

18

Learning, Parsing, and Modularity

917

Stephen Crain,*
Weijia Ni,*
Laura Conway
University of Connecticut
**and Haskins Laboratories*

According to the *modularity hypothesis*, different components of the language apparatus obey different operating principles (cf. Crain & Shankweiler, 1991; Crain & Steedman, 1985; Fodor, 1983; Shankweiler & Crain, 1986). In this chapter, we argue that two subcomponents of the language apparatus are autonomous because of their different operating characteristics. One is the component responsible for language acquisition (i.e., the language acquisition device [LAD]). The other is the component used in resolving ambiguities (i.e., the sentence-parsing mechanism). We identify instances in which operating characteristics of these two components conflict due to the different demands placed on them. The principles of the LAD must be responsive to demands of learnability. To achieve learnability, the LAD must constrain learners' hypotheses to guarantee that linguistic representations that are not derived in the target language will not be formulated or, if formulated, can be disconfirmed by readily available evidence. By contrast, the sentence-parsing mechanism selects among the competing linguistic representations that are derived in a language. Selection is based on considerations of simplicity. It turns out that certain initial representations that are favored on learnability grounds are ones that are dispreferred by the sentence-parsing mechanism. Putting it the other way around, certain linguistic representations that are preferred by the sentence-parsing mechanism are ones that would create problems of learnability if they were initially adopted by learners. The upshot of these deliberations is that the principles of parsing must not guide learners in their formulation of grammatical hypotheses. Therefore, the principles of parsing and the principles of learning must be kept in distinct modules of the language faculty.

We assume the following model of the parser. In resolving ambiguities, the

parser prefers representations that succeed in referring to entities that have already been introduced in the discourse, rather than ones that postulate the existence of entities not previously mentioned. In the absence of context, as in most studies of adult sentence processing, the parser begins to construct multiple representations (mental models) of an ambiguous sentence in parallel, but quickly curtails its operations once a viable representation is constructed. Lacking evidence from the discourse context, parsing decisions are guided by a principle that favors the representation that requires the fewest extensions (or accommodations of presuppositional failure) to the mental model. This is called the *principle of parsimony* (Crain & Steedman 1985).

Our assumptions about the operating principles of the LAD are quite different. These principles must enable learners to successfully converge on the target grammar on the basis of the available evidence. Presumably, the evidence available to learners consists primarily of sentences presented in circumstances that make them true. Sometimes more than one alternative interpretation of a sentence is made available by Universal Grammar (UG).¹ To complicate matters further, the alternative interpretations may sometimes form a subset-superset relationship (i.e., the circumstances that make the sentence true on one interpretation may be a proper subset of the circumstances that make it true on another interpretation). A semantic subset problem arises if the target language includes the subset interpretive option, but not the superset option. To avoid semantic subset problems, the interpretive options for sentences must be ordered in the LAD by a principle instructing learners to initially choose the representation that is true in the smallest set of circumstances.² This is called the *semantic subset principle* (Crain, 1992, 1993; Crain & Lillo-Martin, in press; Crain & Philip, 1993).

This chapter is concerned with conflicts between the basic operating principles of acquisition and parsing. The cases we discuss involve the focus operator—*only*. The next section sets the stage with a brief description of the syntactic and semantic properties of *only*. Following that, we discuss the principle of parsimony, which guides adults in parsing sentences with *only*. Finally, we discuss the semantic subset principle, which guides children in the acquisition of sentences with *only*.

¹However, this does not entail that both options are derived in the target language. Some UG options may be available in the theory, but may not appear in some particular language. Such is the case with parameter settings, for example. Therefore, it is important not to confuse the state of affairs we are describing—where the child selects among competing grammatical options—with the state of affairs that confronts adults in processing structurally ambiguous sentences.

²This argument presupposes that negative semantic evidence is not available to learners. That is, we assume that learners are not informed with sufficient regularity about interpretations that cannot be assigned to sentences in the target language.

SYNTAX AND SEMANTICS OF *ONLY*

Syntactically, the focus operator *only* associates exclusively with elements that it c-commands.³ This means that when *only* appears in preverbal position, it can focus on the entire verb phrase (VP), or on a constituent within the VP (e.g., the direct-object noun phrase [NP]). Because *only* can associate with more than one element on different occasions of use, sentences that contain *only* are often ambiguous. One ambiguity is illustrated in (1), where the boldface **F** indicates the element in focus. The alternative readings of (1) are paraphrased in (2).

- (1) a. The big elephant only eats peanuts.
 b. The big elephant only eats peanuts.
 F
 c. The big elephant only eats peanuts.
 F
- (2) a. -The only thing the big elephant does is eat peanuts.
 b. The only food the big elephant eats is peanuts.

When it appears in preverbal position, as in (1), *only* cannot associate with the subject NP. For this to happen, *only* must appear in presubject position. In that position, it can associate with the entire NP, or with a nominal constituent within the NP. An example is given in (3), with the corresponding interpretations given in (3a–b).

- (3) Only the big elephants eat peanuts.
 a. The only thing who eats peanuts is the big elephant.
 b. The only elephant who eats peanuts is the big elephant.

In presubject position, the c-command domain of *only* does not extend beyond the NP in which it appears, so it cannot associate from this position with the VP or with any element inside the VP.

The basic semantic function of focus operators generally, and *only* in particular, is to signal when the extension of some linguistic constituent is being contrasted with a set of alternatives. An example will clarify the point. Consider Sentence (4) (Willoughby's is a shop in New Haven, Connecticut, that sells really good coffee).

- (4) In New Haven, only Willoughby's coffee is really good.

Sentence (4) is felicitous only if coffee from Willoughby's is being compared to coffee from other shops in New Haven. If the speaker had tried only Will-

³A constituent A c-commands another constituent B if, and only if, there is a path from A up to the first branching node above A and then down to B (cf. Reinhart, 1983).

oughby's coffee, we would be reluctant to say that the statement was appropriate. However, note that the sentence does not assert that a comparison is being made; rather, this is presupposed. The presupposition that coffee from other shops has been tried is triggered by the occurrence of *only*. We call the entities that are presupposed to exist a *contrast set*.

The meaning representation of a sentence with a focus operator has three parts. Two of these parts are concerned with the content of the sentence. Some of its content is background information, B, and some of its content is in focus, F. The third part of the meaning representation is pragmatic: There must be a contrast set of alternatives to the focus element within the domain of discourse. The contrast set, CON, is not mentioned in the sentence; rather, it is presupposed.

The meaning of an individual focus operator includes truth conditions that are specific to it. For sentences with *only* to be true, two conditions must be met. First, the information in background must apply to the focus element (i.e., the background must be true of the focus). Second, the background must not apply to any of the alternatives to the focus element. That is, the background must apply uniquely to the element in focus. The requirement of uniqueness can be paraphrased as follows: If the background applies to any alternative to the focus element, that alternative is the focus element. Formally, the semantic value of the focus operator *only* can be stated using the following rule (adapted from Krifka, 1991; also see Jackendoff, 1972; Rooth, 1985).

