Conflict resolution between semantic and syntactic cues in language acquisition

Marion Beretti[1], Naomi Havron[1], Anne Christophe[1]

[1] Laboratoire de sciences cognitives et de psycholinguistique, ENS, EHESS, CNRS, PSL University

Running title: Conflict resolution language acquisition

Conflict of Interest Statement: The authors declare no conflict of interests

Acknowledgements: This work was supported by the French Agence Nationale de la Recherche (grants ANR-13-APPR-0012 LangLearn, ANR-17-CE28-0007-01 LangAge, and ANR-17-EURE-0017 FrontCog); a postdoctoral grant to Naomi Havron from the French Embassy in Israel and the Victor Smorgon Charitable Fund; a CIF grant to Marion Beretti from the Fongecif Ile De France. We thank the school and children for their participation.

Conflict resolution between semantic and syntactic cues in language acquisition

Research highlights:

- Children and adults use the reliability of different sources of information when inferring the meaning of ambiguous sentences
- Children and adults can adapt their semantic knowledge and learn novel homophone meanings when they consider syntactic cues to be reliable enough to do so.
- Children and adults use the reliability of different sources of information to adjust the weight of each source, for both processing and learning novel words.

Abstract

A central challenge in language acquisition is the integration of multiple sources of information, potentially in conflict, to acquire new knowledge and adjust current linguistic representations. One way to accomplish this is to assign more weight to more reliable sources of information in context. We test the hypothesis that children will adjust the weight of different sources of information during learning, considering two specific sources of information: semantic and syntactic. 4-5-year-old French children listened to sentences which were in conflict between these two cues, and were asked to choose between two videos that illustrate the sentence. One video presented the reasonable choice if the sentence is assumed to be syntactically correct, but familiar words refer to novel things (e.g., *une mange* - "an eat" describes a novel object). The other video was the reasonable choice if the familiar meanings of the words are favored and the sentence is assumed to be syntactically *in*correct – using a familiar noun as a verb or a familiar verb as a noun (e.g., *une mange* - "an eat" describes a video of a girl eating). We varied the reliability of sources of information through an induction phase (reliable-syntax or reliable-semantics). The proportion of syntactic choice (e.g., interpreting *mange* - "eat", as a novel noun) was found to be higher in the reliable-syntax condition than in the reliable-semantics condition. This shows that children can adapt to the current context to integrate local information, and use this adaptation to acquire new knowledge (in this case, novel homophones).

Keywords: *language acquisition, adaptation, reliability, conflict resolution, syntactic bootstrapping, cue integration*

Introduction

From an early age, infants and young children need to rapidly integrate different cues, directly from the context they are immersed in, to construct their linguistic representations. To acquire new words, they have to choose between different referent possibilities. Simultaneously, they need to learn the syntax and the inventory of the sounds of their language. Each of these elements needs to be acquired, but also to be used as a cue in itself. This presents a type of a chicken and egg problem. Children can hardly wait until they have acquired all their lexicon to deduce syntactic categories from that lexicon, or wait until they know all their native syntax before they start acquiring and categorizing vocabulary items. How could an infant choose when she needs to learn about syntax from lexical items, and when she needs to learn about lexical items from syntax? One reasonable way to make this decision is to assign more weight to the more reliable source of information, in context. Instead of assuming that syntax is always more reliable than lexical items, or that lexical items are always more reliable than syntax – an infant could take into account information about the reliability of the context.

While such a model has not yet been applied to language acquisition, it has been used to explain successful language processing. The *noisy channel* model proposed by Gibson, Bergen, & Piantadosi (2013) applies this idea to situations where multiple interpretations of a given sentence are possible, depending on which cue is weighted more strongly. The Bayesian probability of any one of two interpretations given a specific cue, can be deduced from the probability of the cue itself, and the probability of the other interpretations. Changing the probability of one piece of information, then affects the relative weight of all others. Gibson and colleagues (2013) tested this by having adults listen to a sentence such as "*The ball kicked the girl*" and answer a yes-or-no comprehension question (e.g. "*Did the girl kick something/someone?*"). In one condition, syntax errors were inserted into filler items in order to reduce the reliability of the literal meaning (e.g., function words were deleted or added); in the other condition, semantically implausible items were inserted into the fillers to limit semantic reliability. When the syntax of the speaker appeared noisy overall, adults preferred to correct the test sentence by choosing the most plausible meaning (e.g. answer "*yes*" to "*Did the girl kick something/someone?*"). When implausible items were common, adults preferred to preserve the literal meaning (e.g. answer "*no*" to "*Did the girl kick something/someone?*"). Thus, the *noisy channel* model allows to integrate multiple sources of information by prioritizing each one depending on its reliability in context.

This model has also been successfully applied to children's language processing. In a study by Yurovsky, Case, & Frank (2017), 4-5-year-old children listened to ambiguous sentences, where the literal phonetic information and the semantically plausible information were in conflict (e.g., "*I had carrots and bees for dinner*"). The phonetic proximity between the presented word ("*bees*") and the most plausible word for this context ("*peas*"), made both interpretations possible. An induction phase manipulated the reliability of the two sources of information, phonetic and semantic: either by increasing the ambient noise to disadvantage the literal meaning, or by increasing the number of implausible sentences to disadvantage the semantic information. On test trials, the child had to choose between two images representing the two interpretations. Children in the phonetically-noisy condition preferred the semantically plausible choice (assuming the speaker has meant the plausible alternative "*peas*"), and replaced the word by its plausible alternative; but children who were exposed to speakers who produced semantically implausible utterances were more likely to adopt the literal interpretation (assuming the speaker has really meant "*bees*"). These results seem to indicate that children's preference for one source of information over the other, in comprehension, is dependent on the reliability of each cue in the context rather than on a default weighting scheme (with phonetics always weighted higher than semantic plausibility, or vice-versa). We ask whether these results also apply to language acquisition, when children are learning new words or building syntactic representations, rather than processing familiar sentences.

As mentioned above, for children, different sources of information (e.g., syntax, the lexicon) which are involved in language acquisition are also in the process of being learnt, so that they must help refine each other: children use each source of information to both understand sentences, and to learn other aspects of language, while refining and replacing the representations they currently use. For example, on the one hand, syntax can help learn word meanings: according to the *syntactic bootstrapping* hypothesis (Gleitman, 1990), children use syntactic cues to reduce the set of possible meanings for novel words and thus acquire new semantic knowledge. Thus, if they hear an unknown target-word like "*dax*" preceded by a pronoun as in "*he daxes*", they can deduce that "*dax*" is a verb and refers to an action. Similarly, if the same word is encountered preceded by an article as in "*the dax*", they will think that "*dax*" is a noun and refers to an object (for experimental evidence from 23-month-old children, see Bernal, Lidz, Millotte, & Christophe, 2007, and Waxman, Lidz, Braun, & Lavin, 2009). Even at 14-18-month-old, infants are already expecting familiar nouns to appear in frequent noun contexts, and verbs to appear in frequent verb contexts, but not vice-verse

(Brusini et al., 2017; Cauvet et al., 2014; Kedar, Casasola, Lust, & Parmet, 2017; Shi & Melançon, 2010). On the other hand, before it can help to acquire lexical items, syntax itself has to be acquired. Since syntax can be thought of as allowing to organize the different semantic categories in the sentence, it is only logical that knowing some basic words is needed in order to learn it. Thus, children have to be able to prioritize, at least temporarily, one source of information over others, in order to then adapt their knowledge about these other sources, when they appear to be in conflict.

