Reanalysis in Sentence Processing: Rebuilding or Repair?*

Sungki Suh
(The Catholic University of Korea)

Suh, Sungki. (2002). Reanalysis in sentence processing: Rebuilding or repair? *Language Research* 38(4), 1235-1255.

This paper concerns the distinction between rebuilding and repair in the process of reanalysis in sentence processing. Fodor and Inoue (1994, 2000) propose that recovery from garden paths is repair rather than rebuilding, and that the difficulty of repair is tied to the cost of diagnosing the error in the first-pass processing. While the diagnosis model in Fodor and Inoue (1994, 2000) can account for a reasonable range of garden path phenomena, it is questionable whether its operating principles can apply in a consistent way. In particular, it doesn't seem clear why lexical reaccess is not possible in certain cases and the diagnosis of the error in those cases fails. Meanwhile, the rebuilding mechanism of a parser based on structural determinism consistently distinguishes between conscious and unconscious reanalyses by considering the structural configurations in the reanalysis procedures. Provided that both rebuilding and repair mechanism show comparable empirical coverage, the former is more desirable due to its conceptual adequacy and consistency.

Key words: reanalysis, garden path, repair, rebuilding, diagnosis model, structural determinism

1. Introduction

The way a reanalysis component is implemented in human sentence processing mechanism is directly related to the core architecture of the human parser, and hence considered as an important criterion for the classification of various parsing models. Traditionally, reanalysis module has been considered to be tied to whether the parsing model in question

* This research was supported by a 2002 research grant from the Catholic University of Korea. I am grateful to the three anonymous reviewers for their comments. Any errors in this paper are solely mine.

pursues multiple structural possibilities simultaneously or one structural possibility at a time when it faces structural ambiguity.

Recently, however, there have been some proposals emphasizing the distinction between rebuilding and repair in the process of reanalysis. In particular, Fodor and Inoue (1994, 2000) claimed that the traditional parsing models which regard reanalysis as rebuilding procedure are not adequate empirically, and hence 'repair' rather than 'rebuilding' should be the mechanism for reanalysis. The essence of Fodor and Inoue's 'repair model' is that the difficulty of the recovery from a garden path is basically tied to the cost of diagnosing the error in the first-pass processing. This proposal is contrasted to the central claim found in the recent deterministic models, in which difficult recovery from a garden path results from a specific kind of structural alteration of the previously computed structure.[1)]

In this paper, we compare Fodor and Inoue's diagnosis model with a standard deterministic model. Then we focus on the deduction steps necessary for the repair operations in Fodor and Inoue's model, and see whether such steps are conceptually adequate. We want to show that while Fodor and Inoue's model may handle enough empirical data, such empirical coverage can't be a strong argument for their repair model, since it lacks conceptual adequacy in certain stages.

This paper is organized as follows. In section 2, we overview the mechanism of a typical deterministic parser, in which reanalysis basically means rebuilding the structure. In section 3, we consider how the diagnosis model works and why reanalysis should mean repairing rather than rebuilding in this mechanism. In the next section, we compare the Diagnosis model with a typical deterministic model, and consider whether the strategies and conditions governing the Diagnosis model can be generalized. Finally, our discussion is wrapped up in section 5.

2. Serial Parsing and Determinism

2.1. Serial vs. Parallel Parsing

At the onset of a structural ambiguity, the human parser can do either

1) In particular, there has been a consensus among the recent psycholinguistic theories assuming determinism that structural alteration amounting to node raising always creates conscious processing difficulty, whereas node lowering is hardly problematic.

of the following: (i) it constructs more than one structure, and abandons particular structure(s) as they become incompatible with subsequent material; (ii) it constructs only one structure, and revises it later when it is not compatible with subsequent string. (i) is a typical parallel parsing model and (ii) is a serial parsing model.

There are some arguments against a strict parallel parsing model. Consider the following.

(1) #While Mary was mending the sock fell.

(2) #The horse raced fell.

A strict parallel parser will deal with the structural ambiguity in (1) in the following way: Considering the fact that *mend* can be either transitive or intransitive, a parallel parser will compute more than one structure for *the sock* at the moment, i.e., *the sock* as the object of *mend*, and as the subject of the following matrix clause. Given this, it becomes mysterious why the parser is garden-pathed at the very next moment, i.e., when *fell* is encountered; since a parallel parser can consider the possibility of positing *the sock* as the matrix subject, there shouldn't be any processing difficulty. The fact that (1) is a severe garden path sentence implies that a strict parallel parser is not psychologically real.

Similar remarks hold for (2): Since *raced* can be either a past-tense verb or a past participle, a parallel parser will construct more than one syntactic structure at the moment. Then it becomes puzzling why a garden-path results immediately after the next word *fell* is encountered.

