4

Attach Anyway*

Janet Dean Fodor & Atsu Inoue

CUNY Graduate School and University Center &
Kantogakuin University

Abstract

The diagnostic model of garden path recovery that we have advocated in previous work holds that no repair processes are intrinsically costly. Repair costs depend entirely on the difficulty of establishing what revisions to make. The diagnosis process does not require a special-purpose inference system as long as the parser abides by the *Attach Anyway* principle: when it encounters an input word that doesn't fit into the current structure, it attaches it in the least unacceptable way. The attachment creates a conflict internal to the phrase marker, which is then resolved in consultation with the grammar by a process we call *Adjust*. In this chapter we propose a principled constraint on the operations of *Adjust*: the *Grammatical Dependency Principle*. We show that this clarifies some previously noted phenomena such as the Thematic Overlay Effect and the differential difficulty of different types of steal operations. The examples we present show that neither raising repairs nor semantic revisions are difficult *per se*.

1 Background

A garden path occurs when the parser selects the wrong analysis at a point of ambiguity in an input sentence, and discovers later that subsequent words of the input do not fit into the structure it has been building. In previous work (Fodor & Inoue, 1994) we presented a model in which recovery from a garden path consists in repairing the incorrect sentence structure to make it compatible with the current input. A repair system contrasts with a reparsing system, which gives up the current structure, returns to some earlier point in the sentence, and

* We are grateful to Paul Gorrell and Patrick Sturt for their perceptive comments on an earlier draft.

tries parsing it in a different way. A repair system must have some way of deciding what structural alterations to make. We proposed that recovery costs are entirely attributable to the difficulty of this diagnostic process. If it can be defended, this is a very desirable assumption on grounds of parsimony, since there must be some such process in any repair system, and it seems clear that the ease of deducing what error has occurred and how it can be corrected is bound to differ from case to case. It would not be necessary, then, to assume greater costs of carrying out one kind of repair than another (e.g., raising a node versus lowering a node), or even a greater inherent cost for repairs than for any other parsing processes.

Central to the diagnostic process is the nature of the misfit between the current partial phrase marker (CPPM) and the input word that is incompatible with it. This is necessarily the starting point, the premise for deducing which prior parsing decision was the source of the current problem. If the relation between this mismatch and the prior error responsible for it is transparent, recovery is predicted to be easy; if the connection is obscure, recovery would be difficult or impossible. To go from this general idea to predictions about particular examples we must say more about what makes this relationship transparent or otherwise. In other words, we owe an account of the reasoning capabilities of the revision component of the parser. Different reasoning algorithms would exhibit different profiles of success or failure for different examples. More interestingly, turning this around the other way gives us our methodology: observed patterns of difficulty can be informative of the nature of the diagnostic mechanism.

On this basis, we argued in our earlier work (Fodor & Inoue, 1994) against the idea of a repair agent which stands back and studies the whole CPPM and tries to calculate what changes would allow the current input to be integrated with it. The data we considered suggested that the repair system has no insight into clever solutions and that it is unable to combine facts productively; rather, it is successful only where the necessary inference can be made in small steps quite mechanically. For this reason we proposed a repair system that does not reason abstractly about the problem, but just sets about altering the CPPM one step at a time, starting at the specific mismatch that presents itself at the point of breakdown. In the ideal case (but not in every case: not all garden paths are revisable) these successive small adjustments will lead the parser back to the site of the original error. The final revision will be at that point, and the parser will re-attach the item that was incorrectly attached on the first pass; all subsidiary adjustments consequent on that repair will have already been made en route to it.[1] This account of how diagnosis and cure are effected has the great advantage that it does not presuppose abstract reasoning capabilities or require postulation of special-purpose revision routines; it can in large part draw on simple practical mechanisms already in place for first-pass parsing. This is so, given one crucial assumption about how the CPPM and the input are made to confront each other by the parser at the very outset of the repair process. This assumption is stated in the parsing principle that we call *Attach Anyway*, discussed in section 2.

Attach Anyway is a close relation of *Attach*, the fundamental principle of first-pass parsing in a full-attachment (non-delay) parser:

(1) *Attach*: On receiving a word of the input sentence, connect it to the current partial phrase marker (CPPM) for the sentence in such a way that the resulting CPPM is syntactically well-formed though possibly incomplete at its right edge.[2]

This simply enjoins the parser to integrate each input item into the sentence structure as it is received. If there is more than one way of creating a new well-formed CPPM, we assume the parser adopts whichever one is computed most rapidly. This is the *First Analysis Constraint* of Frazier & Rayner (1988), and Frazier & Fodor (1978). We follow this earlier work also in assuming that the simplest possible attachment wins the race. The simplest attachment is one that obeys both *Minimal Attachment* (MA) and *Right Association* (RA) or *Late Closure* (LC). These principles are specific instantiations of the *Minimal Everything* principle, which embodies the basic assumption that the human sentence parsing mechanism is a least effort device. Borrowing from unpublished work by Fodor & Frazier (1983) we assume here that MA and RA are jointly satisfied by the shortest possible path of linking nodes between the rightmost terminal symbol $i-1$ and the input item i that is now being attached, where this path goes up from $i-1$ through existing nodes in the tree, and then down through newly created nodes to i.[3] Minimizing the up-path satisfies RA;

---

[1] As will become clear, this 'backtracking' to the source error is very different from the word by word backtracking of an ATN (Wanner & Maratsos, 1978). We note also that in addition to the adjustments leading back to the source error, there may be others that stem from those. There is an example in Fodor & Inoue (in press), but none in this chapter.

[2] Semantic and pragmatic acceptability might also be a condition on *Attach*, but this will not be discussed here. We also lack space to discuss repairs of errors signalled by pragmatic conflicts. See Fodor & Inoue (1994) for examples where pragmatic symptoms are ineffective as pointers to the site of the original error. See Ni *et al.* (submitted) for a suggestion as to how pragmatic anomaly can initiate appropriate repairs in some circumstances.

[3] In left-branching languages such as Japanese, the path above $i-1$ may have new nodes above old nodes. In the 1983 manuscript we considered the possibility (from Cooper & Paccia-Cooper,

minimizing the down-path satisfies MA. (We called this the *Minimal Connections* principle. See Phillips, 1995, for a new proposal along similar lines.) We lack space here to give further details of *Attach*. Fortunately, exactly how it is implemented is not germane to our concerns in this paper, as long as the mechanism is such that the preference principles (i.e., *Minimal Everything* and its subsidiaries) follow from it naturally without stipulation.

Attach Anyway is similar to *Attach* but applies in garden path constructions when the parse has broken down. The parser discovers that the current input is incompatible with the CPPM, and cannot be legitimately attached into it. *Attach Anyway* requires the parser to continue functioning as much like normal as it can in these circumstances. It is not allowed to quit, but must attach the problematic input into the structure despite the ill-formedness of the new CPPM it thereby produces. Then the principle *Adjust* directs the parser to set about correcting this ill-formedness. In the next section we give a more detailed presentation of both *Attach Anyway* and *Adjust*. They account for a range of garden paths. We present here some new examples not previously noted in the literature. We also propose a linguistically motivated locality constraint on the operations of *Adjust*, and show that this permits interesting refinements of previously noted phenomena such as the Thematic Overlay Effect, and the differential difficulty of different types of steal operations. As we proceed we will relate this system, in which repair is driven by *Attach Anyway*, to other repair models, noting both similarities and differences. Despite the number of competing proposals that have been made in recent years, it seems that we may be moving towards a consensus on at least some of the questions about why garden paths differ as they do in ease of recovery, and how this might follow from the way in which needed repairs are identified and carried out.

2 Attach Anyway and Adjust

We claim that there are just two principles of repair parsing: *Attach Anyway* and *Adjust*. *Attach Anyway* applies when the parse runs into a problem and it appears that there is no way for it to continue. *Attach Anyway* forces the parser to keep going.

1980) that new nodes are costlier than old ones. It could also be that clausal nodes (or verbal projections) on this path cost more than other nodes. This would account for the general preference for attachment within the lowest clause; see Igoa (1995) and discussion in Fodor & Inoue (in preparation a). Other aspects of the Minimal Connections model are not considered here (e.g., the assumption that the parser establishes the minimal connection between adjacent words before checking whether this is consistent with how word *i* is already configured in the CPPM).

(2) *Attach Anyway*: Having established that there is no legitimate attachment site in the CPPM for the current input word,[4] attach the input word into the CPPM wherever it least severely violates the grammar, and subject to the usual preference principles that govern *Attach*.[5]

An attachment made by *Attach Anyway* necessarily creates an ill-formed sub-tree within the CPPM. In a typical case, one of the new nodes introduced to link the input word *i* into the tree conflicts with the properties of some other node already in the tree: its mother or its sister, or, as we shall see, a node with which it enters into a long-distance dependency. This conflict within the CPPM is then resolved by the parser under guidance by *Adjust*.[6,7]

(3) *Adjust*: When a grammatical conflict has been created between two nodes or features X and Y in the CPPM, by either *Attach Anyway* or *Adjust*, eliminate the problem by altering minimally (i.e., no more than is necessary for conflict resolution) whichever of X and Y was less recently acted on, without regard for grammatical conflicts thereby created between that node and other elements in the CPPM.

---

[4] *Attach Anyway* does not always wait for complete breakdown of the first-pass parse, though for simplicity we will assume here that it does. If it did, that would be equivalent to a strict Revision as Last Resort principle (Fodor & Frazier, 1980). But in Fodor & Inoue (in preparation a) we give examples which show that RALR is not an absolute constraint: the wellformedness requirement and the attachment preference principles interact in determining the attachment made. (See also Stevenson, this volume.) This would follow from the First Analysis Constraint if illegal attachments sometimes, but not always, take longer to compute than legal ones, and if *Attach* and *Attach Anyway* are folded into a single principle *Attach (Anyway)*, as we argue they must be.

[5] Some evidence that *Attach Anyway* exhibits the same attachment preferences as first-pass *Attach* comes from examples presented by Sturt & Crocker (1996). For instance, in *I know the man who believes the countess killed himself* the first-pass analysis makes *the man who…* the direct object of *know*, and *the countess* the direct object of *believes*; then for the correct revision, the VP *killed himself* should steal the higher NP, but instead the parser opts to steal the lower one. We presume this is due to RA (LC), which strongly disfavors attachment above the current clause (cf. Igoa, 1995).

[6] Lewis (1993, this volume) proposes a similar process of attachment despite illformedness, followed by adjustment processes. In this respect our repair model is closer to Lewis' than to any other model, though it differs from Lewis' model with respect to which adjustments are permitted; see Fodor & Inoue (in press) for discussion. Also closely related is Stevenson's competitive attachment parser (this volume), which juggles the claims of a number of potential attachments, not all of which are compatible with the CPPM.

[7] There are some special cases in which *Attach Anyway* does not result in a specific conflict between elements in the tree (e.g., where the symptom is the absence of a needed constituent). These are discussed in Fodor & Inoue (in preparation b).

Some aspects of (3) should be noted. The requirement that the adjustment be minimal, though made explicit here, actually follows from the least effort principle *Minimal Everything* (see Frazier, 1990; Inoue & Fodor, 1995). It precludes, for example, detaching a node from its current position if it could have been reconciled with the context by changing just one of its feature values. Nothing in (3) requires that there be only one node that a new or newly adjusted node is incompatible with, but the Grammatical Dependency Principle discussed in section 3 below will normally single out a unique node for the adjustment process to work on. *Adjust* is defined recursively. One application of *Adjust* may restore the whole CPPM to well-formedness. Alternatively, it may cure one grammatical violation only by creating another one elsewhere. In that case, *Adjust* applies to the new violation, and to any violation which it thus creates, continuing until a fully grammatical equilibrium is achieved. A stable state is not always reached, however, because the necessary series of revisions may block at some point (for reasons outlined below) or possibly because time runs out before the necessary computations can be completed (see Konieczny *et al.*, 1994; Ferreira & Henderson, this volume).

