GRAMMATICALITY JUDGMENTS

Why Does Anyone Object to Subject Extraction?

Jacquelyn Schachter
Virginia Yip
University of Southern California

Grammaticality judgments reflect a compound product of both grammatical and processing factors. But because they interact in a symbiotic way, very often grammatical and processing constraints are difficult to separate. According to generally accepted grammatical theory, (a) Who do you think John told Mary he fell in love with? and (b) Who do you think John told Mary fell in love with Sue? are equally grammatical. We have observed, however, that native speakers strongly accept sentences like (a) as grammatical but react quite variably to sentences like (b). A possible explanation is that native English speakers exhibit a processing preference, in searching for the extraction site for the wh-word, for object position over subject position. Proficient nonnative judgmental data offer additional support for a processing account. Nonnatives whose L1 grammars do not bias them toward objects also show preferences similar to those of natives. We provide a processing account based on Frazier's Minimal Attachment principle.

A research question that has generated much interdisciplinary research in linguistics concerns the relationship between the grammar and the parser. In this article, we report on an experiment that compared native and nonnative judgments on certain English structures involving a close interaction between grammatical and processing

We are grateful to the many students who took part in this project, and to Stephen Matthews whose perceptive comments during the course of this experiment were invaluable.
effects. We have found differential responses to wh- extractions from object and subject positions (both without the complementizer), as in examples (1) and (2):

(1) Which book did you say John believes Bill should read? (t = object trace)
(2) Which book did you say John believes t offended many people? (t = subject trace)

Both the object and subject extractions should be grammatical since no grammatical constraints are violated. However, we found a consistent preference for wh- movement from the object position over the subject position across native speakers and nonnative speakers alike. Subjects seemed to have a hard time assigning the correct grammatical analysis to the subject extractions, even though both subject and object extraction sentences were constructed in such a way as to allow only one grammatical analysis: there is only one place to put the gap corresponding to the wh- phrase. We attribute the difficulty involved in subject extraction to the nature of left-right processing and potential on-line garden path effects\(^1\) caused by such sentences.

**GRAMMAR AND PARSER**

A central issue that lies at the interface of linguistics and natural language processing concerns how grammar and parser are related. An assumption widely adopted by current theories of language processing is that in spite of the close relationship between the grammar and the parser, they are nonetheless separate and distinct entities (Frazier & Fodor, 1978; Pritchett, 1988). While it is uncontroversial that grammar and parser must bear some kind of relationship since they have to assign compatible structures, it is not clear exactly how they are related.

One major approach to the characterization of this relationship is to try to come up with processing principles that are independently needed for general human cognitive functions, to the effect that what are considered “grammatical” constraints should derive from such processing principles. The idea is that the grammar is the way it is because of the way the parser operates. Marcus (1980) and Berwick and Weinberg (1984) attempted to derive grammatical constraints such as subjacency and c-command from parsing principles.

Following the mainstream position in generative grammar, we assume, nevertheless, that there are certain grammatical constraints that cannot be reduced to parsing principles, as well as parsing effects that cannot be reduced to grammatical principles. Miller and Chomsky's (1963) characterization of multiple self-embedded sentences being grammatical but unprocessable remains a classic example of this position.

Assuming that both the grammar and the parser are independently motivated, we take the position that the effects of each are in principle separable, even though in practice teasing these effects apart is often difficult. In this article we present a case in which the effects of the grammar are controlled for, allowing a processing effect to be isolated.
JUDGMENT DATA IN L2 RESEARCH

Evidence from metalinguistic judgment tasks provides an important source of data, in particular, for research concerned with the relevance of Universal Grammar (UG) to second language acquisition. For example, Bley-Vroman, Felix, and loup (1988), Coppieters (1987), Felix (1988), Ritchie (1978), Schachter (1989b), and White (1985, 1989) made use of such data to support their claims about the role of UG in constructing the second language grammar. Despite the differences in their methodological orientations, each assumes that such metalinguistic tasks are a means of tapping into a speaker's linguistic competence. The argumentation of this line of research essentially goes like this: if L2 learners are able to judge correctly ungrammatical sentences ruled out by universal constraints just on the basis of positive evidence from the target language, then they have access to UG.

