Referential and Syntactic Processes: What develops?

John C. Trueswell, Anna Papafragou and Youngon Choi
University of Pennsylvania

Send correspondence to:
John C. Trueswell
Department of Psychology
University of Pennsylvania
3815 Walnut Street
Philadelphia, PA 19104-6196

Referential and Syntactic Processes: What develops?

1. Introduction

A primary purpose of language is to permit individuals to communicate their perceptions and conceptions of the world. The linguistic system that underlies this communication must therefore be designed for fairly intricate interactions with the human perceptual and conceptual machinery. The study of sentence comprehension abilities in adults shows quite clearly that this is the case. For instance, it has been found that the recognition of a word includes detailed linguistic information about how that word is likely to combine syntactically and semantically with the current representation of the sentence. In addition, the referential implications of these analyses are computed in real-time and appear to exert a simultaneous influence on the ongoing structural analyses, allowing the listener to pursue referentially plausible analyses and exclude referentially implausible ones. This rapid 'dance' between syntactic, semantic and referential factors over the course of interpreting a sentence has led researchers to the conclusion that the recognition of a word in a sentence exerts immediate effects on multiple tiers of linguistic and nonlinguistic representation, i.e., *phonological, syntactic, semantic, and referential.* These representational systems, though distinct, mutually constrain each other in a dynamic fashion (e.g., Jackendoff, 2002; MacDonald, Pearlmutter & Seidenberg, 1994; Trueswell & Tanenhaus, 1994).

In this chapter, we explore the question of how this sentence processing machinery develops in children. The study of sentence processing development has attracted renewed interest with the advent of a methodology for studying children's language comprehension abilities in real-time. In these studies, children's eye gaze patterns to objects in the world are recorded as they hear spoken utterances about this world, with the measure providing a moment-by-moment window into their interpretation process. We will suggest from these and other data that the system is incremental and interactive at a relatively early stage in development, showing sensitivity to a variety of constraints on computing sentential meaning. But at the same time we will argue that there are systematic changes over developmental time in the reliance on certain sources of linguistic and nonlinguistic evidence depending on the validity and reliability of this evidence in the learner's past experience. The dynamic abilities of this processing system itself are also found to change and mature in time; this interface system, like many others, is subject to developmental changes in information processing control, especially changes in selectional and attentional abilities.

In order to begin our discussion of this theory of parsing development, we must first discuss what is known about the adult end-state, that is, what we believe to be true about the sentence comprehension abilities of adults and how *they* use multiple constraints to shape their online structuring of the input. With this account in hand, we will use it to motivate a developmental theory of sentence processing. We will then turn to experimental evidence that we believe justifies our claims.

1.1 Real-time sentence processing in adults

Let us first consider what are likely to be inescapable truths about sentence comprehension in adults. First, given the way natural languages work, it is almost

certainly the case that listeners must recover much or all of the intended syntactic structure of an utterance. This is because the structural characteristics of an utterance, when combined with the semantics of verbs and other lexical items, convey essential role assignments (i.e., who-is-doing-what-to-whom). This of course has been the bread-and-butter of not only linguistics but also most psycholinguistic research on sentence processing carried out over the last thirty years. Less often discussed, but arguably equally important, is the fact that the structure of an utterance simultaneously conveys intended discourse operations (e.g., discourse status, focusing). Grammatical choices made by a speaker (whether to passivize, whether to include a restrictive modifier, etc.) reflect discourse considerations and are designed to communicate what the speaker is referring to. Listeners therefore need this syntactic information to infer the intended meaning of an utterance and its reference to the world. Thus listeners must look for evidence in the linguistic input about the syntactic operations that gave rise to the utterance.

Exactly how syntactic and semantic structure is recovered by a listener/reader has been of course a central topic of some debate (e.g., Altmann & Steedman, 1988; Frazier & Fodor, 1978; Frazier, 1987; MacDonald et al., 1994; Trueswell & Tanehaus, 1994). In this chapter, we will sketch only our own view because it is used to motivate the particulars of our developmental account. Specifically, we assume that during the comprehension of a sentence, listeners are engaged in the recovery of phonological, syntactic and semantic characterizations of the input, which, importantly, are each maintained within partially independent representational systems (*representational modularity*). These representational systems dynamically constrain each other over time as the sentence unfolds (*dynamic interactive processing*).

These three characterizations of the input (phonological, syntactic and semantic) should be thought of as 'interim' representations whose primary use is to allow listeners to update their mental model of the world (including what they believe speakers are trying to communicate). Importantly, we assert that the recovery of these interim representations is done in real time via probabilistic mechanisms. The process of recognizing a word within a sentence activates probable phonological, syntactic and semantic structures in parallel, including if necessary multiple alternatives within each subsystem. In turn, interface mechanisms act in real-time as the sentence is unfolding to converge on the most consistent and probable solution across these domains (see Trueswell & Tanenhaus, 1994; Kim, Srinivas & Trueswell, 2002).

It follows from this account that the frequency-based accessibility of structural alternatives will play an important role in a comprehender's ability to converge on the intended meaning of an utterance. Perhaps the best evidence for this claim comes from adult studies of temporary ambiguity during reading and listening. For instance, in the fragments below, temporary ambiguities arise as readers/listeners attempt to rapidly structure the input:

1. *The man sliced the loaf with…*
2. *The child believed the doctor…*

In (1) *with* could be linked to the verb *sliced* denoting an instrument (e.g., *with the sharp knife*) or linked to the noun phrase *the loaf*, denoting a modifier (e.g., *with the burnt crust*). In (2) the noun phrase *the doctor* could be structured as the direct object of

believed (e.g., ending the sentence there) or as the start of an embedded sentence that is a complement of the verb (…*believed the doctor was lying.*).

A wealth of experimental findings now exist to suggest that the structural and semantic analyses that comprehenders assign *at the point of ambiguity, mid-sentence,* is determined by detailed lexical factors, including the probability that the given verb takes particular complements, as well as the semantic fit of constituents into the intended roles assigned by the verb (e.g., Britt, 1994; Garnsey, Pearlmutter, Myers & Lotocky, 1997; Trueswell, Tanenhaus & Kello, 1993; Trueswell, Tanenhaus & Garnsey, 1994). For instance, the tendency for the verb *slice* to include an Instrument role in the form of a PP and the tendency for *believe* to include a Patient in the form an NP, predict initial 'parsing' preferences by readers and listeners encountering these phrases (e.g., Garnsey et al., 1997; Taraban & McClelland, 1988).[1] All of this suggests that word recognition processes are often the engine that drives the structuring of input – indeed studies exist which show that covert priming of a verb with different syntactic and semantic tendencies can unconsciously affect comprehenders' parsing preferences for ambiguous phrases (Novick, Kim & Trueswell, 2003; Trueswell & Kim, 1998).

This probabilistic recovery of structure is sensitive to other contingencies as well. In particular, we assert that the referential implications of these interim representations are computed in real-time and can serve as an important top-down constraint on sentence processing. For instance, Altmann and Steedman (1988) found that readers structure phrases like *sliced the loaf with the…* differently depending on the contents of a prior story. As an example, when there are two different loaves, readers prefer to interpret the *with*-PP as a modifier of the preceding NP (*the loaf*). The idea here is that a definite NP must uniquely specify a referent within the current referential domain (Crain, 1980; Crain & Steedman, 1983). If the simple NP *the loaf* fails to do this, further linguistic information is expected in the form of a post-NP modifier. Indeed, indefinite NPs (*a loaf*) alter this parsing preference (Spivey-Knowlton & Sedivy, 1998) and even other referential factors contribute to parsing decisions (Trueswell & Tanenhaus, 1991).

Thus it seems at least from this evidence that comprehenders must also be dynamically tracking what is under discussion and what is within the current referential

[1] The explanation of lexically-specific effects on parsing is more complicated than this. For instance, one could alternatively say that listeners (and language learners) use known semantic properties of the verb *slice* to compute the likelihood of an instrument phrase, rather than saying that listeners/learners track the syntactic contingencies given the word itself. It is very difficult to disentangle these two explanations since semantic and syntactic properties of verbs are so intimately related (e.g., Fisher, Gleitman & Gleitman, 1992). Also, one might argue that an individual's compiled 'corpus' of their native tongue is too small to offer lexically-based biases. However, even a conservative estimate of language exposure (2/words a second for 5 hours/day) yields 13.1 million words per year (see Kelly & Martin, 1994, for a similar estimate). This estimate is clearly conservative when one considers that television exposure alone is an average 3.6 hours/day by age 3 for American children (Christakis, Zimmerman, DiGiuseppe & McCarty, 2004). Nevertheless, it is an obvious necessity for any theory of language comprehension to include semantic-contingent predictors of structures given the way that language works creatively. Another way of stating this is that the learned semantic classes and properties of known words are also used to constrain parsing and are especially important when a word is uncommon. That is, the parser must 'back-off' to larger categories when subcategories have a small N and even infer a likely category when the word is novel to the language user (i.e., N=1). For further discussion and experimentation related to this issue see Juliano and Tanenhaus, (1994), Naigles, Gleitman & Gleitman, 1993, Naigles, Fowler, and Helms (1992, 1995), Trueswell, Tanenhaus and Kello (1993), Trueswell and Gleitman (in press).

domain, since these factors rapidly influence the structuring of the input.[2] Importantly, however, reading studies of this sort also indicate that the effectiveness of this contextual factor depends upon the availability of the structural options at issue. For instance, in the PP ambiguity above in (1), the effectiveness of the two-loaf story in supporting a modifier interpretation depends upon the kind of verb that is used in the stimuli: verbs that frequently include an Instrument role show substantially delayed and reduced contextual effects (see Britt, 1994; Spivey-Knowlton & Sedivy, 1994; see also Garnsey et al., 1997, and Trueswell, 1996, for lexically determined accessibility in other structures).

The picture emerging from these data is one in which the recognition of the word within a sentence automatically triggers linguistic representations at multiple levels. This triggering though is probabilistic in nature: given the evidence at hand, a listener is engaged in a 'guessing-game' in which the linguistic procedures that gave rise to the utterance are recovered. The referential implications of these representations are also computed in real-time and, when possible, used to constrain the representational hypotheses that the listener is considering.

Finally, it cannot be emphasized enough that local ambiguity is pervasive, *the norm*, in real-time language comprehension. Indeed, computational linguists have recognized the pervasiveness of ambiguity, especially once they started to implement 'wide-coverage' parsers and interpreters that were designed to handle naturally produced text (e.g., Marcus, Santorini & Marcinkiewicz, 1993). It has even been claimed that local ambiguity of the sort found in highly-lexicalized formalisms can actually provide a processing advantage because it permits greater flexibility in recovering structure and meaning (Steedman, 2000; Srinivas & Joshi, 1999; see also Kim, Srinivas & Trueswell, 2002). The implication here is that the accessibility of structure is an unavoidable issue in the study of language comprehension, whether one is interested in the issue of syntax, semantics or reference. We strongly suspect the same is true for the study of language comprehension in children.

1.2 A developmental account of sentence comprehension

Given this picture of the adult system, let us now turn to our proposal of how the child learns to implement these dynamic sentence-processing abilities, focusing especially on how referential contingencies are learned and used. To begin we must spell out some basic assumptions of this account. First, we will assume a great deal of *processing continuity* over development. That is, the types of processes used for language comprehension remain constant throughout language learning and into adulthood. In particular we will assume:

1. **Real-time processing continuity:** From the outset, a language learner/listener is attempting real-time, incremental processing of the input speech stream.
2. **Probabilistic processing continuity:** From the outset, the detection from the speech stream of already acquired linguistic elements (including syntactic and

[2] Research suggests that adults, when engaged in conversations and/or other related tasks, coordinate and track the referential domain of utterances, especially when this domain is determined by such factors as common knowledge/ground. Debates exist however about the timing of these effects (e.g., Hanna et al., 2003; Keysar et al., 2000; Keysar et al., 1998).

phrasal elements) is achieved via probabilistic pattern-recognition and pattern-completion processes.

There is good reason to believe these assumptions hold, especially as they pertain to the processing of sub-lexical and lexical elements. In particular, experimental results from Aslin, Newport and Saffran indicate that 8 to 12 months olds are sensitive to the distributional properties of syllables, allowing them to discover likely lexical / morphological candidates from continuous speech (e.g., Aslin, Saffran & Newport, 1998; Saffran, 2001, 2002; Saffran, Aslin & Newport, 1996). These results suggest that from the outset language learners are attempting to extract potentially relevant linguistic elements from the input via probabilistic mechanisms. In turn, these elements serve as candidates for word learning. That is, language learners attempt to map these newly discovered elements onto known conceptual representations.