To avoid unproductive repairs going around in small circles, *Adjust* includes the requirement that when a conflict has been created in the tree, the parser is to hold onto whatever it most recently did under the direction of *Attach Anyway* or *Adjust*, and alter the *other* end of the incompatible relationship.[8] This condition in (3) gives *Adjust* the courage of its convictions: the adjustment made to a node is often at odds with other aspects of the tree, but this must not scare the parser into giving up on it. Rather, the parser needs this new conflict to drive further repairs. Holding tight to the most recent attachment is the main difference between revision and first-pass parsing. First-pass parsing is subject to the *Revision as Last Resort* principle (RALR, Fodor & Frazier, 1980) which requires the parser to hold on to already built structure and make new input conform to it; structure built at the right side of the tree is constrained by decisions made on the left. By contrast, *Adjust* requires the parser to hold on to its ungrammatical emergency attachment at the right frontier of the tree, and let this determine what must give way further to the left.

A reanalysis system governed by *Attach Anyway* and *Adjust* first makes an illegitimate attachment and then makes changes to the CPPM to legitimize it.

---

8 Informally we call this the *Hold Tight On The Right* principle (HTOTR), though it could be the case that the most recently revised element in a dependency is on the left rather than the right. The only necessary truth about linear direction is that the source error must be to the left of the symptom. We note that even HTOTR does not exclude all circular revision sequences, though if necessary it could be strengthened to do so.

The alternative would be to first change the CPPM, postponing attachment of the conflicting input until it can be attached legitimately. The two strategies may seem to differ only trivially, but in fact the difference is critical. The illegitimate attachment demanded by *Attach Anyway* plays a pivotal role in the repair process because it converts a mis-match between the CPPM and subsequent input into a mis-match internal to the CPPM. At that point the *grammar* can be called on to identify the nature of the problem. The grammar knows what is an acceptable tree and what isn't, and it knows what is wrong with unacceptable ones. But without *Attach Anyway*, the grammar's ability to detect ungrammaticalities could not be put to work here. Without *Attach Anyway*, the lack of a well-formed attachment for the current input word *i* would prevent *i* from being attached. The CPPM would therefore remain incomplete, but it would be perfectly well-formed in accord with the grammar. It would be an acceptable CPPM, just like any other; it would be 'incorrect' only in the sense that it is not the CPPM *for this input string*. Internally it could not be faulted, so the grammar could not get a grip on it. The repair mechanism would therefore need a special diagnostic routine to exploit the information in the grammar in a reasoning process quite unlike first-pass tree building.

By contrast, an *Attach Anyway* parser puts the grammar to work to identify the nature of the problem that presents itself in a garden-path sentence at the point at which the parse breaks down. In diagnostic terms: the grammar identifies the error signal or *symptom* of the garden-path. The grammar can also be used to track from this point to the source error (i.e., the original attachment error earlier in the parse which caused the garden path). The grammar shows the parser where to look for the source error. This is another advantage that comes from starting the repair with attachment of the problematic current word, as opposed to trying to put things right in the tree before making the current attachment. The error locus is unknown; finding it is the *goal* of diagnosis. But the site of *Attach Anyway* is known; so the grammar is targeted efficiently to just the right spot. The parser will fix whatever needs fixing there, then shift attention to the damage done by that repair and fix that up, and so forth. This is how it finds its way to the true source of the trouble.

This successive repair process must have a way of knowing, at each step, where to shift its attention next. This, too, is information that the grammar can provide. The grammar can tell the parser which other nodes in the tree stand in a grammatical dependency with the node *n* currently being attended to by the parser. Necessarily, these are all and only the nodes that *n* could be in conflict with. So they are all and only the nodes which must be checked to see whether or not there is a conflict, and which must be altered if it turns out that there is. Checking the existence and well-formedness of dependencies between nodes in the tree may seem to be an enormous amount of work, but it is something that the first-pass routines must do all the time. An input word *i* can be acceptably integrated into the CPPM only if every potential relationship that *i* would thereby enter into in the tree is checked and found to be acceptable. Thus, every parsing model must do this or something equivalent to it, even if normally no mention is made of it; it is a process that is usually buried in implementation routines left unspecified. (But see Abney, 1986.) How heavy the computational burden actually is depends on the extent to which Universal Grammar and the grammar of the language constrain the relations that can hold between nodes: the fewer possible relations there are, the fewer relations the parser needs to check. In any case, we know that in the normal course of first-pass parsing the grammar is continuously providing the parser with information about dependencies. So the fact that *Adjust* needs this same information in order to know which node to consider revising next imposes no additional burden.

To summarize: By converting the mis-match between the input and the CPPM into a mis-match internal to the CPPM, *Attach Anyway* delivers a number of valuable consequences. It converts garden-path diagnosis from an open-ended reasoning problem into the familiar and concrete task of ensuring that all sub-trees of a CPPM obey the grammar. Specifically, an *Attach Anyway* parser can rely on the grammar to direct it toward exactly that piece of the CPPM that the newly attached input is incompatible with. The parser modifies that element to resolve the inconsistency. If that creates a new problem, the grammar will once again show the parser where and what the trouble is. Thus diagnosis and cure are not separate activities but are intercalated. And diagnosis is practical rather than abstract; it is simply the parser looking for its next repair to attend to. In the next sections we illustrate these general points with some examples.

3 The Grammatical Dependency Principle

If the current input item *i* is attached forcibly by *Attach Anyway* rather than *Attach*, then necessarily one of the new nodes introduced in attaching *i* is incompatible with some property of the prior CPPM. In most cases (though not quite all, see fn. 7) the offending new node stands in an improper grammatical relation with a node in the prior CPPM: a relation of mis-agreement, or mis-selection, or mis-match with an antecedent, etc. As noted, a conflict between two nodes can exist only if they stand in a relationship that is regulated by the grammar. This is so both for the first conflict, involving a new node introduced in the emergency attachment, and for later steps in the adjustment process; the adjustment mechanism is the same in all cases. Many grammatical relationships are quite local (structurally, if not linearly), e.g., subject-verb agreement. So very often, when a node conflicts with some other element in the CPPM, the two are in the same local sub-tree. But there are also long-distance grammatical dependencies such as pronoun binding, antecedent government, parallel structure requirements as in Gapping, and so forth. The scope of these various linguistic phenomena determines the size of the steps by which the parser makes its backward progress from the error signal to the original error site. We make this point explicit in the *Grammatical Dependency Principle*.

(4) *The Grammatical Dependency Principle* (GDP): When a grammar violation has been created in the CPPM by an action on node *n* in accord with *Attach Anyway* or *Adjust*, attempt to eliminate the problem by acting on a node that is grammatically incompatible with *n*.

The GDP is supported by empirical observations, such as the difficulty of revision in examples like (6) and (9) discussed below, compared with others such as (7) and (13)-(18). It is also in keeping with our general working principle that the revision routines draw on independently needed parsing mechanisms as far as possible. Together, these strongly suggest that the parser in *Adjust* mode tracks from node to node along pathways defined by grammatical relationships, and does not roam freely around the tree altering nodes wherever it pleases.

The GDP is (almost) implicit in the statement of *Adjust*, but since it is not a necessary property of a repair system, it is worth spelling out separately. (See section 5 for a version of the GDP which points out good repair sites but does not absolutely prohibit other revisions.) The GDP does have some empirical bite. For instance, it limits the movement of items from one phrase into another that happens to be nearby in the terminal string, if the phrases are not related to each other by any kind of grammatical dependency. As we discuss in section 5 below, this rules out some kinds of stealing repairs, in which an element in the tree grabs a linearly adjacent but unrelated constituent. This kind of stealing we call 'theft'. As we will show, there are contexts in which the parser attempts it, but with unhappy consequences because it violates the GDP. The GDP also rules out the complementary operation of stuffing unwanted material from one constituent into an unrelated one nearby. But the effects of a ban against stuffing are less drastic because stuffing of an element *x* (unlike stealing of *x*) is propelled by unacceptability of *x* in its current position; so the effect of stuffing *x* into a constituent *c* can also result from expelling *x* from its current local tree and subsequently reattaching it beneath *c*. Stealing cannot be thus simulated by two GDP-compatible operations, because there would be motive only for re-attachment, not for the initial detachment step.[9]

The GDP is a locality principle of sorts, but it does not impose a narrow definition of locality in terms of minimal distance. It condones some quite wide-ranging shifts—in fact, exactly as wide-ranging as grammatical dependencies can be. How great the distances are depends to some extent on the theory of grammar that is assumed, which may entail that what looks like a long-distance dependency is made up of a number of shorter links. In Government Binding (GB) theory, for instance, a 'long-distance' *wh*-movement dependency may be mediated by several coindexed empty categories; each link in the chain must satisfy Subjacency, the ECP and perhaps other locality conditions, though there is no limit on the length of the total chain. In HPSG a *wh*-dependency is composed of even more local links, since the SLASH feature that links filler and gap must be present on every node that dominates the gap but not the filler. On the other hand, HPSG also admits unmediated binding relations, as between an anaphor and an antecedent which is several nodes distant in the tree. GB does likewise as far as overt syntax is concerned, though it imposes a strict locality requirement on anaphors at LF (see Cole & Wang, 1996). Exactly how constraints at LF (or the Conceptual-Intentional interface) affect the parsing of overt syntactic structure is not yet well understood; though it is important, we cannot address it here. Also, it will suffice in what follows to assume basic mechanisms of GB theory, though translation into the Minimalist Program and into non-transformational frameworks should be possible.

In (5) there is an antecedence dependency between the object *himself* and the subject *the king's son*. In the garden-path sentence (6), a similar dependency exhibits a grammatical conflict: the anaphor is masculine but the subject is mis-parsed as feminine (cf. Fodor & Inoue, 1994, example (5)).

(5) The king's son admires himself.

(6) The daughter of the king's son admires himself.

In the first-pass parsing of (6) the grammar alerts the parser to the ungrammaticality of attaching *himself* as the object. However, for lack of a better option, *Attach Anyway* adopts this attachment despite the improper anaphor binding dependency. The grammar then directs the parser towards the

---

9 Stealing is thus a primitive revision operation, not composed of actions familiar from first-pass parsing. Expulsion, ousting, and other varieties of detachment are also specific to repair. Thus, as Paul Gorrell (p.c.) reminds us, it is not the case that repair needs no new mechanisms. We can claim only that diagnosis by *Adjust* borrows heavily from the first-pass routines for grammar application.

subject NP at the other end of the dependency. This is the site where repair must be made, as required by *Adjust* and the GDP. Thus in (6) the parser knows that *the daughter of the king's son* must somehow be transformed into a masculine NP.

