Research in the theory of human sentence processing can be characterized by three styles of explanation. Researchers taking the first track have tried to motivate principles of structural preference from extralinguistic considerations like storage capacity in working memory or bounds on complexity of incremental analysis. Frazier and Rayner’s (1982) Minimal Attachment and Right Association Principles, and Gorrell’s (1995) simplicity metric, are examples of this type of theory.

The second track eschews “parsing strategies,” replacing them with a fairly complex tuning by speaker/hearers to frequency in the hearer’s linguistic environment. The difficulty of recovering an analysis of a construction in a particular case is a function of how often similar structures or thematic role arrays appear in the language as a whole. The work of Trueswell and Tanenhaus (1994), Jurafsky (1996), and MacDonald, Perlmuter, and Seidenberg (1994) exemplifies frequency- or probability-based constraint satisfaction theories.

The third track takes a more representational view and ties processing principles to independently needed restrictions derived from competence and language learning. This approach claims that the natural language faculty is extremely well designed in the sense that the same principles that govern language learning also contribute to a theory of sentence processing. This track is represented by the work of Gibson (1991), Gorrell (1995), Pritchett (1992), Phillips (1995, 1996), and myself (Weinberg 1992), who argue that processing can be seen as the rapid incremental satisfaction of grammatical constraints such as the θ-Criterion, which are needed independently to explain language learning or language variation. A variant of this approach, developed by Crain and Steedman (1985) among others, retains the grammatical source for parsing principles but locates these principles within a discourse or semantic, rather than syntactic, component.
In this chapter, I propose a model of the last type. I argue that a particular version of the Minimalist Program (Chomsky 1993; Uriagereka, this volume) both provides principles needed to explain initial human preferences for ambiguous structures and provides a theory of reanalysis, explaining when initial preferences can be revised given subsequent disconfirming data and when they lead to unrevisable garden paths. I then argue that this type of theory is to be preferred to theories motivated on extralinguistic principles.

In sections 11.1 and 11.2, I discuss the Minimalist Program for syntax upon which my parsing proposals are based. The following features distinguish this theory from its precursors:

(1) The theory is derivational, providing principles that determine how an analysis is constructed rather than filtering conditions that constrain output representations. The main derivational constraints are the so-called economy conditions (Chomsky 1993).

(2) The theory applies constraints strictly locally. Derivations are evaluated at each point in the analysis. They are optimized with respect to how well they satisfy constraints of a given item that is a candidate for integration into the structure at each point. How a proposed structure satisfies constraints imposed by the derivation as a whole is irrelevant.

(3) The theory incorporates a claim about a one-to-one mapping between precedence order and structural hierarchy or dominance that is embodied in the Linear Correspondence Axiom (Kayne 1994; Uriagereka, this volume).

In sections 11.3 and 11.4, I show how to interpret minimalist principles as a parsing algorithm. Specifically, I show that the economy conditions define a crosslinguistically attested theory of preference judgments. (2) and (3) combined distinguish cases where an initial preference can be reanalyzed from those cases where reanalysis into the appropriate structure is impossible, resulting in a garden path.

In section 11.5, I compare my model of sentence processing with Colin Phillips's. Phillips shares my view that the theory of sentence processing should be based on principles of grammatical theory. The processing principles that he invokes are based on a slightly different grammatical theory, one that he claims is identical to the theory of linguistic compe-
tence. I first discuss what I see as strengths of his theory and then look at three types of problems with his approach.

In section 11.6, I argue that this type of theory has advantages over theories relying on extralinguistic frequency or parsing strategy principles, and in section 11.7, I present conclusions.

11.1 Some Minimalist Assumptions

Readers of this volume are already familiar with many of the features of the minimalist system. Here, I briefly review the features that are important for the construction of my parsing algorithm.

The two most salient features of this system are its derivational character and the role that economy conditions play in regulating possible derived structures. At least at the level of competence, the model has moved away from the overgeneration-and-filtering character of its Government-Binding precursor. Structures that do not satisfy the economy conditions are simply not generated. The two major grammatical operations (Merge and Move) used to generate structure are seen as feature checking. Categories are input from the lexicon with features such as Case and θ-role that must be checked. Checking is satisfied when a category needing a feature is in construction with some other element in the sentence that can supply that feature. Movement or merger operations are licensed only if they allow feature checking to occur. Movement or merger serves to allow an element to transfer a feature necessary to satisfy some constraint. The relevant conditions that rule out overgeneration are the following:

(4) *Last Resort*

Operations do not apply unless required to satisfy a constraint. The minimal number of operations is applied to satisfy the constraint.

(5) *Greed*

"The operation cannot apply to α to enable some different element β to satisfy its properties. . . . Benefiting other elements is not allowed."

(Chomsky 1995, 201)

11.2 Multiple Spell-Out

A corollary assumption that has been incorporated into the Minimalist Program is the derivation of a correlation originally due to Kayne (1994). Previous grammatical formalisms had argued that restrictions on linear
precedence and immediate dominance were the product of two separate subsystems. Kayne (1994) suggested that these two systems are linked, and that one can derive precedence information from information about dominance. This conjecture is known as the Linear Correspondence Axiom (LCA). The initial definition derived precedence from dominance. I invert the claim to make the definition relevant for parsing.

(6) Linear Correspondence Axiom
   a. Base step: If $\alpha$ precedes $\beta$, then $\alpha$ c-commands $\beta$.
   b. Induction step: If $\gamma$ precedes $\beta$, and $\gamma$ dominates $\alpha$, then $\alpha$ precedes $\beta$.

C-command is defined as in Epstein, this volume.

(7) $\alpha$ c-commands all and only the terms of the category $\beta$ with which $\alpha$ was paired by Merge or by Move in the course of the derivation.