- (5) Meaning rule for *only*: $B(F) \ \& \ \forall X\{X \in \text{CON}(F) \ \& \ B(X)\} \rightarrow X = F$, where X is variable of type F, and CON(F) is a set of contextually determined alternatives to F.

To unpack the contents of the meaning rule for *only*, let us examine it one piece at a time. The first conjunct, B(F), states the requirement that the background must apply to the focus element. The second conjunct is the statement of uniqueness: $\forall X\{X \in \text{CON}(F) \ \& \ B(X)\} \rightarrow X = F$. The universal quantifier ranges over a metavariable, X. The metavariable X may be replaced by real variables of different types, depending on the kind of entity that is being contrasted with the focus element. This gives the meaning rule the flexibility to handle alternative interpretations that may be assigned to sentences with *only*. Different interpretations are rendered by replacing the metavariable, X, with a variable of one kind or another. If the focus element is an individual, the contrast set will be individuals; therefore the metavariable in the meaning rule will be replaced by an individual variable: x, y, and so on. If the focus element is a property, as with a VP, the contrast set will consist of properties rather than individuals, therefore, the metavariable will be replaced by a variable of this type: P, Q, and so on. The meaning rule ends by guaranteeing the uniqueness of the focus element—for

each member of the contrast set, if the background applies to it, that member is the focus element.

To illustrate how the meaning rule in (5) works, let us apply it to some examples, beginning with (6).

- (6) The big elephant only eats peanuts.
 a. The only food the big elephant eats is peanuts.
 b. The only thing the big elephant does is eat peanuts.

We saw that when *only* is in preverbal position it can associate with at least two different elements, as indicated by the paraphrases in (6a–b). Therefore, we should be able to derive different logical representations from the meaning rule in (5), corresponding to these interpretations. We take them in turn. Interpretation (6a) can be stated informally as follows: There is something under consideration that the big elephant eats, namely, peanuts; if there is anything else that the big elephant eats, that thing is peanuts. This is a close paraphrase of the logical representation that follows from the meaning rule:

Logical Representation for Reading (6a)

The-Big-Elephant-Eats' (Peanuts') &
 $\forall y\{y \in \text{CON}(\text{Peanuts}') \ \& \ \text{The-Big-Elephant-Eats}'(y)\} \rightarrow y = \text{Peanuts}'$

(In the formula, the focus, F, is the direct object, whose semantic value is Peanuts'. The background, B, is represented as The-Big-Elephant-Eats'. Because the element in focus is an individual, the metavariable X from the meaning rule in (5) is replaced by an individual variable, viz. y.)

Next we derive interpretation (6b). Stating it informally, there is some activity the big elephant is engaged in, namely, eating peanuts; if there is any other activity that the big elephant is engaged in, then that activity is one of eating peanuts:

Logical Representation for Reading (6b)

The-Big-Elephant' (Eats-Peanuts') &
 $\forall P\{P \in \text{CON}(\text{Eats-Peanuts}') \ \& \ \text{The-Big-Elephant}'(P)\} \rightarrow P = \text{Eats-Peanuts}'$

(Here, the element in focus is the VP, Eats-Peanuts'. Because this is a property, a variable that ranges over properties, viz. P, is substituted for the metavariable, X, in the meaning rule in (5).)

To end this section, we acknowledge another construction containing *only* that gives rise to ambiguity. This construction is illustrated in (7).

- (7) The big elephant is the only one eating peanuts.
- a. The only thing eating peanuts is the big elephant.
 - b. The only elephant eating peanuts is the big elephant.

As the paraphrases (a, b) show, the contrast set for (7) can either be the set of elephants or the entire domain of discourse. Although both interpretations of (7) are quite accessible for most people, there are reasons for thinking that the interpretation in (7b) is preferred. In the next section, we explain why this should be the case.

PARSING AMBIGUOUS SENTENCES WITH *ONLY*

This section outlines a model of how adults resolve semantic ambiguities. The model is called the *referential theory*. First, we show how the referential theory explains the differences that are found in parsing so-called "garden path" sentences with the definite article, such as Sentence (8), and ones with the focus operator *only*, such as Sentence (9).

- (8) The horses raced past the barn fell.
 (9) Only horses raced past the barn fell.

According to the referential theory, sentences like (9) should not evoke garden path effects. We explain why and provide some empirical support for this prediction. Then we turn to other constructions containing *only*, which we have recently investigated in a series of experiments, that compare the interpretations assigned by adults and those assigned by children.

The Referential Theory

The referential theory maintains that the preferred reading of an ambiguous sentence is determined by comparing the discourse representations associated with the alternative interpretations (Altmann & Steedman, 1988; Crain & Steedman, 1985; Ni & Crain, 1990). On this account, multiple syntactic analyses are computed as an ambiguous sentence is parsed. These analyses are input to the semantic-pragmatic processor, which selects the interpretation that best conforms to the surrounding context. The referential theory contends that the primary responsibility for resolving structural ambiguities does not rest with structural mechanisms, but with immediate, word-by-word evaluation of the alternative analyses by the semantic-pragmatic processor. The context in which the sentence is presented is checked, and the reading of the sentence that best fits the context is selected, as stated in the following principle (Crain & Steedman, 1985):

The Principle of Referential Success: If there is a reading that succeeds in referring to an entity already established in the perceiver's mental model of the domain of discourse, then it is favored over one that does not. (p. 331)

In most experimental work reported in the literature, ambiguous sentences are presented in the so-called null context. Therefore, the principle of referential success is not applicable. A related principle is operative, however. According to the referential theory, outside of any linguistic or nonlinguistic context, the perceiver actively attempts to construct a mental representation of a situation that is consistent with each of the alternative interpretations of the sentence. In addition to the characters and events depicted in the sentence, the construction of a mental model sometimes requires the perceiver to represent information that the sentence presupposes, not just what it asserts. The process of augmenting one's mental model to represent the presuppositional content of sentences has been called "accommodation for presupposition" by Lewis (1979, p. 340).

According to the referential theory, the accommodation of presuppositional failure plays a critical role in explaining which interpretation of an ambiguous sentence is preferred outside of a discourse context. The parser begins constructing all permissible representations of a sentence, but, due to limited computational resources, quickly settles on just one interpretation—the one that requires the fewest modifications in establishing a coherent discourse representation. Crain and Steedman advanced the following principle to explain such parsing preferences.

The Principle of Parsimony: If there is a reading that carries fewer unsatisfied but consistent presuppositions than any other, then that reading will be adopted and the presuppositions in question will be incorporated in the perceiver's mental model. (p. 333)

To wrap up our review of the referential theory, we note one further property of the principle of parsimony: It pertains only to ambiguous sentences. In processing unambiguous sentences, the parser has no choice but to construct a mental representation that accommodates all unmet presuppositions. This is why sentences like "The horse ridden past the barn fell" do not pose difficulties for the parser, even outside of context.

