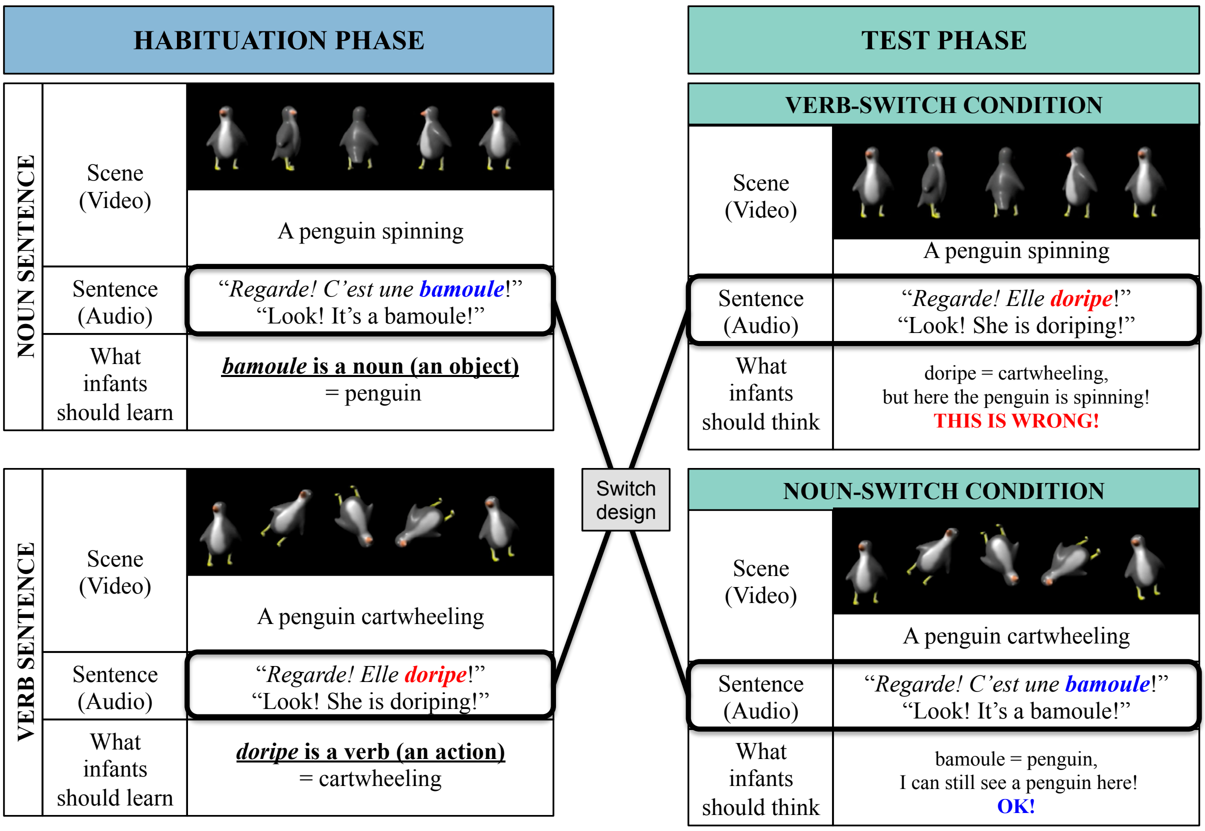


Figure 3: Experimental Design: Habituation-Switch paradigm, used in Experiments 1 and 2. All the infants were habituated with the same two video stimuli showing a penguin doing two different actions (e.g., spinning, cartwheeling, one in each video), while listening to noun sentences in which a novel word was used as a noun, and to verb sentences, in which another novel word was used as a verb. Then, as soon as infants reached a pre-defined habituation criterion, they were exposed to a test phase with a switch between the sentences and the videos, such that half of the children saw a Noun-Switch condition and half saw a Verb-Switch condition. Given that the noun refers to an object – the penguin – while the verb refers to an action, if children correctly use function words (Experiment 1) and phrasal prosody (Experiment 2) to infer the meanings of the novel noun and the novel verb during the habituation phase, at test, they should be more surprised (and look more toward the videos) in the Verb-Switch condition (because the action changed, which is problematic for the verb interpretation) than in the Noun-Switch condition (where they can still see a penguin in the video, and the fact that the action has changed is irrelevant). Note that the syntactic category of the novel words and the associations with the videos were counterbalanced across participants. Thus half of the children tested had *bamoule* as a noun and *doripe* as a verb, and half had the reverse. Half had *spinning* as the verb meaning, and half had *cartwheeling* as the verb meaning.

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

The presentations of these passages were paired with two video stimuli showing a penguin doing two different intransitive actions (e.g., spinning, cartwheeling), one in each video. These videos had exactly the same duration as the audio tracks (37 seconds).
Additionally, a silent video of a butterfly perched on a leaf was used as an attention-getter to recapture infant's attention when they looked away for more than 2 seconds.

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

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

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

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

2.1.5 Data processing and analyses. Data analyses and graphics were performed with R software version 3.2.2 (R Team, 2015). We used the average looking time of the last two trials of the habituation phase, and of the two test trials, and we compared the increase in looking time from habituation to test in the two experimental conditions (Noun-Switch vs Verb-Switch). If infants are able to exploit function words to access the syntactic structure of the sentences in which the novel words occurred, and if they can use this information to infer the syntactic category of the novel words and constrain their possible meaning, we expect a greater increase in looking time from habituation to test in the Verb-Switch condition than in the Noun-Switch condition. To test this, we performed an ANOVA on log-transformed mean looking times as the dependent measure, with participants as the random factor, Condition (Noun-Switch vs Verb-Switch) as a between-participant factor, and Phase (Habituation vs

Test) as a within-participant factor. The expected effect should appear as a significant interaction between Condition and Phase. Note that looking times were log-transformed before running the ANOVA, because the data did not follow a normal distribution, which is a necessary condition to conduct an ANOVA.

2.2. Results

The results of experiment 1 are shown in Figure 2.

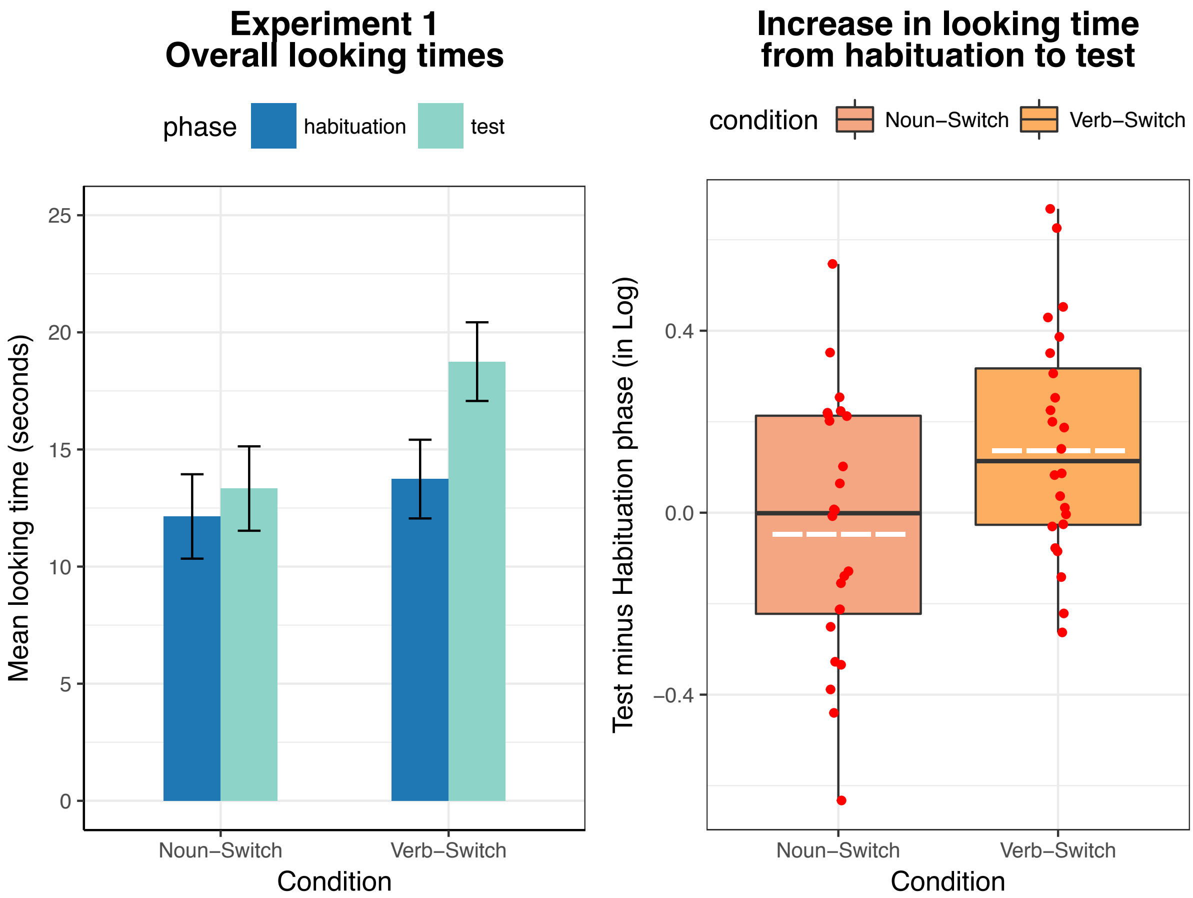


Figure 2: On the left side: Mean looking time in seconds toward the videos during the last two trials of the habituation phase (in blue) and during the two trials of the test phase (in green) for children assigned to the Noun-Switch Condition (N=24) and to the Verb-Switch Condition (N=24). Error bars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

Infants looking times increased more between habituation and test in the Verb-Switch condition than in the Noun-Switch condition: An ANOVA on log-transformed mean looking times revealed a significant interaction between Condition (Noun-Switch vs Verb-Switch) and Phase (Habituation vs Test), $F(1,46)=5.64$, $p<.03$; $d = 0.665$, which confirms that relative to the habituation phase, at test children looked more toward the videos in the Verb-Switch condition than in the Noun-Switch condition. This is consistent with the interpretation that during the habituation phase, toddlers inferred that the novel verb referred to the action (e.g., cartwheeling), while the novel noun referred to the object (the penguin). Thus, at test, infants in the Verb-Switch condition were surprised when watching the penguin performing the other action (e.g. spinning), while listening to sentences containing the verb that they had associated to the cartwheeling action during the habituation phase. Since the penguin was not cartwheeling but spinning, there was a discrepancy between the original meaning they had inferred for this word, and the current situation. Note that this increase in looking time can be explained either by the fact that infants thought that the target word was not used correctly at test (i.e., the speaker was making a mistake and using the word incorrectly), or by the fact that infants realized that they had to broaden the meaning that they had initially inferred for this word (e.g. instead of meaning cartwheeling specifically, *doripe* might refer to a broader class of movements, perhaps involving rotation – which is present in both actions). In contrast, infants tested in the Noun-Switch condition, who listened to noun sentences with *bamoule* while watching the penguin cartwheeling, did not show surprise during test: indeed the meaning they had inferred during habituation (that *bamoule* means penguin), was perfectly consistent with the test video they were watching, since there was still a penguin on the screen.

Given that both groups were exposed to exactly the same videos and sentences in the habituation phase, the only way to explain the asymmetry observed in the test phase, is that infants paid attention to the syntactic context instantiated by function words, to correctly assign a syntactic category to the novel words and constrain their meanings: switching the audio tracks of the videos led to a violation of the inference constructed about the verb meaning (i.e., "cartwheeling" and "spinning" are different actions), but not the inference about the noun meaning (i.e., it's still the same penguin in both videos).

2.3. Discussion

Experiment 1 shows that 18-month-olds are able to use function words to compute the syntactic category of novel words and therefore to constrain their probable meanings (e.g., mapping nouns to objects and verbs to actions, respectively). However, in real-life not all content words are preceded by function words (e.g., in: "The baby flies...", fly can be either a noun or a verb). In such cases, an analysis in terms of syntactic constituents can be very informative (i.e., fly is a noun in: [The baby flies]$_{NP}$..., but a verb in: [The baby]$_{NP}$[flies..“his kite”]$_{VP}$). Since the prosodic structure of an utterance tends to coincide with its syntactic structure, listeners could exploit prosodic boundaries together with function words to constrain syntactic analysis (Morgan & Demuth, 1996; Christophe et al., 2008), a hypothesis that will be tested in experiment 2.

3. Experiment 2: Testing the role of phrasal prosody to constrain the acquisition of word meanings

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

3.1. Method

3.1.1. Participants. Forty-eight French 18-month-olds participated in the study (mean age = 18.2 months, range = 17.8 to 18.8 months; *SD* = 0.2; 24 girls). An additional twenty-three infants came to the lab, but were not included in the sample: because of fussiness not allowing them to finish the experiment (*n* = 11); because they did not meet the habituation criterion within 12 habituation trials (*n* = 4); because of parental interference (*n* = 1); technical problem (*n* = 2); or because they cried during the experiment (*n* = 5). Parents signed an informed consent form. This research was approved by the local ethics committee.

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

Since in this experiment both sentences are composed of the same words in the same order (*regarde-la-petite-bamoule/doripe*), an analysis in terms of which words precede *bamoule* or *doripe* is not sufficient to determine its syntactic category (since they are the same in both conditions): rather, the syntactic difference between these two sentences is reflected in their different prosodic structures. When *doripe* is a verb, there is a prosodic boundary preceding it (i.e., the boundary between the noun phrase and the verb phrase) while when *bamoule* is a noun, it is embedded in a single prosodic unit together with the other words of the sentence, corresponding to the verb 'look' and the following noun phrase. If infants are able to use the information provided by the prosodic structure of a sentence to access its syntactic structure as adults and preschoolers do (de Carvalho, Dautriche, & Christophe, 2016; de Carvalho, Lidz, Tieu, Bleam, & Christophe, 2016; Michelas & D'Imperio, 2015; Millotte, René, Wales, & Christophe, 2008; Millotte, Wales, & Christophe, 2007; Snedeker & Yuan, 2008), and if they can use this information to constrain the meaning of novel words, during the habituation phase they should infer that the novel word used as a noun refers to the penguin (the only object present in the video), and that the novel word used as a verb refers to the action that the penguin is doing (e.g., spinning or cartwheeling, counterbalanced across participants). Thus, as in Experiment 1, during the test phase, we expect infants to be more surprised (looking more toward the video) in the Verb-Switch condition than in the Noun-Switch condition.

3.1.3. Material. The same two novel words in French (*bamoule; doripe*) were used as target words to create minimal pairs of sentences that differed only in their prosodic structures. Thus, for each novel word two sentences were created, one presenting the target word in a noun position within the prosodic-syntactic structure (Noun sentence: [*Regarde la petite bamoule*]! [*Tu vois la petite bamoule*]? – [Look at the little bamoule]! [Do you see the little bamoule]?, and another one presenting the novel word in a verb position (Verb sentence: "[*Regarde*]! [*la petite*] [*bamoule*]! [*Tu vois*]? [*la petite*] [*bamoule*]! - [Look]! [The little one] [is bamouling]! [Do you see]? [The little one] [is bamouling]! (square brackets indicate prosodic boundaries). An example of each kind of sentence is depicted in Figure 4.

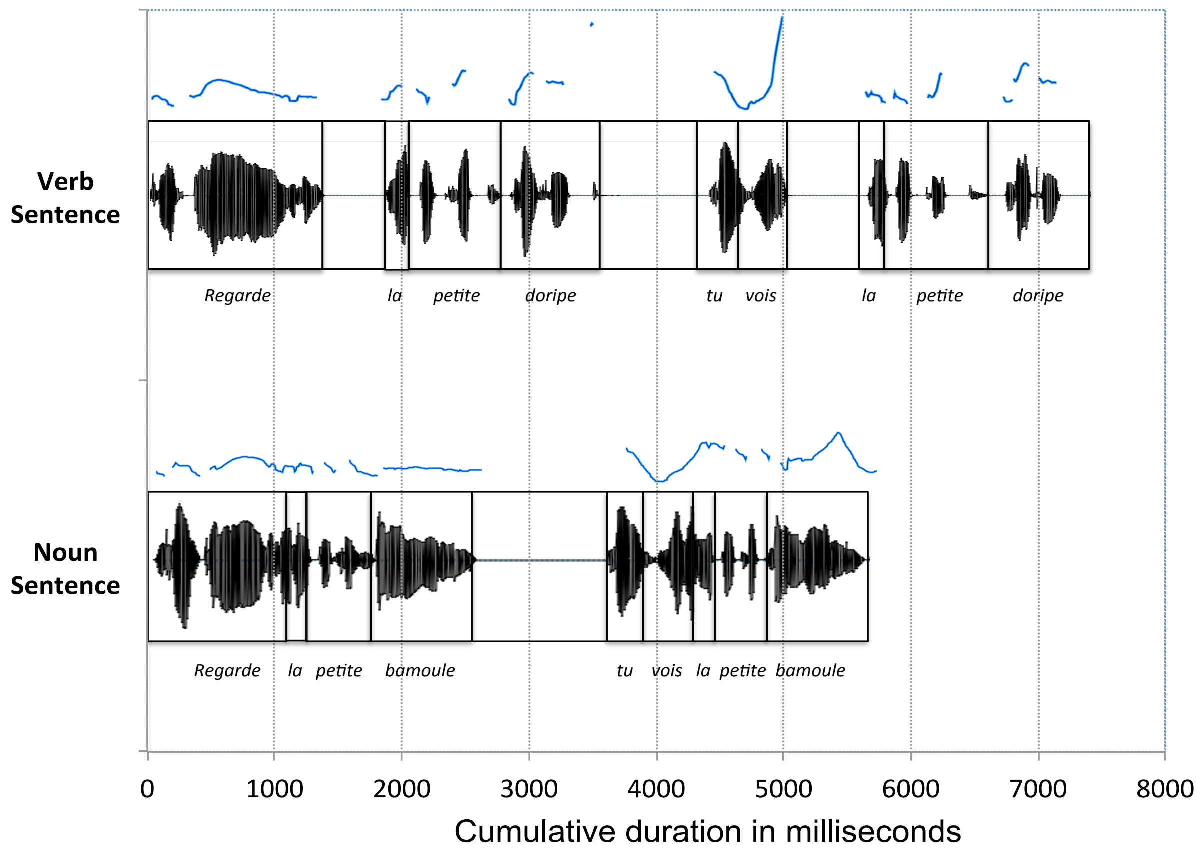


Figure 4: An example of the minimal pair of sentences created for Experiment 2, with the Verb sentence on top and the Noun sentence at the bottom. The x-axis represents the time-course, for each sentence. From bottom to top, we present the transcription, the waveform, and the pitch contour (blue curve) of each sentence. Both sentences were composed of exactly the same words (la-petite-bamoule/doripe) but differed only in their prosodic structure, which reflected their different syntactic structures. In Noun sentences all the critical words were grouped together into a single prosodic unit [*Regarde la petite bamoule*]! - [Look at the little bamoule]!; In contrast, in Verb sentences these words were spread into three different prosodic units [*Regarde*]![*la petite*] [*bamoule*]! - [Look]! [The little one] [is bamouling]!.

Sentences uttered in the verb condition had a phonological phrase boundary before the target word (i.e., corresponding to the boundary between the noun and the verb phrase); in contrast, in sentences uttered in the noun condition all the words were grouped together into a single prosodic unit; these prosodic structures are consistent with theoretical descriptions of the relationship between prosodic and syntactic boundaries (Jun, 2005; Nespor & Vogel, 1986).

In order to create the audio tracks, each sentence was repeated twelve times, resulting in a 50-second-long passage for each target word in each condition; each repetition was introduced by an audio prompt (e.g., “Oh”; “Wow”; “Hey”). As in experiment 1, the assignment of a syntactic category to the two novel words was counterbalanced across participants. All the stimuli were recorded by the same speaker as experiment 1, in child-directed register. The audio tracks were paired with the same video stimuli used in experiment 1 (but with a duration of 50 seconds). The same silent video of a butterfly perched on a leaf was used as attention-getter.

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

3.1.5. Data processing and analyses. Same as Experiment 1.

3.2. Results

The results of experiment 2 are shown in Figure 3.

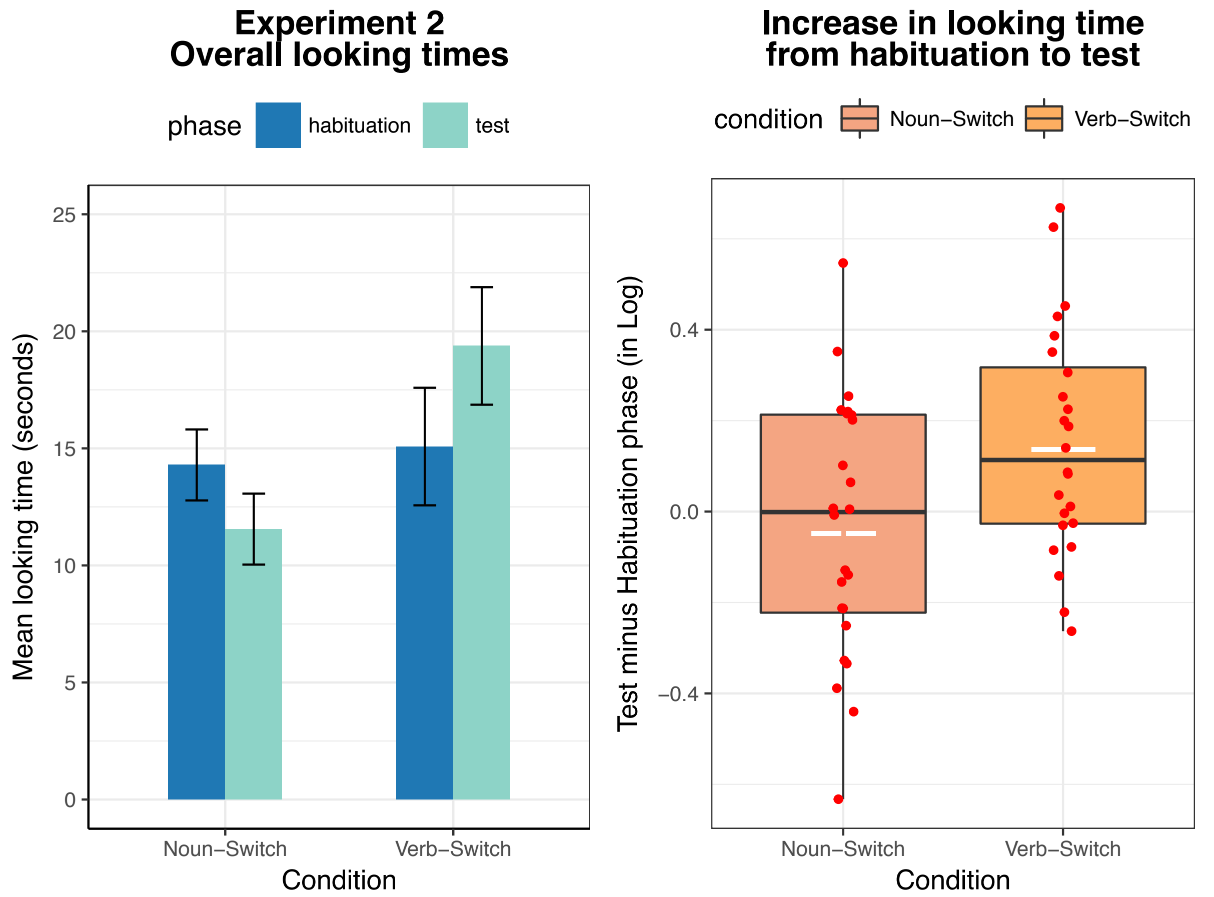


Figure 3: On the left side: Mean looking time in seconds toward the videos during the last two trials of the habituation phase (in blue) and during the two trials of the test phase (in green) for children assigned to the Noun-Switch Condition (N=24) and to the Verb-Switch Condition (N=24). Error bars represent the standard error of the mean of the difference between Test minus Habituation. On the right side: Boxplot of the increase in log-transformed mean looking time from habituation to test in each group. Red dots represent the average for each participant in each group. White dashed lines represent the means of the distributions.

Infants looking time increased more between habituation and test in the Verb-Switch condition than in the Noun-Switch condition: An ANOVA on log-transformed mean looking times revealed a significant interaction between Condition and Phase: $F(1,46)=5.09$, $p < .03$; $d = 0.632$, showing that infants looked longer (were more surprised) when tested in the Verb-switch condition than in the Noun-switch condition. This behavior, just like in Experiment 1, suggests that the switch of the actions led to a violation of the inference constructed about the verb meaning, but not about the noun meaning, consequently infants were more surprised when listening to verb sentences than to noun sentences during the test phase.

This result thus shows that at 18 months, infants are already able to use phrasal prosody as a cue to interpret a novel word as either a noun or a verb depending on its position within the prosodic structure of sentences. When listening to minimal pairs of sentences such as *Regarde-la-petite-bamoule/doripe*, which can be produced either as a single prosodic unit, as in [*Regarde la petite bamoule*]! – [*Look at the little bamoule*]!, where *bamoule* is used as a noun, or as three prosodic units such as [*Regarde*]! [*la petite*] [*doripe*]! – [*Look*]! [*the little one*] [*is doriping*]!, where *doripe* is used as a verb, 18-month-olds correctly interpreted the target word as either a noun or a verb, depending on its position within the prosodic structure of the sentence, and they used this information to constrain the meaning of this novel word. Thus, even in a situation where the information provided by function words alone was not sufficient to compute the syntactic category of the novel words (i.e., the noun and verb sentences had exactly the same words) infants were able to exploit prosodic information to recover the syntactic structure of sentences and, in combination with the other words (e.g., *regarde-la-petite…*), infer the syntactic category of novel words and therefore constrain their meanings.

3.3. Discussion

These results show for the first time that infants use prosody as a cue to constrain the acquisition of noun and verb meanings. They were able to use phrasal prosody to recover the syntactic structure of sentences, interpreting a novel word as either a noun (referring to an object) or a verb (referring to an action) depending on its position within the prosodic-syntactic structure.

4. General Discussion

Across two experiments, we demonstrated that during the first steps of language acquisition, when infants don't know many words yet, they can rely on function words and phrasal prosody to access the syntactic structure of sentences and therefore to constrain the acquisition of novel word meanings. In Experiment 1, French infants were able to exploit the functional elements in a sentence to assign a syntactic category to a novel word and to constrain its meaning. In Experiment 2, where the information provided by function words alone was not sufficient to compute the syntactic category of the novel words, this study showed for the first time that infants were able to simultaneously exploit prosodic information to recover the syntactic structure of sentences, and use this syntactic structure to compute the syntactic category of novel words and therefore constrain their meanings.

Given the chicken-and-egg problem stated in our introduction, here we observe that by providing information about the syntactic structure of sentences, phrasal prosody allows infants to access syntactic information which in turn, allows them to figure out which part of the world is being talked about (e.g., associating nouns to objects, and verbs to actions) and guide their discovery of word meanings.

Remarkably, even in the absence of a very developed lexicon and prior to being able to produce multi-word utterances themselves, 18-month-olds already have the means to go from the sounds of language to the syntactic structure of sentences, and to use this information to infer the probable meaning of a novel word. Despite their reduced vocabulary, 18-month-olds in the present experiment easily exploited the acoustic/prosodic information in speech to parse spoken sentences into groups of words and to identify possible syntactic constituents. They exploited this constituent structure, together with function words, to determine the syntactic nature of these prosodic-syntactic constituents, infer the syntactic category of the novel words, and therefore constrain their probable meaning.

This powerful mechanism would be extremely useful for infants to build their way into language acquisition and to provide them with a tool to construct a first-pass syntactic structure of spoken sentences even before they know the meanings of many words. In this sense, the answer we provide to the chicken-and-egg problem of learning word meanings through syntax, and learning syntax through word meanings, is that phrasal prosody and function words can work as anchors to help infants access syntactic information. Crucially, this ability to jointly exploit phrasal prosody and function words to access syntactic structure seems to be developed even before infants know many words.

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

Author Contributions

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

Acknowledgements

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

Declaration of Conflicting Interests

The authors declare that they had no conflicts of interest.

Funding

This research was supported by a PhD fellowship from the École normale supérieure to Alex de Carvalho. 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), and the ANR 'Apprentissages' (ANR-13-APPR-0012).

References

Bernal, S., Lidz, J., Millotte, S., & Christophe, A. (2007). Syntax Constrains the Acquisition of Verb Meaning. *Language Learning and Development*, *3*(4), 325–341. http://doi.org/10.1080/15475440701542609

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

Christodoulopoulos, C., Roth, D., & Fisher, C. (2016). An incremental model of syntactic bootstrapping. In *Proceedings of the Seventh Workshop on Cognitive Aspects of Computational Language Learning*.

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

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

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

de Carvalho, A., Dautriche, I., & Christophe, A. (2016). Preschoolers use phrasal prosody online to constrain syntactic analysis. *Developmental Science*, *19*(2), 235–250. http://doi.org/10.1111/desc.12300

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

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

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

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. http://doi.org/10.1016/j.cogpsych.2007.12.001

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

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

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

Haryu, E., & Kajikawa, S. (2016). Use of bound morphemes (noun particles) in word segmentation by Japanese-learning infants. *Journal of Memory and Language*, *88*, 18–27. http://doi.org/10.1016/j.jml.2015.11.007

He, A. X., & Lidz, J. (2017). Verb learning in 14- and 18-month-old English-learning infants. *Language Learning and Development*, *5441*(March).

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

Männel, C., & Friederici, A. D. (2009). Pauses and intonational phrasing: ERP studies in 5-month-old German infants and adults. *Journal of Cognitive Neuroscience*, *21*(10), 1988–2006. http://doi.org/10.1162/jocn.2009.21221

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

Michelas, A., & D'Imperio, M. (2015). Prosodic boundary strength guides syntactic parsing of French utterances. *Laboratory Phonology*, *6*(1), 119–146. http://doi.org/10.1515/lp-2015-0003

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

Millotte, S., Wales, R., & Christophe, A. (2007). Phrasal prosody disambiguates syntax. *Language and Cognitive Processes*, *22*(6), 898–909. http://doi.org/10.1080/01690960701205286

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

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

Shi, R., & Melançon, A. (2010). Syntactic Categorization in French-Learning Infants. *Infancy*, *15*(5), 517–533. http://doi.org/10.1111/j.1532-7078.2009.00022.x

Shi, R., Morgan, J., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: a cross-linguistic perspective. *Journal of Child Language*, *25*(1), 169–201.

Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, *72*(2), B11–B21. http://doi.org/10.1016/S0010-0277(99)00047-5

Shukla, M., White, K. S., & Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences*, *108*(15), 6038–6043. http://doi.org/10.1073/pnas.1017617108

Snedeker, J., & Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language*, *58*(2), 574–608. http://doi.org/10.1016/j.jml.2007.08.001

Soderstrom, M., Seidl, A., Nelson, D. G., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language*, *49*(2), 249–267. http://doi.org/10.1016/S0749-596X(03)00024-X

Team, R. C. (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Tomasello, M. (2000a). Do young children have adult syntactic competence? *Cognition*, *74*(3), 209–253. http://doi.org/10.1016/S0010-0277(99)00069-4

Tomasello, M. (2000b). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences*, *4*(4), 156–163. http://doi.org/10.1016/S1364-6613(00)01462-5

Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology*, *59*(1), 67–95. http://doi.org/10.1016/j.cogpsych.2009.02.001

Werker, J. F., Cohen, L., Lloyd, V., & Casasola, M. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology*, *34*(6), 1289–1309.

Chapter 8

“*Look! It is not a bamoule!*” 18-month-olds understand negative sentences

Developmental Science

Developmental Science

"Look! It is not a bamoule!" 18-month-olds understand negative sentences

Journal:	*Developmental Science*
Manuscript ID	Draft
Manuscript Type:	Paper (under 8000 words)
Keywords:	early acquisition of negation, negation understanding, language acquisition, lexical development, infant development
Subject Area:	Infant cognitive development

Running-Head: 18-MONTH-OLDS UNDERSTAND NEGATIVE SENTENCES

RESEARCH HIGHLIGHTS:

- 18-month-olds are able to understand negative sentences
- Although previous studies suggested that children younger than two process negative sentences as affirmatives, here we report opposite behaviors for the two types of sentences.
- This is the first study to report that children under two years old can compute the syntactic function of negation.
- The ability to understand negative sentences early during development should support language acquisition and help infants to constrain the acquisition of word meanings.

Abstract

This study investigated French-learning infants' understanding of negative sentences at 18 months. Although infants start producing the word "no" in their own speech from about 1 year of age, several studies have failed to find any evidence for the understanding of negative sentences before 2 years old. Using a word-learning task, we observed that 18-month-olds already have some understanding of negative sentences. After having learnt that *bamoule* means "penguin" and *pirdaling* means "cartwheeling", infants showed surprise when listening to negative sentences rendered false by their visual context ("*Look! It is not a bamoule!*", while watching a video showing a penguin cartwheeling); in contrast, they were not surprised by negative sentences rendered true by their context ("*Look! She is not pirdaling!*" while watching a penguin spinning). This provides the first evidence for the understanding of negative sentences during the second year of life.

Keywords: early acquisition of negation, negation understanding, language acquisition, lexical development, infant development.

18-MONTH-OLDS UNDERSTAND NEGATIVE SENTENCES

"Look! It is not a bamoule!" 18-month-olds understand negative sentences

1. Introduction

Negation, a universal linguistic concept, presents an interesting and challenging question in language acquisition research. Although infants begin producing the word "no" in their own speech at about 1 year of age, research has failed to find any evidence for the understanding of negative sentences before 24 months, and has even observed that toddlers process negative sentences as if they were affirmatives (Feiman, Mody, Carey, & Snedeker, 2014, with 20-month-olds; Nordmeyer & Frank, 2014, with 2-to-3-year-olds). This inability to correctly interpret negative sentences, at an age when children are still learning their native language, could disrupt their acquisition of word meanings.

It has been demonstrated that well before the age of two, infants start to learn the meaning of words in their language by exploiting the syntactic structures in which the words occur (Bernal, Lidz, Millotte, & Christophe, 2007; de Carvalho, He, Lidz, & Christophe, 2015; Fisher, Hall, Rakowitz, & Gleitman, 1994; Gleitman, 1990; He & Lidz, 2017; Landau & Gleitman, 1985; Waxman, Lidz, Braun, & Lavin, 2009; Yuan, Fisher, & Snedeker, 2012). For instance, from 18 months, infants correctly infer that a novel word such as *doke* refers to a given object when listening to sentences such as *This is a doke*, and to a given action when listening to sentences such as *She is doking* (Bernal, Lidz, Millotte, & Christophe, 2007; de Carvalho, He, Lidz, & Christophe, 2015; He & Lidz, 2017; Waxman, Lidz, Braun, & Lavin, 2009). However, we still do not know whether young children would be able to interpret a syntactic frame such as *This is <u>not</u> a doke*. Negation is an aspect of syntactic structure that has an important impact on the interpretation of sentences, since it can completely change the meaning of that sentence. Thus, if children are truly relying on the syntactic context of words to identify their potential referents, it is important for them to be able to distinguish between affirmative and negative sentences. To investigate this question, the current study tests infants' understanding of negative sentences in a word-learning situation.

Several studies have examined children's production of negative sentences (Bloom, 1970; Cameron-Faulkner, Lieven, & Theakston, 2007; Choi, 1988; Drozd, 1995; Guidetti, 2005; Hummer, Wimmer, & Antes, 1993; McNeill & McNeill, 1967; Pea, 1980; Tam & Stokes, 2001; Vaidyanathan, 1991). These studies found that as early as 13 months, infants start producing the word "no" (or "non" in French), and that before the age of two, infants can produce negative sentences to express concepts such as refusal (e.g., "Je veux pas de lait" - "Don't want milk!") and nonexistence (e.g., "Il y a plus de lumière" - "No more light!", just after the light is turned off; see for instance Choi (1988) for a cross-linguistic longitudinal study on the semantic development of negation in English-, French- and Korean-speaking children).

Surprisingly however, the few existing experimental studies investigating the comprehension of negative sentences in young children failed to evidence any understanding of negative sentences before the age of two (Austin, Theakston, Lieven, & Tomasello, 2014; Feiman et al., 2014; Kim, 1985; Nordmeyer & Frank, 2014; Reuter, Feiman, & Snedeker, 2014). In Austin et al. (2014), for instance, children from three age groups (21-, 24- and 28-month-olds) had to find an object that was hidden either in a bucket or a house; to help them, a second experimenter asked questions such as *Is it in this house?* or *Is it in this bucket?* to the person who had hidden the object. This person replied either with an affirmative sentence, such as *It is in this house* or with a negative sentence, *It's not in this bucket*. While the 28- and 24-month-olds successfully searched for the hidden object in the correct place after listening to a negative statement, the youngest group of 21-month-olds performed at chance. This failure of the youngest group can be interpreted in two ways: either they do not yet understand negative sentences, or they have trouble inhibiting their first move, which is to turn their attention toward the named container (e.g. *the bucket*, which does not contain the object in the case of negative sentences), in order to successfully search for the hidden object in the other "non-named" container (e.g., *the house*). If the first hypothesis is correct, children's early production of negation would reflect the fact that they have a very basic meaning for the word "no", such as "rejection", but that they cannot yet compute the syntactic function of negation (i.e., truth-function negation). In contrast, if the second interpretation about inhibitory abilities is correct, this means that the task used is too complex for young infants, who do not yet possess the inhibitory abilities required to succeed.

Supporting the hypothesis about inhibitory skills, Nordmeyer and Frank (2014) showed that even 3- to 5-year-olds can experience difficulties correctly processing negative sentences, in contexts that require inhibitory abilities. In this study, 2-to-5-year-olds were asked to *Look at the boy who has no apples* while watching two pictures on a TV-screen, one showing *a boy with two apples* and the other *a boy with two boxes*. They observed that 3- to 5-year-olds were able to correctly look toward the image of the boy without apples (although they did not perform as well as adults). In contrast, 2-to-3-year-olds interpreted negative sentences as if they were affirmative sentences, and looked more at the picture of the boy with apples. The authors argued that this task could confuse young children because it requires them to refrain from looking at the boy who has apples, while entertaining a proposition with the concept *apple*. Thus, this task requires them not only to understand the negative sentences, but also to avoid looking at the apples, and instead look at the boy who represents the negation of the proposition (the one who has boxes but no apples). Additionally, Nordmeyer and Frank (2014) argued that fine-grained pragmatic computations might also interfere with the processing of negative sentences in this experiment: in a naturalistic context, if a speaker wanted to ask someone to look at the boy with boxes, she would probably say *Look at the boy who has boxes* rather than *Look at the boy who has no apples*. Thus, the negative utterance is pragmatically infelicitous in this context: Speakers are expected to produce affirmative rather than negative sentences – unless there is a special reason for producing a negative sentence, for instance to mark a contrast with a proposition that has been uttered previously, or to negate an expectation entertained by the listener (Nordmeyer & Frank, 2014, 2015).

All these points about inhibitory abilities required to interpret negative sentences, and the discourse context in which they were uttered, might explain why previous studies have found no understanding of negative sentences before the age of two. It remains unclear however whether infants really lack any understanding of negation before two years old, or whether the experiments used so far were not sensitive enough to reveal their ability. Given that many studies have attested that infants can express negative concepts in their own speech well before two years old, it might be the case that what requires time to develop is not the comprehension of negation *per se*, but rather the sophisticated cognitive abilities that were required to process negative sentences in previous studies, such as pragmatic and inhibitory skills. As a consequence, it is possible that a simpler task, reducing pragmatic and inhibitory conflicts, may uncover an ability to understand negation between 12 and 24 months of age. If at the age of 18 months, as suggested by the studies based on infants' production of negation, infants have already acquired some of the concepts expressed by negation, and if they already know how to interpret syntactic contexts containing negation, they might be able to understand at least simple negative sentences, such as *It is not a doke*, in a context that pragmatically supports the use of a negative sentence. In the current study we tested this hypothesis using a word-learning task.

2. Experiment

To test whether 18-month-olds understand negative sentences, we exploited a paradigm recently used in English and French, showing that 18-month-olds, upon hearing: "Look, it is a *bamoule*!", infer that *bamoule* refers to an object, while from "Look! She is *pirdaling*!", they infer that *pirdaling* refers to an action (de Carvalho et al, 2015; He & Lidz, 2017). In these studies, 18-month-olds were habituated with two video stimuli showing a penguin doing two different intransitive actions (e.g., spinning, cartwheeling), one in each video. During the presentation of one of the videos (e.g., a penguin spinning), participants heard sentences using the novel word as a noun (e.g., *Oh Look! It is a bamoule!*, where bamoule was referring to the object in the video: the penguin), and during the presentation of the other video (e.g., a penguin cartwheeling), participants heard sentences using a novel word as a verb (e.g., *Oh Look! She is pirdaling!*; where *pirdaling* was naming the action that the penguin was doing: cartwheeling). Immediately after this habituation phase, infants were exposed to a test in which the associations between the sentences and the videos were switched: the noun sentences were presented with the video previously associated with the verb (e.g., cartwheeling; the Noun-Switch condition) and the verb sentences were presented with the video previously associated with the noun (e.g., spinning; the Verb-Switch condition). The results showed that when participants heard the noun sentences "*It is a bamoule*" while watching the penguin cartwheeling, they were not particularly surprised by the switch, since there was still a penguin in the video (although it was "cartwheeling" instead of "spinning"). However, when participants heard the verb sentences "*She is pirdaling*" while watching the penguin "spinning" instead of "cartwheeling", they looked longer toward the videos (i.e., were more surprised), because this kind of switch violated the inference they had constructed about the verb meaning during the habituation phase (i.e., "cartwheeling" and "spinning" are different actions).

In the current study, we implemented a slight modification in this experimental design (as summarized in Figure 1) and used negative sentences during the test phase: one negating the verb meaning learned during the habituation phase, and the other one negating the noun meaning. Thus, after the same habituation phase where children learned that *bamoule* refers to "penguin" (when they saw the penguin spinning) and *pirdaling* to 'cartwheeling' (when they saw the penguin cartwheeling), in the test phase, if they hear negative sentences such as "*Oh look! She is not pirdaling!*" while watching the penguin spinning (the Negative Verb-Switch condition) they should not be surprised about the switch of the actions because indeed the penguin is not cartwheeling anymore but rather, spinning. In contrast, if they hear sentences such as "*Oh look! It is not a bamoule!*", while watching the penguin cartwheeling (the Negative Noun-Switch condition) they should be surprised because there is still a penguin in the video. In other words, if infants are able to understand negative sentences, we expect that in the test phase, where only the actions changed, but the penguin is still the same, they should look more toward the video (show more surprise) in the Negative Noun-Switch than in the Negative Verb-Switch condition.

Note that this paradigm neatly avoids taxing children's inhibitory abilities, since here the test phase exposes children to one video at a time, without any competition from an alternative interpretation. Moreover, since the test phase is the result of a switch between the audio sentences and the videos presented during the habituation phase, the context supports the use of negative sentences, to make a contrast between what had been observed before (during the habituation phase) and what infants can currently observe (during the test phase).

Figure 1 summarizes the experimental design (the study was conducted with French infants, and all the sentences used were in French).

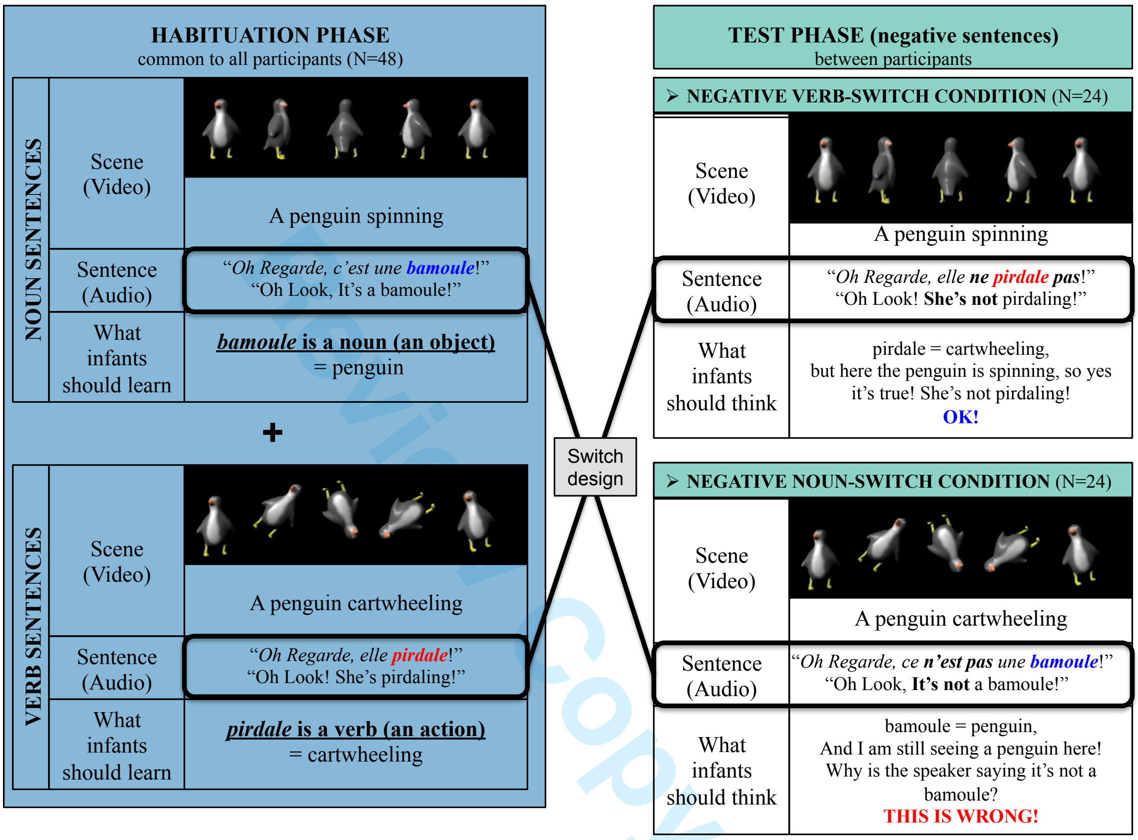


Figure 1: Experimental design (Habituation – Switch design, e.g., de Carvalho et al., 2015; Haryu & Kajikawa, 2016; He & Lidz, 2017; Werker, Cohen, Lloyd, & Casasola, 1998). All infants were habituated with the same two video stimuli. These two videos showed either a penguin doing a spinning action or a penguin doing a cartwheeling action. While watching one of the videos (e.g., spinning) infants listened to sentences in the noun condition, in which a novel word was used as a noun, and while watching the other video (e.g., cartwheeling), they listened to sentences in the verb condition, in which another novel word was used as a verb. At test, all infants heard negative sentences introducing a switch between the videos and the sentences they heard before, such that half of the infants (*N*=24) heard negative sentences in the Negative Noun-Switch-condition and half heard negative sentences in the Negative Verb-Switch-condition. Given that during the habituation phase, the noun always referred to the penguin, while the verb referred to an action (e.g., cartwheeling), if infants correctly understand negative sentences, they should look more toward the video (show more surprise), in the Noun-Switch-condition (because the speaker says that the penguin is not a *bamoule*, when in fact it is) than in the Verb-Switch-condition (where they hear that the penguin is not *pirdaling* when indeed it is now spinning rather than cartwheeling).

If at test, infants look more toward the video when listening to the negative sentences in the Negative Noun-Switch condition than to the negative sentences in the Negative Verb-Switch condition, this study will show that 18-month-olds are able to understand negative sentences and to evaluate whether they are used appropriately or not given their context. In contrast, if infants show the same behavior in both conditions, so that both groups look longer to the video in the test phase than in the habituation phase, this will suggest that they were surprised by the contrast between the negative sentences used at test versus the affirmative sentences presented during habituation phase, without computing the match between the semantic content of the negative sentences and their visual context. Finally, if as in the study by Nordmeyer and Frank (2014), infants are unable to interpret negative sentences and instead process negative sentences as if they were affirmative sentences, we should observe exactly the same results obtained in de Carvalho et al., (2015) and He and Lidz (2017), who used affirmative sentences in the Noun-Switch and Verb-Switch conditions. That is, infants should look longer in the negative Verb-Switch condition than in the negative Noun-Switch condition, at test. If we observe this pattern of results, we will have to conclude that 18-month-olds interpret negative sentences as having the same meaning as affirmative sentences (as did 20-month-olds in Feiman et al., 2014, and 2-to-3-year-olds in Nordmeyer & Frank, 2014).

3. Method

The study reported in this paper, including the entire method, analysis and criteria for exclusion of participants, was pre-registered on the OSF (Open Science Framework) database before running the experiment (the preregistration can be accessed with the following link: https://osf.io/hgjs6/?view_only=7870bdfd90be4e6c8c3a1f847dedaebd). The stimuli used, the collected data, and the data analysis are freely available to readers.

Participants. Forty-eight French 18-month-olds participated in the study (mean age = 17.9 months, range = 17.5 to 18.6 months; *SD* = 0.25; 18 girls). An additional twenty-six infants came to the lab, but were not included in the final sample for one of the following reasons:

because of fussiness not allowing them to finish the experiment (*n* = 8); because they did not meet the habituation criterion within 12 habituation trials (*n* = 5); because of parental interference (*n* = 5); technical problem (*n* = 1); or because they cried during the experiment (*n* = 7). The number of participants tested was chosen on the basis of a power analysis conducted on the effect size found in de Carvalho, He, Lidz and Christophe (2015) and He and Lidz (2017).

Material. Two novel words in French (*bamoule; pirdale*) were used as target words. For each novel word, four sentences were created: two affirmative sentences and two negative sentences. In the two affirmative sentences, one sentence presented the target word as a noun (e.g., "*Oh regarde! C'est une bamoule! Tu la vois la bamoule?*" – "Oh look! It is a bamoule! Do you see the bamoule?") and another one presented it as a verb (e.g., "*Oh regarde! Elle bamoule! Tu la vois qui bamoule?*" – "Oh look! She is bamouling! Do you see her bamouling?"). The two negative sentences were negative versions of the affirmative sentences, such that one negated the target word used in a noun position (e.g., "*Oh regarde! Ce n'est pas une bamoule! Tu vois? Ce n'est pas une bamoule!*" – "Oh look! It is not a bamoule! Do you see? It is not a bamoule!") and the other one negated the target word used in a verb position (e.g., "*Oh regarde! Elle ne bamoule pas! Tu vois? Elle ne bamoule pas!*" – "Oh look! She is not bamouling! Do you see? She is not bamouling!"). In order to create the audio tracks of the videos, each sentence was repeated twelve times, resulting in a 50-second-long passage for each target word in each condition; each repetition was briefly introduced by an audio prompt (e.g., "Oh"; "Wow"; "Hey") to relieve monotony and keep infants listening to the sentences while watching the videos. All the passages (noun and verb sentences, affirmative and negative sentences), had exactly the same audio prompts. The assignment of target words to syntactic categories (noun vs. verb) and the associations with the videos were counterbalanced across participants, such that half of the participants had the target word *pirdale* as a noun and *bamoule* as a verb, and half had the reverse. Half had "spinning" as the verb meaning, and half had "cartwheeling" as the verb meaning. All the stimuli were recorded by a female native speaker of French (the last author), in child-directed register.

Apparatus and procedure. The infants were seated on their parent's lap, facing a 27-inch monitor and a loudspeaker in a sound-attenuated booth. Parents wore headphones and listened to masking music during the experiment. On top of the monitor, a video camera connected with an LCD monitor placed outside the cabin allowed the experimenter to observe the infants behavior and code their looking behavior online. The presentation of the stimuli and the online coding were controlled by the Habit software, version 1.0 (Cohen, Atkinson, & Chaput, 2004). Each trial started with the presentation of an attention-getter (a silent video of a butterfly perched on a leaf) to attract infant's attention. When the child oriented toward the monitor, the experimenter pressed a computer key to start the video. The experimenter pressed a computer key when the toddler looked toward the screen, and released it when the toddler looked away. If the toddler reoriented toward the screen within 2 seconds, the video continued to play, but the time spent looking away was subtracted from their looking time. Each trial lasted until the child looked away for more than two seconds, or until the maximum length of the trial was reached (i.e., 50 seconds). The experimenter was blind to the stimuli.

During the habituation phase, infants were presented with the two video stimuli, showing a penguin doing two different intransitive actions (spinning and cartwheeling, one in each video). During the presentation of one of the videos (e.g., spinning), they heard affirmative sentences presenting a novel word as a noun, and during the presentation of the other video (e.g., cartwheeling), they heard affirmative sentences presenting another novel word as a verb. These videos were presented repeatedly one after the other, until the child reached a pre-defined habituation criterion (the average looking time during any block of 3 consecutive trials had to drop to less than 65% of the average looking time for the 3-trial block that had the longest total looking time). Habituation lasted at least four trials and no more than twelve trials. This habituation phase (as established in de Carvalho et al., 2015 and He & Lidz, 2017) gave infants the opportunity to guess a possible meaning for the two novel words: inferring that the novel word employed as a noun referred to the penguin (the only object present in the video), and that the novel word employed as a verb referred to the action that the penguin was doing (spinning or cartwheeling, counterbalanced across participants).

At test, infants were divided into two different groups and were presented with a fixed number of 2 trials, in which negative sentences were presented, and the noun and verb sentences switched videos (see Figure 1). Participants assigned to the negative sentences in the Noun-Switch condition went through two Negative Noun-Switch trials in which they listened to negative sentences featuring the noun they had learnt during the habituation phase, as they watched the video previously associated with the verb (e.g., a penguin cartwheeling). Participants assigned to negative sentences in the Verb-Switch condition went through two Negative Verb-Switch trials in which they listened to negative sentences featuring the verb they had learned during habituation, as they watched the video previously associated with the noun (e.g., a penguin spinning). Half of the participants were assigned to the Negative Noun-Switch Condition and half to the Negative Verb-Switch condition.

Data processing and analysis. The dependent variable for analysis was the time that infants spent looking at the visual stimulus. To test for increased interest, we compared the average looking time of the last two trials of the habituation phase with that of the two test trials, in the two experimental conditions (Negative Noun-Switch vs. Negative Verb-Switch). If infants are able to understand the meaning of negative sentences, we expect a greater increase in looking time from habituation to test in the Negative Noun-Switch condition than in the Negative Verb-Switch condition. To test this, we performed an ANOVA on log-transformed mean looking time as the dependent measure, with participants as the random factor, Condition (Negative Noun-Switch vs Negative Verb-Switch) as a between-participant factor, and Phase (Habituation vs Test) as a within-participant factor. The expected effect should appear as a significant interaction between Condition and Phase. Note that mean looking times were log-transformed before running the ANOVA because the data did not follow a normal distribution.

Results

Infants' looking times toward the visual stimuli during the last two trials of the habituation phase and during the two trials of the test phase are represented in Figure 2.

An ANOVA on log-transformed mean looking times revealed a significant interaction between Condition (Negative Noun-Switch vs Negative Verb-Switch) and Phase (Habituation vs Test) $F(1,46)=4.24$, $p=.04$; $d=0.615$; this interaction reflects the fact that infants' looking times increased more between habituation and test in the Negative Noun-Switch condition (with 15 out of 24 children increasing their looking between habituation and test) than in the Negative Verb-Switch condition (with only 8 out of 24 children increasing their looking between habituation and test). This result shows that children looked more toward the videos (i.e., were more surprised) when listening to negative sentences that were rendered false by their visual context (the Negative Noun-Switch condition) than when listening to negative sentences that were true relative to their visual context (the Negative Verb-Switch condition). Thus, by 18 months of age, infants exhibit a different behavior for correctly employed over incorrectly employed negative sentences, indicating that they were able to interpret negative sentences correctly.

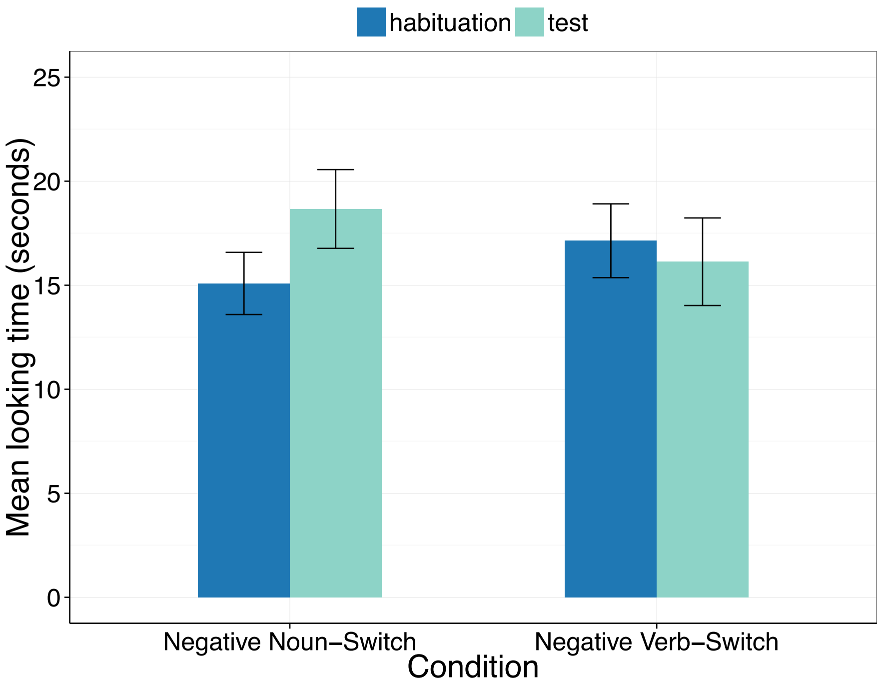


Figure 2: Mean looking time in seconds toward the videos during the last two trials of the habituation phase (in blue) and during the two trials of the test phase (in green) for children assigned to the Negative Noun-Switch Condition (on the left; *N*=24) and to the Negative Verb-Switch Condition (on the right; *N*=24). Error bars represent the standard error of the mean.

Given that both groups were exposed to exactly the same videos and sentences in the habituation phase, the only way to explain the asymmetry observed in the test phase, is that infants were able to learn the meaning of the novel words during habituation (as established in de Carvalho et al., 2015 and He & Lidz, 2017), to correctly interpret the meaning of the negative sentences, and to evaluate whether they were true or false in the context in which they were uttered. Since at test, the associations between the actions and the target words were switched, but the penguin was always present in both videos, a negative sentence saying that the penguin was not doing the previously learned action was true in that context (cartwheeling and spinning are indeed different actions), but negative sentences saying that

"It is not a bamoule" (i.e., in which bamoule = penguin), were false, since a penguin was still present on the screen. Accordingly, infants in the Negative Noun-Switch condition increased their looking time between habituation and test, showing surprise, relative to infants in the Negative Verb-Switch condition.

In addition, contrary to previous studies investigating the acquisition of negation with infants younger than 2, here we can clearly conclude that infants did not process negative sentences as if they were affirmative sentences: the pattern of results that we obtained with negative sentences is the exact opposite of what was observed with affirmative sentences in de Carvalho et al. (2015) (see Supplementary Figure S1 for a direct comparison of the experimental results of these two experiments). A joint analysis of the two experiments (affirmative vs negative sentences), reveals a significant triple interaction between Condition, Phase and Experiment ($F(1,92)=9.85$, $p < .003$; $d = 1.015$). This reflects the fact that while infants tested with affirmative sentences were more surprised by the Verb-Switch condition (they heard *She is pirdaling*, but the action had changed), than by the Noun-Switch condition (they heard *It is a bamoule* and although the action had changed, there was still a penguin on screen), in the current study, the negative versions of the test sentences produced the reverse pattern of results: infants were more surprised by the Negative Noun-Switch condition than by the Negative Verb-Switch condition.

Discussion

These results show for the first time that 18-month-olds are able to understand negative sentences and evaluate whether they are used appropriately or not, depending on context. In the current experiment, after having learnt that *bamoule* means "penguin" and *pirdaling* means "cartwheeling", infants were surprised when listening to negative sentences that were rendered false by their context, such as *Look! It is not a bamoule!*, while watching a penguin cartwheeling. In contrast, they were not surprised when listening to a negative sentence that was true in its context, such as *Look! She is not pirdaling!*, while watching a penguin spinning (instead of cartwheeling).