Two prominent sources of information which have been discussed in the literature as playing an important role in language learning are semantic and syntactic information: some have proposed a default prioritization of bottom-up lexical cues compared to syntactic ones, which would be involved only when semantic cues are insufficient or ambiguous. Other researchers side with prediction-error models, where the weight of each source varies depending on its reliability in the context. There is currently little evidence to discriminate between these two suggestions in the literature on adults and young children (as discussed by Federmeier, 2007; Rabagliati, Pylkkänen, & Marcus, 2012). It has, however, been shown that children are capable of evaluating the relevance of the syntactic source of information, or of the semantic source of information, if each one is tested separately. For example, when comparing an induction condition where a syntactic cue is reliable, to another where the same cue is not, children can stop using syntactic cues for prediction (van Heugten, & Christophe, 2014). Children can even ignore the meaning of a familiar word, to learn a novel meaning for it (therefore creating a homophone), in a case in which the familiar meaning is inappropriate, for either syntactic or semantic reasons (i.e., if children hear *"it's an eat"* and see a stuffed animal, they can learn that it is *an eat*; Dautriche, Fibla, Fiévet, & Christophe, 2018). The goal of the current study is then to understand if children are able to learn novel meanings and adapt their semantic and syntactic knowledge, when the sources of information to integrate (semantics and syntax) are in conflict. Previous studies have never put these two cues in conflict – children would ignore the syntax when it was unreliable, or ignore the familiar word meaning if it was unlikely, but never did they have syntax provide one answer and lexicon a different one. When faced with a conflict, if children adapt to the reliability of different cues in context, this would support a prediction-error model of word learning, whereby children can adjust the weight of different cues in context. If they cannot, then this would mean that one source of information (either syntax or semantics) is inherently stronger in children's inferences, regardless of the current context, and children show little flexibility in their learning.

We follow a paradigm similar to the one used in Yurovsky et al. (2017) to test the *noisy channel* hypothesis in language acquisition. To examine how children adapt either one or the other of their prior sources of knowledge (semantics or syntax) - we place these two sources of information in conflict. We compare two conditions, which vary in terms of which source of information is reliable. Children are presented with an induction phase, where either syntax is reliable (*reliable-syntax condition*) or semantics is reliable (*reliable-semantics condition*). In both conditions, children hear the same sentence (e.g., *elle pomme* - "she apple(s)"). In the reliable-syntax condition, children see a novel action or object, which does not have a name in French, and are thus led to believe that the speaker is using a familiar word form to refer to something the participant has not seen before (in the example *pomme* is a novel verb, that happens to sound like the noun *pomme* – it is a homophone). In the reliable-semantics condition children hear the same sentence. Now there is no novel action or object, but instead the person seems to be using the syntax incorrectly. In the example above, there would be a video of a girl holding an apple, and the speaker can then be assumed to have made a syntactic error, and to have meant "it is an apple" and not "she apples". Thus, in the induction phase, the cues are not in conflict, since the presented videos make only one interpretation of the sentence possible.

In the test phase, participants hear a short sentence composed of a syntactic and a semantic cue that are in conflict (e.g. *elle balle* - "she ball(s)"), and they have to choose between two videos corresponding either to the syntactic cue (a girl performing a novel action) or to the semantic cue (a ball). When you hear an utterance such as "she balls", if you think that this person's syntax is impeccable, but that she tends to use novel words, then you will be tempted to interpret "ball" as a verb you did not know yet, a homophone of the noun "ball". In that case, the correct referent of the utterance would be the girl performing the novel action. In contrast, if you think that this person's syntax is uncertain (because she has been employing known words in incorrect syntactic contexts before), while her use of familiar words is always perfectly appropriate, then you will tend to rely on your recognition of the content word "ball", and deprioritize its syntactic context ("she"), which would lead you to choose the picture of the girl holding a ball as the correct referent of the utterance. Thus, while the test sentences are constructed in the same way as the induction sentences, it is the choice between two reasonable options (the syntactic one and the semantic one) that now makes the sentences ambiguous, and the cues in conflict.

We tested whether 4-5-year-old children and adults will be capable of using the reliability of one source of information compared to the other in order to interpret these

sentences, or whether they preserve a default weight for each source of information, regardless of its reliability (either always follow the syntax or always follow the semantics). If participants have adapted to the reliability of the sources of the information, they will choose the syntactic choice more often in the reliable-syntax condition than in the reliable-semantics condition (either explicitly by pointing, or implicitly by looking at the video). Since the decision to follow the syntactic choice in the reliable-syntax condition forces them to learn novel words (e.g., a novel verb *baller* - "to ball"), then if participants in the reliable-syntax condition choose the syntactic choice more than participants in the reliable-semantics condition, they will be shown not only to be able to rely on the reliability of a source for comprehension, but to also rely on their adapted knowledge to learn novel words (even when this includes the creation of homophones).

Method

All data and materials have been made publicly available via the Open Science Framework (OSF) and can be accessed at:
https://osf.io/fsb92/?view_only=000d7f4b94d64428a29aff67dd9ce583.
The design and analysis plans were preregistered at the OSF and can also be accessed with the same link.

Experimental Design

The experiment is based on exposure to ungrammatical sentences, like *elle balle* - "she ball(s)" or *une parle* - "a speak". This forces participants to resolve a conflict: either syntax has been used incorrectly (if one is talking about the round object, one should have said *une balle* - "a ball"), or lexical items are used unexpectedly (one does not want to talk about the round object, but about a novel action *baller* - "to ball"). To observe which interpretation participants inferred, we present two videos: one illustrates the semantic choice (where the known lexical item "ball" is shown in the video), and the other the syntactic choice (where a novel action provides a good fit for the novel content word). We ask them to point to the video corresponding to the sentence they heard (see figure 1).

To test participants' expectation adaptation to the reliability of different cues, we use induction items to weaken their confidence in the way the speaker is using either syntax or semantics, depending on the experimental condition (see figure 2). Thus, in the reliable-semantics condition, induction items encourage participants to follow the semantics of the

content word, and to ignore the preceding function word. For example, when hearing *elle pomme* - "she apple(s)", they are given a choice between an apple and a distractor, i.e. a familiar action like read, which is not a good candidate for the meaning of a novel verb *pommer* - "to apple", since the action of reading already has a familiar label in French (*mutual exclusivity*, Markman & Wachtel, 1988). In the reliable-syntax condition, participants are encouraged to trust the function word to assign a new syntactic category to the following content word, while ignoring the fact that the form of the content word already has a meaning in French. For example, when hearing *elle pomme* - "she apple(s)", they are faced with a novel action (a good candidate for novel verb *pommer* - "to apple"), and a shoe (which has nothing to do with an apple) as a distractor. In this condition, participants are encouraged to create a new homophone for the familiar word (*pomme* - "apple") – a word that would come from a different syntactic category (here, a verb). Infants as young as 20 months are able to create homophones (Dautriche et al. 2018) when taught a novel word from a different syntactic category (e.g., they can learn that a novel noun "a give" represents a funny animal, even though they already know the verb "to give"). Note that in Dautriche's experiment, there was only a furry animal present as a potential candidate for "a give" – there was never also someone giving something as in our test phase. We use a *between-subject* design, with each participant randomly assigned to one of these conditions. Each participant hears examples of both nouns and verbs, in a counterbalanced manner (e.g. *elle balle* - "she ball(s)" for verb syntactic cue and noun semantic cue; *une parle* - "a speak" for noun syntactic cue and verb semantic cue).

