10

Children's Eye Movements during Listening: Developmental Evidence for a Constraint-Based Theory of Sentence Processing

JOHN TRUESWELL
LILA GLEITMAN

Many comprehension studies of grammatical development have focused on the ultimate interpretation that children assign to sentences and phrases, yielding somewhat static snapshots of children's emerging grammatical knowledge. Studies of the dynamic processes underlying children's language comprehension have to date been rare, owing in part to the lack of online sentence processing techniques suitable for use with children. In this chapter, we describe recent work from our research group, which examines the moment-by-moment interpretation decisions of children (age 4 to 6 years) while they listen to spoken sentences. These real-time measures were obtained by recording the children's eye movements as they visually interrogated and manipulated objects in response to spoken instructions. The first of these studies established some striking developmental differences in processing ability, with the youngest children showing an inability to use relevant properties of the referential scene to resolve temporary grammatical ambiguities (Trueswell, Sekerina, Hill, & Logrip, 1999). This finding could be interpreted as support for an early encapsulated syntactic processor that has difficulty using non-syntactic information to revise parsing commit-

ments. However, we will review evidence from a series of follow-up experiments which suggest that this pattern arises from a developing interactive parsing system. Under this account, adult and child sentence comprehension is a "perceptual guessing game" in which multiple statistical cues are used to recover detailed linguistic structure. These cues, which include lexical-distribution evidence, verb semantic biases, and referential scene information, come "online" (become automated) at different points in the course of development. The developmental timing of these effects is related to their differential reliability and ease of detection in the input.

INTRODUCTION

This chapter describes much of what is currently known about how young children go about interpreting the sentences that they hear against their surrounding real-world environments. As we will describe further, we have the advantage of being able, by recording children's eye movements, to measure their moment-to-moment visual attention to objects in the world while they hear spoken sentences unfolding over time. This eye-gaze-during-listening paradigm was originally developed by Tanenhaus and colleagues to study adults' sentence comprehension abilities (e.g., Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Sedivy, Tanenhaus, Chambers, & Carlson, 1999; cf. Cooper, 1974). As discussed extensively in this volume, the basic idea behind this paradigm is that by measuring how visual-attentional states line up in time with the successive arrival of words and phrases, researchers can gain insight into the real-time processes by which listeners organize sentences structurally and semantically and how they map these representations onto the events and objects that they denote (Tanenhaus & Spivey-Knowlton, 1996). To accept this link between data and interpretation, one need only believe that, to a useful approximation, the mind is going where the eye is going.[1]

How the mechanics of "language processing," so studied and described, relate to more static descriptions of "language knowledge" or "knowledge of grammar" is a hotly debated topic. At one extreme, some investigators have held that these real-time processes (how the listener comes to understand particular instances of language use) impose only minimal constraints on the theory of mental grammar (a person's relatively stable knowledge of the design features of a language). At the opposite end of the theoretical continuum, other investigators hold that linguistic representations are highly constrained by the form of their use (linearly in time, word-by-word), such that, at the extreme, a single theory describes both "knowledge" and "use" of language.

Until quite recently, such questions concerning the architecture of language knowledge could not even be raised realistically for the case of

young children. After all, most sentence-processing techniques relied on the experimental subject's ability to deal comfortably with written text, techniques that are transparently unsuitable for preschoolers and even novice readers in the early school years. Eye-movement techniques that link responses to *speech events* have the potential to revolutionize how we examine the child's emerging understanding of language. Accordingly, we and our colleagues have adapted this online experimental paradigm for use with children as young as 4 years of age.

In the present paper we will first briefly review the results and theoretical approaches that have emerged from the adult sentence-comprehension literature, based on such real-time measures as eye tracking. We will then describe the extension of such techniques and findings to children—to the question of how one learns to parse. In doing so, we review both the original study that introduced eye-tracking techniques and their rationale for this age range (Trueswell et al., 1999) and several later studies that have clarified and refined interpretation of the original findings. Our most general ambition is to understand how the processes of language use and language learning interact in the young child to yield the mature state that all normal language-using humans attain.

HOW ADULTS RECOVER GRAMMATICAL INFORMATION FROM AN UTTERANCE: THE CONSTRAINT-BASED LEXICALIST ACCOUNT

In order to understand the intended meaning of a sentence, a reader or listener must detect much or all of the sentence's grammatical structure. This is because the grammatical operations of a language convey complex combinatory and referential meaning that single words cannot. By finding clues in the sentence about the grammatical operations that gave rise to it, a reader or listener is in a position to recover the intended meaning of the sentence as a whole. Adults are known to be marvelously adept at this process. In fact, numerous studies show that adult listeners and readers are so skilled that they typically achieve sentence interpretations in real time, packaging words into phrases and making provisional commitments to interpretation as each word is perceived (e.g., Altmann & Steedman, 1988; Frazier & Rayner, 1982; Marslen-Wilson & Tyler, 1987; Trueswell, Tanenhaus, & Garnsey, 1994). In this straightforward sense, sentence understanding comes about through a process that is "immediate" and "incremental." Each word makes its contribution to the interpretation at the point of its occurrence in the flow of speech and influences the interim structure and interpretation of the sentence that is being built.

An important implication of this online nature of processing is that readers and listeners sometimes make interpretation errors that require revision

when further words in the sentence are encountered. In this sense they are "led down the garden path" at a point of ambiguity and only later make their way back to the intended interpretation. This garden-path phenomenon is something of a gift from nature to the researcher interested in sentence processing. This is because it can be used as a means for examining which kinds of evidence (syntactic, semantic, discourse context) inform initial commitments to interpretations, thus providing insight into the internal organization of the comprehension system. Investigations of this kind comprise a broad experimental effort in which sources of evidence are parametrically manipulated and their effect on parsing preferences is measured. For example, consider the following sentence:

(1) Anne hit the thief with the stick.

There are two primary ways that we can interpret the prepositional phrase *with the stick:* either as the Instrument with which the action is performed, in which case it is structurally linked to the verb phrase (VP) (*to hit using a stick*—VP attachment[2]), or as a modifier of the direct object, in which case it is syntactically linked to that noun phrase (NP) (*the thief that has the stick*—NP attachment). It has been confirmed experimentally that readers and listeners have a tendency to commit to the so-called instrument interpretation when encountering the ambiguous preposition *with* in sentences like (1) (e.g., Rayner, Carlson, & Frazier, 1983; Taraban & McClelland, 1988). This finding is consistent with our intuition of a misinterpretation, or garden-path, when we encounter the final word in such sentences as:

(2) Anne hit the thief with the wart.

Many early studies of readers' eye-movement patterns for this and other ambiguities suggested that the comprehension device had general structural biases (e.g., for VP-attachment) that were initially uninfluenced by nonsyntactic facts such as plausibility (warts vs. sticks) or situation-specific discourse cues (e.g., Rayner et al. 1983; Ferreira & Clifton, 1986; Ferreira & Henderson, 1991). The findings supported modular theories of parsing, whose initial stage rapidly structured the input words (each represented by its lexical class label, e.g., noun, preposition) based solely on phrase structure rules of the language and resolved ambiguities using a syntactic simplicity metric —the Minimal Attachment Principle (Frazier, 1987, 1989; Frazier & Fodor, 1978).