This said, it is highly unlikely that the human parser employs a strict parallel fashion. Of course, it is not totally impossible for a parallel parser to deal with the above garden path phenomena. In fact, the parallel parsers proposed by Kurtzman (1985) and Gorrell (1987) adopted the idea of ranking and reanalysis, and hence were able to handle the garden path sentences such as (1) and (2). However, since the idea of ranking or reanalysis is actually the central notion of serial parsing, the parsing model proposed by Kurtzman (1985) or Gorrell (1985) can hardly be considered as a typical parallel parser. Given that such a parallel parser is not really distinguished from serial parsing models, it seems reasonable to assume that the human sentence processing mechanism is basically serial. (See Gibson & Pearlmutter, 2000 for the related discussion.)

2.2. Determinism

2.2.1. Marcus Parser

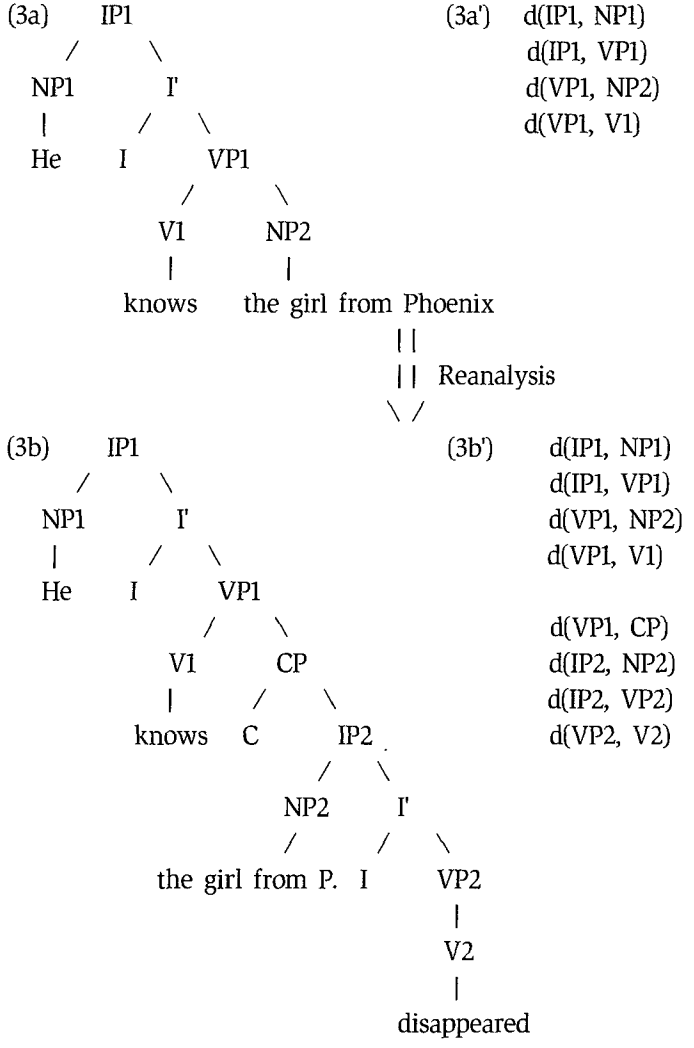
Deterministic parsing was first proposed by Marcus (1980). Marcus hypothesized that the sentences not creating conscious processing difficulty can be parsed strictly deterministically. A strict deterministic parser chooses a single analysis at every ambiguous point, and does not backtrack. The merit of a deterministic parser is that the parsing architecture itself accounts for garden path phenomena: Parsing is deterministic, and when it cannot be deterministic, i.e., when it should retract structural commitments made earlier, it breaks down.

The core problem for a strict deterministic parser is how to deal with the pervasive ambiguity of natural language. Since there exist temporarily ambiguous structures which do not lead to garden-path, it was necessary for a deterministic parser to employ some lookahead device in order to handle such ambiguities. However, adopting a lookahead device resulted in many incorrect predictions. Hence, for the sake of the psychological plausibility of the parser, it became inevitable to modify the strict determinism of the Marcus parser and to remove the lookahead device.

2.2.2. Description Theory

Marcus' (1980) parser has been further developed by Marcus, Hindle and Fleck (1983), which retains the determinism hypothesis but assumes a different parse output. In their system, the output of the parse consists not of a tree structure but of a syntactic description, which is statements about 'dominance' relations among categories. Due to such characteristics, their theory is called D-theory (=description theory). D-theory can also be considered as an underspecification theory since structural descriptions can be underspecified due to the fact that the predicate 'dominance' encodes constituency, but not necessarily immediate constituency. Crucially, employing 'dominance' rather than 'direct dominance' as the predicate in structural descriptions makes it possible to lower a constituent in a parse tree without backtracking. The following example is relevant to this point.

(3) He knows the girl from Phoenix disappeared.