Crucially, this pattern of results was the exact opposite of what was observed with affirmative sentences in de Carvalho et al. (2015) (in French) and He and Lidz (2017) (in English), in which 18-month-olds were more surprised by the Verb-Switch condition than by the Noun-Switch condition. Thus, this experimental procedure allowed us to observe opposite patterns of behavior for affirmative and negative sentences, showing that 18-month-olds process negative sentences correctly, when they are presented in a supportive context.

Note that by using novel words, our study allowed us to test not only whether infants can interpret negative sentences, but also whether they can generalize their interpretation of negation even to sentences containing words that they have never heard in a negative context before. Given the type of negation that children produce with familiar words at this age (i.e., "refusal" and "non existence"), with simple syntactic frames such as "no-juice or no-cat", it is difficult to establish whether children really know how to use the negative operator, or whether they may have simply associated a specific meaning to the chunk "no-cat", as if it was one single word that is uttered in the absence of cats. Crucially, to succeed in our task with novel words, children cannot have used a low-level strategy based on their experience with the novel words 'bamoule' in affirmative versus negative contexts, noticing for instance that 'une bamoule' was produced in the presence of a penguin, and 'pas une bamoule' in its absence. Instead, they really need to have an understanding of the syntactic function of negation, together with the meanings of the novel words, otherwise "pas une bamoule" and "pirdale pas" would have triggered the same behavior during test.

Our study thus provides the first evidence for the understanding of negative sentences during the second year of life, and suggests that previous failures in the literature might not be due to infants' inability to process negative sentences *per se*, but rather to the complexity of the tasks used that required sophisticated inhibitory skills from young children, or to the fact that negative sentences were not felicitous in their context.

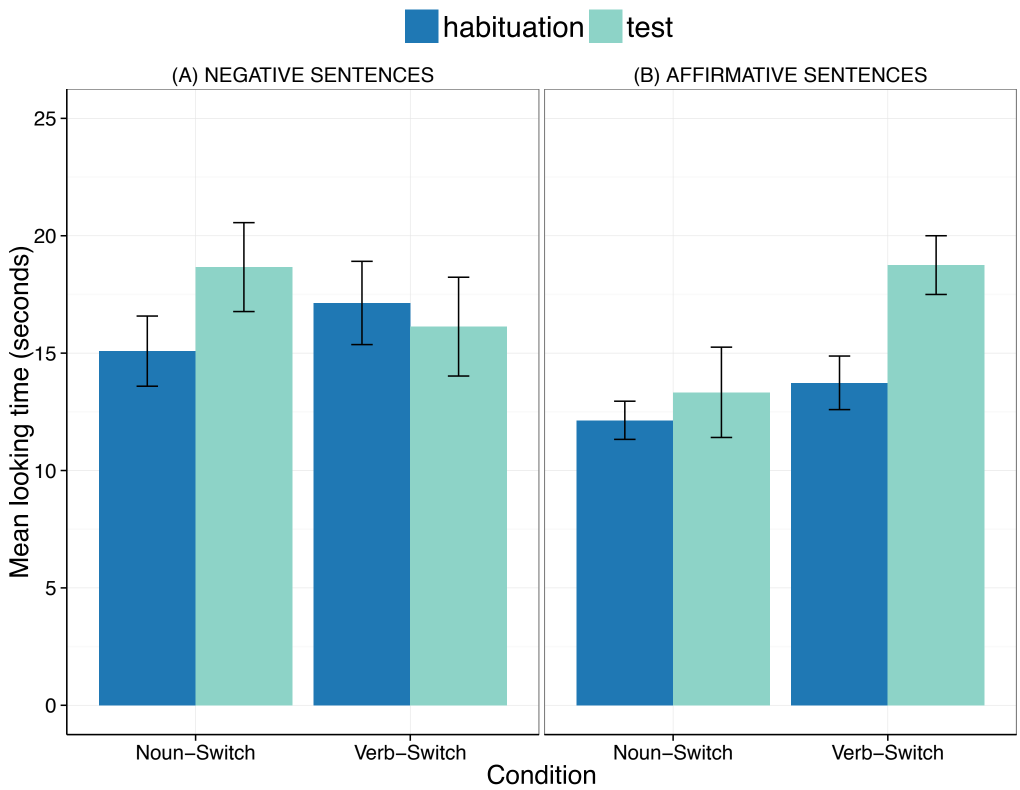
Our paradigm improved on that of earlier studies, since it did not tax young infants' inhibitory skills: the test phase presented a single video at a time, thus avoiding competition between a scene that would make the negative sentence true (target response) and a scene that would merely display objects mentioned in that same sentence (attractive - but wrong - alternative). Moreover, negative sentences were felicitous, because their positive counterpart was uttered earlier in the habituation phase, and because in the test phase, negative sentences introduced a contrast between what had been seen during habituation and what was presented in the test phase (i.e., the switch between actions). Negative sentences were thus negating expectations that the listener had constructed during the habituation phase with the affirmative sentences. Overall, this situation created a supportive context for the processing of negation. Another feature of the current experimental procedure which might have made the task easier, is that the sentences were repeated several times during the habituation and test trials, which gave infants ample time to process them correctly.

Previous failures in the literature posed a developmental mystery in the acquisition of negation: infants seem to start to produce something that they do not understand yet. The causes of this apparent delay between the production of negation and its understanding remained however unknown. This was a serious problem given that a lack of understanding of negative sentences during the second year of life could impact infants' language acquisition. If infants were unable to parse negative sentences, they would face difficulties in learning word meanings when listening to negative sentences such as "This is not a cat. This is a wolf" If the child interprets the negative sentence as an affirmative one (as suggested by previous results) she would probably interpret these sentences as "*This is a cat. This is a wolf"* and would then be confused about what the correct name of the object is. Similarly, if a caregiver tries to forbid a child to do something and says "This is not possible", if the child processes this negative sentence as affirmative, she will interpret the sentence as "This is possible". The present study shows that by 18 months, infants do not process negative sentences such as *This is not an X* as being affirmative, otherwise they would not be surprised when hearing *This is not a penguin* while seeing a penguin on screen.

In the current study, we tested the hypothesis that the previously attested delay between the production and understanding of negation, rather than representing an inability to understand negation between 12 and 24 months, might be related to the fact that the experimental tasks used so far were too demanding for infants. Using a word-learning task that requires no inhibitory skills from infants, we found evidence that 18-month-olds already understand negative sentences: infants were surprised when listening to negative sentences which were false in their context, but they were not surprised when listening to negative sentences which were true in their context. This provides the first evidence for the understanding of negative sentences before two years of age. The ability to understand negative sentences may impact infants' education, since it gives them access to what parents do or do not allow them to do; in addition, since negative sentences can be used to narrow down the space of hypotheses/possibilities in a given situation, understanding them may represent an important tool for infants to constrain their interpretation of different situations, and to support language acquisition.

SUPPORTING INFORMATION

Supplementary Figure S1:



Supplementary Figure S1: Comparison of our current results with negative sentences (A) and the results obtained in de Carvalho et al., (2015- Experiment 1) with affirmative sentences (B). Blue bars represent mean looking time in seconds toward the videos during the last two trials of the habituation phase and green bars represent mean looking time toward the videos during the two trials of the test phase, for children assigned to the Noun-Switch Condition (on the left; *N*=24 in each experiment) and to the Verb-Switch Condition (on the right; *N*=24 in each experiment). Error bars represent the standard error of the mean. Each trial has a maximal duration of 50 seconds in the experiment with negative sentences (A) and 37 seconds in the experiment with affirmative sentences (B). Note that while in both experiments infants were habituated with the same kind of videos and sentences, in Figure (A) we observe that infants tested with negative sentences were more surprised by the Noun-Switch condition than by the Verb-Switch condition, while in Figure (B) we observe the reverse pattern of results: infants were more surprised by the Verb-switch condition than by the Noun-Switch condition. Note that the study of de Carvalho et al., 2015 did not investigate the processing of affirmative sentences *per se* but rather whether infants would use morho-syntactic cues (Experiment1) and prosodic cues (Experiment2) to constrain their syntactic analysis and learn word meanings.

References

Austin, K., Theakston, A., Lieven, E., & Tomasello, M. (2014). Young children's understanding of denial. *Developmental Psychology*, *50*(8), 2061–2070. http://doi.org/10.1037/a0037179

Bernal, S., Lidz, J., Millotte, S., & Christophe, A. (2007). Syntax Constrains the Acquisition of Verb Meaning. *Language Learning and Development*, *3*(4), 325–341. http://doi.org/10.1080/15475440701542609

Bloom, L. (1970). Syntactic and Semantic Development of Early Sentence Negation. In L. Bloom (Ed.), *Language Development: Form and Function in Emerging Grammars* (Research M, pp. 170–220). Cambridge, Massachussetts and London, England: MIT Press.

Cameron-Faulkner, T., Lieven, E., & Theakston, A. (2007). What part of no do children not understand? A usage-based account of multiword negation. *Journal of Child Language*, *34*(2), 251–282. http://doi.org/10.1017/S0305000906007884

Choi, S. (1988). The semantic development of negation: a cross-linguistic longitudinal study. *Journal of Child Language*, *15*(3), 517–531. http://doi.org/10.1017/S030500090001254X

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

de Carvalho, A., He, X. A., Lidz, J., & Christophe, A. (2015). 18-month-olds use the relationship between prosodic and syntactic structures to constrain the meaning of novel words. In *Paper presented at the 40th Boston University Conference on Language Development*.

Drozd, K. F. (1995). Child English pre-sentential negation as metalinguistic exclamatory sentence negation. *J Child Lang*, *22*(3), 583–610. http://doi.org/10.1017/S030500090000996X

Feiman, R., Mody, S., Carey, S., & Snedeker, J. (2014). The acquisition of verbal negation: early comprehension and the emergence of a combinatorial language of thought. In *Paper presented at the 39th Boston University Conference on Language Development*. Boston - MA.

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

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

Guidetti, M. (2005). Yes or no? How young French children combine gestures and speech to agree and refuse. *Journal of Child Language*, *32*, 911–924. http://doi.org/10.1017/S0305000905007038

Haryu, E., & Kajikawa, S. (2016). Use of bound morphemes (noun particles) in word segmentation by Japanese-learning infants. *Journal of Memory and Language*, *88*, 18–27. http://doi.org/10.1016/j.jml.2015.11.007

He, A. X., & Lidz, J. (2017). Verb learning in 14- and 18-month-old English-learning infants. *Language Learning and Development*, *5441*(March).

Hummer, P., Wimmer, H., & Antes, G. (1993). On the origins of denial negation. *Journal of Child Language*, *20*(3), 607–618. http://doi.org/10.1017/S0305000900008503

Kim, K. (1985). Development of the concept of truth-functional negation. *Developmental Psychology*, *21*(3), 462–472. http://doi.org/10.1037/0012-1649.21.3.462

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

McNeill, D., & McNeill, N. B. (1967). A question in semantic development: What does a child mean when he says "no"? In *Studies in Language and Language Behavior, Progress Report IV*. Ann Arbor, MI: Center for Research on Language and Language Behavior, University of Michigan.

Nordmeyer, A. E., & Frank, M. C. (2014). The role of context in young children's comprehension of negation. *Journal of Memory and Language*, *77*(C), 25–39. http://doi.org/10.1016/j.jml.2014.08.002

Nordmeyer, A. E., & Frank, M. C. (2015). The pragmatics of negation across contexts. In *Proceedings of the 37th Annual Conference of the Cognitive Science Society*.

Pea, R. D. (1980). The Development of Negation in Early Child Language. In D. R. Oslon (Ed.), *The Social Foundations of Language and Thought* (pp. 156–186). New York and London: W W Norton & Company.

Reuter, T., Feiman, R., & Snedeker, J. (2014). Young children's comprehension of negation. In *Paper presented at the 39th Boston University Conference on Language Development*. Boston - MA.

Tam, C. W. Y., & Stokes, S. F. (2001). Form and function of negation in early developmental Cantonese. *Journal of Child Language*, *28*(2), 373–391. http://doi.org/10.1017/S0305000901004688

Vaidyanathan, R. (1991). Development of forms and functions of negation in the early stages of language acquisition: a study in Tamil. *Journal of Child Language*, *18*(1), 51–66. http://doi.org/10.1017/S0305000900013295

Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology*, *59*(1), 67–95. http://doi.org/10.1016/j.cogpsych.2009.02.001

Werker, J. F., Cohen, L., Lloyd, V., & Casasola, M. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology*, *34*(6), 1289–1309.

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

Chapter 9: General Discussion

Infants acquiring language face the challenging task of having to learn by extracting information directly from the speech stream. Given the complexity of the spoken input, for instance the fact that in fluent speech there are no pauses between two consecutive words and that for each spoken sentence the world offers a wide array of possible referential intentions, it is impressive that around the age of two or three years, toddlers already show a deep knowledge about aspects of the phonology, semantics and syntax of their native language. Motivated by the studies of the mechanisms that would allow infants to build this impressive knowledge, the current thesis investigated how young children could map the sounds of words they hear to a possible meaning in their environment.

I focused my attention on the perception and use of two cues that seem to be available early to infants: phrasal prosody and function words. I explored whether these two sources of information would provide infants with the means to access the syntactic structure of sentences and whether this information in turn would allow them to constrain their interpretation of word meanings. More precisely, I investigated whether phrasal prosody may serve as a cue to segment the speech stream into groups of words that correlate with syntactic constituents, and whether function words and morphemes may serve as a cue to determine the syntactic nature of the units delimited by phrasal prosody. The empirical studies conducted during this thesis show that during the first steps of language acquisition, when infants do not know many words yet, they are able to rely on function words and phrasal prosody to access syntactic information from spoken sentences, and that they use this information to constrain the acquisition of novel word meanings.

In Part 1, I showed that French and American children are able to exploit phrasal prosody and function words together to access the syntactic structure of sentences, which in turn allows them to identify the syntactic category of an ambiguous word and recover its meaning. We observed that from 20 months onwards, infants are able to correctly assign a grammatical category to an ambiguous word (a noun or a verb) depending only on its position within the prosodic structure of sentences. When presented with ambiguous sentences that were phonemically identical but syntactically and prosodically distinct, toddlers were able to exploit the prosodic structure of sentences to infer their syntactic structure, and to use this information to decide whether an ambiguous target word was a noun or a verb. Young children interpreted the ambiguous target word as a noun when it was embedded in a noun sentence and as a verb when it was embedded in a verb sentence, even though the only cue to syntactic structure was given by phrasal prosody. These studies confirm that young children are able to exploit phrasal prosody and function words together in order to recover the syntactic structure of sentences, and that they use this representation of syntactic structure to infer the syntactic category of an ambiguous word and to access its meaning.

These findings suggest that at an age when children's knowledge of content words is rather limited, but phrasal prosody and function words are available, toddlers can rely on phrasal prosody and function words to retrieve a partial syntactic representation of spoken sentences and attribute meaning to words, depending on their position in the prosodic-syntactic structure of sentences. This mechanism might be extremely important during the first stages of language acquisition, because in the absence of knowledge about the meanings of words, children could exploit the fact that an unknown word occuring in a noun context refers to an object, while words occurring in verb contexts probably refer to actions. In other words, phrasal prosody and function words may be important for infants to achieve the sound-to-meaning mapping.

This is what I tested in part 2: whether phrasal prosody and function words could help infants access syntactic information about novel words, and whether they use this information to constrain their acquisition of the meaning of these novel words. I showed that from 18 months, infants are able to exploit the morpho-syntactic information carried by the function words in the sentence (e.g., articles, determiners, pronouns) to assign a syntactic category to a novel word and to constrain its meaning. Strikingly, when the information provided by function words alone was not sufficient to compute the syntactic category of the novel words, we observed that infants were able to simultaneously exploit prosodic information and function words to recover the syntactic structure of sentences, and that they used this representation of syntactic structure to compute the syntactic category of novel words and therefore constrain their meanings.

The strength of the syntactic computations provided by function words was tested even in situations where nouns and verbs contexts appeared in negative sentences (i.e., "Ce n'est pas une bamoule" – It is not a bamoule - vs. "Elle ne bamoule pas" – She is not bamouling). Our results demonstrated that 18-month-olds were able to understand these negative sentences, therefore to compute the syntactic function of negation. The ability to exploit function words and understand negative sentences early during development should support language acquisition and help infants to constrain the acquisition of word meanings.

This thesis provides empirical evidence in favor of the *prosodic bootstrapping* and the *syntactic skeleton* hypotheses for language acquisition. We demonstrated that by providing information about the syntactic structure of sentences, phrasal prosody and function words allow infants to access syntactic information from the spoken input. This information in turn, as predicted by the *syntactic bootstrapping hypothesis*, allows infants to figure out which part of the world is being talked about (e.g., associating nouns to objects, and verbs to actions) and guides their discovery of word meanings.

Our results show that even in the absence of a very developed vocabulary and prior to being able to produce multi-word utterances, 18-month-olds already have the means to go from the sounds of words in their language to the syntactic structure of sentences, and that they can use this information to infer the probable meaning of a novel word. Infants exploited the prosodic information in speech to parse spoken sentences into groups of words and to identify possible syntactic constituents. They exploited this constituent structure, together with function words, to determine the syntactic nature of the constituents delimited by phrasal prosody, infer the syntactic category of novel words inside these units, and thus constrain their probable meaning.

This mechanism would be extremely useful for infants to get their lexical and syntactic acquisition started and to provide them with a tool to construct a first-pass syntactic structure of spoken sentences, even before they know the meanings of many words. Thus the current thesis provides an important contribution to our understanding of the chicken-and-egg problem stated in the introduction: Infants seem to need syntax to learn word meanings, and to need word meanings to learn syntax. In the current thesis we provide experimental evidence that phrasal prosody and function words can work as anchors to help infants access syntactic information even before they know the meanings of words in their language. These findings suggest that phrasal prosody and function words play an important role in language acquisition, since they provide access to a first-pass syntactic structure of sentences which may help infants to bootstrap language acquisition.

Our results give rise to several new questions and perspectives that will have to be addressed in future research. Given our participants' remarkable ability to recognize noun and verb contexts, we now need to find out how infants learned noun and verb contexts during their development. Their ability to exploit phrasal prosody to recover syntactic structures begs the question of whether infants need to learn that prosodic boundaries map to syntactic boundaries, and if yes, how infants might learn the mappings from prosody to syntax. Additionally, we need to investigate whether this mechanism to exploit phrasal prosody and function words may also apply to other languages, and serve as a universal cue to access syntactic information. Our results show that phrasal prosody is useful for syntactic analysis and that young children can use this information to constrain their syntactic analysis and therefore to guide their learning of word meanings. However, given that in real life not all syntactic boundaries are marked prosodically, there might be situations in which listeners need to constrain their syntactic analysis without the help of phrasal prosody. As such, it is possible that listeners would rely most often on morpho-syntactic cues to constrain their syntactic analysis and only in the few cases where morpho-syntactic information is not sufficient to constrain their interpretation, they would resort to prosody as a supplementary cue to constrain their parsing. This begs the question of whether phrasal prosody and function words/morphemes may play different roles in bootstrapping the acquisition of syntax, and whether phrasal prosody is really necessary for lexical and syntactic acquisition. I will discuss these questions in the following sections.

9.1. How do children learn noun and verb contexts during their development ?

The studies presented in this thesis suggest that children as young as 18 months are able to use phrasal prosody and function words to identify nouns and verbs contexts. The question that arises is how infants managed to learn in which contexts nouns and verbs occur.

As we reviewed in our introduction, infants are sensitive to the function words of their native language before their first birthday (e.g., Hallé et al., 2008; Kedar et al., 2017; Shafer et al., 1998; Shi & Lepage, 2008; Shi, Werker, et al., 2006), and between 12 and 18 months they use this information to speed up their lexical access and to recognize content words: they know for instance that determiners are followed by nouns, and that pronouns are followed by verbs (Cauvet et al., 2014; Kedar et al., 2006; Van Heugten & Johnson, 2011; Zangl & Fernald, 2007). Infants presented to a novel word such as “blick” preceded by function words that frequently co-occur with nouns in their language (e.g., the blick, some blick), then prefer to listen to this novel word “blick” in other noun contexts such as “a blick” rather than in verb contexts such as “I blick” (Höhle et al., 2004, in German; Shi & Melançon, 2010, in French).

Around 18 months of age, infants exploit the syntactic context provided by function words to infer the plausible meaning of unknown content words. For instance, a novel word presented in a noun context (i.e., after a determiner as in "It is a *doke*") was assumed to refer to an object, while in verb contexts (i.e., after a pronoun and an auxiliary verb as in "It is doking) "doke" was assumed to refer to an action (in French: Bernal et al., 2007; de Carvalho, He, Lidz, & Christophe, 2015; in English: He & Lidz, 2017; Waxman et al., 2009; and in Japanese: Oshima-Takane, Ariyama, Kobayashi, Katerelos, & Poulin-Dubois, 2011). These studies show that infants can exploit function words to recognize noun and verb contexts and that they exploit this information to map unknown words with their most probable referent: words are mapped to objects or to actions, depending on their contexts. However, how do infants manage to figure out which contexts correspond to specific syntactic categories?

One hypothesis suggests that to learn nouns and verbs contexts infants have to be able to analyze the distributional information of their input in order to identify which words or sets of words co-occur with words from specific categories (e.g., Redington, 1998). Supporting this hypothesis, computational studies up to date have shown that models relying on distributional information such as the position of articles, determiners and pronouns in the sentences are useful for categorization (Chemla, Mintz, Bernal, & Christophe, 2009; Mintz, 2003; Mintz, Newport, & Bever, 2002; St. Clair, Monaghan, & Christiansen, 2010). In Mintz (2003) for instance, the notion of frequent frames as distributional patterns based on co-occurrence patterns of words in sentences was exploited to create a model for grammatical categorization of words in child-directed speech. That model grouped together all words X appearing in a context or "syntactic frame" (defined as two jointly occuring words with one word X intervening) of the type [A X B], where A and B were two words that frequently occurred together. The results showed that the model builds highly accurate syntactic categories for words. For instance, frames such as [the X is] accurately categorized words as nouns, while frames such as [you X the] categorized words as verbs. In addition to these findings, empirical studies demonstrated that very young infants can use distributional information to categorize words both in artificial and natural languages (Gomez & Gerken, 1999; Marchetto & Bonatti, 2013; van Heugten & Johnson, 2010).

These computational models show that a considerable amount of information concerning the syntactic categories can be obtained from distributional information: they all performed better-than-chance in a categorization task. However, all these models faced some limitations in this task, such as how to categorize words that can appear both in noun or verb contexts (i.e., I watch vs. the watch) or how to handle the fact that several groups of words (frames) correspond to each syntactic category.

In an attempt to solve some of these problems, Brusini and colleagues (Brusini, Amsili, Chemla, & Christophe, 2011; Christophe et al., 2016) presented a model that gets trained on a corpus of child-directed speech in which a few words are initially categorized. Specifically, they hypothesized that when children start addressing the categorization problem, they already managed to learn the meaning of a few highly frequent content words, and to group them into distinct categories: words referring to objects vs. words referring to actions. Infants would then rely on this handful of familiar words, the *semantic seed*, to learn the syntactic contexts in which they appear.

In other words, the *semantic seed* is a small lexicon that children have learned from their environment (i.e., a few highly frequent content words) and that they would use to infer some information about other (unknown) words. For instance, by knowing that a word such as "bottle" refers to a concrete object and that this word is often uttered in a context such as "the bottle", "your bottle", etc, infants could infer that any other word appearing in the same context (e.g., "the blicks", "your blicks") may also refer to a concrete object.

In favor of this hypothesis, several studies found evidence that from 6 months onwards infants already know the meaning of some frequent nouns in their language (Bergelson & Swingley, 2012; Parise & Csibra, 2012; Tincoff & Jusczyk, 1999, 2012) and that around 10 months they already know the meaning of some verbs (Bergelson & Swingley, 2013). In addition, the idea of the *semantic seed* for syntactic bootstrapping is that infants could group words together according to their semantic category (i.e., words referring to objects versus words referring to actions). For instance, infants could start grouping together words such as *bottle, banana, mouth,* and *spoon* because they all refer to concrete objects in their environment, and other words such as *kissing, drinking,* and *eating* could be grouped together because these words refer to actions.

Several studies show that infants are able to group concepts semantically, and form categories such as objects, agents, artefacts, or actions (Carey, 2009; Ferry, Hespos, & Waxman, 2010; Havy & Waxman, 2016; Saxe & Carey, 2006; Waxman & Lidz, 2006; Waxman & Leddon, 2011). For instance, there is evidence showing that from three months onwards, infants distinguish between animate and inanimate objetcs (e.g., Woodward, Sommerville, & Guajardo, 2001) and between agents and non-agents in actions (e.g., Newman, Keil, Kuhlmeier, & Wynn, 2010; Opfer & Gelman, 2010) and that the acquisition of words support their acquisition of concepts (e.g., Gelman, 2003; Graham, Kilbreath, & Welder, 2004; Waxman & Booth, 2001).

Thus, given that nouns are likely to refer to objects and verbs to actions, these two basic semantic categories provided by highly frequent familiar words may represent a seed that infants would be able to use as a prototypical form of noun and verb grammatical categories in their native language. If infants are able to notice in which contexts the words refering to objects often appear (e.g., after a determiner), and in which contexts the words referring to actions often occur (e.g., after a pronoun, or after a full noun phrase) they might be able to decide that an unknown word appearing in the same contexts, also belongs to the object category (as "blicks" in the sentence "the blicks slept") or to the action category (as "blicks" in the sentence "the bottle is blicking").

Taken together, all these approaches suggest that a minimal semantic knowledge, together with an analysis of the distribution of functional element can be extremly important for infants to learn the most useful contexts for categorization. In the current thesis, we added phrasal prosody as an important ingredient for the syntactic categorization of words. In accordance with our findings, recent computational work shows an excellent performance of models relying on a combination of phrasal prosody, function words and a minimal semantic knowledge, to access the syntactic category of unknown words (Brusini, Amsili, Chemla, & Christophe, 2011; Christodoulopoulos et al., 2016; Fisher, 2015; Gutman et al., 2015).

9.2. Do children need to learn the mapping between prosody and syntax ?

The empirical findings in this thesis raise the question of whether infants need to learn that prosodic boundaries map to syntactic boundaries, and if yes, how they might learn the mappings from prosody to syntax. Our results *per se* do not allow us to provide a definite answer to this questions. I will nonetheless discuss some possibilities.

One possibility is that children would not need to learn that prosodic boundaries coincide with syntactic boundaries because they would assume that the prosodic boundaries they hear in the speech stream also delimit all other linguistic elements (words, syntactic constituents, sentences, etc). For instance, we could assume that in absence of any information about what are the words in their language, infants would pay attention to patterns of intonation in the speech (e.g., silent pauses, final lengthening, pitch discontinuity, etc), and that they could use this information to segment the speech stream into chunks of words. Later on during their development and together with the information provided by morpho-syntactic cues and some familiar content words, infants would start to discover what the elements inside the prosodic units are.

Note however that sensitivity to some prosodic information for parsing seems to require infants experience with their native language before they can exploit these cues as reliable boundary markers in their language. As I mentioned in the introduction, there is experimental evidence showing that while infants around 6-months-old (even in different languages) tend to rely on strong and universal prosodic markers, such as pauses when processing speech, it is only after the age of 8 months that infants become able to use more subtle prosodic cues such as pitch contours and lengthening (in the absence of pauses) to recognize some other types of prosodic boundaries (Johnson & Seidl, 2008; Seidl, 2007; van Ommen et al., 2017; Wellmann et al., 2012). Given these results, one could assume that infants would start to segment the speech stream with the information provided by the strong prosodic markers (i.e., marking the end of the sentences for instance) and during their development they will begin to generalize that other kinds of prosodic cues in absence of pauses could also mark some boundaries in the middle of the sentences.

Another possibility, which is not mutually exclusive with the first, is that infants can learn the mappings from prosody to syntax through their experience with unambiguous utterances that they hear in their everyday lives. In this view, the relationship between prosodic and syntactic structure needs to be robust enough to be learned by infants through exposure to their native language. Given that the prosodic boundary we exploited in our studies can be observed in several sentences that infants will hear in their everyday lives, even when the sentence is not ambiguous (e.g., [The little cat] [jumps really high]), we believe that the relationship between prosody and syntax can be robust enough to be learned by infants during their development.

Although in this thesis we needed ambiguous sentences to test infants' abilities to rely on phrasal prosody to recover syntactic structure, we believe that listeners can indeed learn the relationship between syntax and prosody from unambiguous sentences. In favor of this idea, Gutman and colleagues (2015) demonstrated that in a corpus of child-directed speech (Demuth & Tremblay, 2008) containing conversations with four children aged between 1 and 4 years, 246.013 prosodic phrases (as defined by Nespor & Vogel, 1986) might be present for a total of 180.000 utterances. This implies a ratio of more than one (i.e., 1.4) prosodic phrase per utterance in average. This suggests that prosodic boundaries signaling syntactic boundaries (both at the intonational phrase level – often coinciding with the end of sentences; or at the phonological phrase level – typically within sentences) may be present in children's everyday spoken input and could be learned through exposure to language.

9.3. Can phrasal prosody and function words support access to syntactic information in other languages?

If phrasal prosody and function words play an important role in bootstrapping word meaning acquisition, as I claimed in this thesis, then one would expect this to be true across the world's languages. In other words, is it the case that phrasal prosody and function words are universal cues for syntactic parsing? The answer to this question is extremly hard, because we know that although phrasal prosody and function words/morphemes are present in all the world's languages, prosodic information and functional elements can surface differently across languages.

With regards to prosodic information, for instance, prosodic cues can surface differently across languages. At the lexical level for instance, languages can use tone, stress, lexical pitch-accent, a combination of one or more of these cues, or none of these cues. For example, at the word level, Mandarin has tone and stress, but Cantonese has tone and no stress. Swedish and Chickasaw have stress and lexical pitch accent, but Japanese and Basque have only lexical pitch accent. English and other West Germanic languages have only stress, while Korean, West Greenlandic and Halh Mongolian have none of the above mentioned cues at the word level. At the utterance level, languages can also differ in the way the intonational contours appear in speech (see Jun, 2005, 2014, for a full description of the prosodic typology of languages).

