

The Processing Cost of Reference Set Computation: Acquisition of Stress Shift and Focus

Tanya Reinhart
Utrecht University

Reference set computation—the construction of a (global) comparison set to determine whether a given derivation is appropriate in context—comes with a processing cost. I argue that this cost is directly visible at the acquisition stage: In those linguistic areas in which it has been independently established that such computation is indeed at work, experiments have consistently found group performance at the range of 50% (in dual-choice tasks). The proposed explanation is that children are aware of the innately required computation, but they cannot carry it out because of their limited working-memory resources, and they resort instead to strategies that enable bypassing it.

Previous studies have established already the 50% range of performance in the acquisition of 2 areas requiring reference set computation—coreference (Condition B) and implicatures. In this study, I examine the acquisition of stress shift and focus. I argue that computing the focus obtained by stress shift requires constructing a comparison set. Contrary to some prevailing views, there is no evidence that children have general problems with stress, but still, one finds the 50% range of performance when stress shift applies as predicted by the processing cost hypothesis. Analysis of the explanations children give for their answers has revealed that they are attempting to construct the relevant comparison derivation, but they get stuck at that stage. Combined with the analysis of individual responses, 2 bypassing strategies are found in this area: One is simple guessing, dominant in tasks involving switch reference with stress shift in which one finds individual performance at the range of 50%. The other, dominant in tasks involving semantic disambiguation, is the selection of an arbitrary default, which may be fixed for a given child across tasks. However, because the choice of the default is itself arbitrary, the group results remain at the 50% range.

1. INTRODUCTION

An idea that got much attention in linguistic theory in the 1990s is that the well-formedness of sentences is not always determined by absolute conditions, but it may be determined by a selection of the optimal competitor out of a relevant reference set. A restricted version of this was assumed (and gradually abandoned) at the early stages of the minimalist program, and simultaneously, it has been the central notion developed in Optimality Theory (OT).

The type of computation required by global reference set economy is inconsistent with what is known about the processing ability of the human parser. If it is broadly at work in the computational system (CS; grammar), then processing must be guided by independent algorithms that enable skipping the actual computation; that is, special devices are needed to enable the CS actually to be used, given the limitations on human working memory.

The introduction of the competence/performance distinction in the 1960s (Chomsky (1965)) inadvertently led to a trend in theoretical linguistics that ignored complexity considerations in the grammar. On this view, because the grammar is distinct from the parser, complexity problems that may slow down parsing are irrelevant for the research of linguistic competence because the parser operates by an independent set of rules, which may bypass the computations required by the grammar. This view precludes the possibility of a “transparent” parser in which sentence processing simply follows the rules of the grammar, with minimum parser-specific adjustments. Along with conceptual problems, this is methodologically problematic, as it means that we cannot learn anything about human linguistic competence through processing. If the parser is nontransparent, information about actual language use only gives insight into the nature of the parser.

Chomsky (2000) offered a new perspective on the relations between the CS and the processing systems (what he called “the empirical conditions” processing and acquisition). The basic idea is that an optimally designed CS—a system that optimally connects sound and meaning, namely, that satisfies the interface conditions—would also turn out optimal for the processing systems. If true, we would expect the parser to be as transparent as possible, with minimum parser-specific rules or algorithms. This would entail that highly complex computations be associated with increased processing load. It is this last entailment that I pursue here.

Global reference set computation reflects, as I argue, an imperfection of the system. If the whole CS consists of such computations, as in OT, then obviously there can be no transparent parser, and the idea of optimal design is empirically wrong. However, restricted imperfections still enable a transparent parser, with the cost of local processing difficulties where the imperfections are.

I argue that restricted instances of reference set computation are operative at the interface, precisely at those areas where the outputs of the CS do not meet the

(contextual) interface needs and adjustments are needed, that is, areas where there are, indeed, imperfections in the CS. As such, we may expect that there should be also some observable processing cost associated with these imperfections.

Reference set computation requires greater load on working memory than local computation. Adults, apparently, can cope with this load, but there is growing evidence that this load is too big for children, whose working memory is not yet as developed. Grodzinsky and Reinhart (1993) argued that the relatively rare chance pattern found in the acquisition of coreference indicates processing difficulties. The reason is that the relevant coreference strategy involves reference set computation, and children are unable to execute the computation, which, as they know innately, is required for this task. Here, I argue (in section 3) that more generally, in those linguistic areas where it has been independently established that such computation is indeed at work, one should consistently find group performance at the range of 50% (in dual-choice tasks). If this is not found, it is dubious that global reference set computation is indeed involved in the given area.

I center on the acquisition of stress shift and focus. In section 4, I argue that computing the focus obtained by stress shift requires constructing a comparison set, namely, reference set computation. In section 5.1, I show that a prevailing hypothesis that children have general problems with stress cannot explain why this difficulty surfaces only in comprehension and not in production. More generally, there is no evidence for a general stress deficiency of children, but still the 50% range of performance is found when stress shift applies. The analysis of the explanations children give for their answers, in section 5.3, reveals that they are attempting to construct the relevant comparison derivation, but they get stuck at that stage. The analysis of individual responses reveals that two bypassing strategies are found in this area. One is simple guessing, dominant in tasks involving switch reference with stress shift in which one finds individual performance at the range of 50% (section 5.2). The other, dominant in tasks involving semantic disambiguation, is the selection of an arbitrary default, which may be fixed for a given child across tasks (section 5.4).

To establish the type of computation involved, I begin with a survey of the development of the concept of reference set economy in the minimalist program, the reasons it was abandoned, and the way it relates to the hypothesis of optimal design.

2. QUESTIONS OF OPTIMAL DESIGN

2.1. Some History of Reference Set Economy in Syntax

At the transit stage from the Principles and Parameters framework to the Minimalist Program, following Rizzi's (1990) Relativized Minimality, it was noted that certain, apparently distinct, constraints on syntactic movement have something in

common that could be characterized as “least effort.” It was felt that what the bad derivations in (1) have in common is that the (italicized) moved element skips a potential landing site, which is closer; so in some sense, the movement is “longer” than necessary.

Relativized minimality

- (1) a. Head movement (HMC): *Where *find* Max *will* t the book.
 b. A-movement (super-raising): *Max seems [that *it* is certain [t to arrive]]
 c. A'-movement (*wh*-islands): *I wonder *what* you forgot *from* *whom* you got t.

In the superiority cases of (2b) and (3b), there is no intervening landing site. Still, the derivations seem longer than necessary because to check the *wh*-features of C (the head of the CP projection), the *wh*-element closer to it could move, as in the (a) derivations.

Superiority

- (2) a. Who *e* discussed what with you?
 b. */?What did who discuss *e* with you?
- (3) a. Whom did Lucie persuade *e* [PRO to visit whom]?
 b. */?Whom did Lucie persuade whom [PRO to visit *e*]?
 c. Whom did Lucie persuade Max [PRO to visit *e*]?

A striking property of the superiority restriction is that there is no way to state it as an absolute constraint on (overt) movement such as the number of barriers crossed. The distance between *whom* and its trace is precisely identical in the bad (3b) and the good (3c). So well-formedness appears here to be a relative matter: For (3c), there is no shorter movement that could check the C features, whereas for (3b) there is.

In the first implementation of the Minimalist framework (Chomsky (1993), reprinted as chap. 3 of Chomsky (1995)), movement was motivated by the need of the moved element to check its features (“greed”). Under this implementation, it was not possible to state the “shorter link” intuition locally. For example, in (1b), once one selects and merges *it* in the second cycle, there is no shorter way for *Max* to check its Case or determiner phrase (DP) features. Similarly, from the perspective of the *wh* that moved in (3b) (or (2b)), the route it took is the only (hence, the shortest) way to check its own features. Capturing this intuition required, therefore, comparing a set of competing derivations, which was later labeled “the reference set.” (A modified version of this first formulation is given in (4).)

- (4) Minimal Link Condition (the version of Chomsky (1993))
 Given two convergent derivations D1 and D2 [out of the same numeration¹], D1 blocks D2 if its links are shorter.

For the super-raising case in (1b), the relevant reference set is (5), which contains two possible derivations from the same numeration (the same “deep structure” in the previous model).

- (5) a. [(F) It seems that [(F) Max_i is certain [t_i to arrive]]]
 b. *[(F) Max_i seems that [(F) it is certain [t_i to arrive]]]

In (5a), *Max* moves in the second cycle (of *certain*) to check the feature F, and in the higher cycle, *it* is merged. (5b) is (1b). Because the link between *Max* and its trace is shorter in (5a) than in (5b), (5a) blocks (5b). Similarly, the reference set for (3b) is the pair (3a,b). (3a), with the shorter link, blocks (3b). For (3c), there is no alternative derivation that will satisfy the *wh*-feature (no alternative convergent derivation), so it is the single member in its reference set, hence, the shortest possible derivation in this set. From this perspective, the previously unrelated prohibitions on super-raising and superiority are just different realizations of one and the same principle.

The characteristic properties of reference set economy are, first, that it assumes a relative concept of well-formedness (as I showed) and, next, that it requires global computation. In (5), for example, it is useless to construct a reference set locally, at the second (*certain*) cycle, because the effects of either inserting *it* or moving *Max* are only noticeable at the next cycle. So the whole derivation must be kept open and available at that top cycle. As pointed out by Collins (1997), the problem is more general. Because (4) requires comparing only convergent derivations, the construction of the reference set is only possible at the very end where nonconvergent derivations can be filtered out.

OT, which developed around the same period, is based on the same notion of reference set economy with these two properties, although the technical details of the implementation are different. However, the OT system is much richer, assuming, first, that what needs to be checked against a reference set is not just which derivation is shorter but also a currently open list of constraints and, next, that these constraints are ranked, with possible variations of the ranking across languages.

The global nature of reference set economy disables a transparent parser. If I translate it to terms of actual processing, it requires, first, holding all nodes in the

¹The concept of numeration was introduced later. (In the original formulation, the specification was “with the same LF output,” which is obviously not what was intended. See Reinhart (1994/1998) for further discussion.)

derivation accessible in working memory until the full derivation can be completed and at the same time constructing (or attempting to construct) alternative derivations with which to compare the held material. The type of load on working memory assumed here exceeds what is known to be realistic for the human parser. The assumption shared by all processing studies (since, at least, Fodor, Bever, and Garrett (1974)) is that given the limitations of working memory, the human processor attempts to close constituents as soon as possible. Chunks of the derivation that are closed are assigned some abstract representation, and the nodes they dominate are no longer available for subsequent processing. Opening a closed constituent to access its subparts is possible but highly costly, leading to a garden-path effect. If the parser requires global reference set computation, then either nothing gets closed and eventually the overload is too big for processing or constituents constantly close and reopen (garden path). Neither is consistent with the fact that, in actual use of language, sentences usually get processed smoothly. The least we can infer is that the human parser does not operate, in processing, by computation of this kind.

A line developed to address this problem, particularly in the OT framework (although it is still found also in the earlier parts of Chomsky (1995)), is that one should not attempt to deduce the properties of the CS (competence) from properties of the parser (performance): The actual processing of derivations need not literally compute optimality, but rather some heuristic strategies, or algorithms, are developed by speakers for a quick assessment. (For some algorithms proposed for acquisition, see Pulleyblank and Turkel (1998) and Tesar (1998).) Presently, this line is hardly falsifiable given that the algorithms guiding the parser still need to be specified. But rather than dwelling on this point, we may note that if this turns out to be true, then language is not optimally designed in the sense of Chomsky (2000).

The hypothesis of optimal design (proposed at least, but not only, as a guideline for linguistic research) is that an optimal language system, namely, a system that meets the elementary interface conditions for being useable (connecting sound and meaning), also turns out to satisfy other empirical conditions such as processing and acquisition. Suppose we found a CS that is an optimal solution to the elementary interface conditions, and it still fails the other conditions (e.g., it is not fit for processing with limited working memory), then we have to add special devices such as processing algorithms. This would mean either that the strong optimal design hypothesis is false, and human language is not an optimal solution, or that our theory—the CS we found—is not, in fact, the optimal one. (For more on the hypothesis of optimal design, see Reinhart (in press-a, introduction).) Stated this way, the question is whether the problem at hand justifies this departure from the hypothesis of optimal design.

In fact, it turned out that there was no real motivation to assume the complex computation in (4) because whatever is correct about the intuition of least effort or “shortest move” can also be captured by a local computation: In chapter 4 of

Chomsky (1995), both the views of what triggers movement and of the Minimal Link Condition (MLC) are revised. “Greed” is replaced with “attract”: Movement is not triggered by the needs of the moving element but by the higher (functional) category that needs this element to be interpreted or deleted. This enabled building the MLC into the definition of *attract*:

- (6) Attract (combines Last Resort and MLC): K attracts F if F is the closest feature that can enter into a checking relation with a sublabel of K. (Chomsky (1995, 297))

From the perspective of the attracting target, there is nothing complex about finding its nearest candidate. Suppose we reached a stage in the derivation where a functional category (a feature) has been merged. At this point, we search in the chunk of the derivation we have just built for the necessary element to check it, and the search stops as soon as the first such element (going from top to bottom) is found. For example, in the superiority cases of (3), repeated in (8), the relevant state of the derivation is (7), where the *wh*-feature has just been merged at the matrix.

- (7) Q+*wh* [Lucie persuades whom [PRO to visit whom]]
- (8) a. Whom did Lucie persuade e [PRO to visit whom]?
 b. */?Whom did Lucie persuade whom [PRO to visit e]?
 c. Whom did Lucie persuade Max [PRO to visit e]?

This feature attracts now the nearest *wh*-element that it can find, which is in the complement of *persuade*. Hence, (8a) is derived, and there are no further options of continuing the search that could derive (8b). (In (8c), the first *wh* that can be found is the complement of *visit*; hence, it is this one that is attracted.)