This example is of special interest because, though the grammar can point the parser to the right neighborhood in the CPPM, and can tell it very precisely what the outcome of the revision must be, it cannot help the parser find a way of achieving that outcome. The repair trail leads, in accord with the GDP, from *himself* to the NP node over the subject, whose [–MASC] feature must be changed to [+MASC]. This in turn requires the N′ beneath it to change to [+MASC], again in accord with the GDP, following the lines of vertical relations within a syntactic projection. The next step is from the N′ to the N beneath it, the noun *daughter*. Its gender feature is changed from [–MASC] to [+MASC]. But when the parser checks the lexicon it finds that there is no item *daughter* [+MASC]. Therefore the repair path dead-ends here. In general, lexical specifications are non-negotiable; revision attempts are stopped short by contrary lexical facts. (Though see Konieczny *et al.*, 1994, and Fodor & Inoue, 1994, on cases where the parser may override the lexical input.) However, there is an alternative path to be explored in (6): the subject NP could become [+MASC] if *daughter* EITHER could become masculine, OR could abdicate its headship of the subject NP in favor of a masculine noun. Though the former is impossible, the latter is not. The offending noun *daughter* could be expelled from the structure. Assuming some housekeeping operations (see section 5.3), we may suppose that *daughter* would take *the* and *of* with it, leaving just *the king's son admires himself*. This would look like progress: the gender problem is thereby solved. But it creates a new problem: *the daughter of* must then be re-attached into the tree, and there is no obvious way to do so. The *Re-attach* operation, we assume, is exactly comparable to first-pass *Attach*. So if the parser tried to rejoin *the daughter of* and *the king's son*, the result would be [*the daughter of* [*the king's son*]], headed by *daughter* just as before. In some cases detaching and reattaching material brings about a useful reconfiguration of the tree (see Fodor & Inoue, 1994, fn. 17), but in this case it yields no gain.[10] An alternative way of achieving a masculine subject NP would be to promote either *king* or *son* to a position above *daughter*. In fact it is *son* that needs to be raised.

[10] Since *Re-attach* is governed by MA and LC, as *Attach* is (note that the GDP is not relevant), it might have been expected that *the daughter of* would reattach at the lowest site, to *the king* rather than *the king's son*, resulting in the correct analysis. However, we assume that the cliticization of the possessive *'s* onto *king* forms a surface word which cannot be split, so there is in fact no structural position between the node over *king* and that over *the king's son* that could host attachment of *the daughter of*.

But the GDP blocks this, because there is no grammar-defined route of dependencies leading to *son* from the anaphor (or from the current focus of *Adjust*). There is no selection relation between the two positions, or agreement relation, or case or theta assignment relation, or coreference requirement, etc. Because of this, the parser does not know to shift its focus to *son*; and the grammar offers no guidance to the parser on where else to look. Thus garden path recovery fails in (6), despite the very obvious feature mismatch that set the repair process going. The GDP predicts, correctly, that this is a very difficult garden path to recover from.

Example (6) contrasts with other cases of agreement feature mismatch in which revision is successful. Consider (7).

(7) The sick sheep cannot easily find food for themselves in the winter.

Here, a singular interpretation of *sheep* on the first pass prevents attachment of the plural anaphor. *Attach Anyway* will force the anaphor into the tree as the complement of *for* despite the number mismatch, since this at least violates no X-bar principles, case principles, etc. The grammar then directs attention to the subject position; the parser must find a plural NP there as antecedent for the anaphor. The parser runs down the N-projection path to the head noun *sheep*, changes its [–PLU] feature to [+PLU], consults the lexicon to see if there is such an item, discovers that there is, and so the revision goes through. Thus in this case there is no need for the parser to flounder around without GDP guidance looking for an alternative solution. Note that the ease with which (7) is reanalyzed demonstrates clearly that the leap from object to subject position is easy, and that it is the juggling of *daughter*, *king* and *son* within the subject NP that makes (6) hard. Evidently the distance across which the parser's attention must shift is not the primary determinant of difficulty. Thus the GDP is superior to a fixed distance restriction.

Though the comparison of (6) and (7) rules out a distance-based difficulty metric, it is not sufficient by itself to show that the GDP is the correct generalization. After all, (6) requires more extensive repairs than (7) does. Given just these examples, then, the generalization might be quite different, e.g., that the parser has the capacity to revise only one node or feature per garden path. We have no space to consider this possibility here, but in Fodor & Inoue (in press) we show that it is far from true. We present examples there which show that even extensive revisions requiring several tree changes and spanning considerable distances can be easy as long as they conform to the GDP.

4 The Thematic Overlay Effect

We now show that the GDP sheds light on the nature of stealing and its relation to the Thematic Overlay Effect. The Thematic Overlay Effect (TOE) is not a parsing principle but just a statement of the observation that "a thematic role once satisfied in a sentence will resist letting go of its syntactic realization" (Inoue & Fodor, 1995, p. 47). That is, once the parser has associated some input item(s) with a certain thematic role, it will resist any revision which re-assigns those words elsewhere without substituting others for them or eliminating the role entirely by overlaying a new argument structure over the old one. Why this should be so, we discuss below. In our earlier work we speculated that the parser may forget to erase the stolen argument from the initial representation of the sentence, resulting in confusion as one word string is illegitimately assigned two thematic roles;[11] but we will show later that this cannot be all that is going on.

Examples where TOE has been claimed to occur are shown in (8)-(11); (8) is from Frazier (1978), and (9) from Ferreira & Henderson (1991a); (10) and (11) are from Inoue & Fodor (1995).

(8) While Mary was mending the sock fell off her lap.

(9) While the boy scratched the big and hairy dog yawned loudly.

(10) John warned the children at the day care center were noisy.

(11) John promised the children at the day care center would be good.

We will concentrate on (8) and (9) here, and return to (10) and (11) in section 5.2. Examples (8) and (9) are Late Closure garden paths. In (9) the NP *the big and hairy dog* is initially attached as object of *scratched*. Then the disambiguating word *yawned* is encountered, and the impression is of a tug of war between *yawned*, which clearly needs the NP as its subject, and *scratched*, which does not want to let go of the NP as its object.[12] *Scratched* resists giving

---

[11] The reference to thematic roles would need to be reconsidered if TOE applies also to adjuncts. Whether or not it does has never been quite clear. An example such as *Because Ann wanted to study quietly in the kitchen was not a good place for her to sit* does not seem quite as difficult as examples like (8) and (9) below. If this is so, the proposals in section 5.3 offer an explanation: adjuncts are more labile because they are not required by other elements in the tree. But the revision of adjunct attachments deserves considerably more attention. On other differences between argument and adjunct processing, see Frazier & Clifton (1996).

[12] Stevenson (1993) has proposed a *competitive attachment* parsing model which exhibits a tug of war between the two verbs for essentially the same reason as TOE, i.e., if the first verb were to relinquish the NP it would have no object to satisfy its lexical selection feature. Like the original

up the NP despite the fact that this verb can easily be parsed as intransitive in other contexts where it is not temporarily assigned an object. For instance, in (12) the transitive misanalysis is blocked by the adverb *yesterday*, and intransitive *scratched* is not effortful to process. (The adverb doesn't fully block a Heavy NP Shift analysis, but we doubt that the parser favors it.)

(12) While the boy scratched yesterday the big and hairy dog yawned loudly.

What has to be explained, then, is why this same verb *scratched* is unhappy about being converted from presumed transitive to intransitive during reanalysis in (9).

Many other kinds of argument structure revision are tolerated comfortably by the human parser. Thematic reanalyses as in (13)-(15) do not cause any noticeable difficulty. (See Carlson & Tanenhaus, 1988.)

(13) a. We brought the homeless children to the orphanage.
b. We brought the homeless children some candy.

(14) a. John loaded the truck with bricks.
b. John loaded the truck onto the ship.

(15) a. The old man pushed to the front of the crowd.
b. The old man pushed to the front of the crowd the little child clutching a fistful of flowers.

Clearly, the revision of meaning in general, or of thematic roles in particular, cannot be the major reason for the resistance to revision of examples like (8) and (9). One possibly significant differentiating factor is that the thematic shifts in (13)-(15) are within-clause, while in (8) and (9) the ambiguous NP shifts across clauses. We will return to this point below (section 5.1). Note, however, that not all shifts across a clause boundary are equally hard: there is a striking difference in difficulty between the stealing of the object to become the subject of another clause in (8) and (9), and the stealing of an object to become the subject of another clause in (16) and (17):

(16) The boy noticed the big and hairy dog limped badly.

TOE, this account could be extended to account for (18) below, but apparently it mispredicts that the construction in (10) is as difficult to revise as that in (9).

(17) The host was expecting his best friends to refuse the invitation.

In these examples there is no undue persistence of the old argument structure. Even when the meaning encourages confidence in the first-pass simple NP object analysis, as in (17), the clausal complement analysis is easy to spot, and is stable once computed. By contrast, the misanalysis in examples like (8) and (9) is persistent. It varies in strength a little, but is generally quite robust. Ferreira & Henderson (1991a) have shown that it depends in part on the position of the head noun in the ambiguous NP: if the NP is head-final, revision is easier than if the head is several words prior to the disambiguation. We return to this interesting fact in section 5. Clifton (1993) and Pickering & Traxler (1994) have shown that revision of these constructions is faster if the ambiguously attached NP is semantically or pragmatically incompatible with the first verb. Also, the TOE is diminished in this construction if the first verb is an intransitive-preference verb (Ferreira & Henderson 1991b). The extreme case of this is a clause with a pure intransitive verb, as in the example *After the child sneezed the doctor prescribed a course of injections* from Mitchell (1987), where a spurious and fleeting transitive analysis is apparently computed. It is plausible to suppose that the parser gives in initially to the considerable pressures to attach the NP low on-line (see Phillips, 1995; Fodor & Inoue, in preparation a) but then sets about expelling the NP from object position if the NP violates the selection feature or the transitivity preference of the verb, or if it is implausible as object of the verb. This expulsion by the first verb (perhaps with energy and timing proportional to the degree of the violation) would reinforce the subsequent efforts of the matrix verb to steal the ambiguous NP from the first clause, and could thus make this stealing more successful than it otherwise would have been.

The contrast between examples like (9) and examples like (16) has attracted theoretical attention, and several explanations of it have been proposed besides TOE. Unlike TOE, most explanations focus on the fact that the stealing in (9) raises the NP in the tree structure, while the stealing in (16) lowers the NP. A long tradition in deterministic parsing inspired by the D(escription)-theoretic approach of Marcus *et al.* (1983) assumes that raising revisions are always costly, because they involve destruction of domination relations between nodes. The NP *the big and hairy dog* in (9) was dominated by the VP and IP nodes of the subordinate clause in the first-pass analysis but not in the revised analysis. By contrast, in (16) this NP was dominated by the VP, IP, and CP nodes of the main clause in the first-pass analysis, and still is in the revised analysis; it is also dominated now by the subordinate IP and CP, but addition of domination relations does not violate D-theoretic determinism. Gorrell (1995a, this volume) adopts a modified version of this approach, called Structural Determinism, in which both domination and precedence statements are indelible. Pritchett (1992) also prohibits raising a node, though for somewhat different reasons: his On Line Locality Constraint (OLLC) requires a revised node to be either governed or dominated by the position it occcupied in the first-pass analysis. Gibson's (1991) Recency Principle attributes the difficulty of examples like (9) to the fact that the matrix clause node was hypothesized earlier than the subordinate VP node, a fact that is partly correlated with the greater height of the matrix clause; the analysis with matrix attachment of the NP is therefore pruned from the set of analyses being computed in parallel, so that when the disambiguation requires it, it is unavailable.[13] Lewis (1993) ascribes responsibility not to the fact that the ambiguous NP must be raised, but to the fact that it is too distant structurally from the error signal to be acted on at all.[14] (See Sturt & Crocker, this volume, for a classification of models that penalize certain types of repair operations.)

There are examples, however, which show very clearly that neither raising nor a locality violation is what inhibits revision in this construction. These examples provide dramatic support for the TOE generalization that loss of the lexical realization of an argument is at fault. Consider sentence (18), which is like Ferreira & Henderson's example (9) except that the ambiguously attached medial NP consists of two NPs coordinated.