However, during the past 20 years, an accumulating body of evidence on the pattern of garden paths supports a different characterization of the processing system, in which far more detailed and probabilistic linguistic properties of the input are tracked, detected, and used incrementally. First, it has been found that parsing commitments are often guided by the syntactic pref-

erences of individual lexical items, such as whether a particular verb frequently denotes an instrument via a prepositional phrase (Taraban & McClelland, 1988; Trueswell & Kim, 1998; MacDonald, 1994; Trueswell, Tanenhaus, & Kello, 1993). To see this point, compare the likely interpretation of sentence (1) when the verb *noticed* is substituted for *hit*. Second, the semantic fit of preceding complements also rapidly constrains initial parsing commitments, especially when the information precedes the ambiguity (e.g., readers assume a main clause analysis of *The defendant examined…*, but a relative clause analysis for *The evidence examined…*, see Tabossi, Spivey-Knowlton, McRae, & Tanenhaus, 1994; Trueswell et al., 1994).

Third, constraints from the immediate referential context can influence the course of parsing decisions. For instance, in sentence (1), a two-thief context encourages an initial modifier interpretation of *with the stick* (e.g., Altmann & Steedman, 1988; Crain & Steedman, 1985; Trueswell & Tanenhaus, 1991). This is because listeners realize that NPs are most often modified so as to distinguish between referential alternatives made available by either the discourse or situational context. Uttering a plain vanilla *thief* is sufficient for identification when there's only a single criminal in sight; but if there are two or more, then the modified *thief with the wart* is needed to select the guilty party. Finally, the measured effectiveness of these probabilistic constraints on interpretation depends upon whether the verb makes available the appropriate syntactic alternatives (e.g., Britt, 1994; Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Trueswell, 1996).

The theory that we would argue best captures the existing evidence is the constraint-based lexicalist theory (henceforth, CBL: see MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus, 1994). This theory assumes a constraint-satisfaction approach to ambiguity resolution (Marslen-Wilson & Tyler, 1987; McClelland, 1987), in which multiple sources of information are used to converge as rapidly as possible on a single interpretation. The central component of this theory is a grammatical processing system that is highly tuned to the structural preferences of individual lexical items; hence "lexicalist." The recognition of a word includes activation of rich argument structures that define the initial set of possible interpretations. For example, the preference for VP attachment in sentences 1 and 2 is explained as arising from a system that is sensitive to the grammatical preferences of the verb "hit," which include the use of the instrument role, typically headed by the preposition "with" (for related evidence, see Taraban & McClelland, 1988; Garnsey et al., 1997; MacDonald, 1994; Trueswell et al., 1994; Britt, 1994). It is significant that linguistic (Bresnan & Kaplan, 1982; Pollard & Sag, 1987) and computational linguistic (Joshi, Vijay-Shanker, & Weir, 1991; Steedman, 1995; Srinivas & Joshi, 1999) theorizing have also in the same period been converging toward a lexicalized picture of the organization of language (Kim, Srinivas, & Trueswell, 2002). Relatedly, recent theories of the acquisition of word

meaning give independent support for a lexically organized grammar (Landau & Gleitman, 1985; Gleitman, 1990; Fisher, 1996; Gillette, Gleitman, Gleitman, & Lederer, 1999).

DEVELOPMENTAL PREDICTIONS FROM THE CONSTRAINT-BASED LEXICALIST THEORY

With the adult processing patterns and CBL account in mind, we sketch here a developmental version of the CBL theory of sentence processing.[3] First, this theory assumes that the computational procedures involved in sentence comprehension remain constant over development. This means that from a very early age, the learner's comprehension machinery can best be characterized as a perceptual guessing game in which multiple probabilistic cues are used to converge on the grammatical operations that gave rise to the utterance. The assumption that the analysis of the input is driven by statistical learning mechanisms has experimental support. Very young infants, when exposed to 2 to 3 minute sequences of artificially generated syllable sequences, begin to recover the phonological and grammatical tendencies of the language (e.g., Saffran, Newport, & Aslin, 1996; Marcus, 2000; see also Hudson & Newport, 2003). It seems plausible, therefore, especially in the light of adult processing findings above, that statistical tracking is accomplished at multiple levels of utterance representation, and this forms a deep continuity between learning and comprehension processes over the course of the language system's development.

Under this assumption, the CBL theory can be used to make developmental predictions about the development of sentence processing abilities, in particular about developmental differences in how children might resolve temporary syntactic ambiguity. Below, we sketch three such predictions, which will be examined in the child eye movement studies described in this chapter.

Prediction 1: Early Reliance on Lexical Information. Given the probabilistic nature of grammatical processing in CBL, the theory most naturally predicts that the degree of informativity of various sources of evidence (e.g., lexical, semantic, referential) should predict the order in which these cues "come online," i.e., become automatized for use in syntactic ambiguity resolution. Of the sources of evidence we have discussed, the CBL theory most naturally predicts a developmentally early reliance on lexical cues to structure. The reason for this is that lexical information (in the form of verb-argument information, for instance) highly determines the likelihood of local structure, a fact that has been noted time and again by linguists (e.g., Carlson & Tanenhaus, 1988; Jackendoff, 2002), psycholinguists (e.g., Fisher, Gleitman, & Gleitman, 1991) and computational linguists constructing lexically contingent parsing systems (e.g., Srinivas &

Joshi, 1999). Indeed, in the adult system we see evidence for a bias to rely on such information when resolving syntactic ambiguities, which is of course rapidly constrained by other evidence, such as plausibility and referential information (Britt, 1994; Garnsey et al., 1997). Thus, we should expect that the distribution of lexical items (and their meaning) should overly constrain children's parsing commitments to ambiguity, more so than they would for older children and adults, who have learned to recruit referential cues to structure. As experience builds knowledge of how contextual facts predict structure in the utterance, these sources of evidence should begin to exert effects on online parsing commitments.

Prediction 2: Interactive Processing When Cues Are Highly Constraining. Given our interactive processing account, it should in principle be possible to find children using multiple evidential cues, including nonlinguistic cues, to resolve ambiguity, provided that they highly constrain the possible structure of the utterance. That is, evidence for reliance on a particular cue to structure, such as lexicosyntactic information, should not be understood as an early ban on other sources of evidence for resolving ambiguity. Easy to track nonlinguistic cues that are highly constraining should be predicted to combine with lexical evidence, even in younger children.

Prediction 3: Comprehension-Specific Deficits. Our comprehension theory assumes that multiple sources of evidence automatically combine to influence a child's estimate of the intended meaning of a sentence. Thus, it makes the somewhat counterintuitive prediction that children who readily produce particular syntactic structures in appropriate referential contexts could very well show an inability to detect and understand these very same structures in a comprehension setting where evidential sources support an alternative analysis.

