Meghan E. Armstrong*, Llorenç Andreu, Núria Esteve-Gibert and Pilar Prieto

Children's processing of morphosyntactic and prosodic cues in overriding context-based hypotheses: an eye tracking study

Abstract: This research explores children's ability to integrate contextual and linguistic cues. Prior work has shown that children are not able to weigh contextual information in an adult-like way and that between the age of 4 and 6 they show difficulties in revising a hypothesis they have made based on early-arriving linguistic information in sentence processing. Therefore we considered children's ability to confirm or override a context-based hypothesis based on linguistic information. Our objective in this study was to test (1) children's (ages 4–6) ability to form a hypothesis based on contextual information, (2) their ability to override such a hypothesis based on linguistic information and (3) how children are able to use different types of linguistic cues (morphosyntactic versus prosodic) to confirm or override the initial hypothesis. Results from both offline (pointing) and online (eye tracking) tasks suggest that children in this age group indeed form hypotheses based on contextual information. Age effects were found regarding children's ability to override these hypotheses. Overall, 4-year-olds were not shown to be able to override their hypotheses using linguistic information of interest. For 5- and 6-year-olds, it depended on the types of linguistic cues that were available to them. Children were better at using morphosyntactic cues to override an initial hypothesis than they were at using prosodic cues to do so. Our results suggest that children slowly develop the

*Corresponding author: Meghan E. Armstrong, University of Massachusetts Amherst, Amherst, MA, USA, E-mail: armstrong@umass.edu
Llorenç Andreu, Grup de Recerca en Cognició i Llenguatge (GRECIL), Universitat Oberta de Catalunya, Barcelona, Spain
Núria Esteve-Gibert, Aix Marseille Université, CNRS, LPL UMR 7309, 13100, Aix-en-Provence, France; Universitat Pompeu Fabra, Barcelona, Spain
Pilar Prieto, Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain; Universitat Pompeu Fabra, Barcelona, Spain

ability to override hypotheses based on early-arriving information, even when that information is extralinguistic and contextual. Children must learn to weight different types of cues in an adult-like way. This developmental period of learning to prioritize different cues in an adult-like way is consistent with a constraint-based model of learning.

Keywords: Intonation, prosody, L1 acquisition, eyetracking, Catalan

1 Introduction

As listeners, we are constantly confronted with a wealth of information that we exploit in a dynamic way in order to arrive at a given interpretation. Consider the situation in (1).

(1)   Joanna is in the kitchen washing her hands during a dinner party with some of her husband's friends from work. Her husband has just gone into the dining room to serve a rack of lamb. While Joanna has never met any of the guests before, she is sure her husband's colleague Mark will have something to say about it since he is a vegan. Joanna hears a male voice incredulously say *"Lamb?!"*. Joanna has no doubt that it was Mark.

In (1) Joanna uses her world knowledge about vegans (they do not eat meat) in combination with discourse context (meat being served), but also in combination with linguistic knowledge (ability to identify prosodic marking of incredulity). By integrating all of these linguistic and extralinguistic resources, Joanna concludes that Mark must have been the one who incredulously said *"Lamb?!"*. But linguistic resources would lead Joanna to a different conclusion were the linguistic input different, as in (2).

(2)   Joanna is in the kitchen washing her hands during a dinner party with some of her husband's friends from work. Her husband has just gone into the dining room to serve a rack of lamb. While Joanna has never met any of the guests before, she is sure her husband's colleague Mark will have something to say since he is a vegan. Joanna hears a male voice say *"Laaamb!"*, excitedly. Joanna is sure it was not Mark who exclaimed *"Laaamb!"* in such an excited way. She concludes that it must have been one of the other guests.

In (2), although Joanna expects Mark to have had something to say given her knowledge about vegans and the serving of the lamb, she surmises that it is not

Mark that would have uttered the word "lamb" in such a way, since the prosodic information indicated some affinity for lamb.[1] Thus the linguistic information would have overridden her initial hypothesis based on world knowledge and discourse context. Clearly, as listeners (and even as mere overhearers), adults must integrate a wealth of information sources, making conclusions based on the dynamic integration of these information sources. As far as adults are concerned, there is plenty of evidence from language-processing showing that listeners integrate world knowledge and discourse context (Hagoort et al. 2004; Hald 2003; Hald et al. 2007; Nieuwland and Van Berkum 2006). In one such study using ERP methods, Hald et al. (2007) claimed that "world knowledge and discourse information conspire during the ongoing sentence processing as it unfolds in real time. That is, very rapidly and within the same time window both local discourse and general world knowledge information are taken into consideration during the construction of a sentence interpretation" (p. 215). Many researchers agree that many different types of information are simultaneously recruited while an utterance is being interpreted – including word meaning, syntax, discourse and world knowledge (Marslen-Wilson and Tyler 1980; Macdonald et al. 1994; Tanenhaus and Trueswell 1995; Clark 1996; Jackendoff 2002). Therefore both extra-linguistic and linguistic information are used in sentence processing.

At some point, then, children acquire the ability to integrate linguistic information with discourse context. Eventually, they would be able to use this linguistic information to override hypotheses they might have formed based on discourse context, as Joanna does in (2). In this research, we asked the following questions: (A) Are children able to form hypotheses based on contextual information? (B) Is linguistic information sufficient for children to override such a hypothesis? and (C) Does the type of linguistic information matter? Specifically, we compared how children used prosodic versus morphosyntactic information to override context-based hypotheses. In general, little attention has been paid to children's ability to integrate different types of linguistic information (lexical, morphological, syntactic, prosodic) as well as discourse context information during sentence comprehension. As children become savvy comprehenders of the language(s) they are acquiring, they must also become adept at integrating different types of information.

---

1 We acknowledge that an ironic or sarcastic interpretation could also be available for the hearer in such a situation.

But children between the ages of 4 and 6 have been shown to have difficulty controlling their automatic responses (i. e. inhibitory control) (Davidson et al. 2006). Choi and Trueswell (2010) also find that children at these ages show difficulty revising a hypothesis they formed based on early-arriving information. Given these findings, we chose to include children between the ages of 4 and 6 in our own study. Children between these ages were also of interest in light of recent work showing that 5- and 6-year-olds are tolerant of violations of Gricean Maxims (Grice 1975). Katsos and Bishop (2011) found that in tasks where children must accept or reject an utterance, they often accept pragmatically infelicitous utterances. In earlier work (Katsos and Smith 2010) propose a "pragmatic tolerance hypothesis", which predicts that in certain cases, even though children are pragmatically competent both as speakers and listeners, their metalinguistic awareness about accepting or rejecting pragmatically infelicitous utterances may still be developing. Evidence for this is that while violations of the Gricean Maxim of Quantity (overinformativeness and underinformativeness) may be tolerated by children around the age of five, no such tolerance is found for semantic falsity or syntactic ungrammaticality. While these studies have looked specifically at violations of the Maxim of Quantity, pragmatic tolerance could plausibly apply to other maxims. Relevance Theory (Sperber and Wilson 1987) holds that listeners will search for an optimal meaning in a given situation, and once that meaning is found, they stop processing. If a listener exhausts the possible optimal meanings and cannot find a good fit, then an utterance could be considered infelicitous or pragmatically strange. For example, Joanna reasons that the positive-sounding comment about lamb would not be a relevant thing for Mark to say in (2), given what she knows about him. The most relevant conclusion would be that someone else said it. Thus according to the pragmatic tolerance hypothesis a child might be more accepting of an utterance rendered infelicitous due to relevance. We thus reasoned that pragmatic tolerance could play a role in children's ability to revise hypotheses through linguistic information.

1.1 Children's hypothesis revision through linguistic cues

The bulk of the work on children's problems with revising initial interpretations deals with resolving structural ambiguities in processing sentences such as the one in (3).

(3)   Put the frog on the napkin into the box.

While parsing this sentence in real time, the Prepositional Phrase (PP) is potentially ambiguous since it is not clear whether it is a destination (the listener is being told where to put the frog: onto the napkin) or a modifier (the listener must put the frog that is on the napkin into the box). The preferred interpretation while the sentence is being parsed is to assume a destination interpretation (the frog should be placed on the napkin), and then revise this hypothesis as the sentence unfolds (Choi and Trueswell 2010). That is, once the listener hears the PP *into*, a modifier interpretation is forced for the previous PP *on the napkin*, and the destination interpretation is assigned to the PP headed by *into*. Previous work has shown that 4- and 5-year-olds, like adults, start with an initial destination interpretation (Trueswell et al. 1999). Differently from adults, however, children of this age group tend not to revise this initial destination hypothesis. A typical task to test this uses the visual world paradigm, presenting children with four objects: a frog on a napkin (target) a frog not on a napkin (competitor), an empty napkin (incorrect destination) and an empty box (correct destination). Instead of moving a frog that was already on a napkin to an empty box (as we would expect if *on the napkin* has been interpreted as a modifier and not a destination), children tend to move the frog to the empty napkin and then to the box instead of moving the frog that was already on the napkin to the box. However, children have no problem placing the frog on the napkin into the box with unambiguous phrases such as *Put the frog that's on the napkin into the box*.