With regards to function words, functional elements can differ in the way they are combined to indicate grammatical relationships across languages, and in the positions that they occupy in the utterances (as we explained before, i.e., function words/morphemes can appear before or after content words). Languages like Chinese, for instance are considered as Isolating languages because they have a very low morpheme per word ratio. In Chinese, each word tends to contain a single morpheme. Languages classified as Inflectional/fusional can also overlay many morphemes to denote grammatical, syntactic, or semantic changes.

Examples of those languages are Italian, Russian, French, Portuguese and Spanish where we can fusionate a morpheme into a word and/or change a single morpheme into the word to express grammatical features (e.g., in Italian, they use "*ragazzo*" to say – 'boy', but "*ragazzi*", changing the last vowel –o for –i, to say 'boys'; in Portuguese, "menino" means 'boy', but "meninos" with an –s in the end means 'boys'. In Spanish and Portuguese, "comer" means 'to eat', but "*comi*" means 'I ate'). There are also agglutinating languages such as Hungarian, Turkish, Basque in which words contain different morphemes to determine their meaning, but each of these morphemes remains in every aspect unchanged after their union. For instance in Hungarian, one can say "ház-a-i-tok-ba" (glued together), which means 'into your houses', in which the morphemes –a, -i, -ba, remain unchanged. Finally, there are also Incorporating/polysynthetic languages such as Amerindian languages (Navajo, Hopi) in which they can form "long-sentence words" where words are composed of many morphemes (e.g., the "word" *tavvakiqutiqarpiit* means 'Do you have tobacco for sale?') (Dryer, 1992, see also the World Atlas of Language Structures available online).

These surface differences on how languages use prosodic information and function words/morphemes might lead us to think that the use of these cues during sentence processing might be simpler in some languages than others. For instance, in languages where a function word directly precede or follows a content word, it might be easier to infer the grammatical category of the content words compared with languages where a morpheme will change the word form to express grammatical relationships. Also, in some languages phrasal prosody might mark the boundaries between syntactic constituents more clearly than in other languages.

In a first attempt to test the potential universality of these cues, in the current thesis we tested the use of phrasal prosody and function words to constrain syntactic analysis in two different languages: French and English. This choice was motivated by the fact that French and English are different in the way they use prosodic information. As we explained in Chapter 4, while English uses prosody for certain phenomena (e.g., focus), French does not, relying instead on alternative devices such as word order, morphological cues, etc, to express the same phenomena. For instance, to mark focus, English typically relies on prosodic prominence (e.g., JOHN ate the apple), while French prefers to rely on syntactic devices such as fronting (e.g., *C'est Jean qui a mangé la pomme* – 'It's John who ate the apple'). Within the word level, English and French also have distinct types of prosody: English has contrastive stress, such that for each word a given syllable has to bear stress, while in French stress is fixed and word-final (typically marked by lengthening), which leaves more freedom for implementing a pitch-contour for a word, since it is not constrained by stress.

This situation led us to test the assumption that although both French and English use prosodic cues to mark prosodic units, the perception of these prosodic units might be easier in French than in English, because prosodic cues serve for many other purposes in English than in French. This suggestion has often come up at conference presentations of the results that we obtained in French, but we have not found any written discussion of this hypothesis in published work. Nevertheless, Chapter 4 provided evidence against this hypothesis. We observed that although French and English can use prosodic information for different purposes in each language, in the current thesis, both American and French children were able to use phrasal prosody to constrain their parsing.

These results lend support to the hypothesis that phrasal prosody cues syntactic structure in early language development, and likely in different languages. As we explained in the discussion of Chapter 4, we believe that the previous difficulties detecting children's ability to exploit the relationship between prosodic and syntactic structures for parsing were not a due to a weakness of prosody *per se*, but rather due to the fact that the link between prosodic and syntactic structure in previous studies was not sufficiently systematic in the structures that were tested. In cases where this relationship was more systematically marked, as in our studies, we observed that children were just as sensitive to prosody as one might expect. Taken together, our results provide cross-linguistic evidence for the role of phrasal prosody in children's syntactic analysis.

One criticism that could be drawn however is that the differences between French and English were not so strong, and that our results cannot really provide evidence in favor of the universality of phrasal prosody and functional elements for parsing. For instance, the functional elements that children exploited in conjunction with prosodic information in English and in French were very similar. In both languages they could rely on the articles or pronouns that appeared in the prosodic units, and in both languages, function words appear before content words. The open question then is whether children would be able to exploit phrasal prosody and function words in other kinds of languages. For instance, we may wonder whether children could use this information in a language that does not use pronouns and articles before a content word, but rather uses morphemes inside the content word to indicate grammatical relationships (e.g., in agglutinating languages), languages that have different canonical order for words, or languages that have a very different intonation than French and English.

In an attempt to investigate this question, we tried to design an experiment in Japanese, which is not a Indo-European language (as English and French), has a different word order and has very different prosodic countours than English and French (since it has pitch accent at the lexical level). In collaboration with Reiko Mazuka in Japan, we tried to design an experiment to test more directly whether Japanese-speaking children would be able to exploit phrasal prosody to constrain their syntactic analysis of noun-verb homophones (as we tested in English and in French). However, we did not manage to find sentences that would allow us to test the impact of prosody in parsing. Note that to test the impact of phrasal prosody in syntactic parsing, we need to find sentences in which only prosodic information can guide listeners intepretation of sentences. This means that all the words (i.e., function words, morphemes, and content words) in the sentence have to be the same, and only the prosodic structures should reflect the different interpretations that the sentence can have. This situation led us to use noun-verb homophones in our experiments. In addition, those homophones need to be known by young children, and they need to have a meaning that we can easily illustrate in a picture, to test their online interpretation of sentences while inspecting the images.

In Japanese, although we found some homophones that could be known by young children, we noticed that several homophone pairs (e.g., *ha'ru*$_{\text{Noun}}$ – 'spring' vs. *haru*$_{\text{Verb}}$ 'to post'; *tsu'ru*$_{\text{Noun}}$ 'crain' vs. *tsuru*$_{\text{Verb}}$ 'to fish'; *yo'ru*$_{\text{Noun}}$ 'night' vs. *yoru*$_{\text{Verb}}$ 'to stop by') could be disambiguated by pitch accent: the nouns are accented on the first mora and the verbs are unaccented. Moreover, since Japanese has a different word order for nouns and verbs, homophones often appear in different positions in the sentence. Although we did not manage to find a good test case to study Japanese children's ability to use phrasal prosody to constrain their syntactic analyses, this does not suggest that phrasal prosody cannot help them to access syntactic information. Maybe, in the future, we can take advantage of these pitch-accent minimal pairs, and test at least whether Japanese children can use these differences in pitch to figure out the syntactic role of the ambiguous word (noun vs. verb).

One piece of information to answer the question about whether phrasal prosody could constrain syntactic analysis in Japanese and in other different languages comes from recent studies using an artificial language (Hawthorne & Gerken, 2014; Hawthorne, Mazuka & Gerken, 2015). Hawthorne and Gerken (2014) tested whether phrasal prosody could help 19-

months-old English-learning infants to constrain the syntactic organization of sentences composed exclusively of non-words. For instance, when familiarized with non-word sentences exhibiting the prosodic pattern of two syntactic constituents, such as "[*bup div kagî*] [*feb zaf vot*]", infants then preferred to listen to sentences like [*feb zaf vot*] [*bup div kagî*] – where the order of the two constituents is changed – than to sentences like [*zaf vot bup*] [*div kagî feb*] in which words were moved around irrespective of the constituent structure. Hawthorne and Gerken (2014) explained that toddlers behaved in this way, because in natural languages, only words that are grouped in the same constituent can 'move' together.

Hawthorne and colleagues (2015) extended these findings to strings of nonwords with non-native prosody. Testing English- and Japanese-acquiring 19-month-olds with sentences from an artificial language with a non-native prosodic contour (i.e., English infants with Japanese prosody and Japanese infants with English prosody), the authors demonstrated that both groups were able to use phrasal prosody to parse the speech into cohesive and re-orderable syntactic constituent-like units. This finding suggests that the cues that marked syntactic constituent boundaries in these experiments (e.g., pauses, pitch variation and phrase-final lengthening) are important prosodic cues that toddlers can exploit even in non-native prosody.

It is important to note that the importance of prosody for parsing is not restricted to the identification of syntactic constituents, and several studies in different languages provide evidence that children can use prosodic information to constrain their interpretations of different aspects of language. Ito, Jincho, Minai, Yamane, & Mazuka (2012) demonstrated that Japanese adults and 6-year-old children exploit intonation (i.e., pitch accent) in contrast resolution. Zhou, Crain, & Zhan (2012) demonstrated that 4-year-old Mandarin-speaking children can use prosodic information to decide whether the speaker is asking a question or making a statement and in another study they also demonstrated that Mandarin-speaking preschoolers can use prosody to resolve ambiguities involving speech acts (Zhou, Su, et al., 2012). In German, 3-year-olds were able to use prosodic information to disambiguate the interpretation of focus particles (such as for the word '*auch*' – also ; Höhle, Berger, Müller, Schmitz, & Weissenborn, 2009). Very recently, it was demonstrated that Japanese children (from 6- to 7-year old) can exploit pitch accent information to resolve compound processing ambiguities (Hirose & Mazuka, 2017). In Japanese, similar to what was observed in Snedeker and Yuan (2008), it was also shown that five-year-olds can use prosodic information to interpret structurally-ambiguous modifiers when the prosody manipulation was blocked in a between-participants design, but they perseverate in failing to use prosody for parsing in a within-participants design (Mazuka & Tanaka, 2006).

With regards to function words, as we reviewed before, there is cross-linguistic evidence showing that infants can exploit function words for parsing in different languages. From 8 months of age infants use frequent functors to segment potential word forms from continuous speech in English, French, and in Dutch : languages where function words appear before the content word (Höhle & Weissenborn, 2003; Shi & Lepage, 2008; Shi, Cutler, Werker, & Cruickshank, 2006) ; but crucially, infants can also use bound morphemes (such as noun particles) in Japanese to constrain word segmentation (Haryu & Kajikawa, 2016). With regards to the use of function words to categorize content words, there is also cross-linguistic evidence showing that infants can use function words to determine the syntactic category of content words in French (e.g., Bernal et al., 2007; Shi & Melançon, 2010), in English (e.g., He & Lidz, 2017; Waxman et al., 2009), in German (e.g., Höhle et al., 2004), and in Japanese (e.g., Oshima-Takane et al., 2011).

Taken together, all the studies mentioned in this section, together with our current results in English and French, lend support to the hypothesis that phrasal prosody and function words may represent a universal and extremely useful tool for infants to access syntactic information through a surface analysis of the speech stream, and to bootstrap their way toward successful language acquisition.

9.4. Does phrasal prosody and function words/morphemes play different roles in bootstrapping the acquisition of syntax?

Another suggestion that has often come up at conference presentations and that was also raised by some reviewers of our papers was whether we could distinguish between the role played by phrasal prosody and the role played by function words to constrain parsing. In all of the studies presented in this thesis, we explicitly said that phrasal prosody was used as a cue to segment the speech into units (correlating with syntactic constituents), but that what allowed children to determine the syntactic nature of these constituents were function words.

A reviewer suggested that it was hard to buy our arguments about function words (in the studies conducted in Part 1) because he/she believed that similar effects for parsing could have been observed even without function words. For instance, with proper nouns, we could say [Mary Jane shoes] [are in style again], where “shoes” will be interpreted as a noun, or we could say [Mary Jane] [shooes the flies away], where "shooes" would be interpreted as a verb. We agree that in absence of function words, but in the presence of a full noun phrase containing proper names (delimitated by phrasal prosody) and a prosodic boundary, signaling a syntactic boundary between the noun and the verb phrases, the syntactic analysis of sentences would still be possible. In the same vein, we know that in several situations the boundary between syntactic constituents will not be marked prosodically and in these situations, children will have to rely only on function words to constrain their syntactic parsing (e.g., when an entire sentence is pronounced as one single prosodic unit: [She jumps]). These two situations (i.e., constraining syntactic analysis with prosody and without function words, or with function words but without prosody) lead to the question of which cue would be more important for children during language acquisition? Prosody or morpho-syntactic information?

As far as I can tell, the only piece of information to answer this question comes from a very recent study conducted by Massicotte-Laforge & Shi (2015). In this study, 20-month-old toddlers were initially shown to use phrasal prosody and morpho-syntactic cues from function words to identify syntactic constituents and to constrain word categorization even in absence of knowledge about the meanings of content words. For instance, 20-month-olds familiarized with jabberwocky sentences such as [*Ton*$_{\text{Det}}$ *felli*$_{\text{Adj}}$ *crale*$_{\text{N}}$]$_{\text{NP}}$ [*vur*$_{\text{V}}$ *la*$_{\text{Det}}$ *gosine*$_{\text{N}}$]$_{\text{VP}}$, where the novel word "*crale*" should be considered as a noun, listened longer to test trials (i.e., were surprised) when listening to short phrases presenting this novel word as a verb (e.g., "*Tu*$_{\text{Pron}}$ *crales*$_{\text{V}}$" – "You crale"), but they "were not surprised" when the novel word appeared in the expected syntactic context, as a noun "*Le*$_{\text{Det}}$ *crale*$_{\text{N}}$" – "The crale" (Massicotte-Laforge & Shi, 2015).

Surprisingly however, in a follow-up study (Massicotte-Laforge & Shi, 2016), the authors investigated whether phrasal prosody alone would enable 20-month-olds to analyze the same syntactic structures in the absence of function words. The familiarization sentences were all non-words, with the function words replaced by nonsense-functors such as in [*guin felli crale*$_{\text{N}}$]$_{\text{NP}}$ [*vur*$_{\text{V}}$ *ti gosine*$_{\text{N}}$]$_{\text{VP}}$, where "*crale*" could be considered as a noun or [*guin*$_{\text{Det}}$ *felli*$_{\text{N}}$]$_{\text{NP}}$ [*mige*$_{\text{V}}$ *vur ti gosine*$_{\text{N}}$]$_{\text{VP}}$, where "*mige*" should be considered as a verb. The results showed that infants familiarized with the novel word at the end of the first prosodic unit (i.e., [*guin felli crale*]) did not discriminate between grammatical (e.g., *Le crale* – The crale) versus ungrammatical (e.g., *Tu crales* – You crales) trials at test. Note that in the absence of information provided by function words (i.e., an article or a pronoun in the first prosodic unit), a word appearing at the end of a prosodic unit could be either a noun as in "[Mary Jane shoes]", but it could also be a verb as in "[Mary Jane eats]". Crucially however, infants familiarized to the structure presenting the novel word after a prosodic boundary and at the beginning of the second prosodic unit (e.g., [*guin*$_{\text{Det}}$ *felli*$_{\text{N}}$]$_{\text{NP}}$ [***mige***$_{\text{V}}$ *vur ti gosine*$_{\text{N}}$]$_{\text{VP}}$) listened longer to ungrammatical test trials presenting the novel word as a noun (e.g., *Le mige* – The mige) than grammatical test trials where the novel word appeared in the expected syntactic context, as a verb (e.g., "*Tu*$_{\text{pron}}$ *miges*$_{\text{V}}$" – You miges). Given that infants were able to build at least some syntactic computations even in absence of morpho-syntactic cues, the authors concluded that infants can use prosody to learn about syntax even in absence of function words. It is plausible that infants were able to infer that the novel word occupied a verb position in the sentences (even without function words), because of the same situation that we explained for the sentence "[Mary Jane] [shooes the flies away]". They might have infered that after a full noun phrase and a prosodic boundary, the word following is much more likely to be a verb than a noun. It would be ungrammatical in French to have a noun in that position, since nouns are obligatorily preceded by a determiner.

Although we could be tempted to answer the question about which cue might be more important for infants to constrain lexical and syntactic acquisition (i.e., phrasal prosody or function words), I think that the more reasonable way to think about this question is that infants will always use the most they can to bootstrap their way into language acquisition. If both phrasal prosody and function words provide useful cues for them, they should use both, rather than discounting information in the input. However, in situations where only one source of information is available, and the other one is absent or not reliable enough to guide their interpretations, infants would simple rely on the available cue in order to recover what is missing.

9.5. Is phrasal prosody useful for lexical and syntactic acquisition or is it necessary?

One of the criticisms that could be raised against our hypothesis about the mechanism according to which infants would use phrasal prosody and function words to bootstrap language acquisition, is with regards to the precise role played by phrasal prosody during parsing. Since not all the boundaries between noun phrases and verb phrases are marked prosodically, one would assume that the opportunities that infants will have to learn (or to use) the link between prosodic phrase boundaries and the constituents boundaries between subject noun phrase and verb phrase, for instance, may be too infrequent (see Kraljic & Brennan, 2005; Snedeker & Yuan, 2008, for a related discussion). If so, infants will have to rely on other more reliable cues to noun and verb phrases such as morpho-syntactic cues to learn about syntax rather than using a strategy based on prosody.

This idea that prosodic information might be too "weak" to constrain parsing in infants rests also on the fact that the literature has been mixed on whether speakers spontaneously and reliably produce prosodic cues to resolve syntactic ambiguities (see Kraljic & Brennan, 2005; Kraljic, Samuel, & Brennan, 2008; Michelas & D'Imperio, 2012; Millotte et al., 2008, 2007; Snedeker & Trueswell, 2003, for discussions about this aspect). If so, one may expect that listeners will give more weight to the information provided by morpho-syntactic cues rather than prosodic cues during parsing, since prosody would be supposedly less reliable than lexical content. Nevertheless, phrasal prosody is an inextricable component of language and it is present in all the world's languages. Thus, it is hard to imagine a situation in which listeners would process speech without prosody. Although the current thesis cannot provide direct evidence in favor of this hypothesis, I suspect that phrasal prosody may well be an important component of syntactic parsing rather than only an optional/additional cue for cases of ambiguity. Note that in addition to its importance to syntax, prosody also helps listeners to segment the speech stream into words and to constrain their lexical access. So without prosody, even some very basic levels of linguistic processing could be affected, such as word segmentation, lexical access and therefore parsing.

In an attempt to investigate these questions, we are planning to conduct two new lines of research. One line of research will study adult listeners and the other one will test young children. To evaluate whether phrasal prosody would be necessary (rather than only useful) for syntactic analysis, we will construct an artificial language (jabberwocky) in which the function words in French will be used (to preserve morpho-syntactic cues for parsing), and all the content words will be replaced by novel words (to avoid any top-down strategy for lexical access and word segmentation). Sentences will either have the typical prosodic information that characterizes phrasal prosody in French (resynthesized with the MBROLA software - Dutoit, 1997) or will not have any prosodic information (i.e., all syllabes will have the same duration and pitch contour). The idea behind these manipulations is to test whether adult listeners will still be able to conduct syntactic computations even in the absence of prosodic information (based only on morpho-syntactic cues) or whether success in parsing will only be observed in sentences containing prosodic information (in addition to morpho-syntactic cues).

In the second line of research, with children, unfortunately we cannot use jabberwocky sentences because we suspect the task would be too abstract for them. However, we can test how children would interpret the ambiguous noun-verb homophones (used in Part 1 of the current thesis) when the prosodic structure and the lexical content in the sentences provide conflicting information. If prosodic information is only a useful source of information and if morpho-syntactic cues and lexical content provide more reliable information than prosody, then we would expect that when prosodic and lexical information conflict, children would treat prosody as an inferior source of information to constrain their parsing, and rely on lexical content instead.

To test this, in a project that is still ongoing, we used the same homophones (as in Part 1) to create temporary ambiguities in French (e.g., *ferme*, meaning a farm or to close). Sentences began with prosody supporting either a noun or verb interpretation for the homophones, but ended with lexical content supporting the opposite interpretation. These sentences were created by cross-splicing two sentences which featured the homophone as a verb (e.g. [*la petite*]$_{NP}$ [*ferme*$_{V}$ *sa boîte à jouets*] – the little girl closes her toybox), and as a noun (e.g. [*la petite ferme*$_{N}$] [*sera pour les enfants*] – the little farm will be for children; brackets indicate prosodic boundaries). Thus a Verb+Noun sentence was [*la petite*]$_{NP}$ [*ferme*$_{V}$ [*sera*$_{V}$ *pour les enfants*]$_{VP}$ – where listeners first heard [The little girl] [closes..., and then had to revise this interpretation, when only the noun interpretation was consistent with the ending of the sentence (..[*sera pour les enfants*]). In a first experiment, 4-to-6-year-olds and adults listened to these cross-spliced sentences while watching two images side-by-side on a TV-screen: one illustrating the noun interpretation (e.g., a farm) and the other the verb interpretation (e.g., a girl closing something). Both eye-gaze and pointing towards the pictures were recorded. The results showed that adults and children initially looked to the picture supported by the sentence-initial prosody, but only adults switched to look at the ‘correct’ picture upon hearing disambiguating lexical evidence. Children selected the picture consistent with their initial interpretation supported by prosody.

In a second experiment, I asked whether children would revise their initial interpretation (based on prosody) if they first repeated the entire sentence aloud before responding. This also allowed us to assess whether children would repeat the same prosody, or whether they would change the prosody to accommodate the later-arriving lexical information. Children's repetitions revealed that they more often repeated the sentence 'as is', and still chose the initial interpretation provided by prosodic information. In the few cases where they revised, we observed that they tended to correct the prosody of the cross-spliced sentences.

Altogether, these ongoing experiments suggest that both adults and preschoolers rely heavily on the initial information provided by prosody to constrain their parsing, even in a situation where prosodic information is misleading. Contrary to adults, children do not recover from their initial commitment (even though they remember and they can repeat the whole sentence). This behavior replicates the so-called kindergarten-path effect (e.g., Trueswell, Sekerina, Hill, & Logrip, 1999): children give more weight to the information that arrives first, even though that information is prosodic, and supposedly less reliable than lexical content. These results suggest that prosody is an important component of syntactic analysis in children and that it is not treated as an inferior source of information, to be ignored in case of conflict.

9.6. Conclusions

The studies conducted during this thesis show for the first time that infants use phrasal prosody and function words to constrain the acquisition of noun and verb meanings. In our studies, we demonstrated that, despite their reduced vocabulary, infants are able to exploit phrasal prosody to segment spoken sentences into groups of words (i.e., prosodic units). They used these units to infer the presence of syntactic constituent boundaries in a sentence and to identify possible syntactic constituents. Infants exploited this constituent structure, together with function words, to determine the syntactic nature of the prosodic-syntactic constituents, and infer the syntactic category of ambiguous words (Part 1) or of the novel words (Part 2), and therefore constrain their probable meaning. I demonstrated that infants interpreted novel words appearing in noun contexts as referring to objects, and words appearing in verb contexts as referring to actions.

Given the chicken-and-egg problem stated in our introduction, in this thesis we observed that even in the absence of knowledge about the meaning of words, phrasal prosody and function words allowed infants to access the syntactic structures which in turn allowed them to use the “zoom lens” provided by syntax to figure out which part of the world was being talked about and to discover the meaning of novel words. This suggests that as young as 18 months of age, infants already have the means to go from the sounds of language to the syntactic structure of the sentences and therefore that they can use this information to infer the probable meaning of novel words. This powerful mechanism for parsing sentences based on phrasal prosody and function words would be extremely useful for infants to build their way into lexical and syntactic acquisition.

Most of our results were obtained with French-learning children, but we expect that phrasal prosody and function words should support an early access to syntax in many different languages. This hypothesis rests on the fact that although prosodic information and functional elements can surface differently across languages, this information is present in all the world’s languages. The literature until now shows that phrasal prosody and function words are useful sources of information to distinguish between languages, segment the speech stream into words and to constrain lexical access. Our new results bring evidence that this information is also important to constrain syntactic analysis and therefore to discover the meaning of words. Thus, we suggest that phrasal prosody and function words may represent a universal and extremely useful tool for infants to access syntactic information through a surface analysis of the speech stream, and to bootstrap their way toward successful language acquisition.

References

Armstrong, M. E., Andreu, L., Esteve-Gibert, N., & Prieto, P. (2016). Children's processing of morphosyntactic and prosodic cues in overriding context-based hypotheses: an eye tracking study. *Probus*, *28*(1), 57–90. http://doi.org/10.1515/probus-2016-0004

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

Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition*, *114*(3), 442–446. http://doi.org/10.1016/j.cognition.2009.10.015

Bergelson, E., & Swingley, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(9), 3253–8. http://doi.org/10.1073/pnas.1113380109

Bergelson, E., & Swingley, D. (2013). The acquisition of abstract words by young infants. *Cognition*, *127*(3), 391–7. http://doi.org/10.1016/j.cognition.2013.02.011

Bernal, S., Dehaene-Lambertz, G., Millotte, S., & Christophe, A. (2010). Two-year-olds compute syntactic structure on-line. *Developmental Science*, *13*(1), 69–76. http://doi.org/10.1111/j.1467-7687.2009.00865.x

Bernal, S., Lidz, J., Millotte, S., & Christophe, A. (2007). Syntax Constrains the Acquisition of Verb Meaning. *Language Learning and Development*, *3*(4), 325–341. http://doi.org/10.1080/15475440701542609

Bernard, C., & Gervain, J. (2012). Prosodic cues to word order: What level of representation? *Frontiers in Psychology*, *3*(October), 451. http://doi.org/10.3389/fpsyg.2012.00451

Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me: familiar names help launch babies into speech-stream segmentation. *Psychological Science*, *16*(4), 298–304. http://doi.org/10.1111/j.0956-7976.2005.01531.x

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

Brusini, P., Dehaene-Lambertz, G., Dutat, M., Goffinet, F., & Christophe, A. (2016). ERP evidence for on-line syntactic computations in 2-year-olds. *Developmental Cognitive Neuroscience*, *19*, 164–173. http://doi.org/10.1016/j.dcn.2016.02.009

Brusini, P., Dehaene-Lambertz, G., van Heugten, M., de Carvalho, A., Goffinet, F., Fiévet, A. C., & Christophe, A. (2017). Ambiguous function words do not prevent 18-month-olds from building accurate syntactic category expectations: An ERP study. *Neuropsychologia*, *98* (August 2016), 4–12. http://doi.org/10.1016/j.neuropsychologia.2016.08.015

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

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

Chemla, E., Mintz, T. H., Bernal, S., & Christophe, A. (2009). Categorizing words using "frequent frames": What cross-linguistic analyses reveal about distributional acquisition strategies. *Developmental Science*, *12*(3), 396–406. http://doi.org/10.1111/j.1467-7687.2009.00825.x

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

Christodoulopoulos, C., Roth, D., & Fisher, C. (2016). An incremental model of syntactic bootstrapping. In *Proceedings of the Seventh Workshop on Cognitive Aspects of Computational Language Learning*.

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

Christophe, A., Mehler, J., & Sebastián-Gallés, N. (2001). Perception of prosodic boundary correlates by newborn infants. *Infancy*, *2*(3), 385–394. http://doi.org/10.1207/S15327078IN0203_6

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

Christophe, A., Nespor, M., Guasti, M. T., & Van Ooyen, B. (2003). Prosodic structure and syntactic acquisition: the case of the head-direction parameter. *Developmental Science*, *6*(2), 211–220. http://doi.org/10.1111/1467-7687.00273

Christophe, A., Peperkamp, S., Pallier, C., Block, E., & Mehler, J. (2004). Phonological phrase boundaries constrain lexical access I. Adult data. *Journal of Memory and Language*, *51*(4), 523–547. http://doi.org/10.1016/j.jml.2004.07.001

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, *31*(2), 218–236. http://doi.org/10.1016/0749-596X(92)90012-M

de Carvalho, A., Dautriche, I., & Christophe, A. (2016). Preschoolers use phrasal prosody online to constrain syntactic analysis. *Developmental Science*, *19*(2), 235–250. http://doi.org/10.1111/desc.12300

de Carvalho, A., He, X. A., Lidz, J., & Christophe, A. (2015). 18-month-olds use the relationship between prosodic and syntactic structures to constrain the meaning of novel words. In *Paper presented at the 40th Boston University Conference on Language Development*.

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

Decasper, A., & Spence, M. (1986). Prenatal Maternal Speech Influences Newborns' Perception of Speech Sounds. *Infant Behavior and Development*, *9*, 133–150.

Demuth, K., & Tremblay, A. (2008). Prosodically-conditioned variability in children's production of French determiners. *Journal of Child Language*, *35*(1), 99–127. http://doi.org/10.1017/S0305000907008276

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

Dutoit, T. (1997). *An introduction to text-to-speech synthesis 3*. Boston - MA: Kluwer Academic Publishers, Dordrecht.

Echols, C. H. (1993). A perceptually-based model of children's earliest productions. *Cognition*, *46*, 245–296.

Echols, C. H., & Newport, E. (1992). The role of stress and position in determining first words. *Language Acquisition*, *2*, 189–220.

Endress, A. D., & Hauser, M. D. (2010). Word segmentation with universal prosodic cues. *Cognitive Psychology*, *61*(2), 177–199. http://doi.org/10.1016/j.cogpsych.2010.05.001

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

Ferry, A. L., Hespos, S. J., & Waxman, S. R. (2010). Categorization in 3- and 4-month-old infants: An advantage of words over tones. *Child Development*, *81*(2), 472–479. http://doi.org/10.1111/j.1467-8624.2009.01408.x

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

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

Gelman, S. A. (2003). *The essential child: origins of essentialism in everyday thought*. USA: Oxford University Press.