Participants

57 middle to high SES 4-5-year-old children (25 girls, 32 boys, M age = 57.36 months, range = 48.20-66.61 months), and 51 adults (18 women, 33 men) were randomly assigned to the two conditions. Children were tested either in their kindergarten or in our laboratory. Adults were tested either in a high-tech company where they work or in our laboratory. Eyetracking data was filtered with the following criteria: we excluded trials with more than 25% of missing data, and participants with more than 50% of excluded trials (10 children and 14 adults excluded). We kept pointing data for these trials/participants leaving us with 29 children in the reliable-semantics condition (24 also have eyetracking data) and 28 in the reliable-syntax condition (23 also have eyetracking data); and 26 adults in the reliable-semantics condition (15 also have eyetracking data) and 25 in the reliable-syntax condition

(22 also have eyetracking data). In addition, 4 additional children were tested but their data was not included in the final analyses, because they were bilingual (2), fussy (1) or had a psychological disorder (1). Data of one additional adult was excluded because he was bilingual.

Materials

Sentences and videos. All sentences are composed of two words presenting a conflict between syntax and semantics (e.g. *elle balle* - "she ball(s)" or *une parle* - "a speak"): (1) A function word for the syntactic cue (i.e. article *une* - "a" for nouns or pronoun *elle* - "she" for verbs). (2) A familiar content word for the semantic cue.

Stimuli were videos, always paired to present an action and an object on opposite sides of the screen. Objects and persons were all feminine to avoid the additional complexity of grammatical gender. For actions, the video presents a woman performing the action; for objects, the same woman is seen holding an object in her hands while looking at it and at the camera alternately - to encourage participants to look at it instead of at herself. All target actions corresponded to intransitive verbs and all objects were inanimate. Both verbs and nouns were selected from Lexique.org (New, Pallier, Ferrand, and Matos, 2001) and Wordbank (Frank, Braginsky, Yurovsky, & Marchman, 2017) with the criteria that they were well-known to four-year-old children, and were performable on a video for verbs, and available and clear from video for nouns. Since the number of available items was restricted, we sometimes used actions which included objects (e.g. reading, which included a book), but only for distractors. Labels associated with novel words had to be plausible homophones. For this reason, we chose monosyllabic words belonging to a dense phonological neighborhood (Piantadosi, Tily, & Gibson, 2012). To contrast the induced choice, distractor videos should not be too close to the original meaning of the word, in order to avoid a possible polysemy with the distractor video. For example, if the distractor for the sentence *elle pomme* - "she apple(s)", had been a woman eating, *elle pomme* could be interpreted as meaning "eating an apple", such that "to apple" and "an apple" would share a core meaning, a case of polysemy. We therefore presented a girl reading as the distractor, so that it would be unreasonable to suspect that *elle pomme* means "she reads". For each video, target words always appeared in the same time window: first onset around 2.4s and second around 7.5s.

A typical trial. For each item, each of the videos is first presented alone for eight seconds, then there is a fixation point and then both videos are presented together for ten more seconds and the critical sentence is played (see fig.3). The target side is random, but with an equal total number of left and right targets for a given participant. At the end of the trial, the experimenter asks the participant to point at the video corresponding to the sentence. Pointing is recorded by the experimenter on the computer, and gaze is recorded for whole duration of the trial. In one-video induction trials, only one video is displayed on the screen with no distractor. This makes sure participants draw the correct conclusions about the meaning of the sentences (depending on the condition they are in, see Appendix for the items), and shortens the overall length of the experiment.

Procedure

Global organization. The experiment was launched from Matlab (Version R2011b - 7.13.0.564, MathWorks, Inc., 2011) on a portable computer, and videos were displayed on a 27'' screen placed 70cm from the participant. An Eyelink II eyetracker (camera: EyeLink CL Version 1.4 Sensor BIB) was positioned under the screen and connected to the computer to record participants' gaze during the whole task.

Following a calibration phase, each participant watched a series of 14 video items (see fig. 4), distributed in the following phases (phase transitions are transparent to the participant): *warm-up* (no conflict but always an object video + an action video), *one-video induction items* (first half of induction items presented with only one video, without a distractor, to reduce the length of the experiment and reinforce learning), *induction* (pairs of target + distractor), *conflict test* (pairs of semantic vs. syntactic interpretations). Conflict-test items were interspersed within the induction items in order to reinforce the induction before the tests. For each item, participants were asked to point to the video they thought corresponded best to the sentence they heard in the headphones. The dependent measures are pointing and proportion of looks towards the syntactic choice – the case where the target word is interpreted as a homophone to a familiar word. If participants are adapting to the reliability of the speaker, then we expect children from the reliable-syntax condition to point more and look more at the syntactic choice, the homophone, than participants in the reliable-semantics condition.

Only *one-video induction items* and *induction* items change between conditions, *warm-up* and *test* items are the same for all participants. All participants in a given condition saw

exactly the same items, the order of which was randomized within each phase. The pairing of the two videos for each sentence and condition was fixed.

Additional items. Our hypothesis about learning situations is mainly tested with the main conflict-test items. This tests whether there is a difference between the group that heard syntactically reliable sentences in the induction phase, and the group that heard semantically reliable sentences in the induction phase, for test items where sentences can be interpreted as either semantically or syntactically reliable. These test items cannot tell us whether participants who did not select the homophone interpretation, that is, did not follow the syntax, were ignoring the syntax entirely, or instead were creating novel syntactic predictions, going against their life-long experience: that nouns will follow pronouns and verbs will follow determiners. Thus, two additional test items, novel-word tests, were added at the end of the task in an exploratory way. The idea was to examine what syntactic knowledge the participants had created or retained at the end of the task (the prior one or a novel adapted one) by presenting sentences like *elle dase* - "she dase(s)" or *une nuve* - "a nuve". Since these are novel words, that could be interpreted either as a noun or a verb, participants who have kept their prior predictions should expect *dase*, which is preceded by a pronoun, to be a novel verb and *nuve*, which is preceded by a determiner, to be a novel noun. If, however, participants have created new predictions based on the distribution of words in the induction phase, then we expect only the reliable-syntax group to entertain these interpretations. Thus, the participants in the reliable-semantics condition should be less likely than the participants in the syntax-reliable condition, to interpret novel words appearing after determiners as nouns and novel words appearing after pronouns as verbs. These two items were tacked at the end of the experiment, so that they did not disrupt the main experiment: they are underpowered to draw firm conclusions, and we only report the results in the supplementary materials.