CHILDREN'S COMPREHENSION IN REAL TIME

The launch point of our research endeavor is a pair of eye movement studies that we conducted with 5-year-olds (N = 16) and 8-year-olds (N = 16), first reported in Trueswell et al. (1999). In this study, which was modeled after an adult sentence-parsing study by Spivey, Tanenhaus, and colleagues (Tanenhaus et al., 1995; Spivey, Tanenhaus, Eberhard, & Sedivy, 2002), children's eye movements were recorded as they acted upon spoken instructions to move objects in an array (see Figure 10.1). The participants in Trueswell et al. (1999) heard spoken instructions, which on critical trials contained a temporary prepositional phrase- (PP-) attachment ambiguity, as in

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

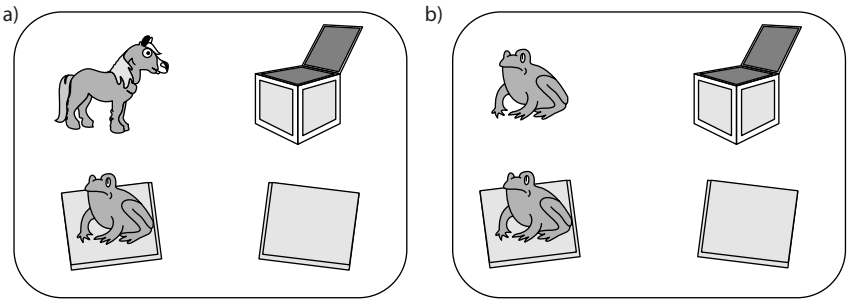

FIGURE 10.1. Illustrations of (a) 1-Referent scene and (b) 2-Referent scene. In both cases, children look to the empty napkin upon hearing "on the *napkin*", showing a strong VP-attachment bias. The initial commitment also influences the final interpretation: most children move a frog to the empty napkin, regardless of context. (Truswell et al., 1999)

Here the phrase *on the napkin*, when first encountered, could in principle link to the verb *put* as a goal, indicating where to put a frog, or link to the NP *the frog* as a modifier, specifying a property of a particular frog. The phrase though is disambiguated toward the modifier interpretation by the presence of a second goal phrase (*in the box*).

The striking finding was that 5-year olds showed a strong preference to interpret *on the napkin* as the goal of *put*, even when the referential scene supported a modifier interpretation (e.g., two frogs, one on a napkin, see Figure 10.1b). Upon hearing *on the napkin*, 5-year olds typically looked over to a potential goal in the scene, the empty napkin, regardless of whether there were two frogs present (supporting a modifier interpretation) or one frog present (supporting a goal interpretation). In fact, 5-year olds' preference for VP-attachment was so strong that they showed little sign of revising it: upon hearing *napkin,* children would look to the empty napkin as a potential goal and then frequently move a frog to that location. In two-referent cases, children were at chance even when selecting which frog to move, suggesting they never considered a modifier interpretation.

Importantly, this child parsing behavior was localized to the ambiguity and not to the complexity of the sentence. This is shown by the fact that 5-year-olds' eye movements and actions became adult-like when the temporary ambiguity was removed, as in the unambiguous modifier form:

(4) Put the frog that's on the napkin in the box.

The near-perfect performance on unambiguous forms rules out a more mundane explanation of our results, namely that long "complicated" sentences

puzzle young children. Here, sentence 4, an even longer sentence with the same intended structure, doesn't flummox the child. Why? Because we have removed the temporary ambiguity.

In contrast to 5-year-olds, adults' responses to the temporarily ambiguous stimuli were found to depend on the referential scene provided. In particular, the mere presence of a two-referent scene eliminated measurable signs of syntactic misanalysis of the ambiguous phrase: there were few looks to the potential goal and few incorrect actions, as compared to one-referent scenes. This finding is consistent with the earlier work on adults with this ambiguity (Tanenhaus et al., 1995; Spivey et al., 2002).

Thus, as CBL theory predicts, younger children appear to be relying more heavily on local lexical cues than referential scene cues to resolve syntactic ambiguity. Specifically, younger children appear to be relying heavily on the grammatical preferences of the verb *put* which almost always denotes a goal, typically as a prepositional phrase.[4] The finding displays a striking paradox: Young children are often presumed to be very well tuned to the semantics and pragmatics of scenes and events in the extralinguistic world and far less attuned to the formal (presumably "boring") formal facts about words and sentences. Yet these initial findings have been that it is the adults who are more sensitive to the referential scene as a guide to parsing commitments, whereas the children seem—on the surface of the matter—to be little lexicalist grammarians, parsing the heard sentence in fine indifference to the reference world.

THE ROLE OF VERB BIAS IN CHILD- AND ADULT-PARSING PREFERENCES

Snedeker, Thorpe, & Trueswell (2001), and Snedeker & Trueswell (submitted) explored in detail the claim that children's parsing preferences are driven by their verb-specific syntactic and semantic knowledge. Such an account predicts that child-parsing preferences for ambiguous phrases should be heavily influenced by manipulations of the type of verb, pursuing the analysis that is consistent with the most likely argument structure for that verb. Recall that this is not a guaranteed outcome, for a general structural bias could also account for the original Trueswell et al (1999) child-parsing outcomes. Certain acquisition theories make this prediction, proposing that the child parser might avoid more complex syntactic structures in favor of simpler ones (e.g., Minimal Attachment Principle, Goodluck, & Tavakolian, 1982; Frazier & Fodor, 1978) or place a general ban on complex syntactic operations (e.g., the No-Adjoin Principle, Frank, 1998), predicting little or no influence of verb bias.

We were also interested in better understanding a paradox in the adult literature regarding verb-bias information. Whereas adult-reading studies

indicated an important role for verb biases in the initial consideration of syntactic alternatives (Britt, 1994), studies of adult listeners in world-situated, eye-gaze studies suggest an almost exclusive role for referential cues in determining initial syntactic choices (Tanenhaus et al., 1995; Spivey et al., 2002). Almost all constraint-satisfaction views of parsing, including our own, would expect some difficulty with *put* instructions, like the ones used in those studies. If, in the adult system, both lexical and referential facts were weighed simultaneously (as supported by reading studies), why would referential cues completely eliminate adult difficulty with *put* instructions, which are heavily lexically biased toward the incorrect destination analysis?

To meet these two exploratory goals, Snedeker et al. (2001) followed the lead of the prior adult-reading studies that have, in a single experiment, fully crossed verb bias with manipulations of referential context (e.g., Britt, 1994), except that we performed these manipulations in the world-situated, eye-gaze task. Such manipulations should reveal the relative contributions of these several factors under all possible combinations, for both adults (N = 36) and 5-year olds (N = 36). Target constructions contained a PP-attachment ambiguity (e.g., "Feel the frog with the feather") in both two-referent and one-referent contexts (see Figure 10.2 below; clip-art images are provided for illustrative purposes only—real-world, beanie-baby objects were actually employed, as in our previous study). Linguistic materials were prenormed, and we compared three different types of verbs: those that typically take an instrument phrase (*hit*), those that rarely do so (*choose*), and equibiased verbs (*feel*). The semantic fit of the instrument noun was controlled across conditions via normative ratings: all nouns, e.g., *fan, feather,* and *stick,* were rated as being approximately equally good-or-poor instruments for their respective verbs.