(8) illustrates the relationships licensed by these definitions.

```
   IP
     /\    /
    /  \  /  \    /
   DP I' I
  /\   /\   /\   /\   /
 D NP I VP
  the man tense slept
```

The precedence relations among elements in the subject are licensed because the determiner c-commands and precedes the NP (*man*). The second part of the definition is needed since the terminal elements in the subject position did not directly combine with the elements in the VP by either Merge or Move. Therefore, they do not c-command these VP elements even though the terminals in the subject precede those in the VP as required by the base step of the LCA. Their presence is allowed, however, by the second clause in the definition because the DP dominating both these terminals precedes the VP and dominates both the determiner and the NP, which inherit precedence by a kind of transitivity. Uriagereka (this volume) argues that the base step of the definition follows from the kind of “virtual conceptual necessity” inherent in the Minimalist Program. The simplest kind of mapping between precedence and dominance is one to one, and therefore we might expect a grammar that specifies
linear and dominance order to have this simplifying restriction (see Uriagereka, this volume, for details). The induction step forces trees to obey a “no tangling” condition as proposed by Partee, Wall, and Ter Meulen (1990) and Chametzky (1996). General goals of the Minimalist Program, which try to derive features of the grammatical system from “virtual conceptual necessity,” force us to either derive the induction step from other considerations or eliminate it from the system. Uriagereka adopts the latter course.

Uriagereka claims that we can maintain the simple relationship between command and precedence given by the base step in (6) if we allow Spell-Out to apply many times during the course of the derivation. Spell-Out is the operation that removes material from the syntactic component and feeds it to the interpretive components of Logical Form (LF) and Phonetic Form (PF) when it is ready for interpretation. Uriagereka points out that since the minimalist system dispenses with a global level of S-Structure as the conduit to the interpretive components, there is nothing to stop material from being passed for interpretation multiple times.

I assume that Spell-Out applies whenever two categories cannot be joined together by the Merge operation. If Merge does not apply, then the category currently being built is spelled out or reduced. I retain the notion from earlier theories of grammar (see Jackendoff 1972) that Spell-Out is a conduit between the syntax and the phonology. It is well known that the constituency established by the syntax is not relevant for phonological processes. Spell-Out turns a syntactic structure with relevant constituent relationships into a string ready for phonological interpretation. Uriagereka uses Spell-Out as a repair mechanism to retain one-to-one correspondence between dominance and precedence. He assumes that both precedence and dominance must be established between terminal elements at all points of the derivation. Precedence implies merger, and merger is possible only when a chain of dominance can be established. When merger is not possible, the string is linearized (turned into an unstructured string where only previously established precedence relations are preserved). Since the elements that have been linearized are invisible in the syntax, precedence does not have to be established between them and other items in the structure. Thus, when two categories cannot be combined through merger or movement (the only syntactic operations) to form a dominating category, the material that has been given structure so far is “spelled out” or linearized.
(9) "L is an operation \( L(c) = p \) mapping command units [units that can be formed through merger] \( c \) to intermediate PF sequences \( p \) and removing phrasal boundaries from \( c \) representations." (Uriagereka 1998)

This idea preserves the one-to-one mapping between precedence and dominance but at the cost of never building single phrase markers. Instead, the system builds blocks (Uriagereka calls them "command units") where all elements stand in a c-command relation to each other. When this c-command relation is interrupted, the unit is spelled out, an unstructured unit being shipped to the phonology for phonological interpretation and a structured unit being shipped to LF for semantic interpretation. The result of Spell-Out is an unstructured string (a syntactic word) with no further internal phrase structure. Within the context of the Minimalist Program, Spell-Out is a grammatical operation, on a par with movement transformations. As such, it is governed by conditions on transformations, in particular by the economy conditions discussed above. These conditions establish a preference for derivations that utilize the fewest operations possible. An operation is applied only to satisfy some independent grammatical condition. In this case, this means that the system spells out or linearizes only when it cannot otherwise establish a chain of precedence.

11.3 Minimalist Principles as a Parsing Algorithm

I will now apply a theory incorporating economy conditions and multiple Spell-Out to parsing. I assume that the algorithm applies left to right and evaluates ambiguities with respect to the economy conditions. As in minimalist theory, items are inserted into the derivation (or moved) with the goal of checking features. The feature-checking aspect of the theory will impose an argument-over-adjunct-attachment preference along lines suggested by Pritchett (1992) and Gibson (1991) on the assumption that 0-roles are relevant features for checking. Attachment as an adjunct will never lead to receipt or transfer of 0-, Case, or other features, whereas insertion into an argument position will allow this transfer to occur. I will show that this preference is well attested. Unlike Pritchett (1992) and Gibson (1991), I assume that feature transfer is optimized locally. Pritchett and Gibson allow the parser to scan the entire derivation up to the position of an item's attachment and to compare whether the attachment
of a category optimizes the assignment of features over all elements of the
tree built so far. By contrast, since feature checking is subject to Greed in
the minimalist system, this theory allows optimal feature checking only on
the particular category that is being attached, irrespective of whether this
optimizes feature checking across the derivation as a whole. This will be
crucial for some of the examples below.

Insertion or movement is governed by the economy conditions dis-
cussed above. The preference to attach a category using minimal structure
follows immediately from this notion of economy. At each point, a cate-
gory is inserted using the fewest operations necessary for feature transfer
or merger. This ban on unnecessary operations subsumes Frazier and
Rayner’s (1982) Minimal Attachment Principle and Gorrell’s (1995) sim-
plicity metric and has the advantage of following from independently
motivated grammatical principles. Minimal Attachment or simplicity
follows because the fewest mergers will be performed that are needed to
incorporate an incoming category.

Following Uriagereka, I assume that Spell-Out occurs whenever a der-
ivation would otherwise violate the LCA (now containing only the base
step). The spell-out conditions thus also provide an independently moti-
vated theory of reanalysis. If a preferred reading induces a precedence/
dominance mismatch, the category that precedes but does not dominate
will be spelled out. Again following Uriagereka, this means that the
material inside the spelled-out category is linearized and all internal syn-
tactic structure is removed, creating a nondecomposable syntactic word.
Given this, reanalysis from the preferred to the dispreferred reading that
requires either extraction of material from, or insertion of material into,
this syntactic word will be impossible. As a lexical item, the spelled-out
material is an atomic unit, which can no longer be decomposed into its
component pieces. However, if reanalysis occurs within a domain where
Spell-Out has not applied, then material can be accessed and the preferred
reading can be transformed into the dispreferred reading. Incorporating
Spell-Out and economy conditions into the grammar also explains the
preference for right-branching derivations without the need for extra
explicit principles that favor this type of derivation.4

As a grammatical operation, Spell-Out is governed by economy. Since
it does not allow the checking of any features, it is a last resort operation.
As such, it is invoked only when no other feature-checking operation can
apply, and the minimal number of spell-outs to guarantee satisfaction of
the LCA operate at each time step in the derivation. A right-branching
structure ensures that an element that precedes also dominates a category, thus minimizing the need for Spell-Out. Therefore, right-branching structures are preferred because they economize on the need for Spell-Out.