The MLC on this view is not a relative condition but an absolute one. The first relevant element must be selected regardless of whatever other considerations may have tempted us to do otherwise. On this formulation, no reference set is constructed at all—(8b) is not ruled out by comparison to alternative options, but it is underivable. The MLC is also local in the sense that it applies as soon as the attracting node has merged, with no need to know about any potential future steps in the derivation. On the other hand, under this formulation, it is possible to observe that this absolute condition is indeed a least-effort condition in terms of actual processing: It minimizes the search, thus enforcing the quickest possible conclusion of the given step in the derivation and freeing working memory for the next task.

Conceptual issues aside, the reasons why the global reference set approach was discarded in the Minimalist framework are also empirical. Even for the small corpus examined here, one can see that the version of the reference set MLC, as

stated in (4), yields the wrong results in the case of *wh*-islands (as was pointed out in Reinhart (1994/1998)): The reference set for (1c), repeated in (10b), is (10a,b). (In terms of deep structure, both (10a,b) are derived from (9).) Recall how this was determined: With the numeration used in (10b), one could obtain all three derivations in (9, 10a,b), as well as several others. Only the derivations in (10a,b), however, converge: In (9), as well as in the other conceivable options, the *wh*-feature is not checked:

(9) I wonder [Q+*wh* [you forgot [Q+*wh* [you got what_j from whom_i]]]]

(10) a. *I wonder [from whom_i [you forgot [what_j [you got t_j t_i]]]]

b. *I wonder [what_j [you forgot [from whom_i [you got t_j t_i]]]]

Given (4), there are now two possible conclusions: either we decide that the two derivations have equally short links, or one of them is shorter than the other. (Computing here is not simple, but nothing hinges on deciding this.) In the first case, both derivations should be allowed, and in the second, one of them (the shorter one). Both these conclusions are wrong.²

It is not crucial for this discussion to go over the details of how *wh*-islands and super-raising are handled within this more recent view of least effort as a local and absolute condition. Suffice it to conclude that the properties of the CS that emerge out of this view provide no evidence for a need to impose “imperfections” such as an altogether separate parser or processing algorithms. So far, it seems that this system happens to also match the processing limitation of human users, namely, the limitation of working memory. The computation is local, which means that only chunks of the derivation that are actively at work need to be retained in working memory; the MLC imposes further acknowledgment of this limitation, forcing the quickest conclusion of operations required in a given step in the derivation, and syntactic islands define the absolute limit for search operations.³

²This in itself does not prove that the idea of reference set economy is wrong for syntax, as one may argue, very reasonably, that *wh*-islands are governed by an independent absolute constraint. Nevertheless, the problem illustrates the danger in using such strategy freely.

³For example, once the Q feature is merged in (ia), it starts the search for a *wh* it can attract. However, the search cannot reach into the syntactic island (which is why (ib) is out). Hence, no *wh*-feature can be attracted, and a derivation starting with this numeration has no way to converge. The same is true for the condition on extraction domain (CED) island in (ii):

(i) a. Q+*wh* [you resign [after Max behaved (in) what way]

b. *In what way did you resign after Max behaved t.

(ii) *Which shelf did you borrow the books on t.

In terms of processing, islands correspond to units that have been closed and stored at the stage of the derivation where the attractor is introduced. Their unavailability, again, decreases the load on working memory.

2.2. Reference Set Computation at the Interface

Although it was found irrelevant for syntax, the concept of reference set computation, in the early Minimalist Program, inspired a line of research of its role at the interface of the computational and the conceptual systems. (The development of this line can be tracked, in chronological order, in Golan (1993), Reinhart (1994/1998; 1995), and Fox (1995; 1998), among others.) The basic idea has been that in those (restricted) areas where reference set economy is active, the reference set consists of pairs $\langle d, i \rangle$ of derivation and interpretation, and it is motivated by interface needs: A given $\langle d, i \rangle$ pair is blocked if the same interface effect could be obtained more economically (i.e., there is a more economical $\langle d, i \rangle$ competitor in the reference set). Reinhart (1995) suggested that this computation is involved when an uneconomical procedure is needed to adjust a derivation for use at the interface. So, reference set computation is triggered *only* by the application of such uneconomical procedures. Obviously, this still raises the same questions of optimal design I discussed previously, and my main goal here is to address these questions. But let me first review briefly the areas where this type of complex computation is argued to be operative in Reinhart (1995; in press-a).

The first question is what is the sense of economy involved here, specifically, what counts as a noneconomical way to satisfy an interface need (what the metric is, in OT terms). I know of two types of procedures requiring reference-set formation. One is applying a superfluous CS operation (not needed for convergence)—this is the case with Quantifier Raising (QR) and stress shift for focus; the other is applying a superfluous interpretative procedure, witnessed in coreference (Conditions B, C) and in scalar implicatures. The latter has had a longer history and accumulated substantial evidence, so let me review here some of the motivation for assuming that economy considerations are involved in the first.

There is a certain resemblance in the history of the view of quantifier scope, and of focus, in theoretical linguistics. (Note that I am talking here about the question of how scope and focus are computed, or identified, and not about their semantics.) At the earlier stages, for example, Chomsky (1971), focus was viewed essentially as a property defined on Phonetic Form (PF) structures. The basic idea was that sentence stress is assigned independently by the phonological rules (Nuclear Stress Rule, or NSR) and the interface systems make use of this available stress in relating a sentence to its context, namely, signaling the focus and presupposition structure. The *focus* was defined as any constituent containing the intonation center of the sentence. This view rests on the notion of “normal, or neutral, intonation” and assumes a distinction between neutral stress and marked stress that is at times forced by discourse needs. In Keenan and Faltz (1978) and Reinhart (1983), the same has been assumed for the scope of quantifiers: Scope is determined by the syntactic configuration at the overt structure. A rule such as QR is used only when it is necessary to derive scope construal wider than the overt c-command domain, and it is viewed there as a marked operation.

The concept of markedness was problematic, however. It appears easy to find examples of covertly determined wide scope that sound perfectly natural (e.g., Hirschbühler's (1982) example, *An American flag was hanging in front of every building*). If it can be as easy to get the marked derivation as the unmarked one, it is not clear what empirical content the concept could have. Similarly, the distinction between marked and neutral stress has been challenged. It has been repeatedly argued against the NSR, or Chomsky's (1971) focus analysis, that in the appropriate context, main stress can fall anywhere, with effects hardly distinguishable from those of the neutral stress. (For an overview, see Selkirk (1984).) The crucial problem here as well is whether any content can be given to the concept of markedness. If there is no obvious way to distinguish neutral from marked stress, then one runs into the danger of vacuity—having a theory that excludes nothing regarding stress. The facts that follow from its rules are labeled *neutral* and everything else *marked*. A more realistic conclusion appeared to be that there is no sentence-level generalization governing the selection of possible foci, and any expression can be a focus, subject only to discourse appropriateness. Hence, main stress cannot be assigned at PF independently of the semantics of the sentence, and it must be the other way around: Sentence intonation reflects its independently determined focus structure.

The prevailing solution (since Chomsky (1976), where Logical Form (LF) movement was introduced) has been that both scope and focus are identified at the covert structure (LF), requiring QR and “focus movement.” But this solution is problematic as well. First, although focus movement does eliminate indeed the problem of markedness, the relations between stress and structure get to be a complex issue, raising questions of visibility of the covert structure to PF rules (stress). Next, a broader problem came to light at the stage of the minimalist program. The basic assumption is that syntactic movement is triggered only by needs of convergence (technically implemented as feature checking), and it is blind to any interface considerations. In an optimally designed language, the bare minimum needed for convergence should enable also meeting the interface conditions. This is not the case with QR and focus movement: The derivation would converge without them, so this is superfluous movement. Although technically, we can always introduce focus and quantification features, feature checking is only the implementation, and the problem remains the same. Any way one goes about it, this is a case of imperfection—one must depart from optimal design to enable the CS to meet the interface conditions.

Admitting an imperfection, we may still wonder whether it must be as sweeping as entailed by this analysis, for example, that the derivation's main stress is uniformly determined at the covert structure. Furthermore, note that this massive imperfection still does not take us very far toward capturing the actual interface conditions. Although no satisfactory content could be given to the notion of markedness, in practice, it is not the case that covert quantifier scope is always as free and easy to get as overt scope and certainly not that the so-called marked stress is

just free. Introducing the machinery of superfluous covert movement is, thus, just the very first step in stating the question when it can actually be used. Answering this question would require introducing more conditions and rules (more imperfections). One may wonder whether it is not possible to start directly with that second question, skipping the massive imperfections introduced just to state it.

In an influential work, Cinque (1993) offered a new perspective on the NSR and argued that the earlier view can be maintained. If so, then focus identification is a PF matter. Independent considerations of the CS determine that stress must be assigned to a derivation as part of its phonetic Spell-Out. At the interface, this property of sentences is used to facilitate communication, using stress to identify the focus as in the view of focus in Chomsky (1971). This line was extensively developed by Zubizarreta (1995; 1998), although she still assumed that focus is also marked at LF (Neeleman and Reinhart (1998), Reinhart (1995; 1998), Szendrői (2001)).

Note, however, that the analysis is based on a revival of the distinction between neutral and marked stress: When the stress assigned by the NSR is not appropriate to the context, special stress shift operations apply, yielding marked stress. So the question, “How does one know it is marked?,” is relevant again.

In Reinhart (1995; 1998), I argued that it is a mistake to search the evidence for markedness in the realm of direct intuitions. A marked derivation is a derivation that involves a superfluous—hence uneconomical—operation. When this is done with no reason, the result is visibly awkward. But if using the uneconomical derivation is, decisively, the only way to satisfy a certain interface need, the result sounds perfectly fine, and it is only indirectly that we can see that it is nevertheless marked, or uneconomical. (In the case of QR, Fox (1995) provided ellipsis evidence consistent with the claim that QR does not take place when not needed for interpretation.) I clarify the details of how this works with stress shift shortly.

In more precise terms, what is claimed in the last paragraph is that computing QR and stress shift involves constructing a reference set and checking whether it does not contain a more economical $\langle d, i \rangle$ pair, namely, a pair derived without the superfluous operation. If it does, the derivation is blocked (*viz.*, if one nevertheless produces it, it has a visibly marked air). This brings me back to the question of optimal design.

If correct, this analysis still means that this is an instance of imperfection in the language system: That superfluous operations can apply at all is, as I argued, an imperfection of the core system. However, it is substantially less massive than what had to be assumed before. First, PF procedures like stress operate, as they should, on the overt structure. Next, the superfluous QR and stress shift cannot apply all over the place but are restricted by reference set computation. On the other hand, as I argued, reference set computation has a serious processing cost, which is an imperfection on the secondary requirement of meeting the empirical conditions of use. Here too, the problem is far less massive than that demonstrated

by the MLC because reference set computation is triggered only if a superfluous operation applies. Nevertheless, in these restricted cases, there is an imperfection.

The strongest interpretation of the concept of imperfection is that, if we have to admit it into our theory, there should also be some way to observe the imperfection in the (use of) language itself. What I turn to show is that computing reference sets, as needed here, is indeed an imperfect way to meet interface needs—it exists, but it has a heavy observable processing cost.

3. THE PROCESSING COST OF REFERENCE SET COMPUTATION IN ACQUISITION

The theoretical prediction of this framework is that whenever reference set computation is involved, there should be some evidence of processing complexity. Apparently, the processing load posed here is not beyond the processing ability of adults (given, e.g., that stress shift is used easily and frequently). However, unexpected evidence that the processing required may exceed children's processing ability came from the studies of the acquisition of Condition B. (This is unexpected because the studies that first found it were not set up with this purpose in mind, which is often the strongest evidence there is.)

Chien and Wexler (1990) pioneered in establishing the basic generalization regarding the acquisition of anaphora. Based on experiments with a large number of children (177, ages 2;6–7;0), Chien and Wexler showed that acquisition delays are found only with coreference and not with the Binding Theory in general. Children perform well on the variable-binding aspects of the Binding Theory. Thus, they rule out successfully the anaphora interpretation in (11), where the pronoun can only be interpreted as a bound variable. But they performed poorly on coreference in Condition B environments as in (12):

(11) Every bear touches him.

(12) Teddy bear touches him.

This conformed with Reinhart's (1983) theoretical conclusions that variable-binding and coreference are governed by different types of linguistic conditions and that the coreference aspects of Condition B (and C) require more complex computation than variable-binding. In present terminology, the coreference procedure proposed in Reinhart (1983) or in Grodzinsky and Reinhart's (1993) Rule I requires that in configurations that, in principle, also allow variable-binding (*viz.*, when one of the arguments *c*-commands the other), a reference set of $\langle d, i \rangle$ pairs must be constructed to determine whether coreference is also allowed. Thus, to determine whether (12) allows the coreference construal (13a), the representation (13b) must be constructed for the same derivation:

- (13) a. Teddy bear touches him and *him* = *Teddy bear*.
 b. Teddy bear (λx (x touched x)).

If (13a) and (13b) are semantically indistinguishable in the given context (as is the case here, where the context is null), coreference is blocked. The underlying assumption in these studies is that variable-binding is a more efficient way to obtain anaphora than coreference. So whenever possible (in configurational terms), it is the only anaphora option, unless this interpretation is inappropriate to the given interface needs, that is, unless the intended reading cannot be expressed with variable-binding.⁴

The acquisition question has been, why should this difference entail a delay in children's performance on coreference? In Reinhart (1983), the condition on coreference was perceived as belonging to pragmatics. It involved an inference based on knowledge of grammar, meaning, and appropriateness to context, and I believed it could be viewed as an instance of Gricean generalized implicatures. Chien and Wexler (1990) formulated a similar intuition and argued that children's coreference performance reflects a delay in acquiring the context considerations underlying a pragmatic principle.