Statistical Analysis

Statistical analyses were conducted with R software (RStudio 1.0.136), using packages *lme4* and *eyetrackingR* (Dink & Ferguson, 2016), and plotted with *ggplot2* (Wickham, 2016). Only trials from the test phase were considered for analysis. If children are able to adapt to the reliability of a source of information in order to interpret sentences, they will choose the syntactic choice more in the reliable-syntax condition than in the reliable-semantics condition. We also examine whether participants were overall biased towards the semantic or syntactic choice by looking at the intercept of the analyses of their pointing responses.

We ran mixed-effects logistic regression analyses[1] on pointing and eyetracking data (Jaeger, 2008). The first dependent variable was the proportion of pointing towards the syntactic choice. We call syntactic choice the one corresponding to the *elle/une* - "she/a" cue (e.g. the girl performing the novel action for *elle pomme* - "she apple(s)"), both for pointing and eyetracking data. Syntactic and semantic choices are complementary. We had condition (sum coded), age (children vs adults, sum coded) and an interaction between age and condition as independent variables. The second dependent variable was the mean proportion of looks towards the syntactic choice (from after the end of the first time the participants hear the sentence until the end of the trial), with condition, age and interaction between age and condition as independent variables. The models included random intercepts for participant and item, and the model for the eyetracking data also included a by-item random slope for condition - the maximal random effect structure that allowed the models to converge.

To better understand the online processes, we also ran a cluster-based permutation analysis on mean proportion of looks towards syntactic choice as a function of time. This analysis determines whether there are time windows where a difference between the two conditions is significant. For each time sample, the analysis runs a t-test testing for the effect of condition (on the arc-sin transformed proportion of looks toward the syntactic choice video). Adjacent time points with a t-value greater than a predefined threshold ($t = 1.5$) 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. To compute the probability of observing a cluster of that size by chance, the data is reshuffled randomly between the two conditions a thousand times, and the same cluster-finding procedure is applied to each of the 1000 simulations. The p value of the effect of condition for a cluster is calculated as the proportion of simulations that produce a cluster larger than the cluster from the original data. Thus, a cluster is considered significant if it is larger than 95% of the clusters generated by chance.

Results

Pointing

As hypothesized, we found a main effect of condition ($\beta = 2.22$, $SE = .56$, $z = 3.97$, $p < .001$): participants in the reliable-syntax condition pointed more often towards the syntactic choice than participants in the reliable-semantics condition, suggesting that participants used the respective reliability of syntax and semantics as a cue to weigh the different sources of

information. The main effect of age was marginally significant (β = -1.06, SE = .55, z = -1.92, p = .055) with adults selecting the syntactic choice slightly less often than children (adults: mean 29.41% (SD = 42.04%), children: mean 32.89% (SD = 30.03%)). Interestingly, there was a significant interaction between age and condition (β = 1.65, SE = .55, z = 2.99, p = .003): the induction had a stronger effect on adults, who chose the reliable video more often than children, in both conditions (semantics and syntax). In the reliable-semantics condition, semantics seem particularly strong, since adults made this choice on almost every trial (fig.6). When running the same model separately for children and adults, the effect of condition is significant for both groups (children: β = .46, SE = .22, z = 2.09, p = .04; adults: β = 4.07, SE = .93, z = 4.39, p < .0001). See Figure 5.

Eyetracking

Overall proportion of looks. There was a main effect of condition (β = 0.32, SE = 0.06, t = 5.48, p < .0001): participants in the reliable-syntax condition looked more at the syntactic choice than participants in the reliable-semantics condition. There was a main effect of age (β = 0.18, SE = 0.06, t = 3.21, p = .002): children were looking at the syntactic choice more than adults. There was also a significant interaction between age and condition (β = -0.26, SE = .08, t = -3.34, p = .001): When analyzing children and adults separately, an effect of condition was found for adults only (adults: β = 0.31, SE = 0.07, t =4.75, p < .0001; children: β = 0.06, SE = 0.05, t = 1.36, p = .18): adults look more to the syntactic choice in the reliable-syntax condition than in the reliable-semantics condition (fig. 6).

Time-course analysis. For adults only, a time-window was identified in which the proportion of looking times towards the syntactic choice differed significantly between conditions: between 4020ms and 6720ms, just after the first content word onset, adults looked significantly more towards the syntactic choice in the reliable-syntax condition than in reliable-semantics condition (p = .002).

Figure 7 shows that the general gaze behavior is consistent with the presented conflict: when the function word appears (end of pronoun or article from 2.4s), the mean proportion of looks towards the syntactic choice increases (i.e. slight peak up towards the syntactic choice), and then decreases just after that (from 3s), when the content word appears (i.e. peak down towards the semantic choice). However, this second semantic peak seems visually smaller in the reliable-syntax condition than in the reliable-semantics condition, which is consistent with the pointing results.

To test for a possible bias towards semantics or syntax, we compared pointing results to chance (50%), in all conditions taken together. A bias is observed towards semantics (the intercept of the mixed-effect model is negative and significant, i.e. semantic choice: $\beta = -1.71$, $SE = 0.39$, $z = -4.42$, $p < .00001$, see figures 5, 6 and 7).

Discussion

We hypothesized that children will adapt to the reliability of different sources of information in their input to prioritize over these sources of information when they are in conflict. This hypothesis is verified by our pointing results, showing a difference in favor of syntax in the reliable-syntax condition relative to the reliable-semantics condition. Both children and adults tended to interpret sentences where the syntactic and semantic cues were in conflict more towards the syntactically reasonable choice when syntax was always reliable, and towards the semantically reasonable choice when semantics was always reliable. These results show that both children and adults are able to use the relative reliability of syntactic and semantic information to resolve a conflict between them. The integration of these sources of information is done by evaluating the relevance of each type of information and adapting their weight in context. This extends the results of Yurovsky et al. (2017), to the question of adaptability of syntactic and semantic sources of information. More importantly, this also indicates that adaptation is a mechanism operating in the context of learning, and not only in comprehension. In this task, children not only use their language knowledge to comprehend ambiguous sentences, as in Yurovsky et al., they also acquire novel meanings for familiar word forms – implying adaptation could be an important mechanism in language learning.

In language acquisition, research has shown that children can adapt their syntactic representations in context by relying on the words they know (e.g., Havron, de Carvalho, Fievet & Christophe, 2019), and that children can adapt their semantic representations, inferring that a word-form they know refers to a new semantic entity (learn a homophone, Dautriche, Swingley, & Christophe, 2015). However, until now it was unclear how children know which of these representations to change when they are in conflict. We have shown that children are able to adjust the weight of different sources of information to prioritize between them, and adapt their representations accordingly.

While both age groups showed an effect of the manipulation, we found that this effect was stronger for adults than for children. This could be due to task demands: Adults might have been able to integrate the induction phase more thoroughly, understanding and

integrating the novel meanings for the homophones, and/or better realizing the scope of the syntactic violations. According to error-driven learning models, children are predicted to adjust more than adults to violations of syntactic expectations, as their syntactic representations are supposedly less robust than adults' (e.g., Peter, Chang, Pine, Blything, & Rowland, 2015). However, previous research has also shown smaller adaptation effects in children than in adults (Havron, de Carvalho, Fiévet, & Christophe, 2019), probably due to their slower cognitive processing and lesser likelihood of integrating the induction sentences. Another way in which task demands might impact children and adults differently is the amount of the noise in our measure: adults tend to focus all their attention on the task, and perform it as error-free as they are able; children are typically less focused, and even when the task is relatively easy for them, there are always more mistakes. As a result, children's performance may move towards the center, resulting in smaller observed effects.