As language learners build up this repository of word-meaning pairs, it appears that they are faced with temporary ambiguity from the start and that they deal with this ambiguity in an adult-like manner, i.e., in real time, as the speech unfolds. For instance, the eyetracking research of Fernald, Swingley and colleagues shows that 18-to-24-month-olds process phonological word cohorts (*dog/doll; tree/truck*) in much the same way as adults, with the major difference being that adults know more words (Swingley, Pinto & Fernald, 1999; Allopenna, Magnuson & Tanenhaus, 1998). Upon hearing *doll* in a sentence like *Look at the doll,* 18-24 months will look to cohort referents of *doll,* such as a picture of a dog, but not to non-cohort referents such a picture of a mouse. Suggestive evidence also exists indicating that children in this age range are even beginning to engage in real-time syntactic/semantic structuring of these utterances. For instance, 27-month-olds hearing *Let's roll the ball* restrict looks to a ball over a non-rollable object in view, as the word *ball* is being heard (Fernald, 2001). All of these patterns have been observed in real-time studies of adult listeners (e.g., Allopenna et al., 1998; Altmann & Kamide, 1999) suggesting considerable continuity over development in dynamic processing abilities.

The question of interest here though is whether child sentence processing abilities show the same sort of processing continuity over development, which is exactly what we wish to assert. One obvious obstacle, of course, is that sentence comprehension is arguably orders of magnitude more complex than word recognition. We therefore must make the following further assumptions about language processing development.

3. **Representational modularity.** The language processing system is innately predisposed to organize linguistic input into three partially independent representational domains: phonological, syntactic and semantic.
4. **Representational interfacing.** The language learner expects systematic correspondences between these representational systems. For instance, the number and type of phrasal constituents present in an utterance will have a systematic mapping onto the number and type of participants denoted in the conceptual representation of an event (Gleitman, 1990; Gleitman, Cassidy, Nappa, Papafragou & Trueswell, in press).
5. **Interactive processing:** From the outset, the language system is detecting and taking advantage of probabilistic tendencies between phonological, syntactic and

semantic elements computed from the input stream so as to constrain possible analyses of this input.

6. **Assume reference.** The language learner is innately predisposed to assume that communicative acts refer to the world. Hence the referential implications of interim linguistic characterizations of speech input are attempted from the outset.

These assumptions, when combined with what we believe are properties of the adult comprehension system, allow us to derive some predictions about how child-listeners ought to resolve temporary syntactic ambiguity during sentence comprehension. First, like the adult system, the child sentence comprehension system is engaged in the recovery of known syntactic and phrasal categories from the input, which is accomplished via pattern recognition processes. These higher-order syntactic and phrasal elements are likely to be discovered via distributional/statistical mechanisms similar to those proposed for lexical discovery by Saffran, Newport and colleagues (i.e., Mintz, Newport & Bever, 2002, cf. Harris, 1957; see also Gerken, 2002; Gómez, 2002; Gómez & Gerken, 2000). Crucially though we assert that certain sorts of categories are preferred by the linguistic processing system and are assumed to map onto semantic and conceptual representations in systematic ways (assumptions 3 and 4 above).

Once a repository of syntactic representations has been learned, we would expect a processing situation somewhat similar to the one characterized in early lexical processing (and documented by Swingley and colleagues): namely the child parsing system must also deal with syntactic ambiguities and must resolve these ambiguities in real-time. Since the adult syntactic parsing system is a probabilistic device that weighs multiple contingencies, it follows that the child processing, though organized and operating in the same way, must *gradually discover and learn these contingencies.*

For illustration, consider the following phrasal category, the Prepositional Phrase (PP). The PP in English is associated with a range of semantic functions, often ambiguously for any particular lexical head. PPs can be used for temporal and spatial specification of events or entities (e.g., *sang on Tuesday, sang on the stage, the book on your left*). PPs are also used as arguments of events and entities (e.g., *given to Susan, put on the table, governor of California*).

Consider a child who has *already* learned that PPs serve these semantic functions.[3] As the child detects a particular instance of a PP in the input stream (e.g., *I really like your doll <u>with the</u>…*), how does he/she decide which of these semantic (and syntactic) operations to compute? Based on the adult literature outlined above, we can list the possible sources of evidence:

- Lexical head (e.g., *with, on, in*) and lexical semantics/syntax
- Syntactic structure up to that point (e.g., verb-syntactic-projections)
- Semantic structure up to that point (e.g., verb-thematic-projections)
- Referential operations up to that point (e.g., currently insufficient referential specificity)

[3] This of course is a *miracle* in our theory as laid out here. See Gleitman et al. (in press) for a theory of learning 'hard words,' of which we include *on, in, under, of, within,* etc.

- Prosodic structure up to that point (e.g., currently open intonational phrase)

First, the presence of a particular lexical head (*with, on, of*) constrains the range of syntactic and semantic alternatives. But this is only probabilistic evidence for the listener. For instance, *with* is an ambiguous word that denotes, roughly, either 'accompaniment' or 'instrumental-manner,' and is associated with several syntactic operations (see Figure 1). Like any ambiguous word, the probability of activating any of these linguistic characterizations will depend upon past history with the word (i.e., its dominant and subordinate meanings). Second, the current syntactic structure and semantic structure (points 2 and 3) constrain alternatives probabilistically as well: *Let's color with…* requires VP-attachment but either as an accompaniment (*Let's color with your friends*) or as an instrument (*Let's color with your crayons*). And like the examples above the presence of a NP object permits other syntactic-semantic options (*Let's color the book with the torn cover*). The probability of these alternatives is determined largely by the subcategorization and thematic tendencies of the verb, again as discussed above. Third, referential implications up to that point will play a role in resolving a particular PP occurrence, since as discussed above the lack of referential success of a definite NP supports modification (e.g., *the book with…* when there is more than one book under discussion). Finally, although not discussed above, prosody serves as probabilistic evidence, since the presence or absence of a major prosodic break has a correspondence to phrasal breaks—though again probabilistically.

Insert Figure 1 about here

Now for the developmental question. Given these possible constraints on parsing and interpreting PPs, which are going to be more valid predictors of semantic/syntactic choice, and hence appear to dominate child parsing and interpretation processes? If the adult sentence-processing literature is any guide, we should expect that lexical constraints on structural analyses (points 1-3) will play an early and potent role developmentally. Adults track subcategorization and thematic preferences to such a great extent that they immediately constrain parsing options. If children build such databases as they learn words, it follows that this information will appear as an early determinant of child parsing.

Beyond the adult literature, there is additional evidence for this conclusion. Research on verb learning strongly suggests that children track the number and types of phrases that occur with verbs *so as to assist in learning the meaning of these verbs* (e.g., Fisher, Hall, Rakowitz & Gleitman, 1994; Gillette, Gleitman, Gleitman & Lederer, 1999; Gleitman, 1990). Said another way, children from an early age track subcategorization and argument-taking properties of verbs as they learn them. Our assertion here is that this probabilistic evidence, which was tracked and developed so as to discover the meanings of verbs, isn't 'thrown away' after the verb is learned. Rather it is used to recognize the intended structure of an utterance every time that particular verb is

encountered again later in life. Moreover, children like adults deploy this knowledge-of-probabilities *on the fly* as a sentence unfolds.

What about potential top-down influences of referential implications (point 4)? Will children show similar early influence of such factors? Given the assumptions sketched above, the answer should be 'yes', as long as the particular contextual evidence is valid, accurately computed by the child, and reliably constrains the structural analysis. After all, we argue that it is a great advantage for a listener (adult, child or infant) to discover the referential conditions under which instances of particular linguistic elements are occurring. However, we have reason to believe from the adult parsing literature mentioned above that the referential constraints on PP modification are substantially weaker than relevant lexical predictors of these same structures (e.g., Britt, 1994). Moreover, it seems reasonable to assume that the ability to track what is currently relevant in the contextual setting of an utterance could be hindered in children, since this often requires dynamically tracking what the speaker is thinking about (e.g., Clark, 1993). Given children's difficulties in taking other individuals' perspectives, as studied in the theory of mind literature, it follows that the referential evidence children build for the purposes of sentence comprehension is going to be noisy and even contaminated. As a result, it should be expected, somewhat counter-intuitively, that lexical-syntactic and lexical-semantic predictors to structure will exert an early and potent influence on child processing, whereas the sorts of contextual factors thus far studied in the adult literature will be developmentally delayed in children's ambiguity resolution abilities.

2. The Kindergarten-Path Effect

Much of the initial impetus for developing this account comes from the results of a child sentence processing study first reported in Trueswell, Sekerina, Hill and Logrip (1999). In that study, 5-year-olds, 8-year-olds and adults were given spoken instructions to move objects around on a table while their eye movements were tracked.[4] Eye position was used to infer ongoing referential commitments that listeners make, which are believed to be derived from provisional syntactic and semantic analyses of the spoken utterance (Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995; Spivey, Tanenhaus, Eberhard & Sedivy, 2002).

On critical trials, participants were presented with one of two types of visual scenes, examples of which are shown in Figure 2.

Insert Figure 2 about here.

For both scene types, the child was given an instruction that was designed to refer to a particular stuffed animal, in this case the frog that was sitting on a napkin, as in:

[4] We realize that the decision to study 5-year-olds skips over much of language learning. And indeed the initial choice of this age range was driven largely by methodological limitations (younger children were less willing to wear the eye tracking visor). Nevertheless, several findings exist to suggest that important developmental changes in parsing and interpretation processes extend well through this age range (e.g., Chomsky, 1969) -- findings which have not been adequately studied or explained.

3. *Put the frog that's on the napkin into the box.* (Unambiguous Modifier)[5]

Consider hearing this sentence in the context of the scene on the left (Figure 2a), where there is only one frog but two napkins, one of which is under the frog. If children were simply looking to possible referents based on the individual words they heard, we might expect them to look to the frog upon hearing *frog* but look to either napkin upon hearing *napkin.* However, if referential commitments (and the resulting eye movements) are derived from real-time structural analyses, and children of all ages are engaged in these real-time processes (assumptions 1 and 6 above), we should expect that both children and adults would not consider the empty napkin as a referent, since the word *napkin* is part of a relative clause that must modify the NP *the frog.* Indeed, this latter pattern was exactly what was found: children and adults looked to the frog within a few hundred milliseconds of hearing *frog,* and continued looking at the frog (and the napkin under it) when hearing *napkin,* rarely if ever looking over to the empty napkin.

Notice however that hearing sentence (3) in the context of the scene on the right (Figure 2b) the structural position of *the frog* in the sentence permits reference to either frog, and it is only upon hearing *that's on the napkin* that a listener could compute the correct referent. Indeed, eye movements in two-frog scenes supported this expectation. Upon hearing *frog* listeners, regardless of age, launched an eye movement to a frog, but they were at chance as to which frog they looked at. This state of affairs remained until hearing *napkin* upon which participants shifted gaze to the frog on the napkin if they were at that moment looking at the wrong frog. Again, participants rarely if ever considered the empty napkin. It should also be noted that participants were essentially flawless at carrying out the instruction, regardless of age: they moved the intended frog into the empty box.

Thus the data from unambiguous trials provides compelling evidence that syntactic and semantic analyses are engaged in real time for the ages we've looked at here, and that the referential implications of these structural analyses are reflected in eye movements.

We suppose that the pattern could in principle be explained as children engaging in some sort of intersecting set strategy (hearing *frog* causes look to frogs, hearing *napkin* causes looks to frog-plus-napkin). However data from another variant of the instruction rules out this possibility and confirms other assumptions/hypotheses sketched above. In particular, on certain trials participants heard instructions like (4).

4. *Put the frog on the napkin into the box.* (Temporary Ambiguity)

Here the PP *on the napkin* becomes temporarily ambiguous. It could be interpreted as a modifier of the NP *the frog* or it could be interpreted as a Goal for the verb *put*, i.e., where to put the frog.

If participants are engaged in real-time probabilistic estimation of the intended structure regardless of age (assumption 2) and they use lexical evidence to compute these estimations (as hypothesized above), we should expect a strong initial preference to

[5] Trueswell et al. (1999) used *in the box* rather than *into the box.* The use of the preposition *in* introduces another ambiguity, which we wish to avoid. All follow-up studies have used *into* and have replicated the results reported here (e.g., Hurewitz et al., 2000). For simplicity we adopt the use of *into.*

interpret *on the napkin* as a goal (where to put the frog) rather than a modifier. This is because the semantics of the verb *put* requires a goal role, and the goal role is typically realized in the form of a PP headed by *on* or *in.* Even five-year-olds are expected to have ample linguistic experience to know these structural tendencies for this verb.[6] If in fact the NP *the napkin* is interpreted as being part of a goal PP, the empty napkin in the scene now becomes a possible referent, since one could move a frog to it. Thus we should expect looks to the empty napkin upon hearing the word *napkin* in temporarily ambiguous sentences like (4). Indeed, for one-frog scenes (Figure 2a) looks to the frog were immediately followed by looks to the empty napkin upon hearing *napkin.* This occurred for all ages, suggesting that the syntactic and semantic tendencies of verbs are used to estimate structure and compute referential implications in real-time.

Two-frog scenes however were designed such that referential implications of hearing *the frog* actually discouraged interpreting *on the napkin* as a goal. This is because *the frog* could refer to either frog. If listeners can use this referential analysis in real-time to constrain structuring of further linguistic input, we ought to expect them to prefer to interpret *on the napkin* as a modifier of the NP *the frog* allowing them specify a unique referent. This top-down constraint however must battle against the lexical biases that support the goal analyses of that same phrase. Thus, two-frog scenes provide the potential for a top-down constraint on structuring *on the napkin* that ought to reduce looks to the empty napkin in the scene, rendering the pattern more like what was seen for unambiguous modifiers like *that's on the napkin.*

However, we hypothesized above that less reliable evidence for estimating the structure of the input should result in a developmental delay in using this evidence. We also suggested above that this particular sort of referential constraint on structure is less reliable than lexical constraints in part because referential factors appear to be less effective in adult parsing behaviors. Moreover we suggest that in order to discover this sort of referential contingency children must be fairly skilled at tracking what is currently under discussion (i.e., what is relevant to the speaker). Any difficulty in computing the referential domain of an utterance would contaminate the ability to discover this contingency. Thus, if any developmental progression is expected it would be that younger children would fail to take into account that the two-frog scene supports a modifier analyses.