Choi and Trueswell point out that this behavior demonstrates children's difficulty with revising their initial interpretation. They discuss two plausible explanations for this type of behavior: cognitive control, and differences in children's ability to make use of linguistic information. Inhibitory control is a component of children's executive function or cognitive control. It allows one to control automatic responses and behaviors. This can affect any number of human behaviors, but Novick et al. (2005) argue that inhibitory control is relevant in language processing. Since language processing is incremental (Choi and Trueswell 2010), interpretations based on early-arriving information (e. g. the verb *put* results in favoring a destination interpretation), may later be revised based on later-arriving information (e. g. processing of the second PP in (3) results in an alternate interpretation). Inhibitory control, and more generally *cognitive control*, would be the underlying mechanism responsible for the ability to not opt for the initial interpretation, resulting in a revision/overriding of this interpretation.

The second possible explanation for children's difficulty with hypothesis revision in language processing is based on evidence that children find different types of information to be more reliable. For instance, verbs are shown to be

highly reliable in terms of syntactic structure (Trueswell and Gleitman 2004, 2007; Snedeker and Trueswell 2004, Snedeker and Yuan 2008). Snedeker and Trueswell (2004) show that children rely more on verb cues than contextual cues in resolving the meaning of ambiguous sentences. Choi and Trueswell propose that there may be an order of acquisition of cues – children discover the most reliable cues early on and make use of them more than less reliable cues. They apply this reasoning to children's performance on the types of tasks described above, proposing that perhaps the morphosyntactic cue *into* is less reliable than a verb like *put* (since *Put* NP PP (destination) is much more common in English than *Put* NP). Therefore as children receive more experience with their language, they would learn more about the different types of cues available and develop the ability to weight the validity and reliability of cues.

　　These two explanations become intertwined, however, in the *Put the frog on the napkin into the box* examples because it is precisely the verb bias that would lead the child to form a hypothesis to be inhibited. Korean, however, provides an excellent test case, since it is a head-final language. In (4) and (5) the case marker *–ey* is ambiguous until the verb information is available at the end of the sentence:

(4)　*naypkhin-ey　kaykwuli-lul　nohu-se-yo*　　(PP is verb argument)
　　　napkin-Loc　frog-Acc　　put-Hon-SE
　　　napkin-on　　frog　　　　put
　　　'Put the frog on the napkin'

(5)　*naypkhin-ey　kaykwuli-lul　cipu-se-yo*　　　　(PP is noun modifier)
　　　napkin-Gen　frog-Acc　　pick up-Hon-SE
　　　napkin-on　　frog　　　　pick up
　　　'Pick up the frog on the napkin'

　　The verb *put* in (4) allows for a destination interpretation, and the verb *pick up* allows for a modifier interpretation. If it were the case that the reliability of verb information explains the behavior of English-speaking children on the *Put the frog on the napkin in the box* task described above, then Korean children should not have any problem parsing either (4) or (5), because the highly reliable verb information should give children enough information to parse the ambiguous phrases at the beginning of the utterance. However, Choi and Trueswell note that it is the destination interpretation that is most commonly used with these types of *–ey* phrases, and so it was expected that there would be a preference for a destination interpretation that could later be confirmed with the verb *put*, or overridden with the verb *pick up*. Again, using a visual

paradigm, Choi and Trueswell show with eye gaze patterns that children had a hard time blocking the destination interpretation when they heard the verb *pick up*. This finding supports a cognitive control explanation such that early-arriving cues lead children to hypotheses that are difficult to revise based on additional linguistic information. However, the authors entertain the possibility that cue weighting differs cross-linguistically – because Korean has case marking, morphosyntax (specifically noun morphology) could be more reliable in this language than in English. Therefore the noun morpheme *–ey* that the children would have heard at the beginning of the utterance, being a reliable cue, may have contributed to the difficulty in hypothesis revision. Thus different languages would weight cues in various ways, and this would affect a child's developmental trajectory. However, Choi and Trueswell's data strongly suggest that it is quite difficult for children to revise an initial hypothesis when it is overridden by linguistic information.

1.2 Children's use of contextual information

As mentioned above, there has been doubt as to whether children are able to use contextual cues in resolving meaning. Returning again to (1) and (2), information from the discourse context, specifically world knowledge and discourse context are crucial for Joanna to determine who spoke. Weighall (2008) finds that even when 5-year-olds were provided with contextual information to support a modifier interpretation of a PP, the children still experienced difficulty in arriving at the intended interpretation (as in Trueswell et al. 1999). Mode of presentation meant to guide children to the correct pragmatic interpretation did not seem to be helpful for children. She argues that in a model of processing where different information sources compete until the best-fit interpretation of all information is accepted and processed (Macdonald et al. 1994), children have difficulties integrating context. Trueswell and Gleitman's theory of constraint-based learning (2004) holds that the difference between adults and children is the prioritizing of different cues. Weighall maintains that contextual cues are prioritized later in the acquisition process. Under this assumption, children around age 5 would have quite a bit of difficulty correctly surmising who spoke in situations like (1) and (2), since it is imperative to integrate contextual information to do so.

1.3 Child development of prosodic comprehension

In addition to morphosyntactic information, we consider the role of prosodic information in the process of overriding children's hypotheses. Even prior to

birth, the human fetus has access to prosodic information (Mampe et al. 2009). And while it is well-known that infants are quite sensitive to prosodic information (Mehler et al. 1988; Soderstrom et al. 2003; Soderstrom et al. 2011; Esteve-Gibert 2014, inter alia), many studies have shown that successful prosodic comprehension at the sentence level comes relatively late (Hornby and Hass 1970; Cruttenden 1974; Bates 1974; MacWhinney and Price 1980; MacWhinney et al. 1985; Cruttenden 1985; Cutler and Swinney 1987; Vogel and Raimy 2002; Wells et al. 2004). It has been observed that children between 5 and 7 years of age produce meaningful intonation contours, but have difficulty comprehending their meaning and use. This observation is known as "the performance paradox" (Cutler and Swinney 1987). For sentence-level prosody specifically, however, recent work giving children more contextual information than had been provided in previous studies has shown that even four-year-olds are rather adept at comprehending sentence-level meaning through prosody. Ito et al. (2012) argued that early studies presented stimuli to children in a way that was not well-motivated pragmatically, and may have sounded too 'out-of-the-blue' to children. Armstrong (2014) showed that 4-, 5- and 6-year-old speakers of Puerto Rican Spanish can detect incredulous meaning associated with intonation in polar (yes-no) questions. They were able to listen to two types of polar question intonation: one that conveyed disbelief on the part of the speaker, and another that did not. All age groups were found to perform significantly above chance when asked to identify a disbelieving individual based on this intonational difference. This shows that 4-6-year-olds can successfully use sentence-level intonation to guide them to a specific meaning, in this case an epistemic meaning (disbelief). In another recent study, Zhou et al. (2012) tested children's ability to use prosody to disambiguate identical segmental strings that differed in speech act type. In Mandarin, a wh- word may either mark an utterance for questionhood or as an indefinite noun phrase in a statement. The difference is signaled intonationally. For example, the segmental string in (6) can have two interpretations. When there is rising intonation on the wh- phrase it means 'What fruit did John not eat?', but with level intonation on the wh-phrase it means 'John didn't eat any fruit.'

(6)   Yuehan   meiyou   chi   shenme   shuiguoa.
       John   not   eat   what   fruit

Zhou et al. used tasks that measured online and offline comprehension of the speech act meaning based on intonational differences. Older 4-year-olds and younger 5-year-olds were examined. The participants were tested using the visual world paradigm (Tanenhaus et al. 1995). As the authors note, this

paradigm is often used in psycholinguistic experiments based on the observation (Cooper 1974) that relies on the fact that participants tend to look at the referents of the words they are listening to. The paradigm has been applied, among others, to the study of word recognition (e. g. Allopena et al. 1998; Dahan et al. 2001), lexical and syntactic ambiguity resolution (e. g. Dahan and Tanenhaus 2004; Spivey et al. 2002; Trueswell et al. 1999), and language comprehension (e. g. Altman and Kamide 1999; Arnold et al. 2000; Kamide et al. 2003). Studies that have used this paradigm have shown that eye movements can be used to understand the mental processes involved in the comprehension of spoken language and provide information about how visual information is gathered in production tasks.

In Zhou et al.'s experiment, participants heard a sentence while they viewed a picture. Additionally, as an offline (explicit) measure of comprehension, a Kermit the Frog puppet would say something to the child after each sentence. If he asked something, participants had to tell him the answer to the question. If he stated something, children had to tell Kermit whether what he said was right. For the online (implicit) measure, participants looked at a picture of a character and five objects, which were grouped into two categories. For example, the participant might see a picture of a boy with a banana, but also a group of animals and a group of other fruits that the boy had not picked. They heard the Mandarin version of either 'What fruit did Xiaoming not pick?' which would facilitate looks to the fruits Xiaoming was not holding, or 'Xiaoming didn't eat any fruit', which would facilitate looks to the animals. Examination of fixation patterns for both children and adults showed that the children's fixation patterns were similar to the adults'. The effect of intonation was just as strong for children as it was for the adults. The offline task (with Kermit the Frog) also showed that children were sensitive to the intonational difference that set apart the two types of speech acts examined.