In addition to testing children and adults' explicit choice, we measured the gaze of the participants. Gaze data confirm this adaptation process for adults, showing that they look more towards the syntactic choice in the reliable-syntax condition than in the reliable-semantics condition. These results were not statistically significant for children. However, their gaze behavior helps to verify that the conflict situation between the two sources of information is indeed taken into account by the participants, both adults and children. In adults, both groups look more at the syntactic choice right after the syntactic cue, but then participants in the reliable-syntax condition continue to look at the syntactic choice while participants in the reliable-semantics condition switch to look at the semantic choice after the familiar word is pronounced. In children, we see a clear orientation change: Right after the syntactic cue, the function word, children look more at the syntactic choice, then, when they hear the familiar word, they look at the familiar object or action it represents. Then, they hesitate between the two videos, with the reliable-syntax condition recovering from the semantic cue a bit faster, though not significantly so. This suggests that the participants spotted each different cue, and that they were trying to integrate these conflicting cues as the sentence unfolded.

Additionally, we observed a general bias towards the semantic choice in conflict situations: children's results are shifted to the semantic choice, and adults' results in the reliable-semantics condition are almost at 100% of semantic choice, unlike results in the reliable-syntax condition, that are not at 100% of syntactic choice. This observation could mean that semantics is stronger than syntax by default, but it could also be due to a task asymmetry: in addition to ignoring the familiar meaning of the word, participants in the

reliable-syntax condition have another step to perform compared to the reliable-semantics condition – create homophones to be able to follow the syntax. The task is thus more difficult. Moreover, the linearity of language makes one cue appear before the other; syntax could be better retained because the function word is heard first, but semantics could be better retained because of its recency: indeed, on the timeline, one can see a change towards the semantic choice at the onset of the semantic cue (the content word), and then the gaze seems to have difficulty breaking off from it to go back to balance, independently of the condition.

It is possible that the mechanism of adapting to the reliability of different sources of information is active in children from the very first stages of language acquisition, and partly explain how infants manage to decide when syntactic representations need to be adjusted, and when lexical items need to be adjusted. Tests with younger children should be considered in order to examine whether children who are still at the beginning stages of language acquisition also show this effect, but a protocol adaptation would then be necessary. Another interesting future study would be to conduct a more thorough investigation of novel words learning. We had two exploratory items tacked at the end of our long experiment, which we describe in the supplementary materials. These two items were not enough to systematically explore whether children in the reliable-semantics condition have been ignoring syntax, or whether they have actually learned predictions that go against their life-long experience: That in the context of this experiment, determiners predict verbs and pronouns predict nouns. In a recent study, Havron and colleagues (2019) have shown that 3-4-year-old children are able to adapt their syntactic predictions, in an ambiguous syntactic context that permits both a noun and a verb continuation. They induced one group of children to predict nouns following this ambiguous syntactic context, and another group to predict verbs. The children then inferred that a novel word appearing in this context referred either to an object or an action, depending on their induction phase. However, in that experiment both predictions were grammatical, whereas in the current experiment, one group learned predictions that are not grammatical. It is unclear whether children would be able to change their predictions so thoroughly as to create un-grammatical predictions.

Conclusion

The current study supported the *noisy channel* hypothesis (Gibson et al., 2013; Yurovsky et al., 2017), extending it to local adaptation to syntactic and semantic information sources. More importantly, it showed that this hypothesis could also be valid for learning situations,

i.e. language acquisition in this case: children are able to adapt to a context and to integrate multiple linguistic sources of information by using each source's reliability, in order to acquire new knowledge (here, learning novel words). The greater use of some sources of information depending on their reliability in the situation, shows that information sources' weights change with the context and the available cognitive capabilities (Federmeier, 2007). It seems that language acquisition can thus rely on this adaptation capacity. Moreover, this adaptation permits children to learn even when the sources of information are in conflict.

References

Bernal, S., Lidz, J., Millotte, S., & Christophe, A. (2007). Syntax Constrains the Acquisition of Verb Meaning. Language Learning and Development, 3(33), 325–341. https://doi.org/10.1080/15475440701542609

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, 4–12. https://doi.org/10.1016/j.neuropsychologia.2016.08.015

Cauvet, E., Limissuri, R., Millotte, S., Skoruppa, K., Cabrol, D., & Christophe, A. (2013). Function Words Constrain On-Line Recognition of Verbs and Nouns in French 18-Month-Olds. Language Learning and Development, 10(1), 1–18. https://doi.org/10.1080/15475441.2012.757970

Dautriche, I., Fibla, L., Fiévet, A.-C. & Christophe, A. (2018). Learning homophones in context: Easy cases are favored in the lexicon of natural languages, *Cognitive Psychology, 104*, 83-105.

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

Dink, J., & Ferguson, B. (2018). *eyetrackingR*. R package version 0.1.8, http://www.eyetracking-R.com.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension, *Psychophysiology, 44*(4), 491-505. https://doi.org/10.1111/j.1469-8986.2007.00531.x

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of child language, 44*(3), 677-694.doi: 10.1017/S0305000916000209.

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y. S. (2008). A weakly informative default prior distribution for logistic and other regression models. *Annals of Applied Statistics, 2*(4), 1360–1383. https://doi.org/10.1214/08-AOAS191

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. https://doi.org/10.1073/pnas.1216438110

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

Havron, N., de Carvalho, A., Fiévet, A., & Christophe, A. (2019). Three- to Four-Year-Old Children Rapidly Adapt Their Predictions and Use Them to Learn Novel Word Meanings. *Child Development, 90*(1), 82-90. https://doi.org/10.1111/cdev.13113

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. https://doi.org/10.1016/j.jml.2007.11.007.Categorical

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, 13*(3), 317-334. https://doi.org/10.1080/15475441.2017.1283229

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology, 20*(2), 121–157. https://doi.org/10.1016/0010-0285(88)90017-5

New, B., Pallier C., Ferrand L., Matos R. (2001) Une base de données lexicales du français contemporain sur internet: LEXIQUE, *L'Année Psychologique, 101*, 447-462. http://www.lexique.org

Peter, M., Chang, F., Pine, J. M., Blything, R., & Rowland, C. F. (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. *Journal of Memory and Language, 81*(December), 1–15. https://doi.org/10.1016/j.jml.2014.12.002

Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition, 122*(3), 280–291. https://doi.org/10.1016/j.cognition.2011.10.004

Rabagliati, H., Pylkkänen, L., & Marcus, G. (2012). Top-down influence in young children's linguistic ambiguity resolution. *Developmental Psychology, 49*(6), 1076–1089. https://doi.org/10.1037/a0026918

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

van Heugten, M. & Christophe, A. (2014). Learning to accommodate syntactic violations during online speech perception. *International Conference on Infant Studies*, Berlin, 3-7 Juillet 2014

