



# DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

## Effects of Prosodic and Lexical Constraints on Parsing in Young Children (and Adults)

The Harvard community has made this article openly available.  
[Please share](#) how this access benefits you. Your story matters.

<b>Citation</b>	Snedeker, Jesse and Sylvia Yuan. 2008. Effects of prosodic and lexical constraints on parsing in young children (and adults). <i>Journal of Memory and Language</i> 58, no. 2: 574-608.
<b>Published Version</b>	<a href="https://doi.org/10.1016/j.jml.2007.08.001">doi:10.1016/j.jml.2007.08.001</a>
<b>Accessed</b>	February 17, 2015 11:31:44 PM EST
<b>Citable Link</b>	<a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:3207705">http://nrs.harvard.edu/urn-3:HUL.InstRepos:3207705</a>
<b>Terms of Use</b>	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP</a>

*(Article begins on next page)*

**Effects of prosodic and lexical constraints on parsing in young children (and adults)**

Jesse Snedeker

Harvard University

Sylvia Yuan

University of Illinois

Address Correspondence to:

Jesse Snedeker

1136 WJH, 33 Kirkland St.

Cambridge MA, 02138

[snedeker@wjh.harvard.edu](mailto:snedeker@wjh.harvard.edu)

### Abstract

Prior studies of ambiguity resolution in young children have found that children rely heavily on lexical information but persistently fail to use referential constraints in online parsing (Trueswell, Sekerina, Hill & Logrip, 1999; Snedeker & Trueswell, 2004). This pattern is consistent with either a modular parsing system driven by stored lexical information or an interactive system which has yet to acquire low-validity referential constraints. In two experiments we explored whether children could use a third constraint—prosody—to resolve globally ambiguous prepositional-phrase attachments (“You can feel the frog with the feather”). Four to six years olds and adults were tested using the visual world paradigm. In both groups the fixation patterns were influenced by lexical cues by around 200ms after the onset of the critical PP-object noun (“feather”). In adults the prosody manipulation had an effect in this early time window. In children the effect of prosody was delayed by approximately 500 ms. The effects of lexical and prosodic cues were roughly additive: prosody influenced the interpretation of utterances with strong lexical cues and lexical information had an effect on utterances with strong prosodic cues. We conclude that young children, like adults, can rapidly use both of these information sources to resolve structural ambiguities.

Keywords: modularity; parsing; prosody; syntactic ambiguity resolution; children's language comprehension; eye movements

### **Effects of prosodic and lexical constraints on parsing in young children (and adults)**

The present study explores how adults and children combine information about the prosodic structure of an utterance and information from individual words to guide syntactic analysis. Several recent studies have demonstrated that prosody has rapid effects on parsing in adults (see e.g., Kjelgaard & Speer, 1999; Steinhauer, Alter, & Friederici, 1999; Snedeker & Trueswell, 2003). But prior research has found little or no effect of prosody on children's interpretation of ambiguous sentences (e.g., Snedeker & Trueswell, 2001; Vogel & Raimy, 2002; Choi & Mazuka, 2003). In these prior studies, however, the lexical content of the utterances and its effects on parsing were not assessed, raising the possibility that children might be sensitive to prosody, but consult it only when lexical constraints are weak. The experiments that follow explore the interaction of prosodic and lexical information in the comprehension of ambiguous sentences in both children and adults.

Our interest in this question is three-fold. First, determining whether young children use prosody in online sentence comprehension could illuminate the architecture of language processing at a critical stage of development. Second, clarifying the role of prosody in children's online processing may help us understand the developmental origins of the links between prosody and syntax. Finally, while there is ample evidence that adults use prosody to parse spoken utterances, there is little data on how prosody is integrated with other cues (but see Pynte & Prieur, 1996). Does prosody chop up utterances for further processing, overriding subsequent sources of information? Or is it merely used to revise or strengthen analyses that were originally proposed by other information sources? In the remainder of the introduction we briefly review the relevant aspects of adult sentence processing, and then explore each of these issues in turn.

### **Setting the stage**

For decades psycholinguists have explored how adult listeners (and readers) recover the syntactic structure of a sentence from a string of words. Much of this research has focused on understanding the kinds of information that are used in the process, when they become available, and how they interact. These questions have primarily been examined by investigating the way readers initially interpret, and misinterpret, syntactically ambiguous phrases. For example, consider the sentence fragment (1) below:

(1) Allison ate the cake with the...

At this point in the utterance the prepositional phrase (PP) beginning with *with* is ambiguous: It could be attached to the verb *ate* (VP-attachment), indicating an instrument (e.g., *with the fork*); or it could be attached to the definite noun phrase *the cake* (NP-attachment) indicating a modifier (e.g., *with the pink icing*). In adults, several kinds of information rapidly influence the interpretation of ambiguous phrases. Three of these are particularly relevant for our discussion. First, knowledge about the particular words in the sentence constrains online interpretation (see e.g., Taraban & McClelland, 1988; Trueswell, Tanenhaus & Kello, 1993). For instance, the sentence in (2) favors the VP-attachment, but if we change the verb from *hit* to *liked* (as in 3) the preference flips and the modifier analysis, or NP-attachment, is favored.

(2) Allison hit the cake with the...

(3) Allison liked the cake with the...

Second, under some circumstances adults can use intonation to resolve attachment ambiguities. The presence of prosodic break or a pause before the preposition will support the VP-attachment (4) while the presence of a break before the direct object favors an NP-attachment (5) (Pynte & Prieur, 1996; Schafer, 1997).

(4) Allison ate the cake / with a butcher knife

(5) Allison ate / the cake with the chocolate ganache

Finally, the situation or referential context in which the utterance is used can have an effect (Crain & Steedman, 1985). While the context can constrain interpretation in many ways, research in this area has focused on manipulating the number of potential referents that are available for a definite noun phrase. In the example above, if only one cake is present in the discourse then the VP-attachment is often preferred, but if multiple cakes are available then readers are more likely to initially interpret the ambiguous phrase as a modifier specifying the cake in question (e.g., Altmann & Steedman, 1988; van Berkum, Brown & Hagoort, 1999, but see Ferreira & Clifton, 1986).

In reading studies, such referential constraints typically take a back seat to strong lexical constraints (e.g., Britt, 1994; Spivey-Knowlton & Sedivy, 1995). However Tanenhaus, Spivey and colleagues found that, in a world-situated spoken-language comprehension task, referential cues prevailed over strong countervailing lexical biases (Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995; Spivey, Tanenhaus, Eberhard & Sedivy, 2002). When participants heard utterances like (6) in the presence of just one apple, they initially interpreted the first prepositional phrase (*on the napkin*) as a destination. But when two apples were provided (one of which was on a napkin) the participants were able to immediately use the referential context to overcome the strong bias of the verb and avoid this garden path, resulting in eye movements similar to unambiguous controls (e.g., *Put the apple that's on the napkin in the box*).

(6) Put the apple on the napkin in the box.

Much of the research on adult sentence processing has focused on questions about time course: Are initial structural hypotheses influenced by all of these information sources? Or, does the architecture of the comprehension system or the nature of the data source force us to exclude some of this information during the earliest stages of processing? Currently the bulk of the evidence suggests that adults rapidly integrate these different information sources to arrive at the

analysis that best meets the constraints they have encountered (for reviews see Tanenhaus & Trueswell, 1995; Altmann, 1998). But disputes continue about how this integration occurs: Do some sources of information establish candidate analyses while other sources of information weigh in at a later stage (see e.g., Boland & Cutler, 1996; Pynte & Prieur, 1996)?

### **What we know about children's parsing**

Our experiments focus on the parsing abilities of children between four and six years of age. We have chosen this age for several reasons. First, by four children's language comprehension and production appear, to the naked eye, to be almost adult-like. Yet on a number of cognitive dimensions children this age are quite different from adults. For example, they are notoriously poor at tasks which invoke multiple representations of a single entity or require the inhibition of prior or dominant responses (Piaget, 1946; Flavell, 1986; Welsh, Pennington & Groisser, 1991; Permer & Wimmer, 1985). They also have smaller memory spans than adults or older children (for review see Schneider & Bjorklund, 1998). The increase in memory span across development is accompanied by an increase in processing speed across a wide range of tasks, raising the possibility of causal link (Kail, 1991; Kail & Park, 1994).

These cognitive differences could have profound implications for syntactic parsing. For example, in adults individual differences in working memory performance are correlated with qualitative differences in parsing during reading (Just & Carpenter, 1992; MacDonald, Just & Carpenter, 1992). Readers with low memory spans are less able to integrate contextual cues or consider multiple analyses of an ambiguity. Presumably parallel limitations in children could shape their spoken language comprehension. Similarly, a global slow down in processing speed, in face of a constant speech rate, might well limit the amount of processing that is possible as each word is spoken and integrated into the sentence.

The present experiment builds on two prior studies exploring syntactic ambiguity resolution in children (Trueswell, Sekerina, Hill & Logrip, 1999; Snedeker & Trueswell, 2004). Taken together they indicate that children's parsing is rapidly influenced by lexically-specific information but is relatively impervious to referential context. In the first of these studies, Trueswell and colleagues explored the use of referential constraints in a study which closely paralleled the Tanenhaus and Spivey experiment described above. In contrast to the adults, five year olds (but not eight-year olds) blindly pursued the VP-attachment analysis, ignoring referential information.

Trueswell and colleagues offered two explanations for the overwhelming VP-attachment preference on the part of young children. First, this preference could be driven by the children's statistical knowledge of the verb *put*, which strongly supports the presence of a PP-argument. This explanation would be consistent with lexicalist theories and constraint-satisfaction theories more generally (e.g., MacDonald, Pearlmutter & Seidenberg, 1994; Trueswell & Tanenhaus, 1994). Second, five year olds might have been exhibiting a general structural preference for VP-attachment. Such a preference would be predicted by theories of acquisition and parsing that favor simple syntactic structures (i.e., a Minimal Attachment strategy, Goodluck & Tavakolian, 1982; Frazier & Fodor, 1978) or ban complex syntactic operations entirely in the early stages of development (Frank, 1998).

Snedeker and Trueswell (2004) explored these possibilities by fully crossing the statistical preferences of the verb and the number of potential referents for the direct-object noun (S&T 2004 hereafter). The target sentences contained globally ambiguous prepositional phrase attachments, like (7) below. These sentences were presented with sets of toys that provided distinct referents for the prepositional object under each of the two possible analyses. For example in 7b both a large feather and frog holding a feather were provided.



- (7) a. Choose the cow with the fork
- b. Feel the frog with the feather
- c. Tickle the pig with the fan

Both adults and five-year old children were strongly swayed by the type of verb that was used in the instructions. When the verb was one that frequently appeared with an instrument phrase (7c), participants began looking at the potential instrument (e.g., a large fan) shortly after the onset of the prepositional object. When the verb was strongly biased to a modifier analysis (7a), participants focused in on the animal holding the object instead. These lexical biases largely determined the ultimate interpretation that the participants assigned to the prepositional phrase and hence their actions (see Kidd & Bavin, 2005 for converging evidence). While adults also incorporated referential information into their analyses, children showed little sensitivity to this manipulation.

This strong reliance on lexical information clearly rules out the possibility that children's interpretation relies on a general parsing heuristic (e.g., minimal attachment) that diminishes with age or experience. Instead the work to date demonstrates a near-exclusive role for lexical evidence in informing children's parsing decisions. This is compatible with three different accounts of the development of parsing.

1. *The lexical modularity hypothesis*: The observed differences between children and adults could reflect architectural changes brought on by expansions in processing ability. For instance, a limited, single-cue, or encapsulated parsing system might become more interactive as processing ability grows with age. Indeed, several theories of parsing grant an architectural privilege to lexical cues. For example, Boland and colleagues have argued that the lexicon alone proposes syntactic and semantic structures while other cues are used at a later stage to select between the proposed analyses (Boland & Cutler, 1996).

2. *The cue validity hypothesis:* Alternately, it is possible children have a probabilistic multiple-cue comprehension system from the start, but the order in which the cues are acquired depends largely on their relative reliability (for a related proposal see Bates & MacWhinney, 1987). Under this account, children might show an initial reliance on lexical cues simply because they are a highly reliable source of information about syntactic structure. Work in computational linguistics demonstrates that lexical cues are highly predictive of local structure (e.g., Collins, 1997; Collins & Brooks, 1995; Marcus, 1994) while studies of infant-directed speech demonstrate that this information is robustly present in children's input (Lederer, Gleitman, and Gleitman, 1995). Other constraints on syntactic structure, such as the need to resolve referential ambiguity, could simply take longer to acquire because they are less reliable in the input database as a whole, and arguably more difficult to track than lexico-syntactic contingencies. Both experimental and theoretical work suggests that the number of referents in a scene is a poor predictor of structure. Although adults understand that a definite NP almost always requires a unique (and agreed-upon) referent, disambiguation of the referent need not be accomplished linguistically, since the local discourse and the goals of the interlocutors often provide the necessary information (e.g., Hawkins, 1978; Prince, 1981). In a referential communication task, Brown-Schmidt and colleagues found that almost half of all definite NPs uttered (e.g., "Now, move the triangle") did not have a unique referent in the scene (Brown-Schmidt, Campana & Tanenhaus, 2005). Participants, however, often had no difficulty identifying the correct referent, presumably because prior statements and the goals of the task had narrowed the field of possibilities down to one. Conversely, Engelhardt and colleagues found that speakers describing unique referents provided unnecessary post-nominal modifiers on almost one third of the trials (Engelhardt, Bailey & Ferreira, 2006). Surprisingly, naïve listeners judged these over-informative descriptions to be as adequate as the more concise simple nouns.

3. *The bottom-up hypothesis:* Finally, the results to date are consistent with a hypothesis that lies between the two extremes discussed so far. Children could have a probabilistic multiple-cue comprehension system but be unable to make use of some information types—regardless of their validity—due to architectural or processing constraints. While constraint satisfaction models propose that multiple types of information have rapid and converging effects on parsing, many such models recognize that different kinds of constraints stem from distinct levels of representation which emerge from distinct processing paths (e.g., Trueswell & Tanenhaus, 1994). Thus developmental differences in the use of different cues could arise from differences in the maturation or efficiency of these processing paths or differences in the relation between the constraining representation and the syntactic representation itself.

Referential context is most readily conceived of as a top-down constraint on syntactic structure. The relevance of reference world depends on several aspects of the linguistic analysis that is under construction: 1) whether a definite determiner was used with the noun (“a bird” is sufficient even when two are present); 2) which noun was produced (“the heron” may be adequate even when “the bird” is not); and 3) whether a prenominal modifier has already disambiguated the referent, and 4) whether the ambiguous phrase is in a syntactic context where NP-attachment is possible (e.g. “pick the dog up with the hat” vs. “pick up the dog with the hat”). In short, it is difficult to imagine how referential constraints could be calculated without some assembly—and semantic evaluation—of the structural hypotheses under consideration.

Top-down constraints may pose problems for young children. To make use of top-down information the child must rapidly construct the syntactic alternatives, information about these alternatives must propagate up to the higher level representation, and then constraints from this higher-level representation must filter back down to the syntax. If children are slower to activate the alternatives, or to pass information from one level to the next, then the syntactic competition

is likely to be resolved before the top-down information has arrived. There is ample evidence that processing speed increases in middle childhood across a variety of tasks (see e.g., Kail, 1991).

Admittedly, there is limited experimental evidence to support our suggestion that top-down influences on bottom-up processing are generally late to develop (but see Chase & Tallal, 1990; James, 2001). Our analysis, however, receives some support from research on the effects of time manipulations in adult language processing. A global slow-down in processing speed (in the presence of a constant speech rate) is roughly parallel to an experimental manipulation in which time pressure is increased by creating a response deadline or increasing the speech rate. In both cases less processing can occur before the next word is encountered or production is initiated. Dell (1986) found that time pressure resulted in a decrease in the effect of a higher level representation (lexical items) or a lower level process (phoneme selection). These findings were tidily captured by a system of word production that shares many of the features of the constraint satisfaction models described above (e.g., distinct levels of representation, interactive processing, and bi-directional connections between levels).

The experiments that follow test the lexical-modularity hypothesis, by exploring children's use of a third constraint on syntactic structure: the prosodic structure of the utterance. There are several reasons why prosodic cues might become available earlier in ontogenetic time than the referential cues explored in the prior experiments. First, while informative prosodic breaks are often absent from short utterances, when they are present they are highly reliable, making them a useful predictor of structure (Snedeker & Trueswell, 2003, see discussion below). Second, while referential context exerts a top-down influence on syntactic parsing, prosodic cues are arguably a bottom-up constraint. By this we merely mean that, from the perspective of the listener, the prosodic structure may constrain syntactic structure but it is not dependent upon it (for discussion see, Kjelgaard & Speer, 1999). This asymmetry could influence the parser's ability to gain

timely access to this information during comprehension. Finally, referential and prosodic representations may develop at different rates, which could influence the age at which they become integrated with online parsing. Many have argued that five-year olds are still struggling to understand the referential demands and goals of various communicative situations (Glucksberg & Krauss, 1967; Robinson & Robinson, 1982). In contrast, even young infants show a well-developed sensitivity to the prosodic structure of their language.

### **Do children use prosodic structure to resolve syntactic ambiguities?**

Prosody clearly plays a central role in infant speech perception. Newborns discriminate between languages on the basis of their rhythmic properties (Mehler, Jusczyk, , Lambertz, Halsted, Bertoni & Amiel-Tison, 1988; Jusczyk, Frederici, Wessels, Svenkerud, & Jusczyk, 1993). Half a year later, infants rely heavily on prosodic cues to segment the speech stream into words (Jusczyk, Culter, & Redanz, 1993; Morgan & Saffran, 1995; Johnson & Jusczyk, 2001). Critically, infants are also sensitive to the coalition of cues that mark the prosodic boundaries between groups of words (Hirsh-Pasek, Kemler Nelson, Jusczyk, Wright-Cassidy, Druss & Kennedy, 1987; Jusczyk, Hirsh-Pasek, Kemler Nelson, Kennedy, Woodward & Piwoz, 1992; Christophe, Mehler & Sebastián-Gallés, 2001). Because the prosody of an utterance depends in part on its syntactic structure, many theorists have suggested that infants might use their extensive experience with prosody to bootstrap their way into syntax (see e.g., Hirsh-Pasek et al., 1987; Morgan, 1996; Christophe, Guasti, Nespor, van Ooyen, 2003). If so, we might expect that prosody would continue to serve as parsing cue for preschoolers (Morgan, 1996; Choi & Mazuka, 2003). However, others have pointed out that the relation between prosody and syntax may be too variable or weak to support a prosodic comprehension strategy (Gerken, 1996; Fernald & McRoberts, 1996).

Recent research favors the skeptics; several studies have found little or no effect of prosody on children's interpretation of structurally ambiguous sentences (Halbert, Crain, Shankweiler & Woodams, 1995; Snedeker & Trueswell, 2001; Vogel & Raimy, 2002; Choi & Mazuka, 2003).<sup>1</sup> For example, Choi and Mazuka tested Korean speaking children on two kinds of ambiguous utterances: word-segmentation ambiguities and syntactic grouping ambiguities. Both types of utterances were disambiguated with the same kind of prosodic cues (an intonational phrase boundary marked by a pause and a boundary tone). Three to four year olds performed well on the word-segmentation ambiguities but were at chance on the syntactic ambiguities, while adults performed well in both tasks. Choi and Mazuka concluded that while children can clearly detect prosodic boundary cues they fail to use them to resolve syntactic ambiguities.

These results are consistent with our own initial exploration of children's ability to use prosody to resolve syntactic ambiguity (Snedeker & Trueswell, 2001, see Appendix A). In our study, a referential communication task was used to simultaneously explore whether mothers would provide systematic prosodic cues to the structure of ambiguous utterances and whether their children (ages 4-6) would use them in comprehension. The mothers, like the college students in our previous studies, varied their prosody systematically depending on the intended attachment of the prepositional phrase. Their children, however, performed at chance, suggesting that they were unable to use prosody to constrain parsing (see Appendix A).

But two features of this experiment lead us to question our results, and by extension, the prior findings. First, all of the children in our study showed a systematic preference for a single response type—some preferred the instrument analysis, others the modifier analysis, but they were all quite consistent. In our study prosody was manipulated within each subject and was not blocked. In such a design, a strong tendency to perseverate across trials could easily wipe out a small or fragile effect of prosody. Second, in this initial study the lexical biases of the target

sentences were not systematically controlled, raising the possibility that strong biases in individual items may have overwhelmed the effects of prosody. To the best of our knowledge these two features were shared by the other studies that have failed to find an effect of prosody on children's interpretation of globally ambiguous utterances (Halbert et al., 1995; Vogel & Raimy, 2002; Choi & Mazuka, 2003). Experiment 1 examines children's use of prosody when these two factors are controlled.

### **Prosody in adult sentence processing**

In contrast with children, adults have a robust ability to use prosodic cues to interpret globally ambiguous utterances (Lehiste 1973; Lehiste, Olive & Streeter 1976; Cooper & Paccia-Cooper 1980; Price, Ostendorf, Shattuck-Hufnagel, & Fong, 1991; Schafer, 1997; Carlson, Clifton & Frazier, 2001; see Cutler, Dahan & van Donselaar, 1997 for review). In fact, there is now a substantial body of evidence that prosody has a rapid effect on online sentence processing as well (Marslen-Wilson, Tyler, Warren, Grenier, & Lee, 1992; Nagel, Shapiro, Tuller, & Nawy, 1996; Pynte & Prieur, 1996; Kjelgaard & Speer, 1999; Steinhauer, Alter & Frederici, 1999; Snedeker & Trueswell, 2003; Weber, Grice & Crocker, 2006).