Indeed, only older children and adults interpreted the ambiguous *on the napkin* as a modifier in 2-frog scenes. That is, they looked randomly to either frog in the scene upon hearing *frog* but rapidly converged on the intended frog upon hearing *napkin*, rarely if ever looking at the empty napkin. In contrast, five-year-olds showed a strong preference to interpret *on the napkin* as a goal, just like in one-frog scenes. Upon hearing *frog,* they were at chance looking at either frog. And upon hearing *napkin* they were still at chance looking at either frog, suggesting they often failed to realize that *on the napkin* could be a modifier of the NP *the frog.* Instead, hearing *napkin* triggered increased looks to the empty napkin just as much as one-frog scenes. Thus, five-year-olds tended to think *on the napkin* was the goal of *put* rather than a modifier, despite the presence of two frogs. This was in striking contrast to how these same children behaved on unambiguous sentences like (3), where they realized the modifier analysis.

---

[6] *Put* is a very common verb in child-directed speech and is frequently accompanied by a Goal PP (see Trueswell et al., 1999, for corpus evidence).

We are fairly confident of our account of five-year-olds' interpretations because of the resulting overt actions that were made by these children. In particular, five-year-olds were so strongly committed to the goal analysis on ambiguous trials that they often carried out actions that involved moving a frog to the empty napkin. On over 60% of the trials, they performed such an action, regardless of the type of referential scene. On unambiguous trials, they were nearly flawless, making errors on about only 5% of trials. Thus the second PP, *into the box,* did not always block the goal analysis of *on the napkin.* It did however for older children and adults: they were nearly flawless on all trials regardless of ambiguity or referential scene, though as expected one-referent scenes did induce some confusion and errors on ambiguous trials, even for adults.

2.1 Questions about our account of the kindergarten-path effect

Although the data presented in Trueswell et al. (1999) are consistent with our account, certain aspects of this account are as of yet inadequately motivated. We articulate these concerns here by posing two questions. The first question pertains to children's insensitivity to the referential scene manipulation:

- Why are two-referent scenes less reliable and less effective at resolving this modifier-argument ambiguity?

We have explained developmental differences in the ability to use 2-referent-scenes by appealing to the poorer reliability of this evidence. In particular, we suggested that this situation (i.e., hearing a definite NP that could refer to multiple visually co-present objects) is not as good a predictor of modifier use as the local lexical evidence (e.g., *put…on*). Moreover, it was suggested that part of the reason for the poor reliability of this referential-scene evidence arises from it being difficult to discover that it can be informative.

However, we have not offered much by way of an explanation for why such a referential setting would be difficult for the child to discover as informative. Moreover, the claim that this information is less reliable than certain lexical information leaves partially unexplained the end-state of this developmental pattern. After all, if this referential situation only weakly predicts the need for linguistic modification, why did Trueswell et al. (1999) and Tanenhaus et al. (1995) find that adults could use the implications of a 2-referent-scene to override the strong lexical bias to interpret *on the napkin* as a goal of *put*? Trueswell et al. (1999) did note that adults had some difficulty with temporary ambiguities in 2-referent scenes, but this difficulty was surprisingly small given our claims.

We address this issue below in Section 3 (*Referential scenes, definite reference and restrictive modifiers*). Specifically we will look to the pragmatics literature on definite reference and discuss findings from a recent referential communication study, all of which suggests that a speaker's choice about linguistic specification (e.g., saying "the frog" vs. "the frog on the napkin") is not strongly determined by the mere presence of multiple identical objects of the relevant type (e.g., multiple frogs). With this knowledge in hand, we will take a second look at the adult listener's ability to use these referential situations to resolve syntactic ambiguity. Here we find that this contextual evidence is not as effective as the original *put* studies might suggest (Snedeker & Trueswell, in press). Moreover, we show that young children, while exquisitely sensitive to experimental manipulations of the lexical biases, are insensitive to the referential scene

manipulation even under conditions of weakly biased lexical evidence. Taken together, the data paint a very reasonable picture about the listener's use of referential situations to constrain the structuring of linguistic input.

The issue of reference and parsing will also be discussed in relation to a recent alternative account of the Trueswell et al. findings provided by Wexler (this volume). Based on earlier data from young children's misuse of the definite determiner *the* in their own productions (Maratsos, 1976; Karmiloff-Smith, 1979), Wexler (2003) proposes that children in the relevant age range lack a complete understanding of the semantics of *the*. In particular, he proposes that children lack the notion of maximality: the requirement that the definite determiner must apply maximally to the current referential domain. In Section 4 below (*Pragmatic vs. Semantic Accounts)* we evaluate this hypothesis with respect to a broader range of data from Maratsos and Karmiloff-Smith and suggest that the real problem lies in a child's understanding of what the referential domain is at any given moment (roughly, the original conclusions of Maratsos and Karmiloff-Smith). Thus Wexler is correct in concluding that the Maratsos and Karmiloff-Smith data shed light on the child parsing studies, though perhaps not in the way he proposes.

A second important question that we wish to address here involves the large difference Trueswell et al. observed between younger children and adults in their final interpretation of these temporary ambiguities.

- Why did younger children fail to revise initial parsing commitments?

Trueswell et al. (1999) offered a developmental reason for why five year olds failed to revise their goal interpretation upon hearing *into the box.* In particular, it was suggested that children's difficulty with revising was the result of limited processing resources or limited working memory, both of which expand with age. That is, maturational differences rather than evidential discovery explained this phenomenon. Since then, we have refined this claim to suggest specifically that this change in revision abilities is related to developmental differences in general executive function processes, specifically the ability to select competing representations (Trueswell & Gleitman, in press; Novick, Trueswell & Thompson-Schill, submitted). In Section 5 (*Revision and lingering garden-paths*), we discuss this proposal. We begin by pointing out that children are not the only ones who sometimes fail to revise their interpretation of a temporarily ambiguous phrase. Christianson, Hollingworth, Halliwel and Ferreira (2001) have found that normal adults can also hold onto beliefs consistent with a rejected interpretation. And importantly, Mendelsohn (2002, 2003) has recently found that individual differences in measures of general inhibition and executive control correlate with this ability to revise parses. We will mention preliminary data of our own lab that suggest these factors play a role in whether adults act like children in the *put* task, making child-like errors in their actions.

We now turn to each of these issues, discussing them in detail.

3. Referential scenes, definite reference, and restrictive modifiers

Given the issues discussed above, a crucial question becomes whether a listener can deduce from scene information alone the need for restrictive modification, specifically, the need for a speaker to provide restrictive modification of a definite NP. Can a person look out into the visual world and anticipate from this information alone a

speaker's need to utter *the little star* (and not *the star*), *the toy closest to you* (and not the toy), or *the frog on the napkin* (and not *the frog*)?

The answer from the pragmatics and psycholinguistics literature on linguistic reference is *no*. For instance, Lyons (1980) sketched a hypothetical situation in which two people were working on a motorcycle and one says "Pass me the spanner" (British English, for wrench). In a situation in which there are two wrenches present, one near the speaker and one near the listener, the listener is likely to 'infer' that the wrench closest to the listener is the one intended. And it seems unlikely that the speaker would have said "Pass the spanner that is closer to you" (see also Lyons, 1999). Similarly, consider the scene in Figure 3, adapted from Stone and Webber (1998). A person could refer to a particular rabbit in this scene by saying "Pull the rabbit out of the hat" but it would be ludicrous to say "Pull the rabbit that's in the hat out of the hat."

Insert Figure 3 about here.

Thus, a referentially ambiguous NP might be disambiguated by the goals of the utterance or the situation in which it is uttered, rather than linguistically via restrictive modification. The referential domain, when applied to a scene, is determined by what is relevant given the goals of the interlocutors. In fact, a recent referential communication study (Brown-Schmidt, Campana & Tanenhaus, 2002) suggests that these 'Lyons-esque' situations are relatively common in conversations about visually present objects. It was observed that adults do not utter restrictive modifiers every time there is more than one potential referent; nearly half of all definite NPs uttered (48%) did not have a unique referent in the scene (e.g., "Okay, pick up the square" might be uttered in the presence of multiple squares.) However, listeners' eye movements, actions and vocal responses all showed that they routinely achieved referential success under these conditions (e.g., picking up the correct square). Obviously, this success isn't evidence for psychic abilities on the part of the subjects. Rather, success occurred because the shape of the discourse and the goals of the task had narrowed the field of possible referents down to one (e.g., only one of the squares was currently a plausible referent). Definite NPs containing restrictive modifiers were uttered only when multiple potential referents were currently under discussion.

Now consider the child, who is trying to learn how deictic reference works for definite and indefinite NPs. If the child has difficulty understanding how the goals of their interlocutor restrict referential domains (a reasonable assumption), they should have specific difficulty with these Lyonsesque situations. Perhaps not surprisingly, developmental studies of definite reference show that young children (3-6 yrs) tend to behave egocentrically in these situations. In particular, in the absence of information that might guide a child's referential domain to the one intended by the speaker, young children's comprehension and productions suggest that what they assume to be the referent must also be what their interlocutor assumes to be the referent (Maratsos, 1976; Karmiloff-Smith, 1979). We will return to this issue in more detail when we discuss Wexler's (2003) recent reinterpretation of this data, but for the time being we simply relate the egocentric account to the current *put*-instruction data.

In particular, we now have better reason to believe our account of the child *put*-study. That is, children receive only sporadic (probabilistic) evidence that a definite NP (*the frog)* that deictically refers to a member of a set of objects of the same type will require restrictive modification (e.g., *the yellow frog, the frog on the napkin*). Moreover, children's discovery of the *actual* contribution of restrictive modifiers requires an understanding of common ground, the dynamics of discourse, and a decent model of shared goals with interlocutors (see Clark, 1993; Tanenhaus, Hanna & Chambers, in press). An older child could understand that a definite NP must apply to the current domain of reference, but his or her estimation of the domain of reference should be expected to get out of line with an interlocutor, certainly more often than an adult in a similar situation.

With this in mind, it becomes easier to understand the child-adult differences found in the *put* study. Adults when presented with a set of objects that have been labeled (*a frog, a napkin, another frog, a box, another napkin*) understand that this set of objects reflects the current referential domain. Young children in such contexts should be expected to behave egocentrically, thinking when hearing *the frog,* that the frog they are thinking of is likely to be the referent. After all, this is true in the common deictic referential situation of one entity, and in situations where other factors allow the child to be guided toward an understanding of a referential domain that is the same as the speaker's, i.e., the correct subset. Indeed, as Trueswell et al. (1999) noted, children's eye fixation patterns show this egocentricity; the frog they looked to first is a fairly good predictor of which frog they return to, and act upon, in their action of *putting*.

3.1. Lexical and referential evidence in interaction in adults and children

We have suggested that the referential scene manipulations were inadequate: simply putting in a scene multiple objects of the same type doesn't mean that a definite NP referring to a particular member of this set will be referentially disambiguiated via linguistic means (such as adding a restrictive modifier). Unfortunately, this account does not completely explain the adult behavior in the *put* study. That is, even in two referent scenes, we should have expected adults to show some temporary consideration of the goal interpretation of *on the napkin,* because the semantics and syntax of *put* supports this interpretation (contra the referential context). We know from years of research in adult sentence comprehension that strong lexical biases that run against contextual/plausibility biases result in temporary misanalysis, or garden-pathing (e.g., Britt, 1994; Garnsey et al., 1997; Spivey & Tanenhaus, 1998; Trueswell, 1996). Indeed, this is a fundamental prediction of constraint-satisfaction accounts of comprehension, which motivated our account of the developmental patterns.

One possibility worth considering quite seriously is that adults in fact did experience some difficulty with "Put the frog on the napkin into the box" in 2-frog scenes, but other factors conspired to reduce this difficulty and perhaps even measurement of this difficulty. The presence of the second prepositional phrase *into the box* increases the likelihood that the first PP is an NP modifier and not a goal. Adults may be especially good at using late post-ambiguity information to revise parses. Also, the primary measurement of garden-pathing in adults, i.e., looks to the empty napkin, occurs during and after the perception of this second prepositional phrase *into the box*. It is also possible that the prosodic structure of these utterances supports NP modification

even during the perception of *on the napkin*. Adults might be better at using this information as well. These factors alone could have reduced signs of considering the goal interpretation across the board in ambiguous trials (in both 1-referent and 2-referent scenes), with 2-referent scenes approaching the floor of the measure: essentially no looks to the empty napkin.