The algorithm in (10) embodies these principles.

(10) A derivation proceeds left to right. At each point in the derivation, merge using the fewest operations needed to check a feature on the category about to be attached. If merger is not possible, try to insert a trace bound to some element within the current command path. If neither merger nor movement is licensed, spell out the command path. Repeat until all terminals are incorporated into the derivation.

11.4 Some Cases

11.4.1 Argument/Adjunct Attachment Ambiguities
Cases of argument/adjunct attachment ambiguity illustrate the role of optimizing feature checking relative to economy conditions. In all cases, attachment as an argument is preferred because it allows assignment of features.

11.4.1.1 Direct Object/Complement Subject Ambiguity The sentences and the relevant structures that illustrate direct object/complement subject ambiguity are given in (11).

(11) a. the man believed his sister to be a genius
    b. the man believed his sister
    c. VP
        V
        believed
    d. VP
        V
        believed
        IP
        DP
        D
        NP
        his
        N
        sister
        D
        NP
        his
        sister

The DP his sister will be assigned both Case and θ-features by the preceding verb if it is attached as the direct object. Appropriate θ-features can only be assigned by the θ-assigner, the head of the complement.
clause. Since this category has not yet been processed, no features will be assigned if his sister is attached as the subject of the complement clause. Therefore, (11c) is the preferred structure. It is also the more economical structure, involving fewer operations, although this is not a crucial determinant of attachment for this case. In neither case is Spell-Out necessary at the site where his sister is attached.

Since there has been no spell-out within the VP, both the verb and the object are available when the embedded verb is encountered in a case like (11a). Therefore, the object NP is available for reinsertion as the embedded subject in (11d) even though the initial structure chosen for this case is (11c). All elements remain on the command path.

11.4.1.2 Proposed Object/Matrix Subject Ambiguity Next consider (12a), where there is a preference to treat the word following the first verb as an object in the preposed adverbial (see (12b)), rather than as the subject of the matrix sentence.

(12) a. after Mary mended the socks fell off the table
   b. after Mary mended the socks they fell off the table

Again, incorporation as an object allows Case and θ-features to be checked off from the phrase the socks. Incorporation as the matrix subject does not allow any Case or θ-feature checking, again because the Case- and θ-assigning head of the IP has not yet been incorporated into the structure. The relevant structures are given in (13).

(13) a. 

```
   IP
   PP
   P
   IP
   after
   DP
   VP
   Mary
   V
   DP
   mended
   D
   NP
   the
   socks
```
We do not expect reanalysis to be possible given the algorithm (10). After the optimal structure in (13) is built, the phrase *fell* cannot be incorporated into the preposed adverbial clause. A globally optimizing algorithm might look to see what series of transformations could be applied to incorporate this category. However, algorithm (10) is a dumb one that acts only to incorporate local material. Since the second VP cannot incorporate into any node within the preposed adverbial, the adverbial is spelled out in a phrase-by-phrase manner, leaving the structure in (14). This structure respects the LCA.

(14)

However, there is no way to incorporate the structure into this remnant either. Since the preceding material has been spelled out, there is no way to retrieve anything from this phrase to be inserted as the necessary matrix subject. Since no further operations apply, and there is remaining unincorporated terminal material, the parse fails and a garden path is detected.

11.4.1.3 Ditransitive/Complex Transitive Object Ambiguity

Next consider (15).
(15) a. John gave the man the dog for Christmas
    b. John gave the man the dog bit a bandage

The preferred reading for (15b) is to treat the dog as a ditransitive object, as in (16), rather than as the subject of a relative clause modifying the man, as in (17).

(16)  
```
  VP
  / \   \
  V   VP
     / \  \
  gave_1 DP V'  \
     \    /  \
      the man  V
                / \  \
               e_i D   NP
                 \    /  \
                   the  dog
```

(17)  
```
  VP
  / \   \
  V   VP
     / \  \
  gave_1 DP V'  \
     \    /  \
    DP CP V  \
       / \  \
  the man  C   e_i
          /   \
         IP  \
           \  
             the dog
```

Clearly, (17) is more complicated and requires more mergers than (16), violating economy. This is again not crucial because the analysis as an indirect object allows features to be checked on the DP the dog whereas
attachment as material in the embedded subject does not allow feature transfer.

Reanalysis is not possible in this structure. To explain why, I crucially assume the Larsonian shell structure in (16). Reanalysis would involve incorporation of the category in the indirect object position originally part of the relative clause on the direct object. However, this cannot be accomplished while the trace of the moved V remains in the structure because a relative clause inside the direct object would not command the verb trace. Therefore, maintaining the terminals of the preceding relative and the verb trace in the same tree would violate the LCA. Therefore, the V' in (17) must be spelt out. If this category is spelt out, however, there is no host site for subsequent attachment of the true indirect object because all structure under the V' node is no longer accessible.

11.4.1.4 Subcategorized PP/NP Modifier Ambiguities Consideration of (18a,b) shows that there is a preference to treat the PP on the table as an argument of the verb put rather than as a modifier of the NP the book. I will assume (noncrucially) the Larsonian analysis of PP complements as well. Whatever the structure is, attachment as an argument allows the PP to receive and the V to discharge features. The structures are given in (18c,d).