NATURE OF GRAMMATICALITY JUDGMENT TASKS

While grammaticality judgments are generally assumed to be a means of tapping competence, they in fact reflect a complex interaction between competence and performance variables inherent in the judgment task. One's judgment is essentially a product of both grammatical and extragrammatical factors. A sentence may be judged to be bad due to a variety of extragrammatical factors. It may be semantically deviant or pragmatically odd, or may contain socially nonstandard forms; care must be taken to prevent these from influencing the judgments. One set of extragrammatical factors that is both particularly interesting and particularly vexing in dealing with grammaticality judgments is the set of factors arising out of processing constraints and preferences—interesting in that it is a set that interacts systematically with grammatical factors, and vexing in that it is the set most often difficult to tease apart from grammatical factors. Since processing and grammatical factors exert an effect on one's judgment simultaneously, it is often difficult, but we believe not impossible, to separate out the different effects on a principled basis. Given any set of judgmental data, one should be aware of the possible misinterpretations that can arise if processing factors are not considered. Our findings point to the need to take into consideration this complex interplay of factors in interpreting the data.

GRAMMAR AND PARSING IN L2 RESEARCH

One example of such misinterpretation is Ritchie’s (1978) characterization of the results of testing Japanese speakers' knowledge of the Right Roof Constraint, a putative universal constraining rightward movement. Sentences involving extraposition such as the following were used in Ritchie's experiment:

(3) That it was surprising was obvious / that Mary had built a boat.
(4) That it was disturbing was clear / that Bill had sent a letter.
Ritchie's subjects consistently rejected these sentences, and since Japanese does not have rightward movement, and thus the effect cannot be due to language transfer, Ritchie attributed these judgments to the Right Roof Constraint of Universal Grammar, which, he claimed, was still available to adult L2 learners. Schachter (1989a), however, called into question Ritchie’s conclusion—that his Japanese subjects’ consistency in rejecting these sentences (and others like them) was due to the fact that they still had the Right Roof Constraint as part of their grammars. She pointed out that these sentences are in fact garden path sentences—sentences that induce an initially incorrect parse which then has to be abandoned and reanalyzed. (See Schachter, 1989a, for a detailed analysis.) The fact that Japanese subjects behave consistently in rejecting these garden path sentences cannot therefore be taken as evidence that the subjects have the Right Roof Constraint.

What this phenomenon reveals is rather significant: when both a grammatical principle such as the Right Roof Constraint and a processing effect such as those proposed for garden path sentences (cf. Frazier, 1978; Pritchett, 1988) predict essentially the same results, that is, that both ungrammatical and unprocessable sentences will be rejected in a judgment task, it becomes virtually impossible to tease apart the two effects. Given an input string like the above extraposition sentences, the parser suffers from a breakdown before a grammatical analysis can be assigned.

The argumentation of the present study parallels that of the case just described, but the facts, as we shall see, are different. In the Ritchie (1978) study, with regard to the extraposed sentences, both the grammar and the parser ruled out the sentences in question, and their separate effects cannot be isolated. In our study, we have a case where the grammar says “yes” and the parser says “no,” or “yes” only with difficulty. Our subjects' difficulty in judging subject extractions, which are predicted to be just as grammatical as object extractions, was attributed to the inherent processing load imposed by the subject extractions over and above that presented by the object extractions.

**BACKGROUND OF THE PRESENT STUDY**

Schachter (1989b) observed a rather surprising phenomenon regarding judgments on sentences in which a *wh = subject* is extracted (but not when a *wh = object* is extracted). She reported that both native and nonnative speakers found the subject extraction sentences difficult to judge, often evaluating them as “ungrammatical.”