The results were systematic and striking. Five-year olds' eye movements and actions showed a sole reliance on the verb preferences, almost perfectly

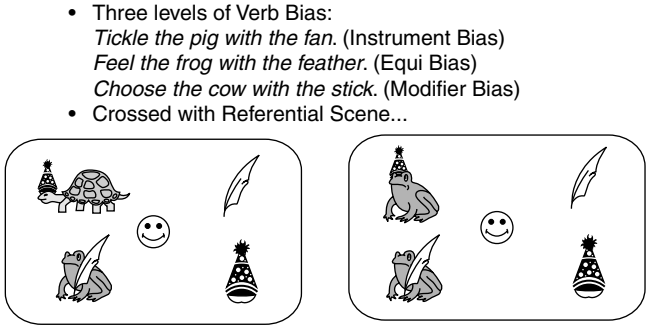

FIGURE 10.2. Manipulating both verb-bias and referential scene. From Snedeker, Thorpe, & Trueswell, 2001.

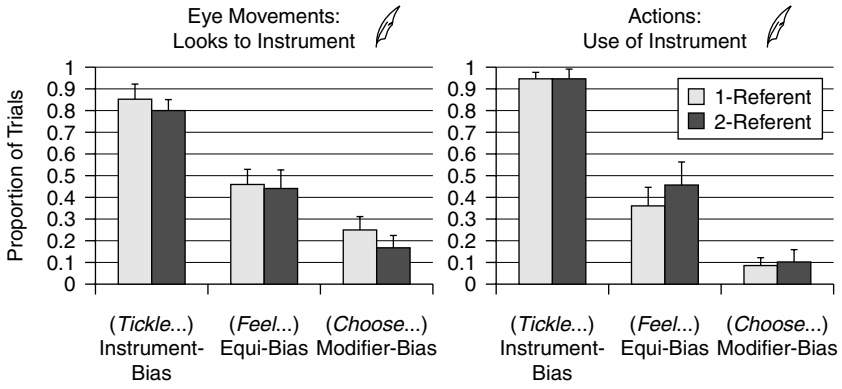

FIGURE 10.3. Five-year-olds (N = 36) show sensitivity to verb bias and not referential scene manipulation (Snedeker et al., 2001).

matching rates of instrument use to the semantic preferences of individual verbs. As shown in Figure 10.3, the proportion of looks to the potential instrument upon hearing "with the X" systematically decreased across instrument-biased, equibiased and modifier-biased conditions. Additionally, absolutely no sensitivity to the referential scene was observed, even for equibiased verbs. These means reflected a reliable effect of verb type, no effect of referential scene, and no interaction in subject and item ANOVAs. In contrast, adults' initial eye movements and actions showed simultaneous sensitivity to both verb-bias manipulations and referential context in the expected directions: Two-referent scenes and modifier-biased verbs both reduced looks to, and use of, a potential instrument (e.g., a large feather), resulting in reliable effects of both the verb type and referential factors.

These findings are in line with what is predicted by a CBL parsing system that gradually recruits reliable information over development. As the processing exposure with modifier expressions increases, children become increasingly sensitive to scene-contingent factors for their use.

It is important to note that although two-referent scenes greatly decreased instrument looks and instrument actions in adults, there were still effects of verb bias in two-referent scenes. That is, instrument-biased verbs resulted in adults considering the contextually *inappropriate* instrument interpretation significantly more often than in unambiguous forms. This stands in contrast with adults in our earlier *put…* study and the study of Tanenhaus, Spivey and colleagues, where the same two-referent scene removed all signs of the contextually inappropriate VP-attachment interpretation. As discussed in the paper, we suggest that the apparent ease that adults show in *put…* instructions comes from the immediately following

disambiguating material (*in the box*), which is integrated into ongoing parallel estimates of parse likelihoods.

Lexical Modularity or Informativeness of the Scene?

Thus far, our experimental work has shown a near-exclusive role of lexical evidence for informing children's parsing decisions. In particular, the child parser does not appear to make consistent use of all potentially relevant cues in the scene to inform parsing decisions. Five-year olds relied on language-internal information such as the syntactic preferences of verbs and the linkages of these preferences to argument-structure interpretation. In contrast, 8-year-olds and adults were measurably influenced not only by these language-internal cues, but also by several other sources of information that were available in the utterance and the situation.

These results in hand, we can now reapproach our original questions: What accounts for the restrictions on early child processing? And what are the mechanisms (or expansions of the knowledge base) that allow these restrictions to be lifted? At least two viable and distinguishable developmental accounts exist.

The Modular/Single-Cue Hypothesis First, contra the account given thus far, it is possible that the observed child-adult differences regarding the child's reliance on lexical cues arise from changes or expansions in processing ability. For instance, a limited, single-cue, or encapsulated, parsing system might become more interactive as processing ability grows with age. Indeed, several current modular theories of parsing propose a special status for lexical cues to structure for architectural reasons, where the lexicon exclusively proposes syntactic and semantic structure (Boland & Cutler, 1996). Our observed 5-year-old preference to use lexical cues might very well reflect an encapsulated, lexicalist parsing system "in the raw," which, as it increases in processing power, becomes increasingly interactive. Similarly, it is possible that even when children control multiple cues, they may only be able to use one cue at a time for ambiguity resolution, perhaps due to an early inability to coordinate multiple cues.

Multiple Cue System from the Start In contrast, our own account assumes a probabilistic multiple cue comprehension system from the start, with the ordering of cue use over development reflecting changes in each cue's relative reliability. As various evidential databases are built and discovered to be informative, they come into use in the comprehension process.[5] Under this account, the child-parsing system shows an earlier reliance on lexical sources (above and beyond other relevant information sources such as referential

sources) precisely because of the high degree of reliability of lexical sources for syntactic structuring. By age 5, the child has learned a great deal about the possible and probable syntactic/semantic environments in which particular verbs can appear—especially, for the common verbs we have tested thus far. Other sources, such as referential scene constraints on syntax, might simply take longer to acquire and use because they are less reliable over the input database as a whole and arguably more difficult to track in particular instances of use than lexicosyntactic contingencies. Thus, the child- parsing system can, in principle, use multiple evidential sources to guide a syntactic choice, but, in practice, the usefulness of particular sources of evidence is a matter of discovery and, hence, changes with experience.

Under this multiple-cue, informativity account, a crucial question becomes whether a child listener can deduce from scene information alone the need for referential specificity, that is, the need for a speaker to provide restrictive modification of a definite NP. In particular, can the child look out into the visual world and easily anticipate 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*)?

Our assumption thus far has been that a speaker's use of a bare definite NP (*the frog*) in the presence of multiple entities of that sort (multiple frogs) ought to be a near-perfect predictor of the need for further linguistically specification, i.e., a post-NP restrictive modifier (e.g., *the frog you caught yesterday*). But is this the case? A recent adult-to-adult referential communication study suggests that there would be only sporadic evidence for this scene-contingent inference (Brown-Schmidt, Campana, & Tanenhaus, 2002). It was observed that adults do not utter restrictive modifiers every time there is more than one potential referent. In particular, 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, conversants' eye movements, actions, and vocal responses all showed that they routinely achieved referential success under these conditions. Obviously, this success isn't evidence for psychic abilities on the part of the conversants. 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 more than one potential referent was currently under discussion.

What this means is that the discourse and its goals relative to the scene, rather than the scene itself, ought to be a far better predictor of restrictive modifier use for the younger child. In prior referential scene-parsing studies (Trueswell et al., 1999; Spivey et al., 2002), adults and older children might have shown a proclivity to use scene evidence as *a proxy for the discourse*, when

such a discourse is absent. Or perhaps humans develop an understanding of how scene cues partially predict structure gradually over developmental time, given the sporadic nature of the cue.