Despite this robust evidence that prosody can affect parsing, there is considerable controversy over the exact role that it plays. One controversy centers around the question of whether the prosodic manipulations used in psycholinguistic experiments are reflective of the prosodic cues provided in natural speech. Naïve readers often fail to prosodically disambiguate globally ambiguous sentences (Allbritton, McKoon & Ratcliff, 1996; Wales & Toner, 1979) raising the possibility that prosodic cues to structure are infrequent and perhaps unreliable. More recent studies using referential communication tasks have been somewhat more optimistic (Snedeker & Trueswell, 2003; Schafer, Speer & Warren, 2005; Kraljic & Brennan, 2005). All of these studies explored the disambiguation of PP-attachment ambiguities. While they differ in

their conclusions, the results converge in three respects. First, all three studies find evidence for reliable prosodic disambiguation when the situational context supports both readings of the ambiguous utterance. Second, despite differences in the length and structure of the utterances there is remarkable consistency in the nature of the prosodic cues that the speakers produce. VP-attachments were generally accompanied by a strong prosodic break immediately before the prepositional phrase. These breaks were reduced or absent in NP-attachments. In both the Schafer and Snedeker studies, NP-attachments were often produced with a substantial prosodic break earlier in the utterance, though the location of this break was variable in the Schafer study and its presence was less reliable in the Snedeker study. Finally, all three studies demonstrate that adult listeners can use these prosodic cues to arrive at the correct interpretation of these otherwise ambiguous utterances. In referential contexts in which utterance is unambiguous, however, the findings of the three studies diverge. Snedeker and Trueswell found that prosodic cues were substantially weaker in unambiguous contexts (perhaps because the speakers failed to notice the structural ambiguity). In contrast the other two studies found reliable prosodic disambiguation even when the utterance was unambiguous in context. These divergent findings may be attributable to differences in the length of the utterances that were used, their syntactic complexity, the nature of the communication task, and the way in which referential ambiguity was manipulated (for discussion see Snedeker & Trueswell, 2003). Nevertheless, the research to date clearly indicates that in referentially ambiguous contexts, like those in the current study, naïve speakers will produce predictable patterns of prosodic phrasing which effectively disambiguate PP-attachment ambiguities, even when those ambiguities are in short simple utterances, much like those in the current study.

A second controversy centers on the relation between prosodic cues and structural parsing preferences or lexical biases. While some observers suggest that prosody is used in the earliest



stages of syntactic analysis (Marslen-Wilson et al., 1992; Kjelgaard & Speer, 1999), others claim that it is used at a later stage (Pynte & Prieur, 1996; Marcus & Hindle, 1990). Only one study has directly explored the interaction of prosodic and lexical information in online parsing (Pynte & Prieur, 1996). In these experiments, a word detection task was used to examine the processing of structurally-ambiguous prepositional phrases. The utterances contained either ditransitive or monotransitive verbs, and verb type was fully crossed with both the disambiguation of the utterance (NP-attached or VP-attached) and the presence or absence of a prosodic break before the preposition. Both prosody and argument-structure preferences affected the ease with which participants interpreted NP-attached and VP-attached prepositional phrases. The prosodic effects, however, occurred only when the argument structure cues conflicted with the resolution of the ambiguity. When the lexical cues were consistent with the disambiguation, the effects of prosody disappeared. This led the authors to suggest that lexical information proposes an analysis, while prosody merely plays a role in revision.

There are two reasons to be cautious in accepting this conclusion. First, the conditions in which lexical cues were consistent with the disambiguation had substantially lower reaction times than the others, suggesting that the interaction may have been caused by a floor effect. Reaction times in a dual task paradigm like this depend both on the difficulty of the comprehension task and on the time it takes to complete the overt task. When syntactic processing is relatively easy, variations in difficulty could be absorbed into the slack introduced by the demands of the word detection task. Second, subsequent studies have found early effects of prosody on both the preferred and dispreferred interpretation of closure ambiguities, suggesting that the role of prosody cannot be limited to revision (Kjelgaard & Speer, 1999).

As this example illustrates, interpreting the literature on prosody and online processing is complicated by the nature of the paradigms that are used. The most common are cross-modal

lexical decision, cross-modal naming, and speeded judgment tasks, all of which have been criticized for their poor temporal resolution and artificiality (Carlson et al., 2001). Even those experiments using more naturalistic tasks (e.g. ERPs, Steinhauer et al., 1999) employ designs which provide limited information about the time course of prosodic influence. The experiments to date manipulate the consistency of the prosodic contour with subsequent morphosyntactic information, and then measure effects of prosody at or after the disambiguation point (three to ten syllables after the onset of the ambiguity). While these measures can clearly support inferences about the outcome of processes that occurred in the ambiguous region, they are mute about the temporal structure of those processes.

The current study uses the visual word paradigm to explore the interaction between prosodic and lexical constraints in online processing (Tanenhaus et al., 1995). While this technique also has its limitations—most notably all utterances must refer to depictable objects and events in a tightly constrained reference world—it has the advantage of providing a fine-grained measure of online interpretation that can be continuously monitored from the onset of the ambiguous region through the completion of the sentence. In earlier work, we employed this technique to examine the use of prosodic cues by listeners in a referential communication task (Snedeker & Trueswell, 2003). Surprisingly, effects of prosody on online interpretation appeared before the onset of the ambiguous prepositional phrase, suggesting that in some circumstances prosody can be used to predict the content of an upcoming phrase.

Experiment 1 explores whether children can use prosodic cues to syntactic structure when lexical biases are controlled and their perseveratory tendencies are harnessed. In Experiment 2 we return to the question of how adult and child listeners combine lexical and prosodic information in online comprehension.

## **Experiment 1**

The initial goal of this experiment was quite modest. We planned on giving prosody one last chance by testing children with a more sensitive measure (eye-movements), in a design that would allow us to distinguish between a failure to use prosodic cues and strong bias to perseverate across trials and with materials that contained no strong lexical constraints that might compete with prosodic constraints. The critical sentences contained structurally ambiguous with-PPs (“You can feel the frog with the feather”) which could be interpreted as either VP-attached instrument phrases or NP-attached modifiers.

We recorded two versions of each sentence, one with Instrument Prosody (an intonational phrase break after noun) and one with Modifier Prosody (an intonational phrase break after verb). Participants received one block of trials in each prosody condition, with the order of the blocks counterbalanced across participants. Thus prosody was manipulated between subjects in the first block, but within subjects across the two blocks. This design was used so we could eliminate possible interference effects in the first block and investigate them in the second block.

Because the sentences used in this study are never definitively disambiguated, we expect continuity between the listeners' online attachment preferences and their ultimate interpretations. If listeners can use prosody, then the placement of the intonational break should affect whether they interpret the with-phrase as an instrument or a modifier and this preference should be reflected in both their eye movements and their actions.

## Methods

### Participants

Twenty-four English-speaking children participated in the study. The children were divided into two age groups. The preschool group was approximately the same age as the five-year-olds in Snedeker and Trueswell (4;2 to 5;8,  $M = 4;10$ ), while the kindergarten group was about a year older (5;10 to 6;7,  $M = 6;2$ ). Parents were contacted from schools and daycares in the

Cambridge area and from a database of children who had participated in research at the Laboratory for Developmental Studies. All the children who began the experiment completed it and were included in the analyses. Half were male.

### Procedure

Children were tested individually in a quiet room in our lab or their school. They were told that they were going to play a game about following instructions. During the experiment the child was seated in front of an inclined podium. At the center of the podium was a hole for a camera which was focused on the participant's face. In each quadrant of the podium was a shelf where one of the props could be placed. At the beginning of each trial, the experimenter laid out the props and introduced each one using indefinite noun phrases. Any object held by a toy animal was introduced separately rather than as part of a complex NP to ensure that we did not prime the modifier analysis of the target sentences. For instance, the objects shown in Figure 1 would have been introduced by saying: "This bag contains a candle, a feather, a leopard, another candle [referring to the miniature one], a frog and another feather [the miniature one]." This procedure ensured that participant knew the labels for toys and that subsequent reference to the objects using definite noun phrases (e.g., "the frog") was felicitous.

After each object had been labeled twice, the experimenter played prerecorded sound files from a computer connected to external speakers. The trial began with an instruction to look at a fixation point at the center of the display. This was followed by two commands. The child heard the first command, performed that action, and then heard the second (an unambiguous filler). A camera placed behind the child, recorded her actions and the locations of the props, while the camera under the podium recorded her gaze direction. The experimenter moved out of the child's view before the first sentence began and remained there until the action was completed.

If the child refused to respond, the sound file was played again but the eye movements were taken from the initial presentation of the sentence. Children were praised for all responses.

### Stimuli

On the critical trials, the first command contained an ambiguous prepositional phrase (8).

(8) You can feel the frog with the feather.

All the ambiguous prepositional phrases were headed by *with* and could be interpreted as either a modifier of the noun or an instrument (and hence an adjunct of the verb, or perhaps an argument, see Koenig, Mauner & Bienvenue, 2003 for discussion). To increase the probability that prosody would play a decisive role in the interpretation of these utterances, we used sentences in which the *with*-phrase was equally apt as a modifier or an instrument. In S&T 2004, prosodically neutral versions of these sentences, presented with the same referential contexts, resulted in a mix of instrument and modifier responses (37% and 33% instrument responses for five-year olds and adults respectively). These eight sentences contained verbs that had been found to have no strong bias in a sentence completion study and had prepositional objects (“feather”) which had been rated as moderately plausible instruments for the action in question (see S&T, 2004, Appendices B & C). In this earlier study the root sentences were produced as direct commands rather than indirect commands. In the present study we added the carrier phrase for the purpose of making the disambiguating prosodic break in the Modifier condition more natural, by creating more balanced intonational phrases (contrast “Feel.....the frog with the feather” with “You can feel...the frog with the feather”).

As Figure 1 illustrates, the set of toys that accompanied the critical sentences always contained the following objects: 1) a Target Instrument, a full-scale object that could be used to carry out the action (for Figure 1 the large feather); 2) a Target Animal, a stuffed animal carrying a small replica of the Target Instrument (the frog holding a little feather); 3) a Distractor

Instrument; a second full-scale object (the candle); and 4) a Distractor Animal, a stuffed animal of a different kind carrying a replica of the Distractor Instrument (the leopard carrying a candle). Contexts with just one potential referent for the direct-object noun (one frog) were used because they allow us to directly compare the performance of children and adults (see Experiment 2). In prior studies manipulations of referential context have not had reliable effects on attachment preferences in five-year olds: in one-referent contexts children have attachment preferences that are similar to adults, but in two-referent contexts children produce fewer modifier responses (S&T, 2004).

-----  
Insert Figure 1 about here.  
-----

Two versions of each sentence were digitally recorded by a female actor, one with Instrument Prosody and one with Modifier Prosody. The utterance with Instrument Prosody had an intonational phrase break after the direct-object noun. The utterance with Modifier Prosody had an intonational phrase break after the verb. This prosody manipulation was modeled on the utterances produced by the mothers in the referential communication task (Appendix A) and by college-aged adults in a parallel experiment (Snedeker & Trueswell, 2003) and is consistent with other production studies using prepositional phrase attachment ambiguities (Schafer, et al., 2005) as well as theoretical descriptions of the relation between prosody and syntax (Watson & Gibson, 2004). In the instrument condition, the presence of a prosodic break before the with-phrase suggests that there is a major syntactic break between the noun and the prepositional phrase, and increases the likelihood that the with-phrase is a VP-attached instrument phrase. In contrast, in the modifier condition the presence of a prosodic break before the noun phrase suggests that there is a heavy constituent after the verb (Ferreira, 1991; Watson & Gibson, 2004), while the

absence of a break between the noun and the preposition provides evidence that the prepositional phrase is a part of this constituent, and hence a modifier of the noun. Breaks in these locations have been shown to influence prepositional phrase attachment in adults in both online and offline tasks (Pynte & Prieur, 1996; Snedeker & Trueswell, 2003; Schafer, et al. 2005).

The digital waveforms were examined to verify the phrase break and ensure that there were no other detectable pauses in the utterance. The length of each word was measured and paired t-tests were conducted to verify the differences between the two types of utterances (see Figure 2 and Table 1). Instrument utterances had shorter verbs, shorter post-verbal pauses, longer direct-object nouns, longer post-nominal pauses, and longer prepositions.

-----  
Insert Figure 2 and Table 1 about here.  
-----

Prosody was manipulated within participants but was blocked. This allowed us to explore whether response perseveration might explain prior failures to find effects of prosody on children's syntactic parsing. Two counterbalanced presentation lists were constructed. The first half of one list contained sentences with Instrument Prosody while the first half of the other list contained sentences with Modifier Prosody. The critical trials were interspersed with 10 filler trials. Both filler and target trials consisted of two commands and the second command was always an unambiguous filler sentence. Thus, each child heard 28 unambiguous sentences (the first instruction of the 10 filler trials and the second instruction of all 18 trials) and 8 ambiguous ones. Each list was presented in two orders (forward and reverse). The filler sentences contained a variety of constructions but the same fillers were used in all lists and all conditions. They were selected with the goal of not biasing the participants' response on the target trials. For example,

half the filler sentences requested actions involving one object (like the modifier reading), while half requested actions involving two objects (like the instrument reading)

### Coding

Trained coders watched the videotape of the participant's actions and coded them into four separate categories: (1) Instrument Responses: participant used the Target Instrument to perform action on Target Animal; (2) Mini-Instrument Responses: participant used miniature object attached to the Target Animal to perform action on Target Animal; (3) Modifier Responses: participant performed the action on Target Animal without using the target or mini-instrument; (4) Other: participant failed to perform the target action or performed it on the wrong entity.

Because Mini-Instrument and Modifier Responses should both lead to exclusive fixation on the Target Animal, these responses would weaken any effect of attachment on eye movements. To minimize this problem we explicitly discouraged participants from manipulating the miniature objects during the demonstration trials and filler trials (no feedback was given after target sentences). As a result, these responses were infrequent in this experiment (4.2% of target trials). Instrument and Mini-Instrument Responses were combined for all analyses.

Eye movements were coded from the videotape of the participant's face, using frame-by-frame viewing on a digital VCR. The coder noted the onset of the sentence and the onset of each change in gaze, and the direction of the subsequent fixation. The direction of a fixation was coded as being in one of the quadrants, at center, or away from the display. If the subject's eyes were closed or not visible, the frame was coded as missing and the data were excluded from the analysis (only 1.9% of the coded frames were missing). Twenty-five percent of the trials were checked by second coder, who was given the list of onset times for the eye movements. The two coders agreed on the direction of fixation for 95.7% of the coded frames. Disagreements were resolved by a third coder. One test trial was excluded from further analysis due to experimental



errors. With displays of this kind, this method of collecting and coding eye-movements produces data that is comparable to that produced by a head-mounted eye-tracker. For example, S&T 2004 simultaneously collected data using both methods and found that the coded location was identical for 93% of the video frames.

### Results and Discussion

The results are divided into four sections below. First, we present the children's actions in response to the target instructions, analyzing whether an instrument was used to carry out an action. This measure reflects their final interpretation of the ambiguous phrase. Second, we present data on the proportion of trials that included looks to the target instrument. This provides a coarse-grained measure of eye movements that could presumably reflect both early and late interpretive processes. In the third section we analyze how the children's fixations change over time to explore how prosodic information is used online. Finally, we explore alternate explanations for the fixation patterns.

For all measures, we initially conducted an analysis of variance (ANOVA) on the participant means containing three between-participant factors (Age, List, and Order) and one within-participant factor (Prosody). Equivalent ANOVAs were conducted on item means containing one between-item factors (Item Group) and three within-item factors (Prosody, Age and Order). Because our manipulation of prosody was blocked, we were also able to examine the effects of prosody in a between subjects design by limiting our analysis to the initial block of trials. In these ANOVAs, Age and Prosody were between participant and within item factors.

### Actions

Figure 3 plots the proportion of trials in which the participants performed instrument actions, thus indicating that they had interpreted the ambiguous prepositional phrase as a VP-attached argument or adjunct. Table 2 lists the results of the ANOVA for the critical variables. Prosody

had a moderate but reliable effect on interpretation; participants performed instrument actions more often when Instrument Prosody was used. The performance of the kindergarteners and preschoolers was similar, resulting in no effect of Age or interaction between Age and Prosody. However there were interactions between Prosody and Order ( $F1(1,16) = 6.42, p < .05$ ;  $F2(1,6) = 6.23, p < .05$ ;  $minF'(1,17) = 3.16, p = .09$ ), and Prosody, Order and List ( $F1(1,16) = 9.80, p < .01$ ), suggesting that the pattern of performance changed over the course of the experiment.

-----  
 Insert Figure 3 and Table 2  
 -----

As Figure 3 illustrates, there was a strong effect of Prosody in the first block of trials but no effect on the second block. A comparison of the two blocks suggests an intriguing asymmetry between the modifier and Instrument Prosody conditions. Participants who received Modifier Prosody in the first half of the experiment switched to instrument responses when the prosody of the utterances changed (compare the inside bars). But those who started out with Instrument Prosody perseverated after the switch to Modifier Prosody, continuing to produce instrument actions (compare the outside bars). This resulted in a reliable effect of Block for the modifier utterances ( $F1(1,20) = 10.13, p < .005$ ;  $F2(1,7) = 23.25, p < .005$ ;  $minF'(1,27) = 7.06, p < .05^*$ ) but not the instrument utterances ( $F's < 1$ , all  $p's > .5$ ). These findings suggest that the prior failures to find effects of prosody on parsing may be attributable to perseveration across trials.

#### Coarse Grained Analysis of Fixations

For each trial we determined whether the participant looked at the Target Instrument any time between the onset of the prepositional object and the beginning of their action (or 1.5 seconds after the prepositional object onset, whichever came first). Figure 4 plots the proportion of trials with instrument fixations in each of the conditions, while Table 3 lists the results of the

ANOVAs. Participants tended to look at the Target Instrument when they were going to use it to perform the action but seldom fixated on it otherwise. Thus the results for the fixation analysis closely echo those of the action analysis.

In the analysis of both blocks, the performance of the kindergarteners and preschoolers was similar, resulting in no effect of Age or interaction between Age and Prosody. Prosody had a modest effect on Target Instrument fixations. But once again there were interactions between Prosody and Order ( $F(1,16) = 18.01, p < .001$ ;  $F(1,6) = 6.59, p < .05$ ;  $minF'(1,11) = 4.82, p = .05$ ), and Prosody, Order and List ( $F(1,16) = 10.01, p < .01$ ), suggesting that the effect of Prosody changed over the course of the experiment.

-----  
Insert Figure 4 and Table 3  
-----

As Figure 4 illustrates, there was a strong effect of Prosody in the first block of trials but no effect on the second block. As in the action analysis, this reflected a difference between the Modifier and Instrument conditions. Participants who received Instrument Prosody in the first half of the experiment, persisted in looking at the Target Instrument after the switch to Modifier Prosody in the second block. This resulted in a reliable effect of Block for the modifier utterances ( $F(1,22) = 5.51, p < .05$ ;  $F(1,7) = 19.70, p < .005$ ;  $minF'(1,29) = 4.30, p < .05$ ) but not the instrument utterances ( $F's < 1, all p's > .5$ ).

The perseveration of instrument responses and looks indicates that some representation or process is being primed by the child's experiences earlier in the study. But we cannot determine the level at which this priming occurs: the effect could be mediated by syntactic priming of the VP-attached prepositional phrase, priming of a semantic category like instrument, or priming of an action plan that incorporates both an animal and an object (see Thothathiri & Snedeker, in

press for evidence of structural priming in preschoolers' spoken language comprehension). The lack of perseveration when participants switch from modifier to instrument utterances suggests either that the complementary category (complex noun-phrase, modifier, or action on a single object) is less readily primed or that the prosodic cues for VP-attachment are more potent than cues for NP-attachment and thus more apt to override the effects of perseveration. This apparent asymmetry is explored further in Experiment 2. However, since our primary interest is in children's ability to use prosodic cues in online comprehension, and not in the nature of this perseveration, subsequent figures and analyses will focus on the data from the first block of trials.

The dashed line in Figure 4 indicates the proportion of trials during which participants looked to the Distractor Instrument. This object was not mentioned in the sentence and thus provides a rough baseline for the Target Instrument looks. During the first block of trials, participants who received Modifier Prosody were no more likely to look at the Target Instrument than they were to look at the Distractor Instrument ( $F's < 1$ , all  $p's > .5$ ). Thus this analysis provides no evidence that participants in the modifier condition are initially considering the VP attachment of the ambiguous prepositional phrase. In contrast those who received Instrument Prosody were far more likely to look at the Target Instrument ( $F1(1,11) = 21.62, p < .001$ ;  $F2(1,7) = 12.78, p < .01$ ;  $minF'(1,14) = 8.03, p < .05$ ).