Experiment 1: Garden-Path Sentences with *Only*

This section presents the results of a recent experimental investigation of the principle of parsimony: it applies to garden-path sentences with the focus operator *only* presented without contextual support. The experiment manipulated the referential content of NPs by substituting the word *only* for the definite determiner *the*. Test materials included sentences like (10)–(12).

- (10) The students furnished answers before the exam received high marks.
 (11) Only students furnished answers before the exam received high marks.
 (12) Only dishonest students furnished answers before the exam received high marks.

These sentences are structurally identical at the point during processing at which the ambiguity arises; namely, at the word *furnished*. Sentence (10) is a variant of the classic garden-path sentence "The horse raced past the barn fell," therefore it should be expected to induce garden-path effects, on any theory. Accordingly, the referential theory predicts that Sentence (10) will produce garden-path effects. This expectation does not hold for Sentence (11), however. Because the modifier *only* has replaced the definite determiner, it follows from the referential theory that the parser will opt for the reduced-relative clause analysis of the ambiguous phrase beginning with "furnished answers. . . ."

As we discussed earlier, the focus operator *only* requires the perceiver to form a discourse representation that contrasts the element in focus with another set of entities. In (11), this requirement is most readily satisfied if the reduced-relative clause analysis of the ambiguous phrase is selected. Having already encountered the NP, "only students," the discourse representation makes reference to a set of students. At the point of the ambiguity, the parser has two options. One option is to analyze the ambiguous verb *furnished* as part of the main clause (i.e., to assign it the grammatical feature, past tense). This puts in focus the head of the NP containing *only*, as indicated in the following.

Only students . . .
 F

If this option is pursued, however, the parser still has to establish the set of alternatives to the focus set—*students*. The problem is that there is no information about the nature of the contrast set either from the context (because we are assuming that there is no discourse context) or from within the sentence. It follows from the principle of parsimony that creating a contrast set from scratch should be avoided if possible (see Heim, 1982, for a similar suggestion).

However, the option remains to interpret *furnished* as a past participle, beginning a reduced-relative clause.

Only students furnished answers . . .
 F

If this option is chosen, the problem of constructing a contrast set is solved. The reduced-relative clause provides the information needed to form the contrast set; the set in focus is subdivided. In the present example, the set of students is partitioned into those who were furnished answers before the exam and those

who were not. In short, the requirements of *only* are satisfied more easily on the reduced-relative clause analysis of the local ambiguity. Consequently, garden-path effects should not occur.

In (12), by contrast, the appearance of the adjective *dishonest* in the subject NP helps the perceiver to establish the focus set (dishonest students) and its corresponding contrast set (honest students) before the ambiguity is encountered.

Only dishonest students . . .
 F

Hence, at the point of ambiguity, there is no longer any reason to favor the reduced-relative clause analysis. The same conditions obtain at this point in (12) as obtain at this point in (10). Therefore, the same analysis should be pursued and we should expect the reemergence of garden-path effects in sentences like (12). In summary, the referential theory makes the following predictions about garden-path effects for these sentences.

- (a) The students—garden-path effects
- (b) Only students—no garden-path effects
- (c) Only dishonest students—garden-path effects

These predictions were tested in an experiment using a self-paced reading paradigm. Subjects read sentences one word at a time on a computer screen. They called up each new word by pressing one of two keys marked "yes" or "no." Subjects were instructed to press the "yes" key as long as each word could be grammatically incorporated into the material they had previously read. Either key press continued to bring up new words, which accumulated on the screen. The computer recorded the duration in milliseconds between the onset of each new word and the following key press.

There were 32 test sentences and 16 control sentences with unambiguous verbs. They were interspersed among 92 filler sentences. Four lists of stimuli were composed so that each version of each test and control sentence appeared only once in any list. Eight subjects were tested on each list, thus 32 subjects participated in the experiment.

To analyze the data, mean reaction times for the test sentences and the controls were divided into four regions. The first region spanned the initial NP; the second region consisted of the first VP, which contained the ambiguous verb in the test sentences; the third region was the main verb; and the final region was the remainder of the sentence. Figure 18.1 presents the data for sentence of type (a), (b), and (c) in each region. The points in each region represent the average time subjects took to read each word in that region. Only those responses that correctly recognized the sentences as grammatical were used in the analysis.

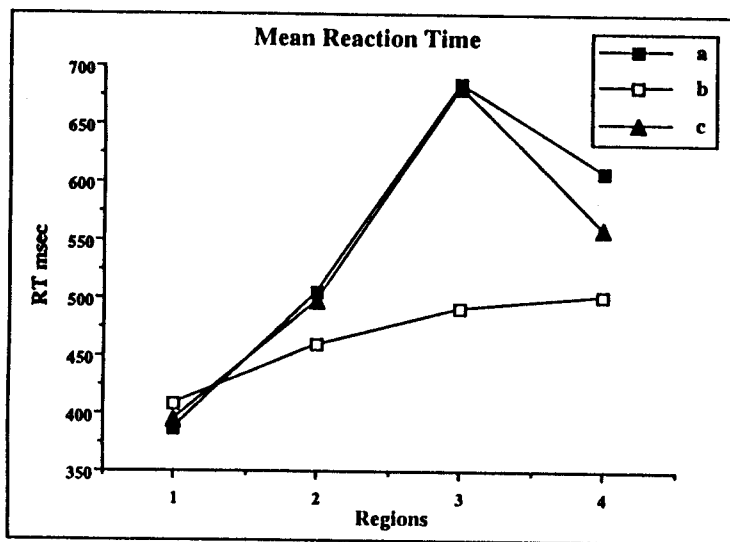


FIG. 18.1 Mean reaction time by region for three types of locally ambiguous sentences: (a) The N . . . , (b) only N . . . , (c) only adj N . . .

a. <u>The students</u>	<u>furnished . . .</u>	<u>received</u>	<u>high . . .</u>
b. <u>Only students</u>	<u>furnished . . .</u>	<u>received</u>	<u>high . . .</u>
c. <u>Only dishonest students</u>	<u>furnished . . .</u>	<u>received</u>	<u>high . . .</u>
1	2	3	4

These were the findings: There were no significant differences in reaction times among any of the sentence types in region 1. In region 2, type (a) and (c) sentences grouped together and yielded slightly longer reaction times than type (b) sentences. Region 3 presented significant differences among the sentence types. Type (a) and (c) sentences took significantly longer to read than type (b) sentences, which patterned like the unambiguous controls without any elevation in response times. Thus, the parser was not disturbed in the on-line processing of type (b) sentences in region 3. By contrast, responses times were elevated in region 3 for sentence types (a) and (c). We interpret these data as showing that type (a) and (c) sentences induced garden path effects at the point of disambiguation, but sentences of type (b) did not. To circumvent the criticism that word-by-word monitoring tasks do not tap rapid on-line processing, we replicated the present study using the technique of eye-movement recording (Ni & Crain, 1993). The findings were equally robust.