In an eye tracking study on the effect of discourse-embedded pitch expansion on prenominal adjectives in Japanese, Ito et al.'s (2012) results revealed that 6-year-olds are indeed sensitive to the marking of contrast through pitch prominence. Both younger and older 6-year-olds were able to use pitch prominence to guide them to a specific referent based on discourse context, but there were still evident differences when compared to adults. While garden path effects with the pitch accent used for contrastive focus in American English, L + H* were shown for 6- and 7-year-olds, their response was delayed when compared to adults. Even children above age 11 have been shown to respond slower than adults (Bibyk et al. 2009; Ito et al. 2014). An eye tracking study by Sekerina and Trueswell (2012) showed that in Russian 5- and 6-year-olds are not adult-like in their interpretation of contrastive pitch accent. Early accent for split

constituent constructions in Russian yielded only a small facilitation effect for identifying a target referent for children of this age group. Snedeker and Yuan (2008), however, found that for a similar age group, English-speaking children can indeed make use of prosodic cues to resolve syntactic ambiguities. This work makes it clear that prosodic development depends on the specific role of prosody being investigated, how prosody is interacting with other parts of the grammar, and the interaction between prosody and discourse context information. The various paths found for prosodic acquisition are thus likely to diverge depending on the specific meaning being conveyed through prosody, but the window between the ages of 4 and 6 appears to be an important one.

1.4 Aims

As in the ambiguity resolution studies we described above, in this paper we are interested in assessing children's ability to revise an initial hypothesis based on early-arriving information. The aforementioned studies have looked at early-arriving linguistic information, but what happens when the early information is contextual, and furthermore extralinguistic? Consider the schematics in Figures 1(a) and 1(b). Figure 1(a) considers the verb as early-arriving information, which
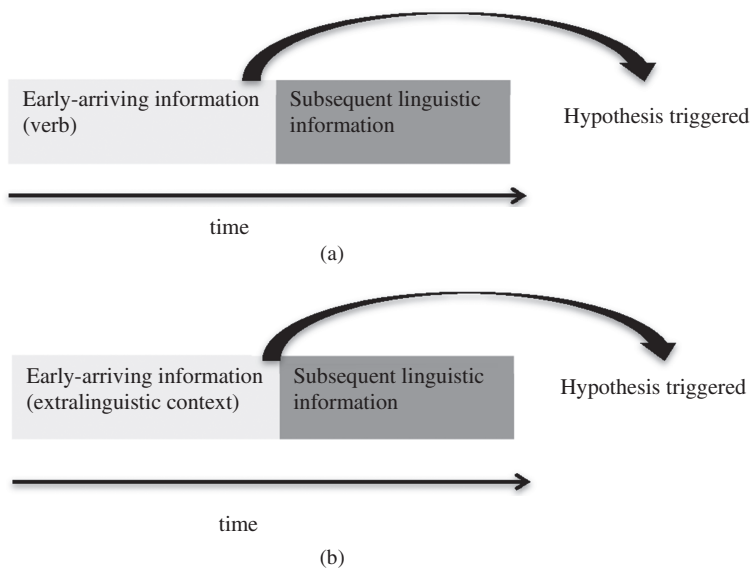

Figure 1: (a) Schematic for an utterance with *verb* as early-arriving information. (b) Schematic for an utterance with *context* as early-arriving information.

would trigger the child's hypothesis as has been shown to happen to sentences like (3). If contextual information became available before linguistic information, then this information, too, could possibly trigger a hypothesis, as depicted in Figure 1(b). Hence the initial hypothesis would be based on information available to them prior to the onset of the utterance.

Given Weighall's (2008) claim that the ability to weight context in an adult-like way is late-acquired, if the early-arriving information is contextual information, will children ignore it? We also consider Choi and Truewell's discussion of the reliability of linguistic cues. For instance, Joanna relied on prosodic information to arrive at her hypotheses in (1) and (2). For the same segmental string [læm], different prosodic patterns can completely change how the intention of the utterance is interpreted. That we know of, no studies have considered how prosodic (namely intonational) cues allow children to revise an initial hypothesis.

In this paper we would like to extend the type of "early information" that might cause children to form a hypothesis. Namely, we consider what happens when children have access to contextual information, and *then* linguistic information. We also compare what happens when a speaker conveys information using different linguistic cues. Specifically, we consider morphosyntactic cues (sentential negation and a verum focus particle) as well as children's comprehension of prosodic cues (intonation contours). If children form hypotheses based on early-arriving contextual information, it would mean that extralinguistic contextual information and not just linguistic information can also contribute to biases formed with early-arriving information. These questions are especially important given claims like those of Weighall, who addresses children's difficulty treating discourse context as a reliable cue. But if children rely on linguistic information only and not contextual information (even though it is early-arriving), it would support the idea that contextual information becomes reliable to children only later in development. Trueswell and Gleitman's theory of constraint-based learning (2004) would predict that different types of linguistic cues are weighted in different ways throughout the acquisition process. If there were differences in how children use the morphosyntactic or prosodic cues, this would support such a theory.

The remainder of the paper is organized as follows: we present methodology used in a comprehension experiment designed to evaluate children's (4-to 6-year-olds') explicit (pointing) and implicit (eye movements) reactions to various discourse context/ linguistic stimulus combinations. We discuss the results of this experiment with respect to the previous literature on the different factors that might be at work when children are doing hypothesis revision.

2 Methods

In order to test children's ability to form a context-based hypothesis, and later to confirm or override it based on linguistic information, we designed a comprehension experiment in which children's responses had both offline (pointing) and online (eye tracking) measures. We refer to the confirming or overriding of a hypothesis as a "pragmatic effect". As stated earlier, there were two different conditions for linguistic type: morphosyntactic and prosodic. These conditions were given in a between-subjects design such that roughly half of the participants for each age group received the morphosyntactic condition, and the other half received the prosodic condition.

2.1 Participants

A total of 160 children participated in the experiment: 50 4-year-olds, 49 5-year-olds, and 61 6-year-olds. Schools in Catalan-dominant areas of Catalonia, Spain were contacted to participate in the study, and parents who wished for their children to participate gave consent. All participants were Catalan-dominant and attended schools where Central Catalan was the language of instruction. A language background questionnaire was administered to determine Catalan dominance. Children who had at least one parent whose L1 was Catalan and used Catalan at least 90 % of the time throughout the day were included in the study. The variety of Catalan spoken by all children in the study is Central Catalan.

2.2 Materials

The trials in our experiment were made up of short videos, at the end of which participants heard an auditory stimulus. Here we describe both the videos (visual materials) and the auditory stimuli heard at the end of the videos (auditory materials).

2.2.1 Visual materials

A total of six short videos (each about forty seconds in duration) were created, two of which were used for familiarization, and four of which were test items. Each video began with two actors (either a pair of adult females or a pair of

adult males), each playing with a doll (identical stuffed rabbits), as shown in Figure 2. Shortly after the video begins, the actors stop playing with the animals, and leave them on their respective chairs. They then leave the room and shut the doors behind them, leaving their dolls on their respective chairs, as in Figure 3.

Figure 2: Videos begin with actors playing with dolls (stuffed rabbits).

Figure 3: Actors leave the room, shutting the door behind them and leaving rabbits on chairs.

While the actors are gone, a "sneaky" puppet appears and replaces just one of the actor's dolls with a different animal, as pictured in Figures 4 and 5.

The sneaky puppet leaves after switching the dolls. The actors then re-enter the room, and approach the chairs to observe the new situation, as shown in Figure 6. The image freezes when each actor is about the same distance from his or her respective chair. The actors did not produce any communicative facial or corporal gestures, but gazed down towards the doll on their respective chairs.

Figure 4: Sneaky puppet steals one of the actor's rabbits.

Figure 5: Sneaky puppet replaces the rabbit with a donkey and leaves.

The audio stimulus was played at this point. It was hypothesized that based on this scenario (and world knowledge), children would expect the actor with the changed doll to react, and would therefore fixate this actor in anticipation of this reaction. World knowledge would tell the participant that the change of the doll violates the actor's beliefs, and as a result the actor would provide some sort of reaction upon witnessing the disappearance of his/her doll.[2]

A unique pair of actors were featured in each of the four test videos. For each participant, there were a total of six trials. The first two trials were familiarization trials, and featured two female actors. The same pair of female

---

2 We recognize that there is an obvious Theory of Mind component in this task, since it involves belief reasoning, specifically the actor's beliefs about events.

Figure 6: The actors re-enter the room and observe the modified situation.

actors appeared in each familiarization trial. For the four test trials, each trial featured either two female actors or two male actors. Participants received two test trials that featured a pair of female actors, and two test trials that featured a pair of male actors. In this way, participants never saw the same pair of actors twice, so as to prevent any kind of voice recognition being used as a cue to who spoke. Additionally, the voice that produced the audio stimulus was never the actual voice of either of the actors, but rather a third party.