This explanation suggests that it would be worthwhile to take another look at 2-referent 1-referent manipulations in visual world studies, in which post-ambiguity and prosodic information do not strongly support the intended interpretation. Under these conditions, we might expect signs of lexical information constraining interpretation, and critically only partial effects of the 1- vs. 2- referent scene

In a very real sense, we are asking here whether we can make adults act like five-year-olds when it comes to parsing preferences. The prediction is that when lexical information supports a particular parsing preference, adults ought to show difficulty and possibly even a total failure to take the referential-scene implications into account. Specifically, in the presence of two possible referents for a definite NP (e.g., *the frog…* in the presence of two frogs) will adults fail to realize that a subsequent phrase should be treated as a restrictive modifier?

Snedeker and Trueswell (in press) addressed this question in a study containing sentences that are globally ambiguous in their structure (not just temporarily ambiguous). In the study, college-age adults heard sentences like those in (5a) through (5c).

5. a. *Tickle the pig with the fan.* (Instrument-biased Verb)
   b. *Feel the frog with the feather.* (Equi-biased Verb)
   c. *Choose the cow with the stick.* (Modifier-biased Verb)

Here, *with the X*, could attach to the verb as an instrument or could attach to the noun phrase as a modifier. Verbs were selected based on a separate sentence completion study, which evaluated how often a with-phrase would be used for these verbs as an instrument. As a result, verbs were operationally defined as: likely to mention an instrument (Instrument-bias), unlikely to mention and instrument (Modifier-bias), or somewhere in-between (Equi-bias). The nouns in the PP (fan, feather, stick) were normed in advance for thematic fit as an instrument for the corresponding verb. Nouns were selected that were rated as poor-to-adequate instruments in each verb class, such that average thematic fit was the same across these verb comparisons. The type of visual scene was also manipulated: 1- vs. 2-Referent scenes as shown in Figure 4.[7]

Insert Figure 4 about here.

The Instrument-biased sentences are most like the *put* items above. This is because the verb biases support a VP-attach assignment of the PP. Just like the empty napkin was a potential goal in the *put*-studies, the large stand-alone object (e.g., the

[7] Verb type was manipulated between subjects such that one group of subjects got Instrument-biased stimuli, another group got Equi-biased stimuli, and a third got Modifier-biased stimuli (embedded in numerous filler trials). For each verb-type, half the trials were 2-Referent scenes and half were 1-Referent scenes (with item-condition pairings counterbalanced across subjects).

feather) serves as the potential instrument here; looks to and use of this object provided a measure of the VP-attach (Instrument) interpretation of the PP (e.g., *with the feather*).

The offline action and eyegaze data from 24 adults are presented in Figure 5. On the left is plotted the proportion of trials in which an instrument action was performed by the subject (e.g., the proportion of trials in which the subject picked up the potential instrument (e.g., the large feather) and used it to act on one of the animals (e.g., the frog wearing a party hat). The right bar graph shows the proportion of trials in which subjects looked at the potential instrument during the course of the trial, regardless of whether they picked it up.

Insert Figure 5 about here.

Consider first the action data from 2-Referent scenes. It was predicted that if these scenes had demanded/required a NP modifier interpretation of '*with the X*' we should have seen essentially none of the subjects using instrument in these scenes. That is, subjects should have opted to use their hand to act upon the target animal, e.g., the frog holding the feather. However, we see this occurs only for Equi-bias and Modifier-bias verbs (0-10% instrument actions). For instrument-biased verbs in two-frog scenes, adults showed nearly 70% instrument actions, far greater than the expected 0-10%.

What appears to be going on here is that it didn't occur to most adults that '*with the x*' could be an NP modifier in this condition, precisely because the verb so strongly suggested the Instrument analysis. If this is so, it means that the definite NP (e.g., *the frog*) in this condition was referentially ambiguous. Indeed, on these instrument action trials, subjects half the time acted on the target animal (e.g., the frog holding the feather) and half the time acted upon the other animal (e.g., the frog wearing the party hat). Subjects who were given these items would sometimes say "which one?" in response to the first 2-referent target trial (to which the experimenter said "Please do your best.").

Moreover, post-experiment interviews showed that subjects were behaving in a pragmatically appropriate manner in response to this referential ambiguity, often in the way Lyons suggested. When explaining (at the end of the experiment) why they acted on a particular frog, they offered the spanner-strategy - "I thought you must have meant the frog closer to me" - or some other strategy - "I picked the frog not holding the feather just to be symmetrical" or "I picked randomly." Averaging across all observations, subjects went for either animal with equal frequency.

Adults are not oblivious to the scene constraint on the need for a restrictive modifier. If it occurs to them, they go for it (as evidenced by 2-referent Equi and Modifier-biased conditions). And of course, one-referent scenes increased the rate of using and looking at the potential instrument in all verb-types. This should be expected since definite reference to a scene containing multiple potential referents only partially predicts whether an NP would be contrastively marked with a modifier.

What if the data we just presented had come from five year olds, rather than from adults? One might have concluded that children do not fully understand definite reference (e.g., they don't understand the meaning of the definite determiner *the*, Wexler, this volume). But these are adults, not children. The logical conclusion is that syntactic

accessibility of forms, as determined by lexical factors, have a potent impact on interpretation and reference.

Of course, the question worth considering is how five-year-olds actually do behave in these very settings. Given that definite reference in these settings poorly constrains the structuring of an upcoming PP, it follows the five-year-olds might be even less influenced by these contextual manipulations. In addition, these same children ought to be quite sensitive to the lexical manipulations since this information is highly predictive of structure and hypothesized to be easy to track.

The data from five year olds (as reported in Snedeker & Trueswell, in press) are shown in Figure 6. Indeed, children perfectly matched the verb-biases (they know which verbs invite an instrument interpretation) but the referential scene manipulation had no effect on their actions and a small (non-significant) effect on their eye fixations.

Insert Figure 6 about here.

These data bode well for the developmental account sketched above. In adults, hearing definite reference in these two visual settings only partially contributes to parsing commitments, and verb-biases strongly constrain structuring of the linguistic input. Children have trouble using this sort of deictic referential information, but have little trouble using the lexical information.[8]

Notice also from the data in Figure 6 that one gets a very different picture of what the child finds referentially ambiguous depending on the lexico-syntactic properties of the utterance. This clearly indicates that study of what children 'know' and 'don't know' about language must be embedded within a theory of how linguistic information is dynamically processed by children. This point is often overlooked in the traditional study of language acquisition (e.g., Wexler, this volume).

3.2 The effects of discourse and pragmatics on child parsing

The linguistic observations and experimental findings sketched in the previous section strongly suggests that the shape of the discourse and the goals of the interlocutors ought to be a far better predictor of the referential domain of a referential expression and hence a better predictor of the level of specificity needed for that expression. For instance, recall that Brown-Schmidt et al. (2002) found that the conversational content when combined with the scene shapes the referential domain and predicts quite well the specificity of definite reference.

[8] It appears that the prior *put* studies may have had additional factors that reduced garden-pathing generally in adults, but not children: specifically post-ambiguity information and prosodic information. Consistent with this, Choi & Mazuka (2003) have provided evidence that prosodic information has little effect on children's interpretation of syntactic ambiguity in Korean, and Snedeker et al. (2003) found significant but weak effects of prosody on PP-attachment ambiguity resolution in children. Also, post-ambiguity (i.e., disambiguating) information might be especially effective for adults in parsing, an issue we return to later in the paper. *Put the X on the…* and *Tickle the X with the…* also differ in the type of semantic properties that are competing with each other in the ambiguity. It is possible that the put ambiguity is easier to revise because both interpretations involve the location of an object (see Snedeker & Trueswell, in press).

If a discourse guides a child listener toward conceiving of the situation in the same way as the speaker talking to the child, we might expect children to use *this* contextual information to guide parsing commitments (egocentrically or otherwise). To this end, we have begun to ask whether potentially potent evidence from the discourse can influence five-year olds' parsing decisions (Hurewitz, Brown-Schmidt, Trueswell & Gleitman, in progress). Here a preceding discourse, established by two conversing puppets, provides the goal to isolate one referent from among multiple referents in the scene prior to hearing an ambiguous PP. If these discourse goals provide a strong constraint on the need for the otherwise ambiguous PP to be a modifier, we might expect even 5-year olds to be sensitive to this fact, combining it with lexical constraints on structure. In other words, we are asking here if we can turn children into adults.

In the study, children (N=24; Age= 4;0-5;6) were tested in a modified version of the Truth Value Judgment task (Crain & Thornton, 1998). On each trial, the child heard a story acted out in the presence of a puppet ("Mr. Walrus", who is known to be not terribly bright). At the end of the story, a second puppet (the clever Ms. Rabbit, who had been hiding under the table listening to the story), appeared and asked Mr. Walrus questions about the story. The child's job was to evaluate and if necessary correct Mr. Walrus's answers to her questions.

On critical trials, each child was always presented with a 2-referent scene (as in Figure 7, two cats, one on a book, one on a fence, a toy barn, another fence, and a turtle; again, real-world objects, not clip-art images, were used).

Insert Figure 7 about here.

The story, pre-recorded and acted out by the experimenter (E) deictically referred to each animal, and established the pair of cats in distinct events, and is paraphrased below:

> *This cat* [E grabs the cat on the book] *and this turtle* [E grabs turtle] *decided to go for a walk, and met up on top of the barn* [E moves animal to barn]. *Suddenly, the turtle tickled the cat. 'Tickle, tickle tickle!' 'Hee! Hee! Hee!'* [E performs appropriate actions with the animals.] *And then they went home.* [E returns each animal to original starting place.] *And, this cat* [E grabs cat on fence] *saw all this and laughed and laughed as well.*

With all the animals back in their original location, Ms. Rabbit returns to ask Mr. Walrus a question. In all conditions, Walrus's answer contains an attachment ambiguity: *I know, the turtle tickled the cat on the fence.* Here *on the fence* can indicate where the tickling happened (Locative VP-attachment) or indicate a particular cat (Locative NP-attachment). Mr. Walrus' utterance, however, was preceded by a question from Ms. Rabbit that either supports the need to contrast the cats (the Contrastive Question condition, *Which cat did the turtle tickle?*) or does not support this goal (the Non-Contrastive Question condition, *Can you tell me something about the story?*). In all cases, both interpretations of the ambiguous sentence <u>are false</u> because the story actually involved the cat on the book being tickled by the turtle, in a different location, i.e., when

they both had been on the barn. Hence, however the child parsed the sentence, she still must correct Mr. Walrus. It is the child's *particular* correction of Mr. Walrus that can reveal the implicit parse-choice (*No! It happened OVER HERE on the barn!* or *No! THIS CAT was tickled, the one on the book!*).

The Question-Type factor (Contrastive vs. Non-Contrastive) was crossed with a verb manipulation. Half the trials involved eventive verbs (such as *tickle*), which easily allow for locative (VP-attach) modifiers such as *on the barn*. The other half involved stative verbs, where the story and the critical sentence involved, e.g., liking (*The turtle liked the cat on the fence.*). Stative verbs do not usually permit locative modifiers, because states typically are not connected to a particular place.

Given the importance of conversation constraints in modifier use (see above), the multiple-constraint account predicts that this sort of discourse manipulation (here, Q-type), as well as lexical information (V-Type) should influence parsing preferences, even in 5-year olds. That is, Contrastive Questions and Stative verbs should both induce greater modifier interpretations and resulting corrections by the child.

Figure 8 plots the proportion of NP modification interpretations exhibited by children in the four conditions. As one can see, the pattern of these corrections across conditions supports our account. In particular, Contrastive questions led to many more NP modification corrections than when the question was not Contrastive. Also, there was an effect of verb-type: Stative verbs led to more NP modification interpretations than when the verb was eventive. This resulted in reliable effects of question-type and verb-type. And, interestingly, adult controls exhibited an even stronger reliance on the discourse-needs of the questions, with adults even coercing stative verbs into eventive readings.

Insert Figure 8 about here.

It should be emphasized that our account of the child parsing phenomena is that it is the product of a system that automatically combines multiple evidential sources, all to serve the purpose of coming up with an estimate of the intended meaning of an utterance. If this is the case, we should expect to observe striking differences in children's use of structures in production and comprehension tasks under particular comprehension conditions. That is, children might correctly produce a particular structure in the right referential context, but then fail to understand this same structure during comprehension, when the lexical sources support an alternative analysis.

To test these claims, Hurewitz, Brown-Schmidt, Thorpe, Gleitman and Trueswell (2000) examined five-year olds' production and comprehension abilities in two-referent scenes. Children heard a story, acted out by the experimenter, which introduced salient differences between the two frogs (by having them doing different things). Afterwards, they were tested by asking them a Specific question: *Which frog visited Mrs. Squid's house?* To answer this question required (a) understanding the story, (b) understanding that the question requires an answer that distinguishes the frogs via locative modification as these frogs were otherwise identical, and (c) producing in the answer a restrictive modifier; namely, *The frog/one on the napkin*. Immediately thereafter, the same child

was asked to perform the *put*-task of Trueswell et al. (1999): *Very good. Touch the Smiley Face. Now put the frog on the napkin into the box.* As a control, another group of children were asked a General question (*Can you tell me something about the story?*) prior to doing the *put* task.