To summarize the findings of Experiment 1, garden-path effects appeared in sentences with *the* and ones with *only* plus an adjectival modifier, but not when *only* was directly followed by a noun. This is exactly as predicted by the referen-

tial theory. These findings underscore the conclusion reached by Crain and Steedman (1985) that there really is no such thing as a "null context." In the absence of an external context, the parser actively engages in internal, mental model building.

Minimal Commitments

The results of Experiment 1 support the claim that the principle of parsimony guides the on-line operation of the parser, outside of context. Crain and Hamburger (1992) suggest that the principle of parsimony is ultimately motivated by the need to minimize cognitive effort in response to limitations in working memory capacity. To conserve effort, unnecessary extensions to the mental model are avoided if possible. The advantage of such a "least effort" strategy for ambiguity resolution is to reduce the risk of making commitments that will need to be changed later. To coin a phrase, the parser is a "minimal commitment" component of the language apparatus.

To avoid unnecessary commitments, the parser selects the interpretation of an ambiguous sentence that makes it true in the largest set of circumstances.⁴ To illustrate this operating characteristic of the parser, consider the ambiguity involving *only* in Sentences like (13).

- (13) The big elephant is the only one playing the guitar.
- a. The only thing playing the guitar is the big elephant.
 - b. The only elephant playing the guitar is the big elephant.

As the paraphrases in (a) and (b) indicate, there are two possible contrast sets. One is the set of elephants; the other is the entire set of individuals in the domain of discourse. On the (a) interpretation, everything in the domain of discourse is in the contrast set. This makes Sentence (13) true only in a limited set of circumstances, as compared with interpretation (b). In fact, the (a) interpretation of (13) is true in a subset of the circumstances that make it true on the (b) interpretation. That is, the (a) reading entails the (b) reading, but not vice versa: If the big elephant is the only elephant playing the guitar, it need not be the only thing playing it; on the other hand, if the big elephant is the only thing playing

⁴Any model of discourse that contains individuals or events whose existence is disconfirmed by new information will have to be modified appropriately to bring the model of the parser in line with that of the interlocutor. Mismatches between the mental models of a speaker and a hearer, or a writer and a reader, are apt to interfere with the flow of information between them. To facilitate the transfer of information, a perceiver must continuously attempt to align their mental model with that of the other participants in the discourse. Following the guidelines of the principle of parsimony, the strategy adopted by the parser is to avoid interpretations of ambiguous sentences that entail additional commitments about individuals and events within the domain of discourse.

the guitar, it must be the only elephant playing the guitar. If our inference is correct—that the parser favors the interpretation that makes a sentence true in the broadest range of circumstances—then adults should prefer the (b) interpretation of sentences like (13).

Another ambiguity occurs when *only* appears in preverbal position, as in example (14). The alternative readings are paraphrased in (a) and (b).

- (14) The dinosaur is only painting a house.
- a. The only thing the dinosaur is doing is painting a house.
 - b. The only thing the dinosaur is painting is a house.

Again, the (a) interpretation is true in a subset of the circumstances that correspond to the (b) interpretation. To minimize commitments, then, the parser should favor the (b) interpretation because this interpretation is consistent with a larger range of possible outcomes.

In contrast to our expectation that adults will prefer the (b) interpretation of sentences like (13) and (14), we argue in the next section that children adopt the (a) interpretation as their initial grammatical representation of such sentences. In other words, children sometimes initially hypothesize semantic representations that are dispreferred by adults. We examine this proposal by looking at children's understanding of sentences with the focus operator *only*.

ACQUISITION OF SENTENCES WITH *ONLY*

The LAD permits children to attain their target grammar(s) solely on the basis of positive evidence. Just as children lack negative syntactic evidence—evidence about the ungrammaticality of sentence forms—it also seems likely that they lack the kind of evidence needed to reject incorrect hypotheses about what sentences may and may not mean. Therefore, if a child were to commit semantic overgeneration (i.e., if he or she assigned sentence meanings beyond those assigned by the adult grammar), he or she would be hard pressed to recover from the error. As with negative syntactic evidence, if every semantic miscue must be corrected on the basis of experience, this would require an enormous supply of negative semantic feedback. To our knowledge, there is little empirical data on the matter. However, we think it highly unlikely that children expunge semantic errors on the basis of experience. Therefore, we explore the possibility that children avoid making semantic errors in the first place. We propose a principle of learnability to account for the absence of errors. The principle orders children's semantic hypotheses in advance, as follows: Default hypotheses are ones that will not subsequently need to be revised (i.e., they are realized universally), and additional (language-particular) hypotheses are added on the basis of positive

evidence from the input.⁵ The principle is called the Semantic Subset Principle (cf. Crain & Philip, 1993).

Semantic Subset Principle: If the interpretative component of UG makes two interpretations, A and B, available for a sentence, S, and if interpretation A makes S true in a narrower range of circumstances than interpretation B does, then interpretation A is hypothesized before B in the course of language development.

Language acquisition is replete with potential learnability problems. A problem arises in principle whenever a sentence can be mapped onto two different interpretations, such that one interpretation is true in a set of circumstances that constitutes a subset of the circumstances corresponding to the other interpretation. To avoid such semantic subset problems, the semantic subset principle arranges grammatical options to ensure that learners initially hypothesize an interpretation that makes a sentence true in the smallest set of circumstances. In this way, learners are assured of formulating falsifiable hypotheses. To make sentences true in the narrowest possible set of circumstances amounts to making the maximal commitments about the entities and events in the domain of discourse. In short, the LAD is a "maximal commitment" component of the language apparatus.

Let us consider some examples of the semantic subset principle in operation. We saw earlier that when *only* appears in preverbal position it can focus on the entire VP or it can focus within the VP, selecting the direct-object NP. This means that (17) is ambiguous, with the interpretations indicated in (17a-b).

- (17) The dinosaur is only painting a house.
- a. The only thing the dinosaur is doing is painting a house.
 - b. The only thing the dinosaur is painting is a house.