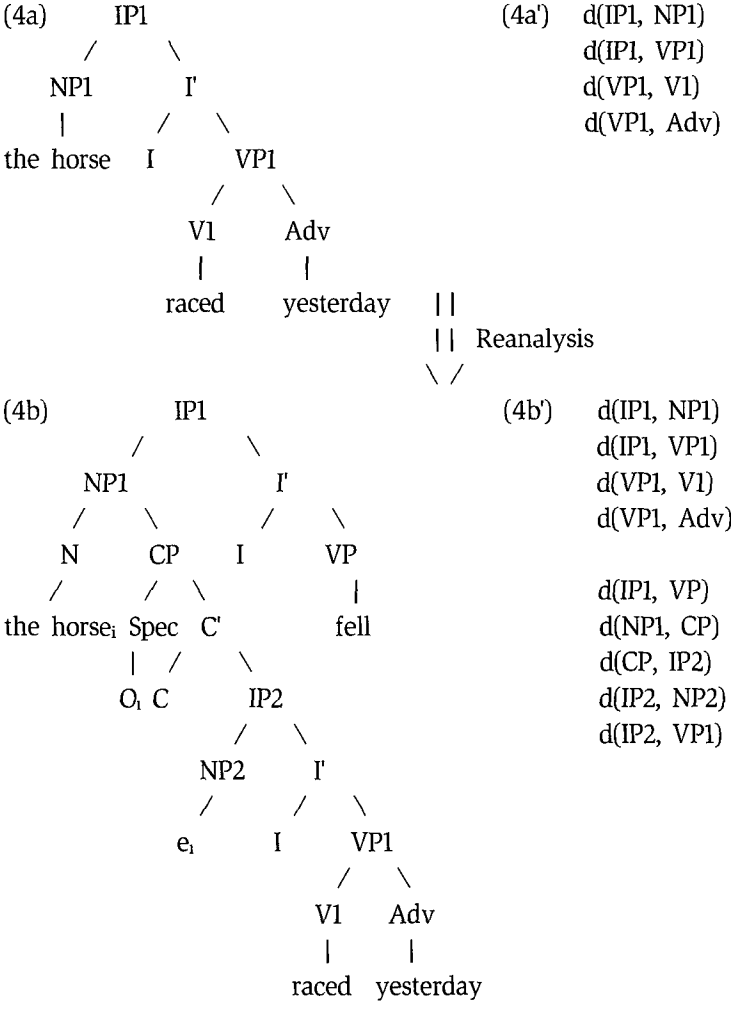


Although *the girl from Phoenix* is initially misanalyzed as the object of *knows* in (3a) and then reanalyzed as the subject of the complement clause in (3b), such a reanalysis is predicted to be unproblematic in the following sense: What has been done in the reanalysis can be considered as adding dominance relations such as 'd(VP1, CP)' and 'd(IP2, NP2)'. As a result, VP1 node does not "directly dominate" NP2 any more, but it still "dominates" NP2. Crucially, none of the dominance relations asserted in

(3a') have been retracted in (3b'). In this respect, there is no violation of determinism, and hence the reanalysis is unprocessable.

Even though D-theory parser brought about improvement over Marcus' (1980) parser, it still faces empirical problems in the respect that the reanalysis amounting to adding dominance relations can create conscious processing difficulty. Consider (4).

(4) #The horse raced yesterday fell.



According to D-theory, the reanalysis in (4b) is predicted to be

unproblematic simply because it can be considered as lowering VP1 into NP1. Notice that all of the dominance relations asserted in (4a') remain true in (4b'). Nevertheless, (4) actually creates severe garden path. Such a result suggests that at least some type of lowering is costly.

2.2.3. Structural Determinism

There have been many proposals for dealing with the problematic data such as (4) within the boundary of determinism. Gorrell (1994, 1995) is one of them, and he has proposed a parsing model governed by a constrained form of determinism, called structural determinism. In his model, only the primary structural relations, i.e., dominance relation *and* precedence relation, are subject to determinism.[2] Such a proposal seems desirable, both conceptually and empirically. Given that both precedence and dominance relations should be computed in order to create a legitimate parse structure, it is rather questionable why only dominance relation is subject to determinism in D-theory and other deterministic parsing theories. Unless there is an independent reason to distinguish dominance from precedence in applying determinism, it is natural to hypothesize that both are constrained by determinism. Such a hypothesis is also supported by empirical evidence. In the previous section, it was pointed out that lowering a constituent sometimes leads to a garden path, as in (4), repeated here as (5).

(5) #The horse raced yesterday fell.

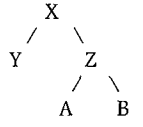
We have seen that the processing difficulty observed in (5) undermines

---

2) Following Wall's (1972) Exclusivity Condition below, Gorrell (1995) considers 'precedence' to be more than left-to-right order.

(i) Exclusivity Condition: In any well-formed tree, either 'x precedes y' or 'y precedes x' is true if and only if neither 'x dominates y' nor 'y dominates x' is true.

According to (i), two nodes are in a precedence relation when they are not in a dominance relation. Hence, Y in the following diagram, for instance, precedes not only Z, but also A and B, which are not in a dominance relation to Y.