2.2.2 Audio materials

The audio materials were designed to either confirm or override the context-based expectation that the actor with the changed doll would speak. Four recordings (a total of eight unique recordings) were made for the confirm condition, with two recordings for each linguistic type (two morphosyntactic and two prosodic) and four recordings were made for the override condition, with two recordings for each linguistic type (two morphosyntactic and two prosodic). Two recordings were made with sentential negation and two with a verum focus particle. One sentential negation recording was made by a female, and the other by a male, and the same was done for the verum focus particle. Likewise, two recordings were made with the prosodic profile depicted in Figure 7 (incredulity question) and two with the prosodic profile depicted in Figure 8 (intensified affirmative statement). An orthographic representation of the stimuli based on linguistic type and pragmatic effect is shown in Table 1.
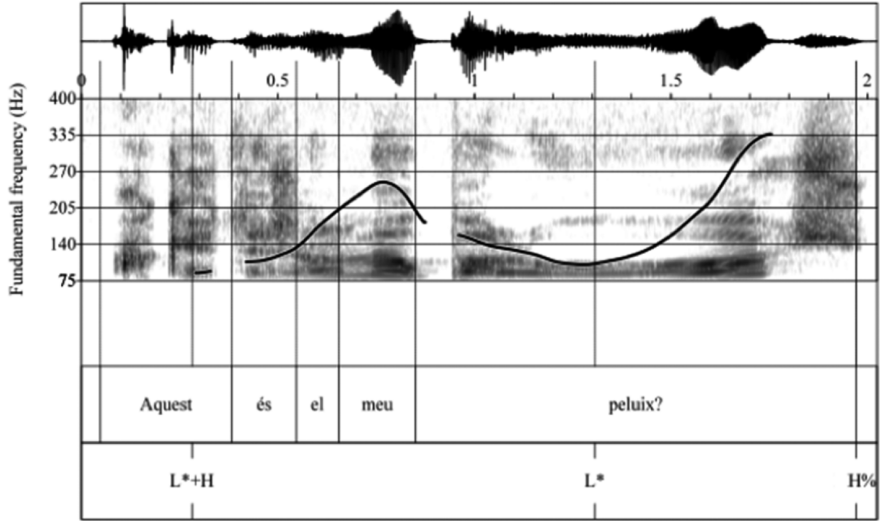
Figure 7: *Aquest és el meu peluix?!* 'This is my doll?!' produced with a L*+H prenuclear accent and a L* H% nuclear configuration in the Cat_ToBI system.
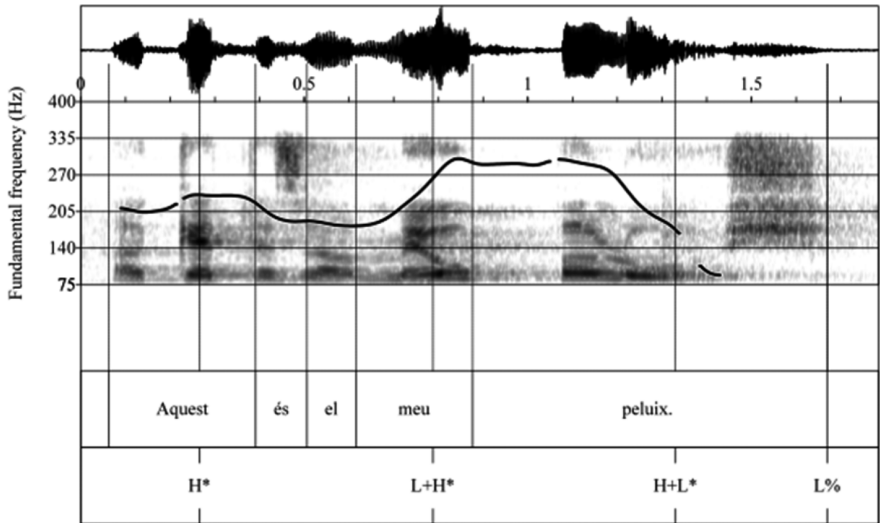

Figure 8: *Aquest és el meu peluix.* 'This is my doll.' produced with two prenuclear pitch accents: H* and L+H*, and a H+L* L% nuclear configuration in the Cat_ToBI system.

Table 1: Stimuli utterances by conditions (linguistic type and pragmatic effect).

Target utterance	Condition: Linguistic type	Condition: Pragmatic effect
*Aquest **no** és el meu peluix* 'This is **not** my doll'	Morphosyntactic (sentential negation particle ***no***) (Broad focus intonation: L*L%)	**Confirm** context-based hypothesis (actor with switched doll will speak)
*Aquest **sí que** és el meu peluix* 'This **sure** is my doll'	Morphosyntactic (verum focus particle ***sí que***) (Broad focus intonation: L*L%)	**Override** context-based hypothesis (actor with same doll will speak)
Aquest és el meu peluix?! 'This is my doll?!'	Prosodic (incredulity intonation: L* + H L*H%)	**Confirm** context-based hypothesis (actor with switched doll will speak)
Aquest és el meu peluix 'This is my doll'	Prosodic (intensified affirmative statement intonation: H* L + H* H + L* L%)	**Override** context-based hypothesis (actor with same doll will speak)

Note that the stimuli recorded for the morphosyntactic condition were all produced with broad focus intonation in Central Catalan (a L* L% nuclear configuration in the Cat_ToBI system, Prieto et al. 2015). As shown in Table 1, for the morphosyntactic condition, the particle *no* negated the VP. This is the strategy used for canonical sentential negation in Central Catalan. This was intended to *confirm* the participant's hypothesis since it would be a pragmatically felicitous utterance for the actor with the changed doll to produce in this context. *Sí que* is considered a verum focus particle since it intensifies and focuses the VP in Catalan. If the participant hypothesized that the actor with the changed doll should be the one to speak, then this form should *override* the participant's expectation that the actor with the changed doll should be the one to speak. These specific forms were desirable since Central Catalan allows for either a negative or a positive particle to precede the VP. Given that we were measuring implicit responses with eye tracking, this structure allowed for a good comparison of how participants processed the negative or positive particle preceding the VP. For the prosodic condition, the segmental string was the same for both the confirm and override conditions, but produced with different intonation contours (Figures 7 and 8). Incredulity intonation was expected to be associated with the actor with the changed doll, and an intensified affirmative statement produced with a child-directed speech style was expected to be associated with the actor whose doll was not changed. Two unique familiarization recordings were also made. In the first, the audio stimulus was *Ostres, ara tinc una granota!* 'Wow, now I have a frog' and in the second the audio stimulus was *Clar! Tinc el meu conill encara!* 'Of course! I still have my rabbit!' The two utterances were produced with different intonational patterns, but neither coincided with the intonational patterns of the test items.

2.3 Setup and procedure

2.3.1 Validation

In order to validate the stimuli and procedure, a group of 12 Catalan-dominant adults (ages 18–23) who were students at the Universitat de Barcelona (Mundet Campus) were given the experiment. The linguistic type condition was given between-subjects. Six of the participants received the morphosyntactic condition, and 6 received the prosodic condition. Each participant received 4 test trials. A total of 24 trials were validated. Crucially, all participants matched 100 % of trials.

2.3.2 Set-up

Participants were seated approximately 22" in front of a Tobii T120 eyetracker with an integrated 17" TFT monitor. Tobii Studio Software was used to present the stimuli, and collect the eye tracking data. Eye position was sampled at 120 Hz (approximately 8 ms intervals). A nine-point calibration was carried out at the beginning of the experiment. The Tobii Studio Software automatically validates calibrations and the experimenter could, if required, repeat the calibration process if validation was poor. Calibration took approximately 20 seconds. The video stimuli were displayed on the monitor of the eyetracker.

2.3.3 Procedure

As mentioned above, before being presented with the test trials, each participant received two familiarization trials. During the familiarization phase, participants were trained not only on how to do the task, but also to teach the children that either of the actors might have something to say. This process also helped train the participants about the expectations of the actors. The participants received the two training trials followed by 4 test trials. Order of presentation and "switch side" (i. e. whether the doll switching occurred on the left or the right) were counterbalanced. Target stimuli were presented from one of two lists (confirm, override, override, confirm (list 1) or override, confirm, confirm, override (list 2)).

The participants watched each video, which did not include any sound until the actors returned (as in Figure 6) and the shot was frozen. At this point participants heard an audio stimulus depending on the condition to which they were assigned. The linguistic type condition was given between subjects

and the pragmatic effect condition was given within subjects. The stimuli were presented to participants via a mono channel split to two loudspeakers positioned on either side of the viewing monitor. Participants were then asked *Qui ho ha dit, això? Assenyala-me'l.* 'Who said that? Point to them for me.' The pointing task constituted the offline (explicit) measure. Responses were measured as correct or incorrect. The recording of the participants' eye movements constituted the online (implicit) measure.

3 Results

3.1 Offline results

Participants' responses were coded for the two linguistic conditions: morphosyntactic and prosodic. A response was considered correct if the participant pointed to the actor associated with either the confirming or overriding response. For example, in Figure 9, if the child pointed to the actor on the left (indicated in Figure 9 with a pointing finger), the actor whose animal had not been changed, s/he was given credit for a correct answer since the linguistic stimulus would override the expectation that the actor on the right (with the changed doll) would speak (recall that both actors started with rabbits).

Audio: *Aquest **sí que** és el meu peluix.* 'That sure is my doll.'

Figure 9: Example of correct pointing response.

Figure 10 shows the mean correct responses for the confirm condition for child participants, by age and linguistic type. 4-year-olds responded correctly 93 % of the time for the morphosyntactic condition and 94 % of the time for the prosodic condition; 5-year-olds responded correctly 97 % of the time for the morphosyntactic condition and 85 % of the time for the prosodic condition; 6-year-olds
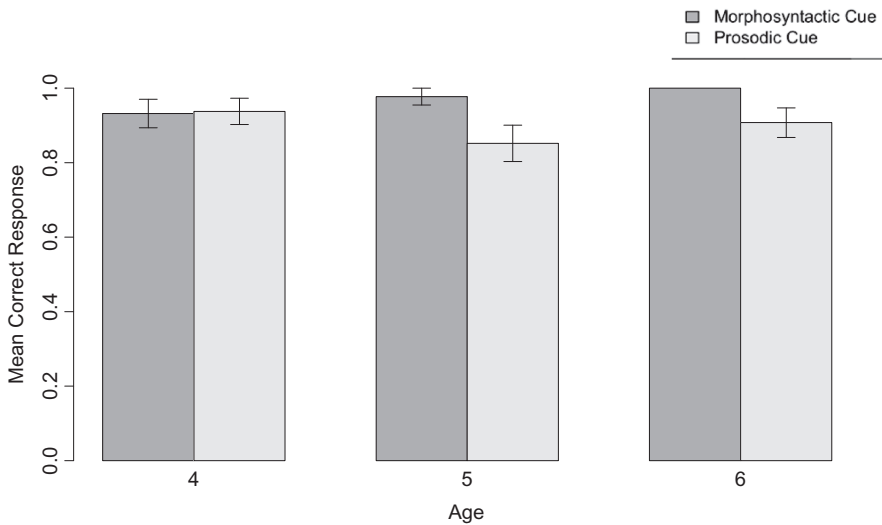
Figure 10: Mean correct responses by Age and Linguistic type, Pragmatic effect = confirm.

responded correctly 100 % of the time for the morphosyntactic condition and 91 % of the time for the prosodic condition. One-sample t-tests showed that all age groups were significantly above chance for both the morphosyntactic and the prosodic conditions.