To this end, we have begun to ask whether potentially potent evidence from the discourse can influence 5-year-old parsing decisions (Hurewitz, Brown-Schmidt, Trueswell, & Gleitman, in progress). Here a preceding discourse, in the form two conversing puppets, establishes the referential goal to contrast 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 cues to structure. If a single-cue, modular system is at work, we would expect to see continued use of only lexical cues to structure.

In the study, children (N = 24; Age = 4;0-5;6) were tested in a modified version of the Truth Verification task (Crain & Thornton, 1998). On each trial, the child heard a beanie-baby 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 two referent scene (as in Figure 10.4, 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). The story, prerecorded and acted out by the experimenter (E)

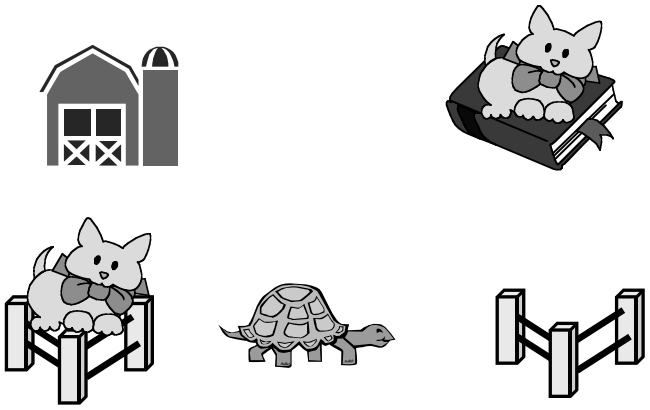

FIGURE 10.4. Illustration of stimulus scene from Hurewitz, Brown-Schmidt, Trueswell, & Gleitman (in progress).

deictically referred to each animal and established the pair of cats in distinct events; it 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 locations, Ms. Rabbit returned to ask Mr. Walrus a question. In all conditions, Walrus's answer contained 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's utterance, however, was preceded by a question from Ms. Rabbit that either supported the need to contrast the cats (the contrastive question condition, "Which cat did the turtle tickle?") or did not support this goal (the noncontrastive question condition, "Can you tell me something about the story?"). In all cases, *both* interpretations of the ambiguity *are incorrect* 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. noncontrastive) was crossed with a verb manipulation. Half the trials involved eventive verbs (such as *tickle*), which easily allow for locative (VP-attached) 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."). These verbs do not usually permit locative modifiers; such modifiers would indicate a particular moment in time that the verb would have to denote going, against the lexical semantics of such verbs.

Given the importance of conversation cues in modifier use (see above), our multiple-cue account predicts that constraining discourse cues (here, question-type), as well as lexical cues (verb-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. In contrast, a modular lexicalist account predicts children will continue to do what they have done in our previous studies: rely on lexical cues. That is, children should show only an effect of the type of verb, with increased modifier responses for stative verbs.

Figure 10.4 plots the rates of VP-attach interpretations exhibited by children in the four conditions. As one can see, the pattern of VP-attach corrections across conditions supports our interactive account. This resulted in reliable effects of question type and verb type (p's < .05). 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. Taken together, these data suggest a general progression toward overcoming local lexical biases that are in conflict with strong discourse requirements.

The Production/Comprehension Dissociation

It should be emphasized that our account of the child parsing phenomenon 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 evidential sources support an alternative analysis.

To test these claims, Hurewitz, Brown-Schmidt, Thorpe, Gleitman & Trueswell (2000) examined 5-year-olds' production and comprehension abilities in two-referent scenes, e.g., two frogs, as in Figure 10.6. Children heard a

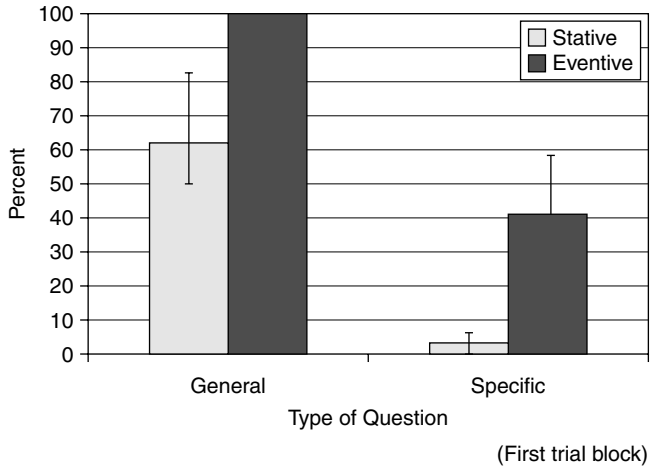

FIGURE 10.5. Proportion of locative (VP-attach) corrections (e.g., "No! It happened in the barn!" (N = 24, Age = 4;0 to 5;6).

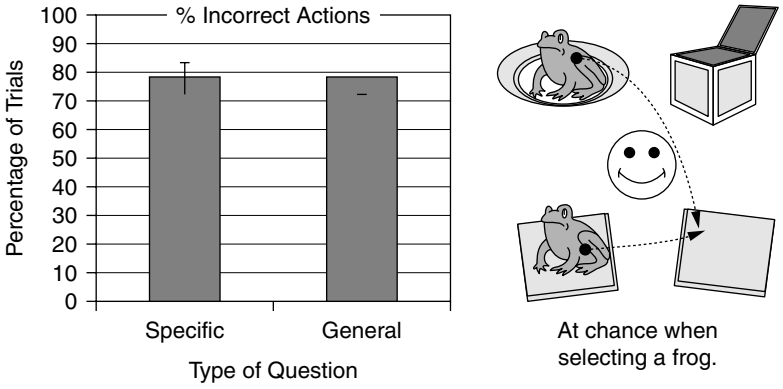

FIGURE 10.6. Children distinguish frogs in their own speech but then garden-path on "Put the frog on the napkin into the box" (Hurewitz et al., 2000).

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 nonlinguistic 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* misanalysis effects as those reported in Trueswell et al. (1999). They typically performed incorrect actions, all of which involved the incorrect destination (see Figure 10.6). 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 eye-tracked the young subjects, with the online results replicating Trueswell et al. (1999).

So, in both of these experiments, we observe, like in the Rabbit-Walrus study of the previous section, 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 swamp the comprehension statistics in favor of VP-attachment interpretations, 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 though 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, Gleitman & Trueswell, in progress). Thus, the data strongly support the automatic use of verb statistics in the young-child parsing system.

THE CONSTRAINT-BASED LEXICALIST LEARNER: A SUMMARY OF FINDINGS

We have presented the results of several experiments which, taken together, support the CBL approach to language comprehension by children during the period when they are constructing the automatic mechanisms for rapid and efficient language understanding, in the age range from 4 to 6 years. All these studies took advantage of the fact, well documented in the adult parsing literature, that the resolution of lexicosyntactic ambiguities can shed light on the internal workings of the comprehension system. An act of comprehension, followed along its course with real-time measures such as eye gaze, gives evidence about how features of the input (e.g., an ambiguous word; a complex visual scene; the preceding discourse) influence the construction of an interpretation and when in this process they are having their effects.