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. http://doi.org/10.1016/j.cogpsych.2007.12.001

Gervain, J., & Werker, J. F. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nature Communications*, *4*, 1490. http://doi.org/10.1038/ncomms2430

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

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

Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*(2), 109–135. http://doi.org/10.1016/S0010-0277(99)00003-7

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

Graham, S. A., Kilbreath, C. S., & Welder, A. N. (2004). Thirteen-month-olds rely on shared labels and shape similarity for inductive inferences. *Child Development*, *75*(2), 409–427. http://doi.org/10.1111/j.1467-8624.2004.00683.x

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

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

Haryu, E., & Kajikawa, S. (2016). Use of bound morphemes (noun particles) in word segmentation by Japanese-learning infants. *Journal of Memory and Language*, *88*, 18–27. http://doi.org/10.1016/j.jml.2015.11.007

Havy, M., & Waxman, S. R. (2016). Naming influences 9-month-olds' identification of discrete categories along a perceptual continuum. *Cognition*, *156*, 41–51. http://doi.org/10.1016/j.cognition.2016.07.011

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

Hawthorne, K., Mazuka, R., & Gerken, L. (2015). The acoustic salience of prosody trumps infants' acquired knowledge of language-specific prosodic patterns. *Journal of Memory and Language*, *82*, 105–117. http://doi.org/10.1016/j.jml.2015.03.005

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

He, A. X., & Arunachalam, S. (2017). Word learning mechanisms. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1–11. http://doi.org/10.1002/wcs.1435

He, A. X., & Lidz, J. (2017). Verb learning in 14- and 18-month-old English-learning infants. *Language Learning and Development*, *5441*(March).

Hirose, Y., & Mazuka, R. (2017). Exploiting Pitch Accent Information in Compound Processing: A Comparison between Adults and 6- to 7-Year-Old Children. *Language Learning and Development*, *5441*(July), 1–20. http://doi.org/10.1080/15475441.2017.1292141

Hirschberg, J., & Pierrehumbert, J. (1986). The intonational structuring of discourse. In *Proceedings of the 24th Meeting of the Association for Computational Linguistics* (Vol. 1, pp. 136–144). http://doi.org/10.1017/CBO9781107415324.004

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

Hochmann, J. R. (2013). Word frequency, function words and the second gavagai problem. *Cognition*, *128*(1), 13–25. http://doi.org/10.1016/j.cognition.2013.02.014

Hochmann, J. R., Endress, A. D., & Mehler, J. (2010). Word frequency as a cue for identifying function words in infancy. *Cognition*, *115*(3), 444–457. http://doi.org/10.1016/j.cognition.2010.03.006

Höhle, B., Berger, F., Müller, A., Schmitz, M., & Weissenborn, J. (2009). Focus Particles in Children’s Language: Production and Comprehension of *Auch* “Also” in German Learners from 1 Year to 4 Years of Age. *Language Acquisition*, *16*(1), 36–66. http://doi.org/10.1080/10489220802584550

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

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

Höhle, B., Weissenborn, J., Schmitz, M., & Ischebeck, A. (2001). Discovering Word Order Regularities: The Role of Prosodic Information for Early Parameter Setting. In J. Weissenborn & B. Höhle (Eds.), *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition* (pp. 246–263). Amsterdam: Benjamins.

Ito, K., Jincho, N., Minai, U., Yamane, N., & Mazuka, R. (2012). Intonation facilitates contrast resolution: Evidence from Japanese adults and 6-year olds. *Journal of Memory and Language*, *66*(1), 265–284. http://doi.org/10.1016/j.jml.2011.09.002

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

Johnson, E. K., & Seidl, A. (2008). Clause segmentation by 6-month-old infants: A crosslinguistic perspective. *Infancy*, *13*(5), 440–455. http://doi.org/10.1080/15250000802329321

Johnson, M., Christophe, A., Demuth, K., & Dupoux, E. (2014). Modelling function words improves unsupervised word segmentation. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, 282–292.

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

Jun, S. (2014). Prosodic typology: by prominence type, word prosody, and macro-rhythm. In

S. Jun (Ed.), *Prosodic Typology II: The Phonology of Intonation and Phrasing*. Oxford University Press.

Jusczyk, P. W. (1997). *The Discovery of Spoken Language*. Cambridge, MA: MIT Press.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in english-learning infants. *Cognitive Psychology*, *39*(3–4), 159–207. http://doi.org/10.1006/cogp.1999.0716

Jusczyk, P. W., & Krumhansl, C. (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology. Human Perception and Performance*, *19*(3), 627–640.

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), 325–38. http://doi.org/10.1111/j.1467-8624.2006.00873.x

Kedar, Y., Casasola, M., Lust, B., & Parmet, Y. (2017). Little Words , Big Impact : Determiners Begin to Bootstrap Reference by 12 Months. *Language Learning and Development*, *0*(0), 1–18. http://doi.org/10.1080/15475441.2017.1283229

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

Knosche, T. R., Neuhaus, C., Haueisen, J., Alter, K., Maess, B., Witte, O. W., & Friederici, A. D. (2005). Perception of Phrase Structure in Music. *Humain Brain Mapping*, *24*, 259–273. http://doi.org/10.1002/hbm.20088

Kraljic, T., & Brennan, S. E. (2005). Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? *Cognitive Psychology*, *50*(2), 194–231. http://doi.org/10.1016/j.cogpsych.2004.08.002

Kraljic, T., Samuel, A. G., & Brennan, S. E. (2008). First impressions and last resorts: how listeners adjust to speaker variability. *Psychological Science*, *19*(4), 332–8. http://doi.org/10.1111/j.1467-9280.2008.02090.x

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Dartmouth Publishing Group.

Ladd, R. (2008). *Intonational Phonology*. *Lingua* (2nd ed., Vol. 104). New York: Cambridge University Press. http://doi.org/10.1016/S0024-3841(98)00007-2

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

Lew-Williams, C., Pelucchi, B., & Saffran, J. R. (2011). Isolated words enhance statistical language learning in infancy. *Developmental Science*, *14*(6), 1323–1329. http://doi.org/10.1111/j.1467-7687.2011.01079.x

Ludusan, B., Cristia, A., Martin, A., Mazuka, R., & Dupoux, E. (2016). Learnability of prosodic boundaries: Is infant-directed speech easier? *The Journal of the Acoustical Society of America*, *140*(2), 1239–1250. http://doi.org/10.1121/1.4960576

Ludusan, B., & Dupoux, E. (2016). The role of prosodic boundaries in word discovery: Evidence from a computational model. *The Journal of the Acoustical Society of America*, *140*(1), EL1-EL6. http://doi.org/10.1121/1.4954652

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

Männel, C., & Friederici, A. D. (2009). Pauses and intonational phrasing: ERP studies in 5-month-old German infants and adults. *Journal of Cognitive Neuroscience*, *21*(10), 1988–2006. http://doi.org/10.1162/jocn.2009.21221

Marchetto, E., & Bonatti, L. L. (2013). Words and possible words in early language acquisition. *Cognitive Psychology*, *67*(3), 130–150. http://doi.org/10.1016/j.cogpsych.2013.08.001

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

Massicotte-Laforge, S., & Shi, R. (2016). Infants use prosody for syntactic analysis and grammatical categorization. In *Paper presented at the 41st Boston University Conference on Language Development*. Boston - MA.

Mazuka, R., & Tanaka, Y. (2006). Children can be better than adults at using prosody to resolve syntactic ambiguity. In *Poster presented at the 19th Annual CUNY Conference on Human Sentence Processing*. New York, NY.

Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences*, *108*(22), 9014–9019. http://doi.org/10.1073/pnas.1105040108

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

Michelas, A., & D'Imperio, M. (2012). When syntax meets prosody: Tonal and duration variability in French Accentual Phrases. *Journal of Phonetics*, *40*(6), 816–829.

http://doi.org/10.1016/j.wocn.2012.08.004

Michelas, A., & D'Imperio, M. (2015). Prosodic boundary strength guides syntactic parsing of French utterances. *Laboratory Phonology*, *6*(1), 119–146. http://doi.org/10.1515/lp-2015-0003

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

Millotte, S., Wales, R., & Christophe, A. (2007). Phrasal prosody disambiguates syntax. *Language and Cognitive Processes*, *22*(6), 898–909. http://doi.org/10.1080/01690960701205286

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

Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech, *90*, 91–117. http://doi.org/10.1016/S0010-0277(03)00140-9

Mintz, T. H., Newport, E. L., & Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children, *26*, 393–424.

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

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

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, *17*(2), 357–74.

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

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

Nespor, M., Shukla, M., van de Vijver, R., Avesani, C., Schraudolf, H., & Donati, C. (2008). Different phrasal prominence realization in VO and OV languages. *Lingue E Linguaggio*, *2*, 1–29.

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

Newman, G. E., Keil, F. C., Kuhlmeier, V. A., & Wynn, K. (2010). Early understandings of the link between agents and order. *Proceedings of the National Academy of Sciences*, *107*(40), 17140–17145. http://doi.org/10.1073/pnas.0914056107

Ohima-Takane, Y., Ariyama, J., Kobayashi, T., Katerelos, M., & Poulin-Dubois, D. (2011). Early verb learning in 20-month-old Japanese-speaking children. *Journal of Child Language*, *38*(3), 455–484. http://doi.org/10.1017/S0305000910000127

Opfer, J. E., & Gelman, S. A. (2010). Development of the animate-inanimate distinction. In U. Goswami (Ed.), *Wiley-Blackwell Handbook of Childhood Cognitive Development* (pp. 213–238). Wiley.

Parise, E., & Csibra, G. (2012). Electrophysiological Evidence for the Understanding of Maternal Speech by 9-Month-Old Infants. *Psychological Science*, *23*(7), 728–733. http://doi.org/10.1177/0956797612438734

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing Syntactic Relations in Language and Music: An Event-Related Potential Study. *Journal of Cognitive Neuroscience*, *10*(6), 717–733. http://doi.org/10.1162/089892998563121

Pierrehumbert, J. B., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. *Intentions in Communication*, 271–311.

Quine, W. V. O. (1960). *Word and object*. Cambridge MA: MIT Press.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax: the nature of early child grammars of English*. Oxford: Wiley-Blackwell.

Redington, M. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science*, *22*(4), 425–469. http://doi.org/10.1016/S0364-0213(99)80046-9

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical Learning by 8-Month-Old Infants. *Science*, *274*(5294), 1926–1928.

Saxe, R., & Carey, S. (2006). The perception of causality in infancy. *Acta Psychologica*, *123*(1–2), 144–165. http://doi.org/10.1016/j.actpsy.2006.05.005

Seidl, A. (2007). Infants' use and weighting of prosodic cues in clause segmentation. *Journal of Memory and Language*, *57*(1), 24–48. http://doi.org/10.1016/j.jml.2006.10.004

Shafer, V., Shucard, D., Shucard, J., & Gerken, L. (1998). An electrophysiological study of infants' sensitivity to the sound patterns of English speech. *Journal of Speech, Language, and Hearing Research*, *41*, 874–886.

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. http://doi.org/10.1111/cdep.12052

Shi, R., Cutler, A., Werker, J. F., & Cruickshank, M. (2006). Frequency and form as determinants of functor sensitivity in English-acquiring infants. *The Journal of the Acoustical Society of America*, *119*(6), EL61-L67. http://doi.org/10.1121/1.2198947

Shi, R., & Lepage, M. (2008). The effect of functional morphemes on word segmentation in preverbal infants: Paper. *Developmental Science*, *11*(3), 407–413. http://doi.org/10.1111/j.1467-7687.2008.00685.x

Shi, R., & Melançon, A. (2010). Syntactic Categorization in French-Learning Infants. *Infancy*, *15*(5), 517–533. http://doi.org/10.1111/j.1532-7078.2009.00022.x

Shi, R., & Werker, J. F. (2001). Six-Month-Old Infants' Preference for Lexical Words. *Psychological Science*, *12*(1), 70–75. http://doi.org/10.1111/1467-9280.00312

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

Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, *72*(2), B11–B21. http://doi.org/10.1016/S0010-0277(99)00047-5

Shukla, M., Nespor, M., & Mehler, J. (2007). An interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology*, *54*(1), 1–32. http://doi.org/10.1016/j.cogpsych.2006.04.002

Shukla, M., White, K. S., & Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences*, *108*(15), 6038–6043. http://doi.org/10.1073/pnas.1017617108

Snedeker, J., & Trueswell, J. (2001). Unheeded cues: Prosody and syntactic ambiguity in mother-child communication. In *Paper presented at the 26th Boston University Conference on Language Development*.

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

Soderstrom, M., Seidl, A., Kemler Nelson, D. G., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language*, *49*(2), 249–267. http://doi.org/10.1016/S0749-596X(03)00024-X

St. Clair, M. C., Monaghan, P., & Christiansen, M. H. (2010). Learning grammatical categories from distributional cues: Flexible frames for language acquisition. *Cognition*, *116*(3), 341–360. http://doi.org/10.1016/j.cognition.2010.05.012

Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*(2), 172–175.

Tincoff, R., & Jusczyk, P. W. (2012). Six-Month-Olds Comprehend Words That Refer to Parts of the Body. *Infancy*, *17*(4), 432–444. http://doi.org/10.1111/j.1532-7078.2011.00084.x

Tomasello, M. (2000a). Do young children have adult syntactic competence? *Cognition*, *74*(3), 209–253. http://doi.org/10.1016/S0010-0277(99)00069-4

Tomasello, M. (2000b). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences*, *4*(4), 156–163. http://doi.org/10.1016/S1364-6613(00)01462-5

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

van Heugten, M., & Johnson, E. K. (2010). Linking infants' distributional learning abilities to natural language acquisition. *Journal of Memory and Language*, *63*(2), 197–209. http://doi.org/10.1016/j.jml.2010.04.001

Van Heugten, M., & Johnson, E. K. (2011). Gender-marked determiners help Dutch learners ' word recognition when gender information itself does not *, *38*(August 2006), 87–100. http://doi.org/10.1017/S0305000909990146

van Ommen, S., Boll-Avetisyan, N., Larraza, S., Wellmann, C., Bijeljac-Babic, R., & Höhle, B. (2017). Language-specific processing of prosodic boundaries: cross-linguistic evidence from German- and French-learning infants. In *Paper presented at the Journée du GDR Neurosciences Cognitives du Développement*. Paris, France: June 27th.

Vogel, I., & Raimy, E. (2002). The acquisition of compound vs. phrasal stress: the role of prosodic constituents. *Journal of Child Language*, *29*(2), 225–50.

Warner, N., Otake, T., & Arai, T. (2010). Intonational structure as a word-boundary cue in Tokyo Japanese. *Language and Speech*, *53*(Pt 1), 107–31. http://doi.org/10.1177/0023830909351235

Waxman, S., & Lidz, J. (2006). Early word learning. In D. Kuhn & R. Siegler (Eds.), *Handbook of Child Psychology* (6th Editio, Vol. 2, pp. 299–335). Hoboken NJ: Wiley.

Waxman, S., & Markow, D. B. (1995). Words as Invitations to Form Categories: Evidence from 12- to 13-Month-Old Infants. *Cognitive Psychology*, *29*(3), 257–302. http://doi.org/10.1006/cogp.1995.1016

Waxman, S. R., & Booth, A. E. (2001). Seeing Pink Elephants: Fourteen-Month-Olds' Interpretations of Novel Nouns and Adjectives. *Cognitive Psychology*, *43*(3), 217–242. http://doi.org/10.1006/cogp.2001.0764

Waxman, S. R., & Leddon, E. (2011). Early word learning and conceptual development: Everything had a name, and each name gave birth to a new thought. In U. Goswami (Ed.), *The Wiley-Blackwell Handbook of Childhood Cognitive Development* (pp. 180–208). Malden, MA: Wiley-Blackwell.

Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology*, *59*(1), 67–95. http://doi.org/10.1016/j.cogpsych.2009.02.001

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), B63-72. http://doi.org/10.1016/j.cognition.2005.07.001

Wellmann, C., Holzgrefe, J., Truckenbrodt, H., Wartenburger, I., & Höhle, B. (2012). How each prosodic boundary cue matters: Evidence from German infants. *Frontiers in Psychology*, *3*(DEC), 1–13. http://doi.org/10.3389/fpsyg.2012.00580

Woodward, A. L., Sommerville, J. a, & Guajardo, J. J. (2001). How infants make sense of intentional action. *Intentions and Intentionality: Foundations of Social Cognition.*, (January 2001), 149–169.

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. http://doi.org/10.1111/j.1467-9280.2009.02341.x

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

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.

Zhou, P., Crain, S., & Zhan, L. (2012). Sometimes children are as good as adults: The pragmatic use of prosody in children's on-line sentence processing. *Journal of Memory*

and Language, *67*(1), 149–164. http://doi.org/10.1016/j.jml.2012.03.005
Zhou, P., Su, Y. E., Crain, S., Gao, L., & Zhan, L. (2012). Children's use of phonological information in ambiguity resolution: a view from Mandarin Chinese. *Journal of Child Language*, *39*(4), 687–730. http://doi.org/10.1017/S0305000911000249

Annexes

The following two papers contain additional work that I have been doing during my PhD. They appear in Annexes because they are not directly related to the content of the current thesis. In de Carvalho, et al. (2016, *Frontiers in Psychology*), I present a study that I did with Anne Reboul and other colleagues from Lyon, in which we studied the impact of pragmatics in the interpretation of word meanings in Adults. In Brusini et al., (2017), we present a study showing that ambiguous function words (i.e., “la” in French can be either an article or a pronoun) do not prevent 18-month-olds from building accurate syntactic category expectations. This collaboration with Brusini and colleagues allowed me to be trained on the EEG-technique and to learn how to test young infants in EEG experiments.

frontiers in Psychology

ORIGINAL RESEARCH
published: 25 October 2016
doi: 10.3389/fpsyg.2016.01500

A. Scalar Implicatures: the Psychological reality of scales

Scalar Implicatures: The Psychological Reality of Scales

Alex de Carvalho[1,2*†], Anne C. Reboul[1*†], Jean-Baptiste Van der Henst[1], Anne Cheylus[1] and Tatjana Nazir[1]

[1] CNRS, Institute for Cognitive Sciences-Marc Jeannerod (UMR 5304), University Claude Bernard Lyon 1, Bron, France,
[2] Laboratoire de Sciences Cognitives et Psycholinguistique (ENS, EHESS, CNRS), Departement d'Etudes Cognitives, École Normale Supérieure, PSL Research University, Paris, France

Scalar implicatures, the phenomena where a sentence like "The pianist played *some* Mozart sonatas" is interpreted, as "The pianist did *not* play *all* Mozart sonatas" have been given two different analyses. Neo-Griceans (NG) claim that this interpretation is based on lexical scales (e.g., <*some, all*>), where the stronger term (e.g., *all*) implies the weaker term (e.g., *some*), but the weaker term (e.g., *some*) implicates the negation of the stronger term (i.e., *some* = *not all*). Post-Griceans (PG) deny that this is the case and offer a context-based inferential account for scalar implicatures. While scalar implicatures have been extensively investigated, with results apparently in favor of PG accounts, the psychological reality of lexical scales has not been put to the test. This is what we have done in the present experiment, with a lexical decision task using lexical scales in a masked priming paradigm. While PG accounts do not attribute any role for lexical scales in the computation of scalar implicatures, NG accounts suggest that lexical scales are the core mechanism behind the computation of scalar implicatures, and predict that weaker terms in a scale should prime stronger terms more than the reverse because stronger words are necessary to the interpretation of weaker words, while stronger words can be interpreted independently of weaker words. Our results provided evidence in favor of the psychological existence of scales, leading to the first clear experimental support for the NG account.

Keywords: lexical scales, masked priming, lexical decision task, scalar implicature, implication, experimental pragmatics, psycholinguistics

OPEN ACCESS

Edited by:
Andrea Moro,
Istituto Universitario di Studi Superiori di Pavia, Italy

Reviewed by:
Robert Fiorentino,
University of Kansas, USA
Vincenzo Moscati,
University of Siena, Italy

****Correspondence:***
Anne C. Reboul
reboul@isc.cnrs.fr
Alex de Carvalho
x.de.carvalho@gmail.com

†These authors have contributed equally to this work and shared joint first-authorship.

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *06 April 2016*
Accepted: *20 September 2016*
Published: *25 October 2016*

Citation:
de Carvalho A, Reboul AC, Van der Henst J-B, Cheylus A and Nazir T (2016) Scalar Implicatures: The Psychological Reality of Scales. Front. Psychol. 7:1500. doi: 10.3389/fpsyg.2016.01500

INTRODUCTION

The notion of *implicature* was introduced by Grice (1975) to account for information that was *communicated* without being, strictly speaking, *said* by the speaker, in other words, for information that was *implicitly* rather than *explicitly* communicated (Grice, 1989). For instance, if the speaker asked where Anne lives, an answer such as "Somewhere in Burgundy, I believe," *conversationally implicates* that he does not know exactly where she lives.

Grice distinguished among conversational implicatures those that (as in the previous example) strongly depend on the **context** (the Particularized Conversational Implicatures: PCIs) from those that depend on the **words** used (the Generalized Conversational Implicatures: GCIs). The paramount examples of GCIs are so-called *scalar implicatures*. For instance, if the speaker says, "The pianist played *some* Beethoven sonatas," she implicates, through the use of *some*, that the pianist did *not* play *all* of them. Note that both GCIs and PCIs are computed through the same mechanism: they are the result of an inference that was made by comparing what the speaker says with what she might have said but did not say. In other words, these inferences are based on alternatives to what was said. A sentence such as "Anne lives somewhere in Burgundy, I believe," leads us to derive a PCI, i.e., the speaker does not know where exactly Anne lives, because we quickly infer that if he knew it, he would have said where she lives precisely. Through the same inferential mechanism a GCI such as "The pianist played *some* Beethoven sonatas," implicates that he did not play all of them, because if he had played all of them (on the assumption that the speaker knew it), the speaker would have used *all* rather than *some*. However, despite the fact that they share the same inferential mechanism, PCIs and GCIs differ in the ways in which "the alternatives" are determined: through the **context** for PCIs and though the **lexicon** for GCIs.

Neo-Gricean and Post-Gricean Accounts

The notion of implicature was quickly incorporated from philosophy of language into pragmatics, but led to two highly different approaches: The Neo-Gricean (NG) and the Post-Gricean (PG) approaches. The NG approaches (see e.g., Horn, 1972, 2004; Levinson, 2000; Chierchia, 2004) claimed, on the basis of scalar implicatures, that GCIs are derived locally and automatically (by default) when the trigger belongs to a linguistic *scale*. Such lexical scales are ordered sets of terms, such as *<and, or>*, where the stronger member, *and*, implies the weaker member, *or*, and the weaker member, *or*, implicates the negation of the stronger member: *p or q* implicates *not (p and q)* (i.e., one or the other but not both). The implicature interpretation can be canceled in favor of a semantic interpretation (*p or q and possibly both*), but this will only come after the pragmatic interpretation has been accessed and at a cost.

The NG approach is a more or less straightforward extension of Gricean theory, in the sense that it considers scalar implicatures to be conversational implicatures, and that it is a development of Grice's intuition that some conversational implicatures are entirely dependent on the context while others are not (and scalar implicatures are a major example of the latter kind). Scalar implicatures—and this is the new development brought about by Horn (1972)—depends on the existence of lexical scales. Horn proposed that, in the case of scalar implicatures, the alternatives we compare to what the speaker said are determined by the lexical scale to which the term that triggers the inference belongs. In other words, when a weaker term of a scale (e.g., "some") is used in a sentence, a comparison is made with the stronger term(s) in the scale (e.g., "all") as alternatives to what the speaker said (e.g., "If she used *some*, it's because it's not *all*"). In a more recent development of this position, Levinson (2000) went one-step further and proposed that the pragmatic interpretation of the scalar term is *lexicalized as its default interpretation*. In other terms, the pragmatic interpretation of scalar items is encoded as a (defeasible) part of its meaning (i.e., "some" also means "not all"), while the semantic interpretation (i.e., "some = at least one") would only be accessible if the pragmatic interpretation is explicitly negated (e.g., "The pianist played some Mozart sonatas and even all of them").

By contrast, PG accounts, such as Relevance Theory (see e.g., Sperber and Wilson, 1986/1995; Carston, 2002), consider scalar implicatures to be explicatures rather than any kind of conversational implicatures. They result from a process of pragmatic enrichment of the linguistic interpretation of the utterance (the so-called logical form), yielding a relevant truth-conditional propositional form. This enrichment process is on a par with what happens for most utterances (e.g., loose talk, metaphors, etc). For instance, a sentence such as "This steak is raw," uttered in restaurant, is usually interpreted as *This steak is undercooked*. This final interpretation is obtained through a contextually driven process of *ad hoc* concept construction (loosening or strengthening) applying locally to the concept raw[1]. In such cases, the *ad hoc* concept construction is not, in any sense, a lexically based process: it is a contextually driven non-linguistic, conceptual process. The claim that scalar implicatures are interpreted through an identical process of *ad hoc* concept construction excludes both any Gricean-style mechanism based on alternatives and any role for the lexical scales as proposed by NG approaches. Additionally, *ad hoc* concept construction is believed to be a cognitively costly process, which implies that scalar implicatures will come at a price and will be accessed only when the context makes them relevant (see Noveck and Sperber, 2007 for a discussion). Thus, the PG approach differs from NG approaches in that it gives a central place to context and sees scalar inferences as the result of a contextual process, not allowing any role to lexical scales.

Previous experimental work on scalar inferences has concentrated on the opposite predictions drawn from the two accounts regarding *processing cost*. According to the NG account, the pragmatic interpretation is less costly than the semantic interpretation. On the PG accoun, the semantic interpretation is less costly than the pragmatic interpretation. Cashing the notion of *cost* in terms of cognitive difficulty, this suggests that the most costly interpretation should come later in cognitive development and that it should take more time to be processed. Thus, NG predicts that the *semantic* interpretation should come later and take more time, while PG predicts that it is the *pragmatic* interpretation that should come later and take more time. Studies that contrasted NG and PG accounts in terms of processing cost have provided robust evidence in favor of the PG account, because there is a clear progression of pragmatic interpretations from the younger age to adults (Noveck, 2001; see also Papafragou and Musolino, 2003; Papafragou and Tantalou, 2004; Guasti et al., 2005; Pouscoulous et al., 2007) and reaction time (RT) measures in adults show that pragmatic interpretations of scalar terms take longer to access than semantic interpretations (Bott and Noveck, 2004; Bott et al., 2012 but see for different results Feeney et al., 2004). Moreover, the proportion of pragmatic answers observed with adults was strongly context-dependent (see also Hartshorne et al., 2015 and Dupuy et al., 2016, for more data on the strong context-sensitivity of pragmatic interpretations for scalar implicatures). This context-dependency contradicts Levinson's default account, which implies that all underinformative sentences with scalar terms should be given

[1] As is usual, small capitals are used here for concepts, not for words.

pragmatic interpretations and that semantic interpretations should only be given when the implicature is explicitly negated.

Thus, all the experimental results up to now strongly favor PG accounts and starkly contradict the predictions of NG accounts. There is nevertheless a crucial and interesting element in NG that has not been empirically investigated: the psychological reality of lexical scales.

Current State of the Debate

While the simple lexical default account proposed by Levinson (2000) has been definitely contradicted by the experimental evidence, a new and more sophisticated NG account has recently been proposed by Chierchia (2013) and has not yet been tested. Chierchia proposes a far-ranging theory, encompassing not only scalars, but free-choice implicatures, polarity items, as well as upward and downward entailing linguistic environments. Regarding scalar implicatures, Cherchia argues that they result from a covert exhaustification operator (roughly equivalent in meaning to *only*) that operates on a set of alternatives determined by the scale the scalar term belongs to. However, this set of alternatives is only available to the exhaustification process if the context makes it mandatory to derive the implicature. For instance, if, in answer to the question "Did the pianist play *all* Mozart sonatas?", the speaker hearer answers "He played *some* Mozart sonatas," the alternative set including *most* and *all* will be available, while if the question had been "Did the pianist play Mozart sonatas?", it would not be. Chierchia (2013, p. 104) notes that relevance to conversational goals is the central contextual factor in the derivation of scalar inferences.

On this new version of NG, quite a few of the differences with PG disappear: Chierchia does not commit himself about the cost of the implicature. He acknowledges a major role of the context, including what he calls "conversational relevance," which determines whether or not the scalar inference will be drawn. However, in Chierchia's theory, the alternatives are entirely due to Horn scales (e.g., <*all, many, some*>, which are lexically determined). It is the psychological reality of such scales that we are interested in testing in the present study.

There is no question that words inside a scale usually form a 'family' in the sense that they have related meanings (e.g., <*all, many, some*>) are all quantifiers. On this, both NG and PG would agree, but there is more to a scale than words with related meanings. In the NG account, the stronger words in a scale (i.e, "all") are necessary for the interpretation of the weaker words (i.e., "some") whenever an implicature is derived (they yield the alternative set: e.g., 'some and maybe all'), while the stronger words can be interpreted without recourse to the weaker words in all circumstances (e.g., 'all' is always all, not less). These two characteristics of scales —that words inside a scale are related, and that there is an interpretive *asymmetry* due to the fact that stronger words are necessary to the interpretation of the weaker words, but the reverse is not true— open a road for behavioral investigations, using a masked priming paradigm (Forster and Davis, 1984).

As scales are supposed to be recovered automatically from the lexicon in NG (the context makes them available or not to the exhaustification mechanism), the simple and automatic nature of masked priming in a lexical decision task seems particularly appropriate to test the question of the psychological reality of Horn scales. Given that one form of priming is semantic in nature (i.e., words belonging to the same semantic fields prime one another more strongly than they prime words from other semantic fields (Perea and Rosa, 2002), we expect that words belonging to the same scale should prime one another. Crucially, as scales are ordered sets of words and given the NG notion that the stronger words are used in the interpretation of the weaker words, while the stronger words can be interpreted regardless of the weaker words, there should be an *asymmetry* in priming: weaker words in a scale should prime stronger words in the same scale more than stronger words would prime weaker words. For instance, in the scale <*all, many, some*>, *some* should prime *many* and *all* more than *many* would prime *some* and more than *all* would prime *many* and *some*.

By contrast, given that PG does not give lexical scales any role in the construction of the *ad hoc* concepts that it sees as the core of scalars, at most it would predict that, as any set of semantically related words, words inside a scale would prime one another more strongly than they would prime other words. However, it would not predict any asymmetry in the strength of priming between weaker and stronger words.

EXPERIMENT: LEXICAL DECISION TASK WITH MASKED PRIMING

In order to test the asymmetry prediction, a lexical decision task with masked priming was conducted. The masked priming paradigm (e.g., Forster and Davis, 1984) consists in presenting a subliminal prime to facilitate the processing of a target word. Note that priming is the phenomenon by which the presentation of a first item (the prime) will influence the processing of a second item (the target). In masked priming, the prime is presented subliminally, that is, too quickly for the participant to be aware that it was presented. These priming paradigms with a simple lexical decision task (where participants have to decide whether the target is a word or a non-word, after they have been presented with another word subliminally) give us a good opportunity to test the psychological reality of scales. Hence, this is a simple experimental paradigm that does not depend on any kind of reasoning and that is largely automatic given that the prime is not consciously perceived (Dehaene et al., 1998).