(5) Who do you think Bill said Mary expected *t* to go to the dance with Mark? (*t* = subject trace; four clauses)

The results were unexpected in the sense that sentences such as this are predicted to be as grammatical as *wh = object* sentences such as

(6) What did the policeman say he thought the thieves would prefer to steal *t* first? (*t* = object trace; four clauses)
Interestingly, not only native speakers of English found these subject extractions difficult; nonnatives behaved just like natives across different language groups, namely, Chinese, Korean, and Indonesian.

The present study is inspired by this subject/object asymmetry, which was observed in both first and second language judgmental data. We pursue an explanation in processing terms, which might account for the difficulty present in subject extractions but absent in object extractions.

**METHODOLOGY**

A questionnaire was designed to test judgments of subject and object *wh*- extractions. All test sentences were constructed so as to be unambiguous, allowing only one grammatical analysis; they contain only one place in which to put the gap corresponding to the *wh*-phrase. Furthermore, we used verbs with a simple argument structure, which allow only one argument. Equivalent sentences with the complementizer *that* were included for later investigation of “that-trace” effects in L1 and L2.

Four types of test sentences were included in the questionnaire: (1) *wh* = object/without “that,” (2) *wh* = object/with “that,” (3) *wh* = subject/without “that,” and (4) *wh* = subject/with “that.” In this article we focus only on types 1 and 3, the sentences of which are listed in the Appendix.

There are three examples of each type at the four-clause, three-clause, and two-clause levels, respectively. The total number of test sentences amounted to 36, with 18 relevant for this article. In addition, 18 distractor sentences involving sentential subjects and rightward movement were included. A total of 54 sentences was randomized to produce the test.

The subjects for the experiment were undergraduates enrolled in freshman writing and elementary linguistics courses at a major West Coast university. All nonnative subjects had passed the English proficiency exams that the university required and were full-time students in English, their nonnative language. The number of subjects was 20 for each of three groups of subjects, namely native English speakers, and proficient speakers of English whose native languages were Chinese and Korean. As is standard in such judgment tasks, subjects were provided with comparable but unrelated example sentences illustrating four possible degrees of grammaticality; they were instructed to ignore possible extraneous problems such as lexical choice, semantic plausibility, and punctuation. Subjects were told that there were no absolute “right” or “wrong” answers for these sentences, and they should feel free to be “uncertain” about any of them. They were given four options according to the following scale: a, clearly grammatical; b, probably grammatical; c, probably ungrammatical; and d, clearly ungrammatical.

**RESULTS AND DISCUSSION**

Table 1 shows the raw scores indicating the number of times options a, b, c, and d were chosen for each sentence type. The a and d options are indicators of certainty
Table 1. Raw score option choices for wh=object and wh=subject

<table>
<thead>
<tr>
<th></th>
<th>2 Clauses</th>
<th>3 Clauses</th>
<th>4 Clauses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a*</td>
<td>42</td>
<td>32</td>
<td>23</td>
</tr>
<tr>
<td>b</td>
<td>13</td>
<td>21</td>
<td>16</td>
</tr>
<tr>
<td>c</td>
<td>4</td>
<td>4</td>
<td>17</td>
</tr>
<tr>
<td>d</td>
<td>1</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Chinese</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>33</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>b</td>
<td>18</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>c</td>
<td>2</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>d</td>
<td>7</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Korean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>34</td>
<td>23</td>
<td>20</td>
</tr>
<tr>
<td>b</td>
<td>12</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>c</td>
<td>7</td>
<td>11</td>
<td>19</td>
</tr>
<tr>
<td>d</td>
<td>7</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

*a, b, c, d represent the following answers: a = definitely grammatical, b = probably grammatical, c = probably ungrammatical, d = definitely ungrammatical.

on the part of the subject, with the b and c options indicating some degree of uncertainty. What we found, in general, was a shift from a to b to c answers paralleling the shift from two- to three- to four-clause sentences, indicating increasing ambivalence or uncertainty as the wh- word became located further and further from its trace.3

The mean scores were arrived at by giving scores of 3 for a judgment of a (the sentence was clearly grammatical), 2 for a judgment of b (probably grammatical), 1 for a judgment of c (probably ungrammatical), and 0 for a judgment of d (clearly ungrammatical). Contributing to each cell were the summed scores from three sentences for each individual, so that a judgment of a on three sentences would result in a score of 9, a judgment of b on the same three sentences would result in a score of 6, and so forth. The means and standard deviations for each extraction type and for extraction over two, three, and four clauses are listed in Table 2.