(18) While the boy scratched the little cat and the big hairy dog yawned loudly.

---

[13] Gibson's (1991) model, like that of Pritchett (1992), is based largely on thematic rather than structural relations, but it cannot account for the difficulty of (9) exclusively in terms of theta assignment. In Gibson's semi-parallel system, alternative analyses are constructed at the choice point in the parse, not at the disambiguation point—in this case, at the point of NP attachment, not after receipt of the matrix verb. At that point, high attachment of the NP is not much less attractive thematically than low attachment is, so the high attachment analysis is not pruned. Hence the tree-sensitive Recency Principle is needed to eliminate it.

[14] The Semantic Cost Principle of Frazier & Clifton (this volume) entails that revisions which alter the meaning assigned to the sentence are costly. This is applicable to (9), but it does not in general discriminate well between easy and difficult revisions. The cost is assumed to increase (though perhaps non-linearly) the longer the misinterpretation has persisted, and the role of an NP is assumed to be assigned only when its head is processed. Therefore the disambiguation in (9) occurs very soon after the semantic commitment is made, and the revision cost should be slight. Furthermore, the distance between semantic commitment and disambiguation is exactly comparable in (16), so the Semantic Cost Principle appears to mispredict that (9) and (16) are equally easy (or difficult) to repair. See also the discussion of the related proposal by Ferreira & Henderson (1991a) in section 5.1 below.

The intuition here is very strong. The whole coordinate NP is initially construed as object of the first verb. Then stealing occurs. Unlike in (9), the steal goes through quite easily, but only the second NP conjunct is stolen. This results in an ungrammatical, or at least incomplete, sentence; the whole string has been analyzed as a *while*-clause with internal coordination, and the main clause is missing. This structure is shown in (19), which has roughly the same meaning as *While the boy scratched the little cat and while the big hairy dog yawned loudly, ...*; it would be acceptable only if the input contained another clause, as in (20). Sentence (20), since it has this preferred reanalysis, is not difficult to grasp despite its complexity.

(19) While [[the boy scratched the little cat] and [the big hairy dog yawned loudly]]…

(20) While [[the boy scratched the little cat] and [the big hairy dog yawned loudly]] Kim slept.

The first-pass analysis with parsing breakdown is shown in (21). The unattained target structure is (22).

(21) [While the boy scratched [the little cat and the big hairy dog]] [yawned... !!!

(22) [While [the boy scratched]] [[the little cat and the big hairy dog] yawned loudly].

The parser goes from (21) to (19) and hangs up there without reaching (22). Clearly the secondary move to (22) from (19) is very difficult. (The reasons for this are not relevant to current concerns, but are quite straighforward.)[15] What matters for present purposes is that the change from (21) to (19) is *not* difficult. This has several important implications.

The ease with which the second NP of the coordinate structure is stolen by the new verb shows clearly that raising a constituent can be easy, as long as the circumstances are right. This is counter to the principles of determinism in all its variants, including D-theory and recent work in the same tradition such as Gorrell's Structural Determinism thesis. It also disqualifies Pritchett's OLLC,

[15] The necessary revision from (19) to (22) is like the difficult revision of (9) in that the *whole* NP object of the first verb must be stolen into the next clause, which will become the main clause. This leaves *scratched* intransitive. Furthermore, (19) (unlike (9)) has a null symptom, and null symptoms commonly fail to trigger reanalysis.

Gibson's Recency Principle and Lewis' locality principle since they too apply with equal force to *the big and hairy dog* in (9) and *the big hairy dog* in (18). But the facts are exactly as the TOE predicts. The contrast between (9) and (18) makes very clear what the parser does and does not like to do. It does not mind raising nodes; it does not mind altering the sentence structure, the sentence meaning, or the theta role assignments; but it does not like to deprive a predicate which has a theta role to assign, and which has successfully discharged that role onto a syntactic constituent, of the chance to discharge it onto *some* syntactic constituent, i.e., the parser does not like to demote a semantic argument role from overt (lexically realized) to implicit (unrealized).[16] In (18), the parser can avoid this by giving up only one conjunct, and retaining the other as object of the subordinate verb so that the verb can continue to discharge its theta role (though not to the same phrase as before). Then, as we see, the other NP conjunct is easy to entice upward.

This contrast between (9) and (18) strongly supports TOE as the right sort of description of the phenomenon. Related accounts are those of Stevenson (1993; see fn. 12), and Ferreira & Henderson (1991a; see section 5.1 below). However, (18) does not support our original explanation of TOE in terms of forgetting. We portrayed the squabble between the two verbs for possession of the NP as a problem of tidying up the sentence representation after stealing has taken place. It arises, we said, because the parser fails to erase the NP as argument of the first verb even while assigning it as argument of the second one; since the grammar knows this is improper, it resists, and confusion ensues. This is the linguistic counterpart of an 'impossible' visual figure in which one line contributes to two incompatible three-dimensional projections. (See also the example *I own the horse raced past the barn* in Gorrell, 1995a, p. 158.) However, if this were the correct explanation it should be so in (18) also. Failure to erase should create a tug-of-war over *the big hairy dog* in (18) just as much as in (9), and yet clearly there is no such effect; once the stealing of *the big hairy dog* is undertaken in (18), the NP snaps into place with no resistance from the lower clause.

---

[16] An intriguing alternative is suggested by a recent proposal by Bader (this volume), who argues that revision is inhibited when the prosodic contour as well as the tree structure must be modified. As Bader notes, this could account for the contrast between the constructions that are exemplified here by (9) and (16). It might also explain the easy revision of (18) if there is a weaker prosodic liaison between the two conjoined NPs in (18) than there is between the verb and its object NP in either (9) or (18). (We thank Patrick Sturt for this suggestion.) The only other account of (9) that we know of which can accommodate the contrast with (18) is that of Ferreira & Henderson, discussed in section 5.1 below. Like ours, it attributes the difficulty of (9) to the need to change the argument structure feature of the first verb.

The underlying idea can be preserved, however, if we make one small change: the spotlight should be not on tidying up the representation assigned to the sentence, but on tidying up the lexical selection feature of the first verb. Examples like (18) show that the CPPM (or the whole sentence representation, including its meaning) can undergo radical alteration quite painlessly (if it is adequately signalled). More particularly, the argument phrase assigned in the CPPM to *scratched* can be freely edited. But the lexical selection feature of the verb *scratch* cannot be freely edited. The verb was transitive before, and cannot easily be made intransitive now. Thus the contrast between (9) and (18) greatly narrows down the delineation of the TOE problem. The ease of revising structure (21) to structure (19) in the parsing of (18) rules out many conceivable explanations of the difficult revision in (9). It is not the case that all stealing is bad. It is not the case that all stealing upward or all raising is bad. It is not the case that meaning cannot be revised, even after the end of the clause or theta domain. The only thing that differentiates the difficult from the easy stealing is that the former requires a change in the selection properties of the first verb.[17] We must conclude that the difficulty is due to inaccessibility of an alternative argument structure feature for that verb.[18] *Secondarily*, this blocks alteration of the tree.

Note that there is one loophole in this lexical inaccessibility explanation of (9): it allows for recovery to be easy if it bypasses lexical revision. This could be

---

[17] This may be too strong. As noted in fn. 16, the possibility that there is a prosodic difference between (9) and (18) that affects revisability should also be considered. We note also that there is a difference in the ultimate height of attachment of the incoming verb and the stolen NP, which could be relevant since the preference for lower clause attachment is strong. However, the parser does not seem to be in doubt that *yawned* in (9) is the matrix verb; thus it appears that the high verb attachment is not the major problem.

[18] The original statement of TOE claimed that difficulty arises only if the theta role assigned to the NP by the first verb persists when the NP itself is stolen away, i.e., if it becomes an implicit argument. But the lexical inaccessibility explanation proposed here for (9) predicts difficulty if *any* change in the verb's argument selection feature is needed. We have been unable to find convincing test cases in which an argument is *added* during a non-GDP-guided revision. But even for stealing, the two accounts make slightly different predictions. The lexical inaccessibility explanation should apply not only to (9) but also to examples like (i)-(iii), in which the intransitive form of the first verb has only one thematic role, overtly realized by the subject NP.

(i) Just as my mother looked up the flight number appeared on the departures board.
(ii) As the water in the pot heated up the potatoes began to sizzle in the frying pan.
(iii) When the photographer moved his camera crashed against the lamppost.

It is not clear to us that these examples are quite as difficult as (9). Pending an experimental test of properly matched examples it would be premature to offer an explanation. However, we note that (9), unlike (i)-(iii), is *also* subject to the 'mild TOE' phenomenon that we discuss below (section 5.2). Thus a slight advantage of these examples over (9) would be explicable.

why revision of (18) by splitting the conjoined structure is easy. In (18), unlike (9), the tree could be reconfigured (switching *and* from being a linker of NPs to being a linker of IPs) without changing a lexical feature. It may be argued that a coordinating conjunction like *and* does not take arguments regulated by a selection feature in the lexicon, as a verb does. It is a functional category whose distribution is regulated by a rule schema in Universal Grammar (comparable to the X-bar schemata). The UG schema for coordination stipulates that a conjunction word must be flanked by categories of the same type, but otherwise it permits all categories to be conjoined (subject to some language-specific parameterization, e.g., only phrasal categories in some languages). When the parser attaches *yawned* in (18) and steals the immediately preceding NP to be its subject, it is creating an IP to the right of *and*. If it can rely on the general UG schema to approve this IP-conjunction analysis, it does not need to access the lexicon to know that it is legitimate. The schema will approve it on condition that the parser find a left conjunct that is also IP. This is possible in (18), so the revision goes through. This account of the difference in revisability between (18) and (9) thus rests on the different linguistic status of conjunction words and verbs. We propose an alternative explanation in section 5.3, but before presenting it we need to establish a more articulated theory of stealing. We do this in the next section.

To summarize: A closer look at TOE has led to a remarkable discovery which overturns most previous thinking about garden path recovery. Despite prior claims to the contrary, it appears that altering the (syntactic and/or semantic) structure that the parser has computed for a sentence carries very little psychological cost. Where major costs are observed, they must be due instead to inability to diagnose the problem or, as we now see, to some specific resistance to carrying out the cure. The lexical resistance that we have observed is actually misplaced—the parser is blind to lexical alternatives, and so it acts as if the current analysis were obligatory even when it is not. The essentially cost-free status of all types of repairs, as long as they are readily diagnosable, was a central claim of the diagnosis model of Fodor & Inoue (1994), but this discovery of lexical inflexibility is something of a surprise. The diagnosis model freely permits re-access of lexical entries, though it requires that, like all other adjustments, lexical re-access must be specifically prompted by a prior step in the diagnostic process. The parser takes actions as needed; it does not flail around at random trying out changes opportunistically. But even so, accessing the lexicon in (9) is an action needed, so why does it not happen? In the next section we address this interesting puzzle about when re-entry to the lexicon can occur and when it cannot.[19] The parser can find the intransitive specification for a verb like *scratched* in first-pass parsing. It can find alternative lexical specifications for verbs in other temporarily ambiguous contexts, like (13)-(15), in revision mode. Yet it cannot find the alternative lexical specification for *scratched* in revision mode in (9).