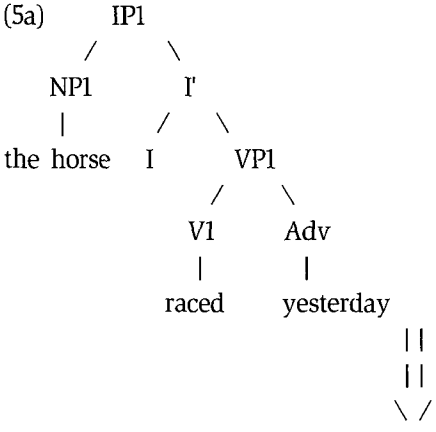


One generalization we can draw from (i) is that if 'x' precedes 'y', then all nodes dominated by 'x' precede all nodes dominated by 'y'.

proposals such as D-theory; the reanalysis necessary for arriving at the final parse stage of (5) is incorrectly predicted to create no garden path in D-theory since all of the dominance relations asserted in the earlier parse stage remain true after the reanalysis.

Meanwhile, Gorrell's theory correctly predicts the garden path effect in (5) by resorting to precedence relation. Consider (4a) and (4b), repeated here as (5a) and (5b).

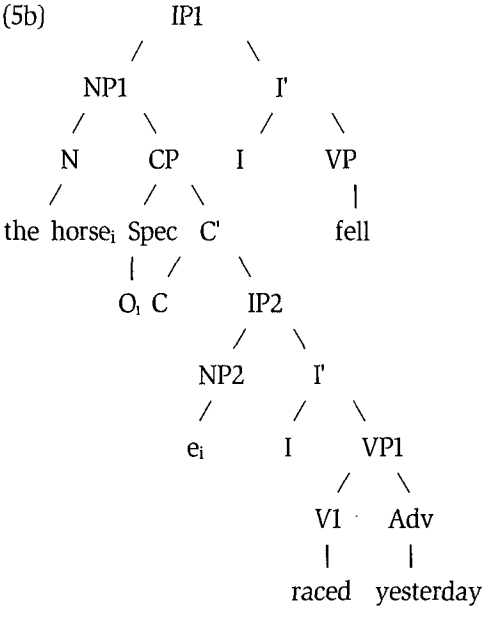
(5a)



(5a')

d(IP1, NP1)
d(IP1, VP1)
d(VP1, V1)
d(VP1, Adv)

(5b)



(5b')

d(IP1, NP1)
d(IP1, VP1)
d(VP1, V1)
d(VP1, Adv)

d(IP1, VP)
d(NP1, CP)
d(CP, IP2)
d(IP2, NP2)
d(IP2, VP1)

The precedence relation between NP1 and VP1 asserted in (5a) no longer holds in (5b), in which NP1 "dominates" VP1. Altering a primary relation in that way is a violation of determinism, and hence a garden path results.

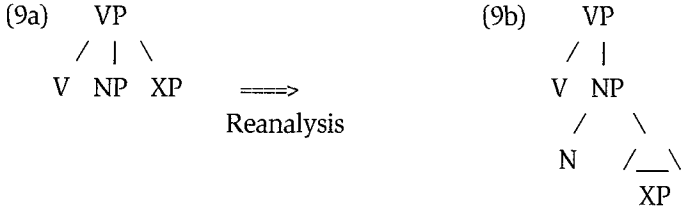
Structural determinism predicts that lowering creates severe processing difficulty when a constituent is lowered into its predecessor precisely because the precedence relation between the constituent and its predecessor is retracted in that case. Such a prediction is borne out by the following data.

(6) #She gave the boy the dog bit a bandage.

(7) #John told the teacher that he hired a story.

(8) #Bill put the candy on the table in his mouth.

What is common in the above sentences is that the verb takes more than one argument; namely, a Goal and a Theme NP in (6) and (7), and a Theme NP and a Locative PP in (8). The processing difficulty in each case results from the initial misanalysis of the phrase immediately following the first argument of the verb: In (6) *the dog* is initially misattached as a Theme of *gave*. In (7) *that he hired a story* is initially misconstrued as a propositional clause. In (8) *on the table* is initially misanalyzed as a locative argument of *put*. In the final analysis, those misanalyzed phrases turn out to be (part of) the adjunct phrase modifying the preceding argument NP. The following diagram represents such a reanalysis procedure.



Since XP is lowered into a preceding NP, the precedence relation between them is altered to a dominance relation, which is a violation of determinism. Hence, a garden path results as predicted.

3. Diagnosis Model and Repair

Fodor and Inoue's (1994, 2000) diagnosis model has originated from Inoue's (1991) information-paced parser. In this section, we review those parsing models and clarify the notion of 'repair' in those systems.

3.1. Information-paced Parsing

Inoue (1991) proposed that parsing head-final languages as well as head-initial languages can be done in a serial fashion, if we assume that reanalysis is not the last resort but a routine part of parsing. The motivation for such a proposal is that there are relatively few garden paths in Japanese, a strict head-final language, even though temporary ambiguities arise frequently due to its head-final property.