As we showed, children's comprehension is already highly nuanced and efficient early in life. Much like adults, children can make use of detailed statistical facts about verbs' individual complementation preferences and the details of the discourse and scene contingencies to converge on an interpretive choice under conditions of ambiguity. As the CBL theory predicts, then, the young learner—like the adult—appears to be a statistics-based incremental device, sensitive to varied cue types as it tries to recover the sense of what is meant by the sequence of words arriving at the ear. This incremental, multicue picture of the child parser contrasts with some interpretations of our initial studies (Trueswell et al, 1999), in which the child parser is perceived as "modular," at least in the sense of being subject to severe limitations on the types of evidence it will recruit at all. Although not fully resolved at this point, because of the early state of this research endeavor, our current evidence lends strong support for a multiple-cue interactive and probabilistic system at work at all ages.[6]

At the same time, these same studies reveal important differences between children and adults that are consistent with our account. As just stated, we do not think these differences are of kind (modular vs. interactive parsing architectures). Rather, the younger language user has yet to discover the full range of evidence pertaining to particular linguistic choices regarding the input. He or she needs to build up relevant linguistic databases, several of which vary cross-linguistically. Minimally, the learner must build up a library of English (or French, or Hindi, etc.) word forms and the sentential contexts (of other words and phrase types) in which each such word occurs, as well as a picture of the language-specific phrasal types and organization (e.g., that in English PP's serially follow their dominating head NP's).[7] This being so, and learning being what it is, it follows that *the more frequent and reliable in the input is an observable property of the system being learned, the sooner a learner will exploit this property in making parsing decisions.* Thus, a youngster may ap-

pear to be deaf to a particular cue, such as contextual scene information, not because his or her comprehension architecture is immune to such information, but because the relevant knowledge base for using such a cue type either doesn't exist (yet), isn't fully automatized, or hasn't been integrated with other cue types.

To take a specific example, earlier we predicted that younger children would be most efficient and adult-like in their use of lexical cues to structure, as indeed they are: As the experiment showed, they are highly sensitive to the frequency facts about complementation type for particular verbs as a clue to the status of a PP as a VP- versus NP- modifier (e.g., *hitting* is more likely than *choosing* to be followed by a VP-attached PP whose noun denotes an instrument).[8] In contrast, learners at early ages are sporadic and quite errorful in their use of a particular situational cue (one frog/two frog) compared to older children and adults. By hypothesis, this is because such cues are harder to come by, and less reliable, in the input. Indeed, the range of complement privileges for common verbs is just about fully attested within single, hour-long, mother-infant (age range approx. 15 months) conversations (Lederer, Gleitman, and Gleitman, 1995; Li, 1994; Geyer, 1997). Maternal usage "errors" in this regard (e.g., saying "Don't come your toys into the living room" or "Put your apple.") are rare to nonexistent. Thus, the information is both bountiful and reliable.

In sharp contrast, the correlation between seeing two entities (let us say, two frogs) and hearing a modified, definite NP referring to a member of the set rather than a bare NP is surprisingly weak ("the green frog," "the frog on the napkin" rather than just "the frog" or even "a frog"). The Brown-Schmidt et al. (2002) referential study of adults indicates that referential specificity is discourse defined, not just scene defined: A modified, definite NP description will happen largely on those occasions when (a) you see a pair of frogs, (b) the two frogs are the topic of conversation, (c) the conversation requires, for communication, that the two frogs be distinguished from each other, and (d) this distinction has not been established in sentences that preceded the one you are now hearing. Thus this information is less available and is less reliable. Accordingly, novices often fail to use scene cues to disambiguate sentences, but show reasonable (but not perfect) understanding when the discourse goals specify the need for modification. By hypothesis, then, the youngest learners come to use syntactic-distributional evidence before this particular contextual evidence because evidence concerning the former has been, in the child's experience, easier to come by. To the naked eye, these early tendencies to make systematic use of only a subset of the cues used by adult comprehenders might masquerade as a modular architecture. In this regard, it is very important to keep in mind that our participants' disinclination to rely on contextual cues is by no means across the board. If such a cue is rendered relevant by the discourse, our child subjects were able to use it.

In sum, the comprehension machinery appears to be incremental and interactive from the earliest developmental times that we can measure it, sensitive to a variety of cues to sentential meaning. But all the same we see a change in the reliance on (or weighting of) several cues depending on the amount and generality of information that the learner has accrued concerning each cue type.

CLOSING REMARKS: THE PLACE OF COMPREHENSION IN A THEORY OF LANGUAGE ACQUISITION

In closing, we want to take a little space here to ask how "learning to parse" fits in with other topics relevant to language learning.

Phrase Learning and Word Learning

We have focused in this paper on the question of how the child and adult listener goes about resolving the meaning of an ambiguous phrase. Several sources of evidence, including scene and distributional evidence, are brought to bear, and the reliance on these over development will differ either because of their informativeness for a particular item, or because the relevant evidence is not yet in place in the young child. In this regard, resolving the meaning of a PP for the child (and the adult) resembles the computational problem facing a language user when he or she is asked to discover the intended meaning of other abstract expressions, notably the computational problem facing a language user when he or she is asked to discover the intended meaning of a new word. Multiple sources of evidence—the observed scenes, the distribution of syntactic structures it can reside in, and its discourse setting—are in this case also potentially available. However, depending on the actual meaning of this word, only some of these sources are likely to be informative; for instance, it is easier to see that somebody is *jumping* than to see that he is *thinking*, and so the observed scene is more informative for the first of these words than it is for the second. Moreover, some potentially informative sources of evidence require time and experience to construct. For instance, the syntactic environment of *think* is highly predictive of aspects of its meaning: This word, like many verbs whose semantic content pertains to mental acts and states, occurs with tensed sentence complements (compare "John thinks that the ceiling is falling" with the nonoccurring "John jumps that the ceiling is falling,"; see Gleitman, 1990). Yet the youngest learners cannot exploit this evidentiary source because they have not yet acquired the requisite syntactic knowledge of the exposure language (Gleitman, 1990 and sources cited earlier in this chapter).

In the case of PPs occurring in an utterance, the child listener is also faced with the problem of discovering the conditions under which a linguistic expression has a particular abstract meaning, that is, when it denotes a contrastive property of an object relative to another object. It is not easy to look out into the world to see when such meanings arise (since relationships like these are typically in the eye of the beholder, who in this case is the speaker). In contrast, distributional facts in the sentence itself are there for the taking by the listener and constrain meaning in the relevant ways. Therefore, these ought to exert an early influence. Discourse evidence, which often points toward how the speaker perceives the world (such as communicative acts that align the speaker's and listener's discourse model of the world), ought to facilitate the discovery of the appropriate conditions for referential specificity. And, indeed, our Rabbit-Walrus study suggests that these discourse functions of aligning perceptions and goals exert their expected influence on this sort of meaning discovery.[9]

"Parsability" as a Factor in Linguistic and Psycholinguistic Theorizing? Parsibly!