(18) a. I put the book on the table
    b. I put the book on the table into my bag
    c. 

\[
\begin{align*}
\text{VP} & \quad \text{VP} \\
\text{V} & \quad \text{VP} \\
\text{put}_i & \quad \text{VP} \\
\text{DP} & \quad \text{V'} \\
\text{D} & \quad \text{NP} \\
\text{the} & \quad \text{book} \\
\text{V} & \quad \text{PP} \\
\text{e}_i & \quad \text{P} \\
\text{on} & \quad \text{the} \\
\text{table} & \quad \text{DP}
\end{align*}
\]
A Minimalist Theory of Human Sentence Processing

Reanalysis is not possible for the same reason as in the ditransitive case above. To reanalyze the PP as part of the direct object as the adjunct to the book requires spell-out of the V’, since material inside the relative would not command this category. If this category is spelled out, though, there is no site for the true locative PP into my bag to merge to.

The final case of an argument/adjunct ambiguity is the famous main clause/relative clause ambiguity exhibited in cases like (19).

(19) the horse raced past the barn fell

These are strict garden paths, for which native speakers prefer a main-clause reading (the horse raced past the barn), being unable to reanalyze them as reduced relative clauses.

Interestingly, Pritchett (1992) and Stevenson and Merlo (1997) have suggested that these types of ambiguities do not always yield garden paths. When transitive and unaccusative verbs replace unergatives like those in (19), the sentences become quite easy to process.5

(20) a. the student found in the classroom was asleep
    b. the butter melted in the pan was burnt

Within the context of the minimalist framework, these subtle facts are accounted for because both transitives and unaccusatives must have traces inserted in the postverbal position, whether or not these structures are analyzed as main clauses or relative clauses. This is because the 0-grid of both transitives and unaccusatives signals to the parser that these verbs both require NP objects. Since there is no overt object in the postverbal position, a trace must be inserted here. So, even if the preferred analysis
for these cases is as main clauses, the structure needed to appropriately interpret them as open sentences, with traces in postverbal position, is built as part of the main-clause analysis, before the spell-out required by the disambiguating main verb for cases that are truly reduced relatives. The initial analyses are given in (21a,b).

(21) a. \([IP \text{ the student}_i [VP \text{ found}_i e_i [\text{ in the classroom}]]]\)
b. \([IP \text{ the butter}_i [VP \text{ melted}_i e_i [\text{ in the pan}]]]\)

The reanalysis proceeds along the lines discussed above. The material preceding the main verb is initially analyzed as a main clause. When the true matrix verb is encountered, everything preceding the verb is spelled out in accordance with the LCA. Now, however, the spelled-out material can be appropriately interpreted as a relative clause, and so no garden path results.

In all of the above cases, economy seemed to redundantly track feature checking in the sense that the most economical structure was also the one that allowed features to be checked. I now turn to cases where local economy is crucial to predicting both preference and reanalysis judgments. Primarily, in these cases the ambiguity is between two different types of adjunct attachment. In neither situation is a feature checked, so economy is the only factor in play.

14.4.2 Adjunct/Adjunct Attachment

14.4.2.1 Adverb or Particle Placement The grammar presents multiple attachment sites directly after the italicized words in all the cases in (22). The parser always chooses the position after the most recently encountered word as the preferred attachment site.

(22) a. I told Mary that I will come yesterday
b. I called to pick the box up
c. I yelled to take the cat out

In (22a), the adverb yesterday is construed with the embedded verb even though this reading is semantically anomalous and even though an alternative attachment to the matrix verb would result in an acceptable reading. (22b,c) show that the particle prefers low attachment as well.

These preferences can be explained on the assumption that Spell-Out, as one of the grammatically licensed operations, is also subject to economy conditions. Therefore, mergers involving fewer spell-outs will be...
preferred. Consider (23) at the point when the adverb *yesterday* enters the derivation.

\begin{center}
\begin{tikzpicture}
  \node (v) {\textit{VP}};
  \node (vp) [below of=v] {\textit{VP}};
  \node (v1) [below of=vp] {\textit{V}};
  \node (v2) [below of=vp] {\textit{V'}};
  \node (dp) [left of=vp] {\textit{DP}};
  \node (v3) [below of=v1] {\textit{\#\# Mary \#\#}};
  \node (v4) [below of=v2] {\textit{\#\# I \#\#}};
  \node (ip) [below of=v3] {\textit{C}};
  \node (cp) [below of=v4] {\textit{IP}};
  \node (dp1) [below of=ip] {\textit{that}};
  \node (i1) [below of=cp] {\textit{I'}};
  \node (vp1) [below of=i1] {\textit{VP}};
  \node (v5) [below of=vp1] {\textit{will}};
  \node (v6) [below of=v5] {\textit{come}};
  \node (v7) [below of=v6] {\textit{V'}};
  \node (v8) [below of=v7] {\textit{V}};
  \node (ap) [below of=v8] {\textit{AP}};
  \node (e) [below of=v8] {\textit{e_i}};

  \draw (v) -- (vp);
  \draw (vp) -- (v1);
  \draw (vp) -- (v2);
  \draw (v1) -- (v3);
  \draw (v2) -- (v4);
  \draw (v3) -- (ip);
  \draw (v4) -- (cp);
  \draw (ip) -- (dp1);
  \draw (cp) -- (i1);
  \draw (i1) -- (vp1);
  \draw (vp1) -- (v5);
  \draw (v5) -- (v6);
  \draw (v6) -- (v7);
  \draw (v7) -- (v8);
  \draw (v8) -- (ap);
  \draw (ap) -- (e);
\end{tikzpicture}
\end{center}

Assuming attachment into a Larsonian shell associated with the lowest verb, where adverbs assume the position of complements, would require no spell-outs at this point. The adverb would simply be merged under the boldfaced phrase. Assuming Uriagereka's version of the LCA, though, attachment as an adjunct to the higher verb would require spell-out of the lower VP, I', and IP, respectively, given the algorithm in (10). This algorithm requires spell-out only of the material that would not c-command the site of a potential merger. Therefore, if the parser has processed everything up to the lowest clause in the preposed position, it will require multiple spell-outs to return to the highest level of the preposed adverbial. In the competence model, one could think of high or low attachment as requiring an equal number of spell-outs, each with a different number of
phrases in the spelled-out component of the analysis. In a parser, however, one does not keep the whole structure in memory at a given point and therefore, one must provide an explicit procedure for dealing with previously processed material. The parser cannot retrieve a site for attachment in this case without successive iterations of Spell-Out, given (10). Since lower attachment involves fewer iterations of Spell-Out, economy considerations favor this attachment choice. This will also be true for (22b,c). Attachment of the particle to the higher verb will cause the phrase remaining on the c-command path of the lower clause (boldfaced in (24) to be spelled out (only the structure of (22b) is illustrated; the structure of (22c) is the same). Attachment as the particle of the lower verb require no Spell-Out and will again be preferred by economy considerations.