Consider the native speaker scores and note the effect of the number of clauses over which the wh- word has been moved. They range from means of 7.8 for one-clause movement to 5.6 for three-clause movement for objects, and from 6.5 to 4.0 for subjects. This is a large effect, and a rather surprising one (but totally consistent with Schachter’s, 1989b, earlier study). The most obvious explanation might be that the longer the sentences, the less acceptable they become, but this does not seem a reasonable explanation for native speakers. These sentences are in fact not particularly long. We argue, rather, that it is the fact of extraction itself that produces this effect, specifically the extra effort it takes to link up the wh- word with its extraction
Table 2. Overall means for wh=object and wh=subject extraction

<table>
<thead>
<tr>
<th></th>
<th>Object</th>
<th></th>
<th>Subject</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Native</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 clause</td>
<td>7.8</td>
<td>1.4</td>
<td>6.5</td>
<td>1.8</td>
</tr>
<tr>
<td>3 clause</td>
<td>7.1</td>
<td>1.5</td>
<td>5.8</td>
<td>2.7</td>
</tr>
<tr>
<td>4 clause</td>
<td>5.6</td>
<td>2.0</td>
<td>4.0</td>
<td>2.6</td>
</tr>
<tr>
<td>Chinese</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 clause</td>
<td>6.9</td>
<td>2.1</td>
<td>5.1</td>
<td>2.2</td>
</tr>
<tr>
<td>3 clause</td>
<td>5.7</td>
<td>1.7</td>
<td>4.6</td>
<td>1.5</td>
</tr>
<tr>
<td>4 clause</td>
<td>5.1</td>
<td>1.5</td>
<td>2.8</td>
<td>1.7</td>
</tr>
<tr>
<td>Korean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 clause</td>
<td>6.6</td>
<td>1.9</td>
<td>4.8</td>
<td>1.8</td>
</tr>
<tr>
<td>3 clause</td>
<td>6.0</td>
<td>1.8</td>
<td>3.9</td>
<td>1.5</td>
</tr>
<tr>
<td>4 clause</td>
<td>5.4</td>
<td>1.7</td>
<td>3.7</td>
<td>1.9</td>
</tr>
</tbody>
</table>

site when there are more clauses involved, even with simple verbs of the type we were careful to use. For four-clause sentences, for example, subjects must keep the *wh-* word in some kind of buffer until they discover the clause in which it fits and the grammatical role it plays in that clause.

Note next that the same phenomenon occurred in the nonnative data, though the nonnative scores are generally less extreme than the native data. Nevertheless, the same tendency exists here.

It seems clear that we would not want to account for any of these differences on the basis of grammatical explanations per se. Processing effects are clear and strong here.

More interesting, from our point of view, is the considerable difference in degree of acceptance of *wh* = object extraction over *wh* = subject extraction in both native and nonnative data. The grammar of English clearly licenses both types of extraction. At first glance, one might assume that this is simply one more example of the many types of subject/object asymmetries discussed in the literature on English (for some discussion of such asymmetries, see van Riemsdijk & Williams, 1986). Those asymmetries, however, have all involved grammatical/ungrammatical pairs (object and subject *wh* word extraction with complementizers, for example). The subject/object asymmetry in the present case clearly does not involve differences in grammaticality.

Furthermore, number of clauses should not be a factor here. Logically, holding a *wh*- word in short-term memory should be no more difficult for those that turn out to be subjects in their respective clauses than for those that turn out to be objects. This difference seems to involve a processing effect that cannot be simply attributed to the length of the sentences.