### Temporal Analysis of Eye-Movements

To explore the relation between the unfolding utterances and the participant's evolving interpretation, we analyzed how the distribution of eye movements changes over time (see Figures 5a & 5b). In each figure, time is displayed along the x-axis in increments of  $1/30^{\text{th}}$  of a second (equivalent to a single video frame). Time is measured relative to the onset of the object of the preposition (e.g., "feather" in "You can feel the frog with the feather"). The lines

represent the proportion of fixations to each of the four types of objects that the subject could look at: the Target Animal, the Distractor Animal, the Target Instrument, and the Distractor Instrument. We expected that prosody would affect the proportion of fixations to the Target Instrument, since this object is the referent of the prepositional object if and only if the phrase is VP-attached. In contrast we expected to see little or no difference in looking time to the Target Animal, since this object is the referent of the direct-object noun phrase regardless of how the ambiguity is resolved.

There are two obvious differences between the instrument and modifier conditions (Figures 5a and 5b). First, participants in the modifier condition shift their gaze from the fixation point to the Target Animal in the 300 ms preceding the PP-Object Onset. In contrast, those in the instrument condition are already looking at the Target Animal prior to this time. This reflects necessary differences in the timing of the two types of utterances. Instrument utterances contain elongated direct-object nouns followed by substantial pauses, giving participants plenty of time to identify this first noun and shift their gaze before the onset of the second noun ( $M = 998$  ms, from the onset of the direct-object noun to the PP-Object Onset). In the modifier utterances, however, there is relatively little time between the two nouns ( $M = 477$  ms from onset to onset).

-----  
Figure 5 and Table 4  
-----

Second and more critically, there is also an increase in looks to the Target Instrument in the instrument condition beginning roughly 300 ms after the PP-Object Onset and plateauing about 900 ms later. In modifier condition, in contrast there are few looks to the Target Instrument during this time period.

To determine when the prosody manipulation began to influence the children's eye movements, we analyzed Target Instrument fixations in two 500 ms time windows following the PP-Object Onset. Previous research demonstrates that lexical information begins to influence eye movements about 200ms after word onset (Allopenna, Magnuson & Tanenhaus, 1998), so we began our Early PP-Object window at 200ms after the onset of the prepositional object. The Late PP-Object window began 700ms after the prepositional object onset. The results of these analyses are presented in Table 4. The effect of prosody was not reliable for the Early PP-Object window, despite a trend towards greater proportion of Target Instrument looks in the instrument condition. In the Late PP-Object window, however, there was a significant difference between the two types of utterances. Participants looked at the Target Instrument more when the sentence occurred with Instrument Prosody than when it occurred with Modifier Prosody. There was no reliable effect of age group in either time window and no reliable interaction between age and prosody.

#### Are These Effects of Prosody or Side-Effects of Time?

In this experiment, we manipulated the prosody of the utterance by having the speaker shift the placement of an intonational phrase boundary. This resulted in systematic changes in the lengths of words and placement of pauses (see Table 1 and Figure 2). Consequently, in our study, as in previous experiments, the effects of prosody could be attributable to differences in the timing of words. There are two ways in which this might occur. First, differences in time could be part of the mechanism by which prosodic variation influences parsing. This hypothesis is consistent with our data and will be explored further in the General Discussion. The second possibility is more worrisome. Perhaps the effects of our prosodic manipulation are attributable to effects of time on our dependant measures that are independent of parsing or prosodic structure.

This alternative provides a prima facie explanation for the effects of prosody on temporal analysis of Target Instrument looks. In this study, as in prior studies, the children tended to look at each object as it was mentioned, beginning with the direct-object noun (*frog*). In the Modifier Prosody condition, the prepositional object rapidly follows the direct-object noun, while in the Instrument condition there is a substantial pause between these words. Thus even if participants had similar parsing preferences for the two types of sentences, we might expect decreased looking to the Target Instrument in the Modifier condition immediately after the PP-Object Onset, simply because participants might still be planning and executing their initial looks to the referent of the direct-object noun (the Target Animal). While this account cannot explain the effects of prosody on the participants' actions, it does raise questions about the interpretation of our online measures.

---

Table 5

---

To explore this possibility we conducted an additional analysis on the trials in which subjects had already shifted their gaze to the Target Animal before encountering the ambiguous prepositional object. If the effects of prosody on instrument fixations in the previous analysis were simply artifacts of delays in looks to the Target Animal then we should see the effects disappear or diminish in this analysis. If however, the differences in instrument fixations reflect the influence of prosody on the interpretation of the ambiguous prepositional phrase, then these effects should persist in this subset of trials. We selected all Block 1 trials on which the participant was gazing at the Target Animal 100ms before the onset of the prepositional object (or 300ms before or first analysis window and the time at which we would expect to see shifts related to the prepositional object). We calculated the proportion of target instrument fixations

during the Early and Late PP-Object Time Windows and conducted one-way ANOVAs with Prosody as a between subjects and within item variable. These analyses, presented in Table 5, closely parallel the analyses for the full data set (Table 4). Once again, there was no reliable effect of prosody during the Early PP-Object window. During the Late PP-Object window participants in the instrument condition spend more than twice as much time looking at the Target Instrument as those in the modifier condition. However, the effect of prosody is only marginal in both the subject and item analyses. This is a divergence from the primary analysis where this effect reached the conventional significance level (see Table 4). Since the means for each condition are quite similar in the two analyses, this difference may reflect an increase in variability in the data due to the reduction in the number of trials used in this comparison (59 of the original 94).

### Summary of Experiment 1

Experiment 1 demonstrates that prosody has an effect on children's interpretation of an ambiguous prepositional-phrase attachment. This effect was apparent both in their actions and proportion of trials in which they looked at the Target Instrument. When children heard Instrument Prosody there was a reliable increase in eye movements to the Target Instrument about 700ms after the onset of the prepositional object. While this effect could be an artifact of differences in timing and their influence on Target Animal looking time, additional analyses suggest that the difference persists even when participants have succeeded in looking at the Target Animal long before the onset of the critical prepositional object.

However, this study also suggests that the effects of prosody are somewhat fragile; they disappear in the second block of trials, swamped by the perseveration of instrument actions. These results raise questions about the relative contribution of prosodic and lexical cues to online sentence processing in children and adults. In a parallel study, S&T 2004 found that lexical



biases had a robust influence on the interpretation of prosodically-neutral prepositional-phrase attachments. In the present study we found that prosody shaped the interpretation of lexically neutral sentences. In Experiment 2 we explore how these cues interact by simultaneously manipulating intonation and lexical biases and children and adults.

This allows us to do three things. First, we can directly compare the relative strength of prosodic and lexical cues to prepositional phrase attachment ambiguity when both information sources are available. Second, manipulating both cues will allow us to explore the time course of prosodic and lexical influences. Third, by examining the full paradigm in both children and adults we can explore whether the relative influence of prosodic and lexical information changes across a period of development in which executive functions blossom, processing speed increases across a variety of tasks, and reading becomes a primary source of linguistic input.

## **Experiment 2**

Experiment 2 examines the effects of prosody and lexical bias in children's and adult's online interpretation of ambiguous prepositional phrase attachments. While there is ample evidence that lexical biases and prosodic phrasing rapidly influence adult parsing, there is little work exploring how these information sources interact over time. Verb Bias was manipulated between subjects. Prosody was blocked so that it could be analyzed as a between subjects variable in the first block but as a within subjects variable across the two blocks.

### Methods

#### Participants

Thirty-six adults and seventy-two children participated in this study. Twelve of the adults and forty-two of the children were male. All were native speakers of American English. The children were divided into two age groups: preschoolers (4;1 to 5;5,  $M = 4;11$ ) and kindergarteners (5;6 to 6;7,  $M = 6;0$ ). Adult participants were recruited from the Harvard community and received

partial course credit or a small payment for their participation. The parents of the child participants were contacted from Cambridge area schools and daycares and from a database of families who had participated in research at the Laboratory for Developmental Studies. Four adults began the study but were not included in the analysis because of experimenter errors (2), because they were not native speakers of American English (1), or because they had a striambus (a wandering eye) that was severe enough to prevent accurate coding of their eye-movements (1). Ten children began the study but were not included in the analysis because of experimenter errors (4) or because the child failed to complete the study (4), produced incorrect actions for control items (1) or was not a native English speaker (1).

### Procedure

For the children the procedure was identical to Experiment 1. For the adults the following changes were made. First, to ease participants' qualms about the childishness of the task, the adults were told that their performance would serve as a point of comparison for a study of how children follow instructions. Second, each prop was labeled only one time. Finally, the adults were not praised for their responses, unless they seemed particularly insecure.

### Stimuli

The critical target sentences were based on those used in S&T 2004. The verbs had been selected in an earlier sentence completion study in which adult participants were asked to complete sentence fragments that ended with the ambiguously attached preposition. The verbs in the Modifier Bias condition were ones for which modifier completions were at least three times as frequent as instrument completions. For the Instrument Bias verbs the opposite rule applied. Equi Bias verbs were those that fell somewhere in between.

- (9) a. You can choose the cow with the stick. (Modifier Bias)
- b. You can feel the frog with the feather (Equi Bias)

## c. You can tickle the pig with the fan. (Instrument Bias)

As in Experiment 1, all of the sentences contained prepositional objects which had been rated as moderately plausible instruments for the action in question. The sentences were presented with prop sets used in S&T 2004, consisting of a Target Animal, a Target Instrument, a Distractor Animal, and a Distractor Instrument. The two prosodic variants of the utterances were prepared in the same way as Experiment 1, using the same actor. The acoustic analyses of the stimuli paralleled those from Experiment 1 (see Table 6). Instrument utterances had shorter verbs, shorter post-verbal pauses, longer direct-object nouns, longer post-nominal pauses, longer *with*'s, and longer prepositional objects.

To determine whether the prosody manipulation was equivalent across the three bias classes, we conducted ANOVA's for each measurement with verb class as a between items variable and prosody as a within item variable. There was a main effect of verb class on the length of the verb, which was unsurprising since the verbs that were assigned to the three classes were not matched for their number of syllables or phonemes ( $F(2,21) = 8.1, p < .01$ ). There was also an interaction between prosody and verb class, with the longer Equi and Modifier Bias verbs showing a greater difference in length ( $F(2,21) = 3.6, p < .05$ ). Since it seems plausible that the effects of prosodic lengthening should be proportional to word length, rather than additive, we conducted a one-way ANOVA with the ratio of the verb length in the Instrument Prosody condition to verb length in the modifier condition as the dependant variable. There was no effect of bias class in this analysis ( $F(2,21) < 1, p > .4$ ), suggesting that the proportional effects of prosody were similar across the three bias classes. There were no other effects of verb-bias or interactions between bias and prosody (all  $F$ 's  $< 2$ , all  $p$ 's  $> .1$ ).

-----  
Table 6

---

To further explore the nature of our prosodic manipulation, a highly trained coder analyzed the utterances using the ToBI labeling system (Beckman & Hirschberg, 1994). The utterances from S&T 2004, which were intended to be prosodically neutral, were also analyzed to serve as a point of comparison. For each utterance we examined the break index after the direct-object noun and after the verb (where 4 = IP break, 3 = ip break, 2 = ambiguous, 1 = phonological word break, 0 = no break) as well as the accents on the verb, noun, preposition and prepositional object (Table 7).

---

Table 7

---

The prosodic transcription verified that the neutral utterances from S&T 2004 consisted of a single intonational phrase (IP). In contrast all the utterances with Modifier Prosody consisted of two IP's with the break occurring after the verb, while the instrument utterances consisted of two IP's with the break occurring after the direct-object noun. In most cases each intonational phrase consisted of just one intermediate phrase (ip). However, there were four neutral utterances with an ip break after the noun, one instrument utterance with an ip break after the verb, and one neutral utterance with an ip break after the verb.

Several researchers have suggested that the relation between syntax and prosody is best captured by examining the entire prosodic structure of the utterance, rather than individual prosodic boundaries (Prince & Prieur, 1996; Carlson et al., 2001; Schafer et al., 2005). In the case of attachment ambiguities, the critical comparison appears to be between the prosodic boundary immediately before the ambiguous phrase and the boundary preceding the phrase that could serve as the lower attachment site (Carlson, et al., 2001). Thus for our stimuli we would

expect instrument utterances to have a larger break before the prepositional phrase (coded here as the noun break index), whereas modifier utterances should have a larger break before the direct-object noun phrase (coded here as the verb break index). Following Schafer et al., 2005, we compared the break indices at these locations and classified the utterances as having instrument prosody (noun break > verb break), neutral prosody (noun break = verb break) or modifier prosody (noun break < verb break). This analysis confirmed that all utterances in the current experiment had the intended prosodic form and that 80% of the utterances from S&T 2004 were in fact prosodically neutral.

All the target sentences had pitch accents on the verb, direct-object noun, and prepositional object. The type of pitch accent on the verb and noun varied across the conditions. All the utterances with Modifier Prosody had a salient L+H\* accent on the verb. This gave the verb greater weight allowing the speaker to make the early IP break in these utterances in a more natural fashion. While the mapping between discourse functions and accent types is controversial, many claim that L+H\* accents signal new information (Pierrehumbert & Hirschberg, 1990; Baumann, 2005) or discourse themes (Steedman, 2000). None of the Instrument or Neutral utterances contained an L+H\* accent. Instead they often had L\*+H or H\* accents, both of which have been argued to be functionally similar to the L+H\* accent but less marked or salient (Steedman, 2000; Baumann, 2005). Thus while the verbal pitch accents in all the utterances suggested, correctly, that the verb was new information, this was emphasized more strongly and consistently in the Modifier utterances.

The pattern of accents on the direct-object nouns was less systematic. The proportion of L+H\* accents did not differ reliably across the three utterance types. However there was a difference in the relative frequency of H+!H\* and H\* accents. The former were common in Instrument and Neutral utterances but absent from the Modifier utterances. The H+!H\* accent

may mark material that is moderately accessible in the discourse model (Baumann, 2005). Thus the direct-object noun is treated as accessible in a minority of the Neutral and Instrument utterances but is treated as new in all the Modifier utterances.

While none of the Neutral or Modifier utterances had a pitch accent on the preposition, 79% of the utterances with Instrument prosody did. A similar pattern was observed by Snedeker & Trueswell (2003) who found that untrained speakers, who were aware of the ambiguity, often placed a pitch accent on the preposition when the instrument reading was intended. There were no reliable differences across the conditions in the distribution of accents on the prepositional object and none of the prosodic variables differed reliably or systematically with verb bias.

Verb Bias was manipulated across participants. Prosody was blocked and fully crossed with Verb Bias. Lists and orders were constructed in the same manner as Experiment 1. The adult participants were given 24 filler trials, while children were given 10. Since the second command for each set of toys was also an unambiguous filler sentence, this meant that each adult heard 56 unambiguous sentences and 8 target sentences and each child heard 28 unambiguous sentences and 8 ambiguous targets.

### Coding

Data was coded in the manner described above. A second person independently coded the eye-movements for 15 of the participants. The two coders agreed on the direction of gaze for 96.0% of the coded frames and disagreements were resolved by a third coder. Both the adults and the children produced few Mini-Instrument responses (1.0% and 2.6% respectively) and these were grouped with the Instrument responses for analysis. Only four trials were lost due to experimenter error (all from children) and just 2.3% of the frames were coded as missing. On four additional trials children refused to carry out the action, though their eye-movements were recorded and analyzed.

## Results

### Actions

*Adults:* The adult's actions were strongly influenced by both Prosody and Verb Bias (see Table 8). Participants who heard utterances with Instrument Prosody performed actions with the Target Instrument on 62% of the trials, while those hearing Modifier Prosody did so only 27% of the time. Similarly, those receiving Instrument Bias verbs primarily performed instrument actions ( $M = 75\%$ ) while those who heard Equi-Bias and Modifier Bias verbs produced fewer ( $M = 37\%$  and  $M = 21\%$  respectively). Although the effect of Prosody appeared to be greatest for Equi-Bias verbs and smallest for Modifier Bias verbs, the interaction between Prosody and Bias failed to reach significance in the subject analysis.

In contrast to the children (see below), our adult participants showed the same pattern of performance in the second presentation block as they had in the first (compare Figures 6a and 6b). In both blocks there were robust effects of Prosody and Verb Bias (Table 8) and there were no effects of presentation block on the adults' responses, and thus no clear evidence of perseveration across trials ( $F_1(1,30) = 1.16, p > .25$ ;  $F_2(1,21) = 3.94, p = .06$ ;  $\min F'(1,44) < 1, p > .3$  for Modifier Prosody and  $F's < 1, p > .5$  for Instrument Prosody).

-----  
 Figure 6 & Table 8  
 -----

*Children:* Like the adults, the children's actions were strongly influenced by both Prosody and Verb Bias (see Table 8). These effects were clearest in the first block. In this block, participants who heard utterances with Instrument Prosody performed instrument actions on 67% of the trials, while those hearing Modifier Prosody did so only 32% of the time (Figure 6c). They were also more likely to produce instrument actions in response to Instrument Bias verbs

( $M = 85\%$ ) than in response to Modifier Bias verbs ( $M = 20\%$ ), while Equi Bias verbs were intermediate ( $M = 44\%$ ).

This pattern changed substantially in the second presentation block (see Figure 6d). While the effect of Verb Bias remained strong, the effect of Prosody disappeared completely. Since for each child Verb Bias was held constant across the presentation blocks while Prosody was varied, this pattern could be caused by perseveration across trials. To explore this further we conducted separate ANOVA's on the Instrument and Modifier Prosody conditions, with Age, Verb Bias and Block as our independent variables. In the Instrument Prosody condition there was an effect of Block and an interaction between Age and Block both of which were marginal in the subjects analysis ( $F(1,60) = 3.34, p = .07$ ;  $F(1,21) = 6.04, p = .02$ ;  $\text{min}F'(1,77) = 2.15, p = .15$  and  $F(1,60) = 3.84, p = .055$ ;  $F(1,21) = 12.21, p < .005$ ;  $\text{min}F'(1,81) = 2.92, p = .09$  respectively). The children, particularly the Preschoolers, were less likely to perform instrument actions in the second block, suggesting that the Modifier Prosody trials in the first block influenced their subsequent responses. In the Modifier Prosody condition there was an effect of Block and no interaction with Age ( $F(1,60) = 9.57, p < .005$ ;  $F(1,21) = 27.88, p < .001$ ;  $\text{min}F'(1,81) = 7.12, p < .01$  and  $F's < 1, p's > .4$ , respectively). Both groups of children were more likely to perform an instrument action on the second block, presumably because they had received Instrument Prosody in the first block.

Thus we replicate the findings of Experiment 1: while children can use prosody to resolve syntactic ambiguity, this ability is easily overridden by their tendency to perseverate across trials. However, the pattern of perseveration in Experiment 2 did not show the same striking asymmetry that was observed in Experiment 1. In Experiment 1, perseveration only affected utterances with Modifier Prosody, while in Experiment 2 it occurred for both prosody types. In Experiment 1, participants who received Instrument Prosody in the first block, showed no



change in their response pattern when they shifted to Modifier Prosody. In Experiment 2, both groups of participants showed a reliable difference between the two blocks ( $t(35) = 3.44, p < .005$  for instrument-to-modifier and  $t(35) = 3.68, p < .001$  for modifier-to-instrument). Thus despite perseveration across blocks we find some sensitivity to prosody in this within subject analysis.

*Comparison:* The action data from the children and the adults was pooled and we conducted separate ANOVA's on each block of trials. In the first block there were no interactions between Age (child vs. adult) and the critical variables, indicating that both groups were equally sensitive to prosodic and lexical cues in their ultimate interpretation (all  $F$ 's  $< 1$ , all  $p$ 's  $> .3$ ). In addition the Bias-Prosody interaction, which was marginal in the adult data and absent in the children's, reached statistical significance ( $F(2,84) = 3.16, p < .048$ ;  $F(2,18) = 5.62, p < .013$ ;  $\text{min}F'(2,83) = 2.02, p = .14$ ). In the second block of trials there was an interaction between Age and Prosody ( $F(1,84) = 5.81, p < .05$ ;  $F(1,18) = 31.15, p = .001$ ;  $\text{min}F'(1,102) = 4.90, p < .05$ ), reflecting the fact that children, but not adults tended to perseverate across trials.

### Coarse Grained Analysis of Fixations

Figure 7 plots the proportion of trials with instrument fixations in each of the conditions, while Table 9 lists the results of the ANOVAs. Participants primarily looked at the Target Instrument when they were going to use it to perform the action, thus the results for the fixation analysis mirror those of the action analysis.

-----  
 Figure 7 & Table 9  
 -----

*Adults:* The adult participants showed the same pattern of performance in the second presentation block as they had in the first (compare Figures 7a and 7b). In both blocks Prosody

and Verb Bias had strong and reliable effects on the proportion of trials with Target Instrument looks (see Table 9). The findings diverged from the action analyses in one respect: there was a reliable interaction between Bias and Prosody reflecting a larger effect of Prosody for the Equi-Bias verbs utterances. The dashed lines in Figures 7a and 7b indicate the proportion of trials with looks to the Distractor Instrument. When Modifier or Equi-Biased utterances were produced with Modifier Prosody the Target Instrument looks did not exceed looks to the Distractor Instrument, suggesting that participants were not considering the VP-attachment (all  $F's < 1$ , all  $p's > .5$ ). In all other conditions, looks to the Target Instrument exceeded looks to the Distractor Instrument (all  $F's > 8$ , all  $p's < .05$ ).