Inoue (1991) claimed that parsing Japanese is done in such a way that structural decisions are made in advance of reliable information and revised later whenever they conflict with the subsequent material containing reliable information. Inoue tried to capture the difference between head-initial languages and head-final languages in terms of the availability of relevant information at the decision point. The following statement is relevant to that point: "The cost of revising a decision is a function of the confidence of the parser in making that decision, which will depend on the quality of the information on which it was based. For a head-initial language such as English, relevant information is usually plentiful, so confidence will be high, and the parser will be quite obstinate about holding to the analysis it has assigned. But for a head-final language like Japanese, access to relevant information is typically delayed, so the parser cannot be confident of its decisions, and so will be flexible about revising them." (Inoue, 1991, p. 148)

An information-paced parsing model seems to provide a direct account of the general parsing phenomena in typical head-final languages such as Japanese; namely, frequent temporary ambiguities but relatively few garden paths. However, the universality of Inoue's parsing model seems questionable when we consider some English garden path sentences such as (1) and (3), repeated here as (10) and (11).

(10) #While Mary was mending the sock fell.

(11) He knows the girl from Phoenix disappeared.

(10) and (11) are analogous in terms of the pace of the information flow for initial analysis and reanalysis: *the sock* in (10) and *the girl from Phoenix* in (11) are initially attached as the object of *mending* and *knows*, respectively. Such attachment decisions could be revised immediately upon receiving the subjectless verb, *fell* in (10) and *disappeared* in (11). Consequently, the cost of recovery from the initial misanalysis is expected to be the same between (10) and (11), if the flow of necessary information is the main factor determining processing difficulty. Yet, the consequence of reanalysis in each sentence is quite different: (10) results in garden path whereas (11) creates no conscious processing difficulty. In order to account for the contrast between (10) and (11), an information-paced parser cannot but be supplemented by some additional conditions.

3.2. Repair in the Diagnosis Model

Fodor and Inoue (1994, 2000) refined Inoue's information-paced model in such a way that the difficulty of a reanalysis depends not on the cost of effecting the structural alterations but on the cost of deducing which alterations are needed. Their theory can be regarded as a descendant of Inoue's information-paced model since the diagnosis of the necessary structural alterations mainly relies on the availability of useful syntactic or semantic information.

Fodor and Inoue's theory is considered to be unique in the following respects: First, unlike most of the previous parsing models, Fodor & Inoue consider the type of necessary structural alteration *not* to be a major factor determining the garden path status of the sentence. Second, in their system, the human parser does *not* "reparse" when the current structure cannot accommodate the new input item. That is, rather than stopping its normal activities and entering a new mode of reanalysis, the human parser simply tries to do some repair, i.e., to continue to parse by assigning the problematic input the least ungrammatical structure at the moment, even though there is conflict between that structure and the previously computed one. And then the parser may try to change some other constituent to get rid of the conflict. The changed element may now conflict with something else in the parse tree, and so the something

else may be altered, and so on. In this case, diagnosing the error easily and immediately is crucial for the parser to do some repair. If the error is not a transparent one and hence cannot be diagnosed immediately, the parser can hardly repair and a garden path results. Also, it is crucial that the repair seems worth attempting to the parser. That is, if the repair can create more serious problems in the next stage, it will not be attempted.

Let us consider Fodor and Inoue's analysis of core data, whose processing difficulty results from complicated diagnosis procedure.

(12) #They told the boy that the girl met the story.

In (12), which is a severe garden path sentence, resolving the mismatch between *met* and *the story* is the major task for the parser. Fodor & Inoue point out that in order to repair the mistake in the first-pass parsing, in which 'that-clause' is attached as a complement of *told*, the parser must be able to deduce the following at once: (i) the object of *met* could be an empty category such as a WH-trace, (ii) 'that-clause' may be part of the indirect object of *told* rather than its complement clause. Fodor and Inoue claim that such lateral thinking is beyond the capacity of the parser, given that it cannot reach the correct analysis without a parsing breakdown.

Fodor and Inoue also account for the contrast between the following sentences in terms of the difficulty of diagnosing the initial misanalysis.

(13) #The horse raced past the barn fell.

(14) They knew the girl at the bakeshop was hungry.

The parser's major task in computing the ultimate structure of (13) and (14) is to find a suitable attachment site for the second verb. To put it differently, the NP which can serve as the subject of the second verb phrase should be sought immediately. Fodor and Inoue claim that the reason why (13) leads to a parsing breakdown is that it is extremely difficult for the parser to reason that the problem arose from the (mis)analysis of *raced* rather than *fell*. In fact, attaching *raced* as the predicate for *the horse* is not problematic at all at the moment, and there seems to be no reason to reanalyze the VP led by *raced*, except for the fact that the second VP *fell* is without an appropriate attachment site.

Fodor and Inoue argue that since there is not enough motivation for reanalyzing the VP led by *raced,* it is not easy at all for the parser to construct the ultimate, correct structure, in which *fell* is the matrix verb and *raced past the barn* is a non-finite VP, modifying *the horse.*

The garden path status of (13) is contrasted with that of (14), where repairing the initial computation of the argument structure of *knew* is not problematic for the parser. According to Fodor and Inoue, even though the second verb *was* appears to be subjectless at the moment since *the girl at the bakeshop* is initially analyzed as the object of *knew,* repair in this case is rather easy due to the fact that *was* can STEAL the preceding noun phrase to make up a full clause.[3] Moreover, that full clause can function as the object of *knew* and hence the overall structure doesn't lead to any problem.