In particular, participants were presented with a subliminal prime word followed by the target and asked to judge whether the target was a word or a non-word. The measure was the RT between the presentation of the target and the participant's answer. The task included two experimental conditions: in one condition, the prime was a weaker term than the target on the informativity scale (Implicature condition: e.g., *SOME — all*); in the other condition the prime was a stronger member than the target (Implication condition: e.g., *ALL — some*). Additionally, two control conditions were designed: one in which the prime and the target were identical (Identical condition: e.g., *SOME — some*); and one in which the prime was a sequence of consonants of the same length (in terms of number of letters) as the target (Consonant condition: e.g., *ZSQW — some*).

The identical condition should yield the shortest average RT because a term maximally primes itself. The consonant condition should have the longest RT response, because there cannot be any priming effect at all in this condition. Thus, these two control conditions should allow us to verify that the experiment worked well and to have a control on whether or not the RT of the participants is the result of the simple processing and reading of the target stimuli. Regarding the experimental conditions, the NG account (which supposes the psychological reality of scales) predicts that the target should be evaluated faster in the implicature condition (e.g., SOME – all) than in the implication condition (e.g., ALL – some).

MATERIALS AND METHODS

Participants

Participants were 48 French native speakers, graduate students from the Ecole Normale Supérieure in Lyon, aged 20–30, right-handed, with normal or corrected-to-normal vision (20 males, mean age 22.4; 28 females, mean age 21.4). They participated on a voluntary basis, with no financial compensation. Five additional participants were tested but their data were not included in our analysis because they were ambidextrous (3), or because they made more than 30 errors (10%) during the test (2).

Design and Stimuli

The experimental material was built on the basis of 129 items: 43 scalar terms, 43 pseudo-words and 43 sequences of consonants. We tested 18 scales: 11 included two words (e.g., *<and, or>*) and 7 three words (e.g., *<some, most, all>*) (c.f., Data Sheet 1 for a complete list of scales). Middle words from the three-word scales were used for both the implication and implicature conditions. The scales we tested were chosen among those mentioned in the NG literature (e.g., Levinson, 2000; Horn, 2004). Given that our purpose was to test the general hypothesis that priming effect would be stronger in the implicature than in the implication condition, we took scales from various syntactic categories, connectives, quantifiers, adverbs, verbs, and adjectives, and scales composed of two or three words, without assuming any particular difference between them. This choice was motivated by the fact that we did not have any specific hypothesis on whether these different categories would trigger stronger or weaker effects of priming or on whether the number of lexical items in the scale (two or three) would modulate the priming effect. French words belonging to scales were used and were controlled for length and frequency of word, letters, bigram and trigram with the LEXIQUE database (New et al., 2004). The pseudo-words were created with an application from the Lexique Toolbox, which is a generator of pseudo-words from the same database. The pseudo-words were controlled for length and bigram frequency. Crucially, note that the frequency of the target words used had a similar range between the two experimental conditions (implication condition: mean = 1.81, median = 1.88, *SD* = 1.10, range: from −0.47 to 4.13; implicature condition: mean = 2.18; median = 1.98, *SD* = 1.24, range: from 0.49 to 4.32).

Each target word in the scales was either primed with itself, its matching consonant or the other word(s) in the scale, resulting in 150 prime-target stimuli (11*2*3 + 7*3*4) and 150 matching pseudo-words conditions for a total of 300 trials. Thus, each word was seen by each subject, as a prime and as a target in the identical condition, as a prime or as target in the implicature and implication conditions, as a prime for the pseudo-word condition and as a target in the consonant condition (where the sequence of consonants was used as a prime). For a better understanding of the way that the words were assigned to the different conditions presented to participants, Data Sheet 1 provides a table of stimuli showing for each target word, the words presented as a prime in each condition.

The entire list of stimuli were presented in a fully within-subjects design, such that all subjects saw exactly the same stimuli in each condition in a different randomized order, for each subject.

Procedure

The experiment was implemented with Neurobehavioural Systems, Inc. Presentation® 14.9 program. The experiment took place on an individual basis in a quiet experimental room. Each trial started with a fixation point presented in the center of the screen for 500 ms. Then a forward mask (######) was presented for 34 ms and was immediately followed by a uppercase prime presented for 34 ms. The prime was replaced by another mask (######) for 34 ms before the target appeared on the screen. Participants were instructed to press one of two pre-defined buttons on the keyboard (the ‘right’ and the ‘left’ key buttons) to indicate whether the lower case letter string was a French word or not. For half of the participants the ‘right’ key corresponded to the ‘yes’ response and for the other half to the ‘no’ response. The target remained on the screen until participant’s response (see **Figure 1**). The lexical decision had to be performed as rapidly and as accurately as possible. The dependent variables were the RTs and error rates. When the participant responded, the target disappeared from the screen. The inter-trial interval was 1500 ms. Participants were not informed of the presence of the prime and in a debriefing after the experiment, none of them have reported detecting the prime words during the experiment.

Data Analyses

Statistical data analysis and graphics were produced with R software version 3.2.2 (R Core Team, 2015) with packages multcomp (Hothorn et al., 2008) and lme4 (Bates et al., 2015). The response time analysis included only correct answers (per subject average 97.83%, median = 98.00%; range: from 93.00 to 99.67%). RTs below 300 ms and above 2000 ms were automatically excluded from the analysis because we assume that responses longer than 2000 ms reflect distraction rather than lexical decision and responses below 300 ms reflect anticipatory responses prior to proper stimulus processing. For the remaining trials, RTs outside of the interval defined by the intra-subject average ±2.5 standard deviation were discarded to minimize the impact of outliers on mean RT. Using these procedures, 5.58%

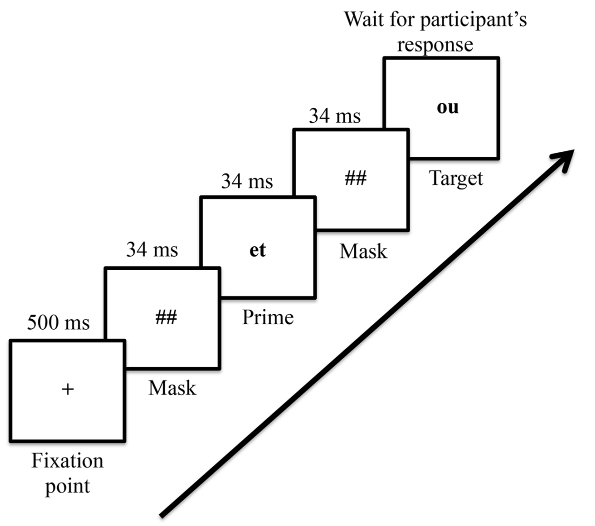


FIGURE 1 | Time course for a single trial. Each trial started with a fixation point in the middle of the screen presented during 500 ms. Then a mask, a prime and a mask appeared for 34 ms each, immediately followed by the target word that appeared in the screen until the participant answered if it was a word or a pseudo-word. Note that for each trial, the mask contained exactly the same numbers of characters of the word used as prime.

of the initial data were discarded from the final analysis. RTs were then averaged for each participant in each of the different conditions prior to the calculation of the grand average over all participants.

RESULTS

Figure 2 shows the averaged RTs for each condition and the standard error across all participants. Average RTs for the target words presented in the identical condition were faster than average RTs for the same targets in the consonant condition. Average RTs for the two experimental conditions were in-between these two control conditions with higher averaged RTs in the implication condition than in the implicature condition.

Note, however, that the target words in the implicature condition (e.g., *SOME* — ***all***) and implication conditions (e.g., *ALL* — ***some***) are not the same. Potential differences in RTs between these two conditions could therefore be related to differences in the default reading time of the target words themselves. We therefore used a linear mixed effect model to analyze our data, with condition as a fixed effect and target word and participants as random effects. Confidence intervals for Tukey contrasts estimated with this model and a 95% family wise confidence level are shown in **Figure 3**. Tests that these contrasts are null based on the model are reported in **Table 1**.

Figure 3 and **Table 1** indicate an estimated 10.12 ms reduction of response time in the implicature condition compared with the implication condition. Confidence intervals based on a 95% family wise confidence level slightly overlaps with zero and the single-step p-value adjustment indicates that this effect is not significant ($p = 0.16$). Note, however, that we are only interested in the single contrast between Implication and Implicature conditions. Correction for multiple testing is therefore not required. The uncorrected p-value is significant ($p = 0.04$). The data can thus be interpreted in favor of the existence of scales.

Following reviewers' suggestions, we also looked at the data scale by scale. As our data does not allow us to conduct statistical analysis (since for each subject we had only one data point of RT for each scale in each condition), this analysis is presented as "Presentation 1" to which the interested reader is directed.

DISCUSSION

As reviewed in our introduction, all the experimental literature has favored PG over NG accounts of how scalar implicatures are derived. However, one issue that has not been experimentally investigated so far is the psychological reality of lexical scales, a central issue for NG. Additionally, given recent developments in NG accounts (see Chierchia, 2013), the existence of scales has become the main point of departure between NG and PG, or at least one that is open to behavioral measures.

Using a masked priming paradigm, we tested the differential predictions of the two accounts. Predictions, based on the NG account, were that RTs in the implicature condition would be faster than in the implication condition because weaker words of a scale should prime stronger words of the same scale more than stronger words prime weaker words. By contrast, following the PG account, one would expect words inside a scale to prime one another (due to their syntactic and semantic proximity), but no such asymmetry of priming would be predicted, as scales are not supposed to play any role in the derivation of pragmatic interpretations for scalar implicatures.

The experiment described in this paper shows that an asymmetric relation holds between the members of lexical scales implicated in scalar implicature computations: weaker terms of a scale (e.g., "seldom") primed stronger terms (e.g., "never") more than the reverse. In a word decision task with masked priming, where participants were asked to judge whether the target presented in a screen was or was not a word in French, they were faster to judge that the stronger term of a scale was a word when it was subliminally preceded by the weaker term of the scale (e.g., "*SELDOM – never*"), than to judge that the weaker term of the scale was a word when it was subliminally preceded by the stronger term (e.g., "*NEVER – seldom*"). This asymmetry suggests, for the first time in the literature, that lexical scales are a psychological reality.

These results do also allow us to distinguish between the different predictions of the two main accounts of the role of lexical scales in the generation of pragmatic interpretation for scalar implicatures. They clearly favor the involvement of scales in the derivation of the pragmatic interpretation for scalar implicatures, in keeping with NG predictions and in contradiction with PG predictions.

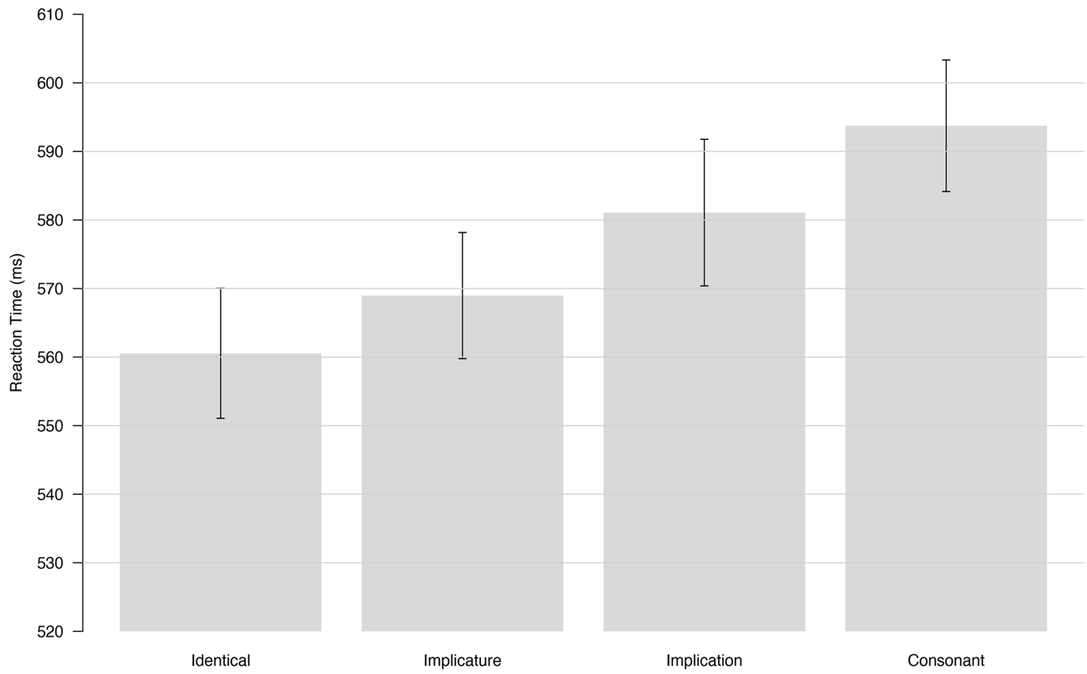


FIGURE 2 | Grand Average of Reaction Time (RT; ms) obtained for each condition. Error bars represent the standard error of the mean across 48 subjects.

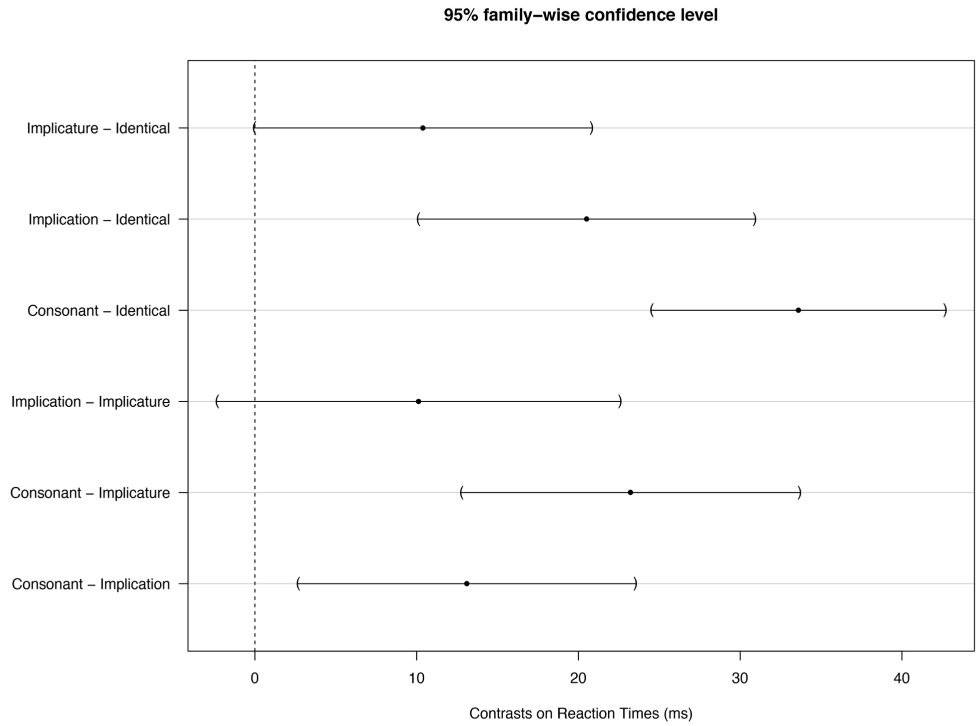


FIGURE 3 | Confidence intervals obtained with a 95% family wise estimate for Tukey contrasts on condition with our linear mixed effect model.

TABLE 1 | Simultaneous tests for general linear hypotheses estimated with a linear mixed effect model with condition as a fixed effect and subject and target word as random effects.

	Estimate	Standard error	*z*-value	Pr(>|t|) single step	Pr(>|t|) Uncorrected
Implicature – Identical = = 0	10.386	4.071	2.551	0.05154	0.01074
Implication – Identical = = 0	20.506	4.062	5.048	<0.001	<0.001
Consonant – Identical = = 0	33.606	3.542	9.489	<0.001	<0.001
Implication – Implicature = = 0	10.119	4.861	2.082	0.15689	0.03738
Consonant – Implicature = = 0	23.220	4.081	5.689	<0.001	<0.001
Consonant – Implication = = 0	13.100	4.070	3.219	0.00703	0.00258

The p-values reported in Pr(>|t|) columns are adjusted either with a single-step method or no correction.

Our results do not address, however, two further questions. The first one concerns the diversity of the scales we tested. As it can be seen in Data Sheet 1, we have tested a somewhat heterogeneous set of lexical scales and it is possible that some scales would induce pragmatic interpretations at a much higher rate than others would. As suggested by the reviewers of this paper, the effect generated by lexical scales with logical connectives (e.g., <*and, or*>), quantifiers (e.g., <*all, many, some*>) and modals (<*allowed, obligatory*>) could be stronger than other scales such as <*bright, intelligent*>. Supporting this hypothesis, Van Tiel et al. (2016) have argued, based on experimental investigations, that some scales (notably <*all, many, some*> and <*and, or*>) induce a much higher rate of pragmatic interpretations for scalar implicatures than do others (e.g., <*small, tiny*>). Although we have checked our results by scale and observed that our significant priming effect from Implication minus Implicature conditions is observed for the majority of the scales we tested, independently of the type of lexical scale or the number of items it contains (see "Presentation 1" for an exploratory analysis by scale), we did not conduct statistical analyses using "scale" as a factor, because for each subject we had only one data point per scale in each condition. So this analysis would be meaningless. Consequently, our data does not allow us to propose an interpretation of the effects derived by each scale individually. However, it might be considered that since the overall pattern of results can be observed for the majority of the scales we tested; despite of their heterogeneity, the asymmetry in RT between the implication and the implicature conditions seems to be robust enough. Further investigations using the same methodology exploited in this paper (masked priming) could, however, be done to address the question of the differences in the magnitude of the effect across scales. For instance, it would be important to investigate more precisely whether the variability between scales that has been evidenced in recent work using other experimental methods (see Van Tiel et al., 2016) and seems to be present in our data (see "Presentation 1"), could be replicated in other studies with specific predictions about how and why some lexical scales can behave differently in they way they induce pragmatic interpretations. Nevertheless, our results are entirely compatible with the idea that scales may differ in how strongly they mandate pragmatic interpretations, or in the degree of automaticity with which they are accessed in the interpretation of scalar implicatures.

The second question that our results do not address concerns the possibility for participants to consider alternatives beyond the lexical items that appear in Horn scales, as it was recently suggested in a computational model of pragmatic inferences developed by Peloquin and Frank (2016). Their model tried to account for the fact that people consider the use of "*some*" inappropriate when the speaker could have used "*one*" or "*two*" and for the fact that when asked to produce alternative words to replace a word in a sentence (e.g., "*some*" in "*Some students came*"), people come up with lexical items beyond the relevant Horn scale (e.g., not only "*many*" or "*all,*" but also "*few*" and "*none*"). This led to the proposition that the alternative set for "*some*" should include "*none,*" "*few,*" "*most,*" "*all*". We think that this does not pose a major problem for Horn scales: the first phenomenon does not lead to a pragmatic interpretation but merely to an infelicity judgment, which does not necessarily entail a pragmatic interpretation to "*not all*"; the second phenomenon does not seem to have anything to do with the derivation of a pragmatic interpretation. So we take it that the present results should be interpreted, quite simply, as a way of adjudicating between the two main approaches to scalar implicatures.

This is thus one of the first empirical results clearly consistent with the new version of NG account as recently proposed by Chierchia (2013). It should not, however, be taken to verify it in its entirety. The process through which scalar implicatures are derived in that account is complex. Additionally, the whole account is wide ranging and cannot be reduced to the interpretation of scalars. However, we provide an important first step in the empirical investigation of that account and bring a new type of data examining to what extent different words on lexical scales prime one another, which allowed us to distinguish accounts of scalar implicature generation.

Note, however, that the present results are obtained from words presented in isolation, while pragmatic interpretations are obtained for scalar terms occurring in sentences, usually in context. Our results do not have much to say about the process itself (notably they shed no light on whether as claimed by Chierchia (2013), it is an exhaustification process using a silent operator on sets of alternatives). However, they strongly suggest, given the asymmetry in RTs between the implication and implicature conditions, that scales must play a role in the interpretation process of scalar implicatures. Otherwise, the asymmetry would not have been observed. It is, indeed, hard if not impossible to explain this asymmetry based on the PG account. Finally, it should be noted that the involvement of scales in the interpretation process of scalar implicatures, which is the conclusion mandated by our experimental results, is compatible not only with Chierchia's (2013) syntactic NG approach, but also with pragmatic Gricean approaches (Geurts, 2009; Geurts and Pouscoulous, 2009; Geurts and Van Tiel, 2013).

In summary, this study reported the first experimental evidence leading a distinction between the two main accounts for the derivation of pragmatic interpretations for scalar implicatures: *NG* versus *PG*. While PGs refuse any role for lexical scales in the derivation of scalar inferences and offer a context-based inferential account for scalar implicatures, NG accounts claim that lexical scales are the core mechanism behind the computation of scalar implicatures, and predict an asymmetry in priming between the implicature and the implication conditions. Supporting this hypothesis, the results that we obtained in a lexical decision task using lexical scales in a masked priming paradigm showed that weaker terms in a scale primed stronger terms more than the reverse. This asymmetry provides then the first experimental evidence in favor of the psychological existence of scales and therefore supports the claim of NG accounts for the role of lexical scales in the computation of scalar implicatures.

AUTHOR CONTRIBUTIONS

AdC and AR have designed the study and written the paper. AC programmed the experiment. AdC ran the experiments and analyzed the data under the supervision of TN, AC and J-BV.

FUNDING

This research was supported by an Ampère Excellence Scholarship from the École Normale Supérieure de Lyon to Alex de Carvalho. This work would not have been possible without the support of the CNRS, and the École Normale Supérieure de Lyon as well.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpsyg.2016.01500

REFERENCES

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01
Bott, L., Bailey, T. M., and Grodner, D. (2012). Distinguishing speed from accuracy in scalar implicatures. *J. Mem. Lang.* 66, 123–142. doi: 10.1016/j.jml.2011.09.005
Bott, L., and Noveck, I. (2004). Some utterances are underinformative: the onset and time course of scalar implicatures. *J. Mem. Lang.* 51, 437–457. doi: 10.1016/j.jml.2004.05.006
Carston, R. (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell Publishing.
Chierchia, G. (2004). "Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface," in *Structures and Beyond*, ed. A. Belletti (Oxford: Oxford University Press), 39–103.
Chierchia, G. (2013). *Logic in Grammar: Polarity, Free Choice and Interpretation*. Oxford: Oxford University Press.
Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., et al. (1998). Imaging unconscious semantic priming. *Nature* 395, 597–600. doi: 10.1038/26967
Dupuy, L., Van der Henst, J.-B., and Reboul, A. (2016). Context in generalized conversational implicatures: the case of some. *Front. Psychol.* 7:381. doi: 10.3389/fpsyg.2016.00381
Feeney, A., Scafton, S., Duckworth, A., and Handley, S. J. (2004). The story of some: everyday pragmatic inferences by children and adults. *Can. J. Exp. Psychol.* 58, 121–132. doi: 10.1037/h0085792
Forster, K. I., and Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *J. Exp. Psychol. Learn. Mem. Cogn.* 10, 680–698.
Geurts, B. (2009). Scalar implicatures and local pragmatics. *Mind Lang.* 24, 51–79. doi: 10.1111/j.1468-0017.2008.01353.x
Geurts, B., and Pouscoulous, N. (2009). Embedded implicatures?!? *Semant. Pragmat. 2,* 1–34. doi: 10.3765/sp.2.4
Geurts, B., and Van Tiel, B. (2013). Embedded scalars. *Semant. Pragmat.* 6, 1–37. doi: 10.3765/sp.6.9
Grice, H. (1975). Logic and conversation. *Syntax Semant.* 3, 41–58.
Grice, H. P. (1989). "Logic and conversation," in *Studies in the Way of Words*, ed. H. P. Grice (Cambridge, MA: Harvard University Press), 16–40.
Guasti, M. T., Chierchia, G., Crain, S., Foppolo, F., Gualmini, A., and Meroni, L. (2005). Why children and adults sometimes (but not always) compute implicatures. *Lang. Cogn. Process.* 20, 667–696. doi: 10.1080/01690960444000250
Hartshorne, J. K., Snedeker, J., Liem Azar, S. Y.-M., and Kim, A. E. (2015). The neural computation of scalar implicature. *Lang. Cogn. Neurosci.* 30, 620–634. doi: 10.1080/23273798.2014.981195
Horn, L. R. (1972). *On the Semantic Properties of Semantic Operators in English*. Ph.D. Dissertation, University of California, Los Angeles, CA.
Horn, L. R. (2004). "Implicature," in *The Handbook of Pragmatics*, eds L. R. Horn and G. Ward (Oxford: Blackwell Publishing), 3–28.
Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous inference in general parametric models. *Biom. J.* 50, 346–363. doi: 10.1002/bimj.200810425
Levinson, S. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicatures*. Cambridge, MA: The MIT Press.
New, B., Pallier, C., Brysbaert, M., and Ferrand, L. (2004). Lexique 2: a new French lexical database. *Behav. Res. Methods Instrum. Comput.* 36, 516–524. doi: 10.3758/BF03195598
Noveck, I. (2001). When children are more logical than adults: experimental investigations of scalar implicatures. *Cognition* 78, 165–188. doi: 10.1016/S0010-0277(00)00114-1
Noveck, I., and Sperber, D. (2007). "Why and how of experimental pragmatics: the case of 'scalar inferences,'" in *Advances in Pragmatics*, ed. N. Burton-Roberts (Basingstoke: Palgrave).
Papafragou, A., and Musolino, J. (2003). Scalar implicatures: experiments at the semantics-pragmatics interface. *Cognition* 86, 253–282. doi: 10.1016/S0010-0277(02)00179-8
Papafragou, A., and Tantalou, N. (2004). Children's computation of implicatures. *Lang. Acquis.* 12, 71–82. doi: 10.1207/s15327817la1201_3
Peloquin, B., and Frank, M. C. (2016). *Determining the Alternatives for Scalar Implicatures*. Available at: http://langcog.stanford.edu/papers_new/peloquin-2016-underrev.pdf
Perea, M., and Rosa, E. (2002). The effects of associative and semantic priming in the lexical decision task. *Psychol. Res.* 66, 180–194. doi: 10.1007/s00426-002-0086-5
Pouscoulous, N., Noveck, I. A., Politzer, G., and Bastide, A. (2007). A developmental investigation of processing costs in implicature production. *Lang. Acquis.* 14, 347–375. doi: 10.1080/10489220701600457
R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Sperber, D., and Wilson, D. (1986/1995). *Relevance: Communication and Cognition*. Oxford: Basil Blackwell.

Van Tiel, B., Van Miltenburg, E., Zevakhina, E., and Geurts, B. (2016). Scalar diversity. *J. Semant.* 33, 137–175. doi: 10.1093/jos/ffu017

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Neuropsychologia 98 (2017) 4–12

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Ambiguous function words do not prevent 18-month-olds from building accurate syntactic category expectations: An ERP study

Perrine Brusini[a,b,*], Ghislaine Dehaene-Lambertz[c,d,e], Marieke van Heugten[b,f], Alex de Carvalho[b], François Goffinet[g], Anne-Caroline Fiévet[b], Anne Christophe[b,g]

[a] *International School for Advanced Studies (SISSA), Language, Cognition and Development Laboratory, Trieste, Italy*
[b] *Laboratoire de Sciences Cognitives et Psycholinguistique (EHESS-ENS-CNRS), Ecole Normale Supérieure, PSL Research University, Paris, France*
[c] *INSERM, Cognitive Neuroimaging Unit, F91191 Gif-sur-Yvette, France*
[d] *CEA, NeuroSpin Center, IFR 49, F91191 Gif-sur-Yvette, France*
[e] *Université Paris XI, F91405 Orsay, France*
[f] *Department of Psychology, University at Buffalo, The State University of New York, United States*
[g] *AP-HP, Université Paris Descartes, Maternité Port-Royal, Paris, France*

A R T I C L E I N F O

Keywords:
Noun-verb categorization
Function words
Event related potentials
Early syntactic processing
Brain

A B S T R A C T

To comprehend language, listeners need to encode the relationship between words within sentences. This entails categorizing words into their appropriate word classes. Function words, consistently preceding words from specific categories (e.g., *the ball*$_{\text{NOUN}}$, *I speak*$_{\text{VERB}}$), provide invaluable information for this task, and children's sensitivity to such adjacent relationships develops early on in life. However, neighboring words are not the sole source of information regarding an item's word class. Here we examine whether young children also take into account preceding sentence context online during syntactic categorization. To address this question, we use the ambiguous French function word *la* which, depending on sentence context, can either be used as determiner (*the*, preceding nouns) or as object clitic (*it*, preceding verbs). French-learning 18-month-olds' evoked potentials (ERPs) were recorded while they listened to sentences featuring this ambiguous function word followed by either a noun or a verb (thus yielding a locally felicitous co-occurrence of *la* + noun or *la* + verb). Crucially, preceding sentence context rendered the sentence either grammatical or ungrammatical. Ungrammatical sentences elicited a late positivity (resembling a P600) that was not observed for grammatical sentences. Toddlers' analysis of the unfolding sentence was thus not limited to local co-occurrences, but rather took into account non-adjacent sentence context. These findings suggest that by 18 months of age, online word categorization is already surprisingly robust. This could be greatly beneficial for the acquisition of novel words.

1. Introduction

Successful spoken language processing requires listeners to compute complex linguistic analyses. Adults typically experience little difficulty correctly applying the grammatical rules of their native language when they communicate with other people. This stands in sharp contrast with young children, whose early telegraphic speech is abundant with omissions. More specifically, function words, such as determiners, auxiliaries, and pronouns –despite their importance for encoding language structure– appear to be consistently lacking from toddlers' early speech patterns (Gerken et al., 1990). As a result, these words have traditionally been thought to be overlooked and not processed by young children. In more recent years, however, experiments testing the perception of these elements have demonstrated that infants and toddlers are sensitive to these items (Gerken and McIntosh, 1993; Hallé et al., 2008; Shi and Gauthier, 2005; Shi et al., 2006b). These studies suggest that the selective production of content words relative to function words is likely strategic (function words convey less semantic content than content words), and that function words may be omitted due to production rather than comprehension constraints (Demuth and Tremblay, 2008; Gerken and McIntosh, 1993).