In order to look at this effect more closely, we performed an analysis of covariance, using the number of clauses over which the *wh*- word was moved as the
covariate. In this way the variation between extraction from subject and object sites would not be confounded with movement over number of clauses. Then, while controlling for depth of embedding (i.e., movement over one, two, or three clauses), we would be able to ascertain whether the difference between grammaticality scores on \( wh = \) object extraction and \( wh = \) subject extraction was statistically significant (see Kerlinger & Pedhazur, 1973, for a discussion of this design). The model to be tested, then, was that the scores could be accounted for by structure type (i.e., subject or object extraction), controlling for depth of embedding.

The overall model for each group (including structure types 2 and 4, not dealt with in this article) is significant. For the English group, \( F(1, 38) = 42.3, p < .0001 \). Native speakers of English clearly treat grammatical extraction from object and subject positions differently. So too do the nonnatives, although the results are not quite as striking due to their general tendency to mark both sentence types as ungrammatical. For the Chinese, \( F(1, 38) = 17.0, p < .0001 \), and for the Koreans, \( F(1, 38) = 16.0, p < .0001 \). The least square means for each group are listed in Table 3 along with the probabilities for differences between extraction sites. In each case, there is a significant preference for object extraction over subject extraction.

We argue that the clear difficulties involved in subject extraction are attributable to the nature of left-right processing and the potential garden path effects caused by subject extractions.

Looking only at the behavior of native speakers of English, our explanation is perhaps plausible but not entirely convincing. Strong supporting evidence for our position comes from the data produced by the Korean and Chinese speakers of English. Their preferences are especially valuable in that both Korean and Chinese are languages that do not have the English type of \( wh- \) question with object or subject extraction. Since there is no \( wh- \) extraction, there does not seem to be anything in either language that would bias speakers of these languages toward objects. The L1 grammar of these speakers of English cannot be a source of bias toward object extraction: at best, one would expect neutrality between subjects and objects. But this does not occur, providing further support for our view that we are dealing with processing effects in all cases.

Recently, Jordens (1989) reported similar findings with Dutch learners of English, who showed the same preference for object extractions, even though the grammar of Dutch allows both object and subject extractions; Jordens independently came up with a processing account for this subject/object asymmetry.

The major explanandum, then, is the pervasive preference for object extraction over subject extraction. While subject/object asymmetries are familiar phenomena in grammar, syntactic theory as currently formulated offers no account of this contrast. In Government/Binding terms, the subject extractions are grammatical because no grammatical principle is violated. The trace left by the extraction of the subject \( wh- \) word in these sentences is properly governed, as required by the Empty Category Principle, which states that all empty categories must be governed (Chomsky, 1981). Object empty categories are lexically governed by the verb: \( take \) in (7). Subject empty categories lack a lexical governor and can only be properly governed by an anteced-
Table 3. GLM procedure—Least square means

<table>
<thead>
<tr>
<th>Language</th>
<th>Structure</th>
<th>Score LS</th>
<th>Student Error</th>
<th>Probability*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>LS Mean</td>
<td>LS Mean</td>
<td></td>
</tr>
<tr>
<td>Native</td>
<td>Object</td>
<td>6.8</td>
<td>0.25</td>
<td>0.00001</td>
</tr>
<tr>
<td></td>
<td>Subject</td>
<td>5.4</td>
<td>0.25</td>
<td>0.00001</td>
</tr>
<tr>
<td>Chinese</td>
<td>Object</td>
<td>5.9</td>
<td>0.26</td>
<td>0.00001</td>
</tr>
<tr>
<td></td>
<td>Subject</td>
<td>4.2</td>
<td>0.26</td>
<td>0.00001</td>
</tr>
<tr>
<td>Korean</td>
<td>Object</td>
<td>6.0</td>
<td>0.23</td>
<td>0.00001</td>
</tr>
<tr>
<td></td>
<td>Subject</td>
<td>4.1</td>
<td>0.23</td>
<td>0.00001</td>
</tr>
</tbody>
</table>

*Indicates the probability that the responses to object extraction differ significantly from the responses to subject extraction.

ent in COMP: the *wh*-phrase in the matrix clause in (8). This is known as antecedent government and is permitted by the grammar.