(24)

The next case is discussed by Phillips and Gibson (1997). Normally, relative-clause attachments are dispreferred, but in this case they are favored.

(25) a. although Erica hated the house she had owned it for years

b. although Erica hated the house she owned her family lived in it for years
Phillips and Gibson presented sentences like these to experimental subjects with either temporal or nontemporal adverbial modifiers in a word-by-word self-paced reading task with a moving window display. At the disambiguation point (either *it* or *her family*), subjects showed a clear preference for attaching the preceding clause as a relative clause modifying the noun phrase *the house*. There was a significant increase in reaction time at the disambiguation point if the ambiguous noun phrase was disambiguated as the matrix subject.

This preference can again be explained with reference to economy of Spell-Out. Again at the relevant point, neither attachment will allow the discharge of a feature. However, attachment as a relative clause permits much more of the preceding material to remain in the derivation, as it will still command the incoming merged material. Attachment as the matrix subject requires spell-out of the entire preposed adverbal. The relevant structure is given in (26); nodes that need to be spelled out for the matrix-subject reading are boldfaced, and those that need to be spelled out for the relative-clause reading are underlined.

(26)

Phillips (1995, 1996) presents very interesting work that argues for an alternative grammatically based processing theory. In fact, Phillips claims that there is no distinction between the parser and the grammar. Derivations in both the competence and performance systems are built up incrementally, left to right.

Given this grammatical underpinning, Phillips tries to link performance preferences to the grammar in the following way. First, he defines a condition called Branch Right.

(27) **Branch Right**

"Select the most right branching available attachment of an incoming item Reference Set: all attachments of a new item that are compatible with a given interpretation." (Phillips 1996, 29)

The preference for right-branching structure is in turn derived from a principle that ensures that the base step of the LCA is incrementally satisfied to the greatest extent possible. As Phillips writes: "I assume that a structure is right branching to the extent that there is a match between precedence relations among terminal elements and c-command relations among terminal elements" (1996, 29).

To handle a variety of bracketing paradoxes, Phillips couples this with the idea that grammatical as well as parsing derivations proceed left to right. Consider (28).

(28) a. John showed the men each other’s pictures  
   b. John showed each other the men’s pictures

These examples suggest that double object constructions have right-branching structures in which the indirect object c-commands the direct object as in (29).

(29)  

```
  VP
   \  /  
  V   VP
     \  /  
    showed, DP V'
     \  /   \  /
    the men V DP
         \  /  
e_i each other’s . . .
```
The fact that (30a–c) are grammatical as VP-fronting structures suggests that the structure for the PPs (given in (31)) should be left branching, allowing the right subparts to be constituents.

(30) I said I would show the men the pictures in libraries on weekends, and
   a. show the men the pictures in libraries on weekends, I will
   b. show the men the pictures in libraries, I will on weekends
   c. show the men the pictures, I will in libraries, on weekends

(31)

Phillips shows that we can derive the effects of a structure like (31) without the need to assume it, by assuming that Branch Right applies from left to right with the apparently left-branching structures actually being intermediate structures in the derivation. For example, Branch Right would first combine show and the men to form a constituent. This constituent would then be reconfigured when subsequent material was uncovered.

Phillips presents a variety of advantages for his approach over other treatments of paradoxical constituency. The definition in (27) suffices to handle all of these paradoxes. Phillips claims that Branch Right can be
used to resolve various parsing ambiguities. In order to do this, he redefines Branch Right as follows (1996, 111):

(32) *Branch Right*

a. Metric: Select the attachment that uses the shortest path(s) from the last item in the input to the current input item.

b. Reference set: all attachments of a new item that are compatible with a given interpretation.

(33), repeated from (11), is a simple illustration of how the principle works.

(33) a. the man believed his sister to be a genius

b. the man believed his sister

Branch Right predicts the preference for construing *his sister* as a direct object because there are fewer branches in the path between *believed* and *his sister* if one construes the postverbal NP as a direct object than if one construes this phrase as the subject of an embedded clause as shown in (34).

(34) a. 

```
VP  
|   
V   DP  
|   
believed   D  NP  
|   his  sister  
Path = 1 step up from V to VP  
1 step down from DP to D
```

b. 

```
VP  
|   
V   IP  
|   
believed   DP  
|   D  NP  
|   his  sister  
Path = 1 step up from V to VP  
3 steps down from VP to IP, DP, and D
```

Since the embedded-subject reading requires more steps on the downward path, it is dispreferred.

Phillips uses this simple principle to handle a wide range of data from English and illustrative cases from German and Japanese. The principle's empirical coverage is impressive. In addition, its use is argued to be independently justified by the LCA or at least by its ability to handle bracketing paradoxes; thus, it appears that we are getting a parsing principle for free from an independently needed competence principle. For these
reasons, Phillips’s approach is quite interesting. Nonetheless, I will argue against it on several grounds.

The problems I will discuss do not focus on the empirical coverage ofPhillips’s theory per se. I note in passing, however, that this theory is intended merely to be a theory of initial preference. It is well known that certain initial preferences, such as the one exemplified in (33), can be overridden given subsequent disambiguating material, whereas cases like (35) are not subject to reanalysis and remain garden paths.⁶

(35) the horse raced past the barn fell

(35) is initially interpreted as a main clause, the horse raced past the barn. Reanalysis as a reduced relative, the horse that was raced past the barn, is impossible. As is well known, the availability of this grammatically licensed interpretation has to be pointed out to naive speakers. Phillips’s theory is silent on the issue of when reanalysis is possible. Phillips claims that reanalysis should not be part of the theory of sentence processing:

[J]t is not clear that one should want Branch Right to account for recovery from error. I assume that Branch Right is a property of the system that generates and parses sentences in a single left-to-right pass, and that reanalyses require backtracking and are handled by other mechanisms. (1996, 161)

I agree with Phillips that the actual mechanisms of reanalysis, particularly in cases where conscious breakdown occurs, may not be the domain of the processor. However, I see no reason not to demand that a full theory of sentence processing distinguish cases where these mechanisms can apply, where the human sentence processor presents the appropriate representations for these mechanisms to operate on, from cases where the sentence processor does not present the appropriate representations for the operation of potentially external general-purpose reanalysis mechanisms. Phillips’s theory is mute on this domain of empirical prediction. Moreover, given that both interpretations in (33) are easily processible, it is hard to see why these reanalyses are not the domain of the human sentence processor. Branch Right predicts that (33a) is the preferred interpretation. Therefore, there must be some principle that is part of normal sentence processing that transforms (33a) into (33b).