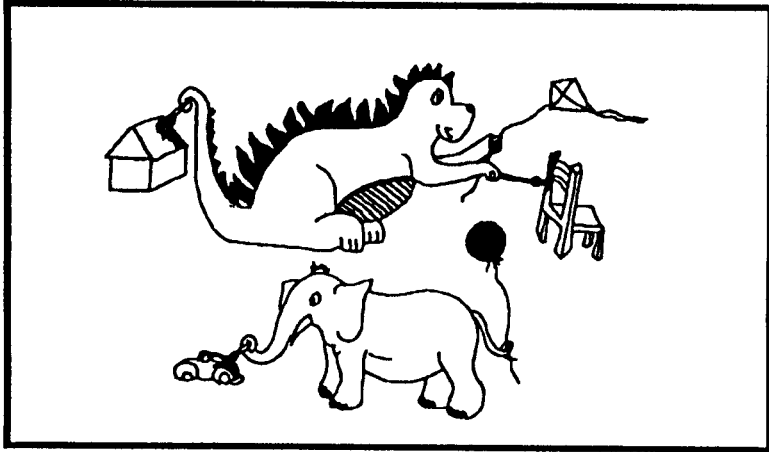
The circumstances in which these two readings are true vary depending on the contrast set for the focus operator *only*. The contrast set of both (a) and (b) include properties of individuals, rather than individuals themselves. The contrast set for the (a) reading includes all of the properties of the dinosaur under consideration. For the sentence to be true on this reading, the dinosaur can only

⁵Even if UG makes alternative interpretive options available for a sentence, it is not necessarily ambiguous for the child. As in parameter setting, children may have a range of options available to them in the theory, but they may nevertheless hypothesize only certain of these values at any given time. There is an important difference between parameter setting and the case we are considering, however. In parameter setting, new parameters supplant old ones. In formulating semantic hypotheses, by contrast, children are seen to begin with a limited set of (universal) interpretive options, which is then extended to include additional (language-particular) options on the basis of positive evidence.

have one property—that of painting the house. The (b) reading is less restrictive. The sentence is true on this reading if the only thing the dinosaur is painting is a house, but this allows the dinosaur to have other properties as well, such as flying a kite, for instance.

Children and adults are both expected to find (17) inappropriate as a description of the situation in (18), but they should have different reasons for rejecting it. Assuming that adults have both readings available, they should adopt the (b) reading of (17) because this reading is consistent with a broader range of circumstances than the (a) reading. Putting it differently, the (b) reading makes fewer commitments than does the (a) reading, so the (b) reading is favored by the principle of parsimony. However, the context in (18) makes Sentence (17) false on this interpretation.

(18)



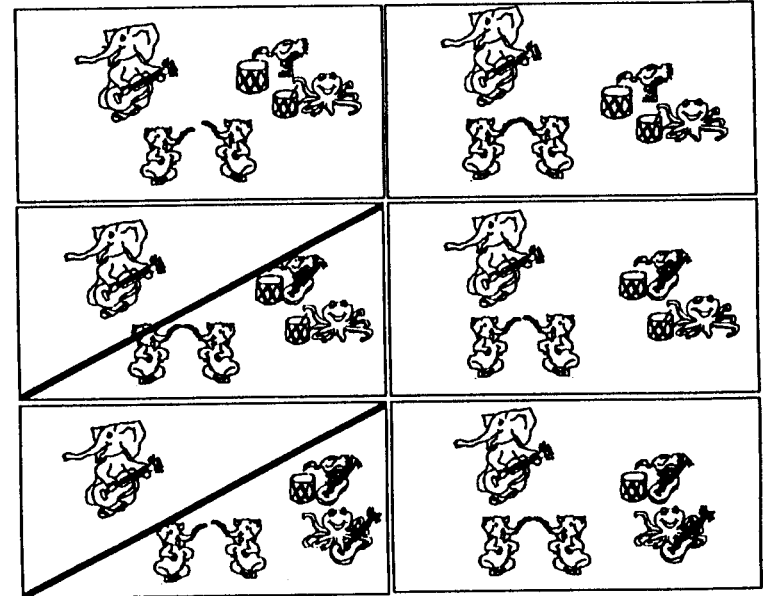
Children should also reject (17), but for a different reason. The circumstances corresponding to the alternative readings of (17) are in a subset-superset relationship. Therefore, the semantic subset principle compels children to initially hypothesize the (a) reading. This reading too is falsified by the context in (18). At a later point in development, the (b) reading will also become available, in response to evidence from the input. For example, one source of evidence would be Sentence (17) "The dinosaur is only painting a house," presented in a situation in which the only thing the dinosaur is painting is a house, but in which he is doing something else as well, say flying a kite. Once children have both readings available to them, they should behave like adults in resolving ambiguities such as (17) both in and out of context. That is, they will appeal to the principle of referential success and the principle of parsimony, respectively.

Another potential semantic subset problem arises with ambiguous sentences like (19). The ambiguity involves the reference of the proform *one*, which varies depending on what is taken to be the contrast set for the focus operator *only*.

- (19) The big elephant is the only one playing a guitar.
 - a. The big elephant is the only thing playing a guitar.
 - b. The big elephant is the only elephant playing a guitar.

The circumstances in which the two readings of (19) are true are depicted in (20). The illustration makes it clear that the set of circumstances allowed by the two readings fall into a subset-superset relationship.

(20)



Maximal Commitment

Minimal Commitment

The (a) reading of (19) is depicted on the left. This is the maximal commitment interpretation. On this interpretation, the sentence is false if anything other than the big elephant is playing a guitar. The (b) interpretation is the minimal commitment interpretation. On this interpretation, the contrast set consists of elephants. The illustrations show that the maximal commitment interpretation (a) makes Sentence (19) true in a narrower set of circumstances than it is true on the minimal commitment interpretation (b). In fact, it is true in only one circum-

stance on the maximal commitment interpretation.⁶ For the sentence to be true on the (b) reading, however, all that is required is that no other elephant besides the big elephant is playing a guitar. The activities of the octopus and the crane are not pertinent on this reading. The crucial observation here is that the (a) interpretation of (19) is true in a subset of the circumstances corresponding to the (b) interpretation. For learners, sentences like (19) represent a potential semantic subset problem. The semantic subset principle resolves the dilemma by compelling children to initially hypothesize the falsifiable interpretation (a) (cf. Hornstein & Lightfoot, 1981).

For the sake of argument, suppose that children initially guess the minimal commitment interpretation. Suppose further that the preferred adult interpretation is the maximal commitment interpretation. On this scenario, the preferred adult interpretation would be unlearnable for the children. This is because the interpretation that children assign will be confirmed in the contexts in which adults generally use sentences like (19). In these contexts, the big elephant will be the only individual that is playing a guitar. According to the children's interpretation, the big elephant just needs to be the only elephant playing a guitar. This requirement is clearly satisfied in the contexts in which adults generally use (19), because if nothing other than the big elephant is playing a guitar, then no other elephant is. This would raise a problem for learnability, however, because the analysis children assign would not require them to examine individuals who are not elephants. Therefore, children will not notice that no other individual besides the big elephant is ever playing a guitar in the contexts that adults use the sentence. But this is precisely what children must notice if they are to achieve the preferred adult interpretation. Hence, that interpretation will remain beyond their grasp. This outcome is clearly contrary to fact.