Waxman, S., Lidz, J., Braun, I. E., & Lavin, T. (2009). 24-Month-Old Infants Interpretations of Novel Verbs and Nouns in Dynamic Scenes. Cognitive Psychology, 59(1), 67–95. https://doi.org/10.1016/j.cogpsych.2009.02.001.24-Month-Old

Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York,

Yurovsky, D., Case, S., & Frank, M. C. (2017). Preschoolers flexibly adapt to linguistic input in a noisy channel. *Psychological science*, *28*(1), 132-140. https://doi.org/10.1177/0956797616668557

Figures

Figure 1. Example of a conflict test

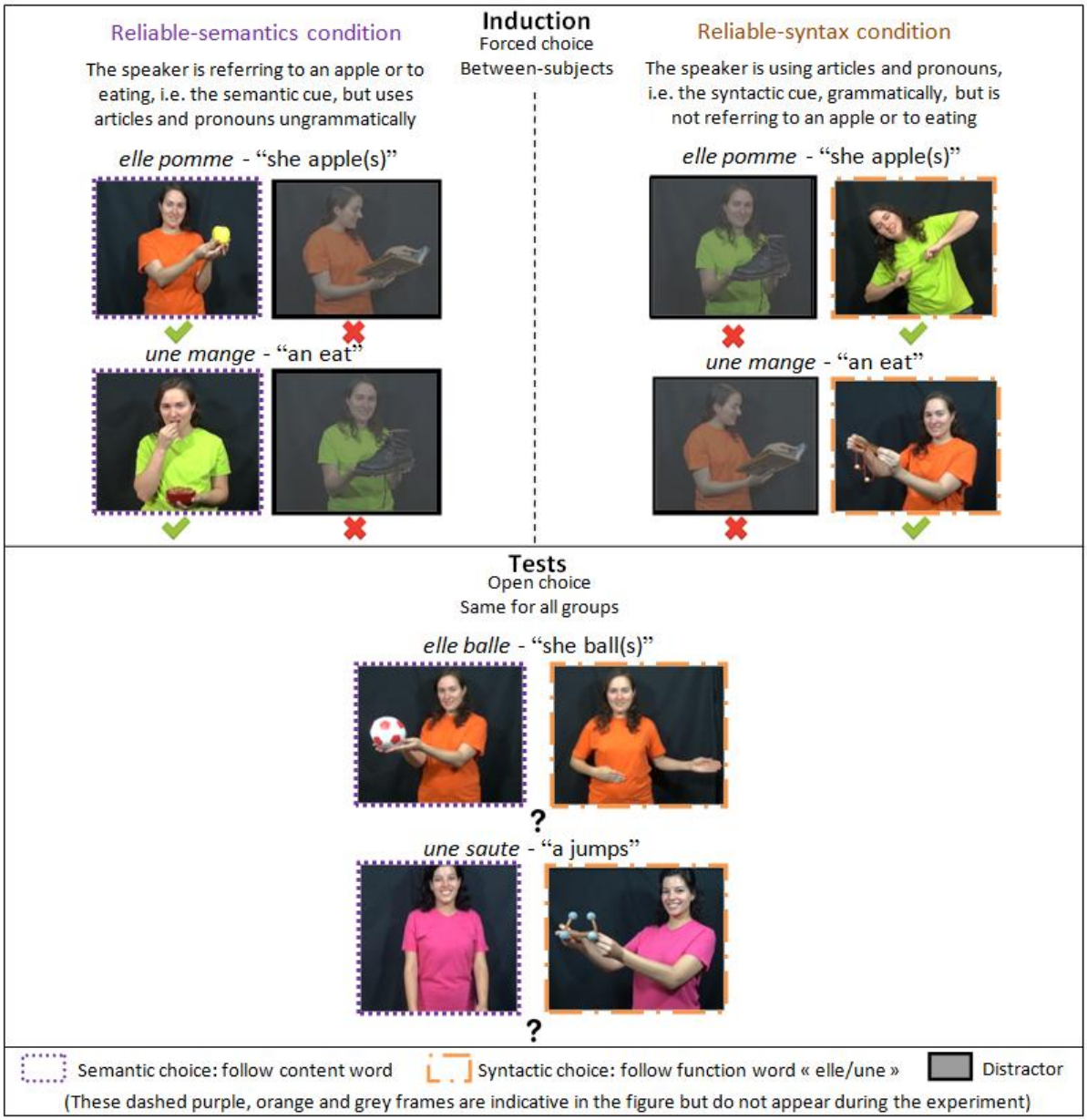


Figure 2. Experimental design: induction phase (2 conditions) and test phase

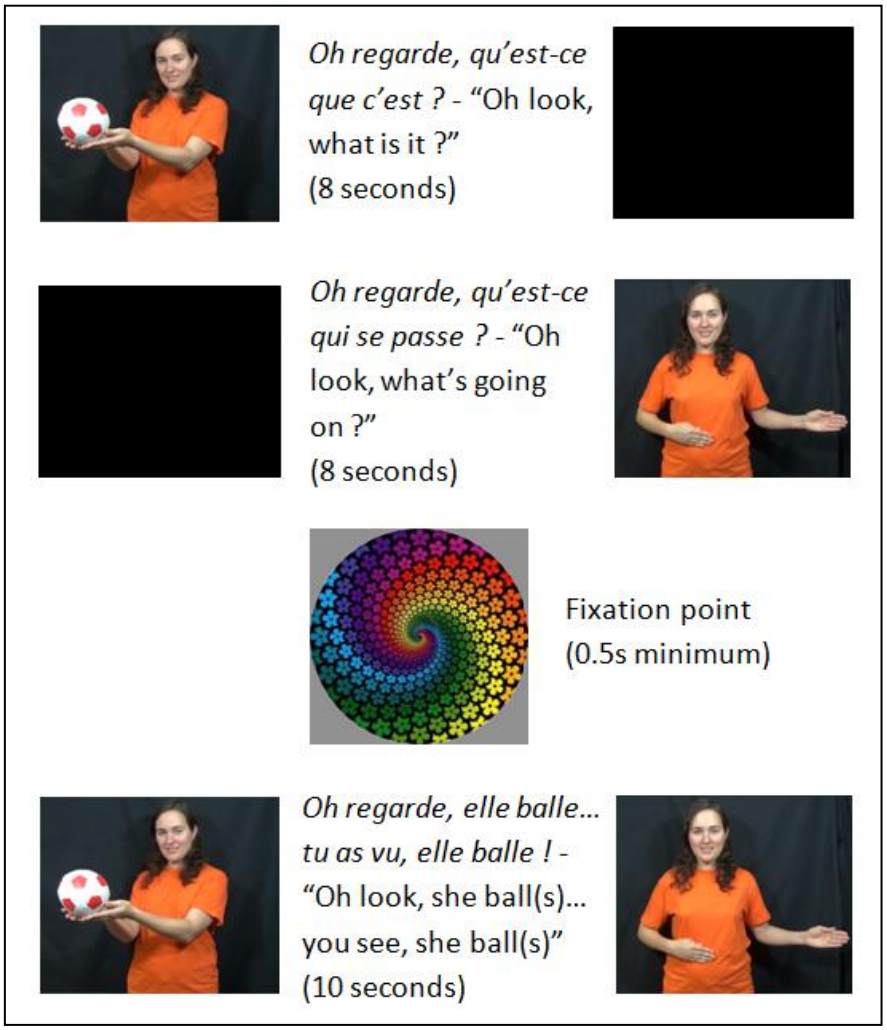


Figure 3. Trial sequence (for each item, example for a test item)