If toddlers do not experience difficulty perceiving function words, then these items may be of great use during the process of language acquisition, since they convey rich morpho-syntactic information. Distributional, phonological and acoustic analyses reveal that function words are highly frequent, typically occur at the edges of syntactic phrases, and tend to be short and unstressed (Shi et al., 1998): characteristics that differentiate them from content words and that

* Corresponding author at: LSCP, Ecole Normale Supérieure, 29, Rue d'Ulm, 75005 Paris, France.
E-mail address: pbrusini@gmail.com (P. Brusini).

http://dx.doi.org/10.1016/j.neuropsychologia.2016.08.015
Received 1 February 2016; Received in revised form 11 August 2016; Accepted 14 August 2016
Available online 17 August 2016

may allow infants to discriminate between the two types of words. Indeed, children as young as a few days can distinguish lists of function words from lists of content words (Shi and Werker, 2001; Shi et al., 1999) and by 10 months of age, infants have gained sensitivity to frequent function words occurring in their native language (Hallé et al., 2008; Shady, 1996; Shafer et al., 1998; Shi et al., 2006a). Given that most function words tend to consistently co-occur with content words from a specific word class, this early acquisition of function words might also allow infants to rely on these words to anticipate the lexical category of a subsequent content word (Christophe et al., 2008; Gervain et al., 2008; Hochmann, 2013; Hochmann et al., 2010). For instance, if infants learned that determiners (such as *the* or *a*) typically precede nouns whereas pronouns (such as *you* or *he*) generally precede verbs, they could use this knowledge to deduce the syntactic category of words that they have never encountered before. This hypothesis is consistent with a growing body of experimental data showing that children can infer the syntactic category of a novel word after auditory exposure to function word – content word co-occurrences. That is, by 14–16 months of age, infants can work out that a novel content word following one determiner can also follow other determiners, but cannot follow pronouns (Hohle et al., 2004; Shi and Melançon, 2010). Moreover, by 18 months of age, lexical access is speeded and more accurate when known nouns and verbs are preceded by a function word from an appropriate category (i.e. determiners and pronouns, respectively, Cauvet et al., 2014; Kedar et al., 2006; Zangl and Fernald, 2007) and by two years of age, toddlers readily exploit the syntactic context of a novel word to infer whether it refers to an object or an action (Bernal et al., 2007; Oshima-Takane et al., 2011; Waxman et al., 2009). This suggests that early on in life, the presence of a function word can trigger children's expectations regarding the word class of the subsequent item.

Children's early reliance on function words to determine the grammatical category of neighboring content words can be greatly beneficial for language acquisition purposes. However, to become a mature language user, it is important to process not only local co-occurrences of function and content words, but to also learn to take into account the contexts in which these dependencies occur. That is, although function words are generally a good predictor of the word class of the following content word, incorporating the broader syntactic context can sometimes improve or fine-tune categorization. This is particularly profitable in the case of function words that can occupy multiple syntactic roles. Consider, for instance, the French functor *la*, which can surface both as a determiner (e.g., *Très gentiment* ***la*** *girafe me prête sa balle* 'Very kindly **the** giraffe lends me her ball') and as a pronoun clitic (e.g., *Alors moi je* ***la*** *donne au crocodile* 'Then I give **it** to the crocodile'). If toddlers only exploit the local co-occurrence between the two classes to deduce the syntactic category of the content word following the function word, such ambiguities could potentially be devastating. A French-learning child hearing a novel verb in a context such as *Elle* ***la*** *dase* ('She dases **it**'), for example, might erroneously infer that *dase* is a noun, because it is preceded by the functor *la*, which occurs as a determiner much more frequently than as an object clitic (more than 80% of the time, as reported in Shi and Melançon, 2010). If, by contrast, the broader syntactic context could be taken into account, toddlers may be able to infer that, despite the overall likelihood as *la* surfacing as a determiner, the presence of the preceding pronoun *elle* ('she') greatly increases the likelihood of a verb to appear.

Recent work has started to explore whether children's online word categorization is local in nature or whether broader sentence structure is taken into account. In particular, French-learning two-year-olds' brain responses were measured while they listened to sentences featuring the ambiguous function word *la*, either used as a determiner or as an object clitic (Bernal et al., 2010; Brusini et al., 2016). Differences in evoked potentials between grammatical and ungrammatical trials revealed that the classification of content words following the functor *la* depended on the structure of the unfolding sentence. That is, children anticipated a noun when the broader sentence context was indicative of *la* being a determiner (e.g., *hier la* ... 'yesterday the ...'), but they anticipated a verb when the broader sentence context was indicative of *la* being an object clitic (e.g., *je la* ... 'I ... it'). This suggests that toddlers use sentence frames to determine the word class of upcoming linguistic material. By their second birthday, children are thus able to compute complex syntactic dependencies during online language processing.

The finding that two-year-olds can take advantage of more than just the immediately adjacent linguistic context during syntactic processing suggests that they possess both the computational abilities and the linguistic experience necessary to accurately compute complex contexts. However, given that children start using function words as a classification method months before their second birthday (e.g. Cauvet et al., 2014, for nouns and verbs in French; Kedar et al., 2006 for nouns in English; Van Heugten and Johnson, 2011, for nouns in Dutch; Zangl and Fernald, 2007, for nouns in English), we may wonder whether learners who are still in the very early stages of exploiting function words as a categorization cue can also incorporate sentence context during online word classification. On the one hand, children may at first rely mainly on local dependencies between word categories, and may not take into account any non-adjacent information. This would align with the finding that the ability to learn adjacent dependencies typically precedes the ability to learn non-adjacent dependencies (Gebhart et al., 2009; Gomez and Gerken, 1999; Udden et al., 2012) and would imply that word classification in the case of ambiguous function words would follow the most frequent adjacent regularity. Only once children can expand their processing window, will they be able to integrate broader sentence context. On the other hand, children may possess more advanced categorization skills, incorporating the wider linguistic context from the beginning. Following this view, children's syntactic categorization would depend not only on the immediately preceding information in the sentence, but also on the more distant information, even during the early acquisition period. To tease apart these two possibilities, it is crucial to test whether 18-month-olds can take into account the non-local syntactic context during the processing of ambiguous function morphemes. To our knowledge, this study is the first one testing whether young infants can use distant contextual information for word categorization in the presence of a consistent adjacent relationship.

To address this issue, we exploited the event related potential (ERP) paradigm developed by Bernal et al. (2010) in which two-year-olds were presented with short video clips containing grammatical and ungrammatical sentences. In this procedure, high-density electroencephalography (EEGs) is recorded without requiring any overt response from toddlers. Thus it is a well-suited paradigm to determine 18-month-olds' spontaneous syntactic abilities by comparing the neuronal response evoked by grammatical and ungrammatical sentences. Although the electrical components induced by syntactic violations described in the toddler literature are more variable than those in the adult literature,[1] a late positivity, which has been related to revision processes in adults (Kuperberg, 2007), is robustly observed across ages for ungrammatical sentences (Bernal et al., 2010; Brusini et al., 2016; Oberecker and Friederici, 2006; Oberecker et al., 2005; Silva-Pereyra et al., 2005a, 2005b). This positivity is sometimes

[1] Variability between studies may depend on the exact nature of the syntactic violation studied, but may also be related to the use of continuous speech in toddler studies (as opposed to serial presentation of written words in many adult studies), which decreases the amplitude and sharpness of the electrical components. The immature syntactic processing of young participants could furthermore lead to additional variability in the electrical components: their onsets might be less precisely time-locked, therefore decreasing the amplitude of the averaged evoked responses. Finally, children's EEG response contains more large-amplitude low frequency components compared to that of adults, which has been shown to lead to an increase in background endogenous noise (Chu et al., 2014).

Fig. 1. Example of a video story. Each story had the same structure: during the test trials only the speaker's face was visible, whereas in the remainder of the video the whole scene was presented, to keep toddlers interested.

preceded by other components (Brusini et al., 2016; Oberecker et al., 2005; Schipke et al., 2011; Silva-Pereyra et al., 2005b), less consistent across experiments.

Our paradigm relies on the comparison of the same sequences of words embedded in longer sentence contexts. These sentence contexts manipulated the status of the sequences: grammatical sentences either contained an object clitic-critical verb or a determiner-critical noun sequence (e.g., *Alors moi je* ***la*** *donne au crocodile,* 'Then I give **it** to the crocodile' or *Très gentiment* ***la*** *girafe me prête sa balle,* 'Very kindly **the** giraffe lends me his ball', critical words are underlined). Ungrammatical sentences were constructed by placing a noun in a verb context or a verb in a noun context (e.g.,**L'animal et* ***la*** *donne sont heureux*, '*The animal and the give are happy' or **Alors il me* ***la*** *girafe en souriant* '*Then he smilingly giraffes it to me'). We used videos (see Fig. 1 for an example) in which sentences containing the critical content words were first introduced with the support of small toys illustrating the story (introductory part); then when the key grammatical and ungrammatical sentences were presented, only the speaker's face was visible (test part). Thus, the grammaticality of these sentences could not be judged from the visual scene. Crucially, at test, all critical words were preceded by the same ambiguous function word ('la' meaning either *the* or *it*). As a result, the simple adjacent co-occurrence of the functor 'la' and the critical content word does not provide any cue regarding the grammaticality of the sentence. For instance, in the ungrammatical sentence **Alors il me la girafe en souriant* '*Then he smilingly giraffes it to me', the two-word sequences '*il me*', '*me la*' and '*la girafe*' are all legal in French, but '**il me la girafe*' is not. The implementation of this design could lead to three possible patterns of results.

First, if 18-month-olds listening to the test sentences are only able to recover content words provided by the previous visual context (e.g. the presence of a giraffe during the introductory sentences) and/or to compute local statistics between pairs of words, then they should process all sentences similarly, since they contain content words which have been mentioned before, and are always locally correct (all pairs of adjacent words occur frequently together within grammatical sentences). Second, if toddlers are able to make more complex analyses based on the frequency of the categories in natural speech outside the lab, and estimate that "la" should be followed by a noun (since "la" is most often followed by nouns), then they should perceive all sentences containing a 'la+noun' sequence as grammatical and all sentences containing a 'la+verb' sequence as ungrammatical (since in our experimental design 'la' is followed equally often by nouns and verbs, in both grammatical and ungrammatical sentences, this will not lead to an overall distinction between grammatical and ungrammatical sentences). A third possibility would be for toddlers to detect the ungrammaticality regardless of whether nouns and verbs are used. If that were the case, this would indicate that 18-months either build an adult-like syntactic representation of the sentence as it unfolds, or at least pay attention to three- or four-word strings. In line with the ERP literature, we would then expect to record a late positivity for ungrammatical sentences relative to grammatical sentences (Bernal et al., 2010; Brusini et al., 2016; Oberecker and Friederici, 2006; Oberecker et al., 2005; Silva-Pereyra et al., 2005a, 2005b), potentially preceded by an early effect (Brusini et al., 2016; Oberecker et al., 2005; Schipke et al., 2011; Silva-Pereyra et al., 2005b).

2. Method

2.1. Participants

A total of 25 monolingual French-learning toddlers (12 boys) were tested (mean age 18.4; range 17.8–19.2). All toddlers were in good health at the time of test, had no detected developmental disorders and no reported hearing deficits. An additional 35 toddlers participated in this experiment but did not provide exploitable data, because they were too agitated, stopped the test prematurely, or the quality of the recorded potentials was not sufficient (only data of toddlers with at least 40 artifact-free trials, and at least 19 in each of the two grammaticality conditions, were analyzed). Another 38 toddlers were recruited for this experiment, but as they refused to wear the EEG net, no data were collected.[2] Families received a diploma as a token of appreciation. This research was approved by the local 'Ile-de-France III' ethics committee.

2.2. Stimuli and design

Four nouns and four verbs that are typically acquired at a young age were selected by Bernal and colleagues (2010) as critical words. None of these critical words were noun/verb homophones (nouns: *une fraise* 'a strawberry', *une balle* 'a ball', *une grenouille* 'a frog', *une girafe* 'giraffe'; verbs: *manger* 'to eat', *donner* 'to give', *regarder* 'to look', *finir* 'to finish'). For each category, two of the words were monosyllabic and two were bisyllabic (for verbs, this was true for the present-tense form that was used in the experiment). Nouns and verbs were embedded in sentences that were grammatical in half of the test trials (e.g., *Alors moi je* ***la*** *donne au crocodile,* 'Then I give **it** to the crocodile' or *Très gentiment* ***la*** *girafe me prête sa balle,* 'Very kindly **the** giraffe lends me her ball', critical words are underlined) and ungrammatical in the other half of the test trials (e.g.,**L'animal et* ***la*** *donne sont heureux*, '*The animal and **the** give are happy' or *Alors *il me* ***la*** *girafe en souriant* '*Then he smilingly giraffes **it** to me' see Table 1 for the full design). All critical words were preceded by the function word 'la', which can either take the role of determiner or that of object clitic. Note that all words were presented in both verb and noun positions across grammatical and ungrammatical sentences. This controls for any possible low-level acoustic differences between the grammatical and ungrammatical

[2] As noted by an anonymous reviewer, the overall rejection rate may seem high (although not unusual for ERP experiments using this age group; e.g. Brusini et al., 2016). Towards the end of the study we reduced the dropout rate by introducing a series of slight modifications to the welcoming procedure of parents and toddlers (playing a cartoon on the screen before the experiment started, and telling the children that the television works only when they have the net on; securing the parents' help by asking them to hold the child's hands while the net was put into place, then asking them to maintain the child firmly on their lap during the experiment, acting as a 'gentle car seat', in order to reduce movement artifacts).

Table 1
Construction of the test sentences: the critical words (nouns and verbs) occurred in noun and verb contexts, yielding grammatical sentences (when the context was congruent with the syntactic category of the critical word) and ungrammatical sentences, marked with a star (when context and critical word syntactic category were incongruent).

	Grammatical	Ungrammatical
Nouns	Très gentiment la girafe me prête sa balle	*Alors il me la girafe en souriant.
	Very kindly the giraffe lends me his ball.	**So he smilingly giraffes it to me.*
Verbs	Alors moi je la donne au crocodile !	*L'animal et la donne sont heureux.
	So I give it to the crocodile!	**The animal and the give are happy*

conditions. Thus, a main effect of grammaticality cannot be due to acoustic properties of the critical words themselves (they are the same on both sides of the comparison), nor can they be due to acoustic properties of the contexts themselves, as they were carefully matched in number of syllables before the critical words, and in the syntactic structures used (see Supplementary materials for a full list of stories, and a detailed description of the matching procedure).

We used the same 16 video clips recorded by Bernal et al. (2010). In these clips, a native French speaker narrated a 30-s story in a child-directed fashion. She used toys to illustrate the stories and maintain the toddlers' interest. All stories were scripted in a similar way (see Fig. 1 for an example). They started with an introduction of the characters present in the story (the critical noun and verb of the story were mentioned in these sentences), followed by two test sentences, a filler sentence to keep the toddlers engaged in the task, and ended with two more test sentences. During the introduction and the filler sentences, the whole scene was visible. During the test sentences, by contrast, only the speaker's face was visible (see the screen-shots in Fig. 1). This way, the visual information was highly similar across test sentences and did not provide any cues regarding the plausibility of the items. Because children tend to focus on the speaker's eyes and mouth (Lewkowicz and Hansen-Tift, 2012), this should also help minimize children's eye-movements.

Within each story, two test sentences contained a critical noun and the other two contained a critical verb. The order of the conditions was counterbalanced across the 16 stories. Overall, there were 64 different test sentences, 32 grammatical and 32 ungrammatical. Across stories, test sentences were counterbalanced for the number of syllables before and after the critical word, and for the syntactic structures used in the grammatical and ungrammatical conditions. Toddlers watched these 16 stories several times, in different orders (at least once and no more than 4 times).

2.3. Procedure

Prior to test, the experimenter positioned a geodesic 128-sensor net (EGI, Eugene, USA) relative to the anatomical markers on the toddler's head. A short play session, featuring the toys that were used in the videos, took place before the experiment and served two purposes: toddlers were reminded of the meaning of the words that would be used in the test session, and they were distracted while the net was put in place. The experiment took place in a sound-attenuated booth. Children were seated on their parent's lap and watched between one and four blocks of 16 video stories, while EEGs were recorded. Parents were asked to remain silent and refrain from distracting their child throughout the experiment. Two computers were used to conduct the experiment; one played the video-clips and the other one selected the clip to be played and sent trial information to the EEG recording system. If necessary, the experimenter paused the session between two stories and restarted it once the toddler appeared to be focusing again.

2.4. ERP recording and data analysis

2.4.1. ERP recording

High-density EEG (128 electrodes, referenced to the vertex) was continuously digitized at 250 Hz during the video presentations (Net amps 200, EGI, Eugene, USA). Recordings were digitally band-pass filtered (0.3–20 Hz) and segmented into 1400 ms-epochs starting 200 ms prior to critical word onset. For each epoch, channels contaminated by eye or motion artifacts (local deviation higher than 80 μV) were automatically excluded, and trials with more than 20% contaminated channels were excluded from the analysis. For each toddler, channels comprising fewer than 50% uncontaminated trials were excluded for the entire session. Excluded channels were interpolated for each trial separately using the linear interpolation method of EEGlab (Delorme and Makeig, 2004). The artifact-free epochs were averaged for each participant in each condition (mean number of artifact-free epochs per toddler: 128.1 in total: 64.5 in the grammatical condition and 63.6 in the ungrammatical condition). Averages were baseline-corrected (−200 to 0 ms) and transformed into reference-independent values using the average of all channels as reference.

2.4.2. Data analysis

Channels at the edge of the scalp, which are generally very noisy in toddlers, were not considered, leaving 91 electrodes for analyses.[3] Given the number of electrodes (here 91) and time samples (here 300), the risk of type I errors (false alarms) is high if each possible comparison (here 91*300) is considered. To avoid this issue and reduce the number of comparisons, three strategies are generally proposed to analyze high-density recordings. The most classic strategy involves constraining the analysis by taking into account the existing literature and computing *t*-tests or ANOVAs on the time windows and scalp regions often reported to be at play in similar experimental conditions. This method has been criticized as being sensitive to biases in the literature and restricting analyses to known effects. Furthermore in less studied populations, the literature may not be sufficiently dense to correctly infer typical time windows and regions. A second strategy consists of first identifying experimental effects in a subset of the data and then checking whether it replicates it using another subset. However, because this strategy requires the data set to be subdivided, it reduces the number of total trials taken into account to establish an effect, which is problematic with a toddler population where it is challenging to obtain a sufficient number of clean trials. A final strategy, the cluster-based permutation analysis (Maris and Oostenveld, 2007), exploits the fact that neighboring channels and time-points are highly correlated. This approach identifies spatio-temporal clusters that exhibit a significant difference between conditions. The statistical value of these clusters is assessed by comparing them to a null distribution obtained through randomized permutations of the initial data. In practice, a *t*-test is computed on each electrode and time-point, then a threshold is applied and clusters are built as the sum of the t-values above threshold in neighboring points in time and space. The same procedure is applied on the shuffled data and the largest clusters from the original data are compared to the distribution of the clusters obtained in the shuffled data. This general method, which is instantiated in several MATLAB toolboxes (SPM Kiebel and Friston, 2004; TFCE Mensen and Khatami, 2013; Fieldtrip, Oostenveld et al., 2011; LIMO, Pernet et al., 2011), is conservative, but its sensitivity depends on how the clusters are constructed (see Mensen

[3] Using the 128-channels Hydrocel Geodesic Sensor Net, the following electrodes, which represent the three outer-most circles of the geodesic net, were removed: 17-126-127-21-14-8-1-125-121-120-119-114-113-107-99-94-88-81-73-68-63-56-49-43-48–128-44-38-32-25-100–95-89-82-74–69-64–57. As a result, 91 electrodes are analyzed.

and Khatami, 2013 for a comparison of the different toolboxes and the different choices to construct clusters). In a nutshell, using this method, there is a trade-off between sensitivity to local but intense effects versus effects with smaller amplitude but more sustained in time and diffuse on the scalp. Here, much like in Brusini et al. (2016), we first use the conservative cluster-based permutation analysis to ensure that a main effect of grammaticality was present in our data. Then, to analyze differences between sub-conditions, we use the more sensitive method based on a selection of regions of interest from the existing literature.

The cluster-based permutation analysis was conducted on the main effect of grammaticality (i.e. comparison between grammatical and ungrammatical sentences, pooling across nouns and verbs) using the Fieldtrip toolbox, with 10,000 iterations and a threshold of p=0.01. For this analysis, we considered two time-windows: an early one (100–600 ms) to capture the early effects typically described in adults, i.e. either a LAN (Left Anterior Negativity), which typically appears between 100 and 400 ms, or an N400 which surfaces around 300–600 ms. The second time window (500–1000 ms) aims to capture the late P600 response whose typical latency is between 500 and 800 ms, but can also occur later, especially in children (Atchley et al., 2006; Schipke et al., 2012).

In the literature-driven analysis conducted next, we then constrained the time windows and clusters of electrodes to be analyzed based on prior findings. We observed a late posterior positivity resembling a P600, when inspecting the grand-average difference between grammatical and ungrammatical sentences. We selected the time window and clusters of electrodes encompassing this effect and, for each subject in each of the 4 conditions (Grammaticality by Word Category) averaged the voltage over the contributing data points. This allowed us to test for potential differences between our sub-conditions (verbs and nouns). As the grammatical and ungrammatical conditions featured the same set of critical words (and the same visual scenes), any difference between conditions will show that 18-month-old toddlers react differently to ungrammatical sentences relative to grammatical sentences. We also analyzed the grammaticality effect separately for the two halves of the experiment to establish whether toddlers might have changed their behavior over the course of the experiment (e.g., learning that 'giraffe' could also be used as a verb, as the experiment proceeds, which would lead to a decrease in the grammaticality effect).

3. Results

3.1. Cluster-based permutation analysis

During the early time-window (100–600 ms), the cluster-based permutation analysis did not reveal any significant effect. In contrast, the analysis of the late time-window revealed a near-significant positive centro-posterior cluster (p=0.051) spreading between 875 and 925 ms and consisting of up to four electrodes around P8 and P4 at its peak, together with a negative cluster that was its counterpart and was significant between 900 and 925 ms, containing 7 electrodes at its peak, around F3 (p=0.02). This effect exhibits the timing and topography typical of a P600, which is almost systematically present in adults when grammatical and ungrammatical sentences are compared. These first analyses reveal that toddlers are able to distinguish between our grammatical and ungrammatical sentences, even in the strictly controlled contexts used here.

3.2. Literature-driven analysis

consistent with the cluster-based permutation analysis reported above, the inspection of the two-dimensional reconstruction of the Ungrammatical-Grammatical difference revealed a late posterior positivity starting 800 ms after the beginning of the ungrammatical words, for a duration of 150 ms. This positivity was located over the parietal area and right-lateralized (Fig. 2). An ANOVA conducted on the average voltage of the 800–950 ms period and over the selected electrodes revealed a significant effect of Grammaticality (F(1,24) =20.0, p=0.0002). This effect remained significant when we considered each half of the experiment separately (first half, t(24)=2.97, p=0.007; second half: t(24)=3.21, p=0.004). The main effect of Word Category was not significant (F(1,24) < 1), nor was the interaction between Word Category and Grammaticality (F(1,24) < 1). Restricted analyses showed that the Grammaticality effect was present for both Word Categories (for Nouns: t(24)=2.0, p=0.057; for Verbs: t(24)=3.52, p < 0.002, see graphs in Fig. S1, Supplementary materials). This suggests that both for nouns and for verbs, 18-month-olds detect the misuse of the critical words in the ungrammatical conditions.

4. Discussion

4.1. On-line identification of noun and verb contexts

In this experiment, we examined 18-month-olds' ability to identify the syntactic contexts in which nouns and verbs occur. More specifically, we observed that toddlers distinguish between contexts that require a noun and contexts that require a verb, even when the function word preceding the critical word was phonemically identical in both cases (i.e., *la* , meaning 'the' or 'it' depending on the context). Phrases such as *Hier* ***la*** X_N 'Yesterday **the** X_N' and *Elle* ***la*** X_V 'She X_V **it**' require different word classes for the critical word X, and toddlers exhibit different evoked potentials for critical content words when they are unexpected (nouns in verb contexts, or verbs in noun contexts) than when they are consistent with their preceding contexts (nouns in noun contexts, and verbs in verb contexts). Thus, by 18 months of age, toddlers' analysis of the unfolding sentence is not limited to tracking two-word co-occurrence patterns of function and content words: if that were the case, they should have based their expectations on the most frequent associations, anticipating a noun after hearing *la*, since this function word occurs more frequently as a determiner (therefore preceding nouns) than as an object clitic (therefore preceding verbs). Instead, their processing is more sophisticated, based on the on-line integration of nouns and verbs within their syntactic contexts. This experiment thus reveals that 18-months know not only the meaning of the verbs and nouns we used, but also the broader syntactic frames in which they occur.

Toddlers may have reacted to inconsistencies between the preceding context and the syntactic category of the critical word in two different ways. First, they may have accessed the lexical entry for the critical word (e.g. *mange* 'eat') and noticed that it occurred in a context that was inconsistent with its syntactic category. Under this interpretation, the ungrammaticality response would reflect a difficulty in integrating a known lexical item within the syntactic-semantic structure of the sentence. On the other hand, toddlers might have attempted to construct a novel lexical entry, with a different syntactic category (and different meaning) for a known phonological form. Under that assumption, upon hearing *mange* 'eat' in a noun context, they would search for a possible meaning for this word, much like they would if they had encountered a novel word form (such as *blicket*). This interpretation is not implausible, since we know i) that there are myriad noun/verb homophones in natural languages; ii) that children under two years of age have already acquired many such homophones (de Carvalho et al., 2014, 2016a; Veneziano and Parisse, 2011); and iii) that toddlers readily learn homophones of well-known words, especially when the two interpretations belong to two different syntactic categories (Dautriche et al., 2015a; Dautriche et al., 2015b). However, in these homophone-learning studies, children were provided with a plausible referent for the second member of the homophone pair. In the present experiment, by contrast, the 'new' homophone (*girafer* 'to giraffe' or *une donne* 'a give') appeared in the absence of any supportive

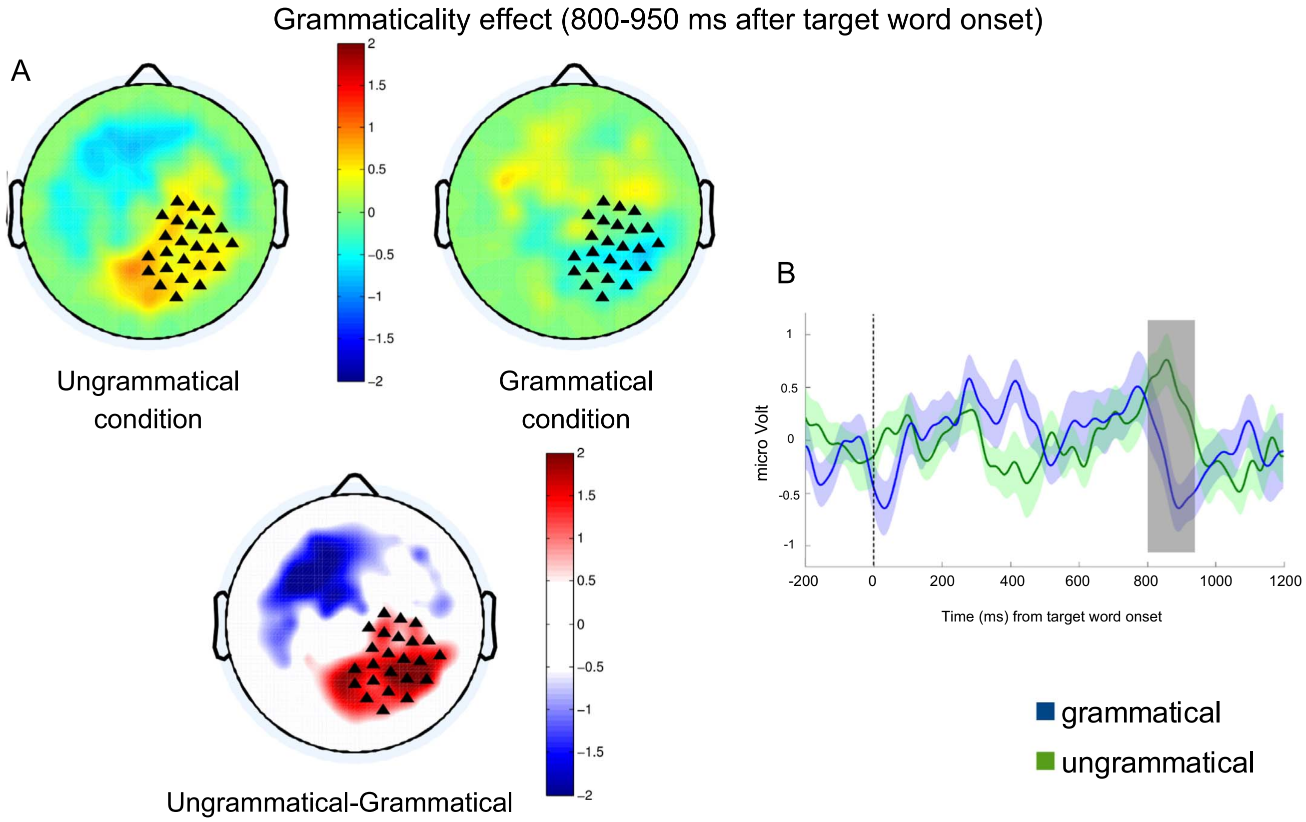


Fig. 2. *Grammaticality effect*: A late positive potential was observed in response to ungrammatical sentences 800 ms after the misplaced noun or verb. (A) Voltage recorded for the ungrammatical and grammatical sentences (on top) together with the map of statistical significance (*z*-score) of the difference Ungrammatical – Grammatical (triangles represent the electrodes used in the ANOVAs). (B) Time course of the activation for the selected cluster of electrodes, over the entire trial (blue curve: grammatical sentences; green curve: ungrammatical sentences); the selected time window is shaded (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

context (as only the face of the speaker was visible in the test sentences), while the 'known' homophone (e.g. *la girafe* 'the giraffe' or *elle donne* 'she gives') appeared in a supportive visual context in which the known meaning was reinforced (during introductory sentences) about half the time. An alternative possibility is thus that this may have encouraged children to access the known lexical item rather than attempt to build a novel lexical entry. The present results do not allow us to disentangle these two interpretations. At any rate, both interpretations lead to the conclusion that 18-month-olds are able to distinguish between noun and verb contexts, and experience difficulty when known nouns and verbs are placed in incorrect contexts.