Grodzinsky and Reinhart (1993), however, took a different perspective on this question. Grodzinsky and Reinhart's point of departure has been another seminal result of Chien and Wexler's (1990) study. Virtually all studies on the acquisition of coreference have found not just vague poor performance but results ranging around 50% adult-like answers. Such figures are, in principle, consistent with chance performance. Chien and Wexler conducted careful statistical analysis, including individual data, and found that many of the children performed at chance level individually, namely, they sometimes answered *yes* and sometimes *no* on the same condition. Based on these findings, Grodzinsky and Reinhart argued that the statistics of children's performance should be taken as a significant part of the data. A 50% range of performance (which in individual experiments can range between 35% and 65%) is not very common in acquisition. In many tasks, children

⁴A relevant question is why variable binding is more efficient than coreference. Reinhart (1983) attempted an answer in terms of the Gricean maxim of manner ("be as explicit as the conditions permit"). Levinson (1987) showed that it cannot be the manner maxim and developed a more plausible line in terms of the Gricean maxim of quantity (active in scalar implicatures). A more recent approach is in terms of semantic processing: Variable-binding is less costly because it enables immediate closure of open properties, whereas coreference requires that the property is stored open until one finds an antecedent for the variable. This approach underlies Fox's (1998) analysis of Rule I as reference set computation of the shortest link. Reinhart (2000) explored the possibility that the reason why coreference is inefficient in Rule I environments is that it enables bypassing a CS restriction: Variable-binding is excluded in these environments, so if coreference just sneaks in precisely the same excluded interpretation in the discourse door, this is not the optimal way to communicate in a more global sense of optimality. (Communication optimality requires reducing, rather than increasing, the competing interpretative options of a given derivation.) In any case, under all these accounts, the computation requires constructing and comparing a reference set.

perform at close to a 100% range (viz., 80% to 95% adult-like responses). In cases in which they set the parameter wrongly, the experimental design should be able to elicit close to 0% correct results. There are other conceivable statistical patterns, but results at the range of 50% should be explained because they are consistent with chance, and if, indeed, they can be shown to be chance performance, this is a significant finding that cannot be reduced to just not knowing the relevant linguistic rule. (In the acquisition literature, 50% findings are often summarized as “children performed poorly” and are not distinguished from below-chance performance.) If children just do not know a rule, they should respond uniformly, most likely following the established *yes* bias.

Grodzinsky and Reinhart (1993) argued that the account for the coreference delay should also explain this specific pattern and suggested that it is the computational complexity, rather than just the appeal to context, that explains children’s difficulties in the relevant tasks. In Grodzinsky and Reinhart’s account, assuming that all linguistic knowledge is innate, children know that they have to construct a reference set and check whether the interpretation needed in the given context justifies selection of coreference. So they start execution. However, the processing poses too big a load on their working memory, which is known to be less developed than that of adults. (For a survey of the literature on children’s working memory, see Gathercole and Baddeley (1993), Gathercole and Hitch (1993).⁵) Specifically, children’s working memory is not developed enough to hold the materials needed to complete the execution of this task—namely, constructing an alternative derivation while holding the previous one and then comparing the two derivations. Failing the execution, they resort to some strategy. In the case of coreference tasks, the strategy turns out to be simple guessing, but as I show, other bypassing strategies are found in other areas.

In subsequent work on Condition B, much effort was invested in proving that the 50% results do not, in fact, indicate guessing (e.g., that some children master the rule and some do not). Nevertheless, by now there is evidence that the individ-

⁵The working memory system should not be confused with memory resources in general (long-term memory). The *MIT Encyclopedia of Cognitive Sciences* defines the working memory system as follows: “Cognitive scientists now assume that the major function of the system in question is to temporarily store the outcomes of intermediate computations when problem solving, and to perform further computations on these temporary outcomes (e.g., Baddeley (1986))” (Smith (1999, 888)). There are no claims that children’s general memory resources are limited. The huge amount of information that children manage to learn suggests, in fact, the opposite. Smith (1999) explained that the view of working memory as a gateway to long-term memory has been undermined by neuropsychological studies that have found that there are patients who are impaired on working memory tasks but perform normally on long-term memory tasks. Note also that the precise details of how working memory develops—whether memory capacity itself increases or only efficiency in allowing more resources to be employed in storage—is a subject of debate. However, these details are not important for this discussion, because either way, children’s working memory has been found not to operate as efficiently as that of adults.

ual performance (of at least most children) on Condition B is by pure chance. To establish whether 50% performance indicates a guessing pattern, statistics of individual child performance should be compared to the binomial model of results expected by pure chance (e.g., in a coin-tossing experiment). As Thornton and Wexler (1999) showed, the statistical analysis of individual performance in their experiments is consistent with this model, at least for most of the children they studied.

Thornton and Wexler (1999) adopted Rule I, under its reformulation in Heim (1998), but they raised several arguments against the processing account and concluded that there is no reason to assume children have any difficulties in processing reference set computation of the type required by Rule I. Thornton and Wexler maintained that the coreference delay reflects a pragmatic deficiency, and they developed an analysis of the pragmatic factors underlying Rule I that children have not acquired yet. I address their arguments in detail in Reinhart (in press-b), including problems raised by the acquisition of coreference in Condition C environments, and argue that they do not provide evidence against the processing account. Furthermore, their pragmatic account cannot explain the chance performance found in their own study. I do not repeat the arguments here, but in any case, this pragmatic account cannot be relevant in the case of stress shift, to which I turn next.

Once a processing failure is established in one of the areas of reference set computation, we can turn to the other areas where this computation has been established. The empirical criterion I have posed here for whether a reference set is indeed involved in a given computation is that, if so, we should find evidence for a greater processing load. Translated to the domain of acquisition, we expect to find out that children are unable to carry out the required computation, and they resort instead to strategies enabling bypassing it. The simplest option that has been observed so far is individual guessing, but in areas involving semantic disambiguation, there is another option of establishing an arbitrary default, a strategy that I examine in detail in section 5.4. When this option is available, it is possible that individual children fix a consistent default across all tasks in a given area. If this happens, one would find a consistent individual response pattern. In the relevant dual-choice experiments, fixing on one of the defaults would consistently yield an adult-like answer and fixing on the other—the wrong adult answer. Even if the child fixes a default in advance for all tasks, however, the choice of the default is still arbitrary. For this reason, the overall group performance would still be at the chance range of 50% (in dual-choice experiments). It may appear possible to interpret such findings as indicating that some children have mastered the required computation, whereas others have not, but then it is not trivial to explain why this division is always in the 50% range (why exactly half the children of the same age group in all relevant experiments reached adult performance and why the same is not found all over the place).

The empirical prediction of this analysis, then, is that in all areas of reference set computation, one should find children's performance at the 50% range, whether this is obtained by the guess or by the default strategy. This is a strong claim, meaning that if no such performance is found, then no reference set computation can be involved. Of the four established areas, evidence is already available, along with Rule I, for the acquisition of scalar implicatures. Chierchia, Crain, Guasti, Gualmini, and Meroni (2001) and Gualmini et al. (2001) found a 50% range of group performance in children's understanding of implicatures. They showed that processing an implicature requires the construction of a reference set and that children fail in the step of constructing the comparison member. This is an area of semantic disambiguation in which the default strategy is possible; hence, the pattern found is of mostly consistent individual response. In the area of QR, not much relevant experimental work has been done to my knowledge.⁶ This is a more difficult area to test because of the interference of other factors, such as the phenomenon of quantifier spreading, in children's interpretation of quantifiers.⁷ A possible experimental design specific to QR is outlined in Reinhart and Szendrői (2003). In any case, the empirical criterion I propose here still holds. If the 50% pattern is not found, once the work is done, this would falsify my analysis of QR (in Reinhart (in press-a)) and will confirm Fox's (1998; 2000) analysis in which QR involves local, rather than global, reference set computation and hence is not a costly operation.

Here I examine the acquisition of the computation involved in stress shift and focus. As we see, the group results on all tasks are at the 50% range, but of the two areas examined, one (switch reference) allows only for the guess strategy, whereas the other (focus disambiguation) is an instance of arbitrary defaults. First, however, I need to be more specific on the details of how reference set computation works in the case of stress shift and focus.

4. AN OVERVIEW OF STRESS AND FOCUS

In this section, I survey the major conclusions of the analysis of stress shift and focus in Reinhart (1998), Neeleman and Reinhart (1998), and Reinhart (in press-a, chap. 3) but without repeating the motivation, argumentation, and evidence.

⁶Some recent experimental attention has been devoted to the relative scope of quantifiers and negation (e.g., Lidz and Musolino (2002); Musolino and Lidz (2003)). However, this problem involves reconstruction rather than QR—the subject quantifier reconstructs back into its Spec VP position, which enables interpreting it in the scope of negation. No reference-set computation is involved in reconstruction, so these issues are irrelevant to this question.

⁷Children reject utterances such as *Every boy is riding an elephant* in contexts in which adults would judge it to be true—for example, three boys are riding three elephants, and there is an additional elephant (Inhelder and Piaget (1964); Philip (1995); cf. Crain et al. (1996)). Another factor argued to exist is children's aversion to wide-scope indefinites (Krämer (2001)).

4.1. Neutral Main Stress (Cinque (1993), Szendrői (2001))

As mentioned in section 2.2, Cinque (1993) revived the view of the 1970s that main sentence stress is determined independently of focus or discourse considerations and that a distinction between neutral and marked stress is therefore feasible. Cinque's implementation rests on reanalysis of the NSR, which determines for each derivation where its main stress falls. The basic framework of his analysis is the Metrical Grid theory of Halle and Vergnaud (1987)—the NSR starts the assignment of stress with the most deeply embedded constituent, which then moves up to the next metrical line. The outcome will be, then, that the most prominent stress falls on this constituent. The gist of Cinque's analysis is that the depth of embedding (in the case of sisters) is determined by the direction of selection (or recursion, as Cinque phrases it) of the given language. In both a VO language such as English and an OV language such as Dutch, the most deeply embedded constituent in (14) is the object. Hence, in both, the object receives main stress (throughout I use bold to indicate the main sentence stress):

- (14) a. I read the **book**.
 b. dat ik het **boek** las.
 that I the book read.

Szendrői (2001) presented an alternative technique for the execution of the neutral main stress rule. She used Liberman's (1979) metrical tree notation. In this method, there are no separate cycles like the NP cycle or the VP cycle assumed in Cinque's system. Rather, stress is assigned to the nodes of the syntactic tree (or alternatively, the prosodic structure). An advantage of this system is that it is fully transparent how it applies to syntactic (or prosodic) trees, and thus, it lends itself to strictly incremental application.

The technical details of the NSR are not crucial for this discussion. What is crucial is that main stress is uniformly assigned to derivations by a rule that is independent of focus considerations. I refer to this rule as the neutral main stress rule but continue to abbreviate it as NSR. In Reinhart (in press-a, chap. 3), I addressed some of the arguments raised against the idea that a uniform neutral main stress rule can be postulated and argued that many of them rest on conflating focus stress and anaphoric destressing, which, as I show next, is independent of the NSR.

4.2. The Focus Set

Cinque's (1993) theory of sentence stress enables reformulation of the idea that focus is marked overtly, at PF, rather than covertly. In Reinhart's (1995) execution, each derivation is associated not with an actual focus but with a set of possible foci, that is, a set of constituents that can serve as the focus of the derivation in

a given context. This set is determined by the CS at the stage where both the syntactic tree and stress are visible. In other words, focus selection applies to a pair $\langle \text{PF}, \text{LF} \rangle$ of sound and configurational structure. The focus set is defined, then, in (15). If stress falls on the object, either in English SVO structures or in Dutch SOV structures, the focus set defined by (15) is the one in (16).

- (15) The focus set of IP consists of the constituents containing the main stress of IP.
- (16) a. $[_{\text{IP}} \text{ Subject } [_{\text{VP}} \text{ V } \text{ Object}]]$
 b. $[_{\text{IP}} \text{ Subject } [_{\text{VP}} \text{ Object } \text{ V}]]$
 c. Focus set: {IP, VP, Object}

This means that in actual use, any of the members of the set in (16) can serve as focus. At the interface, one member of the focus set is selected as the actual focus of the sentence. For illustration, (17), which is generated with stress on the object, can be used as an answer in any of the contexts in (17), with the F-bracketed constituent as focus.

- (17) a. My neighbor is building a **desk**.
 b. Speaker A: What's this noise?
 Speaker B: $[_{\text{F}} \text{ My neighbor is building a desk}]$
 c. Speaker A: What's your neighbor doing these days?
 Speaker B: My neighbor $[_{\text{F}} \text{ is building a desk}]$
 d. Speaker A: What's your neighbor building?
 Speaker B: My neighbor is building $[_{\text{F}} \text{ a desk}]$

If this were the complete story, one could conclude that language is almost perfect in meeting the interface need of associating a focus with derivations because an obligatory PF rule, needed for phonological convergence, is sufficient for signaling the focus. This system still generates semantically ambiguous derivations (each consistent with several focus construals). However, the set of interpretative options is clearly defined and restricted. In reality, however, the set of possible foci the system generates is not sufficient for the interface.

A derivation is inappropriate for a given context if no member of its focus set can be used as an actual focus in that context. (17a), for example, cannot be used as an answer in either of the contexts of (18) (# indicates throughout inappropriateness to context).

- (18) a. Speaker A: Has your neighbor bought a desk already?
 Speaker B: #No, my neighbor is $[_{\text{F}} \text{ building}]$ a **desk**.
 b. Speaker A: Who is building a desk?
 Speaker B: # $[_{\text{F}} \text{ My neighbor}]$ is building a **desk**.

This is so because in the contexts of (18), the F-bracketed constituents should be the foci, but these constituents are not in the focus set generated by (15) for a sentence in which the object bears stress (cf. (16)).

4.3. Stress Shift Operations

For cases such as (18), where the focus set defined by the neutral stress does not contain the desired focus, a special stress-shifting operation applies. For this discussion, it suffices to state it schematically, as in (19), which follows Neeleman and Reinhart's (1998) implementation. It applies to a given output of main stress assignment and while keeping this assignment adds stress to another word. The formulation of the rule, as adding exactly two stars, is for ease of illustration only. For the precise formulation, see Szendrői (2001) and Reinhart (in press-a, chap. 3). In (20), the result is that main stress is on *my neighbor*, but the original stress on *desk* remains as a secondary stress.