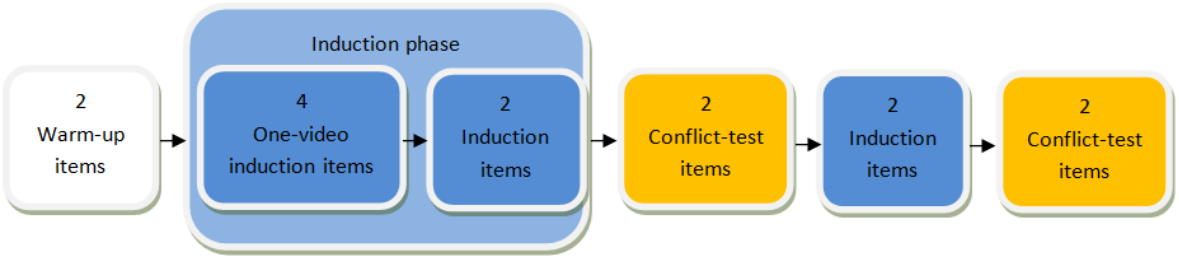


Figure 4. Task phases for each participant

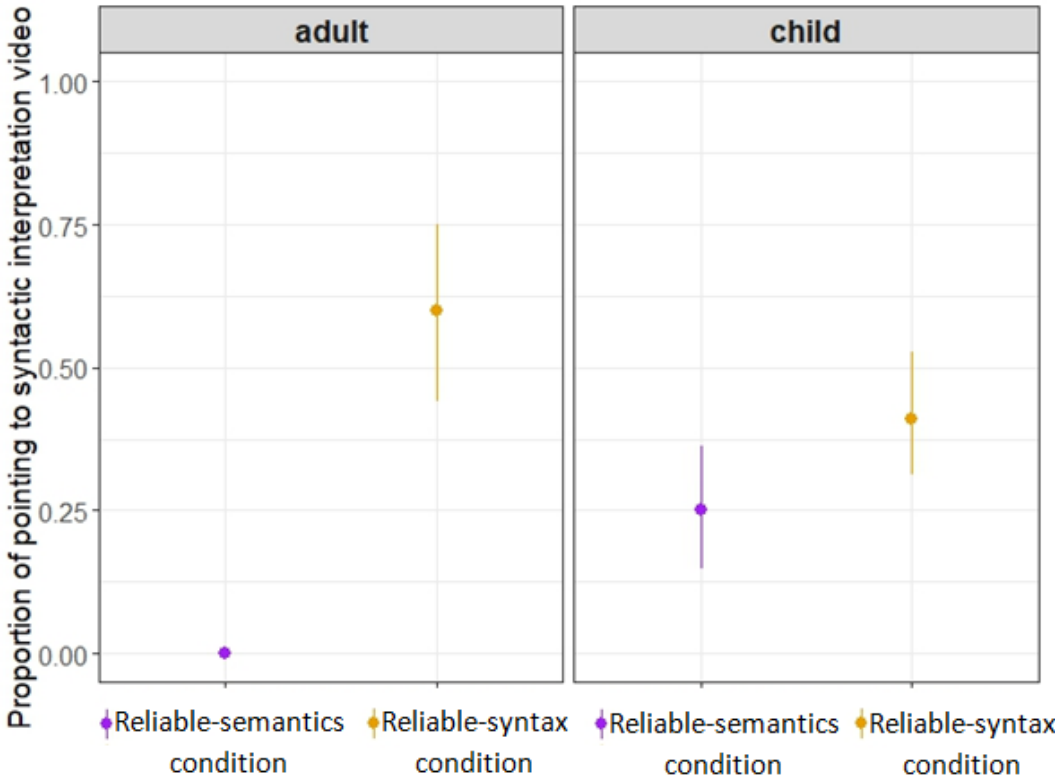


Figure 5. Proportion of points towards the syntactic choice, per condition (reliable-syntax vs reliable-semantics), for adults and children. The dot is the average for each condition and group, lines represent standard errors from bootstrapping as calculated by ggplot2. For both children and adults, there are more syntactic choices in the reliable-syntax condition.

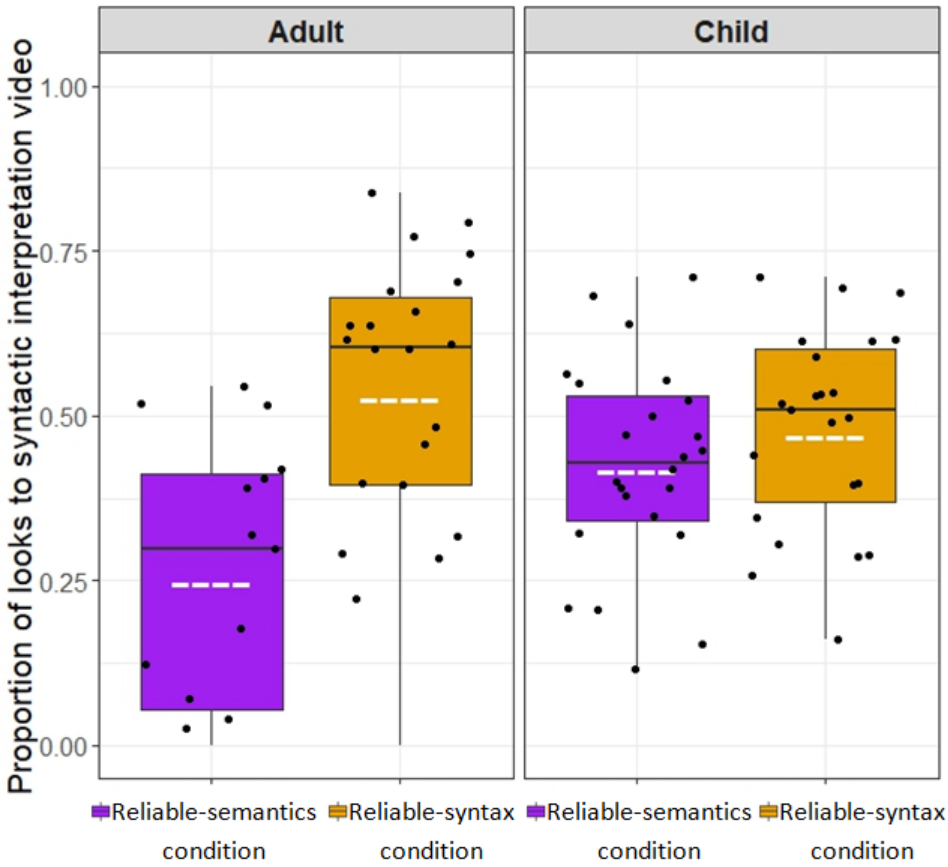


Figure 6. Mean proportion of looks towards syntactic choice per condition (reliable-semantics vs reliable-syntax) for adults and children. Black dots represent individual participants. The lower and upper hinges correspond to the first and third quartiles, the dotted white lines represent the means, and the black lines within the squares represent the median. The top and bottom whiskers denote the minimum/maximum value. As can be seen, adults in the reliable-syntax condition look more to the syntactic interpretation video than those in the reliable-semantics condition. However, there is no difference for children.