Children's production performance on the Specific question showed they were able to perform all of the relevant non-linguistic and linguistic acrobatics to specify uniquely a referent through restrictive modification: 72% of all answers to the specific question were correct, providing answers like *The frog on the napkin*. In striking contrast, these same children's response to the *put* instruction showed the same mis-analysis effects as those reported in Trueswell et al. (1999). They typically performed incorrect actions (over 70% of the trials), which involved the incorrect destination. And, children were at chance in selecting between the two frogs. That is, the very same child who had just correctly responded to the story-question by producing a PP-modified NP (*the frog on the napkin)* might now in response to *Put the frog on the napkin into the box*, pick up the other frog, move it over to the empty napkin and then put it into the box. The sheer differences in *complexity* between the two sentences cannot account for the findings as we know from earlier experimentation (the same children have no difficulty with unambiguous control sentences of equal complexity, e.g., *Put the frog that's on the napkin into the box).*

A further experiment in this line (Hurewitz et al., 2000 Exp. 2) investigated the possibility that children just weren't inclined to notice napkins as salient components of scene description. Making the platforms on which frogs were ensconced more salient (frilly umbrellas and royal thrones) generally increased performance in production (87% restrictive modifiers in production), but still the striking asymmetry between production and comprehension was preserved (60% errors in comprehension). In addition, in this version of the experiment we eyetracked the young subjects, with the on-line results replicating Trueswell et al. (1999).

So, in both of these experiments, we observe, like in the Rabbit-Walrus study earlier, children understanding how the discourse can specify the need for an NP restrictive modifier. In particular, in the case of the Rabbit-Walrus study, we see this discourse-syntax knowledge at work in comprehension: a Contrastive question generates an increased chance of interpreting an ambiguous PP as an NP modifier, though this knowledge must battle against lexical evidence that may support an alternative interpretation (e.g., Eventive verbs generated some Question/Discourse-inappropriate responses in children). The experiments in the present section demonstrate this discourse-syntactic knowledge in children's own productions: Contrastive questions generated a need for referential specificity in the form of a modifier (*The frog/one on the napkin*), which the children often uttered in response to this question type. However, when we then pull out a *put*-sentence from our lexical arsenal, we see that we can tip the scales back to VP attachment, even in the same child who had just a moment ago demonstrated knowledge of the discourse-syntax facts in his or her own productions.

It should be noted though that the discourse conditions are indeed slightly different between our production and comprehension test conditions. The distinction between the frogs had just been made by the child in his or her utterance, and thus the discourse-goal of contrasting the frogs had been achieved by the time we tested for comprehension abilities in our *put*-instruction. We strongly suspect however that *put* was

exerting detrimental effects, since unpublished work from our lab has examined *put*-sentences as part of an answer to a contrastive question (Rabbit: *Which frog should I move?* Walrus: *I know, put the frog on the napkin into the box.*). Here we still find strong VP-attachment preferences despite the immediately preceding contrastive question (Hurewitz, Brown-Schmidt, Trueswell & Gleitman, in progress). Thus, the data strongly support the automatic use of verb-preferences in the young child parsing system.

4. Pragmatic vs. Semantic Accounts

Our account of the child parsing data depends heavily on the pragmatics literature on deictic reference as well as the conclusions drawn by Maratsos and Karmiloff-Smith regarding the development of definite reference. Specifically, faced with a definite NP, children have difficulty calculating the relevant referential domain and frequently do so egocentrically, thinking that what they take to be the referent of a definite NP must also be what their interlocutor believes to be the referent.

Recently, Wexler (2003, and this volume) has offered a reinterpretation of the Maratsos and Karmiloff-Smith data and has gone on to apply his explanation to the kindergarten-path phenomena. For motivation of his position, Wexler points to a specific study by Maratsos (replicated later by Karmiloff-Smith in French). In this experiment, children heard a story that introduced either multiple potential referents of the same type (i.e., several boys and several girls) or just singletons (i.e., one boy and one girl). The story ended with an intentionally vague assertion (e.g., *Someone started giggling and laughing*) followed by a question (*Who was giggling and laughing?*). It was found, especially in the Karmiloff-Smith version of the study, that younger children (ages 3 to 7) overused definite NPs in answering the question. That is, even in a story containing multiple, undifferentiated boys and girls, these children tended to answer "The boy" or "The girl," rather than the expected "I don't know."

On the surface, these data seem consistent with an egocentric account: a younger child has a particular boy or a particular girl in mind, and without realizing the addressee might not, utters "the boy" or "the girl." Wexler questions this account because the stories used in these particular studies did not introduce individual boys and girls but simply established an undifferentiated set of boys and an undifferentiated set of girls. This makes it unlikely that the children could have focused on a particular boy or girl in their mental model.

Wexler proposes an alternative explanation of these findings, according to which the child's deficit is semantic rather than pragmatic. He suggests that children in this age range (3-7 years) lack a full understanding of the meaning of *the*, and hence systematically misuse (and misinterpret) it. In particular, it is proposed that children lack the notion of maximality, which is a crucial part of the linguistic meaning of the definite determiner. Maximality as it applies to the semantics of the definite determiner can be expressed logically as follows. Given a predicate P that takes an argument X, P must be true for each and every member of the set of type X that is in the current referential domain. For instance, in order for the sentence *The boy laughed* to be true, each and every boy in the current referential domain needs to have laughed. In case there is more than one boy in this set, the definite NP needs to be pluralized (*The boys laughed*) and it is necessary for all the boys in the set to have laughed.

Wexler claims that the child version of *the* lacks this notion of maximality: for children, *the* only requires that the predicate applies to some element in the referential domain. Thus in the Maratsos and Karmiloff-Smith stories, children say that the character who laughed was "the boy", even though only one of several boys in the referential domain laughed. In the case of *Put the frog on the napkin into the box*, the number of frogs doesn't matter to 3-7 year olds, and the NP *the frog* on its own can refer to any member of the set of frogs.

Wexler's proposal has some merit because it draws explicit connections between developmental data and semantic theories of definiteness and aims at providing a detailed account of referential development. However, we believe that the theory is unable to capture key intuitions about referential errors and does not handle the range of developmental data already collected about definite reference and syntactic ambiguity resolution.

First, Wexler's semantic account appears to miss intuitions about referential errors of the sort we are interested in. On his account, in a referential domain of several boys (e.g., one crying, one laughing, one sleeping) the sentence *The boy is crying* is <u>false</u>, rather than simply infelicitous. That is, if there is more than one boy in the domain of reference (and the other boys are not crying), the predicate is false because it is not true of all members of the set *boy*. However, the present authors all have the strong intuition that the utterance in this situation is *true* but infelicitous; that is, 'it should have been said in a different way.' Even though adult intuitions about truth vs. felicity are not always robust, it is an advantage for an independently motivated theory to be able to capture such adult judgments.

Second, even though Wexler's account moves away from pragmatic issues in the calculation of the referential domain for definite descriptions, such issues seem inescapable for anyone trying to account for the interpretation of definites in both children and adults (for pointers to a large linguistic and philosophical literature, see Larson & Siegal, 1995). From a psycholinguistic viewpoint, we have already seen numerous cases where the referential domain is calculated not only from the prior context but also from the pragmatic implications of the semantic content of the utterance itself. For instance, consider again the scene in Figure 3: what is the maximal domain for *Pull the rabbit out of the hat* and why is it different for *Pet the rabbit*? In the latter case, we respond *Which one?* - why not in the former? Furthermore, there is now good evidence that the referential domain of an utterance changes for an adult listener based on contextual demands as that utterance unfolds over time. In an on-line comprehension study with adults, Chambers Tanenhaus, Eberhard, Filip and Carlson (2002) showed that adult listeners found the sentence *Put the duck inside the can* to be felicitous in the presence of more than one can so long as one and only one can was large enough to hold the duck; so felicitous in fact, that they applied this pragmatic restriction in real-time, moving their eyes to the appropriate sized can upon hearing …*inside the*… Chambers et al. demonstrated that this effect changes with the goals of the speaker. In the same scene, the sentence *Are you able to put the duck inside the can?* resulted in looks to all of the cans, plus often a vocal response of *Which one?*

In order to capture this flexibility in interpreting referential expressions, any account of reference must include what is plausible given the goals of the interlocutors as a matter of course. Such an account should be able to explain how mismatches of

referential domain may give rise to puzzlement or misunderstandings between interlocutors, even when no semantic problems with definite determiners exist (as in the adult case). A reasonable (and parsimonious) interpretation of children's errors with definite descriptions, then, would be to attribute such errors mostly to pragmatic-referential factors. After all, even the youngest children do not consistently fail in referential studies of production and comprehension; sometimes they have calculated the correct referential domain.

Third, the full range of developmental data on definite reference do not seem to be explained by the semantic-deficit account, even for the specific study used by Wexler to motivate his account. The stories in Maratsos (1976) enumerate the members of the set, which may encourage the establishment of individuals of these sets of boys and girls. This makes it more likely that children had a particular entity in mind when using definite descriptions (*the boy*) in their responses. There are also important puzzles to the data: only a subset of the 4-year-olds behaved in the way predicted by Wexler (those who performed poorly on a separate sentence repetition task). This correlation with memory abilities was only found in this age group, and younger children actually performed better, correctly uttering *a boy* in the context of several boys 83% of the time.

Karmiloff-Smith (1979) in a very similar study in French finds a different developmental pattern, more in line with Wexler's account: when asked to indicate a particular boy or girl in a response to a question about a story containing several boys and several girls, young children (ages 3-7) answered *le garcon/la fille* (the boy/the girl) whereas older children (ages 8-10) responded predominantly with *un garcon/une fille* (a boy/a girl). But interestingly, two out of four adults we asked reacted to translations of these materials by answering with a specific referent ("I don't know, *Mary?*"). These subjects volunteered a name despite the fact that there was no correct answer, presumably because the instructions (*Guess who it was*) required a single, specific referent. Indeed, Karmiloff-Smith noted that children and adults sometimes responded this way (e.g., giving their own name or the name of one of their school mates such as Juliette). This shows that even adults may make quite specific guesses about the identity of referents they cannot truly individuate.

Fourth, under Wexler's view, we would expect children to make more errors than they do with definite descriptions. In recent work that looked at children's comprehension of implicatures, Papafragou and Tantalou (in press) presented children with stories in which a character was supposed to perform a certain action to win a prize. In one case, an elephant was given three oranges and was told *he had to eat the oranges.* When the elephant came back and was asked if he ate the oranges, he answered *I ate some.* A vast majority of 5-year-olds decided the elephant should not get a prize and justified their response by indicating, e.g., that the elephant had not eaten ALL of the oranges. This reveals that children's preferred interpretation for *the oranges* obeys maximality, otherwise the contrast with *some* would be lost. (In section 5 below, we discuss further evidence from Maratsos (1976) that children can behave in an adult-like manner when it comes to definite and indefinite reference under certain circumstances.)

Finally, the full range of data on syntactic ambiguity resolution is difficult to capture under a purely semantic account. In particular, Wexler suggests that a child semantic deficit accounts for the five year olds' failure to realize that a scene containing multiple frogs requires a modifier interpretation of an ambiguous phrase. However, if

this is true, adults seem to also have this semantic deficit. Recall that we showed that adults frequently failed to realize that a modifier interpretation was needed in *Tickle the frog with the feather* in 2-frog scenes (Snedeker & Trueswell, in press). Syntactic accessibility of the alternative meanings of *with the feather* also drives reference patterns, in both adults and children. Another way of stating this is that the semantic deficit account must assume that very similar patterns in adults and children are the result of performance factors in one case and competence factors in the other, and that the similarity between the adult and child data sets is just accidental.

Rather, the picture emerging from this literature on reference and ambiguity resolution is quite different: success on reference assignment requires, among other things, rapid tracking of shifting and flexible referential domains. Children by age 4 or 5 have trouble aligning their referential domain with that of their interlocutor even though they understand the semantics of definiteness, including the maximality assumption. Indeed, such mismatches of referential domains can occur in adults (we return to this issue, and to individual differences in the adult population, in section 5 below). We should be clear here; we are not claiming that maximality is innate and pragmatic factors mask this fact. The pragmatic facts themselves suggest that learning how definite reference behaves requires some work. We do claim, however, that by age five, children understand the semantics of definiteness, including the maximality assumption. Success and failure in child reference in this age appears to be driven by successful match or mismatch in calculations of referential domain.

5. Conflict resolution and garden-path lingering

We close this paper by considering briefly one of the most striking findings reported in Trueswell et al. (1999) which we have not discussed so far. In the original *put*-study and the follow-up studies, children appear to behave impulsively, not revising their initial referential and syntactic commitments. More concretely, after hearing the sentence *Put the frog on the napkin into the box*, children often act upon the frog that they looked to first upon hearing *the frog*. Moreover, the tendency to initially interpret *on the napkin* as a goal is often not rescinded upon hearing the PP *into the box*: five-year-olds tend to carry out actions in which a frog goes to the empty napkin. Adults rarely do this, and neither do five-year olds when the ambiguity is removed (*Put the frog that's on the napkin…*). Why is this so?