The finding that 18-month-olds differentiate between grammatical and ungrammatical sentences, even when the function word immediately preceding the critical word is ambiguous, is consistent with earlier results from 24-month-old children on similar stimuli (Bernal et al., 2010; Brusini et al., 2016). In these two other studies, 24-month-olds exhibited different responses for critical nouns and verbs in correct vs. incorrect contexts. In Bernal et al. (2010), a long-lasting left-temporal positivity was observed starting around 300 ms and extending until 1000 ms after target word onset. In Brusini et al. (2016), by contrast, a late posterior positivity was observed (700–900 ms), preceded by an early left anterior negativity (100–400 ms), a pattern which closely mirrored the results obtained with adult participants tested in the same experiment, who exhibited a LAN-P600 complex typically observed in the processing of ungrammatical sentences in adults (Brusini et al., 2016). The results from the present experiment are very similar to the late response observed in Brusini et al. (2016), only slightly delayed in time (800–950 ms).

This late positivity response is also consistent with prior work from other laboratories examining young children's processing of ungrammatical sentences (Oberecker et al., 2005; Schipke et al., 2011; Silva-Pereyra et al., 2005b). Although the timing of this response is delayed in young children relatively to the typical adult P600 (here by approximately 250 ms), its topography is consistent with the description in adults. Delayed latencies have typically been attributed to greater task difficulty for younger participants (Atchley et al., 2006; Holcomb et al., 1992) and slower higher-level responses due to the weak myelination of long-range tracts and immature cortical areas (Kouider et al., 2013). Following proposals from the adult and child literature, this P600-like effect probably reflects the highest level of sentence integration, unifying the semantic and syntactic levels of analysis (Friederici, 2011; Hagoort, 2005).

Contrary to several studies conducted with older children, the present study did not evoke an early response (Brusini et al., 2016; Oberecker and Friederici, 2006; Schipke et al., 2011; Silva-Pereyra et al., 2005a). The presence of an early response in these other studies was interpreted as an automatic mismatch between the expected and the actual word, thus reflecting children's ability to build on-line expectations regarding the syntactic category of an upcoming word. It may be the case that sentence processing in 18-month-olds is not yet fast enough to allow them to rely on such rapid predictive processing. Alternatively, the onset of this response in children this young may be too variable from one trial to the next, thus hindering the observation of an early effect when trials are averaged.

4.2. Syntactic analysis or multi-word contexts?

The finding that toddlers processed our grammatical and ungrammatical sentences differently is indicative of surprisingly sophisticated categorization skills early on in life. Given that all pair-wise combinations of adjacent function and content words in our sentences (even in the ungrammatical ones) are legal in French, the present experiment demonstrates that 18-month-old toddlers are able to take into account non-adjacent lexical items when computing syntactic structure. What may have underlied this processing of broad sentence context in order to identify noun and verb contexts? One possibility, suggested by a reviewer, is that toddlers may have treated the ambiguous functor *la* as an optional item, thus considering *Alors moi je* ***la*** *X...* ('Then I X **it**...') or *Très gentiment* ***la*** *X...* ('Very kindly **the** X...') to be functionally similar to *Alors moi je X* or *Très gentiment X,* respectively. According to this proposal, toddlers' analysis of the preceding material alone might lead them to anticipate whether the next item is a verb (in the case of *Alors moi je X*) or a noun (in the case of *Très gentiment X*[4]) and *la* may simply be left unanalyzed. While our design indeed leaves open the possibility that the ambiguous function word does not contribute to the prediction of the category of the following word, we believe that this account is unlikely given the vast amount of evidence that 18-month-olds have long gained sensitivity to the co-occurrence between determiners and nouns (Hohle et al., 2004; Kedar et al., 2006; Shi and Melançon, 2010; Zangl and Fernald, 2007). Nonetheless, future EEG work directly targeted at the processing of the ambiguous function word could potentially further clarify this.

This leaves us with two remaining explanations. First, it is possible that 18-month-old toddlers have started to compute the syntactic structure of sentences in an adult-like manner. According to this view, children would know that *giraffe* is a noun and that nouns can follow determiners or adjectives, but not object clitics. In addition, children would assign the correct category to the preceding ambiguous function word –determiner or clitic– depending on the grammatical structure of the sentence. In essence, this would thus involve taking into account the syntactic category of the item preceding the ambiguous function word to build expectations about the syntactic category of the word following this ambiguous function word.

A second possibility is that toddlers relied on the two- or three-word contexts that preceded the critical nouns and verbs. According to this view, the specific brain responses evoked by ungrammatical sentences are due to the fact that the sequence of words *elle+la* has never been heard directly before a noun. Computing two-word contexts has been shown to be an effective way of categorizing nouns and verbs. For instance, a computational model tested on child-directed speech achieved very high precision in noun/verb categorization, simply by using two-word contexts that were extracted during a training phase on the basis of just a handful of nouns and verbs (under the assumption that these words are known by toddlers) to classify unfamiliar content words at test (Brusini et al., 2011; see also Redington et al., 1998). Thus, toddlers may very well have succeeded in the present experiment because they know which two-word contexts are indicative of subsequent nouns and which are indicative of subsequent verbs. Both of these mechanisms, the computation of a full-fledged hierarchical syntactic structure and the computation of two-word contexts, can greatly facilitate language processing. In fact, they are both part of mature behavior, as adult listeners have been shown to use both strategies during language processing (e.g. Ferreira and Patson, 2007).

Although the present experiment does not allow us to distinguish between these latter two interpretations, an indication that 18-month-olds might already be able to compute a rough syntactic structure, rather than solely relying on two-word contexts, comes from very recent behavioral work showing that French 18-month-olds are able to use phrasal prosody to assign two different syntactic structures to the same string of words (de Carvalho et al., 2015). For example, a sentence like 'Do you see the baby blicks' can be produced either as '[Do you see the baby blicks]? ', where the novel word 'blicks' is a noun, or as '[Do you see]? [the baby] [blicks]! ', where 'blicks' is a verb (as in 'Do you see? the baby sleeps! '; brackets indicate prosodic boundaries, reflecting the different syntactic structures). Toddlers correctly attributed a noun or a verb meaning to the critical word 'blicks', depending on its position within the prosodic-syntactic structure that they heard (consistent with earlier work with adults and preschoolers, de Carvalho et al., 2016a, 2016b; Kjelgaard and Speer, 1999; Millotte et al., 2008, 2007). Note that to succeed in this task, processing two-word contexts is not sufficient, since the words themselves are identical: 'Do-you-see-the-baby-blicks'. Thus, only phrasal prosody, reflecting the different syntactic structures, gives an indication as to how the words might be organized into syntactic constituents. This suggests that 18-month-olds pay attention to more than strings of words and take into account the hierarchical syntactic structure of sentences during language comprehension.

To conclude, this study shows that 18-month-old toddlers have gained a thorough understanding regarding the contexts dedicated to nouns and verbs. This knowledge is sufficiently detailed to allow them to compute the syntactic category of the content word following an ambiguous function word. Despite the overwhelming evidence for the local co-occurrence of 'la' and nouns in their input, children – at least by 18 months of age – do not solely rely on the computation of these simple distributional patterns, but also take into account the broader sentence context to deduce the syntactic category of the upcoming word. Toddlers might eventually use this ability to infer the syntactic category of novel words, and constrain their possible meanings. Thus, even at an age when children only produce extremely short utterances, their processing of syntactic structure is surprisingly robust.

Acknowledgements

This work was supported by the French Ministry of Research, the French Agence Nationale de la Recherche (grants n° ANR-2010-BLAN-1901, ANR-13-APPR-0012, ANR-11-0001-02 PSL* and ANR-10-LABX-0087), the Fondation de France (grant n° 2012-00033704), as well as by the Région Ile-de-France (grant n° 2009 IF-09-2095/R), a postdoctoral grant from the Fondation Fyssen to M.vH., and the European Research Council 269502, which supported P.B. while writing this manuscript. We thank L. Legros for help with data collection and V. Ul and M. Dutat for technical help, as well as all children and their families for their participation.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.neuropsychologia.2016.08.015.

References

Atchley, R.A., Rice, M.L., Betz, S.K., Kwasny, K.M., Sereno, J.A., Jongman, A., 2006. A comparison of semantic and syntactic event related potentials generated by children and adults. Brain Lang. 99 (3), 236–246. http://dx.doi.org/10.1016/j.bandl.2005.08.005.

Bernal, S., Dehaene-Lambertz, G., Millotte, S., Christophe, A., 2010. Two-year-olds compute syntactic structure on-line. Dev. Sci. 12, 69–76, http://dx.doi.org/10.1111/j.1467-7687.2009.00865.x.

Bernal, S., Lidz, J., Millotte, S., Christophe, A., 2007. Syntax constrains the acquisition of

[4] Note that common nouns are obligatorily preceded by an article in French, which would make 'très gentiment girafe...' (*Very kindly giraffe...*) ungrammatical. But proper names are bare, and it may happen in children stories that an animal is named after its kind (e.g. '*very kindly, Giraffe...*'); under that reading the sentence would then be grammatical.

verb meaning. Lang. Learn. Dev. 3, 325–341, http://dx.doi.org/10.1080/15475440701542609.
Brusini, P., Amsili, P., Chemla, E., Christophe, A., 2011. Learning to categorize nouns and verbs on the basis of a few known examples: a computational model relying on 2-word contexts. Paper presented at the Society for Research on Child Development Biennial Meeting, Montreal (Canada).
Brusini, P., Dehaene-Lambertz, G., Dutat, M., Goffinet, F., Christophe, A., 2016. ERP evidence for on-line syntactic computations in 2-year-olds. Dev. Cogn. Neurosci. 19, 164–173, http://dx.doi.org/10.1016/j.dcn.2016.02.009.
Cauvet, E., Limissuri, R., Millotte, S., Skoruppa, K., Cabrol, D., Christophe, A., 2014. Function words constrain on-line recognition of verbs and nouns in French 18-month-olds. Lang. Learn. Dev. 10, 1–18. http://dx.doi.org/10.1080/15475441.2012.757970.
Christophe, A., Millotte, S., Bernal, S., Lidz, J., 2008. Bootstrapping lexical and syntactic acquisition. Lang. Speech 51, 61–75. http://dx.doi.org/10.1177/00238309080510010501.
Chu, C.J., Leahy, J., Pathmanathan, J., Kramer, M.A., Cash, S.S., 2014. The maturation of cortical sleep rhythms and networks over early development. Clin. Neurophysiol. 125 (7), 1360–1370. http://dx.doi.org/10.1016/j.clinph.2013.11.028.
Dautriche, I., Fibla, L., Christophe, A., 2015a. Learning homophones: syntactic and semantic context matters. Paper presented at the Boston University Conference on Language Acquisition, Boston (USA).
Dautriche, I., Swingley, D., Christophe, A., 2015b. Learning novel phonological neighbors: syntactic category matters. Cognition 143, 77–86. http://dx.doi.org/10.1016/j.cognition.2015.06.003.
de Carvalho, A., Dautriche, I., Christophe, A., 2014. Phrasal prosody constrains online syntactic analysis in two-year-old children. Paper presented at the Boston University Conference on Language Acquisition, Boston (USA).
de Carvalho, A., Dautriche, I., Christophe, A., 2016a. Preschoolers use phrasal prosody online to constrain syntactic analysis. Dev. Sci. 19, 235–250. http://dx.doi.org/10.1111/desc.12300.
de Carvalho, A., He, A.X., Lidz, J., Christophe, A., 2015. 18-month-olds use phrasal prosody as a cue to constrain the acquisition of novel word meanings. Paper presented at the Boston University Conference on Language Acquisition, Boston (USA).
de Carvalho, A., Lidz, J., Tieu, L., Bleam, T., Christophe, A., 2016b. English-speaking preschoolers can use phrasal prosody for syntactic parsing. J. Acoust. Soc. Am. 139 (6), EL216–EL222. http://dx.doi.org/10.1121/1.4954385.
Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134 (1), 9–21. http://dx.doi.org/10.1016/j.jneumeth.2003.10.009.
Demuth, K., Tremblay, A., 2008. Prosodically-conditioned variability in children's production of French determiners. J. Child Lang. 35 (1), 99–127. http://dx.doi.org/10.1017/s0305000907008276.
Ferreira, F., Patson, N.D., 2007. The "Good Enough" approach to language comprehension. Lang. Linguist. Compass 1 (1–2), 71–83. http://dx.doi.org/10.1111/j.1749-818X.2007.00007.x.
Friederici, A.D., 2011. The brain basis of language processing: from structure to function. Physiol. Rev. 91 (4), 1357–1392. http://dx.doi.org/10.1152/physrev.00006.2011.
Gebhart, A.L., Newport, E.L., Aslin, R.N., 2009. Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. Psychon. Bull. Rev. 16 (3), 486–490. http://dx.doi.org/10.3758/pbr.16.3.486.
Gerken, L., Landau, B., Remez, R.E., 1990. Function morphemes in young children's speech perception and production. Dev. Psychol. 26, 204–216. http://dx.doi.org/10.1037/0012-1649.26.2.204.
Gerken, L., McIntosh, B., 1993. Interplay of function morphemes and prosody in early language. Dev. Psychol. 29, 448–457. http://dx.doi.org/10.1037/0012-1649.29.3.448.
Gervain, J., Nespor, M., Mazuka, R., Horie, R., Mehler, J., 2008. Bootstrapping word order in prelexical infants: a Japanese-Italian cross-linguistic study. Cogn. Psychol. 57 (1), 56–74. http://dx.doi.org/10.1016/j.cogpsych.2007.12.001.
Gomez, R.L., Gerken, L., 1999. Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. Cognition 70 (2), 109–135. http://dx.doi.org/10.1016/s0010-0277(99)00003-7.
Hagoort, P., 2005. On Broca, brain, and binding: a new framework. Trends Cognit. Sci. 9 (9), 416–423. http://dx.doi.org/10.1016/j.tics.2006.07.004.
Hallé, P., Durand, C., de Boysson-Bardies, B., 2008. Do 11-month-old French infants process articles? Lang. Speech 51, 45–66. http://dx.doi.org/10.1177/00238309080510010301.
Hochmann, J.-R., 2013. Word frequency, function words and the second gavagai problem. Cognition 128 (1), 13–25. http://dx.doi.org/10.1016/j.cognition.2013.02.014.
Hochmann, J.-R., Endress, A.D., Mehler, J., 2010. Word frequency as a cue for identifying function words in infancy. Cognition 115, 444–457. http://dx.doi.org/10.1016/j.cognition.2010.03.006.
Hohle, B., Weissenborn, E., Kiefer, D., Schulz, A., Schmitz, M., 2004. Functional elements in infants' speech processing: the role of determiners in the syntactic categorization of lexical elements. Infancy 5 (3), 341–353. http://dx.doi.org/10.1207/s15327078in0503_5.
Holcomb, P.J., Coffey, S.A., Neville, H.J., 1992. Visual and auditory sentence processing - a developmental analysis using event-related brain potentials. Dev. Neuropsychol. 8 (2–3), 203–241. http://dx.doi.org/10.1080/87565649209540525.
Kedar, Y., Casasola, M., Lust, B., 2006. Getting there faster: 18-and 24-month-old infants' use of function words to determine reference. Child Dev. 77 (2), 325–338. http://dx.doi.org/10.1111/j.1467-8624.2006.00873.x.
Kiebel, S.J., Friston, K.J., 2004. Statistical parametric mapping for event-related potentials (II): a hierarchical temporal model. Neuroimage 22 (2), 503–520. http://dx.doi.org/10.1016/j.neuroimage.2004.02.013.
Kjelgaard, M.M., Speer, S.R., 1999. Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. J. Mem. Lang. 40, 153–194. http://dx.doi.org/10.1006/jmla.1998.2620.
Kouider, S., Stahlhut, C., Gelskov, S.V., Barbosa, L.S., Dutat, M., de Gardelle, V., Dehaene-Lambertz, G., 2013. A neural marker of perceptual consciousness in infants. Science 340 (6130), 376–380. http://dx.doi.org/10.1126/science.1232509.
Kuperberg, G.R., 2007. Neural mechanisms of language comprehension: Challenges to syntax. Brain Res. 1146 (23–49). http://dx.doi.org/10.1016/j.brainres.2006.12.063.
Lewkowicz, D.J., Hansen-Tift, A.M., 2012. Infants deploy selective attention to the mouth of a talking face when learning speech. Proc. Natl. Acad. Sci. USA 109 (5), 1431–1436. http://dx.doi.org/10.1073/pnas.1114783109.
Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164 (1), 177–190. http://dx.doi.org/10.1016/j.jneumeth.2007.03.024.
Mensen, A., Khatami, R., 2013. Advanced EEG analysis using threshold-free cluster-enhancement and non-parametric statistics. Neuroimage 67, 111–118. http://dx.doi.org/10.1016/j.neuroimage.2012.10.027.
Millotte, S., René, A., Wales, R., Christophe, A., 2008. Phonological phrase boundaries constrain on-line syntactic analysis. J. Exp. Psychol.: Learn. Mem. Cogn. 34, 874–885. http://dx.doi.org/10.1037/0278-7393.34.4.874.
Millotte, S., Wales, R., Christophe, A., 2007. Phrasal prosody disambiguates syntax. Lang. Cogn. Process. 22, 898–909. http://dx.doi.org/10.1080/01690960701205286.
Oberecker, R., Friederici, A.D., 2006. Syntactic event-related potential components in 24-month-olds' sentence comprehension. Neuroreport 17 (10), 1017–1021. http://dx.doi.org/10.1097/01.wnr.0000223397.12694.9a.
Oberecker, R., Friedrich, M., Friederici, A.D., 2005. Neural correlates of syntactic processing in two-year-olds. J. Cogn. Neurosci. 17, 1667–1678. http://dx.doi.org/10.1162/089892905774597236.
Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011. http://dx.doi.org/10.1155/2011/156869.
Oshima-Takane, Y., Ariyama, J., Kobayashi, T., Katerelos, M., Poulin-Dubois, D., 2011. Early verb learning in 20-month-old Japanese-speaking children. J. Child Lang. 38 (3), 455–484. http://dx.doi.org/10.1017/S0305000910000127.
Pernet, C.R., Chauveau, N., Gaspar, C., Rousselet, G.A., 2011. LIMO EEG: a toolbox for hierarchical LInear MOdeling of ElectroEncephaloGraphic data. Comput. Intell. Neurosci. 2011. http://dx.doi.org/10.1155/2011/831409.
Redington, M., Chater, N., Finch, S., 1998. Distributional information: a powerful cue for acquiring syntactic categories. Cogn. Sci. 22 (4), 425–469. http://dx.doi.org/10.1207/s15516709cog2204_2.
Schipke, C.S., Friederici, A.D., Oberecker, R., 2011. Brain responses to case-marking violations in German preschool children. Neuroreport 22, 850–854. http://dx.doi.org/10.1097/WNR.0b013e32834c1578.
Schipke, C.S., Knoll, L.J., Friederici, A.D., Oberecker, R., 2012. Preschool children's interpretation of object-initial sentences: neural correlates of their behavioral performance. Dev. Sci. 15 (6), 762–774. http://dx.doi.org/10.1111/j.1467-7687.2012.01167.x.
Shady, M., 1996. Infant's Sensitivity to Function Morphemes (Ph.D Thesis). State University of New York, Buffalo.
Shafer, V.L., Shucard, D.W., Shucard, J.L., Gerken, L., 1998. An electrophysiological study of infants' sensitivity to the sound patterns of English speech. J. Speech Lang. Hear. Res. 41, 874–886.
Shi, R., Cutler, A., Werker, J., Cruickshank, M., 2006a. Frequency and form as determinants of functor sensitivity in English-acquiring infants. J. Acoust. Soc. Am. 119, EL61–EL66. http://dx.doi.org/10.1121/1.2198947.
Shi, R., Gauthier, B., 2005. Recognition of function words in 8-month-old French-learning infants. J. Acoust. Soc. Am. 117, 2426–2427. http://dx.doi.org/10.1121/1.4786583.
Shi, R., Melançon, A., 2010. Syntactic categorization in French-learning infants. Infancy 15, 517–533. http://dx.doi.org/10.1111/j.1532-7078.2009.00022.x.
Shi, R., Morgan, J.L., Allopenna, P., 1998. Phonological and acoustic bases for earliest grammatical category assignment: a cross-linguistic perspective. J. Child Lang. 25 (1), 169–201. http://dx.doi.org/10.1017/s0305000997003395.
Shi, R., Werker, J., 2001. Six-months old infants' preference for lexical words. Psychol. Sci. 12, 71–76. http://dx.doi.org/10.1111/1467-9280.00312.
Shi, R., Werker, J., Cutler, A., 2006b. Recognition and representation of function words in English-learning infants. Infancy 10, 187–198. http://dx.doi.org/10.1207/s15327078in1002_5.
Shi, R., Werker, J., Morgan, J., 1999. Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. Cognition 72, B11–B21. http://dx.doi.org/10.1016/S0010-0277(99)00047-5.
Silva-Pereyra, J., Klarman, L., Lin, L.J.F., Kuhl, P., 2005a. Sentence processing in 30-month-old children: an event-related potential study. Neuroreport 16 (6), 645–648.
Silva-Pereyra, J., Rivera-Gaxiola, M., Kuhl, P.K., 2005b. An event-related brain potential study of sentence comprehension in preschoolers: semantic and morphosyntactic processing. Cogn. Brain Res. 23, 247–258. http://dx.doi.org/10.1016/j.cogbrainres.2004.10.015.
Udden, J., Ingvar, M., Hagoort, P., Petersson, K.M., 2012. Implicit acquisition of grammars with crossed and nested non-adjacent dependencies: investigating the push-down stack model. Cogn. Sci. 36 (6), 1078–1101. http://dx.doi.org/10.1111/j.1551–6709.2012.01235.x.
Van Heugten, M., Johnson, E.K., 2011. Gender-marked determiners help Dutch learners' word recognition when gender information itself does not. J. Child Lang. 38 (1),

87–100. http://dx.doi.org/10.1017/s0305000909990146.
Veneziano, E., Parisse, C., 2011. Retrieving the meaning of words from syntactic cues: a comprehension study. Paper presented at the International Conference on the Study of Child Language, (Montreal, Canada).
Waxman, S.R., Lidz, J.L., Braun, I.E., Lavin, T., 2009. Twenty four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. Cogn. Psychol. 59 (1), 67–95. http://dx.doi.org/10.1016/j.cogpsych.2009.02.001.
Zangl, R., Fernald, A., 2007. Increasing flexibility in children's online processing of grammatical and nonce determiners in fluent speech. Lang. Learn. Dev. 3, 199–231. http://dx.doi.org/10.1080/15475440701360564.

Résumé

Des études précédentes démontrent qu'avoir accès à la structure syntaxique des phrases aide les enfants à découvrir le sens des mots nouveaux. Cela implique que les enfants doivent avoir accès à certains aspects de la structure syntaxique avant même de connaître beaucoup de mots. Étant donné que dans toutes les langues du monde la structure prosodique d'une phrase corrèle avec sa structure syntaxique, et que par ailleurs les mots et morphèmes grammaticaux sont utiles pour déterminer la catégorie syntaxique des mots, il se pourrait que les enfants utilisent la prosodie et les mots grammaticaux pour initialiser leur acquisition lexicale et syntaxique. Dans cette thèse, j'ai étudié le rôle de la prosodie phrasale et des mots grammaticaux pour guider l'analyse syntaxique chez les enfants (PARTIE 1) et la possibilité que les jeunes enfants exploitent cette information pour apprendre le sens des mots nouveaux (PARTIE 2).

Dans la partie 1, j'ai construit des paires minimales de phrases en français et en anglais afin de tester si les enfants exploitent la relation entre les structures prosodique et syntaxique pour guider leur interprétation des homophones noms-verbes. J'ai démontré que les enfants d'âge préscolaire utilisent la prosodie phrasale en temps réel pour guider leur analyse syntaxique. En écoutant des phrases telles que [La petite ferme][.., les enfants interprètent ferme comme un nom, mais pour les phrases telles que [La petite][ferme...], ils interprètent ferme comme un verbe (Chapitre 3). Cette capacité a également été observée chez les enfants américains: en écoutant des phrases telles que «The baby flies... », ils utilisent la prosodie des phrases pour décider si flies est un nom ou un verbe (Chapitre 4). Par la suite, j'ai démontré que même les enfants d'environ 20 mois utilisent la prosodie des phrases pour récupérer leur structure syntaxique et pour en déduire la catégorie syntaxique des mots (Chapitre 5), une capacité qui serait extrêmement utile pour découvrir le sens des mots inconnus.

C'est cette hypothèse que j'ai testé dans la partie 2, à savoir si l'information syntaxique obtenue à partir de la prosodie phrasale et des mots grammaticaux permet aux enfants d'apprendre le sens des mots. Une première série d'études s'appuie sur des phrases disloquées à droite contenant un verbe nouveau en français: [ili dase], [le bébéi] qui est minimalement différente de la phrase transitive [il dase le bébé]. Mes résultats montrent que les enfants de 28 mois exploitent les informations prosodiques de ces phrases pour contraindre leur interprétation du sens du nouveau verbe (Chapitre 6). Dans une deuxième série d'études, j'ai étudié si la prosodie et les mots grammaticaux guident l'acquisition de noms et de verbes. J'ai utilisé des phrases comme «Regarde la petite bamoule» qui peuvent être produites soit comme [Regarde la petite bamoule!], où «bamoule» est un nom, ou [Regarde], [la petite] [bamoule!], où bamoule est un verbe. Les enfants de 18 mois ont correctement analysé ces phrases et ont attribué une interprétation de nom ou de verbe au mot bamoule selon sa position dans la structure prosodique-syntaxique des phrases (Chapitre 7).

Ensemble, ces études montrent que les jeunes enfants exploitent les mots grammaticaux et la structure prosodique des phrases pour inférer la structure syntaxique et contraindre ainsi l'interprétation possible du sens des mots. Ce mécanisme peut permettre aux enfants de construire une représentation initiale de la structure syntaxique des phrases, avant même de connaître la signification des mots. Bien que les informations prosodiques et les mots grammaticaux puissent prendre des formes différentes selon les langues, nos études suggèrent que cette information pourrait représenter un outil universel et qui permettrait aux enfants d'accéder à certaines informations syntaxiques des phrases qu'ils entendent, et d'initialiser l'acquisition du langage.

Mots Clés

Prosodie phrasale ; mots grammaticaux ; acquisition du langage ; acquisition de la syntaxe ; acquisition du lexique

Abstract

Previous research demonstrates that having access to the syntactic structure of sentences helps children to discover the meaning of novel words. This implies that infants need to get access to aspects of syntactic structure before they know many words. Since in all the world's languages the prosodic structure of a sentence correlates with its syntactic structure, and since function words/morphemes are useful to determine the syntactic category of words, infants might use phrasal prosody and function words to bootstrap their way into lexical and syntactic acquisition. In this thesis, I empirically investigated the role of phrasal prosody and function words to constrain syntactic analysis in young children (PART 1) and whether infants exploit this information to learn the meanings of novel words (PART 2).

In part 1, I constructed minimal pairs of sentences in French and in English, testing whether children exploit the relationship between syntactic and prosodic structures to drive their interpretation of noun-verb homophones. I demonstrated that preschoolers use phrasal prosody online to constrain their syntactic analysis. When listening to French sentences such as [La petite ferme][…–[The little farm][…, children interpreted ferme as a noun, but in sentences such as [La petite][ferme…] – [The little girl][closes…, they interpreted ferme as a verb (Chapter 3). This ability was also attested in English-learning preschoolers who listened to sentences such as 'The baby flies…': they used prosodic information to decide whether "flies" was a noun or a verb (Chapter 4). Importantly, in further studies I demonstrated that even infants around 20-months use phrasal prosody to recover syntactic structures and to predict the syntactic category of upcoming words (Chapter 5), an ability which would be extremely useful to discover the meaning of unknown words.

This is what I tested in part 2: whether the syntactic information obtained from phrasal prosody and function words could allow infants to constrain their acquisition of word meanings. A first series of studies relied on right-dislocated sentences containing a novel verb in French: [ili dase], [le bébéi] - 'hei is dasing, the babyi' (meaning 'the baby is dasing') which is minimally different from the transitive sentence [il dase le bébé] (he is dasing the baby). 28-month-olds were shown to exploit prosodic information to constrain their interpretation of the novel verb meaning (Chapter 6). In a second series of studies, I investigated whether phrasal prosody and function words constrain the acquisition of nouns and verbs. I used sentences like 'Regarde la petite bamoule', which can be produced either as [Regarde la petite bamoule!] - Look at the little bamoule!, where 'bamoule' is a noun, or as [Regarde], [la petite] [bamoule!] - Look, the little (one) is bamouling, where bamoule is a verb. 18-month-olds correctly parsed such sentences and attributed a noun or verb meaning to the critical word depending on its position within the syntactic-prosodic structure of the sentences (Chapter 7).

Taken together, these studies show that infants exploit function words and the prosodic structure of an utterance to recover the sentences' syntactic structure, which in turn constrains the possible meaning of novel words. This powerful mechanism might be extremely useful for infants to construct a first-pass syntactic structure of spoken sentences even before they know the meanings of many words. Although prosodic information and functional elements can surface differently across languages, our studies suggest that this information may represent a universal and extremely useful tool for infants to access syntactic information through a surface analysis of the speech stream, and to bootstrap their way into language acquisition.

Keywords

Phrasal prosody; function words; language acquisition; syntactic acquisition; lexical development