I turn now from the domain of prediction to that of independent motivation. Part of the main appeal of the Branch Right theory is its independent motivation in terms of the LCA and the bracketing paradoxes. We get a processing principle for nothing. However, I will show that this motivation is partial at best.
Again consider (32). (32) crucially relies on comparing the number of steps needed to derive both possible readings independently of whether either reading causes a precedence/c-command mismatch. Both structures in (33) respect the grammatically relevant version of Branch Right given in (27) where “right-branchingness” is defined in terms of respect for the base step of the LCA. In both structures, the verb both precedes and dominates the following NP whether or not it is construed as the direct object, as in (33b), or the complement subject, as in (33a). Nonetheless, speakers have a clear preference for interpretation (33b) over interpretation (33a). This prediction thus rests on the notion of shortest path. This, however, is not independently motivated by any of Phillips’s grammatical considerations. In effect, Phillips has sneaked in a branch-counting principle similar to the grammatically unmotivated Minimal Attachment Principle of Frazier and Rayner (1982), yielding a principle that is only half motivated by the grammar. Without the “node-counting” part of this principle, the theory is too weak to predict the preference for (33b) over (33a).

In (36), I present a case where the theory without the minimal attachment addendum is too strong.

(36) a. the man told the doctor that he was having trouble with his feet
   b. the man told the doctor that he was having trouble with to leave

Building either structure at the ambiguous point involves creating a precedence/dominance mismatch. Nonetheless, there is a strong preference for (36a) over (36b). Phillips assumes that the preferred structure is analyzed as a VP shell. As such, it would look like (37) (ambiguous material italicized).

(37) 

```
           VP
            |
          told
            |
          VP
            |
          DP
            |
          V'
            |
          D   NP  e_i  CP
            |
          the   doctor   C
            |
                  that
```
In this structure, the direct object *the doctor* dominates neither the trace of the verb *told* nor the complementizer of the complement clause. This structure induces a precedence/dominance mismatch. As (38) shows, the same is true in the less highly valued relative reading.

(38)

```
  VP
   \-- V
       \-- VP
           \-- told
          \--- DP
             \-- V'
                 \-- CP
                     \-- C'
                         the doctor
                         O
                         that
```

Again, then, the difference in these cases is not attributable to the metric of precedence/dominance correspondence or mismatch, but to the length of the path between *man* and the next terminal node. Again, this reduces to the unprincipled “node-counting” portion of Branch Right.

To sum up: I have identified two problems with Branch Right. First, it fails to provide a theory of reanalysis; more precisely, it does not distinguish representations in such a way as to form a basis even for an independent theory of reanalysis. Second, it incorporates a “minimal path” condition as well as a preference for right-branching structure in such a way that the minimal path condition cannot be derived from the latter part of the condition. As a result, a large portion of the constraint is not grammatically motivated. Without this unmotivated portion, the theory is empirically both too strong and too weak.

11.6 Constraint-Based Theories

In this section, I contrast a grammatically based approach, such as the two previously discussed, with frequency-based or probabilistic constraint-based theories. MacDonald, Pearlmuter, and Seidenberg (1994) present a theory of this type, summarizing it as follows:
Processing involves factors such as the frequencies of occurrence and co-occurrence of different types of information and the weighing of probabilistic and grammatical constraints. Our approach has suggested that syntactic parsing, including ambiguity resolution, can be seen as a lexical process. (1994, 700)

Under this view, structural heuristics are replaced with frequency data about use of either a lexical item or, in some theories, a construction type. For example, the "minimal attachment preference" in (33) would not derive from a minimal attachment preference, or from its grammatical derivation through economy. Rather, speakers can tune either to the fact that believe is used much more frequently with a simple NP as its direct object than with a sentential complement, or to the fact that simple sentences occur more frequently in the language than sentences with embeddings. Since this theory is "verb sensitive," it can easily account for the verb sensitivity of a variety of preference judgments. For example, verbs like decide, which occur much more frequently with sentential complements, are correctly predicted to be immune from the "minimal attachment" effect.

(39) John decided the contest was fair

I would like to argue that, although speakers may very likely track frequency, this variable works in tandem with independent grammatical constraints. If a structure is chosen, based on an amalgam of frequency and grammatical variables, the grammatically driven reanalysis principles decide what will or will not be a garden path.

In (20), repeated as (40), I considered a case where lexical choice is also relevant to preference judgments. Stevenson and Merlo (1997) have suggested that unaccusative and transitive cases are much better as reduced relative clauses than are unergative verbs.

(40) a. the student found in the classroom was asleep
    b. the butter melted in the pan was burnt

Table 11.1 gives grammaticality ratings for unaccusative versus unergative single-argument verbs (all tables are from Merlo and Stevenson 1996). Stevenson and Merlo found that unaccusatives were indistinguishable from transitives with respect to grammaticality judgments, yielding a two-way distinction, with unergatives being terrible as reduced relatives, and transitives and unaccusatives being fine. They surveyed corpora with the goal of determining whether grammatical preferences could be derived from frequency of occurrence in a corpus. Using the Wall Street Journal
Table 11.1
Grammaticality ratings (1 = perfect; 5 = terrible)