We discuss here the possibility that this change in revision abilities is related to differences in executive function abilities, specifically the ability to select a subordinate analysis under conditions of representational conflict (Trueswell & Gleitman, in press; Novick, Trueswell & Thompson-Schill, submitted). In particular, Novick et al. (submitted) propose that these patterns represent changes in executive function abilities generally over the course of development, especially those associated with response selection of representations under conditions of conflict (e.g., Diamond & Taylor, 1996; Zelazo & Reznick, 1991). Frontal lobe regions (e.g., Left Prefrontal Cortex) appear to be implicated in this ability (e.g., Thompson-Schill, Jonides, Marshuetz, Smith, D'Esposito, Kan, Knight & Swick, 2002). And indeed, these regions are known to be late developing anatomically, well into years five and six (e.g., Huttenlocher & Dabholkar, 1997). It seems quite plausible then to consider that the reranking of interpretations in garden-path phenomena (inhibiting an initial interpretation, selecting a

new interpretation) would involve these very systems. And indeed, perseveration in response is the hallmark of patients with severe frontal lobe damage -- and normally developing children. For instance, in the Wisconsin Card Sorting task, participants first sort cards by one criteria (e.g., by color) but then are asked to switch to sorting by another criteria (e.g., by shape). Children and frontal lobe patients have great difficulty switching between criteria, presumably because one sort of evidence was developed (color) as relevant and then must be over-ridden by some other evidence (shape).

To address this account of garden-path recovery, we begin by noting that children are not unique in this failure to rescind garden-path interpretations. Christianson, Hollingworth, Halliwell & Ferreira (2001) have shown this in comprehension questions about garden-path sentences. The questions were specifically designed so that the correct answer would be "NO" about the intended meaning of the sentence as a whole, but would be answered "YES" if based (erroneously) on the temporarily considered but rejected interpretation. These questions took significantly longer to answer and showed more errors than the same questions about unambiguous versions of the sentences. That is, the garden-path appears to linger, as if subjects sometimes failed to completely inhibit/reject the intended interpretation (Christianson et al., 2001).

One might expect individual differences in this ability related to executive function generally. Indeed, Mendelsohn (2002, 2003) has offered evidence in favor of this claim. In her study, she found that the size of subjects' lingering garden-path effect correlated with several linguistic and nonlinguistic measures all of which arguably involve inhibition/selection mechanisms. Perhaps the most compelling observation was that a completely non-linguistic task, the so-called anti-saccade task, correlated with lingering garden-path measures. In this task, subjects were to look in a direction opposite of a flash of light (thus inhibiting the reflex to look to the light); difficulty on this task correlated with the ability to reject garden-path interpretations of sentences.

Our own research group, in collaboration with Sharon Thompson-Schill (Novick, January, Trueswell & Thompson-Schill, in preparation) has been exploring similar issues of individual differences in garden-path recovery. We have, for instance, found that adults who make child-like errors of action in the *put*-task also show decreased performance on measures of executive function (tasks that don't involve grammatical processing of linguistic material).

We believe these issues are deeply related to some interesting experimental findings that point to differences in deictic and anaphoric reference. We mentioned above that in some studies from Maratsos (1976) young children behave much better in their use of definite and indefinite NPs. In these studies, Maratsos again compared referential situations in which either multiple entities were present (several boy dolls and several girl dolls) or singletons (one boy and one girl). All of the dolls were placed at the top of a slide, and children played a game with the experimenter in which the experimenter could send a doll down the slide in a toy car. The child had to decide which doll to send down, and tell the experimenter.

Two conditions were compared: one in which the dolls could be seen by the child and experimenter, and another in which the dolls were shown but then placed out of view of the child. When the dolls were in view, many younger children made definite NP errors of the familiar sort (*Okay, send down the boy!* in the presence of multiple boy dolls). Maratsos reported that when children used a definite NP, they clearly had a

particular doll in mind and were looking directly at it. They also made explicit comments on the experimenter's choice of referents such as *That's the one!* or *That one's all right, I guess…* However, when the dolls were out of sight the use of indefinites increased significantly (*Okay, send down a boy!*). Maratsos took this as a sign that children were acting egocentrically, though it was not made entirely clear why hiding objects would discourage egocentric behavior.

Interestingly, Meroni and Crain (this volume) report a related out-of-sight phenomenon in a kindergarten-path *put* study. They report that children in 2-frog scenes do quite well (around 90% correct actions) when they are first asked to close their eyes before hearing the utterance. We take this finding as very preliminary because comparisons have not been made with the same stimuli and same settings in an eyes-open condition (they compare only to the Trueswell et al., 1999, 2-Referent ambiguous condition, where there were approximately 35% correct actions). Indeed, although it seems likely there is an effect of closing the eyes, we are surprised by such a high number of successful interpretations in the Meroni and Crain study and suspect that other differences between studies may explain some of the good performance (i.e., the discourse context and the goals for the child were also different, and we believe further supported the correct interpretation). Nevertheless this finding is potentially quite important.

We take these two out-of-sight phenomena (the Maratsos slide study and the Meroni & Crain eyes-closed study) to be clearly related. And we believe there is good reason to expect these patterns. First, informally, our intuition is that deictic reference and anaphoric reference behave differently in one especially relevant way. That is, it seems likely that restrictive modifiers are used less often in conversations about visually present referents as compared to visually absent ones. After all, a speaker can identify the referent of a linguistic expression (e.g., *the frog*) using a range of behaviors that pertain to his/her attentional state: eye gaze, pointing, posture, etc. These nonlinguistic means for referential disambiguation are not available in anaphoric, non-deictic, reference, thereby increasing the likelihood that a listener will disambiguate a definite NP linguistically (e.g., saying *the frog that you like*, rather than just *the frog*). If this is so, children may expect more simple NPs in deictic situations.

But there is a deeper connection, relating to the issue of differences in selection and inhibition. Meroni and Crain offer one kind of story along this line, suggesting that closing ones eyes gets people out of an "interpret-mode" of language comprehension; children may have trouble inhibiting plans made incrementally during speech perception. However, this doesn't explain the Maratsos language production findings, regarding changes in the use of definite and indefinite NPs when the referents are not visually present. Here the child has a plan in mind and is producing an utterance that describes this plan. Nor does this account adequately explain why the absence of visual input would postpone planning in comprehension.

We suggest there is in fact an underlying connection between visual attention and discourse focus that better explains both of these out-of-sight phenomena, which need not propose that interpretation is postponed when objects are not present. In particular, if one takes seriously the informal analogy that the current referential domain in a discourse model is like an attentional mechanism, these developmental patterns are expected. In the case of attention to entities in a discourse model, this attentional focus is not spatially

restricted: we can think of two boys even if they are not near each other spatially. However, looks to a particular boy in the world necessarily focus our attention on that boy and not other boys if they are spatially distant (the fovea subtends 2-3 degrees visual angle and material that is foveated is typically what is being attended to). Deictic reference reflects an interface between a mental model of the world and these co-present objects, whose perception and visual recognition point to these same entities in the mental model.

These developmental differences could be recast as arising from differences in executive function: visual attention and discourse attention can conflict; younger children may not easily deal with this sort of conflict. Visual inspection of an object should be expected to activate this entity in the discourse model, but in many cases the current referential domain is larger or different than this attentional space, and hence such effects on the model must be inhibited. Indeed, excessive problems on interference tasks (e.g., the Stroop task) are also common in adults with frontal lobe deficits (e.g., see Thompson-Schill et al., 2002).

Interestingly, Maratsos noted that even some adults behaved in a childlike fashion when the dolls were visually co-present, saying "send down the boy" in the presence of multiple boys. This is to be expected since this sort of disambiguation by eyegaze or other means can occur. However, we would predict that some of this behavior could very well correlate with individual differences in executive function, as has been found with lingering garden-paths in comprehension in adults (Mendelsohn, 2002, 2003).

Finally, we also make the following prediction: in addition to an eyes closed effect, there should be an effect of visual angle in children: that is, pairs of objects that are close together spatially (close enough to be foveated simultaneously) are likely to improve performance in 2-frog put-studies.[9] That is, if visual attention can activate the relevant referential domain (i.e., both frogs simultaneously) we should expect children to more easily discover the felicitous restrictive modifier reading of …*the frog on the napkin*… Plans are underway to explore this hypothesis.

6. Closing Remarks

Based on the work reviewed here, it should be clear that young children who are trying to comprehend language are faced with a processing problem of considerable complexity. From a sequence of words, a child must rapidly glean detailed grammatical information in order to determine not only 'who is doing what to whom' but also how an utterance relates to their current conception of the world. There exist multiple probabilistic sources of evidence for constraining the grammatical structure of an utterance, and the child must discover, weigh, and combine this evidence. We have suggested elsewhere that the child tracks and builds detailed syntactic and semantic representations of words which allow for the efficient recovery of the structure of the sentence as a whole (e.g., Trueswell & Gleitman, in press). However, this only gets a child or adult listener so far. The referential implications of these analyses must also be considered since some analyses are very unlikely given the referential setting.

Here, the problem for the developmental theorist and for the child is the same: they must grapple with the fact that the interlocutors' goals and shared conception of the

[9] Previous studies in our lab had the animals approximately 10 degrees apart, which makes it relatively easy to collect accurate eyegaze measures for the objects.

world guide what is considered to be the relevant referential domain. Within the realm of deictic reference, the scene itself only partially constrains the choice of referential expressions (i.e., the level of specificity / linguistic economy offered in an utterance, *the frog, the frog closer to you*). What matters more is which aspects of this world are under discussion and relevant to the intended goals of the utterance. We have discussed numerous cases in adult conversation and adult comprehension where this is shown to be the case. We then turned our attention to the implications these observations might have for a child who is learning how deictic reference works, and how reference might impact relevant syntactic choices during comprehension.

The story we offer is long, and admittedly complicated, but how could it be any other way? After all, reference is amazingly sensitive to a vast array of considerations, from the linguistic to the nonlinguistic, and the child must learn these facts presumably bit by bit. Indeed we believe the complexity of the referential process explains in part why 'bottom-up' sources of structure (i.e., the lexical evidence) so strongly constrain syntactic ambiguity resolution at earlier stages of development. Even after a child understands how reference works within a referential domain, the computation of that domain, which rapidly changes over time, is expected to be frequently misaligned with the domain currently entertained by his or her interlocutor. Thus, fairly systematic errors can be observed in definite reference both in production and comprehension, especially with regard to the syntactic concomitants of reference.

Nevertheless we have documented several instances in which the child successfully computes the referential domain, and understands the level of specificity needed for reference within this domain. Most of these successes we would argue arise from circumstances in which there is a clear reliable indicator from the conversation about the shape of the current referential domain, and the goal of the utterance (e.g., *which*-questions need to contrast members of a set). Under a view in which comprehension is a 'guessing game' in which listeners are continuously guessing the intentions of the speaker, we might expect such a pattern: highly constraining, easy to discover sources of evidence trump, in developmental time, the less reliable and more complex. We strongly suspect that other demonstrations of non-egocentric referential behaviors in young children also arise out of simpler mechanisms that unambiguously predict the attentional state of the interlocutor (e.g., eyegaze and physical constraints on perception, see Baldwin, 1991; Nadig & Sedivy, 2002). Indeed, there is now growing evidence that 'theory of mind' abilities emerge out of converging cues of this sort, and simpler use of some of these cues arises in part in other species, and grows during development in children (Leslie, 2000; Call & Tomasello, in press).

Finally, we have considered attentional development in children and have suggested that during the presence of physical objects under discussion, children are overly sensitive to their own visual/perceptual attentional state when it comes to calculating what to attend to in their discourse model. Of the age of interest here (4-6 years), we have concluded they understand that a definite NP (produced or heard) must apply maximally to the current referential domain, but this domain is skewed by their attentional state. This explains a range of phenomena in their own use of definite and indefinite reference, but also explains their difficulty applying referential facts to parsing procedures. Moreover, developmental changes in a listener's ability to rescind syntactic commitments may reflect this same processing change, in which competing structure(s)

must be rapidly inhibited, and the correct structure must be activated. This points to an explanation of individual differences in adults regarding errors in production and comprehension, which are skewed in the direction of the child-like behaviors.

It is clear that we are only at the early stages of understanding how the processing demands of language production and language comprehension impinge on the embedded language learning process. However, methods like those discussed here, which examine how children interpret speech in real-time, offer new insight into these developing processing abilities. The data suggest that there is hope for building a more unified theory of language acquisition and language use over a lifetime, in which we recognize that multiple sources of evidence must be discovered, and sometimes built by the child, in a way that allows for immediate integration of this evidence into comprehension and production mechanisms.

Author Notes

Some of the research reported herein was supported by grants from the National Institute of Health, R01- HD37507 to Trueswell and 5F32-MH6502 to Papafragou. We would like to thank Peter Gordon of Columbia's Teacher College and Colin Phillips of University of Maryland College Park for insightful comments, specifically about the development of definite reference and Wexler's semantic account of these phenomena. We also thank David January and Samantha Crane for their assistance in the preparation of this chapter. Finally, we thank Lila Gleitman for her support and helpful comments on earlier drafts.

References

Allopenna, P.D., Magnuson, J.S. & Tanenhaus, M.K. (1998). Tracking the time course of spoken word recognition: Evidence for continuous mapping models. Journal of Memory and Language, 38, 419-439.

Altmann, G. & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. Cognition, 73, 247-264.

Altmann, G. & Steedman, M. (1988). Interaction with context during human sentence processing. Cognition, 30, 191-238.