(19) **Main stress shift:** Add two stars.

* * *
* * *

(20) My neighbor is building a desk → My neighbor is building a desk.

In the context of (18a), repeated in (21a), extra stress is assigned to the verb. As a result, the verb is in the focus set defined by (15), and the derivation is appropriate in this context. In (21b), the same operation applies to the subject.

- (21) a. Speaker A: Has your neighbor bought a desk already?
 Speaker B: No, my neighbor is [_F **building**] a desk.
 b. Speaker A: Who is building a desk?
 Speaker B: [_F **My neighbor**] is building a desk.

The output of (19) is what I called *marked stress*. Although they sound perfectly natural in their context, the foci in (21) are marked because their derivation violates the neutral main stress rule (NSR)—it is obtained by a superfluous operation that undoes the results of the NSR. I return directly to the symptoms of markedness.

The effects of (19) are often confused in discussions of marked stress with the effects of a different process of anaphoric destressing. This distinction, proposed in Selkirk (1984), is discussed in Reinhart (1998) and Neeleman and Reinhart (1998) who argued that the latter is completely independent of considerations of the focus set. For reasons discussed there and in Reinhart (in press-a, chap. 3), it does not generate the same markedness effects. (In terms of the next subsections,

it allows focus projection, and it does not require reference set computation).⁸ The destressing operation can be stated schematically as in (22). It applies locally, to any anaphoric constituent, independently of the NSR. This can be captured by assuming that it applies at the word level prior to the NSR. Thus, the relevant D- (discourse) linked or anaphoric expressions do not carry an intonational star when the NSR applies as in (23):

(22) **Anaphoric destressing:** Remove a star (prior to the NSR).

- (23) a. Destress: Max [saw her]
 * *
 *
 b. NSR: Max [saw her]

The NSR, then, just operates in the standard way, turning the most embedded star into the main stress. Because in (23a), the lowest star is on the verb, it is the verb that will carry main stress, as in (23b).

Destressing can apply to units larger than a word. Typically, when it applies to a whole VP, the destressed VP can also be unpronounced, giving rise to VP ellipsis. This is illustrated in (24). Because the VP in the second conjunct is anaphoric, it is destressed. Main stress then is assigned to the only possible candidate, namely, the subject, as in (24b). The VP could either be pronounced, as in (24b), or mispronounced (“deleted at PF”), as in (24c).

- (24) a. Destressing: First Max [touched Felix_i] and then Lucie [touched him_i]
 * * * *
 * *
 b. NSR: First Max [touched **Felix**_i] and then **Lucie** [touched him_i]
 c. PF deletion: First Max [touched **Felix**_i] and then **Lucie** did [e]

- (25) Stress shift: First Max_i [touched **Felix**] and then *Lucie* [touched **him**]_i
 * * *
 * * *

Note finally that, if that is what the context requires, it is possible to make an anaphoric destressed constituent the focus. This would of course require applying stress shift. If we apply this to the derivation (24), we obtain (25). Foci obtained this way are often described as “contrastive,” but for all of my purposes, they are just standard applications of the stress shift operation.

⁸In this discussion, I do my best to abstract away from anaphoric effects, which should be helped by the use of a verb of creation such as *build*.

4.4. Reference Set Computation

The widely acknowledged characteristic of the focus obtained by shifted (marked) stress is that it “does not project” (it can only be “narrow focus”). As was seen, stress obtained by the neutral NSR allows any projection containing it to serve as focus, for example, the whole IP in (17b), repeated. The shifted cases of (21), by contrast, cannot be used in the same context, as in (26), which means that they do not project IP as focus. Similarly, stress shifted inside the VP does not project VP as focus as seen in the comparison of (17c) and (27).

- (17b) Speaker A: What’s this noise?
 Speaker B: [_F My neighbor is building a desk]
- (26) Speaker A: What’s this noise?
 Speaker B: #[_F My neighbor is building a desk]
 #[_F My neighbor is **building** a desk]
- (17c) Speaker A: What’s your neighbor doing these days?
 Speaker B: My neighbor [_F is building a desk]
- (27) #My neighbor [_F is **building** a desk]

Although widely discussed, such facts did not receive a satisfactory account. Standard approaches postulate a special focus projection rule for contrastive focus. However, it is far from obvious how one can distinguish (in a noncircular way) the “contrastive” (27) from the “standard” (17d) (*Speaker A: What’s your neighbor building? Speaker B: My neighbor is building [_F a **desk**]*), given that in both, the focus is narrow.

Within the present framework, the difference between neutral and contrastive stress is that the latter is derived through a violation, or extension, of the NSR. Generally, violations of core principles of the CS are never allowed to apply superfluously. They are permitted just in case this is the only way to satisfy an interface requirement. In the specific case of focus, I have already shown that the stress system is not perfect for the interface, as it does not generate the full set of foci that may be needed in context. To overcome this, stress shift applies to generate a focus not available otherwise. However, the price for resorting to such extension of the system is always the same: Reference set computation is required to determine whether indeed this is the only possible way to reach the interface goal. In the case of focus, the outcome is that shifted stress cannot project if the given projection is also available with neutral stress. Let us see how this works.

I assume, first, just the one definition of the focus set in (15), repeated, which is blind to how stress is assigned. Hence, for the derivations at hand, the focus sets defined are those in (28b), (29b), and (30b).

- (15) The focus set of IP consists of the constituents containing the main stress of IP.
- (28) a. My neighbor is building a **desk**.
b. Focus set: {IP, VP, Object}
- (29) a. My neighbor is **building** a desk.
b. Focus set: {IP, VP, V}
- (30) a. My **neighbor** is building a desk.
b. Focus set: {IP, subject}

The focus sets of (28) and (29) intersect in the case of IP and VP. Suppose that in a given context, one wants VP (or IP) to be the focus. One could obtain this result by using (28), without applying the superfluous stress shift. Hence, (29) is ruled out for that context. The only focus of (29) not already in the focus set of (28) is the verb. Hence, it is only the need to use this focus that can motivate the stress shift. Similarly, (30) intersects with (28) on IP. Hence, (30) can only be used with the subject as focus.

Computing this type of reasoning requires a construction of a reference set, which consists of <d,i> pairs of a derivation and interpretation. In this case, the relevant interpretation is a selection of a focus out of the focus set. So, suppose the task is to decide whether (29) can be used with the IP as focus (as in the context of (17b)—*What's this noise?*). The reference set is (31).

- (31) a. d: My neighbor is building a **desk**. →
My neighbor is **building** a desk.
i: Focus: IP
b. d: My neighbor is building a **desk**.
i: Focus: IP

Because the pair in (31b) does not involve the extra-operation, it blocks (31a).

Suppose now one wants to use (29) with the verb as the focus, as in the context of (17c). Because stress shift is involved, one has to construct a reference set there as well. However, the reference set is (32), which contains only this one member because no other derivation (of the same numeration) has the verb as focus. Hence, this derivation is allowed.

- (32) d: My neighbor is building a **desk**. →
My neighbor is **building** a desk.
i: Focus: V

On this analysis, stress shift imposes a costly complex computation. At first glance, it may appear stipulative to postulate such a complex solution to the prob-

lem of focus projection with stress shift. In all cases under consideration so far, the actual focus ends up the narrow one, namely, the smallest projection of the stressed node. If this is the correct descriptive generalization, several simpler accounts may suggest themselves. One would be to change the definition of the focus set. Rather than a unified definition for all forms of stress, one could assume that when the stress shifts—namely, when the nondefault stress is used—the focus set contains only this one member. In frameworks assuming a focus feature, the same could be obtained by associating the feature with the type of stress. However, this descriptive generalization is not, in fact, true. There are configurations that allow shifted stress to project beyond the smallest projection.

(33) [The man with the **martini**] is the murderer.

- (34) a. [Did the man with the apron] commit the murder?
 b. Who is the murderer?

In (33), main stress falls on a nonneutral position (the last noun in the subject). By the generalization under consideration, the focus set for (33) should contain then only the lower DP, *the martini*. The derivation can certainly be used with this constituent as a focus, for example, as an answer to the question (34a). But it can also be used as an answer to (34b) in which case the whole DP, *the man with the martini*, is the focus. So the nondefault stress does project up to the top DP. However, it is still impossible to use this derivation with the full IP as focus. If focus projection is determined by reference set computation, this is the expected outcome because, unlike IP, the top DP is not in the focus set of the derivation with neutral main stress (which would fall on *murderer*).

In conclusion, whenever stress shift applies, a reference set must be constructed to check its appropriateness. This means that this operation entails a computational complexity whether the final outcome is “in” or “out.” On the other hand, there is no reason to assume any reference set computation in derivations involving no stress shift (as would be assumed in some optimality approaches). So, the difference between neutral and marked stress is that the latter requires computational complexity not involved in the former.

5. GUESS AND DEFAULT IN THE ACQUISITION OF STRESS SHIFT

5.1. Preliminaries

The question whether reference set computation must be assumed for stress shift foci resembles that posed by Rule I. In both, the empirical evidence that it is indeed needed is not huge. In most instances, an alternate simpler theory would ap-

pear to capture the facts just the same. For this reason, acquisition findings may have a crucial role in evaluating the proposed analysis.

In the case of Rule I, the actual contexts that allow a pronoun to corefer in apparent violation of the binding requirement of condition B are rather minimal, so the assumption that a complex reference set computation is involved in coreference resolution may seem speculative. Similarly, in the case of focus, there are only very limited contexts that motivate the assumption that such computation takes place, namely, the cases where the stress shifted, but the focus still projects (e.g., in subject DPs). The analysis of stress shift as requiring reference set computation rests on the assumption that there is a fixed definition of a focus set. So even when the stress is shifted, any constituent containing the stress is a potential focus of the derivation. This is what enforces constructing a comparison set to determine whether any of the candidates could have been derived also with neutral stress, in which case it is ruled out as a focus of the derivation with stress shift. As I mentioned, this assumption may seem the most speculative part of the analysis because it appears easy to capture most of the facts without it. It turns out, however, that this aspect of the analysis is easiest to test experimentally, and as I show in section 5.3, it gets direct confirmation: Given children's failure in processing sentences with stress shift, the experimental design enables examining which interpretations they are considering. There is clear evidence that when the stress shifts to the direct object in, for example, *Lucie sold a car to Max*, children still access the VP focus (*sold a car to Max*). This, then, is consistent with the hypothesis that a computation of the sort assumed here is at work and furthermore, that children are aware of the computation required by stress shift and initially attempt to carry it out.

If reference set computation is indeed at work here, the most immediate expectation is that we should find not just some vague acquisition delay but performance at the range of 50%. This general expectation is confirmed in the available studies. Several studies mentioned below have noticed a correlation between children's performance on the coreference aspects of condition B, or Rule I, and their performance on tasks involving contrastive stress. In the latter as well, group statistics have been found to be around 50% adult-like answers. As mentioned, in the present framework, there is no special status to the so-called contrastive stress, and it is viewed as an instance of the stress shift operation. The findings, then, bear directly on the question how stress shift is acquired.

McDaniel and Maxfield (1992) studied two contexts of what they viewed as contrastive stress. One involves focus selection and the other stressed pronouns of the switch reference type that I examine in section 5.2. The focus selection task is illustrated in (35).

- (35) Experimenter: Bert doesn't want to eat the **big** strawberry. What do you think he wants to eat?

Props: Bert, a big strawberry, a little strawberry, a big tomato, a little tomato, a pear, a carrot, an orange, and a green pepper

Under the analysis in section 4, the contextually appropriate answer to the experimenter's question rests on selecting a focus out of the focus set of the first clause, given partially in (36b).

- (36) a. Stress shift: Bert does not want to eat the **big** strawberry.
 b. Focus set: {big, the big strawberry, eat the big strawberry . . . }
- (37) a. Neutral stress: Bert does not want to eat the big **strawberry**.
 b. Focus set: {strawberry, the big strawberry, eat the big strawberry . . . }

The question clause of (35) restricts attention to the first two members of (36b) (because it presupposes Bert's wanting to eat something). So the child needs to select one of these two in order to pick an object in the props inventory. If the narrow (adjective) focus *big* is selected, as in the adults' response, then the only choice is the little strawberry. If the full DP focus, *the big strawberry*, is selected, then Bert does not want to eat any strawberry, so any of the nonstrawberry objects in the props inventory can be selected. The choice of focus, however, is not free. Because stress shift applied, a reference set must be constructed with the neutral stress derivation (37a). Because the wide DP focus is also in the focus set of (37), it is excluded as a candidate for the focus of (36). On my hypothesis, children know that this computation is required, but they are unable to execute it. Hence, their choice will be arbitrary.

In McDaniel and Maxfield's (1992) experiments, children were given 1 point for each correct answer, with the maximum score of 10 points. The mean score for the two experiments together was 5.5, meaning that 55% of the answers were correct, a result consistent, in principle, with chance performance, although this pattern cannot be established without the individual data. McDaniel and Maxfield did not score the results for (35) separately but rather combined the results of their two experiments (which dealt with very different contexts). They also reported an earlier experiment (Tavakolian (1974)) using similar tasks and methodology as (35) with similar results. In the next subsections, I examine later experiments in both these contexts, and we will see that the same overall performance in the range of 50% was found in all studies on the acquisition of stress shift. In one of the instances of contrastive stress, the switch reference contexts discussed in section 5.2, analysis of the individual results reveals a pattern of guess. In the other, a default is involved.

Nevertheless, such results are not yet sufficient to establish the analysis under consideration. In principle, it cannot be excluded that there are other linguistic tasks at which children perform at the range of 50%. (The entailment goes only one way—that if reference set computation is involved, there is performance at

50% but not that performance at 50% entails reference set computation.) So, in principle, other accounts for this performance range are possible.