Suppose, then, that children start out with the maximal commitment hypothesis, (a). In this case, there will be positive evidence that can prompt them to add the minimal commitment reading (b) as a possible interpretation of the sentence. According to the maximal commitment hypothesis, nothing in the domain of discourse besides the big elephant can be playing a guitar. On some occasions, however, children will witness adults using a sentence like (19) in a situation that renders it false according to this interpretation. For example, they might encounter Sentence (19) in a situation where a crane is playing a guitar as well. Children will observe that such a situation is inappropriate on their hypothesis, which precludes the crane from playing a guitar. Given the reasonable assumption that children take adult sentence-meaning pairs as positive evidence for grammatical change, contexts of this kind will cause children to add semantic interpretation

⁶There is a positive correlation between the size of contrast set and the number of commitments being made. The interpretation with the smallest contrast set makes the fewest commitments, whereas the larger the contrast set the more commitments that are being made. The consequence of a large number of commitments is maximal falsifiability.

(b) to their grammars. Sentences like (19) will be ambiguous for children from that point on. Presumably, when more than one interpretive option becomes operative for a given sentence type (i.e., when that type of sentence becomes ambiguous for children), the same parsing preferences that characterize the adult processing system will also be invoked by children to resolve ambiguities.

EXPERIMENTS COMPARING LEARNING AND PARSING

We have reached quite different conclusions about the operating characteristics of the adult parser and those of the LAD. According to the referential theory, the parser favors representations that are true in the broadest range of circumstances (i.e., ones that make minimal commitments). By contrast, the semantic subset principle encourages learners to initially hypothesize representations that are true in the narrowest range of circumstances (i.e., ones that make maximal commitments).

Earlier we reported the results of an experiment examining on-line responses by adults to local structural ambiguities. We saw that the parser is sometimes able to circumvent garden path effects that would otherwise occur by pursuing the minimal commitment analysis of garden path sentences. There is an even more direct way to compare the decisions of the parser and those of the learner, however. In this section, we report the findings of two experiments that take advantage of this more direct approach—by considering how children and adults respond to globally ambiguous sentences with the focus operator *only*. The findings from two experiments are reported, demonstrating striking dissimilarities between children and adults.

Experiment 2: The focus of Preverbal *Only*

Experiment 2 was designed to assess the interpretations that children and adults assign to sentences like (21) in which the focus operator *only* can associate with either the entire VP or can focus more narrowly within the VP, associating with the direct object NP.

- (21) The dinosaur is only painting a house.
- a. The only thing the dinosaur is doing is painting a house.
 - b. The only thing the dinosaur is painting is a house.

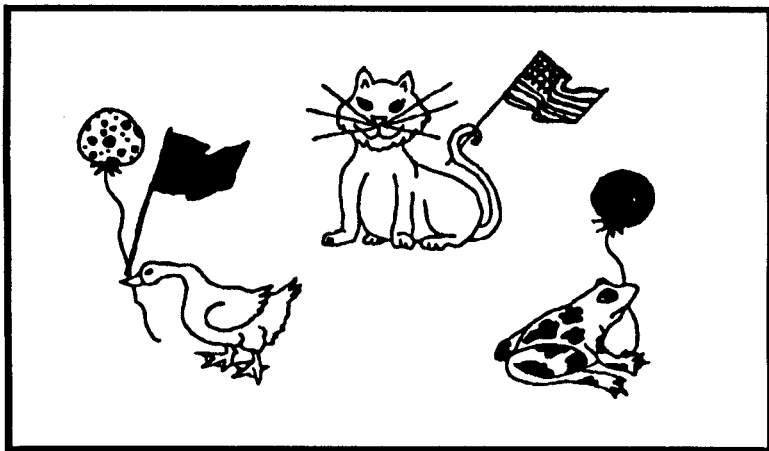
A result from previous research on children's understanding of sentences with *only* played a vital role in the experiment. The result was obtained in a study by Crain, Philip, Drozd, Roeper, and Matsuoka (1992). The study asked children to respond to sentences in which the focus operator *only* preceded the subject NP

and ones in which it preceded the VP, as in (22) and (23). The test sentences were used to describe pictures such as (24). If children interpreted the test sentence in the same way as adults do, they should have accepted (23) as a correct description, (24), but they should have rejected (22).

(22) Only the cat is holding a flag.

(23) The cat is only holding a flag.

(24)



The main finding was that 35 of the 38 three- to six-year-old children tested in this study assigned a nonadult interpretation to the test sentences. The majority of children ($n = 21$) consistently interpreted *only* as if it were construed with the VP, regardless of its surface position in the test sentence. These children correctly interpreted sentences like (23) with *only* in pre-VP position, but they interpreted sentences like (22) in the same way (i.e., they interpreted "Only the cat is holding a flag" as if it meant "The cat is only holding a flag").⁷ Children who gave this pattern of responses are called VP-oriented children.

VP-oriented children are of special interest to us because six such children served as subjects in Experiment 2. Use of these children made it possible to avoid a potential problem in presenting sentences like (23), namely, the possi-

⁷Although adults adopt different semantic representations depending on the surface position of *only*, the findings show that children initially hypothesize just one of the interpretative options from UG. As a consequence, they are forced to ignore surface position. The interpretations that are lacking in early child grammars are simply added on the basis of positive evidence, however. For example, VP-oriented children will encounter sentence-meaning pairs that are false on the interpretation they assign, but true on an alternative interpretation that is consistent with the circumstances they encounter.

bility that the prosodic contour of sentences could favor one reading or another. This problem was avoided because VP-oriented children would assign focus on the VP even when *only* preceded the subject NP. This allowed us to present sentences like (25) auditorily, with *only* in pre-subject position, to test children's assignment of focus within VPs.

(25) Only the dinosaur is painting a house.

Prosodic information from sentences like (25) could not provide cues about the speaker's intended focus of *only* because any stress by the experimenter would fall on the subject NP, and not on or within the VP.⁸ Crucially, however, Sentence (25) has the same meaning for VP-oriented children as "The dinosaur is only painting a house." In the experiment, six VP-oriented children (mean age: 4 years, 9 months) were given four test trials in a picture-verification task.⁹ The picture in (18) is representative of the test materials.

The main finding was that three of the six children always associated *only* with the entire VP of the test sentences, such as (25) not with the direct object NP. The response of these children clearly conform to the semantic subset principle. The circumstances corresponding to the alternative readings of (25) are in a subset-superset relationship. Therefore, the semantic subset principle compels children to initially hypothesize the reading that makes the maximal commitments. In the present example, this is the reading in which the only activity being performed by the dinosaur is that of painting a house. Three children's responses were exactly of this form. For example, they rejected (25) on the grounds that the dinosaur was flying a kite and painting a chair, as well as painting a house.

The responses of the remaining three children were difficult to interpret. The majority of their responses were rejections of the test sentences, but, in explaining their reasons for rejection, these children mentioned every event depicted in the picture, regardless of which character was involved. Although such responses are not inconsistent with our claims about the stepwise acquisition of VP focus, neither do they offer support for these claims. We speculate that these children were adopting a nonlinguistic strategy to derive their responses.