5 Capture and theft

Though lexical resistance can explain why the stealing of one conjunct in (18) is easier than the stealing of the whole NP, we actually lack an explanation of why any stealing is possible at all in this construction. At first sight, (18) appears to be a clear counterexample to the GDP. There is no grammatical relation between the stealer (the verb of the second clause) and the material stolen (part of the object NP of the first clause). There is also no relation between the subject position to which the stolen material must move, and the position of the NP that is to be stolen. Thus the GDP apparently prohibits stealing of the object NP or any part of it in both (9) and (18). Yet it is clear that stealing does occur in (18), and at least an *attempt* at stealing is made in (9)—this is the tug on one side of the tug of war between the two verbs. This is particularly clear by comparison with an example where there is no sense of any stealing being initiated at all, such as (23). In (23) the top clause ought to steal *in the library* from the lower clause, but it does not even try; see Fodor & Inoue (1994) for discussion.

(23) Susan put the book that she had been reading all afternoon in the library.

The facts thus seem to be insisting that in both (9) and (18) the parser *tries* to steal—in spite of the GDP—but it fails to make off with the booty when there is lexical resistance from the subordinate verb, as in (9). Why does the verb resist change? So far we have concluded that it must be because the parser cannot locate the lexical feature for *scratch* which says that it can be intransitive. And though there are alternatives, we have assumed that this is because the parser cannot make contact with the lexicon in the course of this repair (though it evidently can in the case of (13)-(17) and other examples). But *why* can the

[19] For expository simplicity here we have adopted the assumption that initial lexical access retrieves only one argument structure for a verb, and revision is blocked by inability to reaccess the lexicon. As discussed below (section 5.1), an alternative is to assume there is parallel lexical access of all the argument structures of a verb, and revision is blocked if the one that is needed has faded before the garden path is recognized. The explanation we propose goes through on either basis.

parser not re-access the lexicon while revising (9)? The fact that the stealing in (9) violates the GDP provides an explanation.

Because there is no grammatical dependency along which the parser can travel from the matrix to the lower clause, the parser is forbidden by the GDP to steal the object NP node or make any other *structural* change in the lower clause. But it could try stealing at a different level where access to structural facts is not at issue. That is, it could grab some nearby words in the word string to fill up the hole in the matrix clause where the subject ought to be. This is to steal linearly and very superficially from the word string, rather than structurally and legitimately from the tree.

As a parse proceeds, the words of the input string are assigned category labels and attached as the leaves of the tree. These terminal elements of the CPPM are what we assume superficial non-GDP repairs operate over, when the parser is denied access to the significant grammatical relations between nonterminal nodes. Note that though the stealing is superficial it is initiated by a structural fact. In (9) the finite verb needs a (nominative) NP as its subject. But if this structural need cannot be satisfied by *Adjust* under control of the GDP, it can be converted into a word string search. The grammar specifies that an NP has an obligatory noun head (at least normally; the noun might be phonologically empty, or the NP could be a pronoun). So the parser could hunt for a noun in the preceding terminal string. If it finds one, it will make off with it together with all the words that follow it. It may also need to drag along some preceding words (e.g., a determiner) to make a complete NP in the new position. (See discussion in section 5.3.)

This stealing of terminal elements is what we call *theft*. It contrasts with *capture*, which is the stealing of a structural node in the tree. Theft of a string of words (even if they add up to an NP) does not address the NP node, and so it does not set in motion the train of subsequent repair operations that would normally result when an NP node is removed. This differs from the capture of an NP node, as in (16) or (17). Because this is a tree structure revision, *Adjust* does tidy up after it; it makes all necessary further repairs (if it can) until the CPPM is fully grammatical again. The GDP directs *Adjust* toward positions grammatically related to the captured NP. In (16) one of these positions is the verb's argument structure feature. There is a dependency between this feature and the arguments actually present in the VP; if they don't match up, the sentence is ungrammatical. So when capture of the NP occurs in (16), the GDP prompts the parser to change the verb's argument structure feature. The lexicon is checked to see whether this is acceptable, and if it is, the revision is successful.

The GDP does not predict that the human parser engages in word theft. Indeed, the occurrence of theft indicates that the GDP is not an absolute law, but is more like a recommendation for orderly and successful revision. The repair system can break away from the grammar-guided *Adjust* mechanism (though we assume only if it has been tried and failed); but if it does, its actions are less well-motivated linguistically, and success is chancy. The assumption that *Adjust* normally respects the GDP predicts that there will be cases where structural adjustments are needed but cannot be made, thus creating a situation in which the parser is tempted to resort to word-level theft. Since word theft is not an instance of *Adjust*, it is not followed up by systematic further applications of *Adjust*. Among other things, it is not followed up by a return to the lexicon to ratify needed changes in selection features. This predicts exactly the three-way split between (9), (16), and (18) that we have observed. The CPPM resists tree alteration just in case it implicates lexical property alteration, in cases which do not conform to the GDP; that is, in cases like (9).

All that remains to account for is which cases of stealing are node capture and which are word theft. Why can *limped* capture *the big and hairy dog* in (16) while *yawned* can only commit theft in (9)? The examples are repeated here for comparison.

(16) The boy noticed the big and hairy dog limped badly. CAPTURE

(9) While the boy scratched the big and hairy dog yawned loudly. THEFT

The answer has to do with where the verb is placed in the CPPM by *Attach Anyway*, which is determined by the interplay of preference principles and degrees of ungrammaticality mentioned in section 2. Without entering into a full discussion here, we sketch what seems reasonable for these two constructions. In the NP-versus-clause garden path in (16), the CPPM contains no attachment site for the finite verb *limped* when it is encountered: there is no finite VP awaiting its verb. The parser must create a VP somewhere. It could try positing a higher VP, with the current clause as its subject; it could try a medium-height attachment in a VP conjoined to the *noticed*-VP; or it could try a low attachment in a clause subordinate to the current one, a complement or relative. Low attachment is the usual preference. In any case, if the higher sites were attempted the revision would subsequently fail. The only thing that will succeed in (16) is low attachment into a complement clause, so let us track it through. The finite verb *limped* projects a subject [SPEC,IP] position to its left which it needs to fill. It can achieve this by tucking in next to *the big and hairy dog* in the CPPM; the projected subject NP can unify with this NP in the CPPM, thereby importing the IP (and CP) projections also. Two things then need to be checked: that the preceding verb *noticed* can accept a complement clause and give up its object NP; and that *the big and hairy dog* can change its feature values appropriately now that it is governed by *limped*. Its case feature must change from ACC to NOM, and its agreement features would need checking also, though in this example they happen to be neutralized on the verb. In (16) all outcomes are positive so no further adjustments are required; the repair is satisfactory.

The stealing of the NP in (16) satisfies the GDP because the stealer (the incoming verb *limped*) was not previously in the tree; and when it attaches it does so right at the site of the NP it steals, and enters into a subject-predicate relation with it. So there is no question as to whether its position in the CPPM is appropriately related to the NP it needs to steal; the process of attachment makes it so. This is the classic case of capture: the current input item *i*, in process of being attached into the CPPM, steals a node X by attaching itself to X as its sister, beneath X's mother, with secondary adjustments as necessary (addition of dominating nodes, revision of features such as case, agreement and selection features). Now compare the stealing in (9). Here the incoming verb *yawned* attaches far from the NP it needs to steal. It cannot snuggle in right next to the NP. (It may try to, since this is a desirably low attachment site, but it will be rejected when the lexicon is checked, because *scratch* does not take a clausal complement.) So *yawned* cannot capture *the big and hairy dog* on-line. The only place for *yawned* to attach is in the anticipated structure of the matrix clause. This attachment site is less preferred because it is less low, and has the additional drawback that the subject projected by the verb is missing; there will be a dangling NP node that lacks lexical content. But this just means that the verb must steal some lexical content for it. Note, however, that the verb *yawned* is now attached some distance away in the tree from the NP it needs to steal, so it has to reach across the tree from one position to another. Reaching across to the NP in the subordinate clause is blocked by the GDP, however. The two positions are not grammatically related, hence no legal stealing can occur. The parser can only resort to superficial theft of words (which, as we have seen, does not go well in this example). To summarize: The essential fact that forces theft rather than capture in examples like (9) is that there are two separate stages: attachment somewhere, then reaching out to steal the NP from somewhere else. Unlike capture in (16), the attachment stage drags the parser's attention away from the NP to be stolen. If the parser can't get back to the NP with the blessing of the GDP, word-level theft is its only option.

It is unparsimonious to postulate two kinds of stealing. But the contrast between theft and capture that we have outlined explains several other noteworthy facts. It predicts the effect of head position discovered by Ferreira

& Henderson (1991a). It explains what has long seemed a strange and suspicious variability in the strength of TOE across different construction types. And it offers an interesting explanation of why a coordinate NP may be stolen without resistance.

5.1 The Head-Position Effect

Ferreira & Henderson (1991a) presented sentences to subjects word by word centered on a screen at a rate of 250 msec a word, and subjects gave timed grammaticality judgments. Acceptance of a garden-path sentence as grammatical was taken as evidence that the perceiver had succeeded in revising the incorrect first-pass structure. The mean "grammatical" judgments for four types of sentence are given below with typical examples. It should be noted that these results are from a single series of experiments with similar procedures, but not from the same experiment; therefore comparison of the outcomes can only be approximate. Ferreira & Henderson's paper should be consulted for more precise information.

(24) While the boy scratched [the *dog*] yawned loudly. ("grammatical": 61%)

(25) While the boy scratched [the big and hairy *dog*] yawned loudly. (51%)

(26) While the boy scratched [the *dog* that Sally hates] yawned loudly. (24%)

(27) While the boy scratched [the *dog* that is hairy] yawned loudly. (29%)

The recovery rate is quite low even at best, but we will be concerned here primarily with differences across the sentence types. In (24) the ambiguous NP is short and the recovery rate is the highest. In (25) the ambiguous NP is long but recovery is almost equally easy. In (26) and (27) the NP is long and reanalysis is more difficult than in (25), even though (27) roughly matches (25) in meaning. Ferreira & Henderson propose the generalization that "head location affects ease of reanalysis… the greater the distance from the head of the misanalyzed phrase to the syntactically disambiguating word, the harder the process of reanalysis becomes" (1991a, pp. 741-742). (We have italicized the head noun of the misanalyzed NP in (24)-(27).)

This effect of head position is an important finding, not anticipated in investigations based on intuitive judgments of processing difficulty. It is explicable given the GDP prediction that only theft can supply the missing subject in this construction. The parser is focused on the second verb and is seeking a noun in the terminal string preceding it. When the relevant noun immediately precedes the verb it should be easier to spot than when it is further back. For capture operations the position of the noun within the NP should not matter, since in terms of tree structure the NP node immediately precedes the verb in all cases. Thus it is predicted that comparable manipulations of head position for a garden path cured by capture, such as (16), would have little or no effect on revisability. But theft must check backward through all the intervening terminals between the verb *yawned* and the noun *dog*, so the success of theft should decline as the distance in words increases, as Ferreira & Henderson discovered.[20]

The head position effect might be accounted for in a number of other ways (e.g., head-driven parsing as defined by Pritchett, 1991), but we will not review them here since most of them fail to account for the ease of stealing the second conjunct in (18). The only exception, besides the superficial theft explanation, is Ferreira & Henderson's own proposal. They suggested that the possible argument structures for a verb are accessed in parallel though only one is adopted, and that unadopted ones start to decay once theta roles have been assigned. Assuming that a theta role is assigned to a phrase only when its head is processed, the difficulty of revising the attachment of a phrase will begin to increase once its head and its theta-assigner have both been processed, though only for revisions that entail change in the argument structure feature of the verb. In (26) and (27) the noun is followed by several other words before the disambiguator, so the intransitive argument structure has time for significant decay and there is a high probability that it will have vanished entirely by the time it is needed. In (24) and (25) the NP is head-final so the intransitive argument structure has little time to decay before the disambiguating verb is encountered. Note that the initial decay rate must be assumed to be quite steep since the acceptance rate even in these cases of immediate disambiguation is only 50-60%; this is the basic TOE phenomenon, which an early head position simply exacerbates.