We believe our work also has implications for psycholinguistic theorizing on grammatical formalisms that are designed for language use. As mentioned in the introduction, recent experimental work on adult sentence processing has had much to say about grammatical representation in the mind. In particular, the data have been most consistent with the notion that grammatical information is highly *lexicalized,* in the sense that a lexical event triggers the computation of detailed grammatical structures that can permit the local processing of what (on the surface) appears to be longer-distance grammatical and semantic relationships (e.g., Srinivas & Joshi, 1999; MacDonald et al. 1994; Trueswell & Tanenhaus, 1994; Kim et al. 2002). The fact that, as predicted, young children show an early greater reliance on lexical information of this sort lends support to the notion that grammars are organized along these lines and implicitly implemented as such in the comprehension system.

Indeed, we believe it is no accident that the grammatical formalisms most compatible with this psycholinguistic account have been independently developed within computational circles, especially among those interested in formalisms for natural language parsing. Here, many have noted the computational advantages of lexicalized/localized structure (CCG, Steedman, 2000; LTAG, Joshi et al., 1991; HPSG, Pollard & Sag, 1987; LFG, Bresnan & Kaplan, 1982) and the need for and success of statistical mechanisms in parsing (Srinivas & Joshi, 1999; Collins & Brooks, 1995; Marcus, 1995; Kim et al., 2002). This consistency of theory suggests that linguistic and psycholinguistic formalisms are causally related to an extent not appreciated a decade or two ago. A particular contribution from the psycholinguistic inquiries of the past

two decades has been to add *incrementality* as a desideratum: Language design is what it is, in part, because the sentence has to be interpreted in real time. Syntactic and semantic formalisms, if they are to be psychologically plausible, need to allow for the incremental building of structure and the rapid computation of contextual dependencies reflected in the meaning of these structures (for some of the most elegant and linguistically sophisticated discussions of these matters, see Altmann & Steedman, 1988; Sedivy et al., 1999; Sedivy, 2002).

In the present paper, we have taken the position, in common with many prior researchers, that the problems of learning vocabulary and language-specific syntax also fit into the same broad picture; that is, that language design is also constrained by aspects of how it is learned (e.g., the introduction of "learnability constraints" such as those proposed by Wexler and Culicover, 1980; Wexler & Hamburger, 1973, Osherson & Weinstein, 1982; Pinker, 1984; but see also Seidenberg, 1997, and Elman, 1993, for related, but quite different, perspectives on how learning may constrain language design). Here again we have tried to show that, for psychological plausibility, the notion of incrementality enters into the account but in new ways. First, there is a natural ("intrinsic") ordering among features of the evolving language system. As one crucial instance, the child receives a database of reliable lexical evidence that can be deciphered, stored, and organized early on. The individual verbs have their discernable distributions (e.g., instruments for hitting acts) and are similarly selective in their syntactic structures. The lexically specific learned facts leave their footprints in the design of the mature system, which, as we have stated above, continues to operate from lexically derived information. Second, there are practical constraints on which evidence will come on line when, as the reliability of the evidence for constructing databases differs in the input. Overall, children's comprehension ability looks the way it does because they are building a linguistic ladder, so to speak, as they are climbing it.

NOTES

1. Such an assumption seems even less radical and more familiar when we reconsider the often unspoken assumptions behind such measures as reaction time, as assessed by the stroke of a finger on a key or lever. Nevertheless, we are not advocating that psycholinguistics or any other research field rely on a single experimental technique. The linking assumptions of this new measure certainly need to be more carefully stated and tested as we proceed with its development.
2. Throughout, we use the abbreviations NP (Noun Phrase), VP (Verb Phrase) and PP (Prepositional Phrase).
3. In certain ways, our developmental account (and the CBL account generally) is reminiscent of the Bates & MacWhinney's (1987) Competition Model. For instance, both theories assume constraint-satisfaction mechanisms for language discovery and

use, and therefore emphasize information reliability when accounting for developmental patterns. However, most of the similarities end there. One crucial difference is that the CBL account assumes a central role for detailed linguistic representations in language use along multiple, partially independent dimensions (phonology, syntax, semantics). Representational modularity in the presence of interactive processing, a key assumption of CBL, is crucial for accounting for a broader range of phenomena (Trueswell & Tanenhaus, 1994, see also Jackendoff, 2002).

4. Alternatively, this could be evidence that children are "mini-minimal-attachers," showing a general structural bias toward VP-attachment. This issue has been addressed below and reported in Snedeker, Thorpe, & Trueswell (2001).
5. This means that some inherent ordering to cue use is imposed even on a constraint-based system, since the building of certain linguistic representations (i.e., evidential sources), serve as prerequisites to building other, often higher level, linguistic representations (see Fisher & Gleitman, 2002).
6. We say this because, for instance, online measures need to be collected on our discourse-based (Rabbit-Walrus) effects, to assure us that children are not using discourse cues at a later stage of processing, such as during revision. A pattern of this sort would be consistent with a modular developmental account.
7. If the highly "incremental" approach to parsing is correct - i.e., if the listener makes use of information as soon as it becomes available - then parsing in different languages can look materially different. For instance, in English, certain information that can be gleaned from the subject NP usually becomes available first, temporally preceding information about the verb and about the verb's complement structure. In other languages, object NP's or main V's canonically capture first serial position. Depending on the scope and reliability of "first received" information, later-arriving information is often redundant. Parsing in one language might be more "verb-based" and in another more "noun-based" depending on just when, given serial ordering facts, each reaches the ear. Thus, a uniform system might be at work across languages but differ in how information is weighed and used (Kaiser & Trueswell, in progress). To the extent that this is so, it follows that the infant is building comprehension machinery that is specifically tailored to exploit surface properties of the exposure language.
8. In a reasonable world, understanding what these verbs mean equivalently predicts these complementation likelihoods. To that extent one might hold that the language learner never has to record the frequency with which each verb occurs with a VP-attached PP complement. This expectation is a consequence of the verb's logical structure (for a relevant linguistic discussion, see e.g., Baker, 2001; Jackendoff, 2002), a matter of knowing the semantics of the verb rather than a matter of knowing probabilistic facts about its distributional contexts (which latter could arise as an artifact of the meaning). However, these two approaches to how children come to acquire verb complement relations (via the semantics or via the observed distributions) are not as different as they may seem at first glance. They are different only insofar as the claim can be sustained that the verb meanings are acquired independently of and prior to acquiring knowledge of their distributional structure. And this does not seem to be the case. Current evidence is consistent with a word-learning process that uses situational and distributional evidence interactively to converge on the semantics of verbs—much like the interactive comprehension-learning machinery

that is the topic of the present paper; for many kinds of words, it just isn't possible to learn their meanings before and independently of acquiring information about their distributional structure (Gleitman, 1990; Gillette et al., 1999; Gleitman & Trueswell, forthcoming).

9. One might wonder whether verb learning is also influenced by how discourse and joint attention shape a listener's perception of the world. We have been asking this very question in ongoing research, where we find that the alignment of the speaker's and listener's perception of the scene does indeed partially influence a listener's guess of verb use (Nappa, Trueswell, & Gleitman, in progress).

REFERENCES

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

Baker, M. C. (2001). *The atoms of language: The mind's hidden rules of grammar Basic Books.*

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), Mechanisms of language acquisition. (pp. 157-193) Hillsdale, NJ.: Lawrence Erlbaum Associates.