Aslin, R.N., Saffran, J.R. & Newport, E.L. (1998). Computation of conditional probability statistics by 8-month old infants. Psychological Science, 9, 321-324.

Baldwin, D.A. (1991). Infant contribution to the achievement of joint reference. Child Development, 62, 875-890.

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. (2002). Reference resolution in the wild: On-line circumscription of referential domains in a natural, interactive problem-solving task. Proceedings of the 24th Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Lawrence Erlbaum.

Call, J. & Tomasello, M. (in press). Reasoning and thinking in nonhuman primates. In K. Holyoak & B. Morrison (eds.), Cambridge Handbook of Thinking and Reasoning. Cambridge: Cambridge University Press.

Carlson, G.N. & Tanenhaus, M.K. (1988). Thematic roles and language comprehension. In W. Wilkins, (Ed.), Syntax and Semantics, Vol. 21: Thematic Relations, pp. 263-89. Academic Press.

Christianson, K., Hollingworth, A., Halliwell, J. & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. Cognitive Psychology, 42, 368-407.

Chambers, C. G., Tanenhaus, M.K, Eberhard, K.M; Filip, H. & Carlson, G.N (2002). Circumscribing referential domains during real-time language comprehension. Journal of Memory & Language. 47(1), 30-49

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

Chomsky, C. (1969). Acquisition of syntax in children from 5 to 10. Cambridge, MA: M.I.T. Press.

Christakis, D.A., Zimmerman, F.J., DiGiuseppe, D.L. & McCarty, C.A. (2004). Early Television Exposure and Subsequent Attentional Problems in Children. Pediatrics, 113(4), 708-713.

Clark, H.H. (1993). Arenas of Language Use. The University of Chicago Press. Chicago, Ill.

Crain, S. (1980). Pragmatic constraints on sentence comprehension. Unpublished PhD. Dissertation, University of California at Irvine.

Crain, S. & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Karrattunen & A. Zwicky (Eds.), Natural language parsing: Psychological, Computational, and Theoretical Perspectives. Cambridge, MA: Cambridge University Press.

Crain, S. & Thornton, R. (1998). Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics. Cambridge, MA: MIT Press.

Diamond, A. & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "Do as I say, not as I do." Developmental Psychobiology. Vol. 29, 315-334.

Fernald, A. (2001). How two-year-olds look as they listen: the search for the object begins at the verb. Talk presented at the 14th Annual CUNY Conference on Human Sentence Processing. March 15-17, 2001, Philadelphia, PA.

Fisher, C., Hall, G., Rakowitz, S. & Gleitman, L. (1994). When it is better to receive than to give. Lingua, 92, 333-375.

Fisher, C., Gleitman, H. & Gleitman, L. (1992). On the semantic content of subcategorization frames. Cognitive Psychology, 23(3), 331-392.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), Attention and Performance XII: The Psychology of Reading. Hillsdale, NJ: Erlbaum.

Frazier, L. & Fodor, J.D. (1978). The sausage machine: A new two-stage parsing model. Cognition, 6, 291-325.

Garnsey, S. M., Pearlmutter, N.J., Myers, E. & Lotocky, M.A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. Journal of Memory and Language, 37, 58-93.

Gillette, J., Gleitman, L., Gleitman, H. & Lederer, A. (1999) Human simulations of vocabulary learning, Cognition, 73, 153-90.

Gleitman, L. (1990). The structural sources of verb learning. Language Acquisition, 1, 3-35.

Gleitman, L.R., Cassidy, K., Nappa, R., Papafragou, A. & Trueswell, J.C. (in press). Hard Words. Language Learning and Development.

Gerkin, L.A. (2002). Early sensitivity to linguistic form. Annual Review of Language Acquisition, 2 1-36.

Gómez, R.L. (2002). Variability and detection of invariant structure. Psychological Science, 13, 431-436.

Gómez, R.L. & Gerkin, L.A. (2000). Infant artificial language learning and language acquisition. Trends in Cognitive Sciences, 4, 178-186.

Hanna, J. Tanenhaus, M.K. & Trueswell, J.C. (2003). The effects of common ground and perspective on domains of referential interpretation. Journal of Memory & Language, 49, 43-61.

Harris, Z. (1957). Co-Occurrence and Transformation in Linguistic Structure. Language 33, 283-340.

Huttenlocher, P. R. & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. Journal of Comparative Neurology, 387, 167-178.

Hurewitz, F., Brown-Schmidt, S., Trueswell, J. & Gleitman, L. (in progress). The contribuation of conversation goals and verb-preferences to children's syntactic decisions.

Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L. & Trueswell, J. (2000) One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. Journal of Psycholinguistic Research, 29, 597-626.

Jackendoff, R. (2002). Foundations of Language. Oxford: Oxford University Press.

Juliano C. & Tanenhaus, M. (1994). A constraint-based lexicalist account of the subject/object attachment preference. Journal of Psycholinguistic Research, 23(6), 459-471.

Karmiloff-Smith, A. (1979). A functional approach to child language : a study of determiners and reference. New York: Cambridge University Press.

Kelly, M. H; Martin, S. (1994). Domain-General Abilities Applied to Domain-Specific Tasks: Sensitivity to Probabilities in Perception, Cognition, and Language Lingua, 92, 105-140.

Keysar, B., Barr, D.J., Balin, J.A. & Brauner, J.S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. Psychological Science, 11, 32-37.

Keysar, B., Barr, D.J. & Horton, W.S. (1998). The egocentric basis of language use: Insights from a processing approach. Current Directions in Psychological Science, 7, 46-50.

Kim, A., Srinivas, B. & Trueswell, J.C. (2002). The convergence of lexicalist perspectives in psycholinguistics and computational linguistics. In P. Merlo and S. Stevenson (Eds.). Sentence Processing and the Lexicon: Formal, Computational and Experimental Perspectives, pp. 109-135. Philadelphia, PA: John Benjamins Publishing.

Larson, R. & Siegal, G. (1995). Knowledge of meaning. Cambridge, MA: MIT Press.

Leslie, A.M. (2000). Theory of mind as a mechanism of selective attention. In M. Gazzaniga (ed.), The New Cognitive Neurosciences, 2nd ed.,1235-1247. . Cambridge, MA: MIT Press.

Lyons, C.G. (1980). The meaning of the english definite article. In Johan Van der Auwera, (Ed.), The Semantics of Determiners, pp. 81-96.

Lyons, C.G. (1999). Definiteness. Cambridge: Cambridge University Press.

MacDonald, M.C., Pearlmutter, N.J. & Seidenberg, M.S. (1994). The lexical nature of syntactic ambiguity resolution. Psychological Review, 101, 676-703.

Marcus, M.P., Santorini, B. & Marcinkiewicz, M.A. (1993). Building a large annotated corpus of English: The Penn Treebank. Computational Linguistics, 19, 313-330.

Maratsos, M.P. (1976). The use of definite and indefinite reference in young children : an experimental study of semantic acquisition. New York: Cambridge University Press.

Mendelsohn, A. (2002). Individual differences in ambiguity resolution: working memory and inhibition. Ph.D. Dissertation, Department of Psychology, Northeastern University.

Mendelsohn, A. (2003) The role of inhibition in sentence processing as demonstrated by a new task. Paper presented at the CUNY Conference on Human Sentence Processing, Cambridge, MA, March 27-29, 2003.

Mintz, T. H., Newport, E. L., and Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children. Cognitive Science, 26, 393-424.

Nadig, A.S. & Sedivy, J.C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. Psychological Science, 13(4), 329-336.

Naigles, L.G., Gleitman, L.R., and Gleitman, H (1993) Children acquire word meaning components from syntactic evidence, in Language and cognition: a developmental perspective, E. Dromi (ed.), Norwood NJ: Ablex.

Naigles, L.G., Fowler, A. & Helm, A. (1992). Developmental shifts in the construction of verb meanings. Cognitive Development, 7(4), 403-427.

Naigles, L.G., Fowler, A. & Helm, A. (1995). Syntactic bootstrapping from start to finish with special reference to Down syndrome. In M. Tomasello & W.E. Merriman (Eds). Beyond names for things: Young children's acquisition of verbs. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc., pp. 299-330.

Novick, J.M., Kim, A. & Trueswell, J.C. (2003). Studying the grammatical aspects of word recognition: Lexical priming, parsing and syntactic ambiguity resolution. Journal of Psycholinguistic Research, 32(1), 57-75.

Novick, J.M., Trueswell J.C. & Thompson-Schill, S. (submitted). Toward the neural basis of parsing: Prefontal cortex and the role of selectional processes in language comprehension. Manuscript in preparation.

Novick, J.M., January, D. Trueswell J.C. & Thompson-Schill, S. (in progress). Individual differences in selection, inhibition and syntactic revision. Work in progress.

Rayner, K., Carlson, M. & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing. Journal of Verbal Learning and Verbal Behavior, 22, 358-374.

Papafragou, A. & Tantolou, N. (in press). Children's computation of implicatures. Language Acquisition.

Saffran, J. R. (2001). Words in a sea of sounds: The output of statistical learning. Cognition, 81, 149-169.

Saffran, J. R. (2002). Constraints on statistical language learning. Journal of Memory and Language, 47, 172-196.

Saffran, J.R., Aslin, R.N. & Newport, E.L. (1996). Statistical learning by 8-month old infants. Science, 274, 1926-1928.

Snedeker, J., Thorpe, K. & Trueswell, J.C. (2002). On choosing the parse with the scene: The role of visual context and verb bias in ambiguity resolution. Proceedings of the 22nd Annual Conference of the Cognitive Science Society, Edinburgh, Scotland.

Snedeker, J. & Trueswell, J.C. (in press). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*..

Spivey-Knowlton, M. & Sedivy, J. Parsing attachment ambiguities with multiple constraints. Cognition, 55, 227-267.

Spivey, M. J. & Tanenhaus, M.K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. Journal of Experimental Psychology: Learning, Memory & Cognition. 24(6), 1521-1543.

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.

Srinivas, B. & Joshi, A.K. (1999). Supertagging: An approach to almost parsing. Computational Linguistics, 252(2), 237-265.

Steedman, M. (2000). The Syntactic Process, Cambridge, MA: MIT Press/Bradford Books.

Stone, M. & Webber, B. (1998). Textual economy through close coupling of syntax and semantics. In Proceedings of INLG, pp. 178-187.

Swingley, D., Pinto, J.P., and Fernald, A. (1999). Continuous processing in word recognition at 24 months. Cognition, 71, 73-108.

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.K., Spivey-Knowlton, M.J., Eberhard, K.M. & Sedivy, J.C. (1995). Integration of visual and linguistic information in spoken language comprehension. Science, 268, 1632-1634.

Tanenhaus, M.K. Hanna, J. & Chambers, C. (in press). Referential domains in spoken language comprehension: Using eye movements to bridge the product and action traditions. In In J. M. Henderson & F. Ferreira (Eds.), The Interface of Language, Vision, and Action: Eye Movements and the Visual World, New York: Psychology Press.

Thompson-Schill, S.L., Jonides, J., Marshuetz, C., Smith, E.E., D'Esposito, M., Kan, I. P., Knight, R. T. & Swick D. (2002). Effects of frontal lobe damage on interference effects in working memory. Cognitive, Affective & Behavioral Neuroscience, 2, 109-120.

Trueswell, J.C. (1996). The Role of Lexical Frequency in Syntactic Ambiguity Resolution. Journal of Memory and Language, 35, 566-585.

Trueswell J.C. & Gleitman, L. (in press). Children's eye movements during listening: Developmental evidence for a constraint-based theory of sentence processing In In J. M. Henderson & F. Ferreira (Eds.), The Interface of Language, Vision, and Action: Eye Movements and the Visual World, New York: Psychology Press.

Trueswell, J.C. & Kim, A.E. (1998). How to prune a garden-path by nipping it in the bud: Fast-priming of verb argument structures. Journal of Memory and Language, 39, 102-123.

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. (1991). Tense, temporal context and syntactic ambiguity resolution. Language and Cognitive Processes, 6(4), 303 - 338.

Trueswell, J.C. & Tanenhaus, M.K. (1994). Toward a lexicalist framework for constraint-based syntactic ambiguity resolution. In C. Clifton, K. Rayner & L. Frazier (Eds.), Perspectives on Sentence Processing. Hillsdale, NJ: Erlbaum.

Trueswell, J. C., Tanenhaus, M. K. & Garnsey, S. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. Journal of Memory and Language, 33, 285 –318.

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(3), 528-553.

Wexler, K. (2003). Maximal Trouble. Paper presented at the CUNY Human Sentence Processing Conference, Boston, MA.

Zelazo, P.D. & Reznick, J.S. (1991). Age-related asynchrony of knowledge and action. Child Development. Vol 62(4), 719-735.

Figure Captions

Figure 1. The syntactic alternatives associated with the ambiguous word "with." Shown here using the LTAG formalism (Srinivas & Joshi, 1999).

Figure 2. One-referent (a) and two-referent (b) scenes, from Trueswell, Sekerina, Hill and Logrip (1999).

Figure 3. *Pull the rabbit out of the hat.* (Adapted from Stone & Webber, 1998.)