*Children:* Here again the results of the fixation analysis closely match those of the action analysis. Like the adults the children showed robust effects of Prosody and Verb Bias in the first block. In Block 2 the effect of Verb Bias remains while the effect of Prosody disappears. Since prosody is manipulated within subjects and Verb Bias is manipulated between subjects this could reflect a greater sensitivity to priming in the children (or a decreased sensitivity to *changes* in prosody). The dashed lines in Figures 7c and 7d indicate the proportion of trials during which participants looked to the Distractor Instrument. When modifier utterances were produced with Modifier Prosody looks to the Target Instrument did not exceed looks to the Distractor Instrument, suggesting that participants were not considering the VP-attachment ( $F's < 1$ ,  $p's > .4$ ). For the Equi-Bias and Instrument Bias utterances with Modifier Prosody the data pattern was unclear ( $F1(1,23) = 3.78$ ,  $p = .06$ ;  $F2(1,7) = 2.12$ ,  $p = .19$ ; and  $F1(1,23) = 5.29$ ,  $p < .05$ ;  $F2(1,7) = 2.08$ ,  $p = .19$ , respectively), perhaps reflecting the increase in Target Instrument looks in Block 2 when they were preceded by utterances with Instrument Prosody. In all other conditions, looks to the Target Instrument exceeded looks to the Distractor Instrument (all  $F's > 10$ , all  $p's < .05$ ).

In summary, in both children and adults Target Instrument looks in the first block of trials were strongly influenced by both the bias of the verb and the prosody of the utterance. In the second block of trials, the children showed no reliable effect of prosody. Thus to ensure that the data from the adults and children was comparable and to better equate the Prosody and Bias manipulations, the remainder of our analyses were limited to the first presentation block, in which both Prosody and Verb Bias were manipulated between subjects.

### Temporal Analysis of Eye-Movements

Interpretation of the prepositional phrase was explored by examining the proportion of fixations to the Target Instrument. We analyzed the fixation probabilities in three time windows. To explore the earliest effects of Verb Bias and Prosody we added a time window (the "With" Window) that corresponded to the period from the onset of the preposition until the beginning of the prepositional object (-67ms before to 167 ms after PP-Object Onset). As in Experiment 1 we divided the period after the PP-Object Onset into an Early PP-Object window (200ms-667ms after PP-Object Onset) and a Late PP-Object window (700-1167ms). The results of these analyses appear in Table 10.

-----  
 Table 10 and Figure 8  
 -----

*Adults:* Figure 8 highlights the effects of verb bias by comparing Target Instrument fixations for the three bias classes in each of the prosody conditions. In both graphs, looks in the Instrument Bias condition begin increasing shortly before the PP-Object Onset. These looks begin declining about 400ms after PP-Object Onset and then increase rapidly in the later part of the trial. In contrast, participants in the Modifier Conditions rarely look at the Target Instrument. The effect of the prosody manipulation can be seen by comparing the two panels of Figure 8.

For each bias class there is a sharp increase in Target Instrument looks for the Instrument Prosody utterances that begins directly after the PP-Object Onset and levels out about 500ms later.

In the “With” Window, there was a marginal effect of Verb Bias that was reliable only in the items analysis (Table 7). Participants in the Instrument condition looked at the Target Instrument 9% of the time, while those in the Equi-Bias and Modifier-Bias conditions never did so. In the Early PP-Object Window the effect of Verb Bias remained marginal (21% for Instrument-Bias vs. 7% for Modifier-Bias). There was a robust effect of Prosody in this time window; adults hearing Instrument Prosody looked at the Target Instrument more than those hearing Modifier Prosody (24% vs. 4%). In the Late PP-Object Window there were large and reliable effects of both Bias and Prosody. Those who heard Instrument Bias sentences were more likely to look at the Target Instrument (39%) than those hearing Equi and Modifier Biased ones (17% and 15% respectively). Similarly, participants who heard Instrument Prosody had a larger proportion of Target Instrument looking time than those hearing Modifier Prosody (35% vs. 12%).

*Children:* Figure 9 highlights the effect of lexical bias by comparing Target Instrument fixations for the three bias classes in each of the prosody conditions. In both graphs, looks in the Instrument Bias condition begin increasing about 100ms before the PP-Object Onset, appearing to be roughly time locked with the onset of the preposition (“with”). In contrast, participants in the Modifier Bias conditions rarely look at the Target Instrument. Thus children appear to use lexical information as quickly as the adults. This difference between the verb classes grows rapidly and persists throughout trial, in contrast with the biphasic pattern seen in the adults.

The effect of the prosody manipulation can be seen by comparing the two panels of Figure 9. For all three bias classes there is little difference between the two types of prosody until about 400 – 500 ms after the PP-Object Onset when the looks to the Target Instrument begin increasing

in the Instrument Prosody conditions. This effect increases gradually and persists throughout the trial. While the onset of this effect is about 300ms later than it was for the adults, the maximum proportions in each condition (the “peaks”) are similar across the two groups.

-----  
Figure 9  
-----

In the “With” Window (Table 7), there was no effect of Prosody and an effect of Verb Bias that was reliable by subjects but only marginal by items. Participants in the Instrument Bias Condition looked at the Target Instrument more often than those in the Modifier Bias condition (16% vs. 4%) while the Equi-Bias condition was intermediate (9%). In the Early PP-Object Window the effect of Verb Bias was robust (24%, 11%, and 3% for Instrument Bias, Equi-Bias and Modifier Bias sentences respectively) and there was a marginal effect of Prosody in (15% for Instrument Prosody vs. 10% for Modifier Prosody). In the Late PP-Object Window there were large and reliable effects of both Bias and Prosody, echoing the pattern seen in the action analysis. Participants who heard Instrument Bias sentences spent 35% of their time looking at the Target Instrument while those hearing Equi and Modifier Bias utterances were far less likely to do so (21% and 13% respectively). Similarly, participants who heard Instrument Prosody had a larger proportion of Target Instrument looking time than those hearing Modifier Prosody (33% vs. 13%).

*Comparison:* To compare the children and adults, we pooled their data and conducted ANOVA's for the instrument looks in each time window. The results of these analyses are given in Table 11. In the pooled data the effect of Verb Bias emerged in the “With” Time Window and persisted through both of the PP-Object windows. There were no interactions between Verb Bias and Age (adult/child) in any of these analyses. In contrast, the effect of Prosody did not emerge

until the Early PP-Object Time Window. In this time window, the effect was carried largely by the adults resulting in an Age by Prosody interaction. This interaction disappears in the Late PP-Object Onset Window, suggesting that the later portion of the trial the effect of Prosody is similar for the two age groups.

-----  
Table 11  
-----

#### Are These Effects of Prosody or Side-Effects of Time?

While the results of the temporal analyses strongly suggest that both children and adults rapidly use prosodic structure to interpret ambiguous prepositional phrases, these analyses are open to the same critique as those in Experiment 1. In the Instrument Prosody conditions the direct object noun begins almost 1000 ms before the onset of the prepositional object noun. In the Modifier Prosody conditions the direct object only precedes the prepositional object by about 450 ms. Target Instrument looks may be absent in the Modifier Prosody conditions not because prosody has an early effect on interpretation, but because participants in these conditions are still programming and executing eye-movements in response to the direct object noun.

-----  
Table 12  
-----

As in Experiment 1, we explored this possibility by analyzing the subset of Block 1 trials in which subjects had already shifted their gaze to the Target Animal 100ms prior to the onset of the prepositional object noun. We calculated the proportion of target instrument fixations during the Early and Late PP-Object Time Windows and conducted one-way ANOVAs with Prosody as

a between subjects and within item variable. If the effects of prosody on instrument are artifacts of delays in looks to the Target Animal, then they should diminish or disappear in this analysis. If they reflect effects of prosody on the interpretation of the ambiguous prepositional phrase, then we would expect them to persist.

The findings, presented in Table 12, closely parallel the findings for the full data set (Table 10). In the adults there is a reliable effect of Prosody in the Early PP-Object Windows. Those participants in the Instrument Prosody conditions who were looking at the Target Animal 100 ms prior to the prepositional object spent 30% of their time looking at the Target Instrument 300-800 ms later. While those in Modifier conditions did not look at it at all. This effect persisted in the Late PP-Object Window ( $M = 36\%$  and  $M = 10\%$  for Instrument and Modifier Prosody respectively). In the children, there was no effect of Prosody in the Early Time Window. Instead the effect emerged in the Late PP-Object Window as it had in the analysis of the full set of trials ( $M = 32\%$  and  $M = 11\%$  for Instrument and Modifier Prosody respectively). Thus we find that the effects of Prosody in this subset of trials are quite similar to those seen in the full analysis, suggesting that the differences in eye-movements to the Target Instrument are not a side-effect of differences in timing of initial fixations on the Target Animal.

### Summary of Experiment 2

In most respects the children's pattern of performance mirrored that of the adults. In both groups, the participants' ultimate interpretation of the ambiguous phrase (as indicated by their action) was shaped by both prosody and lexical biases. In both groups, the effect of verb bias emerged early in the sentence, coinciding roughly with the onset of the preposition and preceding the appearance of the critical noun. However, the timing of the prosodic effects differed between the two groups. In adults the effect of prosody was apparent shortly after the onset of the prepositional object, becoming reliable in the Early PP-Object Window. In contrast,

this effect was slightly delayed in the children, becoming robust only in the Late PP-Object Window. The effects of prosody persisted when we limited our analysis to those trials in which the participants had fixated on the Target Animal well before any eye-movements could be made in response to the prepositional object.

### General Discussion

These experiments produced three clear findings which we hope will not be lost among the more tantalizing but ambiguous results. First, children as young as four clearly use prosody in their structural interpretation of spoken utterances. Second, we confirm our earlier finding that young children can rapidly employ lexically-specific information in online parsing. Third, both children and adults combine prosodic and lexical information in a roughly additive fashion: using prosodic information, even in lexically-biased utterances, and lexical information, even in prosodically-biased utterances.

Our prosody manipulation had effects on eye-movements to the Target Instrument which emerged approximately 200 ms after the onset of the prepositional object in the adults and about 300 ms later in the children. As we noted earlier the interpretation of these effects is unclear. Below we explore the possibility that these effects are an artifact of timing and find that this hypothesis fails to explain the full data pattern. In the remainder of the discussion we address how our findings bear on the three issues that motivated this study: the interface between prosody and syntax across development, the interplay between lexical and prosodic cues in adult sentence processing, and the architecture underlying children's parsing.

#### **The missing Target Instrument looks: Artifact or evidence for the rapid use of prosody?**

In both experiments we analyzed the proportion of looking time to the Target Instrument immediately after the onset of the prepositional object. We found robust effects of both prosody and verb bias. This measure was intended to capture the participants evolving interpretation of



the utterance. However, as we noted earlier, our prosody manipulation could potentially influence eye-movements without affecting online interpretation. In the Instrument Prosody conditions participants had ample time to look at the referent of the direct-object noun before they encountered the ambiguous prepositional object, allowing them to quickly switch their gaze to the Target Instrument. In contrast participants in the Modifier Prosody conditions encountered the two nouns in quick succession (see Figure 2 and Tables 1 & 6). Perhaps they failed to look at the Target Instrument because they were still programming or executing their initial looks to the Target Animal. If this were true then these eye-movements would tell us little about temporal dynamics of the syntax-prosody interface. Figure 5 suggests that this is a valid concern. When children heard Instrument Prosody their looking time to the Target Animal was at ceiling 300ms before the critical prepositional object onset. In contrast, when they heard Modifier Prosody Target Animal looking time peaks during the critical analysis windows.

Nevertheless, there are five reasons why we believe that initial Target Animal looks cannot fully account for the early effects of prosody on Target Instrument looks:

1. Target Instrument looking time is quite high in Instrument Bias utterances with Modifier Prosody (see panel B of Figures 8 & 10). Thus it is possible for participants who have heard the two nouns in close proximity to squeeze in early eye-movements to the Target Instrument.
2. If the effects of prosody are artifacts of the timing of Target Animal looks, then we would expect that they would be time locked to the onset of the direct-object noun in the Modifier Prosody conditions. The effects should emerge about 200ms before the prepositional object onset (approximately 200ms after the onset of direct-object noun) and they should decrease over time as participants completed their initial looks to the Target Animal. However, in Experiment 2 we found no effects of prosody in the time

window that preceded the prepositional object onset. In the children the effects are also absent in the Early PP-Object time window. We would have to supplement the artifact hypothesis in some way to account for this delay.

3. The developmental differences are not readily explained by this hypothesis. If the difference between the Modifier and Instrument Prosody conditions is caused by delayed eye-movements in response to the direct-object noun, then the effect should be greatest in participants who process language more slowly. Yet we find that the effects of prosody on Target Instrument looks emerge earlier and more robustly in adults.
4. If, in the Modifier Prosody conditions, Target Animal looks were simply squeezing out or delaying Target Instrument looks, then we would expect to see these looks emerge later in the trial. However, we found the participants who received Modifier Prosody and Modifier or Equi-Bias verbs were unlikely to ever look at the Target Instrument (Figures X and 7).
5. Most critically, when we limit our analysis to trials in which participants were already looking at the Target Animal 100ms prior to the onset of the prepositional object, the effects of prosody on Target Instrument looking time are unchanged. These participants appear to have already identified the referent of the direct-object noun. Nevertheless those who hear Modifier Prosody are not shifting their gaze to the Target Instrument, while those who hear Instrument Prosody are.

These arguments strongly suggest that the effect of prosody in the temporal analyses cannot be reduced to the direct effect of time on eye-movements. However we cannot entirely rule out the possibility that direct effects of this kind play some role in the observed pattern of eye-movements. It is critical to note that this alternative only jeopardizes the interpretation of the

effects of prosody in the fine-grained temporal analyses. The critical confound is not present in our manipulation of lexical bias and it cannot explain the effects of prosody on the participant's actions, the preference for the Target Instrument (relative to the Distractor Instrument) in the Instrument Prosody conditions and the lack of any such preference in the Modifier Prosody conditions. These effects indicate that prosody is constraining the interpretation of the ambiguous phrase, and suggest that it is doing so quite rapidly.

### **The development of interface between prosody and syntax**

In the introduction we noted that children's apparent failure to use prosody to resolve syntactic ambiguity was curious in light of their early sensitivity to prosodic manipulations and theoretical claims that prosody serves as a bootstrap into syntax. If prosodic structure is salient to children, why would they fail to notice its relation to syntactic structure? Our findings clear up this mystery by demonstrating that four to six-year-old children can use prosody to resolve ambiguity. We attribute the previous failure to find such effects to the strong tendency of young children to perseverate on a single interpretation. Recently, Mazuka and Tanaka (2006) have confirmed both of these claims: five-year-old Japanese speakers can use intonation to interpret structurally-ambiguous modifiers when the prosodic cues are blocked, but they perseverate failing to change their analysis when the prosody shifts.

Note that our findings do not provide direct support for the prosodic bootstrapping hypothesis. The children in our studies are in a radically different epistemic situation than infant language learners. They have full access to the lexical content of the utterance and substantial knowledge about syntax of their native language. The mere fact that children were able to use these strong prosodic cues to select between two structural analyses has little bearing on whether infants can use naturally occurring prosodic cues to derive these analyses in the first place (see Fernald & McRoberts, 1996 for a related discussion). Knowing that preschoolers use prosody in

parsing merely eliminates the need for prosodic bootstrapping theories to posit U-shaped development of the interface between prosody and syntax.

Our findings raise the question of how children acquire their sensitivity to the mapping between prosody and syntax. We see three broad possibilities.

1. *Children could learn the mappings from prosody to syntax through experience with unambiguous utterances.*

The first challenge for this proposal is to demonstrate that the correspondences between the syntax and prosody are robust enough to be learned by young children. Our work on the production of attachment ambiguities provides reasons for both skepticism and for optimism. When speakers are in an ambiguous referential context and are motivated to communicate the correct interpretation of a globally ambiguous pp-attachment, they will produce disambiguating prosodic cues, much like that were used in the present study (Snedeker and Trueswell, 2003). Appendix A demonstrates that several of these cues persist when the listener is a young child (critically the IP break after the noun when the instrument interpretation is intended). However, we have also found that these strong prosodic cues are typically absent when the referential context is unambiguous. Under these circumstances, short utterances with PP-attachment ambiguities are generally produced as a single intonational phrase. Thus the opportunities for learning the relation between IP breaks and PP-attachment may be infrequent or limited to longer, more complex utterances (see Schafer et al., 2005; Kraljic & Brennan, 2005). But this concern by no means rules out the possibility that the mapping between syntax and prosody could be learned. It simply suggests that generalization that we are tapping might be broader. If children are tracking a broader pattern of correlation (such as the relative strength of prosodic boundaries as related to attachment height), then they may make use of more subtle prosodic cues (placement of intermediate phrase boundaries) or draw the relevant generalization from

syntactic constructions which are more consistently marked (such as clause closure ambiguities see Carlson et al., 2001; Cooper & Paccia-Cooper, 1980; Marcus & Hindle, 1990).

2. *Children could acquire a mapping between syntax and prosody without having to learn it from the input.*

This hypothesis is a natural extension of developmental theories that link prosody to the acquisition of syntax. While there is little theoretical work in this area, developmental accounts of the syntax-semantics interface suggest two alternate theories. The first theory reduces syntax to prosody (on analogy to Schlesinger, 1971). Prosodic reductionism would posit that at the earliest stages of acquisition the functions of syntactic structure are fulfilled by prosodic structure. As development precedes inconsistencies in this analysis (and perhaps in its mapping to semantics) would force the child to stretch or alter this representation, resulting in a syntactic analysis that is separate from prosody but still rooted in it. In the second theory, prosody is used as a bootstrap into syntax. For example, we might posit a learner with a rich set of innate prosodic-syntactic correspondences (on analogy to Pinker, 1984). On both of these theories the mapping between syntax and prosody is a direct product of the acquisition of syntax rather than the product of separate learning process.

3. *Children could show sensitivity to prosody in their online language comprehension without acquiring a mapping between prosodic structure and syntactic structure.*

Earlier we noted that alterations in prosody typically results in changes in the relative timing of words, thus altering the rate at which information is made accessible to the parser. These changes in the temporal dynamics of processing might be sufficient to account for some of the effects of prosody or parsing. For example, assume that parsing involves: 1) the activation of constituents in the order in which they are encountered; 2) the dissipation of this activation over time; 3) a bias to attach phrases to constituents that are more active. Such a

model might predict the pattern of prosodic effects observed in the present study, without invoking any mapping from prosody to syntax. The IP break after the verb increases the delay between the verb and the ambiguous phrase, allowing activation to dissipate and increasing the probability of NP-attachment. In contrast the IP break before the preposition might have a more profound effect on the noun (which would otherwise be highly active) than on the verb which is further upstream. To the best of our knowledge no one has developed such a model or tested its predictions. While such an account may wane in plausibility when we consider all the data on adult comprehension, it appears to be consistent with the limited information that we have on the role of prosody in children's parsing.

### **The interplay of lexical and prosodic cues in adult parsing**

Our participants used prosodic information, even in lexically-biased utterances, and lexical information, even in prosodically-biased utterances. In adults both cues had effects that emerged about 200ms after the onset of the critical word. Taken at face value this data is inconsistent with theories that suggest that either prosody or lexical information is used to select an analysis which is then evaluated by other cues. There are two limitations to the present findings. First, as we noted early the eye-movement data must be interpreted with caution in light of differences in timing between the two prosody conditions. Second, the relevance of the present data for any particular proposal depends on the nature of the prosodic or lexical information that is said to control the initial selection processes. For example, while Boland argues for privileged role for lexical information, she clearly limits this to verb argument structure (Boland, 1997; 2005). Thus information about adjuncts like instrument phrases is not available in this early more modular stage and the interaction of adjunct preferences with prosodic cues would be irrelevant to evaluating her hypothesis.

Nevertheless, the data from our action analyses allow us to address two specific proposals which have been made about the role of prosody in ambiguity resolution. First, we examine the proposal that prosody is used solely as a cue for structurally complex analyses. This claim originates with Wales & Toner (1979) who found that while prosody could increase the probability that participants would arrive at the more complex reading of an ambiguous utterance, appropriate prosody had no effect on the likelihood of deriving the structurally simpler (and more frequent) analysis. This finding is consistent with a two stage theory in which initial parsing is guided by structural simplicity and prosody serves solely as a cue to revision (but see Kjelgaard & Speer, 1999). If we assume that NP-attachments are more complex (see Rayner & Frazier, 1987), then this hypothesis would lead us to expect that Modifier Prosody should increase the probability of reaching modifier interpretation, but Instrument Prosody should have no effect on the probability of arriving at the instrument interpretation.

To explore this we needed a baseline to determine how these utterances would be interpreted in the absence of strong prosodic cues. We used the actions from the one referent conditions of S&T 2004. This study employed the same procedure, the same base sentences and the same toy sets that were used in the Experiment 2. However the utterances were produced with prosodic contours that contained no IP breaks (see Table 6) and were consistent with both NP and VP attachment (see Snedeker & Trueswell, 2003; Schafer et al., 2005). We conducted ANOVAs comparing each prosody condition against this baseline for both the children and the adults (Table 13). Because this analysis involves a comparison across experiments, our conclusions must be tentative. However, the data suggest that both prosodic forms have an influence the participants' interpretation of the ambiguous phrase.