Figure 11 shows the mean correct responses for each age group by linguistic type with override as the pragmatic effect. Child responses were less adult-like for the override condition. Four-year-olds responded correctly 47 % of the time for the morphosyntactic condition and 33 % of the time for the prosodic condition. One-sample t-tests revealed that they were not significantly below chance for the morphosyntactic condition ($t(43) = -0.29$, $p > 0.05$), but were significantly below chance for the prosodic condition ($t(47) = -2.42$, $p < 0.05$). Five-year-olds responded correctly 77 % of the time for the morphosyntactic condition and 50 % of the time (at chance) for the prosodic condition. Five-year-olds' responses were significantly above chance for the morphosyntactic condition ($t(43) = 4.27$, $p > 0.001$). Six-year-olds responded correctly 84 % of the time for the morphosyntactic condition and 69 % of the time for the prosodic condition. They performed significantly above chance for both the morphosyntactic ($t(63) = 7.51$, $p < 0.001$) and prosodic ($t(54) = 2.90$, $p < 0.01$) conditions. Again, as age increases less variability is found for both the prosodic and morphosyntactic conditions.
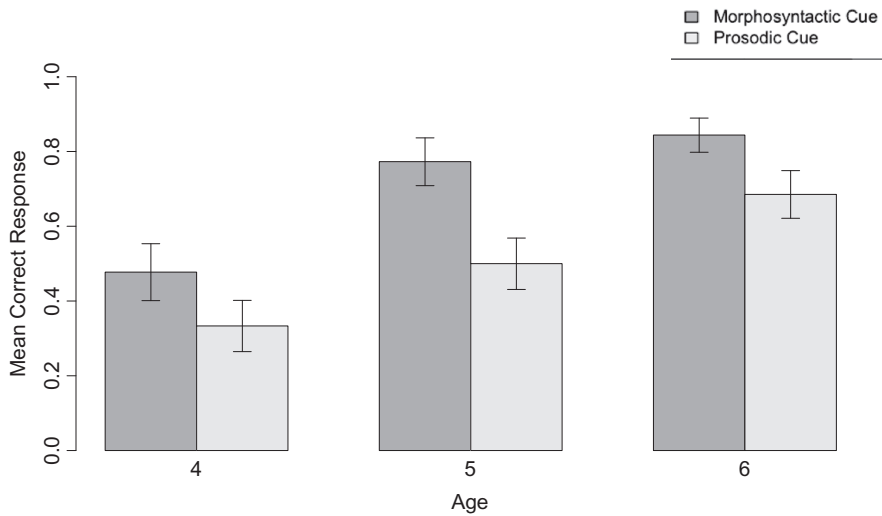
Figure 11: Mean correct responses by Age and Linguistic type, Pragmatic effect = override.

To statistically evaluate the children's response patterns further, mixed-effects logistic regression models were fit using the lmer package in the R software package (R Core Team 2013). We fit the data for the children with RESPONSE (correct or incorrect) as the dependent variable. We included AGE (3 levels: Age 4, 5 and 6), LINGUISTIC TYPE (two levels: morphosyntactic vs. prosodic) and PRAGMATIC EFFECT (two levels: confirm vs. override) as fixed effects, with participant as a random factor, as well as all possible interactions between those factors. The best-fit model included age, pragmatic effect and linguistic types as well as the following interaction: pragmatic effect × age. Each of these fixed effects

Table 2: Fixed effects from best-fit mixed effects logistic regression model of correct responses, Age 5 as reference level.

	Estimate	SE	z value	Pr(>|z|)
Intercept (Age 5)	4.10	0.76	5.43	$p < 0.001$
Age 4	0.32	0.74	0.43	$p > 0.05$
Age 6	0.89	0.75	1.81	$p > 0.05$
Pragmatic effect: override	−2.64	0.77	−3.43	$p < 0.001$
Linguistic type: prosodic	−1.66	0.74	−2.25	$p < 0.05$
Pragmatic effect: override × Age 4	−1.69	0.76	−2.23	$p < 0.05$
Pragmatic effect: override × Age 6	0.00	0.76	0.00	$p > 0.05$

were significant predictors of correct response on the offline task. Table 2 shows the estimates, standard error, z values and *p* values for the model with Age 5 at reference level, so as to observe any differences between 5- and 6-year-olds.

The summaries of the model shown in Table 2 reveal that there were no significant differences in age, when the data are considered as a whole. The negative coefficient for PRAGMATIC EFFECT: OVERRIDE reflects the fact that overall, children performed significantly worse for the override condition. LINGUISTIC TYPE was also selected as a significant predictor of performance on the task. The negative coefficient for LINGUISTIC TYPE: PROSODIC shows that overall, children performed significantly worse for the intonation condition when compared to the morphosyntactic condition. The predictor PRAGMATIC EFFECT interacts with AGE. The positive estimates in Table 2 show that 5- (PRAGMATIC EFFECT: OVERRIDE × AGE 5) and 6- year-olds (PRAGMATIC EFFECT: OVERRIDE × AGE 6) performed significantly better on the override condition than 4-year-olds. However, there appear to be no significant differences between 5- and 6-year-olds for the override condition.

3.2 Online results

Online results were obtained from eye tracking data. These results provided a moment by moment record of where children were looking as they heard the sentences relevant to each scene. These data complement the off-line results by showing the looks to the actor associated with either the confirming or overriding response on a millisecond time scale. Thus, eye tracking results show the process of decisions followed by children in the interpretation of the sentences. Eye tracking results are descriptive and do not include statistical analysis but rather describe the children's visual scan path. Off line results, on the other hand, show what the child's final decision was.

3.3 Data treatment

Proportion of looks to the correct actor was calculated for the implicit measure. The horizontal and vertical eye position data obtained from the Tobii Studio Software were used to assess eye position. A value of one was given to every eye tracking sample that fell within the target area of interest; otherwise it was given a zero. As for the implicit eye tracking measure, the target area of interest (AOI) was set as the actor with a changed doll when the context-confirming stimuli were presented. The target AOI was considered the actor without a changed doll when the context-overriding stimuli were presented. As stated above, conditions were presented between-subjects. Twenty-five 4-year-olds, 23 5-year-olds and 34

6-year-olds received the morphosyntactic condition, while 25 4-year-olds, 26 5-year-olds and 27 6-year-olds received the prosodic condition. We present the data grouped by age in Figures 12–14. The vertical line on all figures indicates the offset of the utterance; information shown after this line indicates post-utterance fixation patterns.

3.4 Results by age

3.4.1 4-year-olds

Note that in Figures 12(a)–12(d), the proportion of looks to the actor with the changed doll is higher than those to the actor with the same doll (black line Figures 12(a) and 12(b), grey line in Figures 12(b) and 12(d)). Thus 4-year-olds seem to expect some sort of reaction from the actor with the changed doll. There appears to be an effect of sentential negation in Figure 12(a), where the proportion of looks to the target begin to rise shortly after hearing the particle, and continue to rise throughout the utterance. But this appears to be the only pragmatic effect/ linguistic type combination where this happens for this age group. While looks to the competitor decrease towards the end of the utterance in Figure 12(b), the proportion of looks to the target AOI never exceed the proportion of looks to the competitor. When a prosodic linguistic cue was meant to confirm the listener's
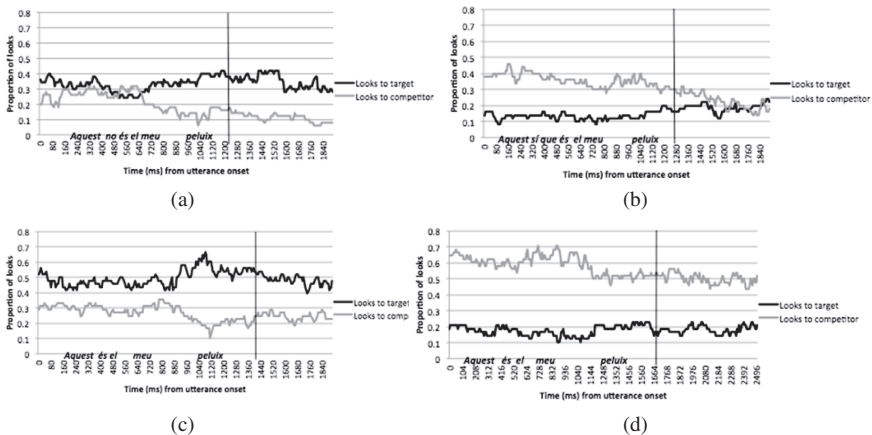

Figure 12: Eye tracking results for 4-year-olds – (a) Proportion of looks to target AOI – confirm/ morphosyntactic – *Aquest no és el meu peluix*. (b) Proportion of looks to target AOI – override/ morphosyntactic – *Aquest sí que és el meu peluix*. (c) Proportion of looks to target AOI – confirm/prosodic – *Aquest és el meu peluix?!?* (d) Proportion of looks to target AOI – override/ prosodic – *Aquest és el meu peluix*.

context-based hypothesis, as Figure 12(c), there is no clear effect of prosody either. Finally in Figure 12(d), it is very clear that the prosodic linguistic cue fails to cause 4-year-olds to revise their context-based hypothesis. They begin with a higher proportion of looks to the actor with the changed doll (just as they did when the linguistic stimulus was meant to confirm their hypothesis) and continue to fixate this actor more after the offset of the utterance.