Figure 4. Clip-art illustrations of 1- and 2-Referent scenes from Snedeker, Thorpe & Trueswell (2001; Snedeker & Trueswell, in press). Physical objects were used in the study.

Figure 5. Data from adults (N=24). Proportion of trials with instrument actions (left) and instrument looks (right). Snedeker and Trueswell (in press).

Figure 6. Data from five year olds (N=48). Proportion of trials with instrument actions (left) and instrument looks (right). Snedeker and Trueswell (in press).

Figure 7. Illustration of objects presented to child in example stimuli from Hurewitz et al. (in prep.)

Figure 8. Proportion of trials in which five year olds gave NP Modification response (e.g., "No, it was this cat, the one on the book.") (Data from first block only.) Hurewitz et al., in prep.

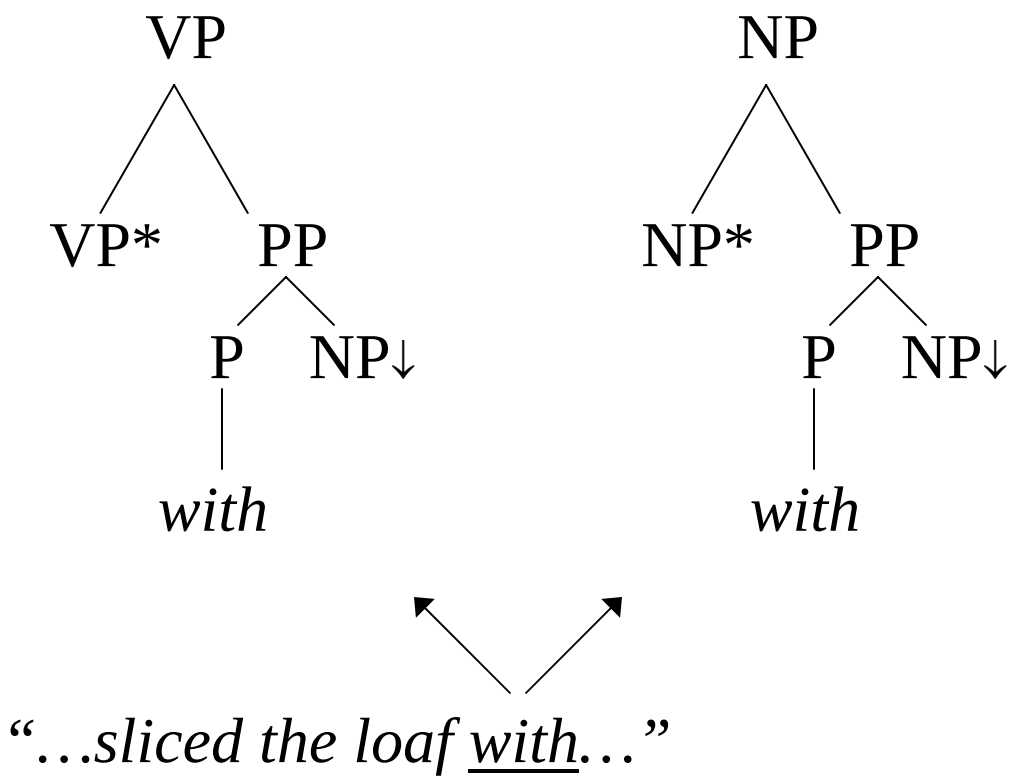

Figure 1

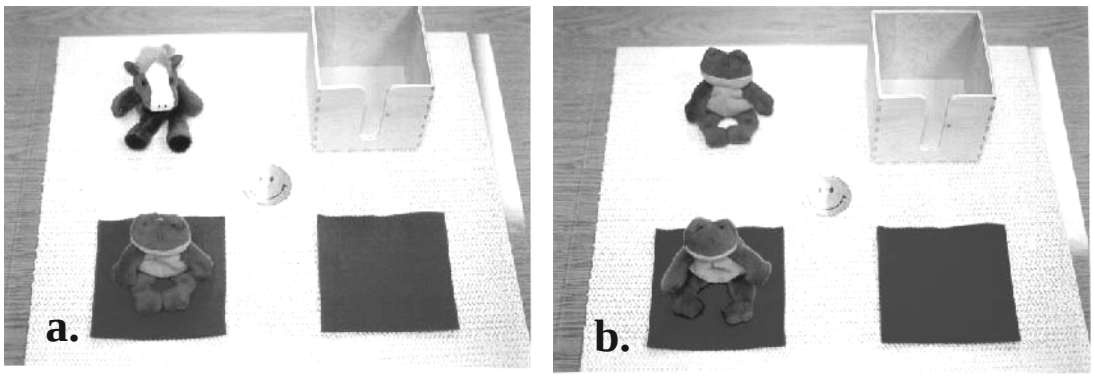

Figure 2

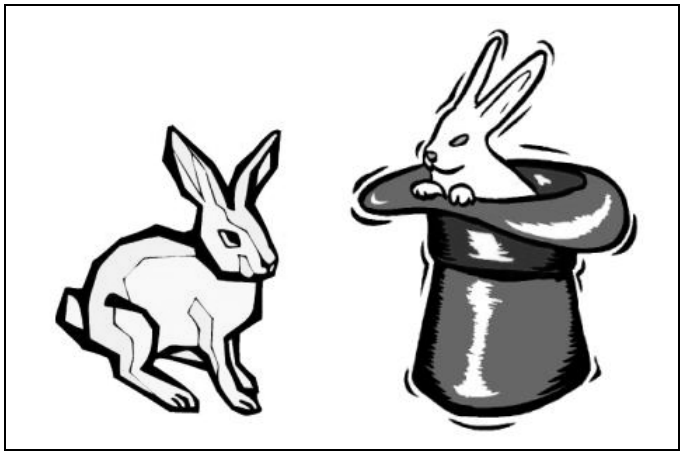

Figure 3

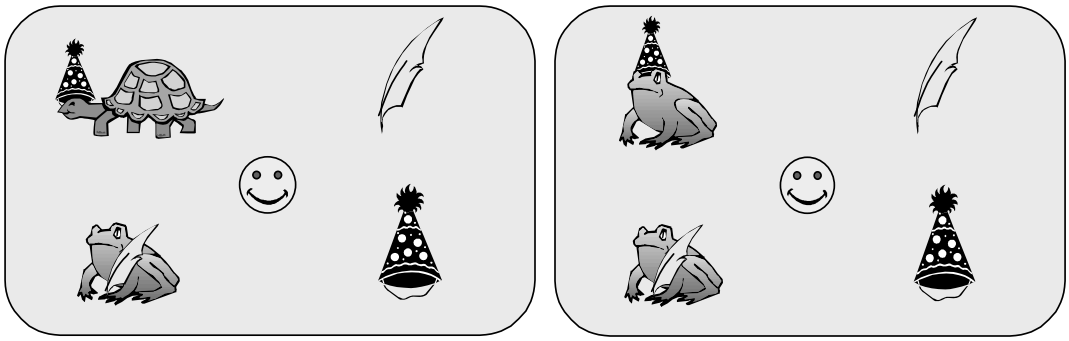

Figure 4

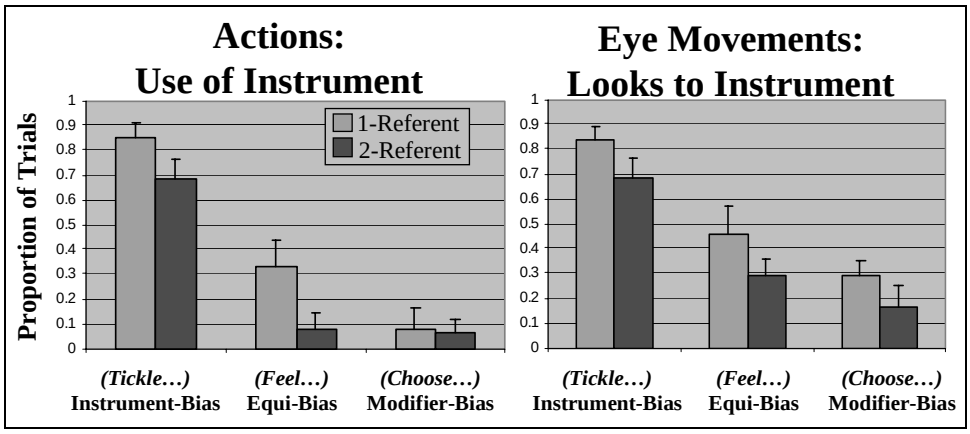
Actions:
Use of Instrument
Eye Movements:
Looks to Instrument
Proportion of Trials
1-Referent
2-Referent
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
(Tickle…)
Instrument-Bias
(Feel…)
Equi-Bias
(Choose…)
Modifier-Bias
(Tickle…)
Instrument-Bias
(Feel…)
Equi-Bias
(Choose…)
Modifier-Bias

Figure 5

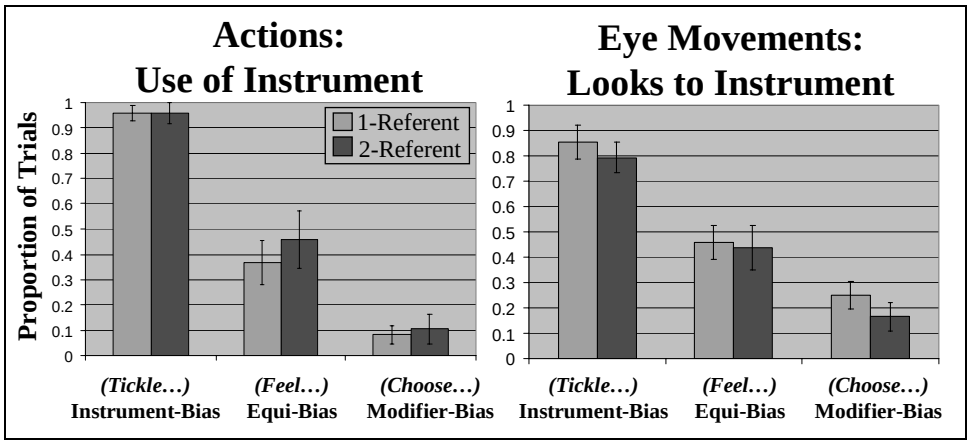
Actions:
Use of Instrument
Eye Movements:
Looks to Instrument
Proportion of Trials
1-Referent
2-Referent
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
(Tickle…)
Instrument-Bias
(Feel…)
Equi-Bias
(Choose…)
Modifier-Bias
(Tickle…)
Instrument-Bias
(Feel…)
Equi-Bias
(Choose…)
Modifier-Bias

Figure 6

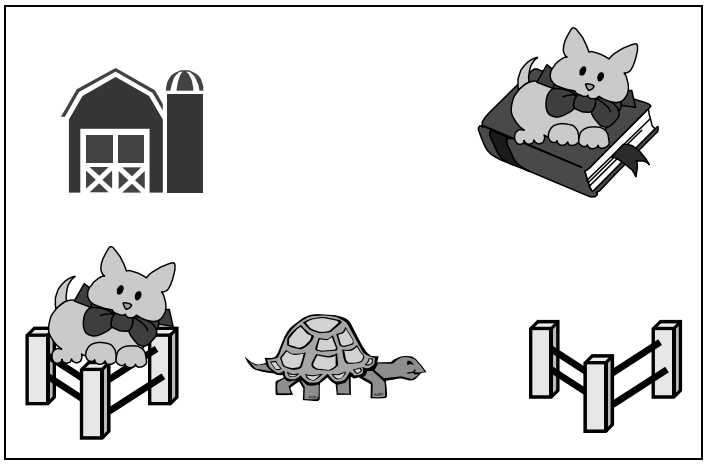

Figure 7

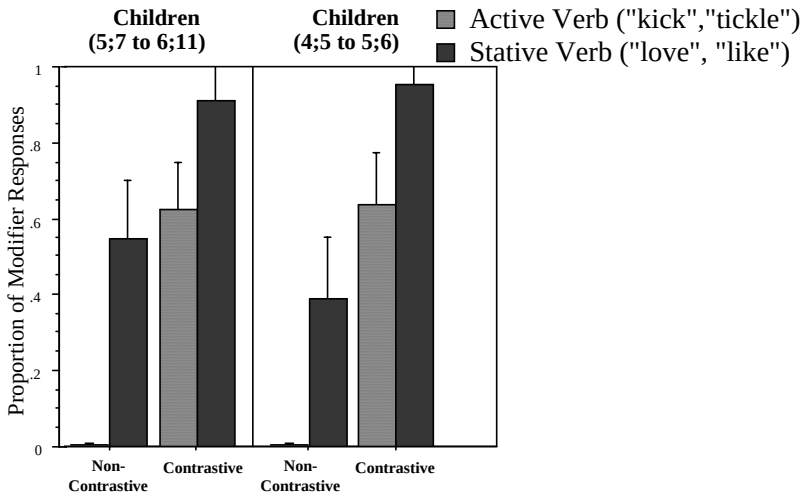
Children
(5;7 to 6;11)
Children
(4;5 to 5;6)
Active Verb ("kick","tickle")
Stative Verb ("love", "like")
Proportion of Modifier Responses
1
.8
.6
.4
.2
0
Non-
Contrastive
Contrastive
Non-
Contrastive
Contrastive

Figure 8