The adults who were given Modifier Prosody, performed fewer instrument actions than the adults in S&T 2004, while those who were given Instrument Prosody performed more (22%,

40% and 61% respectively). The children showed a parallel pattern, performing instrument actions less often in the modifier condition than in the prior study, and more often in the instrument condition (32%, 47% and 67% respectively). This pattern is inconsistent with the hypothesis that prosody merely signals a more structurally complex analysis since Instrument Prosody (signaling the structurally simpler VP-attachment) has robust effect on interpretation. These analyses also suggest that although the prosodic cues signaling NP-attachment may be less common in production (see Snedeker & Trueswell, 2003; Schafer et al., 2005), they still have an effect on the interpretation of ambiguous attachments.

Our action data also allow us to explore the hypothesis that the role of prosody is limited to the revision of hypotheses that are initially generated by lexical cues (Pynte & Prieur, 1996). Were this the case, we would expect to see an interaction between prosody and verb bias in both analyses, due to an elimination of the effects of prosody when it is redundant with lexical bias. While there is no such interaction in any of these analyses (Table 13), this could simply reflect a lack of power in the analysis. In fact, there is no hint in the data that Modifier Prosody has a measurable effect on Modifier Biased utterances ( $M = 8\%$  instrument responses for both Modifier Prosody and the neutral baseline). The findings for the Instrument Bias condition are more compelling: while the adults in baseline study gave 79% of these utterances an instrument interpretation, those who received Instrument Prosody interpreted them as instrument phrases 96% of the time ( $t_1(1,5) = 2.75, p < .05; t_2(1,7) = 2.65, p < .05$ ). Thus prosody affects ambiguity resolution even when it is redundant with strong lexical cues, suggesting that its role is not limited to the revision analyses which are generated on the basis lexical biases.

### **The architecture of early parsing**

In the introduction we pointed out that prior research was consistent with two quite different theories of children's parsing: 1) a modular theory in which parsing relies solely on lexical



information to generate an initial structural analysis and 2) a interactive theory in which parsing is the product of a system that seeks coherence across multiple levels of representation (phonological, syntactic and semantic) by rapidly integrating information across levels as soon as it becomes available (Culicover & Jackendoff, 2005; Trueswell & Tanenhaus, 1995; MacDonald et al, 1994). We believe that the present studies support this second hypothesis. About 700 ms after encountering the object of the ambiguous prepositional phrase, prosody has an effect on children's fixations to the target instrument. Earlier we argued that this difference is a reflection of online interpretation rather than an artifact of temporal differences in the stimuli. Given the 150 to 200ms lag in programming eye-movements, this suggests that prosody is influencing structural interpretation just a few hundred milliseconds after children have enough information to identify the referentially ambiguous noun. Critically these effects occur even in lexically-biased utterances (see Figure 9) raising questions for any theory in which lexical information proposes candidate analyses (see the caveat above).

There is however one respect in which our data appear to support lexical modularity. Children used lexical information to interpret the ambiguous phrase about 500ms before they made use of prosodic cues. We cannot attribute this pattern to differences in the strength, validity or reliability of the two information types: in adults the lexical and prosodic effects emerged at essentially the same time and had roughly equivalent effects on the final interpretation. Nor can we write off the children's slower use of prosody to simple developmental differences in processing speed or the sensitivity of our measures: the early effects of lexical bias were as robust in the children as they were in the adults.

Data patterns like this one are sometimes taken to support two-stage models in which an initial modular parsing process is followed by an interactive revision process. We acknowledge that our data are compatible with such an analysis with the following caveats: 1) in children this

revision process would have to occur within 500ms of encountering the critical word; 2) in both adults and children this revision process must be triggered even when the initial information source (lexical biases) provides a strong constraint on the syntactic analysis; 3) over the course of development this revision process becomes so rapid that it is no longer detectable using these fine-grained temporal measures (with both the effects of verb bias and prosody emerging about 200ms after the prepositional object onset).

However, we see two problems with this analysis. First, there is no evidence that young children are willing and/or able to revise their initial structural commitments. In fact, Trueswell and colleagues have found that five year olds, unlike adults and older children, will persist in analyzing a prepositional phrase as VP-attached even when a subsequent phrase clearly rules out this interpretation (Trueswell et al., 1999; Trueswell & Gleitman, 2004). Second, because our stimuli are globally ambiguous it is not clear what would cause children to initiate a revision process. Both analyses are syntactically well-formed and both are supported by the visual context. One possibility is that revision is triggered when the lexical cues generate two equally viable candidate structures. But this proposal would make the incorrect prediction that revision, and hence the effects of prosody, would be largely limited to utterances without strong lexical biases. Another possibility is that revision is triggered when the initial analysis is found to be semantically implausible. However, this proposal has equal difficulty in accounting for the effect of prosody in the lexically-biased utterances. Instrument Bias verbs should support VP-attachments in the first stage of processing and the global plausibility of this analysis should generally be quite high (see below).

Instead, we tentatively suggest that the difference in timing between the lexical and prosodic cues could be linked the amount of time that elapses between the point at which the relevant information is presented and the point at which the ambiguous prepositional phrase is

encountered. We manipulated lexical bias by varying the verb, which in these utterances appears more than 1000ms before the preposition. Thus there is ample time for the children to retrieve information about the verb's meaning or the distributional environments in which it occurs before the ambiguity arises. In contrast the critical prosodic cue is probably the presence or absence of an intonational phrase boundary before the preposition (Snedeker & Trueswell, 2003; Schafer, Speer & Warren, 2003). While the prosodic boundary before the direct-object noun in the Modifier Prosody condition is also relevant, its interpretation may be unclear for the children until the preposition is encountered, since a strong boundary in that location (as in 10) could render the first boundary uninformative (Snedeker & Trueswell, 2003; Schafer, et al., 2003; Carlson et al., 2001)

(10) (You can feel) (the frog) (with the feather)

Since the critical prosodic break occurs immediately before the ambiguous preposition, children have little time to perceive it, interpret it and determine its implications for the syntactic analysis. Adults may simply be faster at these operations regardless of whether the information is lexical or prosodic.

Thus we conclude that children, like adults, have a parsing system that can make rapid use of at least two types of information. While lexical information had an earlier effect in the present studies, this may reflect differences in the timing of the two cues rather than architectural constraints of the types of information that are used in parsing. These data, however, cannot tell us why children use lexical and prosodic information but fail to use some referential manipulations. As we noted earlier, this could reflect either differences in how readily the cues are learned or an early developmental reliance on bottom-up information. Teasing apart these possibilities will require exploring children use of a top-down constraint that is a highly valid predictor of structure.

**Final Words:**

The results of these experiments clearly demonstrate that children are able to use prosody to constrain their syntactic analyses. In fact, when children do not have to switch between responses, they rely on prosodic cues to the same extent as adults. In both groups these cues are used quite quickly. These results, in conjunction with S&T 2004, demonstrate the children, like adults, rapidly use multiple constraints to arrive at a syntactic analysis. Thus our findings are consistent with interactive constraint-based models and suggest that children's failure to use referential constraints in prior studies reflects properties of that information source or its relation to syntactic structure. Finally our results suggest that both adults and children treat these information sources in a symmetric fashion—prosody influences interpretation even in the presence of strong lexical biases, and lexical biases continue to affect online interpretation of prosodically-biased utterances.

## **Appendix A: Prosodic Disambiguation in Child-Directed Speech.**

To learn more about the prosodic cues that might be available in children's input we conducted a referential communication study (Glucksberg & Krauss, 1967) with young children and their mothers.

### Methods

#### Participants

Sixteen native English speaking mother-child dyads participated. The children were 4;9 to 5;7 (M = 5;3). The mother acted as the speaker, while the child played the role of listener.

#### Procedure

The procedure was closely modeled on the referential communication task employed by Snedeker and Trueswell (2003). During the experiment, the mother and child sat on opposite sides of a screen that prevented them from seeing one another and thus exchanging any information via gesture or eye-gaze. The study was conducted by two experimenters, one who worked with the child and another who worked with the parent. At the start of every trial, each experimenter laid out a set of toys in front of their participant, labeling each as they did so. Next the experimenter on the mother's side of the screen demonstrated an action using the toys. This action could not be seen by the child. The mother was given a card with a written sentence describing the action, which she was told to memorize. After she did this, the experimenter took away the card and demonstrated the action again. The mother then asked the child if s/he was ready and produced the target sentence. The children were told that their job was to listen carefully to their mother and try to do same thing with their toys that the experimenter had done. The mother's utterances were audio taped, and the child's actions were videotaped. Two practice trials occurred before the first critical trial to ensure that the participants had ample

opportunity to learn the game. Children were given corrective feedback on the practice trials if necessary, and were consistently praised for their actions during the critical trials.

### Stimuli

Subjects were tested in one of two conditions. In the ambiguous condition, the target sentences contained an ambiguous prepositional phrase attachment which was disambiguated for the mother by the action that accompanied it (as in 1a and 1b below). In the unambiguous condition, the same two kinds of events were described using an unambiguous structure (1c and 1d). On each trial the mother and child were given identical sets of objects. Each set contained: 1) a Target Instrument (e.g., a large flower); 2) a Marked Animal, a stuffed animal carrying a small replica of the instrument (a frog holding a little flower); 3) an Unmarked Animal (an empty-handed frog); 4) a Distractor Animal, an unrelated animal wearing or carrying a different miniature object and 5) a Distractor Object, an unrelated full-scale object.

The experimenter demonstrated one of two actions for the mother. The Instrument action involved the Instrument and the Unmarked Animal (e.g., the experimenter picked up the large flower and tapped the frog that was not holding anything). The Modifier action involved the Marked Animal and did not involve the Instrument (e.g., using her hand, the experimenter tapped the frog that had the small flower).

1. a. *Tap the frog with the flower.* (Ambiguous, Instrument Demonstration)
- b. *Tap the frog with the flower.* (Ambiguous, Modifier Demonstration)
- c. *Use the flower to tap the frog.* (Unambiguous, Instrument Demonstration)
- d. *Tap the frog that has the flower.* (Unambiguous, Modifier Demonstration)

The sixteen target sentences from Snedeker & Trueswell (2003) were divided into two presentation lists and two versions of each list were constructed to counterbalance demonstration type across items. An ambiguous and unambiguous version of each list was created. Thus

ambiguity was manipulated between subjects and demonstration was manipulated within subject. In each list, the eight target trials were interspersed with the nine filler sentences which used a variety of objects and sentence types.

### Coding

Acoustic analyses of the ambiguous target sentences were performed using speech waveform displays of the mothers' target utterances. Coders, who were blind to the condition, measured the duration of: the verb, the pause after the verb, the direct-object noun, the pause after the noun, and the prepositional phrase. Actions were coded in the same way as in Experiments 1 and 2. Six trials were excluded because the mother did not produce the target utterance and two were excluded due to experimenter error.

-----  
Table 14  
-----

### Results and Discussion

Table 14 compares the Modifier and Instrument utterances produced by the mothers in the ambiguous condition. In a parallel task with college-aged speakers and listeners, Snedeker & Trueswell (2003) found: 1) a dramatic increase in the length of direct-object nouns and substantial post-nominal pauses for Instrument utterances, reflecting the placement of an intonational phrase boundary after the noun; 2) a smaller but robust increase in the length of the prepositional phrases for Instrument utterances; and 3) a reliable but weaker tendency toward longer verbs and post-verbal pauses in Modifier utterances, reflecting the placement of intonational phrase boundaries after the verb. The current study replicates the first two effects quite precisely: both the direct-object nouns and the prepositional phrases in the Instrument utterances were 45% longer than in the Modifier utterances and the probability of a pause

occurring between them increased from 11% to 67%. The third effect was not present in the current study: mothers were equally likely to pause after the verb for Modifier utterances as they were for Instrument utterances.

To determine whether children were able to use this prosodic information to disambiguate the utterances we examined the proportion of Instrument responses in each of the four conditions. In the unambiguous condition the children's actions matched the experimenter's demonstration on nearly all of the trials, resulting in a strong effect of demonstration type ( $M = 100\%$  and  $M = 9\%$  for Instrument and Modifier Demonstrations respectively;  $F_1(1,4) = 168.20$ ,  $p < .001$ ;  $F_2(1,12) = 128.65$ ,  $p < .001$ ;  $\text{min}F'(1,14) = 72.89$ ,  $p < .001$ ). Thus children clearly understood the game and could perform both types of actions. In contrast, when children heard syntactically ambiguous sentences their actions were unaffected by the action that their mother had seen ( $M = 48\%$  and  $M = 51\%$  for Instrument and Modifier Demonstrations respectively;  $F's < 1$ ,  $p's > .5$ ). Thus we find no evidence that children are able to use the prosodic cues provided in this task to guide their analysis of the ambiguous utterance. A closer examination of the data suggests that perseveration may have played a role in the children's failure to use prosody. The children quickly closed in on a single interpretation of the ambiguous with-phrase and rarely deviated from that interpretation (of the eight children in the ambiguous condition 2 children never performed their dispreferred response, 3 children did so once, and 3 did so twice).



**Footnotes**

1. There is one study which suggests that five year olds can use prosody to constrain syntactic analysis. Beach, Katz, and Skowronski (1996) found that children could use prosodic boundary cues to interpret a grouping ambiguity with adult-like accuracy ["pink and (green and white)" versus "(pink and green) and white"]. This task differs from the studies we reviewed in three ways. First, the same sentence was presented on every trial, potentially highlighting the contrast between prosodic variants. Second, children were instructed about how intonation could be used for grouping and given an example of a good prosodic contour for each interpretation. Finally, interpretation was assessed by having participants select a picture in which the order of the colored animals on the page (from left to right) matched the order in which the colors were mentioned and the space between the animals matched the intended prosodic grouping.

**Author Notes**

These experiments grew out of an earlier collaboration with John Trueswell and an idle comment made by Jim Morgan. We are grateful to Isabel Martin, Dina Roumiansteva, Julie Seet, Elaine Chung and Mahvash Malik for their assistance with the eye-movement studies, which were supported by Harvard University. We thank Stacey Rubin, Sarah Brown-Schmidt and Jared Novick for their assistance with the referential communication experiment, which was supported by an NIH grant to John Trueswell (1-R01-HD37507) and a National Science Foundation Science and Technology Center grant to the Institute for Research in Cognitive Science at the University of Pennsylvania (NSF-STC Cooperative Agreement number SBR-89-20230). The ToBI analyses were expertly executed by Alejna Brugos. These analyses and the preparation of this manuscript were supported by an NSF grant to the first author (NSF-BCS 0623945).

**References**

- Allbritton, D., McKoon G. & Ratcliff, R. (1996). Reliability of prosodic cues for resolving syntactic ambiguity. Journal of Experimental Psychology: Learning, Memory, and Cognition 22, 714–735.
- Allopenna, P.D., Magnuson, J.S. & Tanenhaus, M.K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. Journal of Memory & Language. Vol 38(4), 419-439.
- Altmann, G.T.M. (1998) Ambiguity in Sentence Processing. Trends in Cognitive Sciences, 2(4), 146-152.
- Altmann, G. & Steedman, M. (1988). Interaction with context during human sentence processing. Cognition, 30, 191-238.
- Bates, E.A. & MacWhinney, B. (1987). Competition, variation and language learning. In MacWhinney, B., (ed.), Mechanisms of Language Acquisition (pp. 157-194). Lawrence Erlbaum, Hillsdale, NJ.
- Baumann, S. (2005). Degrees of Givenness and their Prosodic Marking. Paper presented at the International Symposium on "Discourse and Prosody as a complex interface", September, 2005, Aix-en-Provence.
- Beach, C.M., Katz, W.F., & Skowronski, A. (1996). Children's processing of prosodic cues for phrasal interpretation. Journal of the Acoustical Society of America, 99(2), 1148-1160.
- Beckman, M. E., & Hirschberg, J. (1994). The ToBI annotation conventions. Columbus, OH: Ohio State University.
- Bock, J. K. (1986). Syntactic persistence in language production. Cognitive Psychology, 18, 355-387.

- Boland, J. & Cutler, A. (1996). Interaction with autonomy: Multiple output models and the inadequacy of the Great Divide. Cognition, 58, 309-320.
- Boland, J. E. (1997). The relationship between syntactic and semantic processes in sentence comprehension. Language and Cognitive Processes, 12, 423-484.
- Boland, J.E. (2005). Visual Arguments. Cognition, 95, 237-274.
- Britt, M. A. (1994). The interaction of referential ambiguity and argument structure in the parsing of prepositional phrases. Journal of Memory and Language, 33, 251-283.
- Brown-Schmidt, S., Campana, E. & Tanenhaus, M.K. (2005). Reference resolution in the wild: Circumscription of referential domains by naive participants during an interactive problem solving task. In J.C. Trueswell & M.K. Tanenhaus (Eds.) Approaches to Studying World Situated Language Use. Cambridge, MA: MIT Press.
- Carlson, K., Clifton, C., & Frazier, L. (2001). Prosodic boundaries in adjunct attachment. Journal of Memory & Language, 45(1), 58-81.
- Chase, C. H., & Tallal, P. (1990). A developmental, interactive-activation model of the word superiority effect. Journal of Experimental Child Psychology, 49, 448-487.
- Choi, Y. & Mazuka, R. (2003). Young children's use of prosody in sentence parsing. Journal of Psycholinguistic Research, 32, 197-217.
- Christophe, A., Guasti, M. T., Nespors, M., & van Ooyen, B. (2003). Prosodic structure and syntactic acquisition: the case of the head-complement parameter. Developmental Science, 6, 213-222.
- Christophe, A., Mehler, J. & Sebastián-Gallés, N. (2001). Perception of prosodic boundary correlates by newborn infants. Infancy, 2, 385-394.
- Collins, M. & Brooks, J. (1995). Prepositional Phrase Attachment through a backed-off model. Proceedings of the Third Workshop on Very Large Corpora.

- Collins, M. (1997). Three Generative, Lexicalized Models for Statistical Parsing. Proceedings of the 35th Annual Meeting of the ACL (jointly with the 8th Conference of the EACL), Madrid.
- Cooper, W.E. & Paccia-Cooper J. (1980). Syntax and Speech. Cambridge MA: Harvard University Press.
- Crain, S. & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological parser. In Dowty, Karrattunen & Zwicky. Natural Language Parsing, Cambridge: Cambridge University Press.
- Culicover, P. & Jackendoff, R. (2005). Simpler syntax. Oxford: Oxford University Press.
- Cutler, A., Dahan, D., and van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. Language & Speech, 40(2), 141-201.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. Psychological Review, 93, 283-321.
- Engelhardt, P.E., Bailey, K.G.D. & Ferreira, F. (2006). Do speakers and listeners observe the Gricean Maxim of Quantity? Journal of Memory and Language, 54(4), 554-573.
- Fernald, A. & McRoberts, G. (1996). Prosodic bootstrapping: A critical analysis of the argument and the evidence. In J. Morgan & K. Demuth (Eds), Signal to Syntax Mahwah, NJ: Erlbaum.
- Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. Journal of Memory and Language, 30, 210-233.
- Ferreira, F. & Clifton, C. (1986). The independence of syntactic processing. Journal of Memory and Language, 25, 348-368.
- Flavell, J.H. (1986). The development of children's knowledge about the appearance-reality distinction. American Psychologist, 41, 418-425.

Frank, R. (1998). Structural complexity and the time course of grammatical development.

Cognition, 66, 249 – 301.

Frazier, L. & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model.

Cognition, 6, 291-325.

Gerken L. (1996). Prosody's role in language acquisition and adult parsing. Journal of

Psycholinguistic Research, 25(2), 345-56.

Glucksberg, S., & Krauss, R. (1967). What do people say after they have learned to talk?

Studies in the development of referential communication. Merrill-Palmer Quarterly, 13, 309-316.

Goodluck, H. & Tavakolian, S. (1982). Competence and processing in children's grammar of relative clauses. Cognition, 11, 1-27.

Halbert, A., S. Crain, D. Shankweiler, and E. Woodams (1995) Children's Interpretive Use of Emphatic Stress, presented at the 8th Annual CUNY Conference on Human Sentence Processing, Tucson, AZ.

Hawkins, J. A. (1978). Definiteness and indefiniteness: A study in reference and grammaticality prediction. Atlantic Highlands, N.J.: Humanities Press.

Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Wright-Cassidy, K., Druss, B., &

Kennedy, L. (1987). Clauses are perceptual units for young infants. Cognition, 26, 269–286.

James, V. (2001). Lexical access in preschool children. Unpublished Doctoral Dissertation, University of South Carolina.

Johnson, E.K. & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. Journal of Memory & Language. 44, 548-567.

Jusczyk, P. W., Friederici, A. D., Wessels, J. M. I., Svenkerud, V. Y., & Jusczyk, A. M. (1993).

Infants' sensitivity to the sound patterns of native language words. Journal of Memory and Language 32, 402–420.

Jusczyk, P. W., Culter, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. Child Development, 64, 675–687.