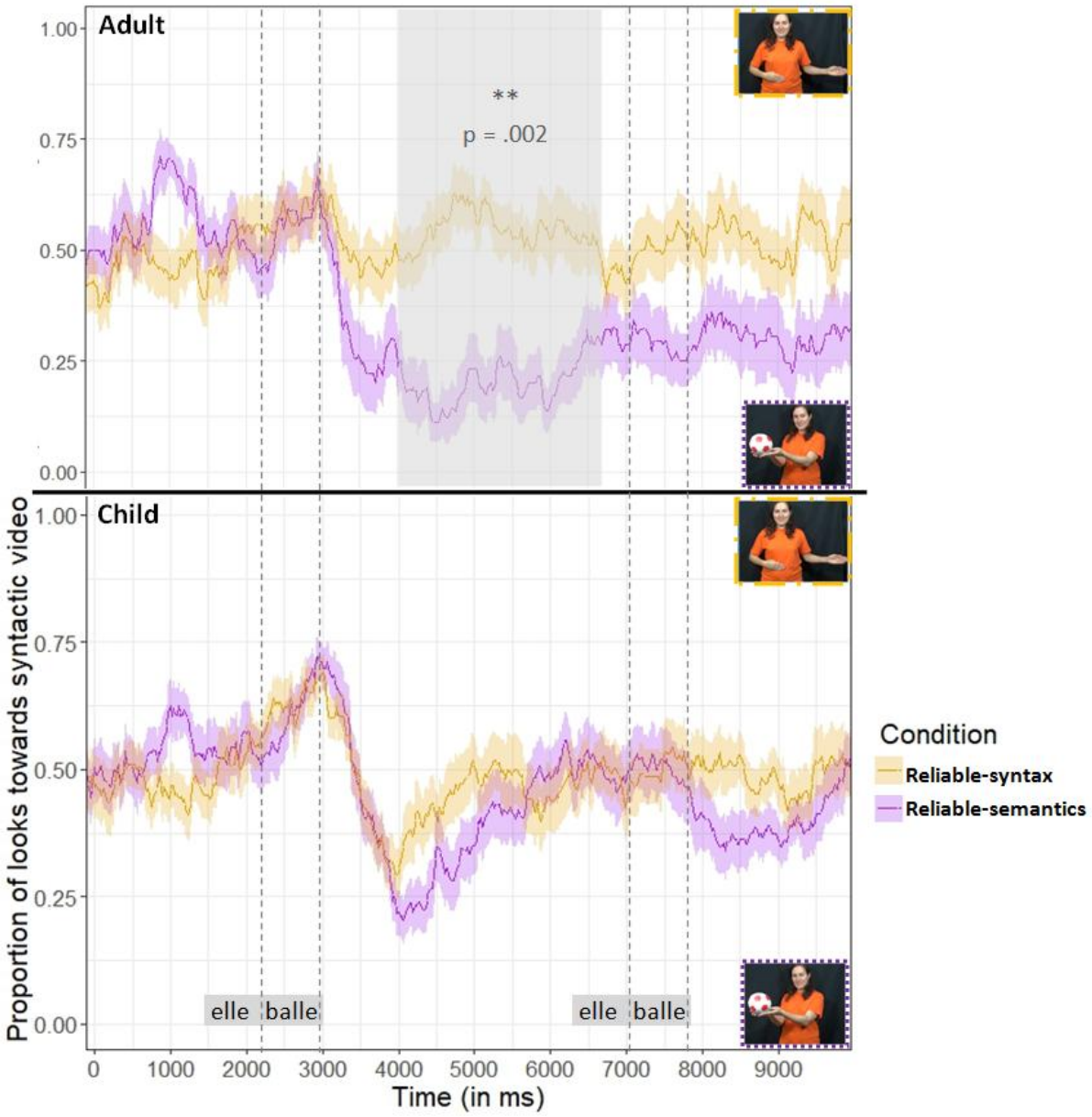
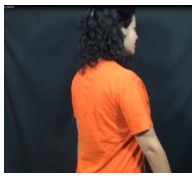


Figure 7. Proportion of looks towards syntactic choice, as a function of time, per condition (reliable-semantics/reliable-syntax) for adults (top half of the figure) and for children (bottom half of the figure).

Appendix: Stimuli

Sentence heard	English translation	Semantic video	Distractor
One-video induction:			
Elle robe	She dresses	 A dress	No distractor
Elle fraise	She strawberrys	 A strawberry	No distractor
Une parle	A speaks	 A girl who speaks	No distractor
Une tourne	A turns	 A girl who turns	No distractor
Induction:			
Elle pomme	She apples	 An apple	 A girl who reads
Elle chaise	She chairs	 A chair	 A girl who reads

Une mange	An eats	 A girl who eats	 A shoe
Une tousse	A caughs	 A girl who caughs	 A hat

Table A1. Induction items in reliable-semantics condition (4 familiar nouns and four familiar verbs)

Sentence heard	English translation	Syntactic video	Distractor
One-video induction:			
Elle robe	She dresses	 Novel action	No distractor
Elle fraise	She strawberrys	 Novel action	No distractor
Une parle	A speaks	 Novel object	No distractor
Une tourne	A turns	 Novel object	No distractor
Induction:			
Elle pomme	She apples	 Novel action	 A shoe
Elle chaise	She chairs	 Novel action	 A hat

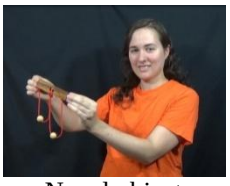
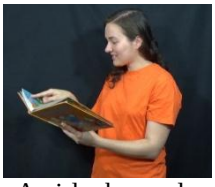
Une mange	An eats	 Novel object	 A girl who reads
Une tousse	A caughs	 Novel object	 A girl who drinks

Table A2. Induction items in reliable-syntax condition (4 familiar nouns and four familiar verbs)

Sentence heard	English translation	Video 1	Video 2
Elle balle	She balls	 A ball	 Novel action
Elle lune	She moons	 A moon	 Novel action
Une saute	A jumps	 A girl who jumps	 Novel object
Une dort	A sleeps	 A girl who sleeps	 Novel object

Table A3. Conflict-test items, the same for both conditions (two familiar nouns and two familiar verbs)

Sentence heard	English translation	Video 1	Video 2
Elle nuve / une nuve	She nuves/a nuves	 Novel object	 Novel action
Une dase / elle dase	She dases/a dases	 Novel action	 Novel object

Table A4. Novel-word test items, the same for both conditions (two novel words, one used as a noun and one as a verb)

[1] For pointing only, instead of *glmer* that did not converge, mixed-effect regression models have been launched using *bglmer*, with default parameters (Bayesian layer adding a weak informative prior on the fixed-effects parameters, as suggested in Gelman, Jakulin, Pittau, & Su, 2008).