McDaniel and Maxfield (1992), for example, suggested that the source of children's difficulties here is their inability to perceive stress and thus to use it as a clue for interpretation. McDaniel and Maxfield argued further that the clue for permitting coreference in violation of condition B (Rule I) is contrastive stress. Because children are unable to perceive stress, this also explains their poor performance on ruling out coreference. In Reinhart (in press-b), I argued that the coreference options in Rule I contexts are not, in fact, determined by stress. So stress problems cannot, anyway, explain the coreference findings. But let us focus here on children's perception of stress. On the present hypothesis, what the coreference and the stress shift tasks have in common is that they both involve reference set computation, although the specific required computations are unrelated. The relevant question for this hypothesis is whether children's difficulties with stress shift could be explained by an independent general problem with identifying stress. Indeed, popular hypotheses in studies of the acquisition of sentential prosody are that children are insensitive to stress: Either they do not perceive it at all, or if they perceive it, they cannot use it as a key to disambiguation in comprehension tasks (although they have no problem in production tasks).

Note, however, that these studies have reached this conclusion based on experiments with stress shift (e.g., Cutler and Swinney (1987); Maratsos (1973); and the references cited there). Under consideration here are two different explanations of children's difficulties with stress shift: one is that the problem is specific to stress shift because only in this case reference set computation is at work; the other is that it reflects general stress problems. It is impossible to decide that the second is right based only on findings of acquisition with stress shift. To settle the issue, children's performance on stress in other areas must be determined.

Regarding the first hypothesis—that children simply do not perceive stress at all—it is interesting to note that in one of their experiments, Cutler and Swinney (1987) checked the perception of stress in a string of arbitrary words that did not form a sentence, but one of them carried stronger stress. They found out that children processed the stressed words significantly faster than the nonstressed words. (In the experiment, Cutler and Swinney measured response time to a target word when it is accented and when it is not.) This contrasts with their experiments with actual sentences. They concluded that children can perceive stress, so the difficulties must lie at the sentence level.

The second hypothesis is that at the sentence level, children cannot use stress for disambiguation or for parsing the information structure. Halbert, Crain, Shankweiler, and Woodams (1995) examined this question with pairs like (38) where no stress shift is involved.

- (38) a. Big Bird threw [the fish] **food**
 b. Big Bird threw the [**fish**-food]

Without stress, the word-string in (38) is ambiguous between the syntactic construal (38a) and (38b). When pronounced (i.e., once the main stress rule applied), there is no ambiguity: In (38a), where *food* is the theme argument, its stress is projected as the sentence stress. In (38b), the most deeply embedded argument is *fish-food*, which gets the standard compound stress. This stress then projects as the main sentence stress. In both derivations, stress is assigned by the NSR, and no stress shift applies. If children have problems either with the NSR or with disambiguating via stress, they should perform badly on distinguishing these sentences. Under my hypotheses that the main stress rule is both innate and computationally simple and that it is only stress shift that requires reference set computation, they should do well here. In Halbert et al.'s study, children were given a story for each sentence. In this context, the sentence would be false for adults. Of the children participants, 89% (16 out of 18, ages 3;0–5;3) indeed gave the correct answer (i.e., judged the sentence as false). This single experiment is not yet sufficient to establish the point, but the results suggest that children have no independent comprehension problems in using stress for disambiguation.⁹

The most serious problem for the view that children have independent difficulties with contrastive stress is that all findings have pointed out that the problem arises only in comprehension tasks, whereas in production, children are able to use contrastive stress. In Hornby and Hass's (1970) experiment, for example, surveyed in Thornton and Wexler (1999), children ages 3;8 to 4;6 were shown pictures and asked to describe them. The pictures came in pairs differing in only one element. When the pictures showed different agents (a boy or a girl patting a dog), children stressed the subject 80% of the time in their description of the second picture. On my analysis, this is obtained by stress shift, so the experiment shows that they are able to apply this procedure in production. As pointed out by Thornton and Wexler, this inconsistency of comprehension and production remains a mystery in subsequent studies.

It is not easy to imagine how this difference can be explained within the view of stress deficiency. However, within the analysis of stress shift I propose here, production and comprehension do not involve the same sort of difficulty, although both require more processing than with neutral stress. In production, language users know which focus they intend for the utterance. Considering all my assumptions, in the derivation that the producers execute, the NSR would still need to apply, first because this is an obligatory rule of the phonological component and next because it is only with stress assigned that they can know whether the focus they intend is in the focus set. If it is not, stress shift must apply. This itself is a costly operation because it requires undoing steps in the derivation. However, the producer's task ends at this stage. For the producer's addressee, on the comprehension side, there is an additional mission. The comprehender is faced with a phonological input, which, by the innate definition of the focus set, allows

⁹Halbert et al. (1995) also found different results when they checked a verb with two obligatory internal arguments such as *give*, and the findings were not further tested and clarified.

several focus construals, that is, the derivation is semantically ambiguous. The comprehender needs to determine which of them the producer intended. Semantically ambiguous derivations always pose greater load on the hearer than on the producer (who always knows which meaning he or she intends). However, in this specific case in which the input derivation contains stress shift, the hearer needs in addition to apply reference set computation, namely, construct an alternative derivation additional to the input derivation. Based on this, the hearer would decide whether the derivation is still ambiguous (as in the case of stress shifted to the subject DP), or only one of the focus construals is actually permitted, in which case, this must be the focus that the producer intended. In actual language use, then, reference set computation of stress shift is involved only in comprehension. If it is this step of constructing and comparing an alternative derivation that children are unable to perform, one should expect problems in comprehension but not in production.

The central hypothesis under consideration is that children are unable to process reference set computation, and they resort to some bypassing strategy instead. So far, I have examined one strategy—guessing. But focus identification involves also semantic disambiguation. Even with neutral stress, a focus is selected in context out of a set of possible candidates. It is known that both children and adults may use semantic defaults to settle ambiguities when possible. Hence, the study of acquisition in this area enables us to examine the relations between guessing and default strategies. To differentiate the two, I start with an area in which semantic defaults do not play a role and indeed, the guessing pattern was found with stress shift.

5.2. Switch Reference Resolution

A contrastive stress context that has attracted much attention is anaphora resolution with contrastive pronouns. This was the second contrastive stress experiment of McDaniel and Maxfield (1992) and was studied before in Maratsos (1973) and Solan (1983). The problem is examined in greater depth in Zuckerman, Vasić, and Avrutin (2001) and Baauw, Ruigendijk, and Cuetos (2004). Here, I examine their findings. The structures under consideration are of the type illustrated in (25) and repeated in (39).

- (39) a. First Max touched Felix and then Lucie touched **him**.
 b. First Max touched Felix and then **he** touched **him**.

* * * *

- (40) a. Destressing: First Max [touched Felix_i] and then Lucie [touched him_i]

* *

- b. NSR: First Max [touched **Felix**_i] and then **Lucie** [touched him_i]

*

* * *

- c. Stress shift: First Max_i [touched **Felix**] and then *Lucie* [touched **him**_i]

Recall, first, that within this system, the stress on the pronouns of (39) is derived by stress shift. The derivation of (39a) is repeated in (40). Because the whole VP of the second conjunct is anaphoric, it is destressed by the anaphoric destressing rule independently of the neutral main stress rule as in (40a). The NSR then applies to assign main stress to the only available word, *Lucie* (40b). In the relevant context, the pronoun needs to be focused, which is obtained by applying the stress shift rule in (40c) (where *Lucie* retains secondary stress). In the derivation of (39b), stress shift applies to both pronouns.

The anaphora peculiarity of such structures is that the stressed pronoun appears to select the reverse value of its destressed counterpart. In (40b), with a destressed pronoun, it picks up *Felix* as reference. In (40c), the stressed pronoun picks up *Max*. For this reason, these structures are sometimes referred to as “switch reference” instances.

In Zuckerman et al.’s (2001) experiment, 28 children ages 4;3 to 6;2 (and 12 adults) were tested on similar sentences (e.g., *First Tinky-Winky hugged Po and then Dipsy hugged **him***). They were presented with sets of four pictures that were accompanied by a prerecorded sentence that described the events in the pictures. The first picture in each set corresponded to the first conjunct of the sentence. Of the three remaining pictures, only one correctly corresponded to the second conjunct of the sentence. The task of the children was to point to the correct picture out of these. There were 24 items for each condition. In 12, the contrastive pronoun was the object, and in the other 12, it was the subject. No difference was found between performance on subject and object pronouns. The percentage of correct answers was 51%. As Zuckerman et al. (2001) pointed out, the result is no different from chance. Zuckerman, Vasić, and Avrutin (2004) presented the individual data of the same experiment, which enables matching the results to the binomial model (of pure chance). Zuckerman et al. (2004) concluded that for most children, the individual results are consistent with that model. The pattern that was established is that most children in the experiment applied the guess strategy when faced with the experimental task.

To assess these results, I should get clearer about the type of computation involved in the resolution of anaphora with contrastive stress. Zuckerman et al. (2001) attributed their account to Akmajian and Jackendoff (1970) in which contrastive stress is used to switch reference. The basic intuition is that identifying coreference requires first establishing the anaphora construal of the same derivation without contrastive stress under parallelism. Parallelism determines that in certain contexts, the DP in a syntactic position parallel to that of the pronoun will be favored as the antecedent of that pronoun. In Zuckerman et al.’s (2001) examples like (41), the subject pronoun must corefer with the parallel subject in the preceding conjunct, and the object pronoun must corefer with the object. (Throughout, italic type marks coreference, and boldface marks main stress.) In both the following examples, the stress is obtained by the neutral main stress rule as was seen for (41b) in (40b).

- (41) a. First *Mary* touched Sue and then *she* touched **Peter**.
 b. First *Mary* touched *Sue* and then **Peter** touched *her*.

The details of the parallelism condition are not crucial to this discussion. (This condition is discussed in greater detail in Reinhart (in press-a, chap. 5.2.3.) What is crucial, under Zuckerman et al.'s (2001) analysis, is that resolving anaphora with stressed pronouns requires first construing the same derivation without stress:

In essence, the stress rule is a two step operation. The first step entails establishing what the reference for the pronoun would be in a similar unstressed structure. The second step then follows, where the reference that was established in the first step is cancelled or switched. (p. 782)

Although Zuckerman et al. (2001) did not explicitly state this, their account means that the resolution of stressed pronouns requires reference set computation. To determine the reference of the stressed pronoun in (42a), one needs to construct another derivation with neutral stress, namely, (41b) repeated in (42b):

- (42) a. First *Mary* touched Sue and then Peter touched *her*.
 b. First *Mary* touched *Sue* and then **Peter** touched *her*.

- (43) First *Mary* touched *Sue* and then *she* touched *her*.

One needs then to compute the pronoun resolution in (42b) according to the parallelism condition. Based on the result of this step, *her* in (42a) cannot be assigned the reference parallelism determines for it in (42b), namely, *her* cannot be *Sue*. So the option left for the pronoun in this context is *Mary*. If there are two stressed pronouns, as in (43), then the procedure applies to each of them. The result is that the object pronoun refers to the subject of the first conjunct, and the subject pronoun refers to its object.

If so, then children's guess pattern in tasks such as (42a) follows directly from the hypothesis that children cannot process the computation required in reference set comparisons. The processing load required by the switch reference rule is at least equal to that posed by Rule I, and accordingly, children's performance is identical—the majority just offer a guess.¹⁰

¹⁰Zuckerman et al. (2001) offered a different explanation for children's difficulties in both the switch reference cases and Condition B. They argued that children at first misclassify pronouns as anaphors, namely, expressions that can be bound locally and cannot carry stress. Therefore, an encounter with a stressed pronoun "creates confusion that leads either to chance performance or to ignoring the stress altogether and performing as if these sentences include an unstressed pronoun" (Zuckerman et al. (2001, 791)). It is not obvious to me why, on this account, children do not have a more general problem with anaphora resolution with pronouns given that pronouns occur all over the place in positions that disallow anaphors.

Although this is sufficient to explain children's performance on switch reference tasks, the account rests on discovering another area of reference set computation independent of that required for focus selection with stress shift. From the perspective of optimal design, discovering a new linguistic rule that requires this computation is not good news. Rather, it makes one wonder how optimally designed language could be if such rules can be so easily discovered. Let me therefore examine briefly the option that the processing cost here is reducible to the problem of focus computation with stress shift for which we know already that reference set computation is required.

Suppose stress shift has a single function of selecting a focus not in the focus set of the derivation with neutral stress. So when stress shifts to a pronoun, this is to mark it as a focus. The interpretation of the focus, regardless of whether it is obtained by neutral stress or by stress shift, always rests on establishing some contrast set or a set of alternatives (leaving maximal IP foci aside for this discussion). In case the focus is a single referential NP, this simply means that the NP must be disjoint in reference from its contrast set. This is, in essence, the intuition underlying Williams's (1997) disanaphora law, which states that the strong (stressed) element is disanaphoric. In standard contexts, the contrast set need not be associated with any specific position in the previous clause or even explicitly mentioned. For example, in (44), both *she* and *Max* are foci. The contrast set for *she* is identified contextually as *everyone*, so the pronoun must be interpreted as not in this set for which Lucie is the obvious candidate. (The interpretative effect is that *everyone* is construed as *everyone but Lucie*.) However, the contrast set for *Max* is not explicitly mentioned. Rather, it is construed as the set of better husbands Lucie could find, from which Max must be excluded:

(44) Everyone thinks that Lucie could marry more wisely. But **she** loves **Max**.

(45) Everyone thinks that Lucie could find a better partner than Max.

a. But she **loves** him.

b. But **she** loves **him**.

In standard contexts, it is often equally possible to use either a destressed pronoun or a stressed one (focus) with interpretative nuances. For example, in (45a), only neutral stress applied. The focus is construed as the whole IP, and the relevant contrast is between everyone's beliefs and Lucie's love for Max. In (45b), both pronouns are marked as foci through stress shift, and the contrast is construed along the lines of (44).