A different experimental design was used to test for adult preferences in assigning focus in sentences with *only*. Because we assume that all interpretations of an ambiguous sentence are available to adults, we thought it ill advised to present pictures that falsified more than one interpretation at the same time, as was true of the pictures we presented to children. Therefore, for adults, test

⁸Subjects were classified as VP-oriented on the basis of their response to intransitive sentences. For example, the intransitive sentences "Only Oscar is dancing," was presented in a context in which someone was dancing in addition to Oscar, and Oscar was drinking a Coke, as well as dancing. A child was classified as VP oriented if he or she rejected such sentences on all three of the test trials.

⁹There were also four intransitive verb controls and two unrelated filler sentence-picture pairs.

sentences were presented in written form, outside of context, and with *only* in preverbal position, as in (26).

(26) The dinosaur is only painting a chair.

We interviewed 10 adults. They were instructed to imagine a situation corresponding to each test sentence, but one in which the sentence was false. Then, they were asked to write down a description of that aspect of their imagined situation that made the sentence false. For example, in response to Sentence (26), we wanted to find out whether subjects would write things like "He also painted a table" or "He also flew a kite." Based on the principle of parsimony, we expected adults to prefer descriptions like "He also painted a table," indicating that they associated *only* with the direct object NP (the minimal commitment interpretation), rather than associating it with the entire VP (the maximal commitment interpretation). This expectation was clearly met. Twenty-eight of the 34 responses that we could interpret repeated the verb of the test sentence with a different direct object. For example, in response to example (26), subjects wrote, "The dinosaur is also painting a table" and "He's going to paint all the furniture." However, five subjects gave at least one maximal commitment response. This indicates that this interpretation is available to adults, although it is not easily accessed in the absence of context.

In summary, Experiment 2 was designed to assess the interpretations that children and adults assign to sentences in which the focus operator *only* can associate with either the entire VP or it can associate within the VP. We predicted that children and adults would associate the focus operator differently for such sentences. Children were expected to associate it with the entire VP, whereas adults were expected to associate it within the VP. The results were largely as predicted. Three of the children gave the maximal commitment interpretation to the test sentences in Experiment 2. The responses of the other children could not be counted for or against the experimental hypothesis. We interpret the positive finding from the three children whose responses were germane to the experimental hypothesis as evidence that they were guided by the semantic subset principle, which encourages learners to initially hypothesize representations that are true in the narrowest range of circumstances (i.e., ones that make the maximal commitments). By contrast, adults strongly favored the minimal commitment interpretation of the test sentences. We interpret this as evidence that adults were guided by the principle of parsimony, which instructs perceivers to favor representations that are true in the broadest range of circumstances (i.e., ones that make minimal commitments).

Experiment 3: The Interaction of *Only* and *One* Substitution

This experiment concerns the interaction between the focus operator *only* and the linguistic phenomenon known as *one substitution*. This term describes the use of

the proform *one* to refer back to the contents of a nominal element mentioned earlier in a sentence. Sometimes more than one referent for the proform *one* is possible, as (27) illustrates. The ambiguity turns on which nominal element the proform *one* is interpreted as replacing.

- (27) The big elephant is the only one playing a guitar.
 a. The big elephant is the only thing playing a guitar.
 b. The big elephant is the only elephant playing a guitar.

On the (a) interpretation, where *one* substitutes for the entire NP, Sentence (27) is true in a subset of the circumstances corresponding to the (b) interpretation, where *one* substitutes for the nominal *elephant*. For learners, these interpretations present a potential semantic subset problem. Consequently, the semantic subset principle compels children to initially hypothesize the maximally falsifiable interpretation—namely, (a). By contrast, adults are expected to favor the (b) interpretation of (27), because this interpretation makes fewer commitments as to who is playing a guitar.

Again, different tasks were administered to children and adults. To test for the influence of the semantic subset principle on children's initial grammatical representations, children were interviewed using the truth-value judgment task developed by Crain and McKee (1985). On a typical trial, a child and a puppet, Kermit the Frog (played by one experimenter), watched stories that were acted out (by a second experimenter) using toy figures and other props. In the course of a story, the experimenter identified each character, as he or she participated in the action. Following each story, Kermit the Frog reported what he thought happened in the story (using a test sentence). The child's task was to indicate if Kermit was correct by rewarding him with a bite of his favorite food. But if Kermit was wrong, the child was to encourage him to pay closer attention by pretending to feed him a bite of an old shoe. Whenever a child indicated that Kermit had said the wrong thing, we asked him or her, "What really happened?" Test sentences were presented with neutral intonation, especially within the subject NP. The stories were constructed so that a "yes" response indicated that the child assigned the minimal commitment interpretation; a "no" response indicated that the child rejected the minimal commitment interpretation. By asking the subject to explain what really happened, we ascertained whether some children at least assigned the alternative maximal commitment interpretation as predicted by the semantic subset principle.

The results were largely as expected. Eight of the twelve 3- to 5-year-old children we interviewed (mean age: 4 years, 8 months) consistently rejected sentences like "The big elephant is the only one playing a guitar" if any character other than the big elephant was playing a guitar. In theoretical terms, these children hypothesized the entire domain of discourse as the contrast set. These children rejected the alternative interpretation, according to which the contrast set consisted of just elephants. If this interpretation had been available to chil-

dren, they would have presumably said "yes." The four remaining children occasionally did accept this interpretation, although three of the four children rejected it more often than they accepted it.

Adults' responses to test sentences like (27) were assessed using the procedure described in Experiment 2. Test sentences were presented in written form outside of context. Subjects were asked to imagine a situation corresponding to each sentence, but one in which the sentence was false. Subjects were asked to write down a description of that aspect of the situation that falsified the sentence. The referential theory predicts that adults should favor those interpretations of ambiguous sentences that are true in the broadest circumstances. In particular, adults should prefer the minimal commitment interpretation of ambiguous sentences such as (27), in contrast to the response's children gave to such sentences. Based on the principle of parsimony, we expected adults to respond to (27) with descriptions that mentioned another elephant, which was being contrasted with the big elephant.

The results confirmed this expectation. Sixteen of the 18 responses that were relevant to the experimental hypothesis indicated that the referent of the head noun of the subject NP was taken as the contrast set (e.g., the set of cats was the contrast set for the sentence, "The gray cat is the only one with a toy"). In response to this sentence, for example, subjects wrote, "The black cat also has a toy," "Black and white cats have toys too," and so on. Besides the predicted pattern of responses, many subjects mentioned some other reason why the sentence was false (e.g., "The cat was not gray," "The gray cat lost its toy"). Only two of the subjects' responses conformed to a maximal commitment interpretation, however. This indicates that the maximal commitment interpretation is highly dispreferred for adults, whereas the minimal commitment interpretation is readily accessible.

To summarize the findings of the two experiments reported in this section, the responses of children and adults to the same global ambiguities reveal a striking dissimilarity in their understanding of sentences with the focus operator *only*. We conclude by discussing the implications of these findings for the architecture of the mind.