Notice that the above repair procedure is closely tied to the possibility of reaccessing the lexicon. That is, in order to STEAL *the girl at the bakeshop* to construct a full clause, it is necessary to check the lexicon for confirming that *knew* can take a clausal complement. Fodor & Inoue argue that lexical reaccess is not necessarily associated with high cost, and it can be as costless as the initial lexical access during the first-pass processing if the parser can be confident of the necessity of prescribing it.

4. Predicting Garden Path in a Diagnosis Model and a Deterministic Model

As discussed earlier, Fodor and Inoue propose that whether recovering from the initial misanalysis creates conscious processing difficulty is tied to the cost of diagnosing the error, rather than to the cost associated with the type of structural alteration necessary for deriving the ultimate structure. Crucially, they regard the traditional 'reanalysis' as a simple 'repairing procedure'. That is, for Fodor and Inoue, reanalysis is not to reparse the problematic input, but to continue to parse by computing the least ungrammatical structure for the problematic input, even though

3) 'STEAL NP' is the strategy employed in Frazer and Rayner (1987), who borrowed the term from Abney (1986). Even though such a strategy is useful in accounting for the lack of processing difficulty in (14), it has the limitation since it fails to account for the garden path effect of the sentences such as (16), which is almost identical to (14) in terms of the availability of an NP for its subjectless verb.

there is conflict between that structure and the previously built one. The parser then may try to change some other constituent to eliminate the conflict, and such a try may continue until there is no conflict.

While Fodor & Inoue's diagnosis model offers reasonable explanations for the garden path status of a wide range of data, there seems to be lack of principled account for some crucial types of data. We will examine those data and consider alternative analyses for them.

4.1. Problematic and Unproblematic Reanalyses in a Diagnosis Model

Consider the following pair again.

(13) #The horse raced past the barn fell.

(14) They knew the girl at the bakeshop was hungry.

As discussed earlier, Fodor & Inoue claim that the contrast between (13) and (14) in terms of their garden path status is rather expected, given that the cost of diagnosing the necessary repair is quite different between the two cases: In (13), searching for the subject NP for the second verb *fell* is a costly task in the sense that even though *the horse* is a possible candidate for the subject of *fell*, such an analysis creates a new problem, i.e., the first verb phrase *raced past the barn* sitting unattached. At that moment, the parser is not able to see what structure should be assigned to the *raced* VP. Consequently, the parser cannot but be reluctant to try such an analysis and hence repairing can hardly begin. Meanwhile, in (14), seeking the subject NP for the second predicate *was hungry* is not a difficult task, since the preceding NP *the girl at the bakeshop* can easily be "stolen" and posited as the subject of *was hungry*. Notice that, as pointed out by Fodor & Inoue, such repair doesn't lead to any new problem: Even though *knew* momentarily loses its object, such a gap is immediately filled by constructing a new complement clause, *the girl at the bakeshop was hungry*.

Notice that Fodor & Inoue's explanation on the contrast between (13) and (14) is crucially based on the possibility of reaccessing the lexicon. That is, in (14), reaccessing the lexicon in order to check the subcategorization framework of *knew* is well motivated since that process is directly related to computing a legitimate clausal structure. Meanwhile,

according to Fodor & Inoue, the parser is not likely to reaccess the lexicon in (13); the parser is not confident of the necessity of checking the subcategorization framework of *raced*, since reanalyzing the *raced* VP doesn't look promising to the parser in the following respects: The major task being to find the subject for *fell*, the parser may try a simple and safe analysis, in which *the barn* is "stolen" and posited as the subject of *fell*. Such an analysis, however, is not acceptable since there is no coordinate conjunction between *the horse raced past* and *the barn fell*. The parser may try another analysis, in which *fell* as well as *raced* takes *the horse* as its subject. That analysis should also be abandoned ultimately since the sentence lacks a coordinate conjunction before *fell*. By the time those analyses fail, the parser may be stuck and not even have a chance to consider the option of going back to the lexicon and checking the subcategorization framework of *raced*.

4.2. Problems with the Diagnosis Model

While Fodor and Inoue's theory can account for the difference between (13) and (14), it doesn't appear to be able to deal with a very similar problem. Consider the following example.

(15) #The horse raced fell.

It doesn't seem obvious how the conscious processing difficulty observed in (15) can be handled within Fodor & Inoue's framework. Since (15) consists of only one noun phrase and two verbs, the parser doesn't really come up with many possibilities to consider. The two possible analyses considered in (13), for instance, don't seem relevant in this case; the first possibility of conjoining two clauses is not applicable here, and the second possibility doesn't seem realistic in the sense that the lack of a coordinate conjunction looks so clear in this short sentence. Then what can the parser do when it encounters the subjectless verb *fell*? The immediate possibility is, of course, to associate *fell* with *the horse*. In that process, the parser will check the lexicon and see the option of using *raced* as a modifier, i.e., the past participle of a transitive verb. Given that, as Fodor & Inoue claim, lexical reaccess can be either easy or difficult depending on whether the need for it is determinable, it is mysterious why (15) leads to conscious processing difficulty; the necessity

of lexical reaccess is quite obvious in this case and hence lexical reaccess should be easy and then the reanalysis shouldn't be problematic.