However, in the specific contexts of switch reference in which the two clauses are identical in all but one element, the contrast set is structurally fixed. This appears to be another aspect of the parallelism requirement. In the spirit of Williams (1997), parallelism can be viewed as a structural condition that matches object with object and subject with subject. Thus, it matches anaphoric (destressed) ele-

ments with their structural counterpart in the antecedent clause, but I argue that it does precisely the same with the foci (stressed) elements. The switch reference or disanaphora outcome of this matching need not be postulated as a specific rule or generalization. In both cases of (42), repeated in (46), the object pronoun of the second clause must be matched with the object of the first and the subject with the subject by the parallelism requirement.

- (46) a. First *Mary* touched Sue and then Peter touched **her**.
 b. First Mary touched *Sue* and then **Peter** touched *her*.

The difference is that in (46b), the pronoun is destressed; hence, it must be anaphoric to its counterpart (by the general destressing convention discussed briefly in section 4.3). In (46a), it is the focus; hence, its matching with the object *Sue* means that this object NP is the contrast set for the focus pronoun, which entails that it is disjoint in reference from this set. In (47), with both pronouns stressed, the object is again matched with the object and the subject with the subject. Because the pronouns are stressed, hence, necessarily foci, the matching fixes their contrast sets. *Her* contrasts with (hence is disjoint in reference from) *Sue* and *she* with *Mary*.

- (47) First *Mary* touched *Sue* and then **she** touched **her**.

As pointed out in Zuckerman et al. (2001), the fact that *she* ends up anaphoric to *Sue* and *her* to *Mary* is not part of the switch reference process; it is the outcome of these being the only two available referents given that the other anaphoric construal is ruled out. Zuckerman et al. (2001) pointed out that this is further confirmed with examples like (48).

- (48) John introduced Bill to Harry and then Mary introduced **him** to Ken.

The theme pronoun must be matched by parallelism with the theme *Bill*. Because the pronoun is the focus, the matching means contrast, so *him* cannot corefer with *Bill*. However, the derivation still allows the interpretation of the pronoun as either John or Harry. The same is witnessed even more sharply in (49), from K. Szendrői (personal communication (2004)):

- (49) Max hit the man near Peter and Felix hit the man near **him**.

The only construal excluded by parallelism matching is that *him* refers to Peter. But in the situation of men fighting described here, almost all other construals of the pronoun are allowed (Max, the man near Peter, or Felix himself).

In this view, then, there is no need to assume that the neutral stress derivation has to be constructed to determine the reference of a stressed pronoun. Rather, it is computed directly on the stress shift derivation, and it follows from parallelism just like in the case of the destressed pronoun. So the phenomenon of switch reference itself does not require any reference set computation, or more generally, there is no special rule of anaphora resolution that needs to be assumed for these cases.¹¹ The outcomes follow from the standard interpretation of focus combined with parallelism. Why should children, then, fail to process switch reference derivations?

The resolution of anaphora in these contexts rests crucially on focus identification, and because the foci are anaphoric pronouns, they can only be derived by stress shift. My hypothesis has been that the identification of foci obtained by stress shift always requires reference set computation, which surpasses children's processing ability. To determine the reference of the stressed pronoun, the child first has to determine if it is the focus or not. By my definition of the focus set, if the subject pronoun is stressed, the focus set also includes the IP, as in (50a). (For simplicity, I ignore here the focus status of *Max*.)

- (50) a. [First Mary touched Sue and then] **she** touched Max.
 Focus set of second conjunct: {*she, she touched Max*}
- b. [First Mary touched Sue and then] she touched **Max**.
 Focus set of second conjunct: {*Max, she touched Max*}

To identify the stressed pronoun as a focus, the child needs to exclude the option that the IP is the actual focus, an option illustrated for an unstressed pronoun in (45a). It is this stage of identifying the actual focus intended that requires the construction of the comparison derivation with neutral stress in (50b) thus forming the reference set (50). Given that the IP focus is also available in the neutral stress (50b), it is excluded as a possible focus of (50a), so the stressed pronoun must be selected as the focus. On the hypothesis under consideration, then, it is in this pre-

¹¹There is one area in which this analysis differs empirically from that proposed in Zuckerman et al. (2001). They argued that the contrast in (i) shows that the neutral stress derivation must be the basis for the computation of contrastive pronouns:

- (i) a. John hit Bill and then Mrs. Smith **punished** him.
 b. John hit Bill and then Mrs. Smith punished **him**.

On their intuition, in (ia) parallelism is cancelled owing to the special context, so *him* must refer to *John*. Consequently, the switch reference rule must reverse this construal, so in (ib), *him* cannot refer to *John*. If so, the switch reference rule must consider what the anaphora construal of the neutral derivation is. Here we differ on the intuition. For me, either of the derivations allows both anaphora construals. I argued (Reinhart (in press-a, chap. 5.2.3)) that such sentences do not fall at all under the parallelism constraint.

liminary stage of focus identification that the children get stuck. This means that they do not even get to the stage of processing parallelism and resort to a guess strategy first.¹²

Nevertheless, faced with a processing failure, the strategy children resort to does not always have to be guessing. The task in the switch reference cases does not involve semantic disambiguation, which is typical of focus selection. All that is required is to identify the focus by forming a reference set as a prerequisite for determining the value of the pronoun. Hence, it is not easy to imagine a strategy, other than guessing, that would enable bypassing the required computation. I now turn to areas in which the selection of the focus has semantic (truth conditional) impacts. Focus identification in the scope of *only* is a clear instance. As I show, in this area, semantic defaults may play a role.

5.3. Focus Identification in the Scope of Only

Let me first review the problem of focus projection in the scope of *only* that was discussed in section 4.4. The standard assumption is that the potential scope of *only* is just its c-command domain where it selects the focus as its scope. In (51), stress is assigned by the main (nuclear) stress rule to *builders*. In this case, the scope of *only* can be either (the narrow focus) *builders* or the whole VP that contains it. Suppose our store sells equipment only to builders, but at the same time we also buy used equipment from builders and others. In this situation, (51a), with the narrow focus, is true, and (51b), with the VP focus, is false.

- (51) a. We only sell equipment [_F to **builders**] (not to the general public).
 b. We only [_F sell equipment to **builders**] (We do not buy anything from anybody).
- (52) a. We only sell [_F **equipment**] to builders (not health insurance).
 b. #We only [_F sell **equipment** to builders] (We do not buy anything from anybody).

In (52), main stress shift applied. The sentence can only be used to exclude the option that we sell anything but equipment to builders but not to exclude anything else as witnessed by the inappropriateness of (52b). This means that the only ele-

¹²It has been found that children have difficulties also with the parallelism condition that, under the analysis presented here, applies once the focus has been identified. However, their performance in this area is not at chance level. In Thornton and Wexler's (1999) experiments, children accepted violations of parallelism 21% of the time. Zuckerman et al. found a much higher failure rate (45%), but they suggested that this may have to do with the experimental design, such as the fact that stressed and unstressed sentences were tested in the same session. Thus, there is no evidence that computing parallelism uniformly surpasses children's processing ability.

ment in the scope of *only* is the narrow focus, the argument bearing the new stress, as in (52a), but not the whole VP.

Computing this type of reasoning requires a construction of a reference set, which consists of <d,i> pairs of a derivation and interpretation. In this case, the relevant interpretation is a selection of a focus out of the focus set. (52b) is ruled out because the focus it selects—the VP is in the focus set of (51b)—is a derivation with no application of main stress shift.

This means that this operation entails a computational complexity whether the final outcome is in or out. On the other hand, there is no reason to assume any reference set computation in derivations involving no stress shift (as would be assumed in some optimality approaches).

Halbert et al. (1995) checked the interaction of *only* with what they called “emphatic stress” as in their (53). Under the analysis in section 4, the main stress rule assigns stress to the dative object *Miss Piggy*, and in (53), stress shift has applied.

(53) Daisy only gave a **cherry** to Miss Piggy.

Halbert et al. tested 33 children ages 3;6 to 6;6. The experiment was a truth value judgment task. One experimenter told a story. Another experimenter, playing a puppet, then played a tape with a sentence about the story (prerecorded to guarantee the correct stress), with the pretext that the puppet has a sore throat, and asked the children if the puppet was right or wrong. For (53), the story was about Daisy Duck, who had a famous restaurant with hot dogs and cherries. Miss Piggy, who spent her day in the gym, came in very hungry. However, Daisy explained that it is not good to eat much after the gym and therefore offered her a cherry, which she accepted. Then a spaceman arrived who was not familiar with Earth food, and Daisy gave him a hot dog.

In this context, (53) is true under the narrow construal of focus (the only thing Daisy gave Miss Piggy was a cherry). However, it is false on the wide construal, with the VP/IP as focus, namely, that the only thing that happened was that Daisy gave a cherry to Miss Piggy. Uttered in this context, no adult will have a problem judging (53) as true (as witnessed also by the judgments of the adult control group in the experiment). This judgment requires first identifying the focus of the sentence (in order to compute *only*). Because stress shift is involved, a reference set needs to be consulted to check whether this is the only way to obtain *cherry* as focus, and this is indeed the case. (The example is precisely analogous to (52) where the computation was discussed.) The results of the experiments were that only 46% of the children judged it true, that is, the results fall within the 50% range. Halbert et al. proceeded to show that children’s problems with such tasks cannot be attributed to a general stress deficiency. (Their experiment on this was described in section 5.1.) So the problem must be related to stress shift. For the processing account assumed here, this is the expected result. The computation required surpasses the processing ability of children, so they resort to a strategy.

Halbert et al.'s (1995) study was conducted independently of the question of whether reference set computation may entail processing difficulties for children. However, three subsequent studies (Gualmini, Maciukaite, and Crain (2003) and Gennari, Gualmini, Crain, Meroni, and Maciukaite (2001) on English; Szendrői (2004) on Dutch) replicated their experiment with this explicit question in mind. These studies compared children's performance on sentences with stress shift, as in (53), repeated in (55), to their performance on sentences with neutral stress such as (54). In the story context, (54) is judged by adults to be false and (55) true. (The experimental design is as described for Halbert et al.'s study previously.)

(54) Daisy only gave a cherry to **Miss Piggy**. (Adult answer in context: no)

(55) Daisy only gave a **cherry** to Miss Piggy. (Adult answer in context: yes)

In terms of the global statistics of the group's performance, these studies found similar results. In Gualmini et al. (2003), children gave an adult-like (no) reply 87% of the time for (54) and an adult-like reply (yes) only 35% for (55). In Gennari et al. (2001), it was 97.5% for (54) and 36.5% for (55). In Szendrői's (2004) experiment on Dutch-speaking children, it was 84.8% for (54) and 52.2% for (55).

It seems pretty well established that, as in the case of Condition B, children show an acquisition delay of foci obtained by stress shift and that here as well their (group) performance is in roughly the range of 50%. It seems safe to conclude based on this data that (at least a large group of) children do not carry out the processing required for the reference set computation of stress shift outputs and resort instead to some strategy. A different question is whether the strategy is to guess as in the case of Condition B and the switch reference cases. The question arises because, in the case of focus interpretation, there are also default strategies available.

Gualmini et al. (2003) argued that the strategy cannot be to guess based on analysis of the reasons given by the children who replied "no" (the non-adult-like response). The story and test sentences I use to illustrate their findings are those of Szendrői (2004) to which I turn next. However, the conditions in the two experiments were identical. The participants of the story are Tigger, Piglet, and Winnie the Pooh. They are playing in a garden where there is a lot of old furniture around. Tigger wants to show how strong he is, and he throws furniture over to the others. The relevant propositions that are true in the story context are given in (56). The target sentence with a stress shift focus is (57). The neutral stress sentence is (58).

- (56) a. Tigger threw a chair over to Winnie.
 b. Tigger threw a table over to Winnie.
 c. Tigger threw a chair over to Piglet.

(57) Tigger only threw a **chair** to Piglet.

(58) Tigger only threw a chair to **Piglet**.

Example (57) with *chair* as the focus is true in this context (the only thing that Tigger threw to Piglet is a chair). But (58) is false whether it is construed with *Piglet* or the VP as focus because Tigger also threw a chair to Winnie. The children who answered “no” to (57) were asked what they think really happened. Gualmini et al. reported that most of them answered that the puppet was wrong because Tigger also threw a chair to Winnie. This is a very surprising result (that would not be found with adults). It appears that children use the falsifying condition of (58) to falsify (57). Gualmini et al. took this to mean that children interpret *Piglet* as the focus of (57). This suggests, Gualmini et al. argued, that children apply a fixed default that takes the indirect object as focus, or, in the similar conclusion reached by Gennari et al. (2001), children do not interpret the stressed NP as the focus but rather the last NP (the NP that would bear the neutral stress of the sentence). If this is the correct interpretation, then children have a bigger deficiency than I assumed so far: They are unable to identify the focus set associated with the derivation. Recall that under the analysis in section 4, the focus set is computed directly on stress and structure, and it consists of all the constituents that contain the main stress regardless of whether it was assigned by the neutral main stress rule or by stress shift. The focus sets of (57) and (58) are given next:

(59) a. Tigger only threw a **chair** to Piglet.
 b. Focus set: {*a chair; threw a chair to Piglet.*}

(60) a. Tigger only threw a chair to **Piglet**.
 b. Focus set: {*Piglet; threw a chair to Piglet.*}

The VP (*threw a chair to Piglet*) is in the focus set of both derivations, but they differ regarding the other member in the set. In (60), either member of the focus set, in principle, can be selected in context. But in (59), a further consideration applies. Because the derivation involves stress shift, members of the focus set that would be found also in the derivation without stress shift are discarded. The upshot is that, for adults, the chair is the only possible actual focus. In any case, *Piglet* is not a member of the focus set of (59). Nevertheless it appears to be selected as the focus by the relevant children. If this is the case, then Gualmini et al. (2003) are correct in pointing out that “this invites the conclusion that children do not make use of prosodic prominence to determine the associate of the focus operator only” (p. 96).