Like the GDP/theft explanation of the TOE, and unlike all others in the literature, this decay proposal implicates inaccessibility of the correct lexical argument structure feature for the verb rather than the cost of making alterations

---

[20] Though less reliable than capture, it seems we must admit that theft does often succeed (up to 60% of the time under ideal circumstances, if the acceptance rates in Ferreira & Henderson's study reflect accurate reanalyses; see discussion below). Success may be due to assistance from expulsion (see section 4), or to lucky word theft followed by a confirmatory reparse. These are compatible with the strong claim that word theft by itself never results in a stable reanalysis involving lexical revision. We think this may be correct, but at present we have no grounds for excluding the possibility that nodes which exclusively dominate thieved words are pied piped along with them, provoking the checking of selection features as in legitimate node stealing, so that if by good luck the right NP node is pied piped, the theft will be approved.

in the syntactic/semantic representation of the sentence. It is therefore fully compatible with our observation that raising an NP is easy if the verb is not thereby entirely deprived of an object. However, the decay explanation (unlike the GDP explanation) apparently predicts that examples like (13)-(17) should be as difficult as (9) if matched for distance between the head noun of the mis-assigned NP and the error signal. We believe this is incorrect. Though there are no experimental data bearing directly on this predicted comparison, intuitive judgments strongly suggest that these other constructions are much less resistant to revision than (9) is. If this proves to be so, the decay model is incorrect as it stands. It would need to invoke additional factors that could either exempt (13)-(17) or else magnify the effect for (9).

Ferreira & Henderson (this volume) have developed a different explanation of this set of facts, in order to accommodate some new experimental findings; but it is of interest nonetheless to consider whether a decay model could plausibly account for the presumed contrast between (13)-(17), and (9). One possible assumption would be that the decay of unadopted argument structure features accelerates sharply when the clause (the argument structure domain) is closed off. We don't know whether this is so, but it is plausible: the parser would not normally have any use for such features once there is no further possibility of finding arguments for the predicate. In (13)-(17) the revision occurs while the parser is processing one of the arguments of the relevant verb. That is, the garden-path symptom occurs prior to any word not in one of the verb's arguments, or to any other signal that the verb's argument domain is closed; and the symptom word itself is such that it keeps that argument domain open. In (16), for example, the symptom *limped* is such that it must attach within the clausal complement of *noticed*. But in (9) the symptom *yawned* is not part of any argument of the first verb; it can only be the matrix verb. So as *yawned* attaches, it closes off the domain of *scratched*, triggering a sharp drop in accessibility of alternative argument structures for *scratched*. Some such discontinuity in the decay rate would explain why, in this particular construction, recovery is difficult even with immediate disambiguation, as Ferreira & Henderson's data show. (The difference in argument domain closure between (16) and (9) is not unrelated to the fact that (16) has the clause order main-subordinate, while (9) has subordinate-main order: a main clause can contain a subordinate clause, but not vice versa. This is the remnant of truth, perhaps, in explanations that point to raising as the source of the problem in (9).)

With some such modification, the decay theory could handle all the examples under discussion. On present evidence we cannot tell whether it or the GDP account of the TOE is correct (though the decay theory makes what we believe to be overly strong predictions for left-branching languages such as Japanese). Though alike in pinning the blame on loss of lexical flexibility, the two accounts differ in other respects.[21] For example, revision by theft predicts that when the head noun is not final in the NP, the parser may be very sensitive to the presence of other superficially appealing items closer to the thieving verb, as in *While the boy scratched the puppies of the big hairy dog yawned loudly*. An incomplete revision might occur in such a case, giving the perceiver the impression of having recovered from the garden path but not in fact achieving the correct analysis.[22] This and other differentiating predictions of the two explanations need to be tested.

5.2 Two Strengths of TOE

The TOE phenomenon has always shown a puzzling variability in its strength. In (8) and (9) the verb strongly resists giving up its object, but in other examples such as (10) and (11), repeated here, the resistance is slight.

(10) John warned the children at the day care center were noisy.

(11) John promised the children at the day care center would be good.

Both types of construction involve loss of the lexical realization of a thematic role. Why should they differ as they do in recoverability? Let us consider

---

[21] As presented here the two accounts differ with respect to whether lexical access puts all of a verb's selection features into the CPPM, or only one. But this difference may be unimportant. The account of (9) in terms of theft that we have proposed is compatible with either serial or parallel access of lexical features. Even if the intransitive specification of *scratch* were present in the tree in (9) when the disambiguation occurred, the parser would be unable to refer to it for much the same reason as given above: theft of terminal items does not target the NP node, so the GDP would not direct attention to the verb's selection feature (with which the NP stands in a dependency), so the parser would fail to recognize the possibility of switching to a different feature.

[22] Ferreira & Henderson (this volume) found no significant effect of a 'decoy' noun close to the verb in sentences such as *When the boy scratched the dog that hates Sally yawned loudly* (34% grammatical judgments) as compared with *When the boy scratched the dog that seems furry yawned loudly* (38% grammatical judgments). This may mean that the theft account is wrong, or it could be that subjects sometimes thought they had succeeded in grasping the sentence structure (and so reported that it was grammatical) when in fact they had not. To check whether the judgements that subjects give are sometimes over-optimistic, it will be important to find out whether ungrammatical sentences such as **While the boy chased the dog barked loudly* are also sometimes 'successfully' reanalyzed. Also, the discussion in section 5.3 below implies that it could matter whether there is a natural 'break point' in the middle of the NP. For instance, with *the puppies of the big hairy dog* there may be a tendency to keep *the puppies* as object of the first clause, and steal *the big hairy dog* as subject of the main clause (but the *of* is an embarrassment that would have to be ignored).

(10).[23] The garden path followed by revision is noticeable here but not devastating; repair is achievable. On a scale of difficulty running from (16) (easy) to (9) (difficult), (10) seems to be closer to the easy end. For more precision let us compare (10) with (28), which is the same construction as (16) but matched to (10) in lexical content.

(28) John knew the children at the day care center were noisy.

The verdict, though only informal and perhaps not universal, is that (10) is not much harder than (28), with the important consequence that (10) is significantly easier than (9). Thus we seem to observe the TOE phenomenon in two strengths: mild in (10) and severe in (9). We argue now that these are in fact two different phenomena, though they have important aspects in common.

An explanation for why (10) does not exhibit the TOE at full strength follows from the GDP and the distinction between capture and theft. Though (10) resembles (9) in that its first verb loses an overt argument, it resembles (28) in *how* it loses it. The stealing in (9) is theft, as discussed above, but in (10) it is capture, as it is in (28). We assume that the post-verbal NP in (10) first attaches as object (Goal) of *warned*. The disambiguating verb *were* is finite and plural, and ideally would attach to the right of a plural NP. It finds no pre-existing verb attachment site in the CPPM. Its least undesirable attachment is as sister to the preceding NP, with introduction of clausal nodes above them, i.e., it captures the object NP. Directed by the GDP, the parser adjusts the lexical properties of the matrix verb (*warn*) to suit this new clausal context, and checks the lexicon to make sure that this is acceptable, which it is. Thus *warn* can be flexible about

---

[23] Some perceivers find (10) difficult to parse because *warn* is not fully grammatical when it has neither an indirect object nor a complementizer. But the points of concern to our discussion can be made equally well with (11), which is generally regarded as acceptable. The construction with *warn* was first noted by Pritchett (1987) who judged it much harder to revise than examples with verbs of the *know* class, such as (28) below. We believe that the difficulty of Pritchett's early *warn* examples was due to their marginal grammaticality, and also to the fact that their revision involved other complications, as in (i) where there is uncertainty between an analysis like (10) and an analysis in which the NP *her mother* has been broken into two NPs with recategorization of *her* from determiner to pronoun.

(i) Mary warned her mother hated her.

When these complications are avoided, as in (11), the difference in revisability between *warn*/*promise* and *know* constructions seems much reduced. However, Pritchett (1992) does not agree with this judgement. He maintains (pp. 82-ff.) that sentences like *Rex warned the ugly little man feared him* and *The teacher promised the students could leave* are (a) well-formed and (b) "unprocessable". Since it is important to our discussion to distinguish between mild and severe TOE garden paths, we must set aside Pritchett's judgments, which do not seem to be representative.

the loss of syntactic realization of its Goal NP, since the GDP reminds the parser to fit the verb's selection feature to it. However, the loss of the NP does meet with some slight resistance in (10); the revision goes through, but it is not quite as painless as in (28). It seems that this mild TOE in (10) must be explained as in past proposals, such as Pritchett's (1987) Theta Reanalysis Constraint, or the original 'forgetting to erase' version of the TOE. In some way or other, the problem is that the bearer of the Goal role was specified in the first-pass analysis of (10) and after the repair it no longer is. Note that it is the *losing* of information that is discomforting; (10) is worse than (29) where the Goal is unidentified from the start.

(29) John warned there were some noisy children at the day care center.

Exactly why this loss bothers the parser is not crucial to our general theory of tree repair, but it is still an interesting question. We suggest, tentatively, a slight variation on our 1995 proposal, an account which implicates confusion though not forgetting. In (10) the parser thought it knew the identity of the Goal, but the revision leaves it uncertain. It is not even clear whether the original incumbent of the role can be retained once the role becomes implicit rather than explicit. The identity of the Goal *can* be changed, but must it be? The answer varies from case to case: it is not a matter of syntax but of semantics and plausibility. Who did John warn that the children were naughty? Probably not the children. Also not himself, so presumably some third party not mentioned. Who did John promise that the children would be good? Probably not the children. But if John promised the children would have cake for lunch, who did he promise? Probably the children. The comprehension system must evaluate the candidates in each case. This is so whether the issue arises on the first pass or in revision. But in revision the usual processes of evaluation could well be disrupted by the initial illusion of having had precise information based on syntax rather than pragmatic guesswork. This is not far from the original idea behind the TOE, but now it is clear how slight the effect is. In Inoue & Fodor (1995) we observed the same for Japanese; after extraneous complications in Pritchett's examples were cleared away, TOE was detectable but weak. In general, then: loss of identification of the bearer of a thematic role in the course of an otherwise successful revision has a detectable but small effect on processing difficulty. This is distinct from the severe TOE in constructions like (9), where with a high probability (see Ferreira & Henderson's data above) the repair process fails to retrieve the lexical material it wrongly assigned to a thematic role. Something else is needed to explain that, and we have given evidence above that it is due to perseveration of a lexical feature.

To summarize: The examples we have compared permit a separation of two aspects of the problem that we dubbed the TOE. One is a mild disappointment or confusion at losing a clear identification of the bearer of a thematic role, after that role has been brought into focus by the first-pass analysis. This is represented by the comparison of (10) with (28) and (29). The other is a persistent rebellion against giving up an overt argument due to the mistaken belief that it is obligatory, where the mistake is due to failure to access lexical alternatives. This is represented by the comparison of (8) or (9) with all other examples discussed here. The two phenomena are not totally dissimilar. In both cases "a thematic role once satisfied resists letting go of its syntactic realization" during a garden-path revision. The difference that we overlooked in our previous work is between resistance that prevents any argument structure substitution at all, and distress after substitution has occurred if the new argument structure leaves a thematic role unclothed without eliminating it entirely. With the distinction between capture and theft, we are now able to explain when and why each of these occurs. The severe TOE occurs in cases of theft, where failure to reaccess the lexicon blocks adoption of any new argument structure. The needed revision appears to the parser to be syntactically ill-formed. The mild TOE occurs in cases of capture, where revision of argument structure is possible.[24] If the revision overlays the role in question (e.g., (28)) it is easy; if it does not there is a mild TOE experience as the parser tries to sort out what effect that has on the sentence meaning.