RESOLVING THE CONFLICT: MODULARITY

Is it time to draw our final conclusions. These are based on theoretical considerations of learning and parsing, as well as the results of our empirical investigations of how children and adults interpret sentences with the focus operator *only*. One set of conclusions concerns language learnability. Languages differ in the availability of particular semantic interpretations, and certain interpretative options that distinguish languages are in a subset-superset relationship. To meet the demands of learnability in the absence of negative evidence from the environ-

ment, the learner must start with the most restricted interpretation and add additional interpretations to his or her grammar based on positive evidence from the linguistic community. To converge on the target grammar, children must initially hypothesize semantic representations that make a sentence true in the narrowest possible set of circumstances. To accomplish this, the interpretations children initially hypothesize are ones that make the greatest number of commitments. This is dictated by the semantic subset principle, used by the LAD to arrange grammatical hypotheses in the order in which they are to be evaluated by the learner.

Another set of conclusions concerns language processing. Every language contains massive ambiguity. To confront this problem, with only limited computational resources, the parser rapidly checks the context surrounding an ambiguous sentence to see which of the alternative interpretations is most appropriate. In the absence of an explicit context, the parser is guided by the principle of parsimony. According to this principle, the language perceiver selects the interpretation that makes the fewest commitments, so as to limit the number of revisions that may be required in the light of impending information. The interpretation that makes the minimal commitments is the one that is true in the broadest range of circumstances.

Given this opposition between the principles of learning and parsing, the final conclusion we draw is that an impasse can be avoided only if the architecture of the mind is modular. An "interaction" between components, according to which the parser influences grammar formation, would have disastrous consequences in cases like those described in this chapter. If learners hypothesized the interpretation preferred by the parser, this would render the target grammar unlearnable. To prevent the parsing device from interacting with the LAD in this fashion, the principle of parsimony and the semantic subset principle must be kept in distinct modules of the mind. Within a modular linguistic system, the principles that dictate adult parsing preferences need not influence decisions that are made within the LAD. Learnability problems will thereby be avoided. As a final comment, we would underscore a point we made earlier. Once the learner has expanded his or her interpretive options beyond the default interpretive options prescribed by the LAD, he or she is at liberty to invoke the same parsing principles as adults do in deciding among competing interpretations. This minimizes the difference between the language apparatus of the learner and the adult language user. The difference is that the learner has an additional component, the LAD. In our view, this is all that distinguishes learners and parsers.

ACKNOWLEDGMENTS

We wish to thank Chuck Clifton, Lyn Frazier, Jane Grimshaw, Nina Hyams, Donald Shankweiler, and Mark Steedman for their comments on an earlier draft

of this chapter. Special thanks goes to Rosalind Thornton for helping with the design and execution of the experiments with children. We also thank the children, teachers, and administrators of the Child Development Laboratories at the University of Connecticut. Portions of this research were presented at Rutgers University, the University of California at Los Angeles, and at the 16th GLOW Colloquium, Lund, Sweden. The research was supported, in part, by a program project grant to Haskins Laboratories by the National Institute of Child Health and Human Development (HD-01994) and by a graduate fellowship from the National Science Foundation (to Laura Conway).

REFERENCES

- Altmann, G. T. M., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, 30, 191–238.
- Crain, S. (1992, July). *The semantic subset principle in the acquisition of quantification*. Paper presented at the Workshop on the Acquisition of WH-Extraction and Related Work on Quantification, University of Massachusetts, Amherst, MA.
- Crain, S. (1993). Continuity and modularity. In E. Clark (Ed.), *Proceedings of the 25th Annual Stanford Child Language Research Forum* Chicago, IL: University of Chicago Press.
- Crain, S., & Hamburger, H. (1992). Semantics, knowledge and NP modification. In R. Levine (Ed.), *Formal grammar: Theory and implementation* (vol. 2, pp. 372–401). British Columbia: The University of British Columbia Press.
- Crain, S., & Lillo-Martin, D. (in press). *Language and mind: An introduction to linguistic theory and natural language acquisition*. Oxford, England: Blackwell.
- Crain, S., & McKee, C. (1985). Acquisition of structural restrictions on anaphora. In S. Berman, J.-W. Choe, & J. McDonough (Eds.), *Proceedings of the North Eastern Linguistic Society* (Vol. 16, pp. 94–110). Amherst, MA: GLSA, The University of Massachusetts.
- Crain, S., & Philip, W. (1993). *Global semantic dependencies in child language*. Paper presented at the 16th GLOW Colloquium, Lund, Sweden.
- Crain, S., Philip, W., Drozd, K., Roeper, T., & Matsuoka, K. (1992). *Only in child language*. Unpublished manuscript, University of Connecticut, Storrs, CT.
- Crain, S., & Shankweiler, D. (1991). Modularity and learning to read. In I. G. Mattingly & M. Stutter-Kennedy (Eds.), *Modularity and the motor theory of speech perception* (pp. 375–392). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Crain, S., & Steedman, M. J. (1985). On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives* (pp. 320–358). Cambridge, England: Cambridge University Press.
- Fodor, J. A. (1983). *Modularity of mind*. Cambridge, MA: MIT Press.
- Heim, I. (1982). *The Semantics of definite and indefinite noun phrases*. Unpublished doctoral dissertation, University of Massachusetts, Amherst, MA.
- Hornstein, N., & Lightfoot, D. (1981). Introduction. In N. Hornstein & D. Lightfoot (Eds.), *Explanation in linguistics: The logical problem of language acquisition* (pp. 9–31). New York: Longman.
- Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.
- Krifka, M. (1991). A compositional semantics for multiple focus constructions. *Linguistische Berichte, Sonderheft*, 4. (Also in *Proceedings of Semantics and Linguistic Theory (SALT) I*, Cornell Working Papers 11).
- Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, 8, 339–359.
- Ni, W., & Crain, S. (1990). How to resolve structural ambiguities. In *Proceedings of the North Eastern Linguistic Society* (Vol. 20, pp. 414–427). Amherst, MA: GLSA, The University of Massachusetts.
- Ni, W., & Crain, S. (1993). Up and down the garden path. Unpublished manuscript, The University of Connecticut, Storrs, CT.
- Reinhart, T. (1983). Coreference and bound anaphora: A restatement of the anaphora questions. *Linguistics and Philosophy*, 6, 47–88.
- Rooth, M. (1985). *Association with focus*. Unpublished doctoral dissertation, University of Massachusetts, Amherst, MA.
- Shankweiler, D., & Crain, S. (1986). Language mechanisms and reading disorder: A modular approach. *Cognition*, 24, 139–168.

PERSPECTIVES ON SENTENCE PROCESSING

Edited by

Charles Clifton, Jr.

Lyn Frazier

Keith Rayner

University of Massachusetts at Amherst



1994

LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
Hillsdale, New Jersey

Hove, UK