In addition to the above case, other standard garden path sentences such as (16) seem to be an embarrassment for the diagnosis model.

(16) #While Mary was mending the sock fell.

(16) is analogous to (13) and (15) in the respect that the "symptom" is a subjectless verb.[4] In (16), after *the sock* is attached as the complement of *was mending*, there seems to be no subject for the verb *fell* at the moment. According to Fodor & Inoue's theory, the necessary reanalysis in this case should not be costly, since *the sock* can easily be "stolen" and posited as the subject for *fell*. Also, checking the subcategorization framework of *mending* in that process is well motivated, since the necessity of lexical reaccess for confirming that *mend* can be used as an intransitive verb seems quite obvious. Consequently, (16) is predicted not to create conscious processing difficulty within Fodor & Inoue's framework. Contrary to such a prediction, (16) is a severe garden path sentence.

One may conjecture that the processing difficulty observed in (16) can be attributed to the lexical property of *mending*. That is, since *mend* seems to be used as a transitive verb more frequently than as an intransitive verb, the parser may be obstinate in trying the above repair. Such a conjecture, however, doesn't seem to be on the right track, given the following data.

(17) #Since Jay always jogs a mile seems like a very short distance to him.

(17) patterns with (16), and the "symptom" here is the subjectless verb *seems*. Note that *jog* is predominantly used as an intransitive verb, and hence (17) can be regarded as a marked structure. The fact that (17) also creates conscious processing difficulty strongly suggests that even when the verb in question is predominantly an intransitive verb, the noun phrase following it tends to be attached as its complement,[5] and

4) The "symptom" is the first word of an input string which can't be attached properly into the current parse tree. It signals that a mistake has been made in the first-pass processing.

reanalyzing the complement as the subject of the upcoming clause is costly.

Our observation thus far suggests that whether the verb is mainly used as a transitive verb or an intransitive verb is not a primary factor affecting the reanalysis procedure.[6] Also, the repair mechanism employing the notion of lexical reaccess doesn't seem to be able to handle some core examples such as (16) and (17). Meanwhile, the contrast between (14) and (17) is accounted for straightforwardly by a deterministic parser; even though the NP preceding the subjectless verb may be stolen to fill the subject position in both (14) and (17), the reanalysis in the latter is predicted to be costly, since the dominance relation between the noun phrase to be stolen and the verb preceding it cannot but be retracted during the reanalysis.

4.3. Some Arguments against the Rebuilding Mechanism in Deterministic Parsing

As for the criticism on deterministic parsing, most of them can be considered to be an argument that there exist garden path sentences which cannot be attributed to the violation of determinism. Such a line of argument, however, is not really persuasive, since the parsing theories based on determinism do *not* assume that every garden path results from violating determinism. In fact, the theories based on determinism admit that certain garden paths can be accounted for by some principles or conditions other than determinism, and different deterministic models indeed postulate additional conditions of their own to deal with such garden paths. For instance, if two sentences with identical syntactic structure show different degrees of processing difficulty, then it is likely that lexical idiosyncrasies play a role in determining their garden path status. In this case, different parsing models employ different devices for dealing with the lexical factors responsible for the garden path.

---

5) That is a typical case reflecting the standard parsing strategies known as Minimal Attachment and Late Closure originating from Frazier (1978).

6) Of course, we cannot totally ignore lexical idiosyncracies when we consider the parser's structural commitment. (See Bresnan, Ford, & Kaplan, 1982 for the cases where lexical idiosyncracies such as the frequency of the usage of the verbs influence parsing preferences.) Nevertheless, as pointed out by many researchers, it is the structural configuration which mainly influences and guides the parser's commitment.

Given this, the valid criticism on determinism should be able to show that there are cases where violation of determinism does not lead to conscious processing difficulty. In Fodor and Inoue's work, we have found one example belonging to such a category. Consider the following.

(18) Our son doesn't tell us what he's thinking about any more.

In this example, *any more* should ultimately be associated with *doesn't* in the matrix clause since it is a negative polarity item. In the first-pass processing, however, *any more* might be attached as an adjunct inside the complement clause, due to the operational principles such as Late Closure or Right Association. Then, the reanalysis in this case cannot but involve raising *any more* from the complement clause to the matrix clause in the parse tree. Since such raising amounts to retracting the dominance relation between the complement clause IP node and *any more*, the reanalysis here is predicted to be problematic within the deterministic framework. Nevertheless, (18) hardly causes conscious processing difficulty.