5.3 Conjunction and Raising

The contrast between (9) and (18) shows that stealing an NP out of a conjunction is quite different from stealing an NP away from a verb. We repeat the examples here for convenience.

(9) While the boy scratched the big and hairy dog yawned loudly.

(18) While the boy scratched the little cat and the big hairy dog yawned loudly.

It is a familiar fact that revision of an NP conjunction into a clausal conjunction, as in (30) and (31), is not particularly difficult. The garden path is perceptible on-line but it does not persist.

[24] Note that *both* phenomena apply to a 'severe TOE' construction like (9). That is, if the parser does spot the correct analysis of such a construction (see fn. 20), the successful revision would then create a mild TOE, adding to perceived difficulty.

(30) John was talking to Bill and the twins were talking to Jane.

(31) The boy scratched the little cat and the big hairy dog yawned loudly.

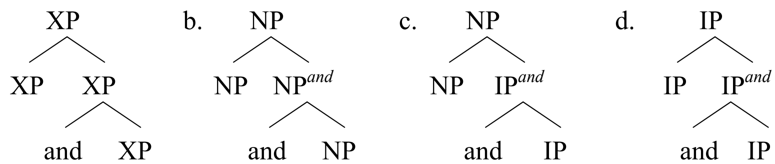
This revision involves raising the NP that follows *and*, with the elimination of several domination relations. The NP moves out of the argument domain of the first verb. The conjunction word tags along with the NP, and is raised from inside the first VP up to the root level. These are the kinds of revisions that on most theories would predict a severe garden path, but that is not what is observed. In examples like this we have long had evidence that not all raising repairs are difficult. (This is not to say there is no coordination garden path at all; see Hoeks *et al.*, 1997.) The only novelty of (18) is that a revisable coordinate structure is put into a context where raising material out of the first clause is notoriously troublesome; but raising a conjunct is found to be easy nevertheless. So far our interest in this has been in what it shows about the Late Closure garden path construction in (8) and (9). But we turn attention now to coordination structures and ask why they are more easily revisable than the internal structure of VP. We suggested in section 4 that this is because *and* needs no lexical subcategorization for the complements it takes; it can take conjuncts of any category as long as they come in matched pairs. However, now that we have distinguished between theft and capture, a more interesting explanation is possible.

Consider (31) at the point at which the verb *yawned* is encountered. It is finite and needs a subject. In principle the parser might build a second clause, attach *yawned* in it, and then steal a preceding NP as its subject. But in general the parser strongly favors low attachment sites, for *Attach Anyway* as well as *Attach*, as noted earlier. So the preference would be for *yawned* to attach low and capture a preceding NP there. It could capture the whole coordinate NP or just *the large hairy dog*. The latter would be favored by the preference for low attachment. (If the higher attachment were tried it would require a three-way coordination of clauses for lack of a conjunction word between the first two; but in any case it would not thrive since it would run into TOE trouble as in (9).) The IP resulting from capture of *the large hairy dog* by *yawned* becomes the complement of *and*, taking the place of the NP complement.[25] Let us assume, as is standard, that coordination is not symmetric but has the structure shown in

[25] There is a terminological problem here hiding a theoretical problem. To call the conjoined element the complement of *and* is to imply that *and* is the head of the coordinate structure. If that is so, the standard theory of phrasal projection needs to be refined, since with NP complements the structure has the properties of an NP, while with IP complements it has the properties of an IP, and so forth; see Kayne (1994). Whichever way this linguistic issue is eventually settled, the substance of our present claims should not be significantly affected.

(32a). (We use XP^{and} in (32) as an arbitrary label to distinguish an XP marked with a coordinating conjunction from a 'plain' XP.)

(32) a. b. c. d.



UNIVERSAL SCHEMA FIRST-PASS ANALYSIS INTERMEDIATE STAGE REVISED STRUCTURE

Once IP is the sister of *and*, the NP^{and} node dominating them must be relabelled as IP^{and}, as shown in (32c). But this IP^{and} is incompatible with its NP sister and mother. The IP^{and} needs an IP left sister. The parser may try to convert the NP on the left into an IP somehow, but this will fail. It can't make *the little cat* into an IP; and there is no way to steal other constituents to combine with *the little cat* to make up an IP. So the IP^{and} does not fit with its context. The only thing the parser can do to continue this line of revision is to detach the IP^{and}. By good fortune this turns out to be just the right thing to do. The IP^{and} once free can find an IP on its left to join with. To do so it must rise up to the top of the CPPM, but evidently this is not costly. The revision is thus successful. Note that it is predicted to be easy whether *and* has a lexical entry of the standard kind or not.[26] If it does not have a standard lexical entry, the explanation goes through as in section 4; if it does, the fact that this repair started with a capture operation means that the lexical feature associated with *and* was changed accordingly, it was then checked against the lexical entry, and the revision was approved, subject to appropriate adjustment of the left conjunct. Thus the difference between stealing an NP from *and*, which is easy, and stealing an NP from a verb, which is difficult (at least in a structural context as in (9)), can be explained in terms of the difference between capture and theft.[27]

---

[26] On balance, we are inclined to suppose that *and* and other coordinating conjunctions do have lexical entries containing selection features. Though part of the instantiation of a UG schema, they differ from each other in ways that the grammar must specify (e.g., *either/or* can take NP or IP complements while *neither/nor* can take NPs but not IPs).

[27] To forestall confusion we note that NP capture followed by expulsion and high re-attachment of the resulting clause is *not* possible in (9). We assume capture is tried in (9), and the selection feature of *scratched* in the CPPM is accordingly changed from selection of NP to selection of IP (or CP). But the lexicon then rejects this combination of *scratched* with an IP selection feature, so this line of revision can go no further. This is unlike the case of conjunction, where the feature change is approved by the lexicon and the repair process can proceed. The revision of (9) could continue along some other lines, but in fact there is no other legitimate action to be taken. The verb cannot expel the IP because there is no conflict between them; the conflict is internal to the verb. Reversion

It is of interest that though revision of NP coordination to IP coordination is not painful, it is not always the preferred revision. Gorrell (1995b) has observed that when a coordinate NP is the ambiguous NP in an NP-versus-clausal-complement garden path, as in (33), the preferred revision retains the NP coordination structure and shifts it bodily from matrix object to complement subject position. Gorrell claims that (33) is preferentially reanalysed as in (34a) rather than (34b).

(33) John knew the men and the women were strangers.

(34) a. John knew [[the men and the women] were strangers]
b. [John knew the men] and [the women were strangers]

We agree with Gorrell's intuition about the preference here. It shows that the object-to-subject revision is less costly in this construction than breaking up the coordinate NP would be. This provides the valuable information that it is not the case that a coordinate NP always crumbles in the face of any parsing conflict. It splits easily in examples like (30) and (31), and it splits in (18) where the alternative revision meets lexical resistance, but it holds out against the alternative revision in (33). We may conclude that splitting a conjunction is harder than easy revisions but easier than hard ones. Though this is true as far as it goes, we can be more specific.

In Gorrell's example (33), a reanalysis like that sketched above for (31) is in competition with one in which the second verb captures the whole coordinate NP. The choice is between the verb *were* capturing the lower NP node over just the right conjunct, and its capturing the higher NP node over the conjoined structure. If it captures the higher NP, the lexical feature of *knew* must be adjusted (as in (16) and (28) above). If the lower NP is captured, the reanalysis continues with adjustment of the structure around *and*, as described for (31). To account for the observed preference for the analysis (34a) we must assume that the usual preference for low attachment is offset by the extra work involved in relabelling and reattaching the IP^{and} for (34b). This is in accord with the Minimal Revision principle (another instance of Minimal Everything; see Frazier, 1990; Frazier & Clifton, this volume; Inoue, 1991). Since all minimizing preferences are implicitly comparative, this carries the interesting

---

to the transitive verb feature would violate HTOTR (see fn. 8). The verb's selection feature cannot be erased to relieve the tension, because even an intransitive verb must have a selection feature. The lexicon cannot volunteer the information that *scratched* has an intransitive form, because each step in the repair process must have a specific trigger. The verb cannot be expelled along with the IP, as *and* can, because it is not now a licit lexical item; and in any case the first clause would not survive without it.

suggestion, which we cannot pursue further here, that alternative repairs are undertaken in parallel, in a race, so that the first one that succeeds in resolving the problem wins. Though a lower attachment site is preferable to a higher one in general, attachment to the higher NP node gives a faster solution in (33) and so that is what the parser does.

In (18) the same choice presents itself except that here the higher attachment option is problematic. Just as for (9), the whole object NP cannot be captured because the verb *scratched* does not accept an IP. Stealing the object by reaching down from the matrix clause is theft and runs up against a severe TOE. Since high attachment is thus spoiled in (18), the option of capturing the lower NP is a clear winner. Finally, we may consider (9) again in light of these other examples ((31), (33), (18)). Here the stealing verb has no choice between lower and higher NPs. There is just one NP to steal; it cannot be captured, and theft is inhibited by the TOE, so the revision is difficult. The differing perceptual profiles of these four examples are thus accounted for by *Attach Anyway* and *Adjust* guided by the GDP.

Our claim has been that raising is as easy as any other repair, except where a construction has some specific property that makes raising difficult. Clearly there is no such factor at work in the coordinate structure revisions discussed here in which an IP^{and} is raised. This needs explaining also. Why can the IP^{and} in (18) and (31) float upward without restraint?[28] The answer we propose is that nothing is trying to hold the IP^{and} down. The NP to the left of *and* exerts no selection pressure on it. Selection is an asymmetric relation. A conjunction word needs its flanking XP conjuncts, but an XP does not need a following conjunction word. Thus the NP left sister to the IP^{and} does not protest the escape of the IP^{and}. Indeed it is glad to see it go, since it is a misfit where it is; the IP^{and} is expelled, rather than stolen by an external attractor. Thus there are

---

28 The IP^{and} can even rise over two clauses as long as the symptom clearly signals it. Consider (i) and (ii).

(i) Susan thinks that Matthew had lunch with his sister and the twins did too.

(ii) Susan thinks that Matthew had lunch with his sister and the twins do too.

When the verb *did* captures *the twins* in (i), the resulting IP^{and} rises just far enough to take the lowest available IP as its left sister. But when *do* captures *the twins* in (ii), the IP^{and} rises one clause further. The effort is perceptible, perhaps, but the revision is successful (more successful, certainly, than NP raising in (9)). There is no resistance to the IP^{and} leaving its current position, and a good symptom (the tense and aspect of *do* in relation to the preceding verbs) indicates clearly where it should reattach. Under these circumstances there appears to be no limit imposed on how high it may rise. (We cannot, however, rule out the possibility that in (ii) the IP^{and} raises in two steps, trying out the lower clause attachment on the way.)

two factors contributing to the ease of this revision, as compared with stealing an object away from a verb. One, noted above, is that the incoming verb can *capture* the preceding NP and does not have to attach high and resort to theft. The other is that the word *and* moves with the ambiguous NP, so that the constituent that raises is not attempting to escape from something which selects it and has an interest in retaining it. Indeed, the reason why the *and* raises along with the IP is that *and* does have an interest in retaining a constituent on its right (though it is flexible about its category). Given that *and* needs that constituent, it has two options: it could block the raising of the constituent or it could raise with it. The first leads to a dead end in (18) and (31). The second is successful and painless, since nothing on the left complains when *and* moves.