Jusczyk, P. W., Hirsh-Pasek, K., Kemler Nelson, D. G., Kennedy, L., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. Cognitive Psychology, 24, 252–293.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. Psychological Review, 98, 122–149.

Kail, R.V. (1991). Development of processing speed in childhood and adolescence. In H. W. Reese (Ed.), Advances in child development and behavior (Vol. 25). New York: Academic Press.

Kail, R., & Park, Y.-S. (1994). Processing time, articulation time, and memory span. Journal of Experimental Child Psychology, 57(2), 281- 291.

Kidd, E. & Bavin, E.L. (2005). Lexical and referential cues to sentence interpretation: An investigation of children's interpretations of ambiguous sentences. Journal of Child Language, 32(4), 855-876.

Kjelgaard, M.M. & Speer, S.R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. Journal of Memory and Language 40, 153–194.

Koenig, J., Mauner, G. & Bienvenue, B. (2003). Arguments for adjuncts. Cognition, 89, 67-103.

Kraljic, T. & Brennan, S.E. (2005). Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? Cognitive Psychology, 50(2), 194-231.

- Lederer, A., Gleitman, L. & Gleitman, H. (1995). Verbs of a feather flock together: Structural properties of maternal speech. In M. Tomasello and E. Merriam (Eds.), Beyond Words for Things: Acquisition of the Verb Lexicon. New York: Academic Press.
- Lehiste, I. (1973). Phonetic disambiguation of syntactic ambiguity. Glossa, 7, 102-122.
- Lehiste, I., Olive, J.P., & Streeter, L.A. (1976). Role of duration in disambiguating syntactically ambiguous sentences. Journal of the Acoustical Society of America, 60, 1199-1202.
- MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. Cognitive Psychology, 24, 56-98.
- MacDonald, M.C., Pearlmutter, N.J., & Seidenberg, M.S. (1994). The lexical nature of syntactic ambiguity resolution. Psychological Review, 101, 676-703.
- Marcus, M. (1994). New Trends in Natural Language Processing: Statistical Natural Language Processing, Proceedings of the National Academy of Science, Vol 92, 10052-10059.
- Marcus, M., & Hindle, D. (1990). Description theory and intonation boundaries. In G. Altmann (Ed.), Cognitive models of speech processing: Computational and psycholinguistic perspectives. Cambridge, MA: MIT Press.
- Marslen-Wilson, W.D., Tyler, L.K., Warren, P., Grenier, P., & Lee, C.S. (1992). Prosodic effects in minimal attachment. Quarterly Journal of Experimental Psychology, 45A, 73-87.
- Mazuka, R. & Tanaka, Y. (2006). Children can be better than adults at using prosody to resolve syntactic ambiguity, Poster presented at the 19th Annual CUNY Conference on Human Sentence Processing, New York, NY.
- Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. Cognition, 29, 143-178.
- Morgan, J. L., & Saffran, J. R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. Child Development, 66, 911-936.



- Morgan, J. L. (1996). A rhythmic bias in preverbal speech segmentation. Journal of Memory & Language. 35, 666-688.
- Nagel, H.N., Shapiro, L.P., Tuller, B., & Nawy, R. (1996). Prosodic influences on the resolution of temporary ambiguity during on-line sentence processing. Journal of Psycholinguistic Research, 25(2), 319-344.
- Permer, J. & Wimmer, H. (1985). "John thinks that Mary thinks that . . .": Attribution of second-order beliefs by 5- to 10-year-old children. Journal of Experimental Child Psychology. 39, 437-471.
- Piaget, J. (1946). *The development of children's concept of time*. Paris: Presses Universitaires de France.
- Pierrehumbert J.B. & Hirschberg J. (1990), The Meaning of Intonational Contours in the Interpretation of Discourse, in Cohen P.R., Morgan J. & Pollack M.E. (eds.), Intentions in Communication, Cambridge, MIT Press, 271-311
- Pinker, S. (1984). Language Learnability and Language Development. Cambridge, MA: Harvard University Press.
- Price, P., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. Journal of the Acoustical Society of America, 90, 2956-2970.
- Prince, E. F. (1981). Toward a taxonomy of given-new information. In Peter Cole (ed.), Radical Pragmatics. New York: Academic Press. Pp. 223-256
- Pynte, J. & Prieur, B, (1996). Prosodic breaks and attachment decisions in sentence parsing. Language and Cognitive Processes, 11, 165-192.
- Rayner, K. and Frazier, L. (1987). Parsing temporarily ambiguous complements. Quarterly Journal of Experimental Psychology 39A, 657-673.

Robinson, E. J. & Robinson, W. P. (1982). Knowing when you don't know enough: Children's judgments about ambiguous information. Cognition, 12(3), 267-280.

Schafer, A. J. (1997). Prosodic parsing: The role of prosody in sentence comprehension.

Unpublished doctoral dissertation, Amherst, MA: University of Massachusetts.

Schafer, A., Speer, S., & Warren, P. (2005). Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. In Tanenhaus & Trueswell (Eds.) Approaches to Studying World Situated Language Use. Cambridge: MIT Press.

Schlesinger, I. M. (1971) Production of utterances and language acquisition. In D. I. Slobin , ed., The Ontogenesis of Grammar. New York: Academic Press.

Schneider, W., & Bjorklund, D. F. (1998). Memory. In D. Kuhn & R. S. Siegler (Eds.), Cognitive, language, and perceptual development, Vol. 2 (pp. 467-521). In W. Damon (General Editor), Handbook of child psychology (5th Ed.). New York: Wiley.

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

Snedeker, J. & Trueswell, J. (2003). Using Prosody to Avoid Ambiguity: Effects of Speaker Awareness and Referential Context, Journal of Memory and Language, 48, 103-130.

Snedeker, J. & Trueswell, J. (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.

Spivey-Knowlton, M., & Sedivy, J.C. (1995). Resolving attachment ambiguities with multiple constraints. Cognition, 55(3), 227-267.

- Spivey, M.J., Tanenhaus, M.K., Eberhard, K.M. & Sedivy, J.C. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. Cognitive Psychology, 45, 447-481.
- Steedman, M. (2000). Information structure and syntax-phonology interface. Linguistic Inquiry 31(4), 649-689.
- Steinhauer, K., Alter, K., & Friederici, A.D. (1999). Brain responses indicate immediate use of prosodic cues in natural speech processing. Nature Neuroscience, 2, 191-196.
- Taraban, R., & McClelland, J. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. Journal of Memory and Language, 27, 1-36.
- Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. Science, 268, 1632-1634.
- Tanenhaus, M.K. & Trueswell, J.C. (1995). Sentence Comprehension. In Eimas & Miller (Eds.) Handbook in Perception and Cognition, Volume 11: Speech Language and Communication. Academic Press.
- Trueswell, J. & Gleitman, L.R. (2004). Children's eye movements during listening: evidence for a constraint-based theory of parsing and word learning. In J. M. Henderson & F. Ferreira (eds.). Interface of Language, Vision, and Action: Eye Movements and the Visual World. NY: Psychology Press.
- Trueswell, J.C., Sekerina, I., Hill, N.M., & Logrip, M.L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. Cognition, 73, 89-134.
- Trueswell, J.C. & Tanenhaus, M. K. (1994). Toward a lexicalist framework of constraint-based syntactic ambiguity resolution. In Clifton, and Frazier (Eds), Perspectives on sentence processing. Hillsdale, NJ: Lawrence Erlbaum.

- Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 528-553.
- van Berkum, J, Brown, C, & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, *41*, 147-182.
- Vogel, I. & Raimy, E. (2002). The acquisition of compound vs. phrasal stress: the role of prosodic constituents. *Journal of Child Language*, *29*, 225-250.
- Wales, R. & Toner, H. (1979). Intonation and ambiguity. In W.E. Cooper and E.C.T. Walker (Eds.), *Sentence Processing: Psycholinguistic studies presented to Merrill Garrett*. Hillsdale, NJ: Erlbaum.
- Watson, D. & Gibson, E. (2004). The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes* *19*, 713–755.
- Weber, A., Grice, M., & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: a study of anticipatory eye movements. *Cognition*, *99*, B63-B72.
- Welsh, M., Pennington, B., & Groisser, D. (1991). A normative-developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology*. *7*, 131-149.

Table 1: Duration analyses for the stimuli in Experiment 1.

Dependent Variable	Mean for Instrument Prosody	Mean for Modifier Prosody	Analysis
verb length	372 ms $CI_{.95} = \pm 54$ ms	625 ms $CI_{.95} = \pm 79$ ms	$t(7) = 8.57, p < .001^{**}$
verb pause	3 ms $CI_{.95} = \pm 5$ ms	115 ms $CI_{.95} = \pm 26$ ms	$t(7) = 6.98, p < .001^{**}$
direct-object noun	478 ms $CI_{.95} = \pm 60$ ms	236 ms $CI_{.95} = \pm 35$ ms	$t(7) = 13.86, p < .001^{**}$
noun pause	235 ms $CI_{.95} = \pm 43$ ms	0 ms	$t(7) = 9.43, p < .001^{**}$
"with"	191 ms $CI_{.95} = \pm 15$ ms	157 ms $CI_{.95} = \pm 12$ ms	$t(7) = 3.46, p < .01^*$
prepositional object	455 ms $CI_{.95} = \pm 92$ ms	473 ms $CI_{.95} = \pm 58$ ms	$t(7) < 1, p > .5$

Table 2: Action analysis for Experiment 1 (children). The dependant variable is in percentage of instrument actions.

	All Blocks	Block 1	Block 2
Mean Instrument Prosody	61%, $CI_{.95} = \pm 10\%$	58% $CI_{.95} = \pm 18\%$	65% $CI_{.95} = \pm 19\%$
Mean Modifier Prosody	40% $CI_{.95} = \pm 10\%$	21% $CI_{.95} = \pm 12\%$	58% $CI_{.95} = \pm 18\%$
Prosody	$F1(1,16) = 9.80, p < .01^*$ $F2(1,6) = 8.82, p < .05^*$ $\text{min}F'(1,17) = 4.64, p < .05^*$	$F1(1,20) = 10.39, p < .005^{**}$ $F2(1,6) = 7.25, p < .05^*$ $\text{min}F'(1,15) = 4.27, p = .06$	$F1(1,20) < 1, p > .5$ $F2(1,6) < 1, p > .5$ $\text{min}F'(1,18) < 1, p > .5$
Age (K or Pre-K)	$F1(1,16) < 1, p > .5$ $F2(1,6) < 1, p > .5$ $\text{min}F'(1,17) < 1, p > .5$	$F1(1,20) < 1, p > .5$ $F2(1,6) < 1, p > .5$ $\text{min}F'(1,18) < 1, p > .5$	$F1(1,20) < 1, p > .5$ $F2(1,6) < 1, p > .5$ $\text{min}F'(1,18) < 1, p > .5$
Prosody * Age	$F1(1,16) < 1, p > .4$ $F2(1,6) = 2.46, p > .15$ $\text{min}F'(1,22) < 1, p > .4$	$F1(1,20) < 1, p > .5$ $F2(1,6) < 1, p > .5$ $\text{min}F'(1,18) < 1, p > .5$	$F1(1,20) = 1.15, p > .25$ $F2(1,6) = 2.54, p > .15$ $\text{min}F'(1,25) < 1, p > .3$

Table 3: Coarse grained analysis of fixations for Experiment 1 (children). The dependant variable is the proportion of trials with looks to the Target Instrument.

	All Blocks	Block 1	Block 2
Prosody	F1(1,16) = 12.05, p < .005** F2(1,6) = 5.23, p = .06 minF'(1,12) = 3.65, p = .08	F1(1,20) = 6.33, p < .05* F2(1,6) = 3.04, p = .125 minF'(1,12) = 2.05, p = .18	F1(1,20) < 1, p > .5 F2(1,6) < 1, p > .3 minF'(1,18) < 1, p > .5
Age (K vs. Pre-K)	F1(1,16) < 1, p > .5 F2(1,6) < 1, p > .3 minF'(1,17) < 1, p > .5	F1(1,20) < 1, p > .5 F2(1,6) < 1, p > .5 minF'(1,18) < 1, p > .5	F1(1,20) < 1, p > .3 F2(1,6) = 1.13, p > .3 minF'(1,20) < 1, p > .4
Prosody * Age	F1(1,16) = 2.15, p > .15 F2(1,6) < 1, p > .4 minF'(1,17) < 1, p > .5	F1(1,20) < 1, p > .5 F2(1,6) < 1, p > .5 minF'(1,18) < 1, p > .5	F1(1,20) < 1, p > .5 F2(1,6) < 1, p > .3 minF'(1,18) < 1, p > .5

Table 4: Temporal analyses of fixations for Experiment 1 (children). The dependant variable is the proportion of looking time to the Target Instrument.

	Early PP-Object	Late PP-Object
Mean Inst Prosody	19%, $CI_{.95} = \pm 15\%$	29%, $CI_{.95} = \pm 16\%$
Mean Mod Prosody	6%, $CI_{.95} = \pm 8\%$	8%, $CI_{.95} = \pm 6\%$
Prosody	F1(1,20) = 2.22, p = .15 F2(1,6) = 2.46, p = .17 minF'(1,19) = 1.17, p > .25	F1(1,20) = 5.02, p < .05* F2(1,6) = 7.56, p < .05* minF'(1,22) = 3.02, p = .10
Age (K or Pre-K)	F1(1,20) < 1, p > .4 F2(1,6) < 1, p > .4 minF'(1,18) < 1, p > .5	F1(1,20) < 1, p > .4 F2(1,6) < 1, p > .4 minF'(1,18) < 1, p > .5
Prosody * Age	F1(1,20) < 1, p > .4 F2(1,6) = 4.38, p = .081 minF'(1,26) < 1, p > .3	F1(1,20) < 1, p > .4 F2(1,6) < 1, p > .4 minF'(1,18) < 1, p > .5



Table 5: Analysis of trials in which participants were fixating on the Target Animal prior to the prepositional object onset for Experiment 1 (children). The dependant variable is the proportion of looking time to the Target Instrument.

	Early PP-Object	Late PP-Object
Mean Inst Prosody	18%, $CI_{.95} = \pm 15\%$	26%, $CI_{.95} = \pm 15\%$
Mean Mod Prosody	6%, $CI_{.95} = \pm 12\%$	10%, $CI_{.95} = \pm 9\%$
Prosody	F1(1,21) = 1.57, p > .2 F2(1,6) = 1.22, p > .3 minF'(1,19) = 1.17, p > .25	F1(1,21) = 3.27, p = .085 F2(1,6) = 4.27, p = .084 minF'(1,21) = 1.85, p = .19

Table 6: Duration analyses for the stimuli in Experiment 2.

Dependent Variable	Mean Instrument Prosody	Mean Modifier Prosody	Analysis (Modifier vs. Instrument)	Mean Neutral Prosody <sup>†</sup>
verb length	306 ms <i>CI</i> <sub>.95</sub> = ± 31 ms	608 ms <i>CI</i> <sub>.95</sub> = ± 59 ms	t (23) = 13.69, p < .001**	354 ms <i>CI</i> <sub>.95</sub> = ± 35 ms
verb pause	14 ms <i>CI</i> <sub>.95</sub> = ± 11 ms	238 ms <i>CI</i> <sub>.95</sub> = ± 41 ms	t (23) = 10.31, p < .001**	7 ms <i>CI</i> <sub>.95</sub> = ± 6 ms
direct-object noun	482 ms <i>CI</i> <sub>.95</sub> = ± 22 ms	212 ms <i>CI</i> <sub>.95</sub> = ± 16 ms	t (23) = 31.23, p < .001**	285 ms <i>CI</i> <sub>.95</sub> = ± 18 ms
noun pause	218 ms <i>CI</i> <sub>.95</sub> = ± 28 ms	0 ms	t (23) = 15.30, p < .001**	1 ms <i>CI</i> <sub>.95</sub> = ± 1 ms
"with"	173 ms <i>CI</i> <sub>.95</sub> = ± 7 ms	127 ms <i>CI</i> <sub>.95</sub> = ± 6 ms	t (23) = 7.70, p < .001**	142 ms <i>CI</i> <sub>.95</sub> = ± 6 ms
prepositional object	523 ms <i>CI</i> <sub>.95</sub> = ± 51 ms	490 ms <i>CI</i> <sub>.95</sub> = ± 31 ms	t (23) = 2.16, p < .05*	475 ms <i>CI</i> <sub>.95</sub> = ± 44 ms

<sup>†</sup> The Neutral Prosody utterances were the stimulus sentences from Snedeker & Trueswell (2004). They included the same root command without the carrier phrase ("You can").

Table 7: Analyses of the prosodic transcriptions for the stimuli in Experiment 2 and Snedeker &amp; Trueswell (2004)

Dependent Variable	Instrument Prosody	Modifier Prosody	Neutral Prosody <sup>†</sup>
verb break index	$M = 1.08$ $CI_{.95} = \pm .16$	$M = 4$	$M = 1.08$ $CI_{.95} = \pm .16$
noun break index	$M = 4$	$M = .92$ $CI_{.95} = \pm .11$	$M = 1.33$ $CI_{.95} = \pm .31$
pitch accent on verb	54% H* 25% L* 21% L*+H	100% L+H*	38% H* 62% L*+H
pitch accent on noun	21% H+!H* 46% H* 33% L+H*	79% H* 21% L+H*	46% H+!H* 42% H* 12% L+H*
presence of pitch accent on <i>with</i>	79%	0%	0%
accent on PP-object	58% H* 42% L+H*	67% H* 33% L+H*	67% H* 33% L+H*
break indices prediction	100% Instrument	100% Modifier	79% Neutral 17% Instrument 4% Modifier

<sup>†</sup> The Neutral Prosody utterances were the stimulus sentences from Snedeker & Trueswell (2004). They included the same root command without the carrier phrase (“You can”).

Table 8: Action analysis for Experiment 2. The dependant variable is in percentage of instrument actions.

	All Blocks	Block 1	Block 2
<i>Adults</i>			
Prosody	F1(1,24) = 37.69, p < .001** F2(1,18) = 101.04, p < .001** minF'(1,38) = 27.45, p < .001**	F1(1,24) = 27.62, p < .001** F2(1,18) = 37.97, p < .001** minF'(1,42) = 15.99, p < .001**	F1(1,24) = 11.00, p < .005** F2(1,18) = 54.00, p < .001** minF'(1,33) = 9.14, p < .005**
Verb Bias	F1(2,24) = 24.23, p < .001** F2(2,18) = 37.00, p < .001** minF'(2,42) = 14.64, p < .001**	F1(2,24) = 25.81, p < .001** F2(2,18) = 32.73, p < .001** minF'(2,42) = 14.43, p < .001**	F1(2,24) = 8.61, p < .005** F2(2,18) = 12.19, p < .001** minF'(2,42) = 5.05, p < .05*
Prosody * Bias	F1(2,24) = 2.85, p = .077 F2(2,18) = 6.07, p = .01* minF'(2,40) = 1.93, p > .15	F1(2,24) = 2.04, p = .148 F2(2,18) = 2.73, p = .093 minF'(2,42) = 1.17, p > .3	F1(2,24) = 1.39, p > .25; F2(2,18) = 5.91, p = .011 minF'(2,34) = 1.13, p > .3
<i>Children</i>			
Prosody	F1(1,48) = 23.17, p < .001** F2(1,18) = 30.64, p < .001** minF'(1,59) = 13.19, p < .001**	F1(1,48) = 29.52, p < .001** F2(1,18) = 60.48, p < .001** minF'(1,65) = 19.83, p < .001**	F1(1,48) < 1, p > .5 F2(1,18) < 1, p > .5 minF'(1,52) < 1, p > .5
Verb Bias	F1(2,48) = 42.08, p < .001** F2(2,18) = 56.94, p < .001** minF'(2,59) = 24.19, p < .001**	F1(2,48) = 35.74, p < .001** F2(2,18) = 37.46, p < .001** minF'(2,53) = 18.29, p < .001**	F1(2,48) = 25.51, p < .001** F2(2,18) = 66.82, p < .001** minF'(2,66) = 18.46, p < .001**
Prosody * Bias	F1(2,48) < 1, p > .25 F2(2,18) < 1, p > .25 minF'(2,52) , 1, p > .5	F1(2,48) = 2.04, p > .2 F2(2,18) = 2.73, p = .093 minF'(2,52) = 1.17, p > .3	F1(2,48) < 1, p > .5 F2(2,18) < 1, p > .5 minF'(2,52) < 1, p > .5

Table 9: Coarse grained analyses of fixations for Experiment 2. The dependant variable is the proportion of trials with looks to the Target Instrument.