[Correction added after online publication 1 April 2016: 5-year-olds was changed to 4-year-olds in the caption of figure 12]

3.4.2 5-year-olds

As in the case of the 4-year-olds, 5-year-olds also show an effect of sentential negation, such that it causes an increase in their proportion of looks to the target AOI, since this strategy was intended to confirm a hypothesis that the actor with the changed doll would be the one to speak. This is shown in Figure 13(a), where we see that prior to the utterance's offset, there is an increase in the already-high proportion of looks to the target AOI. The proportion of looks to the target AOI for 5-year-olds is even higher than we saw for 4-year-olds in Figure 12(a). Figure 13(b) shows that 5-year-olds were indeed able to make use of the verum focus particle *sí que* to revise their hypothesis about who would speak. The proportion of looks
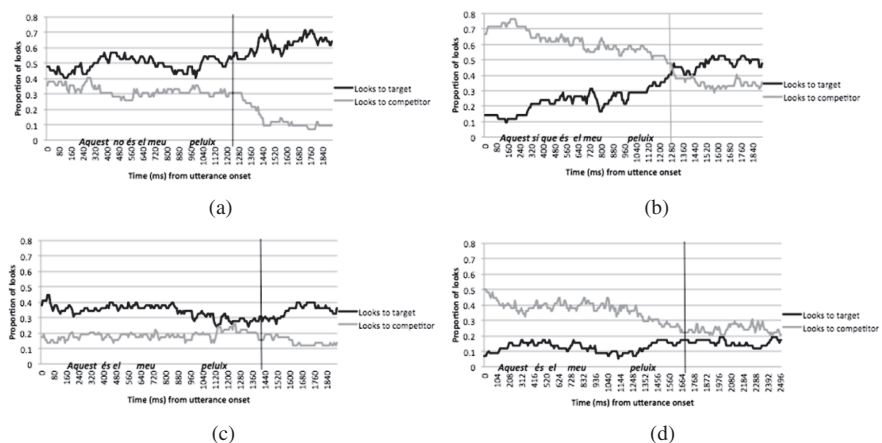

Figure 13: Eye tracking results for 5-year-olds – (a) Proportion of looks to target AOI – confirm/ morphosyntactic – *Aquest no és el meu peluix*. (b). Proportion of looks to target AOI – override/ morphosyntactic – *Aquest sí que és el meu peluix*. (c). Proportion of looks to target AOI – confirm/prosodic – *Aquest és el meu peluix?!?* (d) Proportion of looks to target AOI – override/ prosodic – *Aquest és el meu peluix*.

to the actor with the changed animal (the competitor AOI in this case) start out substantially higher than those to the target. After *sí que* is heard, the proportion of looks to the competitor AOI begin to decrease and those to the target begin to increase, with the proportion of looks to the target exceeding those to the competitor just after the offset of the utterance. When it is the prosodic cue confirming that the actor with the changed doll would speak (Figure 13(c)), we find a less robust effect than we find for sentential negation, as in Figure 13(a). It appears that there is an effect of the linguistic cue, since the proportion of looks to the target increase after the offset of the utterance, but not to the extent that they do when the linguistic type is morphosyntactic. Finally in Figure 13(d), where the listener must revise his/her hypothesis based on linguistic information, we see that although 5-year-olds' looks to the actor with the changed doll decreased as the utterance unfolds, looks to the actor with the same doll never exceed looks to the actor with the changed doll. It does appear that 5-year-olds are somewhat sensitive to the prosodic meaning differences for the confirming vs. overriding intonation however, since in Figure 13(d) we see that the linguistic stimulus did cause the proportion of looks to the competitor to decrease and those to the target increase, unlike what was found for 4-year-olds.

3.4.3 6-year-olds

As expected, Figures 12(a), 13(a) and 14(a) show that like 4- and 5-year-olds, 6-year-olds were able to use sentential negation to confirm an initial context-based hypothesis. Just prior to the last lexical word of the utterance *peluix* 'doll', the proportion of looks to the actor with the changed doll increase and continue to increase after utterance offset, as shown in Figure 14(a). In Figure 14(b), we see that similarly to 5-year-olds, the proportion of looks to the actor with the changed doll (competitor) start to decrease just after *sí que*, which is the morphosyntactic cue intended to revise the context-based hypothesis, though 6-year-olds do not exceed the looks to the competitor until almost 500 ms after the offset of the utterance. It is important to note that neither 5- nor 6-year-olds fixate the target more than the competitor in the override condition (regardless of linguistic type) before the end of the utterance (though we can see a tendency for looks to the target to gradually increase during the utterance). Figure 14(c) shows that prosodic information available during the utterance helps to confirm the context-based hypothesis, with a divergence in looks to the target and competitor shortly before the word *peluix* (and therefore even before the nuclear configuration L⋆ H % was heard). The proportion of looks to the target AOI continues to increase after utterance offset. Finally, as in the case of the 5-year-olds, Figure 14(d) shows
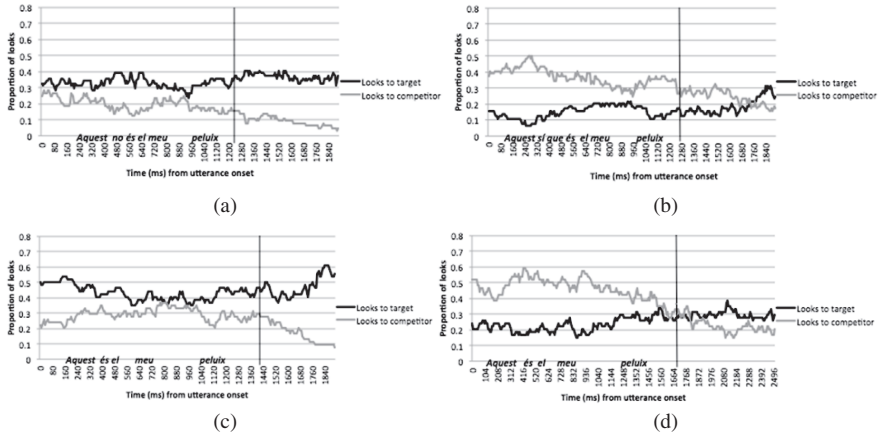
Figure 14: Eye tracking results for 6-year-olds – (a) Proportion of looks to target AOI – confirm/morphosyntactic – *Aquest no és el meu peluix*. (b) Proportion of looks to target AOI – override/morphosyntactic – *Aquest sí que és el meu peluix*. (c) Proportion of looks to target AOI – confirm/prosodic – *Aquest és el meu peluix?!?* (d) Proportion of looks to target AOI – override/prosodic – *Aquest és el meu peluix*.

that the proportion of looks to the target are already increasing mid-utterance when intonation was meant to revise the listener's context-based hypothesis. This continues after the offset of the utterance, and finally the proportion of looks to the target (actor with the same doll) exceed those to the competitor again almost 500 ms after the utterance offset. While this is clearly not by much (roughly 40 % versus 20 % at its largest point of divergence), it is still noteworthy that this difference does not occur with 5-year-olds.

To summarize the eye tracking results, children started off with more looks to the actor with the changed doll than they did the actor with the same doll for all conditions, suggesting that they expected some kind of reaction from this actor. We found differences between how the various age groups respond to the two types of linguistic cues, however. The only linguistic cue that affected 4-year-olds' processing was sentential negation, which increased looks to the target AOI for 4-year-olds prior to the offset of the utterance. With the exception of this cue, 4-year-olds did not seem to make use of the linguistic cues for confirming or overriding their hypotheses. Five-year-olds, on the other hand, were able to make use of the verum focus particle *sí que* to override the initial hypothesis. However, looks to the target AOI did not exceed looks to the competitor until after utterance offset for this condition. Unlike the 4-year-olds, 5-year-olds were able to make use of prosody as a linguistic cue to confirm their hypotheses. But the effect of the prosodic cue for confirming a hypothesis

seemed to be less robust for 5-year-olds when compared to the morphosyntactic cue, since the greatest increase in looks to the target AOI in the prosodic condition did not occur until utterance offset. Differently from the morphosyntactic cue *sí que*, the prosodic cue did not appear to help 5-year-olds in revising their hypotheses. As expected, 6-year-olds were also able to make use of negation to confirm their hypothesis, and like the 5-year-olds also were sensitive to the verum focus particle (though they are a bit faster than 5-year-olds for this cue). Like 5-year-olds, they were also able to make use of prosody to confirm their hypotheses, and also appear to be faster than 5-year-olds in doing so. It seems that 6-year-olds are barely able to make use of the prosodic cue for hypothesis revision, and do not do so until after utterance offset. This behavior is similar to that of the 5-year-olds.