However, as pointed out in Szendrői (2004), this is not the only possible interpretation of the findings. Before I turn to Szendrői’s alternative interpretation, let us pay more attention to the semantic conditions underlying the experiment. The

children's task in explaining why the puppet is wrong is to provide the falsifying condition—the proposition that makes the sentence false under a given focus construal. To analyze the children's responses, we need first to get clear about the falsifying conditions for each member of the focus set. The list of true propositions in the context was given in (56) and repeated here. Recall that the stress shift (59) has two members in its focus set. The interpretation of the sentence for each of these members is given in (61) and (62).

- (56) a. Tigger threw a chair over to Winnie.
 b. Tigger threw a table over to Winnie.
 c. Tigger threw a chair over to Piglet.

(61) Narrow (DP) focus selection for (59):
 The only thing that Tigger threw to Piglet is a chair.

(62) Wide (VP) focus selection for (59):
 The only thing that Tigger did is throw a chair to Piglet.

(63) Narrow (dative DP) focus selection for (60):
 The only one that Tigger threw a chair to is Piglet.

(64) Wide (VP) focus selection for (60):
 The only thing that Tigger did is throw a chair to Piglet.

Example (61) would be falsified if Tigger had also thrown to Piglet something other than a chair. There is no proposition corresponding to this in the context of (56), hence (61) is true. For (62), the falsifying conditions are (56a) and (56b). It is of course enough that one of them holds to falsify this focus construal, but the experimental context happens to have both. If a child selects the VP focus construal of (59), the child would provide one of these propositions as an answer. Gualmini et al. (2003) reported only answers based on (56a), but Szendrői (2004) found also responses containing (56b). The falsifying conditions are, of course, identical if the VP is selected as focus in (60), as in (64). The source of the puzzle that led Gualmini et al. to their interpretation of the findings is that (56a) is also the falsifying condition for the narrow focus construal of the neutral stress derivation in (63). So they took replies with (56a) to indicate that the latter is the focus construal of the relevant children. Szendrői (2004) argued that, in fact, the non-adult-like responses reflect a choice of the VP interpretation. She showed that there is, therefore, no reason to assume that children cannot identify the focus set. Rather, what their responses indicate is that about half of them selected the VP focus out of this set, which would be ruled out by the required reference set computation.

Szendrői (2004) proceeded to argue that there are in fact two distinct hypotheses on which the present analysis rests, both of which can be tested by acquisition performance. The first is that reference set computation is indeed involved in focus resolution with stress shift, and the second is that children are unable to execute this computation. The last finding, that children access the seemingly irrelevant VP focus, is in and of itself strong evidence for the first assumption. As mentioned in section 5.1, the theoretical problem with stress shift foci is similar to that raised by Rule I. In both, the contextual evidence for the need to assume a complex computation is small. In the case of focus, the evidence is the cases where the stress shifted, but the focus still projects (e.g., in subject DPs). Other than these, it appears easy to imagine different approaches that would not require any such computation for the relevant problem of focus projection. Statistically, in an overwhelming majority of cases, such formulation would capture the facts. However, under such views, it would be impossible to explain why children access the VP focus at all in such derivations. (A similar point was made by Gennari et al. (2001), who interpreted what they viewed as the selection of the indirect object focus as an indication that children access a derivation with neutral stress.) The fact that many children get stuck in this first stage of executing the required computation and select the VP focus is the most direct evidence that this focus construal is active in the computation of focus with stress shift.

Accessing all members of the focus set is only the first stage in determining the actual focus of sentences with stress shift, however. The next step is computing whether any of these members could also be obtained without stress shift, namely, the reference set computation. The second hypothesis under examination is that children are unable to execute this computation owing to limitations of their working memory. We saw already that children indeed do not complete this computation the same way that adults do, and I am concerned here with the question of what the strategy is that they use instead. Based on the group statistics examined so far, the strategy could be still guessing. Children have to decide whether the focus is the wide VP or the narrow DP. About half the time they select the one and thus give what appears to be an adult answer; in the other half of the time, they select the other. However, in an individual analysis of children's responses, Szendrői (2004) showed that uniform performance of individuals across trials is substantially larger than could be predicted by a binomial model. Of the 23 children in Szendrői's (2004) experiment, 7 consistently gave the "yes" answer on the stress shift condition such as (59), which means they consistently selected the narrow focus (*a chair*). Nine children consistently rejected the sentence, which means that they selected the wide VP focus. The other 7 children had less consistent responses. It seems that at least two subgroups of the children have fixed a default selection of either the narrow or the wide member of the focus set.

In the case of foci in the scope of *only*, it has been established that both adults and children apply default strategies also with neutral stress where no reference set computation is involved. So, as Szendrői (2004) pointed out, it is possible that

they are just applying their independent strategies also in the case of stress shift. The next question then is whether, indeed, children's difficulties with stress shift indicate a specific inability to compute the reference set.

5.4. Useful and Arbitrary Defaults

Focus resolution provides an opportunity to examine the relations of reference set computation with the resolution of semantic ambiguity. Independent of stress shift, focus resolution (with neutral stress) is a procedure of selecting the relevant focus out of the focus set associated with the derivation. This, then, is an instance of semantic ambiguity. Unlike the case of reference set computation, the competing alternatives are directly locally available in the derivation and need not be constructed or retrieved. Nevertheless, the resolution of semantic ambiguity requires comparing two (or more) semantic representations relative to context. This alone already poses some load on working memory (e.g., Altmann and Steedman (1988); Crain and Steedman (1985)). Returning to the hypothesis of optimal design, semantic ambiguities can be viewed as a sort of imperfection of the CS. A system that would fit perfectly to the limited hardware of the human processor would not require holding two representations in working memory. Indeed, although it is pretty well established that adults have the full ability to carry out these contextual selections, this is also the area in which they develop default strategies enabling them to bypass the comparison and selection procedure when possible. Crain, Ni, and Conway (1994) pointed out that these defaults can be experimentally witnessed when sentences are given with no context, as in some processing experiments, or when the context is consistent with more than one interpretation. It is not established that the required computation exceeds children's processing ability, but, as we shall see, they do develop defaults as well.

Semantic defaults are not computed individually for each derivation but rather provide a ready-made way to make the selection without an actual contextual computation. Nevertheless, what distinguishes a default from mere guess is that there are underlying principles that determine which defaults may be set. Crain and Hamburger (1992) argued that in resolving ambiguity, the semantic parser is motivated by the need to minimize cognitive effort or overload on the limited working memory capacity. The generalization behind the principle Crain and Hamburger proposed is that the semantic parser tries to reduce the risk of making commitments that will need to be changed later. As Crain et al. (1994) put it, the semantic parser is guided by a "minimal commitment" preference.

Crain et al. (1994) exemplified this principle with instances of semantic entailment, namely, one of the members of the set of possible interpretations entails the other. Focus construal with *only*, in neutral stress sentences, is one such instance. Let us examine this with the sentence that we have been following, repeated in (65) to (67):

- (65) a. Tigger only threw a chair to **Piglet**.
 b. Focus set: {*Piglet*; *threw a chair to Piglet*.}
- (66) Wide (VP) focus selection for (65):
 The only thing that Tigger did is throw a chair to Piglet.
- (67) Narrow (dative DP) focus selection for (65):
 The only one that Tigger threw a chair to is Piglet.

The wide focus construal with *only* always entails the narrow focus construal. Specifically, (66) entails (67). This means that the situations at which the wide construal (66) is true are a subset of the situations at which the narrow construal (67) is true. (Example (66) rules out many more situations than (67) does; for example, if (66) is true, it cannot be true also that Tigger threw a table to Winnie as in the story context. Yet (67) remains true in this situation.) Another way to describe the distinction is that (66) is more informative than (67), namely, accepting it is undertaking a bigger commitment. The principle of minimal commitment, or lowering the risks of the semantic parser, entails that adults should prefer in such contexts the less committal narrow scope construal (67). An intuitive way to motivate this preference is that when hearers select this option of interpreting a speaker's utterance, they maximize their chances to be in accord with the speaker's intentions. Crain et al. (1994) conducted several experiments that show that such preference is indeed strongly operative with adults.

Although Crain et al. do not discuss this in this article, the same entailment relations hold in the cases of quantifier scope:

- (68) Everybody in this room speaks two languages.
- (69) a. Wide existential scope
 There are two languages such that everybody in this room speaks these same two languages.
 b. Narrow existential scope:
 For every person *x* in this room, there are two languages that *x* speaks (possibly different languages for different persons).

I have argued elsewhere (e.g., Reinhart, in press-a) that QR is not involved in any of the scope construals of (68) illustrated in (69). Rather, the scope of the existential quantifier is determined by where the choice function that corresponds to the indefinite DP is existentially closed. This means that no reference set computation is required for any of these construals. (Such computation is required only when an illicit covert operation needs to apply.) However, examined from the perspective of the semantic parser or of efficiency of communication, the situation is precisely the same as in (65). Sentence (69a) entails (69b); that is, (69b) is consistent

with more situations in the world than (69a) is, or (69a) is true in a subset of the situations where (69b) is. Although both scope construals are equally permitted, the Principle of Minimal Commitment entails that in the absence of context, as in this example, the default preferred interpretation would be (69b), which closes fewer options than (69a) and thus maximizes the chances of the language user to be correct.

Returning to the focus selection default, Crain et al. (1994) discovered that children too use default strategies in the disambiguation of *only* sentences, but for many children, it is the opposite default of maximum commitment. Rather than the adult choice of the narrow focus interpretation (67), these children select the wide VP focus (66), which excludes more situations. The account Crain et al. (1994) gave for this default is that it is the most efficient strategy to acquire a potential semantic interpretation. The more children exclude, the more they will get the chance to add, based on positive evidence, whereas subtracting from the initial hypothesis may require negative evidence.

Crain et al. (1994) assumed that this strategy reflects an obligatory requirement for language acquisition whenever one interpretation is a subset of the other. This is their *Semantic Subset Principle*, which determines that to avoid subset learnability problems, learners initially hypothesize an interpretation that makes a sentence true in the smallest set of situations. But it is not clear from their experimental study whether indeed all children follow this default. A VP default was identified for 3 out of 6 children in one of their experiments and for 8 out of 12 in the other. It may turn out that some of them apply the narrow default, just like adults. It is also not established whether a majority of children apply the same Principle of Maximum Commitment in other instances such as the quantifier scope example (68). Furthermore, I am not sure that the semantic subset cases pose a learnability problem in the classical sense. This would be the case if the range of interpretative options was to be acquired just based on experience. However, the hypothesis here is that semantic knowledge, just like syntax, is innate. So the child knows innately that a derivation like *Tigger only threw a chair to Piglet* is ambiguous. Specifically, the innate definition of the focus set (combined with knowledge of the c-command scope of *only*) identifies two foci in this derivation. Children also know the truth conditions associated with each of the focus selections. Possibly, what motivates the VP-default children is the drive to discover actual uses of the two interpretations they know are associated with the sentence. If this is the case, then consistently selecting the reading compatible with least situations enables them indeed to identify the contexts associated with the broader readings. Given that it is not yet established that all children always abide by the semantic subset principle or maximum commitment, it may even be the case that the default children settle on reflects other aspects of their personality such as curiosity (maximally exclusive reading) versus the ambition to succeed (minimally exclusive reading). Decision on these matters is not crucial for this discussion.

In any case, the characteristic property of defaults used for semantic disambiguation is that they are useful relative to a goal. Unlike the option of bypassing contextual computation by guess, the selected default has greater chances of meeting the goals, whether these are the standard ones of minimum mistakes and revisions or some special motivation favored for the acquisition stage.

When it comes to reference set computation, however, the semantic disambiguation defaults turn out to be useless. The reason is that what needs to be computed is not just the appropriateness of a given interpretation to context but also whether this interpretation is not independently ruled out. Let us view the task of a child encountering a stress shift sentence like *Tigger only threw a chair to Piglet*. We saw already that the child identifies correctly the focus set and thus the two semantic representations of the sentence, that is, the different sets of truth conditions associated with each. In the given experimental setting, the sentence is false under the wide focus VP reading and true under the narrow DP reading. The child has to determine which of these is intended in order to answer whether the sentence is true or false. I assume that the child can identify stress, that is, knows the main stress rule. Hence, the child realizes that the derivation involves stress shift and thus violates the basic stress rule. I assume next that, as with Rule I, the child knows that when such violation occurs, reference set checking should be executed to determine that the contextual needs could not be satisfied without this violation. This entails constructing an alternative derivation with neutral stress and filtering out any focus construal that is available also at that derivation. However, the child cannot carry out this computation. In the case of coreference or switch reference resolution, the only option left was guessing. In the case of focus, there appears to be another option. Superficially, processing both neutral stress and stress shift derivations involves disambiguation, namely, selecting a member from the focus set, so the child may decide to use his general disambiguation default for *only* foci and thus avoid the required reference set computation.

Suppose, then, a child has a wide VP default. He or she would then apply the same default in the reference set task. However, this default is irrelevant for the task because the VP focus is not available with stress shift. It is impossible to learn anything about the contexts favoring wide VP construals by systematically using this construal when a different module of the language system rules it out independently of the subset problem. Opting for an irrelevant default can thus be viewed as another form of an arbitrary choice or at least as a systematic way to bypass the required computation. However, the arbitrary selection that children with the VP default make is set in advance for all tasks, so their performance is consistent across trials—they will always be wrong in comparison to the adult response.

If a child is set on the narrow focus default for semantic disambiguation, applying this same default to bypass the reference set computation turns out more useful, although again for irrelevant reasons. The narrow default always guides the user to avoid the wide focus option. Because in the relevant instances of stress

shift, it is the wide focus that is excluded by the required reference set computation, children who use this default will always end up with the correct (adult-like) response. Note, however, that this apparent success is nevertheless based on an irrelevant default. The only reason this default is useful is statistical probability. It is indeed the case that in most instances, the reading excluded by the semantic default is also excluded by the relevant reference set computation. So based on probability, one's chances to be correct when applying this default are high. In fact, because this irrelevant default is so successful statistically in contexts of stress shift in the scope of *only*, it is possible that adults as well adopt it to bypass the required computation.