<table>
<thead>
<tr>
<th>Verb</th>
<th>Score</th>
<th>Verb</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>melt</td>
<td>2</td>
<td>begin</td>
<td>2</td>
</tr>
<tr>
<td>mutate</td>
<td>1.66</td>
<td>break</td>
<td>1</td>
</tr>
<tr>
<td>pour</td>
<td>1.66</td>
<td>freeze</td>
<td>1.5</td>
</tr>
<tr>
<td>reach</td>
<td>1</td>
<td>grow</td>
<td>1</td>
</tr>
<tr>
<td>advance</td>
<td>5</td>
<td>sink</td>
<td>3.25</td>
</tr>
<tr>
<td>glide</td>
<td>5</td>
<td>fly</td>
<td>4.25</td>
</tr>
<tr>
<td>march</td>
<td>5</td>
<td>ring</td>
<td>3.75</td>
</tr>
<tr>
<td>rotate</td>
<td>5</td>
<td>run</td>
<td>5</td>
</tr>
<tr>
<td>sail</td>
<td>5</td>
<td>withdraw</td>
<td>3.40</td>
</tr>
<tr>
<td>walk</td>
<td>3.75</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

corpus as the reference, they counted how many times a structure appeared as a reduced relative versus how many times it appeared as a main clause. They looked at the relative frequencies of three lexical classes of verbs: unergative, unaccusative, and “ordinary.” “Ordinary” verbs are distinguished from unergative and unaccusative verbs in that adding the second argument does not invoke a “causative interpretation on the predicate.” A paradigm is given in (41).

(41) a. **Unergative**
    I raced the horse (caused the horse to race) vs. the horse raced
b. **Unaccusative**
    I broke the vase (caused the vase to break) vs. the vase broke
c. **“Ordinary”**
    John played soccer vs. John played
d. **“Ordinary”**
    John kicked the soccer ball vs. John kicked

The results are given in table 11.2. Notice that both unergatives and unaccusatives appear less frequently in relative clauses than do “ordinary” verbs.

In a frequency-based approach, it is reasonable to correlate other factors that are presupposed by a reduced-relative-clause analysis with preference for this analysis to be interpreted as a relative clause. Reduced relative clauses can only be formed from passivized transitive verbs.
Table 11.2
Number of reduced relatives versus main clauses in 1.5-million-word *Wall Street Journal* corpus

<table>
<thead>
<tr>
<th></th>
<th>Reduced relatives</th>
<th>Main clauses</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unergatives</td>
<td>1</td>
<td>327</td>
<td>328</td>
</tr>
<tr>
<td>Unaccusatives</td>
<td>6</td>
<td>358</td>
<td>364</td>
</tr>
<tr>
<td>Ordinary</td>
<td>16</td>
<td>361</td>
<td>377</td>
</tr>
</tbody>
</table>

Table 11.3
Number of transitive versus intransitive frames from Penn treebanked subsection of *Wall Street Journal*

<table>
<thead>
<tr>
<th></th>
<th>Transitives</th>
<th>Intransitives</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unergatives</td>
<td>86</td>
<td>242</td>
<td>328</td>
</tr>
<tr>
<td>Unaccusatives</td>
<td>176</td>
<td>228</td>
<td>404</td>
</tr>
<tr>
<td>Ordinary</td>
<td>268</td>
<td>114</td>
<td>382</td>
</tr>
</tbody>
</table>

Therefore, the ability of a lexical class of verbs to appear in transitive or passive constructions irrespective of whether these constructions are reduced relatives could be used to predict the compatibility of this class with the main-clause or relative-clause interpretation. Stevenson and Merlo (1997) looked at these factors for the three classes of verbs.

First, they looked at the number of times a verb appeared as a transitive or an intransitive verb. Interestingly, as table 11.3 indicates, these data seem to show a three-way distinction: unergatives are normally used with one argument, unaccusatives show a more even distribution, and "ordinary" verbs show a distinct tendency to be transitive. Merlo and Stevenson report that all of these differences are significant.

When Merlo and Stevenson tested frequency of occurrence in the passive voice, they found no significant difference between unergative and unaccusative verbs; both types occur relatively infrequently as passives. However, they observed a significant difference between unaccusative and "ordinary" verbs: the latter occur significantly more frequently as passives. They observed a similar pattern for frequency of occurrence as a past participle. The relative frequencies are summarized in table 11.4. In all cases, the "ordinary" verbs show up significantly more frequently in constructions supporting the reduced-relative-clause interpretation. This would predict that speaker preference for a reduced-relative-clause anal-
Table 11.4  
Relative frequencies for several values of voice in the three verb classes

<table>
<thead>
<tr>
<th></th>
<th>Transitive</th>
<th>Passive</th>
<th>Past participle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unergative verbs</td>
<td>.26</td>
<td>.02</td>
<td>.12</td>
</tr>
<tr>
<td>Unaccusative verbs</td>
<td>.44</td>
<td>.05</td>
<td>.18</td>
</tr>
<tr>
<td>“Ordinary” verbs</td>
<td>.70</td>
<td>.26</td>
<td>.40</td>
</tr>
</tbody>
</table>

ysis should be most strongly correlated with ordinary verbs, less strongly with unaccusative verbs, and least strongly with unergative verbs. Merlo and Stevenson (1998) did not test ordinary verbs in their grammaticality study. Nevertheless, they claim that ordinary verbs “are readily interpretable as an RR [a reduced relative]” (p. 136), citing MacDonald 1994 as support. However, a closer look at MacDonald’s experiments yields the opposite conclusion. MacDonald contrasted the sentences in (42) with respect to self-paced reading times at the underlined position.

(42) a. the dictator fought in the violent coup was hated throughout the country
   b. the dictator chased in the violent coup was hated throughout the country
   c. the dictator overthrown in the violent coup was hated throughout the country

Fought is an “ordinary” verb, as is chased Chased, however, occurs overwhelmingly as a transitive verb, whereas overthrown is an unambiguous passive participle that is obligatorily transitive. MacDonald (1994) reports that verbs from the first class led to significantly greater reading times at the disambiguating region than did the second and third types. This means that verbs like fought were originally treated as main verbs in these constructions, yielding garden paths and increased reading times at the disambiguating main verb. Such verbs must be strongly disambiguated within the relative clause; otherwise, they will lead to a garden path. (43) is much easier to interpret as a relative clause because the by phrase disambiguates the structure as passive before the relative VP is closed.