4 General discussion and conclusions

We first asked whether children between the ages of 4 and 6 would be able to form a hypothesis based on contextual information. The results from both the offline and online measures make clear that children are able to make use of information from the discourse context to form a hypothesis. In general, children often showed a tendency to attribute any utterance heard to the actor with the changed doll. Additionally, the eye tracking results show that children of all age groups show a higher proportion of fixations to the actor with the changed doll at the onset of the target utterance, reflecting an expectation that this actor would be the one to speak/react. In this way children were making logical use of world knowledge – they reasoned that the person who had had their doll secretly changed would be the one to say something. Results from the offline (pointing) task showed that when children needed to make a decision about who spoke, pragmatic effect and linguistic type were significant predictors for their success on the task. Thus, while prior studies (Snedeker and Trueswell 2004; Weighall 2008) have not supported children's ability to integrate contextual information, our work points to the consistent use of contextual information by 4-, 5- and 6-year-olds. Choi and Trueswell (2010) discussed children's reliance on early-arriving structures to resolve ambiguity – and we propose that discourse context be included in the types of early-arriving information that children may rely on disproportionately.

Next, we asked whether children would be able to override such a hypothesis based on linguistic information. Overall (when both confirming and overriding a hypothesis), 6-year-olds significantly outperformed 5-year-olds, and 5-year-olds in turn significantly outperformed 4-year-olds. Generally speaking,

however, children performed significantly worse when they had to revise their hypothesis (the override condition) vs. when they were able to confirm it. The eye tracking data allow us to observe that generally speaking, 4-year-olds struggled in using linguistic information to revise their initial hypotheses. They never performed at above-chance levels for the override condition, regardless of the type of linguistic information. Five- and six-year-olds, on the other hand, showed that they were capable of using linguistic information to revise a context-based hypothesis. None of the age groups approach ceiling for the override condition, however, showing that in general it is difficult for children between 4- and 6-years to revise an initial hypothesis by using either of the linguistic cue types investigated in this study.

Finally, we considered whether the type of linguistic information available to the child would affect their ability to override a hypothesis. Linguistic type was a significant predictor of children's success on the task. Children performed significantly worse on the offline task when information was signaled prosodically compared to when it being signaled by a morphosyntactic cue, in this case sentential negation encoded through the particle *no*, or the verum focus particle *sí que*. As mentioned above, 4-year-olds were not adept at using linguistic information to override hypotheses in general. They never performed above chance for the override condition, regardless of whether the linguistic information was morphosyntactic or prosodic, but they were close to ceiling when linguistic information confirmed their hypotheses. Five-year-olds performed above chance when using morphosyntactic information, but at chance when using prosodic information. Six-year-olds performed significantly above chance for the offline task regardless of the type of linguistic material. Eye tracking results showed that for 4-year-olds, the only noticeable effect of linguistic information on proportion of looks was for morphosyntactic information (the particle *no*), and only when this information confirmed the context-based hypothesis. The eye tracking results provide no indication that prosodic cues helped 4-year-olds to confirm their context-based hypotheses. These results show that for 4-year-olds, the morphosyntactic cue *no* was more reliable than the prosodic cue of incredulity intonation, since the former helped to confirm their initial hypothesis and the latter did not. Five-year-olds were able to make use of both morphosyntactic and prosodic cues to confirm the context-based hypothesis, and morphosyntactic cues to override the context-based hypothesis. Five-year-olds thus demonstrate the cognitive control necessary to revise a context-based bias, but this depends on the type of linguistic cue. Finally, the eye tracking data for 6-year-olds show an ability to make use of both linguistic cue types to help them either confirm or revise their hypotheses (albeit to a weaker extent when a prosodic cue is being used to revise the hypothesis). In terms of how quickly

children can make use of these cues, we point out that when revising a hypothesis, 5- and 6-year-olds only cross the threshold to showing a higher proportion of looks the competitor in the override condition (i.e. the actor with the same doll) just after utterance offset (though for 5-year-olds this is only the case with the morphosyntactic cue). In the confirm condition, the increase in proportion of looks to the target actor happened before utterance offset. These slower responses are likely due to the processing time it takes for hypothesis revision. We summarize the offline and online findings in Table 3.

Table 3: Comprehension patterns for 4-, 5- and 6-year olds.

Age (yrs)	MS-C	P-C	MS-O	P-O
4	+	–	–	–
5	+	+	+	–
6	+	+	+	Emerging

Notes: MS-C, morphosyntactic, confirm; P-C, prosodic, confirm; MS-O, morphosyntactic, override; P-O, prosodic, override.

Table 3 makes clear that the youngest children perform best when they are confirming a hypothesis with morphosyntactic information. Note that the performance of 5-year-olds in the override condition mirrors the performance of 4-year-olds in the confirm condition. The specific morphosyntactic cue that 4-year-olds were able to use to confirm their hypothesis was sentential negation. This is interesting in light of findings by Austin et al. (2014), who found that young children understand the denial of a statement before they are able to understand the affirmation of a statement. Summarizing, we can conclude that overriding a hypothesis with prosodic information is the most challenging combination of conditions in the study. Our findings make it quite clear that the type of linguistic cues available for hypothesis revision make a difference in this developmental trajectory. The information visualized in Table 3 also suggests that during these ages constraint-based learning (Trueswell and Gleitman 2004) is occurring. Children are learning to weight and prioritize different types of linguistic cues in more adult-like ways.

It is important to point out here that none of the children included in the study had problems correctly specifying that the actor with the same doll spoke when they heard *Tinc el meu conill, encara* 'I still have my rabbit' in the familiarization session. This alone shows that children *can* use linguistic information to revise a hypothesis, but as we argue, it depends on what that information is. But just how much information is enough? The phrase 'I still have my rabbit' has the adverb *encara* 'still' and the lexical item *rabbit*, making

the task easier for the child. The target utterances for hypothesis revision were less explicit, and required pragmatic inferences as well as reasoning about pragmatic felicity on the part of the participant. Grice's Maxim of Quantity states that one may give as much information as needed and no more. In real life, hearers must make the most out of the linguistic information that is available to them, and search for the most relevant interpretations (Maxim of Relation). Four-year-olds were generally unsuccessful at using either of the target linguistic cues provided in the test stimuli to revise their hypotheses. This changes at Age 5, when children were able to use the morphosyntactic cue, but not the prosodic cue, to override their hypotheses. Thus 4- and to some extent 5-year-olds may ignore linguistic information in favor of early-arriving contextual information. This exhibits some degree of pragmatic tolerance on the part of these age groups. It is more relevant to associate the speaker with the switched doll with the utterance *This is not my doll*, and more relevant for the speaker with the same doll to utter *This sure is my doll*. In turn it would be pragmatically *odd* for the speaker with the same doll to say *This is not my doll*, or for the actor with the switched doll to say *This sure is my doll*. Children very often tended to attribute utterances that would have been pragmatically odd to the actor with the changed doll. We propose that this is due to children's pragmatic tolerance (Katsos et al. 2010; Katsos and Bishop 2011). However, we propose that pragmatic tolerance may be modulated by the specific linguistic strategy for conveying a given message.

Our findings also add to earlier work investigating children's ability to revise hypotheses in language comprehension (Trueswell et al. 1999; Davidson et al. 2006; Weighall 2008; Choi and Trueswell 2010). It appears that between the ages of 4 and 6, children are still in the process of developing the ability to revise hypotheses based on linguistic information, and learning to inhibit an initial interpretation.[3] However, our results suggest that pragmatic tolerance is also playing a role. On the other hand, our results also support a constraint-based learning theory. With more and more linguistic experience, children gain access to the different types of linguistic (and extralinguistic) cues available, and continue to change the ways in which they weight them. We would like to point out, however, that it is possibly too strong to claim that morphosyntactic cues are more reliable than prosodic cues for the language variety we have

---

3 As an anonymous reviewer points out, bilingual children show enhanced cognitive control (Carlson and Meltzoff 2008; Bialystok et al. 2009). As we note in our methods, our participants used very little Spanish in their every day lives. We do not rule out the idea that our participants had enhanced cognitive control, given their linguistic profiles. Future research could reproduce this work with speakers of a language that is not in a contact situation.

investigated here, Central Catalan. Prosodic cues themselves vary in how reliable they are. The intensified affirmative intonation contour we chose in this study might not be as reliable as other intonation contours in the language when pitted against each other. For instance, it was found that Mandarin-acquiring 4-year-olds are just as good as adults at using intonation to process statements vs. questions (Zhou et al. 2012). Prosodic cues indicating sentence modality, for example, might be more reliable when pitted against each other than other prosodic cues, depending on the semantico-pragmatic meaning and frequency of form-meaning pairs. It is also likely that the one-to-many relationships between intonation contours and their meanings could make prosodic cues less reliable for children, and contextually-activated meanings of intonation could present an acquisitional challenge. It will be important for researchers to test children's comprehension of core, semantic meanings of prosody versus more pragmatic/context-specific meanings. Other aspects of prosody like pauses, speech rate, voice quality and duration should be considered as well. Another limitation of our study is that children did not have access to gestural information in the stimuli. Armstrong et al. (2014) showed that 3-year-olds do not differ significantly from 5-year-olds when using facial gesture as a cue to incredulity, but this occurs only in a video-only condition, when children have no access to audio information. We expect then, that gesture would also factor into the cue-weighting process and could even strengthen linguistic cues. There are indeed a plethora of possible cues children must learn to weight and make use of, and this research has just begun to scratch the surface. Further research exploring the possible permutations of such cues in context can be empirically tested, and the methods presented here offer a way to do so.