Boland, J., & Cutler, A. (1996). Interaction with autonomy: Multiple output models and the inadequacy of the Great Divide. *Cognition, 58*, 309-320.

Bresnan, J., & Kaplan, R. (1982). Lexical functional grammar: A formal system of grammatical representation. In J. Bresnan (Ed.), *Mental representation of grammatical relations*. Cambridge, MA: MIT Press.

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.

Carlson, G.N., & Tanenhaus, M. K. (1988). Thematic roles and language comprehension. In W. Wilkins (Ed.), *Syntax and Semantics, Vol. 21: Thematic Relations*, pp. 263-289. Academic Press.

Collins, M., & Brooks, J. (1995). Prepositional phrase attachment through a backed-off model. *Proceedings of the third workshop on very large corpora .*

Cooper, R.M. (1974). The control of eye fixation by the meaning of spoken language. *Cognitive Psychology, 6*, 84-107.

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.

Ferreira, F., & Clifton, C. Jr. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25,* 348-368.

Ferreira, F. & Henderson, J.M. (1991). How is verb information used during syntactic parsing? In G. B. Simpson (Ed.), *Understanding word and sentence*. North Holland: Elsevier Science Publishers.

Fisher, C. (1996). Structural limits on verb mapping: The role of analogy in children's interpretations of sentences. *Cognitive Psychology, 31*(1), 41-81.

Fisher, C., Gleitman, H., & Gleitman, L. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology, 23*(3), 331-392.

Fisher, C., & Gleitman, L. (2002). Breaking the Linguistic Code: Current Issues in

Early Language Learning. In C.R. Gallistel (Ed.) *Steven's handbook of experimental psychology: Vol. 2, Learning, motivation and emotion.* John Wiley & Sons: New York.

Frank, R. (1998). Structural complexity and the time course of grammatical development. *Cognition, 66*, 249-301.

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. (1989). Against lexical generation of syntax. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process*. Cambridge, MA: MIT Press.

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

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178-210.

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*, 153-190.

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

Gleitman, L., & Gleitman, H. (1970) *Phrase and paraphrase.* New York: Norton.

Gleitman, L., & Gleitman, H. (1997) What is a language made out of? *Lingua 100*, 29-55.

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

Geyer, H. (1997). Subcategorization as a predictor of verb meaning: Evidence from modern Hebrew. Unpublished manuscript, University of Pennsylvania.

Hudson, C. L., & Newport, E. L. (2003). Regularizing unpredictable variation in the transition from pidgin to creole: An experimental investigation of the role of adult learners in pidgin/creole development. Manuscript submitted for publication.

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.

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

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

Joshi, A., Vijay-Shanker, K. & Weir, D. (1991). The convergence of mildly context sensitive formalisms. In P. Sells, S. Shieber, & T. Wasow (Eds.), *The processing of linguistic structure*, 31-91. Cambridge, MA: MIT Press.

Kim, A., Srinivas, B., & Trueswell, J.C. (2002). The convergence of lexicalist perspectives in psycholinguistics and computational linguistics. In P. Merlo, & S. Stevenson (Eds.), *Sentence processing and the lexicon: Formal, computational and experimental perspectives*. Philadelphia: John Benjamins Publishing.

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

Lederer, A., Gleitman, L., & Gleitman, H. (1995) Verbs of a feather flock together: Structural properties of maternal speech. In M. Tomasello, & E. Merriam (Eds.), *Beyond words for things: Acquisition of the verb lexicon*. New York: Academic Press.

Li, P. (1994). Maternal verb usage in Mandarin Chinese. Unpublished manuscript, University of Pennsylvania.

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and cognitive processes, 9*, 157-201.

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

Marcus, G. F. (2000). Pabiku and Ga Ti Ga: Two mechanisms infants use to learn about

the world. *Current Directions in Psychological Science*, 9(5), 145-147.

Marcus, M. (1995). New trends in natural language processing: Statistical natural language processing. *Proceedings of the National Academy of Science, 92*, 10052-10059.

Marslen-Wilson, W. D., & Tyler, L. K. (1987). Against modularity. In J. Garfield (Ed.), *Modularity in knowledge representations and natural language understanding*. Cambridge, MA: MIT Press.

McClelland, J. L. (1987). The case for interactionism in language processing. In M. Coltheart (Ed.), *Attention and performance XII*. London: Erlbaum.

Nappa, R., Trueswell, J., & Gleitman, L. (in progress) Referential cues to perspective verbs.

Osherson, D.N., & Weinstein, S. (1982). Criteria of language learning. *Information and Control 52(2),* 123-138.

Pinker, S. (1984) *Language learnability and language development*, Cambridge MA: Harvard Univ. Press.

Pollard, C., & Sag, I. (1987). *Information-based syntax and semantics*. CSLI Lecture Notes.

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.

Saffran, J., Newport, E., & Aslin, R. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language, 35*, 606-621.

Sedivy, J. (2002). Invoking discourse-based contrast sets and resolving syntactic ambiguities. *Journal of Memory and Language*, 341-370.

Sedivy, J. Tanenhaus, M., Chambers, C., & Carlson, G. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71*, 109-147.

Seidenberg, M .S. (1997). Language acquisition and use: learning and applying probabilistic constraints. *Science, 275,* 1599-1604.

Snedeker, J., & Gleitman, L. R. (in press). Why it is hard to label our concepts. In Hall, & Waxman (Eds.) *Weaving a lexicon*. Cambridge, MA: MIT Press.

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. (submitted). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing.

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.

Spivey-Knowlton, M., & Tanenhaus, M. K. (1994). Referential context and syntactic ambiguity resolution. In C. Clifton, Jr., & L. Frazier, et al. (Eds.) *Perspectives on sentence processing.* Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Steedman, M. (1995). *Computational aspects of the theory of grammar.* Cambridge, MA: MIT Press.

Steedman, M. (2000). *The syntactic process.* Cambridge, MA: MIT Press/Bradford Books.

Tabossi, P., Spivey-Knowlton, M. J., McRae, K., & Tanenhaus, M. K. (1994). Semantic effects on syntactic ambiguity resolution: Evidence for a constraint-based resolution process. In C. Umilta, & M. Moscovitch (Eds.), *Attention and performance 15: Conscious and nonconscious information processing*. Cambridge, MA: MIT Press.

Tanenhaus, M. K., Spivey-Knowlton, M. J. (1996). Eye-tracking. *Language & Cognitive Processes, 11*(6), 583-588.

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.

Taraban, R., & McClelland, J. (1988). Constituent attachment and thematic role assignment in sentence processing: Influ-

ences of content-based expectations. *Journal of Memory and Language, 27*, 1-36.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language, 35*, 566-585.

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. M. (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., & Culicover, P. (1980). *Formal principles of language acquisition*. Cambridge, MA: MIT Press.

Wexler, K., & Hamburger, H. (1973). On the insufficiency of surface data for the learning of transformational languages. In K. J. J. Hintikka, J. M. E. Moravcsik, & P. Suppes, (Eds.), Approaches to natural language. *Proceedings of the 1970 Standard Workshop on Grammar and Semantics* (pp. 167-179). Dordrecht: Reidel.