We admit that (18) can be a powerful argument against deterministic parsing. Nevertheless, we want to point out that there might be an explanation on the peculiarity of (18) within the deterministic framework. Recall that *any more* is a negative polarity item, and hence it is licensed only when it is linked to the negative particle *not*. Provided that a negative polarity item is licensed via syntactic feature checking and syntactic features must be considered for the structural commitment of the parser,[7] there seems to be a fair possibility that *any more* is attached as an adjunct of the matrix clause in the first-pass processing. Then, there will not be any reanalysis, and hence (18) will not be a problem for a deterministic model any more.

To recapitulate, if parsing (18) involves reanalysis with regard to the attachment of *any more*, then a deterministic model will incorrectly predict a strict garden path. But we cannot exclude the possibility that parsing (18) does not involve any reanalysis.

---

7) This is a basic assumption for most principle-based parsing models. See Crocker (1992) and Weinberg (1994), among others, for details.

4.4. Summary

We have shown that whereas the diagnosis model may be able to deal with a reasonable range of data on garden path, it is sometimes difficult to justify its repair mechanism. In particular, it is not obvious at all why lexical reaccess is not possible in certain cases and hence the diagnosis of the error in those cases fails. Notice that this is directly contrasted with the rebuilding mechanism of a deterministic parser, where most conscious garden paths are correctly predicted by considering the structural configurations in the reanalysis procedures. Provided that the diagnosis model and the deterministic model show comparable empirical coverage, it is natural to consider which model is operated by more transparent and conceptually adequate strategies. In this respect, the deterministic model can be regarded as more desirable one.

5. Concluding Remarks

No one believes that a single theory can deal with every garden path sentence satisfactorily. Also, it seems true that the degree of processing difficulty can vary across the sentences even though those sentences have very similar syntactic structure. In order to provide principled and satisfactory accounts for those complicated data, it may be necessary to rely on more than one parsing theory. For instance, processing difficulty may sometimes be accounted for in terms of the frequency of usage, as in Bresnan, Ford, and Kaplan (1982), and sometimes it may be attributed to the difficulty in the diagnosis of misanalysis, as proposed by Fodor and Inoue (1994, 2000). However, despite all of the above complications and variations, there must a central framework by which we can draw a line overall between problematic and unproblematic reanalyses. Such a framework, of course, must be equipped with conceptually appealing principles and able to deal with a wide range of data. We have seen that a constrained form of determinism, i.e., structural determinism, can provide such a framework. The repair mechanism of the diagnosis model doesn't seem quite adequate for playing such a central role.

References

Abney, S. (1987). Licensing and parsing, *Proceedings of NELS* 17. University of Massachusetts, Amherst.

Bresnan, J., Ford, M., and R. Kaplan. (1982). A competence-based theory of syntactic closure, In J. Bresnan, ed., *The Mental Representation of Grammatical Relations*. MIT Press, Cambridge.

Crocker, M. (1992). *A Logical Model of Competence and Performance in the Human Sentence Processor*. Unpublished doctoral dissertation. University of Edinburgh.

Fodor, J., and A. Inoue. (2000). Garden path repair: Diagnosis and triage. *Language and Speech* 43(3), 261-271.

Fodor, J., and Inoue, A. (1994). The diagnosis and cure of garden paths. *Journal of Psycholinguistic Research* 23(5), 407-434.

Frazier, L. (1978). *On Comprehending Sentences: Syntactic Parsing Strategies*. Unpublished doctoral dissertation. University of Connecticut, Storrs.

Frazier, L., and K. Rayner. (1987). Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences. *Journal of Memory and Language* 26, 505-526.

Gibson, E., and N. Perlmutter. (2000). Distinguishing serial and parallel parsing. *Journal of Psycholinguistic Research* 29(2), 231-240.

Gorrell, P. (1995). *Syntax and Parsing*. Cambridge University Press.

Gorrell, P. (1994). Japanese trees and the garden path. In R. Mazuka and N. Nagai, eds., *Japanese Syntactic Processing*. Lawrence Erlbaum Associates.

Gorrell, P. (1987). *Studies in Human Syntactic Processing: Ranked-Parallel vs. Serial Models*. Unpublished doctoral dissertation. University of Connecticut, Storrs.

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

Kurtzman, H. (1985). *Studies in Syntactic Ambiguity Resolution*. Unpublished doctoral dissertation. MIT, Cambridge.

Marcus, M. (1980). *A Theory of Syntactic Recognition for Natural Language*. MIT Press, Cambridge.

Marcus, M., Hindle, D., and M. Fleck. (1983). D-theory: Talking about talking about trees. *Association for Computational Linguistics* 21,

129-136.
Weinberg, A. (1994). Licensing constraints and the theory of language processing. In R. Mazuka and N. Nagai, eds., *Japanese Syntactic Processing*. Lawrence Erlbaum Associates.

Sungki Suh
Department of English Language and Literature
The Catholic University of Korea
Yeokgok-dong, Wonmi-gu
Bucheon 422-743, Korea
E-mail: sksuh@catholic.ac.kr

Received: Sept. 1, 2002
Revised version received: Nov. 23, 2002
Accepted: Nov. 27, 2002