Recall that in Szendrői's (2004) experiment, there were three groups identified: one that systematically fails the task across trials, one that systematically succeeds, and an undecided group. Although further experiments are needed to confirm this, it is possible that the first is the group that borrows the wide VP default for the task; the second is the group that borrows the narrow default, and the third are children who just apply guessing. For tasks of focus identification in the scope of *only*, it turns out to be impossible to check further whether the third of the children in Szendrői's (2004) experiment who systematically responded correctly across trials were successful in carrying out the required reference set computation or if they just borrowed the narrow focus default from the area of semantic disambiguation. For these reasons, it may be more useful to study children's identification of focus with stress shift in different semantic contexts and not in the scope of *only*.

To conclude the discussion of the acquisition of stress shift, we saw that (at least the majority of) children do not carry out the required reference set computation, but they operate by some bypassing strategy. The strategy can be guessing, as in the case of switch reference, or it can be an irrelevant semantic default available for other tasks of semantic disambiguation. Either way, because they do not execute the required computation, whether their answers in the experimental tasks correlate with those of adults or not is purely accidental.

ACKNOWLEDGMENTS

This article is based on lectures given in 1999. Its first draft (Reinhart (1999)) was submitted to *Language Acquisition* in 2000. The theoretical introduction (section 2) appears here as in that draft. I thank Eddy Ruys for extensive comments and discussion of this part. In the 4 years since, experimental evidence was found for what was at the time just a theoretical hypothesis, and in this article, I present that evidence also. The sections on the acquisition of focus and stress shift are largely based on joint work with Kriszta Szendrői, whose constant input and insight have shaped their development.

REFERENCES

- Akmajian, A. and R. Jackendoff (1970) "Coreferentiality and Stress," *Linguistic Inquiry* 1, 124–126.
- Altmann, G. T. M. and M. Steedman (1988) "Interaction with Context During Human Sentence Processing," *Cognition* 30, 191–238.
- Baauw, S., E. Ruigendijk, and F. Cuetos (2004) "The Interpretation of Contrastive Stress in Spanish-Speaking Children," in J. van Kampen and S. Baauw, eds., *Proceedings of GALA 2003, Volume 1*, LOT, Netherlands Graduate School of Linguistics, Utrecht, The Netherlands.
- Baddeley, A. D. (1986) *Working Memory*. Oxford University Press, Oxford, England.
- Chien, Y.-Ch. and K. Wexler (1990) "Children's Knowledge of Locality Conditions in Binding as Evidence of the Modularity of Syntax and Pragmatics," *Language Acquisition* 1, 225–295.
- Chierchia, G., S. Crain, M. T. Guasti, A. Gualmini, and L. Meroni (2001) "The Acquisition of Disjunction: Evidence for a Grammatical View of Scalar Implicatures," in A. H.-J. Do, L. Domínguez, and A. Johansen, eds., *The Proceedings of the 25th Annual Boston University Conference on Language Development, Volume 1*, Cascadilla Press, Somerville, Massachusetts.
- Chomsky, N. (1965) *Aspects of the Theory of Syntax*, MIT Press, Cambridge, Massachusetts.
- Chomsky, N. (1971) "Deep Structure, Surface Structure, and Semantic Interpretation," in D. Steinberg and L. Jakobovits, eds., *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology*, Cambridge University Press, Cambridge, England.
- Chomsky, N. (1976) "Conditions on Rules of Grammar," *Linguistic Analysis* 2, 303–351.
- Chomsky, N. (1993) "A Minimalist Program for Linguistic Theory," in K. Hale and S. J. Keyser, eds., *The View from Building 20*, MIT Press, Cambridge, Massachusetts.
- Chomsky, N. (1995) *The Minimalist Program*, MIT Press, Cambridge, Massachusetts.
- Chomsky, N. (2000) "Minimalist Inquiries: The Framework," in R. Martin, D. Michaels, and J. Uriagereka, eds., *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, MIT Press, Cambridge, Massachusetts.
- Cinque, G. (1993) "A Null Theory of Phrase and Compound Stress," *Linguistic Inquiry* 24, 239–298.
- Collins, C. (1997) *Local Economy*, MIT Press, Cambridge, Massachusetts.
- Crain, S. and H. Hamburger (1992) "Semantic Knowledge and NP Modification," in R. Levine, ed., *Formal Grammar: Theory and Interpretation, Volume 2*, University of British Columbia Press, Vancouver, British Columbia, Canada.
- Crain, S., W. Ni, and L. Conway (1994) "Learning, Parsing, and Modularity," in C. Clifton, L. Frazier, and K. Rayner, eds., *Perspectives on Sentence Processing*, Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey.
- Crain, S. and M. Steedman (1985) "On Not Being Led Up the Garden Path: The Use of Context by the Psychological Parser," in D. Dowty, L. Karttunen, and A. Zwicky, eds., *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives*, Cambridge University Press, Cambridge, England.
- Crain, S., R. Thornton, C. Boster, L. Conway, D. Lillo-Martin, and E. Woodams (1996) "Quantification Without Qualification," *Language Acquisition* 5, 83–153.
- Cutler, A. and D. Swinney (1987) "Prosody and the Development of Comprehension," *Journal of Child Language* 14, 145–167.
- Fodor, J. A., T. G. Bever, and M. F. Garrett (1974) *The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar*, McGraw-Hill, New York.
- Fox, D. (1995) "Economy and Scope," *Natural Language Semantics* 3, 283–341.
- Fox, D. (1998) "Locality in Variable-Binding," in P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky, eds., *Is the Best Good Enough? Optimality and Competition in Syntax*, MIT Press, Cambridge, Massachusetts.
- Fox, D. (2000) *Economy and Semantic Interpretation*, MIT Press, Cambridge, Massachusetts.
- Gathercole, S. and A. Baddeley (1993) *Working Memory and Language*, Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey.

- Gathercole, S. and G. Hitch (1993) "Developmental Changes in Short-Term Memory: A Revised Working Memory Perspective," in A. F. Collins, S. E. Gathercole, M. A. Conway, and P. E. Morris, eds., *Theories of Memory*, Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey.
- Gennari, S., A. Gualmini, S. Crain, L. Meroni, and S. Maciukaite (2001) "How Adults and Children Manage Stress in Ambiguous Contexts," *Proceedings of 1st Workshop on Cognitive Models of Semantic Processing*, University of Edinburgh, Edinburgh, Scotland.
- Golan, V. (1993) "Node Crossing Economy, Superiority, and D-linking," ms., Tel Aviv University, Tel Aviv, Israel.
- Grodzinsky, Y. and T. Reinhart (1993) "The Innateness of Binding and Coreference," *Linguistic Inquiry* 24, 69–102.
- Gualmini, A., S. Crain, L. Meroni, G. Chierchia, and M. T. Guasti (2001) "At the Semantics/Pragmatics Interface in Child Language," in *Proceedings of Semantics and Linguistic Theory 11*, Cornell University, Cornell, New York.
- Gualmini, A., S. Maciukaite, and S. Crain (2003) "Children's Insensitivity to Contrastive Stress in Sentences with ONLY," *Proceedings of the 25th Penn Linguistics Colloquium*, University of Pennsylvania, Philadelphia.
- Halbert, A., S. Crain, D. Shankweiler, and E. Woodams (1995) "Children's Interpretive Use of Emphatic Stress," paper presented at the 8th Annual CUNY Conference on Human Sentence Processing, Tucson, Arizona.
- Halle, M. and J.-R. Vergnaud (1987) *An Essay on Stress*, MIT Press, Cambridge, Massachusetts.
- Heim, I. (1998) "Anaphora and Semantic Interpretation: A Reinterpretation of Reinhart's Approach," in U. Sauerland and O. Percus, eds., *The Interpretative Tract, MIT Working Papers in Linguistics*, 25, Cambridge, Massachusetts.
- Hirschbühler, P. (1982) "VP Deletion and Across-the-Board Quantifier Scope," in *Proceedings of NELS 12*, GLSA, University of Massachusetts, Amherst.
- Hornby, P. A. and W. A. Hass (1970) "Use of Contrastive Stress by Preschool Children," *Journal of Speech and Hearing Research* 13, 387–394.
- Inhelder, B. and J. Piaget (1964) *The Early Growth of Logic in the Child*, Routledge and Kegan Paul, London.
- Keenan, E. L. and L. M. Faltz (1978) "Logical Types for Natural Language," in *UCLA Occasional Papers in Linguistics* 3, University of California, Los Angeles.
- Krämer, I. (2001) *Interpreting Indefinites*, Doctoral dissertation, Max-Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.
- Levinson, S. C. (1987) *Pragmatics*, Cambridge University Press, Cambridge, England.
- Lieberman, M. (1979) *The Intonational System of English*, Garland Press, New York.
- Lidz, J. and J. Musolino (2002) "Children's Command of Quantification," *Cognition* 84, 113–154.
- Maratsos, M. (1973) "The Effects of Stress on the Understanding of Pronominal Reference in Children," *Journal of Psycholinguistic Research* 2, 1–8.
- McDaniel, D. and T. Maxfield (1992) "Principle B and Contrastive Stress," *Language Acquisition* 2, 337–358.
- Musolino, J. and J. Lidz (2003) "The Scope of Isomorphism: Turning Adults into Children," *Language Acquisition* 11, 277–291.
- Neeleman, A. and T. Reinhart (1998) "Scrambling and the PF Interface," in W. Geuder and M. Butt, eds., *Projecting From the Lexicon CSLI*, Stanford University Press, Stanford, California.
- Philip, W. (1995) *Event Quantification in the Acquisition of Universal Quantification*, Doctoral dissertation, University of Massachusetts, Amherst.
- Pulleyblank, D. and W. J. Turkel (1998) "The Logical Problem of Language Acquisition in Optimality Theory," in P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky, eds., *Is the Best Good Enough? Optimality and Competition in Syntax*, MIT Press and MIT, Cambridge, Massachusetts.
- Reinhart, T. (1983) *Anaphora and Semantic Interpretation*, Croom Helm, London.

- Reinhart, T. (1994/1998) "Wh-in-situ in the framework of the minimalist program," *OTS Working Papers in Linguistics* 1994, TL-94-003. [Appeared with slight revisions in *Natural Language Semantics*, 1998, 6, 29-56]
- Reinhart, T. (1995) "Interface Strategies," *OTS Working Papers in Linguistics*, TL-95-002, University of Utrecht, Utrecht, The Netherlands.
- Reinhart, T. (1998) "Interface Economy: Focus and Markedness," in C. Wilder, H.-M. Gaertner, and M. Bierwisch, eds., *The Role of Economy Principles in Linguistic Theory*, Akademie Verlag, Berlin, Germany.
- Reinhart, T. (1999) "The Processing Cost of Reference Set Computation: Guess Patterns in Acquisition," *OTS Working Papers in Linguistics*, 99-001-CL/TL, Utrecht University, Utrecht, The Netherlands.
- Reinhart, T. (2000) "Strategies of Anaphora Resolution," in H. Bennis, M. Everaert, and E. Reuland, eds., *Interface Strategies*, Royal Netherlands Academy of Arts and Sciences, Amsterdam.
- Reinhart, T. (in press-a) *Interface Strategies: Reference Set Computation*, MIT Press, Cambridge, Massachusetts.
- Reinhart, T. (in press-b) "Processing or Pragmatics?—Explaining the Coreference Delay," in E. Gibson and N. Pearlmuter, eds., *The Processing and Acquisition of Reference*, MIT Press, Cambridge, Massachusetts.
- Reinhart, T. and K. Szendrői (2003) "Optimal Design in Language, an NWO Grant Proposal," available at <http://www.let.uu.nl/~tanya.reinhart/personal>
- Rizzi, L. (1990) *Relativized Minimality*, MIT Press, Cambridge, Massachusetts.
- Selkirk, E. (1984) *Phonology and Syntax*, MIT Press, Cambridge, Massachusetts.
- Smith, E. E. (1999) "Working Memory," in R. A. Wilson and F. C. Keil, eds., *The MIT Encyclopedia of the Cognitive Sciences*, MIT Press, Cambridge, Massachusetts.
- Solan, L. (1983) *Pronominal Reference: Child Language and the Theory of Grammar*, Reidel, Dordrecht, The Netherlands.
- Szendrői, K. (2001) *Focus and the Syntax-Phonology Interface*, Doctoral dissertation, University College London.
- Szendrői, K. (2004) "Acquisition Evidence for an Interface Theory of Focus," in J. van Kampen and S. Baauw, eds., *Proceedings of GALA 2003, Volume 2*, LOT, Netherlands Graduate School of Linguistics, Utrecht, The Netherlands.
- Tavakolian, S. (1974) "Contrastive Stress Pattern in Four Year Olds," ms., University of Massachusetts, Amherst.
- Tesar, B. (1998) "Error-Driven Learning in Optimality Theory Via the Efficient Computation of Optimal Forms," in P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky, eds., *Is the Best Good Enough? Optimality and Competition in Syntax*, MIT Press, Cambridge, Massachusetts.
- Thornton, R. and K. Wexler (1999) *Principle B, VP Ellipsis, and Interpretation in Child Grammar*, MIT Press, Cambridge, Massachusetts.
- Williams, E. (1997) "Blocking and Anaphora," *Linguistic Inquiry* 28, 577-628.
- Zubizarreta, M.-L. (1995) "Word Order, Prosody, and Focus," ms., University of Southern California, Los Angeles.
- Zubizarreta, M.-L. (1998) *Prosody, Focus, and Word Order*, MIT Press, Cambridge, Massachusetts.
- Zuckerman, S., N. Vasić, and S. Avrutin (2001) "The Syntax-Discourse Interface and the Interpretation of Pronominals by Dutch-Speaking Children," in *Proceedings of the 26th Annual Boston University Conference on Language Development, BUCLD 26*, Boston, Massachusetts.
- Zuckerman, S., N. Vasić, and S. Avrutin (2004) "Pronominal Reference in Child Language," ms., UiL OTS, Utrecht University, Utrecht, The Netherlands.