	All Blocks	Block 1	Block 2
<i>Adults</i>			
Prosody	F1(1,24) = 56.89, p < .001** F2(1,18) = 78.66, p < .001** minF'(1,42) = 33.01, p < .001**	F1(1,30) = 40.80, p < .001** F2(1,21) = 34.21, p < .001** minF'(1,48) = 18.60, p < .001**	F1(1,30) = 24.59, p < .005** F2(1,21) = 49.47, p < .001** minF'(1,50) = 16.42, p < .001**
Verb Bias	F1(2,24) = 35.56, p < .001** F2(2,18) = 19.42, p < .001** minF'(2,35) = 12.56, p < .001**	F1(2,30) = 28.40, p < .001** F2(2,21) = 21.50, p < .001** minF'(2,46) = 12.24, p < .001**	F1(2,30) = 7.11, p < .005** F2(2,21) = 7.00, p < .005** minF'(2,49) = 3.53, p < .05*
Prosody * Bias	F1(2,24) = 5.23, p < .05* F2(2,18) = 7.29, p < .005** minF'(2,42) = 2.42, p = .058	F1(2,30) = 3.39, p < .05* F2(2,21) = 2.93, p = .076 minF'(2,48) = 1.57, p > .2	F1(2,30) = 3.89, p < .05* F2(2,21) = 7.82, p < .005** minF'(2,50) = 2.60, p = .08
<i>Children</i>			
Prosody	F1(1,48) = 12.24, p < .001** F2(1,18) = 11.61, p < .005** minF'(1,51) = 5.96, p < .05*	F1(1,60) = 12.17, p < .001** F2(1,21) = 13.59, p < .005** minF'(1,66) = 6.42, p < .05*	F1(1,60) = 1.27, p > .2 F2(1,21) = 1.15, p > .2 minF'(1,59) < 1, p > .4
Verb Bias	F1(2,48) = 25.69, p < .001** F2(2,18) = 17.03, p < .001** minF'(2,43) = 10.24, p < .001**	F1(2,60) = 18.68, p < .001** F2(2,21) = 11.90, p < .001** minF'(2,49) = 7.23, p < .005**	F1(2,60) = 13.72, p < .001** F2(2,21) = 17.79, p < .001** minF'(2,70) = 7.45, p < .001**
Prosody * Bias	F1(2,48) < 1, p > .4 F2(2,18) < 1, p > .4 minF'(2,51) , 1, p > .5	F1(2,60) < 1, p > .5 F2(2,21) < 1, p > .5 minF'(2,76) < 1, p > .5	F1(2,60) = 1.19, p > .3 F2(2,21) = 1.36, p > .2 minF'(2,66) < 1, p > .5

Table 10: Temporal analyses of fixations for Experiment 2. The dependant variable is the proportion of looking time to the Target Instrument.

	"With" Window	Early PP-Object	Late PP-Object
<i>Adults</i>			
Prosody	F1(1,24) < 1, p > .4 F2(1,18) = 2.42, p = .14 minF'(1,38) < 1, p > .4	F1(1,24) = 18.92, p < .001** F2(1,18) = 30.27, p < .001** minF'(1,42)=11.64, p < .005**	F1(1,24) = 17.98, p < .001** F2(1,18) = 18.82, p < .001** minF'(1,41)=9.20, p < .005**
Verb Bias	F1(2,24) = 2.06, p = .15 F2(2,18) = 7.72, p < .005** minF'(2,35) = 1.63, p > .2	F1(2,24) = 2.91, p = .07 F2(2,18) = 3.92, p < .05* minF'(2,42) = 1.67, p > .15	F1(2,24) = 8.21, p < .001** F2(2,18) = 4.64, p < .05* minF'(2,36) = 2.96, p = .06
Prosody * Bias	F1(2,24) < 1, p > .4 F2(2,18) = 2.42, p = .12 minF'(2,38) < 1, p > .4	F1(2,24) < 1, p > .4 F2(2,18) = 1.34, p > .25 minF'(2,42) < 1, p > .4	F1(2,24) < 1, p > .4 F2(2,18) < 1, p > .4 minF'(2,41) < 1, p > .5
<i>Children</i>			
Prosody	F1(1,48) = 2.26, p > .1 F2(1,18) = 1.08, p > .25 minF'(1,36) < 1, p > .3	F1(1,48) = 4.03, p = .05* F2(1,18) = 3.36, p = .08 minF'(1,48) = 1.83, p = .18	F1(1,48) = 16.23, p < .001** F2(1,18) = 20.67, p < .001** minF'(1,58) = .09, p < .005**
Verb Bias	F1(2,48) = 6.54, p < .005** F2(2,18) = 2.71, p = .093 minF'(2,34) = 1.92, p > .15	F1(2,48) = 24.43, p < .001** F2(2,18) = 6.60, p < .01* minF'(2,28) = 5.20, p < .05*	F1(2,48) = 7.06, p < .005** F2(2,18) = 9.97, p < .001** minF'(2,60) = 4.13, p < .05*
Prosody * Bias	F1(2,24) < 1, p > .4 F2(2,18) < 1, p > .4 minF'(2,52) < 1, p > .5	F1(2,48) < 1, p > .5 F2(2,18) < 1, p > .5 minF'(2,52) < 1, p > .5	F1(2,48) = 1.85, p > .15 F2(2,18) = 1.69, p > .15 minF'(2,50) < 1, p > .4

Table 11: Temporal analysis of the effects of Age on fixations in Experiment 2. The dependant variable is the proportion of looking time to the Target Instrument.

	"With" Window	Early PP-Object	Late PP-Object
Age (Child or Adult)	F1(1,84) = 7.23, p < .01* F2(1,18) = 6.14, p < .05* minF'(1,53) = 3.32, p = .07	F1(1,84) < 1, p > .4 F2(1,18) < 1, p > .5 minF'(1,59) < 1, p > .5	F1(1,84) < 1, p > .5 F2(1,18) < 1, p > .5 minF'(1,59) < 1, p > .5
Verb Bias	F1(2,84) = 7.77, p < .001** F2(2,18) = 4.85, p < .05* minF'(2,44) = 2.99, p = .06	F1(2,84) = 17.82, p < .001** F2(2,18) = 7.27, p < .005** minF'(2,34) = 5.16, p < .05*	F1(2,84) = 15.96, p < .001** F2(2,18) = 8.94, p < .005* minF'(2,41) = 5.73, p < .01*
Age * Verb Bias	F1(2,84) < 1, p > .5 F2(2,18) < 1, p > .5 minF'(2,59) < 1, p > .5	F1(2,84) < 1, p > .4 F2(2,18) < 1, p > .3 minF'(2,59) < 1, p > .5	F1(2,84) < 1, p > .5 F2(2,18) < 1, p > .5 minF'(2,59) < 1, p > .5
Prosody	F1(1,84) < 1, p > .4 F2(1,18) < 1, p > .5 minF'(1,59) < 1, p > .5	F1(1,84) = 15.46, p < .001** F2(1,18) = 35.34, p < .001** minF'(1,92) = 10.75, p < .005**	F1(1,84) = 36.81, p < .001** F2(1,18) = 36.63, p < .001** minF'(1,59) = 18.36, p < .001**
Age * Prosody	F1(1,84) = 2.08, p = .153 F2(1,18) = 4.30, p = .053 minF'(1,88) = 1.40, p > .2	F1(1,84) = 7.87, p < .01* F2(1,18) = 9.49, p < .01* minF'(1,67) = 4.30, p < .05*	F1(1,84) < 1, p > .5 F2(1,18) < 1, p > .5 minF'(1,59) < 1, p > .5
Prosody * Bias	F1(2,84) < 1, p > .5 F2(2,18) < 1, p > .4 minF'(2,59) < 1, p > .5	F1(2,84) < 1, p > .5 F2(2,18) < 1, p > .4 minF'(2,59) < 1, p > .5	F1(2,84) = 1.47, p > .2 F2(2,18) = 1.28, p > .3 minF'(1,54) < 1, p > .4
Age * Verb Bias * Prosody	F1(2,84) < 1, p > .3 F2(2,18) = 1.90, p > .15 minF'(2,85) < 1, p > .4	F1(2,84) < 1, p > .4 F2(2,18) = 1.13, p > .3 minF'(2,64) < 1, p > .4	F1(2,84) < 1, p > .5 F2(2,18) < 1, p > .5 minF'(2,59) < 1, p > .5

Table 12: Analysis of trials in which participants were fixating on the Target Animal prior to the prepositional object onset for Experiment 2. The dependant variable is the proportion of looking time to the Target Instrument.

	Early PP-Object	Late PP-Object
Adults: Prosody	F1(1,29) = 32.92, p < .001** F2(1,14) = 15.24, p < .005** minF'(1,27)=10.42, p <.005**	F1(1,29) = 7.91, p < .01* F2(1,14) = 17.84, p < .001** minF'(1,43)=5.48, p <.05**
Children: Prosody	F1(1,46) < 1, p > .5 F2(1,16) = 1.5, p > .2 minF'(1,58) < 1, p >.4	F1(1,42) = 4.96, p < .05* F2(1,16) = 10.14, p < .01* minF'(1,60) = 3.33, p =.073



Table 13: Analysis of the actions for each prosody condition of Experiment 2 in comparison with prosodically-neutral versions of the same sentences. The data for the neutral baseline is taken Snedeker and Trueswell (2004). The dependant variable is the proportion of instrument actions.

	Instrument Prosody compared to Neutral Baseline	Modifier Prosody compared to Neutral Baseline
<i>Adults</i>		
Prosody	F1(1,30) = 5.67, p < .05* F2(1,21) = 15.37, p < .001** minF'(1,47)=4.14, p < .05*	F1(1,30) = 4.08, p = .05 F2(1,21) = 10.85, p < .005** minF'(1,47)=2.97, p = .09
Verb Bias	F1(2,30) = 23.01, p < .001** F2(2,21) = 30.00, p < .001** minF'(2,51)=13.02., p < .001**	F1(2,30) = 18.26, p < .001** F2(2,21) = 56.90, p < .001** minF'(2,46)=13.82., p < .001**
Prosody * Bias	F1(2,30) < 1, p > .5 F2(2,21) < 1, p > .5 minF'(2,48) < 1, p > .5	F1(2,30) < 1, p > .5 F2(2,21) = 1.90, p = .19 minF'(2,33) < 1, p > .5
<i>Children</i>		
Prosody	F1(1,48) = 7.92, p < .01* F2(1,21) = 15.07, p < .001** minF'(1,68)=5.19, p < .05*	F1(1,48) = 4.38, p < .05* F2(1,21) = 6.90, p < .05* minF'(1,67)=2.68, p = .11
Verb Bias	F1(2,48) = 40.80, p < .001** F2(2,21) = 60.09, p < .001** minF'(2,66)=24.30., p < .001**	F1(2,48) = 18.26, p < .001** F2(2,21) = 56.90, p < .001** minF'(2,46)=13.82., p < .001**
Prosody * Bias	F1(2,48) = 1.38, p > .2 F2(2,21) = 2.86, p = .09 minF'(2,69) < 1, p > .3	F1(2,48) = 1.22, p > .3 F2(2,21) = 2.86, p = .14 minF'(2,68) < 1, p > .4

Table 14: Effects of Demonstration on maternal speech in the referential communication task

Dependent Variable	Instrument Prosody	Modifier Prosody	95% CI for within SS difference	Analysis
verb length	423 ms	408 ms	± 36 ms	F1 (1,4) < 1, p > .5 F2 (1,11) < 1, p > .5 minF'(1,12) < 1, p > .5
verb pause	66 ms	135 ms	± 59 ms	F1 (1,4) = 3.85, p = .12 F2 (1,11) = 3.87, p = .075 minF'(1,12) = 1.93, p > .15
direct-object noun	499 ms	343 ms	± 54 ms	F1 (1,4) = 23.01, p < .01* F2 (1,11) = 76.40, p < .001** minF'(1,7) = 17.68, p < .005**
noun pause	260 ms	30 ms	± 32 ms	F1 (1,4) = 30.22, p < .005** F2 (1,11) = 42.18, p < .001** minF'(1,10) = 17.60, p < .005**
prepositional object	519 ms	444 ms	± 43 ms	F1 (1,4) = 8.27, p < .05* F2 (1,11) = 6.63, p < .05* minF'(1,13) = 3.68, p = .08

Figure 1: Example of the referential context used in the experiments (for the target sentence "You can feel the frog with the feather").

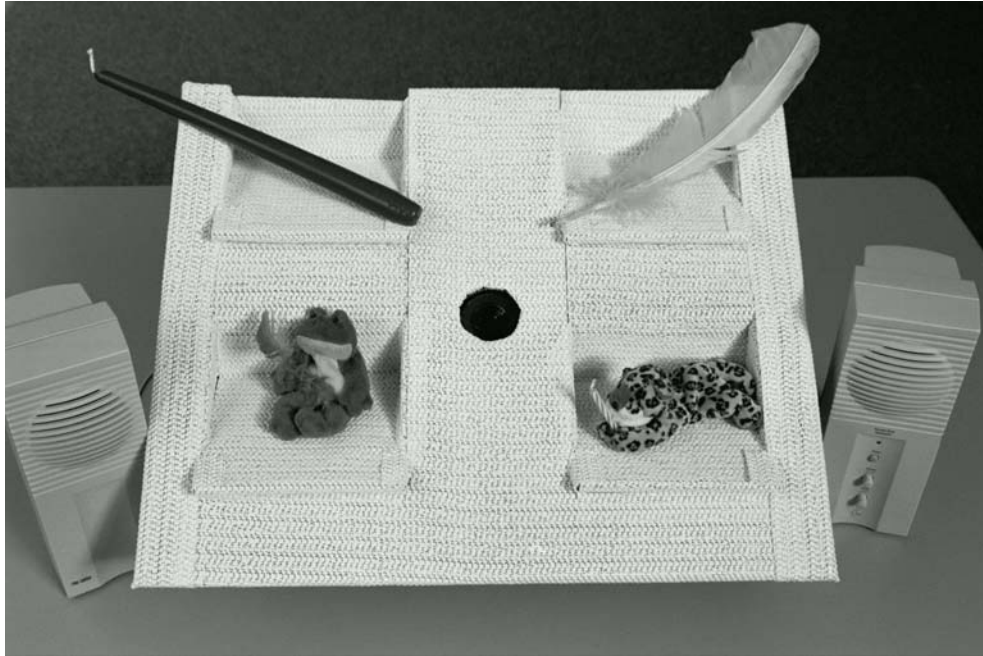


Figure 2: Time course for the critical utterances in Experiment 1.

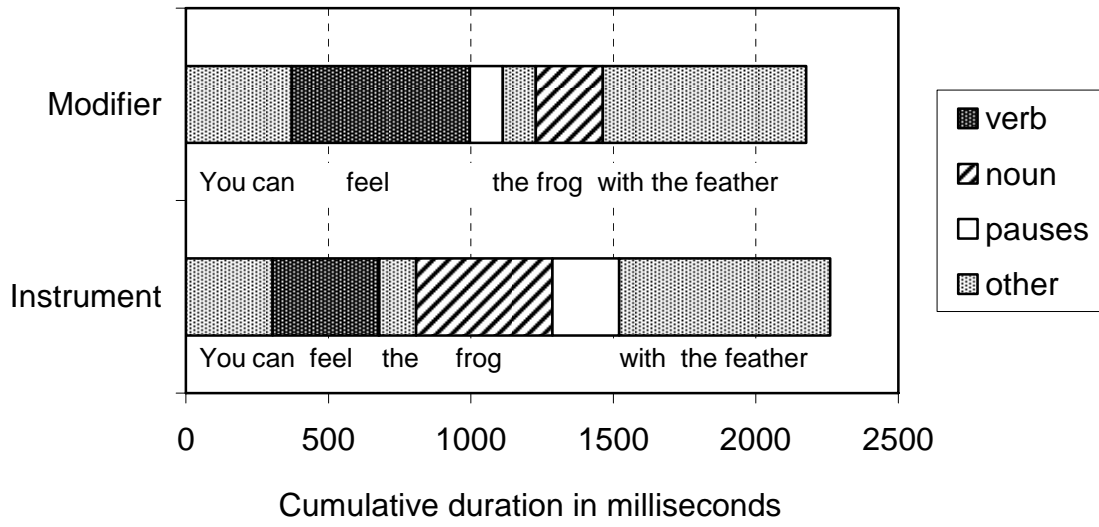


Figure 3: Action analysis. The proportion of instrument responses from children in Experiment 1 by trial block.

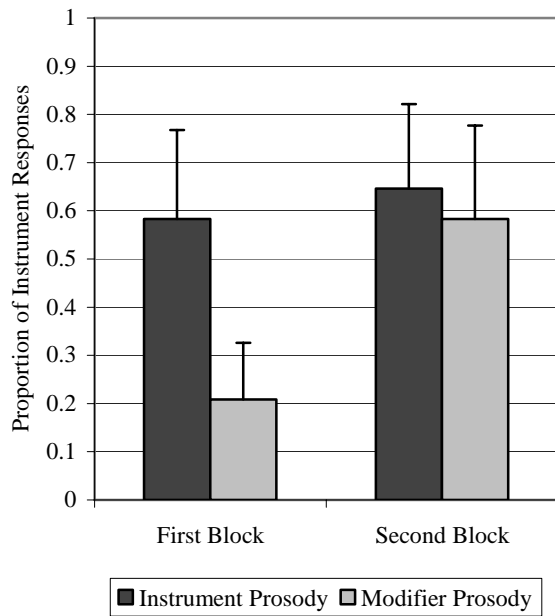


Figure 3: Coarse grained analysis of fixations. The proportion of trials with looks to the Target Instrument for children in Experiment 1 by trial block. The dashed line indicates the proportion of trials with looks to the Distractor Instrument across the four conditions.

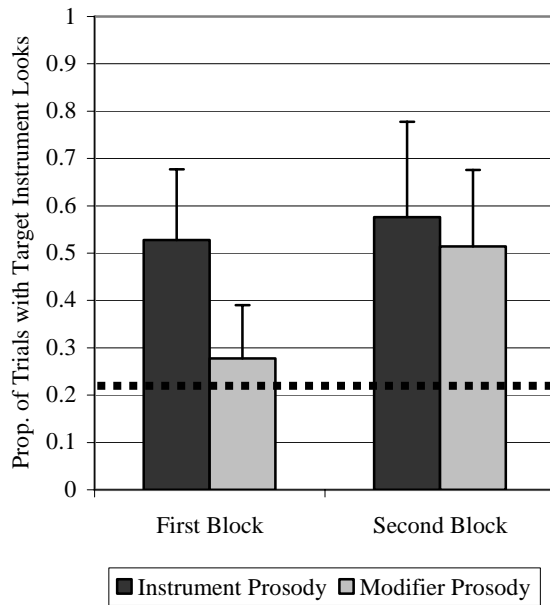
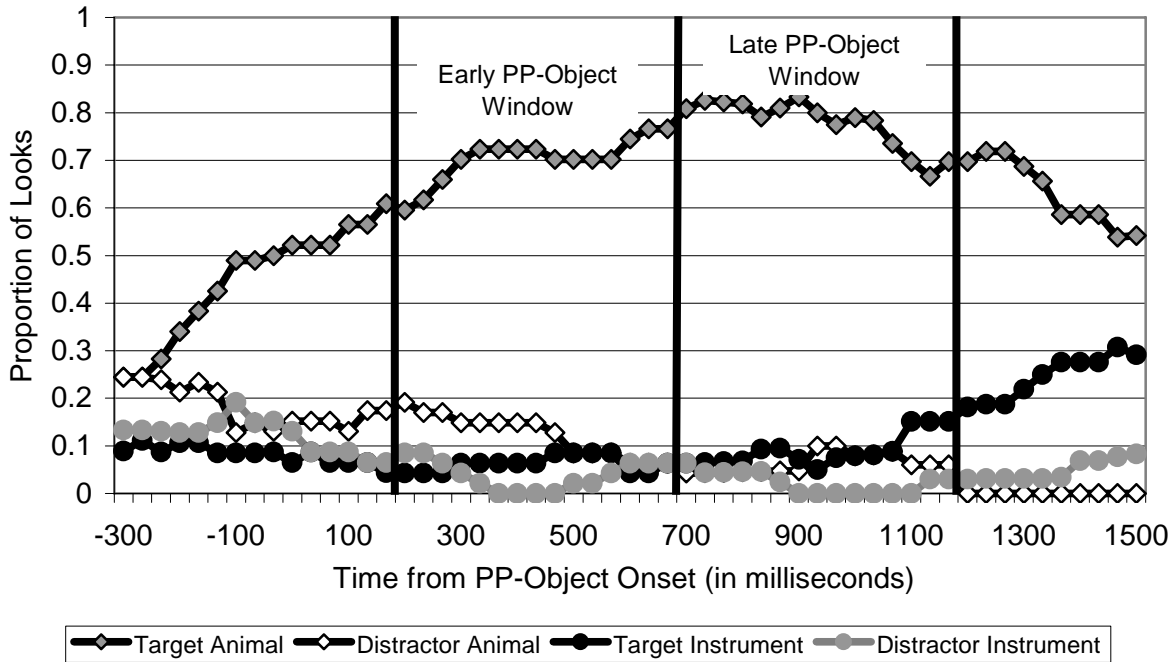


Figure 5: Fixation probabilities relative to the onset of the prepositional object for children in Experiment 1.