(7) What did the nurse say she reported the patient had taken *t*?
(8) What did the nurse say she reported *t* had happened to the patient?

Intuitively, however, the *wh*-subject extraction sentence seems to require a little more effort to process. The task involved in this case consists in figuring out the respective arguments of the verbs, that is, the complement of *report* and the subject of *happened* in (8). Meanwhile, the *wh*-word is waiting to be related to a predicate; there is some tendency to interpret it on-line as the object of *report*, thereby completing the parse—an assignment that must be undone when the next clause is reached.

We now turn to standard accounts of parsing of dependencies to see how these intuitions can be formalized. Following Frazier and Fodor (1978), we assume that as processing of input proceeds from left to right, each new item is incorporated into a structural representation in accordance with a principle of Minimal Attachment:

Attach each new item into the current phrase marker postulating only as many syntactic phrase nodes as is required by the grammar. (Frazier, 1987, p. 520)

The application of this principle is illustrated in (9):

(9) Minimal Attachment analysis
Initially, Minimal Attachment posits an empty category as the object of report. When the sentence continues, the verb had happened sets up an argument structure that calls for a subject; the parser then has to backtrack to posit a subject empty category. Assuming that any backtracking demands an extra processing effort, we have an account of the dispreference for subject extractions.

This initial Minimal Attachment analysis has to be revised in order to incorporate the remaining input string into a grammatical analysis; this entails a Nonminimal Attachment analysis, as illustrated in (10):

(10) Nonminimal Attachment analysis

While the original mis-parse is not a salient reading—presumably because the ambiguity is such a short-lived one—it is clearly available in principle in all our test sentences. In fact, such temporary local ambiguity seems to be inherent in just about any verb one can think of in the same contexts. The reason is that any verb that subcategorizes for a sentential complement also subcategorizes for an NP object:

report: [-S'] e.g., report that time is up
[-NP] e.g., report the latest news
suggest: [-S'] e.g., suggest that it works
[-NP] e.g., suggest a solution

This kind of on-line garden path effect for subject extractions potentially arises in object extractions as well. However, the crucial difference is that no backtracking is called for in parsing object extractions. Consider the following sequence of events: Assume that in parsing (7), repeated here,

(7) What did the nurse say she reported the patient had taken t?

the parser opts for a Minimal Attachment analysis, and the structural representation initially constructed is no different from that in (9), the one for subject extraction:
(9) Minimal Attachment analysis

When the sentence continues with *the patient*, a lexical NP rather than a trace, the parser has to revise the analysis and embed the NP inside a sentence:

(9')

As parsing proceeds, the verb *had taken* sets up an argument structure that calls for an object, and the object trace naturally falls in the postverbal position. Note that the only extra work involved is embedding the NP node inside an added S node, which is a Nonminimal Attachment analysis. Processing of the sentence goes on from left to right without any backtracking.

**CONCLUSION**

The subject/object asymmetry observed in this study as well as in Schachter (1989b) proved to be strikingly consistent across both natives and various L2 speakers of English. This finding clearly calls for an explanation in terms of some universal principle governing extraction. In the absence of a grammatical account, we have naturally looked to processing factors. The processing difficulty we have identified
can be explained by backtracking (costly in terms of left–right on-line processing) coupled with the temporary garden path effect (which has been extensively investigated by Frazier and others).

We noted earlier that grammaticality judgments are a product of grammatical and processing factors. In many cases, it is difficult to separate these two factors in judgmental data. For example, with respect to Ritchie's (1978) experiment, which was designed to test for knowledge of a universal grammatical principle (the Right Roof Constraint), the test sentences are presumably rejected by both the grammar and the parser. A processing explanation is as valid as a grammatical one in this case. Assuming that parsing of an input string has to be completed before the correct grammatical analysis can be reached, one can argue, as do Frazier (1978) and Pritchett (1988), that when the parser suffers an initial breakdown, the grammar may be blocked from assigning a grammatical analysis—as with garden path sentences in general. The fact that Japanese subjects correctly reject the garden path sentences cannot therefore be taken as evidence for their knowledge of a universal grammatical constraint.