This perspective on the 'pied piping' observed in some reanalyses offers clarification of an observation in section 4. We noted there that word theft (to the extent that it succeeds at all) may pull other words along with the targeted one. Since theft is a superficial operation, it seems odd that it should care about taking a complete well-formed phrase. But in fact we can see now that this could be due rather to the other words refusing to let go of the one being stolen. For instance: when a noun is stolen and a determiner goes along with it, this is not because the noun is dragging the determiner, but because the determiner is tagging along with the noun. Before the stealing, the determiner was already in the CPPM, so it was part of a structured tree reflecting grammatical constraints. When its noun is pulled away from it, those constraints set up a protest. The determiner's selection feature requires a noun complement (an overt one, except in special contexts). We assume this selection feature is present in the tree (as a verb's selection feature is), and is left unsatisfied if the complement is ripped away. (Note: this explanation goes through whether the determiner is analyzed as the specifier of NP or as the head of DP; only details differ.) As suggested above for conjunction, a determiner whose noun is the target of attempted theft has two options: it can try to prevent the noun from going, or it can go along with the noun. The latter may well be the easier course. How easy it is depends on whether there is something to the left of the determiner that cannot do without a DP to its right, such as a preposition. If there is, that will have to tag along too or it will block the whole process. For success in thieving there must be some point at which the word string can snap apart, i.e., a point at which the element on the left has no need of the element to its right. For a correct revision, this weak link must be at the point where the structure needs to be divided. In (31) the break to the left of *and* is perfect. In an example like *While John scratched the puppies of the hairy dog yawned loudly* there is a natural break to the left of the preposition *of*, but this is the wrong split from the point of view of the sentence structure, so revision is difficult. In (9) there is no natural snapping point at all (the verb won't let go of its object, the subject won't let go of the verb, the adverb *while* won't let go of the clause); so no reanalysis presents itself, not even a spurious one.

6 Conclusion

The data for a theory of garden path reanalysis are the differential costs of recovering from different kinds of garden paths. The goal is a model of the reanalysis process which explains those facts, i.e., a model which entails true generalizations about when recovery is easy and when it is difficult. In searching for these generalizations it is most reasonable to begin with simple strong hypotheses such as: revision is easy if and only if it does not alter the meaning assigned to the sentence; if it does not require erasure of domination statements; if it involves only local restructuring; etc. However, counterexamples to all of these hypotheses can be constructed. The diagnosis model that we have proposed is flexible enough to cope with the variety of examples observed.

A disadvantage of greater flexibility is that it may imply a more elaborate mechanism; but the diagnosis model, we claim, is exempt from this charge for two reasons. As we have noted previously, every repair model must have some machinery for calculating what repairs are possible in the case at hand and for selecting one repair to try out, as well as for effecting the appropriate changes in the tree. The diagnosis model assumes nothing more than this in order to predict reanalysis costs. But we have also tried to establish in this chapter that diagnosis does not require a rich reasoning component with open-ended deductive capabilities out of keeping with the dedicated modular design that we believe is characteristic of the first-pass routines. We have presented a practical reasoning machine that thinks by doing. It has no general powers of logic, but linguistically it is just as sophisticated as the first-pass parser, since it draws on the grammar in just the same way. It fails only where linguistic sophistication is of no benefit. Specifically, we have argued that the *Adjust* process fails if there is no linguistically motivated route, composed of grammatical dependency links, from the symptom back to the original error. There may be no such connecting path despite the fact that the original error was the cause of the apparent error that the parser is now confronted with. As in the examples discussed here, the original error may have used up some input words which really belong at a later position, a position structurally unrelated to it though necessarily linearly adjacent. In this way, an earlier choice may impinge on later options without there being any grammatical relation between the tree positions which *Adjust* could take advantage of. In such cases the adjustment process will blunder about and may fail to find the right revision, however obvious it might be to a general reasoning system.

References

Abney, S. 1986. Licensing and parsing. In J. McDonough and B. Plunkett (eds.), *Proceedings of NELS*, 17, 1, 1-15.

Bader, M. (this volume). Prosodic influences on reading syntactically ambiguous sentences.

Carlson, G.N. & Tanenhaus, M.K. 1988. Thematic roles and language comprehension. In W. Wilkins (ed.), *Thematic Relations, Syntax and Semantics, Vol. 21*. New York, NY: Academic Press, 263-300.

Clifton, C. 1993. Thematic roles in sentence parsing. *Canadian Journal of Experimental Psychology*, 47, 2, 222-246.

Cole, P. & Wang, C. 1996. Antecedents and blockers of long-distance reflexives: The case of Chinese *ziji*. *Linguistic Inquiry*, 27, 3, 357-390.

Cooper, W.A. & Paccia-Cooper, J. 1980. *Syntax and Speech.* Cambridge, MA: Harvard University Press.

Ferreira, F. & Henderson, J.M. 1991a. Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language*, 30, 725-745.

Ferreira, F. & Henderson, J.M. 1991b. How is verb information used during syntactic parsing? In G.B. Simpson (ed.), *Understanding Word and Sentence*. North Holland: Elsevier Science Publishers B.V., 305-330.

Ferreira, F. & Henderson, J.M. (this volume). Syntactic reanalysis, thematic processing, and sentence comprehension.

Fodor, J.D. & Frazier, L. 1980. Is the human sentence parsing mechanism an ATN? *Cognition*, 8, 417-459.

Fodor, J.D. & Frazier, L. 1983. Local attachment in a one-stage parser. Unpublished manuscript, University of Connecticut, Storrs, CT.

Fodor, J.D. & Inoue, A. 1994. The diagnosis and cure of garden paths. *Journal of Psycholinguistic Research*, 23, 407-434.

Fodor, J.D. & Inoue, A. (in press). Garden path recovery: The Grammatical Dependency Principle. To appear in *CUNYForum*, 20, Fall 1997.

Fodor, J.D. & Inoue, A. (in preparation a). The Diagnosis Model: Revision as Last Resort. To appear in M. De Vincenzi & V. Lombardo (eds.), *Cross Linguistic Perspectives on Sentence Processing*. Dordrecht: Kluwer Academic Publishers.

Fodor, J.D. & Inoue, A. (in preparation b). Null symptoms and empty categories. Manuscript, CUNY Graduate Center, New York, NY.

Frazier, L. 1978. *On Comprehending Sentences: Syntactic Parsing Strategies*. Unpublished doctoral dissertation, University of Connecticut, Storrs, CT. Distributed by the Indiana University Linguistics Club, Bloomington, IN.

Frazier, L. 1990. Identifying structure under X^0. In A. Jongman & A. Lahiri (eds.), *Yearbook of Morphology*, 3, 87-109.

Frazier, L. & Rayner, K. 1988. Parameterizing the language processing system: Left- vs. right-branching within and across languages. In J.A. Hawkins (ed.), *Explaining Language Universals*. Oxford, UK: Basil Blackwell, 247-279.

Frazier, L. & Fodor, J.D. 1978. The sausage machine: A new two-stage parsing model. *Cognition*, 6, 1-34.

Frazier, L. & Clifton, C. 1996. *Construal*. Cambridge, MA: MIT Press.

Frazier, L. & Clifton, C. (this volume). Sentence reanalysis, and visibility.

Gibson, E.A.F. 1991. *A Computational Theory of Human Linguistic Processing: Memory Limitations and Processing Breakdown*. Unpublished doctoral dissertation, Carnegie Mellon University, Pittsburgh, PA. Available as Center for Machine Translation Technical Report CMU-CMT-91-125.

Gorrell, P. (this volume). Syntactic analysis and reanalysis in sentence processing.

Gorrell, P. 1995a. *Syntax and Parsing*. Cambridge, UK: Cambridge University Press.

Gorrell, P. 1995b. Parsing theory and phrase-order variation in German V2 clauses. *Journal of Psycholinguistic Research*, 25, 135-156.

Hoeks, J., Vonk, W., Hagoort, P., & Brown, C. 1997. Processing coordination: Eye movements and ERPs. Poster presented at *Tenth Annual CUNY Conference on Human Sentence Processing*, Los Angeles, CA.

Igoa, J.M. 1995. Parsing decisions and the Construal hypothesis: Attachment preferences in primary phrases in Spanish. Paper presented at *Second Symposium on Psycholinguistics*, Tarragona, Spain.

Inoue, A. 1991. *A Comparative Study of Parsing in English and Japanese.* Unpublished doctoral dissertation, University of Connecticut, Storrs, CT.

Inoue, A. & Fodor, J.D. 1995. Information-paced parsing of Japanese. In R. Mazuka & N. Nagai (eds.), *Japanese Sentence Processing*. Hillsdale, NJ: Lawrence Erlbaum Associates, 9-63.

Kayne, R. 1994. *The Antisymmetry of Syntax.* Cambridge, MA: MIT Press.

Konieczny, L., Hemforth, B., & Scheepers, C. 1994. Reanalysis vs. internal repairs: Non-monotonic processes in sentence perception. In B. Hemforth, C. Scheepers, & G. Strube (eds.), *First Analysis, Reanalysis and Repair*. IIG-Berichte 8/94. Freiburg: Institut für Informatik und Gesellschaft, 1-23.

Lewis, R.L. 1993. *An Architecturally-Based Theory of Human Sentence Comprehension*, Unpublished doctoral dissertation, Carnegie Mellon University, Pittsburgh, PA. Available as Technical Report CMU-CS-93-226 from reports@cs.cmu.edu.

Lewis, R.L. (this volume). Reanalysis and limited repair parsing: Leaping off the garden path.

Marcus, M.P., Hindle, D., & Fleck, M.M. 1983. D-theory: Talking about talking about trees. *Association for Computational Linguistics*, 21, 129-136.

Mitchell, D.C. 1987. Lexical guidance in human parsing: Locus and processing characteristics. In M.Coltheart (ed.), *The Psychology of Reading, Attention and Performance, Vol. 12*. Hillsdale, NJ: Lawrence Erlbaum Associates, 601-681.

Ni, W., Fodor, J.D., Crain, S., & Shankweiler, D. (submitted). Anomaly detection: Eye movement patterns.

Phillips, C. 1995. Right Association: A single strategy for structural parsing. Paper presented at *NELS 26*, MIT, October 1995.

Pickering, M. & Traxler, M. 1996. Plausibility and recovery from garden paths: An eye-tracking study. Unpublished manuscript, University of Glasgow.

Pritchett, B.L. 1987. *Garden Path Phenomena and the Grammatical Basis of Language Processing*. Unpublished doctoral dissertation, Harvard University, Cambridge, MA.

Pritchett, B.L. 1991. Head position and parsing ambiguity. *Journal of Psycholinguistic Research*, 20, 3, 251-270.

Pritchett, B.L. 1992. *Grammatical competence and parsing performance*. Chicago, IL: University of Chicago Press.

Stevenson, S. 1993. A competition-based explanation of syntactic attachment preferences and garden path phenomena. *Association for Computational Linguistics*, 31, 266-273.

Stevenson, S. (this volume). Parsing as incremental restructuring.

Sturt, P. & Crocker, M.W. 1996. Monotonic syntactic processing: a cross-linguistic study of attachment and reanalysis. *Language and Cognitive Processes*, 11, 5, 449-494.

Sturt, P. & Crocker, M.W. (this volume). Generalized monotonicity for reanalysis models.

Wanner, E. & Maratsos, M. 1978. An ATN approach to comprehension. In M. Halle, J. Bresnan, & G.A. Miller (eds.), *Linguistic Theory and Psychological Reality*. Cambridge, MA: MIT Press, 119-161.