a. Block 1 Modifier Prosody



b. Block 1 Instrument Prosody

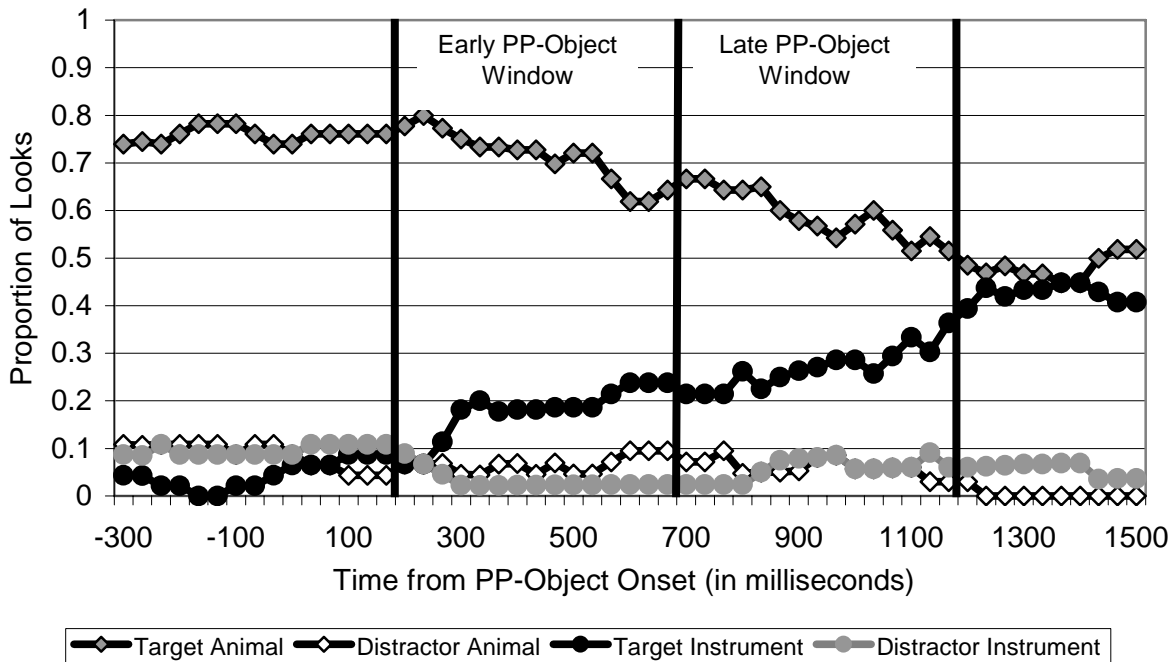
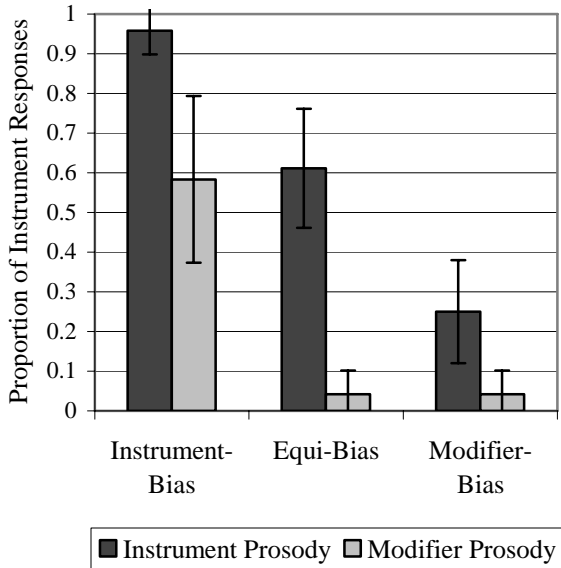
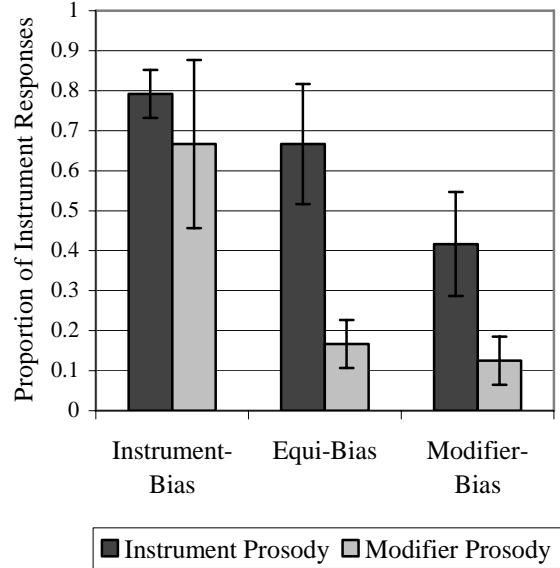


Figure 6: Action analysis. Proportion of instrument responses in Experiment 2 by age, bias class and trial block.

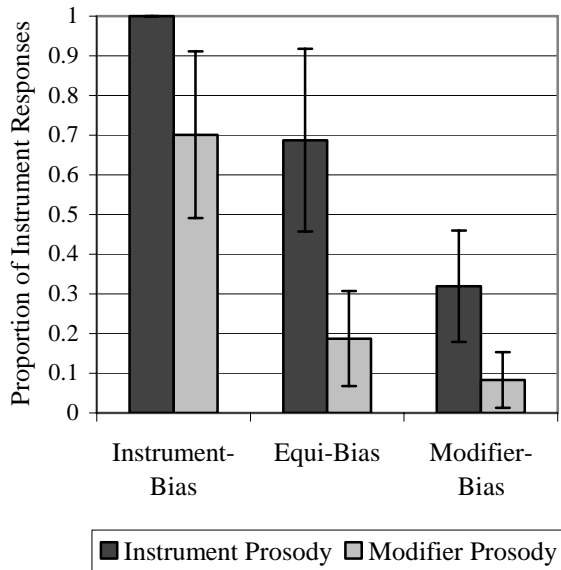
a. Adults Block 1



b. Adults Block 2



c. Children Block 1



d. Children Block 2

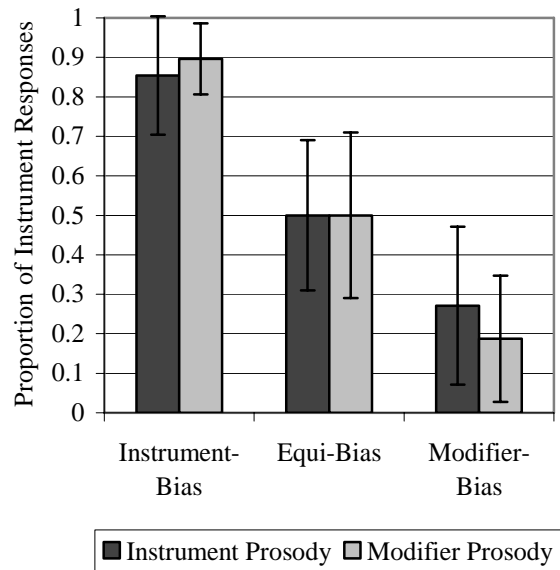
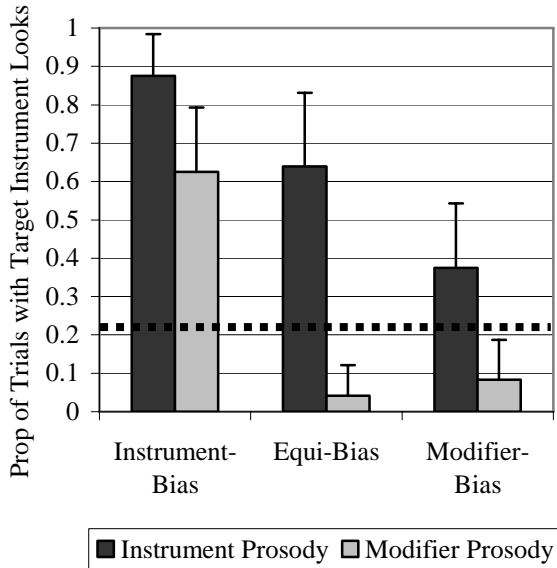


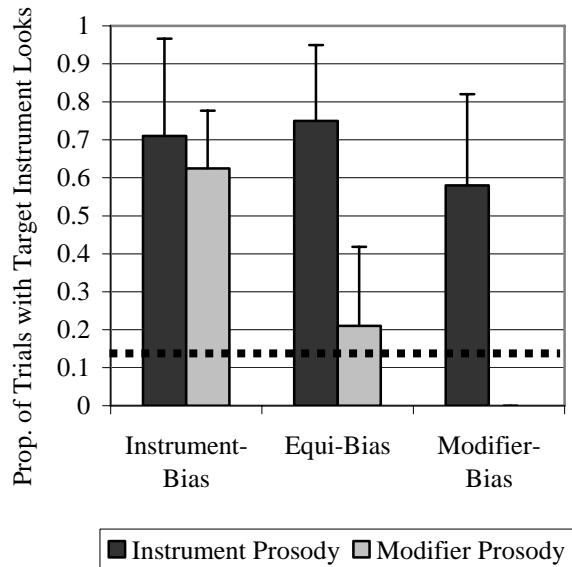


Figure 7: Coarse grained analysis of fixations. The proportion of trials with looks to the Target Instrument for children in Experiment 1 by age, bias class and trial block. The dashed line indicates the proportion of trials with looks to the Distractor Instrument across the six conditions in each panel.

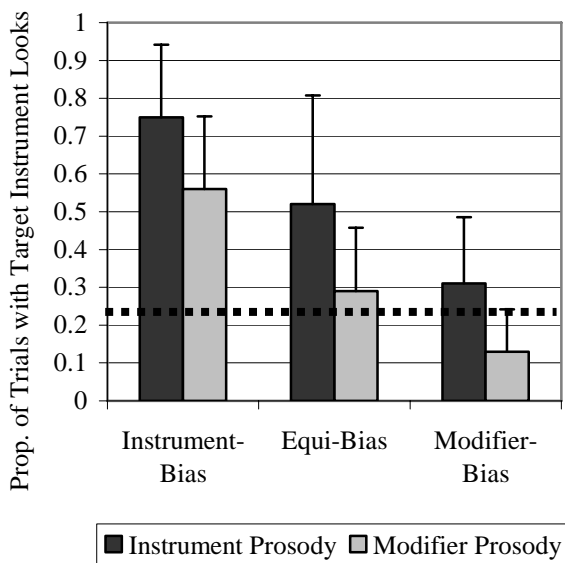
a. Adults Block 1



b. Adults Block 2



c. Children Block 1



d. Children Block 2

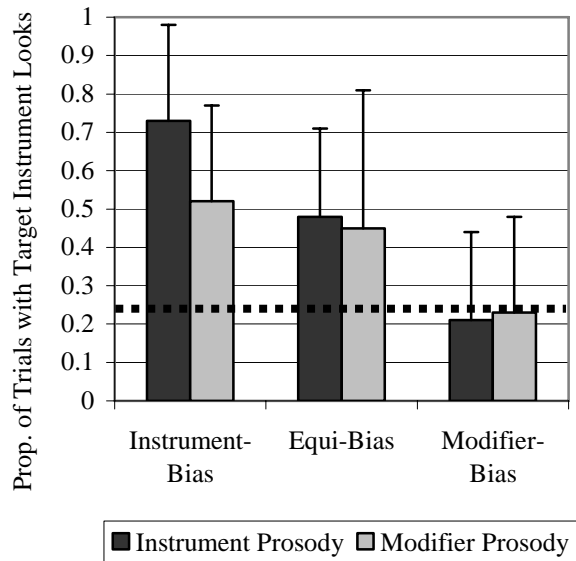
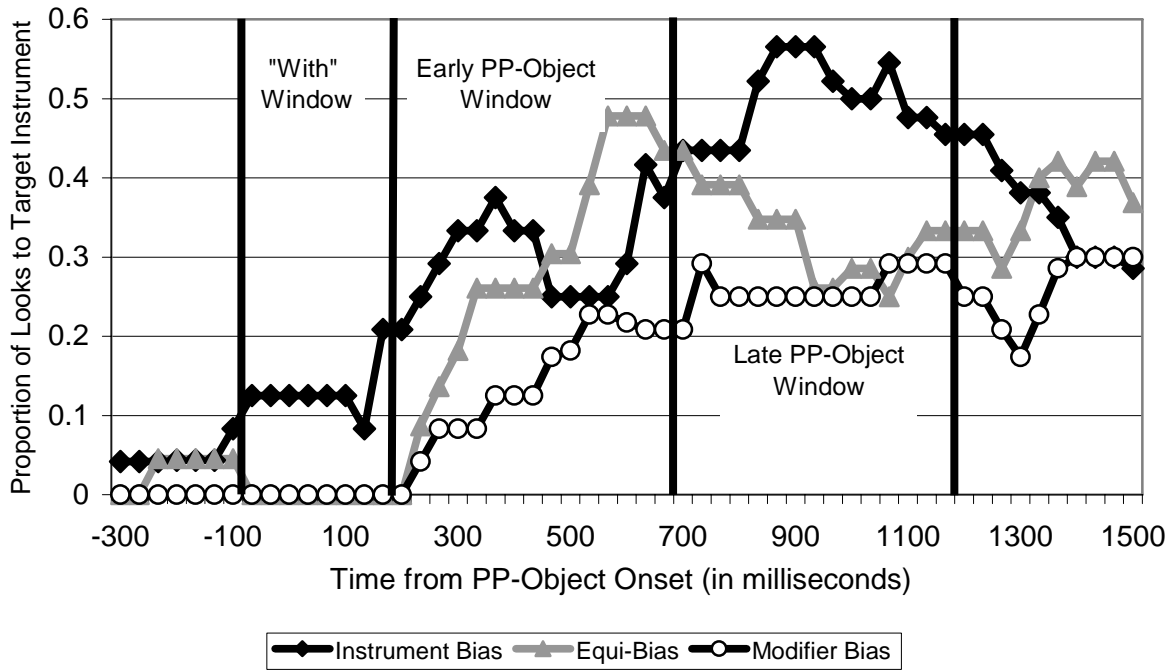


Figure 8: Temporal analysis of fixations to the Target Instrument for the adults in Experiment 2 for: a) the Instrument Prosody conditions and b) the Modifier Prosody conditions (first block).

a. Instrument Prosody



b. Modifier Prosody

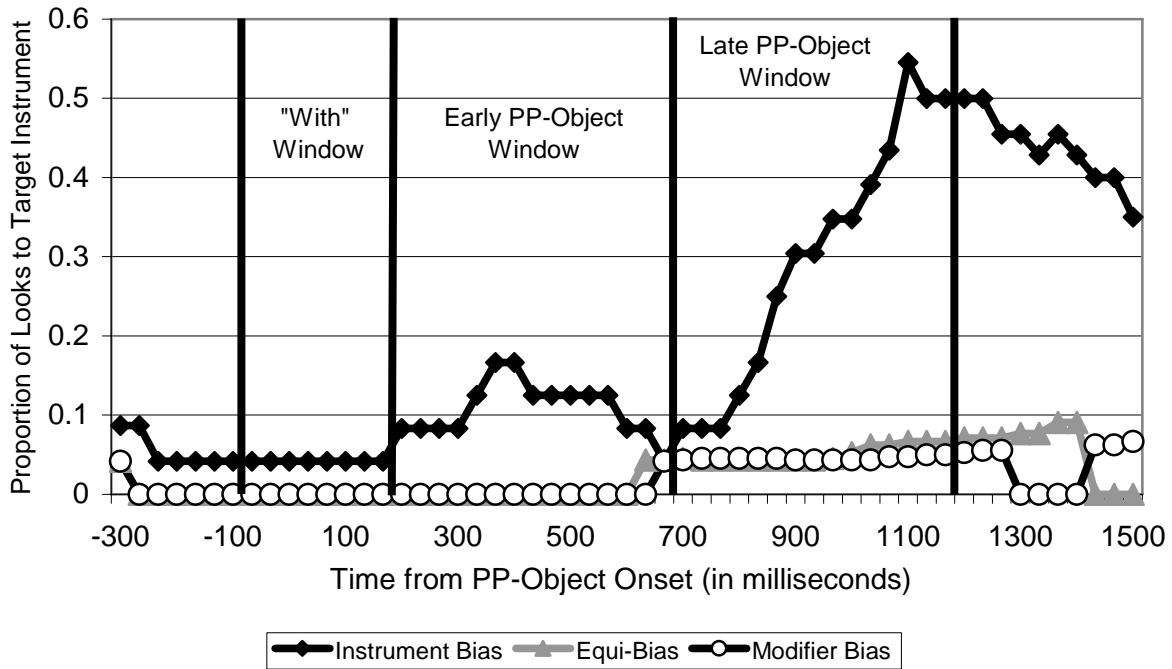
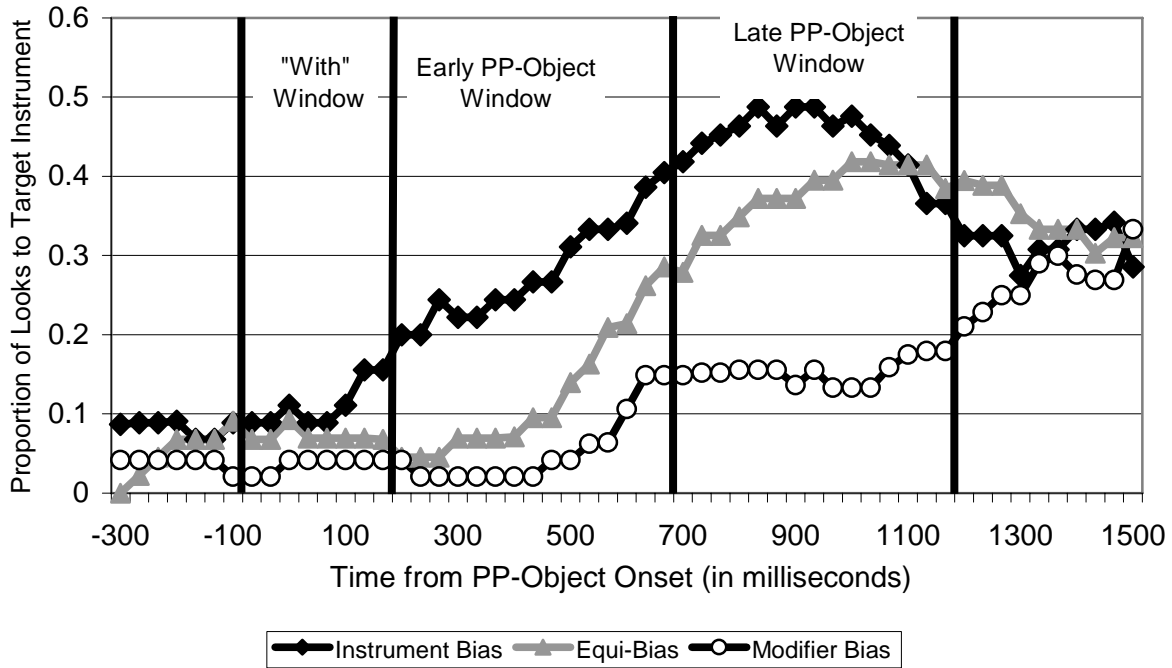
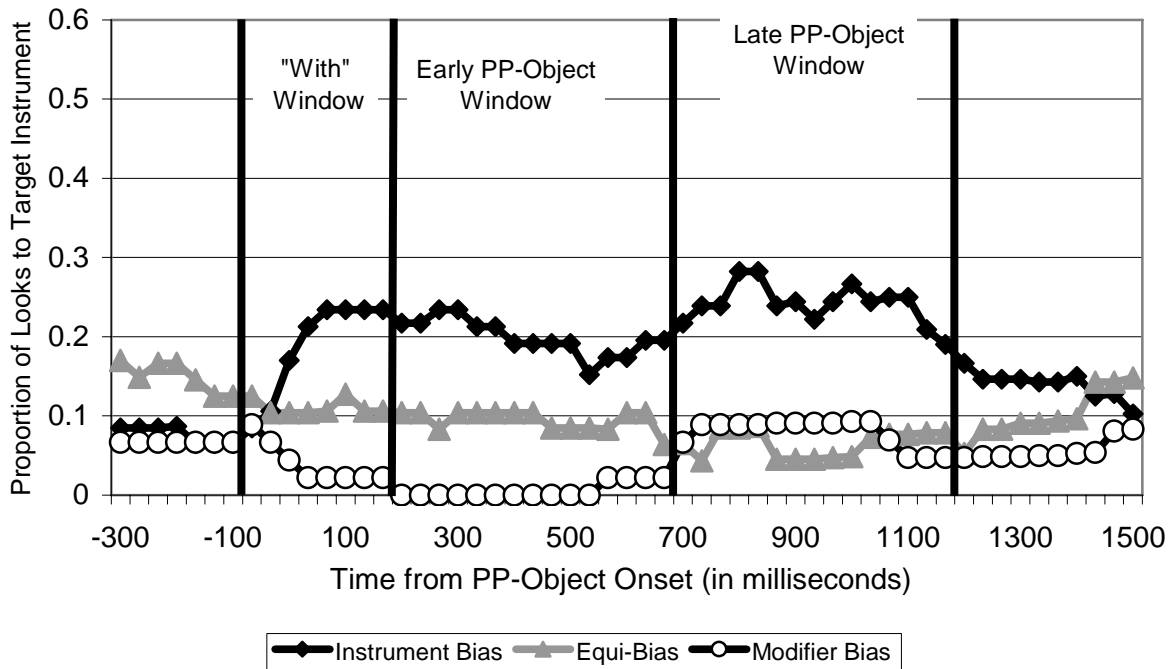


Figure 9: Temporal analysis of fixations to the Target Instrument for the children in Experiment 2 for: a) the Instrument Prosody conditions and b) Modifier Prosody conditions (first block).

a. Instrument Prosody



b. Modifier Prosody



---