The issue raised, then, concerns cases where both the grammar and the parser yield the same negative results, that is, judgments of "ungrammatical" or "unacceptable" on judgment tests. In these cases it becomes impossible to attribute the results merely to grammatical factors. In view of the fact that most research concerned with the role of Universal Grammar makes critical use of judgmental data, we need to be cautious in making any claims about the operation of Universal Grammar on the basis of data that are not unequivocally attributable to the grammar.

In contrast to Ritchie's (1978) data, the data in our study allow us to single out a processing effect, which is distinguishable from the grammatical effects. The grammar is neutral with respect to subject and object extraction, and a processing effect can be isolated. Since both native and nonnative judgmental data converge, the results point to the cross-linguistic validity of the processing factors involved.

In light of our findings, we suggest that a number of earlier judgmental studies may need to be reexamined in terms of the possible effect of processing constraints, which we have shown play a crucial role in shaping judgmental data. In order to interpret judgmental data reliably, we need to enrich our understanding of the interaction of grammar and processor. We suggest that apart from pursuing pure grammatical explanations, the field of second language acquisition research should look to more sophisticated, refined processing explanations, which are equally critical in illuminating the data.

(Received 19 May 1989)

NOTES

1. The phrase "on-line garden path effects" characterizes those real-time processing problems involved when a listener or reader is led to an initial misparse due to the ambiguous structure of the utterance or sentence presented. One of the most famous examples of sentences that produce such effects is The horse raced past the barn fell. It is called "garden path" because one's initial reading is to assign an active rather than passive reading to The horse raced past the barn. Only after one tries to figure out what to do with the leftover word fell does one become aware of the reading in which raced past the barn is interpreted as a reduced
relative clause containing a passive verb; that is, the reading associated with the unambiguous sentence *The horse which was raced past the barn fell.*

2. The other two sentence types involve the presence of "that-trace" effects which involve a difference in grammaticality in subject and object extractions.

3. As is evident, subjects made liberal use of the "probably" options, a clear indication of uncertainty on their part. This phenomenon was not observed in other judgment studies, notably Bley-Vroman, Felix, and Ioup (1988). We believe the four-choice scale used in this experiment is useful in that it forces subjects to make a choice but also allows them to indicate some degree of uncertainty.

REFERENCES


APPENDIX

TEST SENTENCES—GRAMMAR/PROCESSING

Four Clauses

OBJECT EXTRACTION

1. What did the policeman say he thought the thieves agreed they would steal first?
2. Who did you say you thought the nurse reported the doctor had seen last week?
3. Which class did Sam say he thought the professor suspected he would drop this semester?

SUBJECT EXTRACTION

4. What did the policeman say he thought John agreed would prove his innocence?
5. Who did you say Bill thought Sue suspected would go to the dance with Mark?
6. Which doctor did you say you thought the nurse reported had seen the patient?

Three Clauses

OBJECT EXTRACTION

7. What did the nurse say she suspected the patient had taken?
8. Who did the President say he thinks he'll appoint as ambassador?
9. Which book did you say John believes Bill should read?

SUBJECT EXTRACTION

10. What did the nurse say she thought would happen to the patient?
11. Who did you say John suspects fell in love with Sue?
12. Which book did you say John believes offended many people?

Two Clauses

OBJECT EXTRACTION

13. What did Sarah think Lisa bought for my birthday?
14. Who did Hank suspect Lisa liked?
15. Which file did John believe Lisa had under the table?

SUBJECT EXTRACTION

16. What did Paul think entered the computer system last week?
17. Who did Hank suspect liked Lisa?
18. Which file did John believe contained the information?