References

Allopenna, Paul D., James S. Magnuson & Michael K. Tanenhaus. 1998. Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language* 38(4). 419–443.

Altmann, Gerry T. M. & Yuki Kamide. 1999. Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition* 73(3). 247–264.

Armstrong, Meghan E. 2014. Child comprehension of intonationally-encoded disbelief. In Orman, Will & Matthew James Valleau (eds.), *BUCLD 38: Proceedings of the 38th Annual Boston University Conference on Language Development* vol. 2, 25–38. Somerville, MA: Cascadilla Press.

Armstrong, Meghan E., Núria Esteve-Gibert & Pilar Prieto. 2014. The acquisition of multimodal cues to disbelief. *Proceedings of Speech Prosody 7*. Dublin, Ireland, 20–24 May.

Arnold, Jennifer E., Janet G. Eisenband, Sarah Brown-Schmidt & John C. Trueswell. 2000. The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition* 75. B13–B26.

Austin, Keith, Anna Theakston, Elena Lieven & Michael Tomasello. 2014. Young children's understanding of denial. *Developmental Psychology*. Advance online publication. http://dx.doi.org/10.1037/a0037179.

Bates, Elizabeth. 1974. The acquisition of pragmatic competence. *Journal of Child Language* 1. 277–283.

Bialystok, Ellen, Fergus I. M. Craik, David W. Green & Tamar H. Gollan. 2009. Bilingual minds. *Psychological Science in the Public Interest*, 10(3). 89–129.

Bibyk, Sarah, Kiwako Ito, Laura Wagner & Shari R. Speer. 2009. Children can use contrastive pitch accent in on-line processing. Paper presented at the Boston University Conference on Language Development 34, Boston, MA.

Carlson, Stephanie M. & Andrew M. Meltzoff. 2008. Bilingual experience and executive functioning in young children. *Developmental Science* 11(2). 282–298.

Choi, Youngon & John C. Trueswell. 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.

Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press.

Cooper, Roger M. 1974. Control of eye fixation by meaning of spoken language: New methodology for real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology* 6. 84–107

Cruttenden, Alan. 1974. An experiment involving comprehension of intonation in children from 7 to 10. *Journal of Child Language* 1. 221–231.

Cruttenden, Alan. 1985. Intonation comprehension in ten-year-olds. *Journal of Child Language* 13. 643–661.

Cutler, Anne & David A. Swinney. 1987. Prosody and the development of comprehension. *Journal of Child Language* 14. 145–167.

Dahan, Delphine, James S. Magnuson, & Michael K. Tanenhaus. 2001. Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology* 42. 317–367.

Dahan, Delphine & Michael K. Tanenhaus. 2004. Continuous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30. 498–513.

Davidson, Matthew C., Dima Amso, Loren Cruess & Adele Diamond. 2006. Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia* 44. 2037–2078.

Esteve-Gibert, Núria. 2014. *The integration of prosody and gesture in early intentional communication*. Barcelona, Spain: Universitat Pompeu Fabra dissertation.

Grice, H. P. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and semantics volume 3: Speech acts*, vol. 3, 48–57. New York: Academic Press.

Hald, Lea. 2003. *The integration of semantic versus world knowledge during on-line sentence comprehension*. Tucson, AZ: University of Arizona dissertation.

Hald, Lea, Esther Steenbeek-Planting & Peter Hagoort. 2007. The interaction of discourse context and world knowledge in online sentence comprehension. Evidence from the N400. *Brain Research* 1146. 210–218.

Hagoort, Peter, Hald Lea, Bastiaansen Marcel & Peterson Karl Magnues. 2004. Integration of word meaning and world knowledge in language comprehension. *Science* 304. 438–441.

Hornby, Peter A. & Wilbur A. Hass. 1970. Use of contrastive stress by preschool children. *Journal of Speech and Hearing Research* 13. 395–399

Ito, Kiwako, Nobuyuki Jinho, Utako Minai, Naoto Yamane & Reiko Mazuka. 2012. Intonation facilitates contrast resolution: Evidence from Japanese adults and 6-year olds. *Journal of Memory and Language* 66. 265–284

Ito, Kiwako, Sarah Bibyk, Laura Wagner & Shari R. Speer. 2014. Interpretation of contrastive pitch accent in 6- to 11-year-old English speaking children (and adults). *Journal of Child Language* 41(1). 84–110.

Jackendoff, Ray. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

Kamide, Yuki, Gerry T. M. Altmann & Sarah L. Haywood. 2003. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language* 49. 133–159.

Katsos, Napoleon & Nafsika Smith. 2010. Pragmatic tolerance and speaker-comprehender assymetries. In Katie Franich, Kate M. Iserman & Lauren L. Keil (eds.), *Proceedings of the 34th Boston university conference in language development*, 221–232. Somerville, MA: Cascadilla Press.

Katsos, Napoleon & Dorothy V.M. Bishop. 2011. Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition* 20. 67–81.

Macdonald, Maryellen C., Neil J. Pearlmutter & Mark S. Seidenberg. 1994. The lexical nature of syntactic ambiguity resolution. *Psychological Review* 101(4). 676–703.

MacWhinney, Brian & Derek Price. 1980. The development of the comprehension of topic-comment marking. In David Ingram, Fred C.C. Peng & Phillip S. Dale (eds.), *Proceedings of the first international congress for the study of child language*. Lanham MD: University Press of America.

MacWhinney, Brian, Csaba Pléh & Elizabeth L. Bates. 1985. The development of sentence interpretation in Hungarian. *Cognitive Psychology* 17. 378–209.

Mampe, Birgit, Angela D. Friederici, Anne Christophe & Kathleen Wermke. 2009. Newborn's cry melody is shaped by their native language. *Current Biology* 19(23). 1994–1997.

Marlen-Wilson, William & Lorraine K. Tyler. 1980. The temporal structure of spoken language understanding. *Cognition* 8(1). 1–71.

Mehler, Jacques, Peter W. Jusczyk, Ghislaine Lambertz, Nilofar Halsted, Josiane Bertoncini & Claudine Amiel-Tison. 1988. A precursor of language acquisition in young infants. *Cognition* 29(2). 143–178.

Nieuwland, Mante S. & Jos J. A. Van Berkum. 2006. When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience* 18(7). 1098–1111.

Novick, Jared M., John C. Trueswell & Sharon L. Thompson-Schill. 2005. Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective & Behavioral Neuroscience* 5(3). 263–281.

Prieto, Pilar, Joan Borràs-Comes, Teresa Cabré, Verònica Crespo-Sendra, Ignasi Mascaró, Paolo Roseano, Rafeu Sichel-Bazin & Maria del Mar Vanrell. 2015. Intonational phonology of Catalan and its dialectal varieties, 9–62. In Sónia Frota & Pilar Prieto (ed.), *Intonational variation in romance*. Oxford: Oxford University Press.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Sekerina, Irina A. & John C. Trueswell. 2012. Interactive processing of contrastive expressions by Russian children. *First Language* 32(1–2). 63–87.

Snedeker, Jesse & John C. Trueswell. 2004. The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology* 49(3). 238–299.

Snedeker, Jesse & Sylvia Yuan. 2008. Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language* 58. 574–608

Soderstrom, Melanie, Amanda Seidl, Deborah G. Kemler Nelson & Peter W. Jusczyk. 2003. The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language* 49. 249–267.

Soderstrom, Melanie, Eon-Suk Kob & Uliana Nevzorovaa. 2011. It's a question? Infants attend differently to yes/no questions and declaratives. *Infant Behavior and Development* 34(1). 107–110.

Sperber, Dan & Deirdre Wilson. 1987. Precis of relevance: Communication and cognition. *Behavioral and Brain Sciences* 10. 697–754.

Spivey, Michael, Michael K. Tanenhaus, Kathleen M. Eberhard & Julie C. Sedivy. 2002. Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology* 45). 447–481.

Tanenhaus, Michael K. & John C. Trueswell. 1995. Sentence comprehension. In Joanne L. Miller & Peter D. Eimas (eds.), *Handbook of perception and cognition Vol. 11: Speech language and communication*, 217–262. San Diego, CA: Academic Press.

Tanenhaus, Michael K., Michael J. Spivey-Knowiton, Kathleen M. Eberhard & Julie C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268. 1632–1634.

Trueswell, John C., Irina Sekerina, Nicole M. Hill & Marian L. Logrip. 1999. The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition* 73(2). 89–134.

Trueswell, John C., & Lila R. Gleitman. 2004. Children's eye movements during listening: Evidence for a constraint-based theory of parsing and word learning. In John M. Henderson & Fernanda Ferreira (eds.), *Interface of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.

Trueswell, John C. & Lila R. Gleitman. 2007. Learning to parse and its implications for language acquisition. In Mark Gareth (ed.), *Oxford handbook of psycholinguistics*. Oxford, UK: Oxford University Press.

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

Weighall, Anne R. 2008. The kindergarten path effect revisited: Children's use of context in processing structural ambiguities. *Journal of Experimental Psychology* 99(2). 75–95.

Wells, Bill, Sue Peppé & Nata Goulandris. 2004. Intonation development from five to thirteen. *Journal of Child Language* 31. 749–778.

Zhou, Peng, Stephen Crain & Likan Zhan. 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. 149–164