(43) the dictator fought by the opposition was hated throughout the country

Interestingly, five of the eight “ordinary” verbs from table 11.3 were part of the fought class tested by MacDonald. This means that even though
the *fought* class appears more frequently in the transitive construction than either the unergative or the unaccusative class in Merlo and Stevenson's frequency counts, it is difficult for speakers to interpret occurrences of these verbs as a reduced relatives. If this is correct, it poses a problem for frequency-based approaches, which would predict that this class should be the least difficult to intrepret as it most frequently appears in constructions that relative-clause interpretations presuppose.\(^\text{10}\)

These remarks are also relevant to evaluating work by Filip et al. (to appear). This work claims that Merlo and Stevenson's results can be accounted for by a constraint-based theory. The first part of the paper essentially replicates Stevenson and Merlo's finding that a frequency-based approach can predict an unergative/unaccusative split for reduced-relative-clause interpretations since unaccusatives occur significantly more often than unergatives in constructions supporting the reduced-relative-clause interpretation. Since Filip et al. do not contrast these constructions with constructions containing "ordinary" verbs, however, they do not refute the claim that the difficulty of interpreting these constructions is unrelated to their frequency distribution. Nor does their hypothesis (19), shown here in (44), account for the relative difficulty of these cases.

(44) "The acceptability of sentences with reduced relative clauses, headed by passive participles derived from unergative and unaccusative verbs, increases when the passive participle and the main verb of a matrix clause assign their subject-NPs more Proto-Patient and fewer Proto-Agent properties." (p. 17)

*Proto-patient* and *proto-agent* are terms taken from Dowty 1991. Perfect proto-agents are volitional, sentient causers of the event described, and they move and exist independently of the verb descriibed. Perfect proto-patients undergo a change of state, serve as incremental themes, are causally affected by another event participant, are stationary relative to another event participant, and do not exist independently of the event described. Dowty allows a category to serve in either of these roles if it is the argument bearing the greatest number of the role's defining properties, even if it is not perfectly correlated with them. Given (44), this analysis distinguishes (45) from (46). (45) is predicted to be unparsable because both the main verb and the embedded verb assign proto-agent properties to the head of the relative clause (italicized), even though it must be the proto-patient in the reduced-relative-clause interpretation.
(46) is parsable because the embedded verb assigns only proto-patient properties to its subject.

(45) the horse raced past the barn fell
    (Horse is sentient, causes the movement, and exists independently both as the object of race and as the subject of fell.)

(46) the fish fried in the pan made me sick
    (Fish undergoes a change of state, is the incremental theme, and is affected by fry in the reduced-relative-clause reading. It is a proto-agent of the predicate make sick because it is causal, does not undergo a change of state, and so on.)

Notice, however, that even though the head of the relative is a proto-patient in the reduced-relative reading, (47) (from MacDonald 1994) is still a garden path sentence.

(47) the cattle moved into the crowded room were afraid of the cowboys

Both move and be afraid assign the cattle proto-patient properties, because both imply that this argument undergoes a change of state, is the incremental theme, and is causally affected by another participant. Despite being assigned a relatively high number of proto-patient properties, this sentence is a clear garden path.

Results like these suggest that frequency has a role to play, but is filtered through grammatically justified constraints. Given the minimalist theory discussed above, “ordinary” verbs pattern like unergatives because when they are given their preferred interpretation as main-clause verbs, they are pure intransitives with no trace in the object position. The main-clause verb were in a case like (48) triggers reanalysis of moved ... room as a relative clause. However, by the time that happens, the material preceding were is already spelled out, and the trace necessary for interpreting moved ... room as a reduced relative cannot be inserted. The structure is given in (48).

(48)  
```
       # IP
       /   
  # DP      
   /       
  # the cattle moved #  were
```
Frequency, coupled with economy-driven conditions, may drive the initial preference for a given verb to be part of either a main clause or a reduced relative, but if the former, incorrect interpretation is chosen, reanalysis as a reduced relative will be impossible. This contrasts with the unaccusative cases, as discussed above. In these cases, a trace must be inserted in postverbal position whether or not the structure is interpreted as a main clause or as a reduced relative. Therefore, whether or not the main-clause or relative-clause reading is initially chosen (perhaps based on frequency), reanalysis is possible. If this account is correct, we are driven to a theory where frequency information interacts with grammatically based principles, but frequency does not replace these principles.

11.7 Conclusions

In this chapter, I have argued for a theory of processing preference and reanalysis that is heavily based on independently needed conditions within Chomsky’s grammatical theory. There are no independent “parsing principles.” In this case, the theory of preference is grounded in the economy conditions of Chomsky’s (1993) Minimalist Program.

I contrasted my approach with one proposed by Phillips (1995, 1996). The two theories are similar in that principles are all independently motivated by grammatical consideration. I argued, however, that the economy conditions allow us to derive the unmotivated “shortest path” portion of Phillips’s Branch Right. The Last Resort principle discussed above favors feature passing that involves the minimal number of steps.

I followed Uriagereka (this volume) in eliminating the induction step of the LCA in favor of a theory involving multiple instances of Spell-Out. I showed that multiple Spell-Out, when combined with the independently motivated economy conditions, also accounts for the preference for right-branching structures and an independently motivated theory of reanalysis.

Finally, I argued that these principles interact with frequency-derived parsing constraints in interesting ways and can explain subtle differences between the garden path status of reduced relatives derived from unergatives, unaccusatives, and transitives that are otherwise mysterious. This argues in turn for a theory where grammatical principles are supplemented but not replaced by considerations of frequency or probability.

Notes
I gratefully acknowledge support from the following agencies: the Army Research Laboratory under contract DAAL0197K0135, the National Security Agency
under contract MDA90496C1250, and MicroAnalysis and Design under SBIR contract 7030. Thanks to audiences at the CUNY Sentence Processing and Computational Psycholinguistics Conferences, to audiences at Yale University and the Basque Country Summer Courses, and to Robert Berwick, Stephen Crain, Samuel Epstein, Norbert Hornstein, Colin Phillips, and Juan Uriagereka for helpful discussion. Errors remain my own.

1. Thanks to Colin Phillips for pointing this out to me.

2. But see Steedman 1996 for a different view within the framework of Combinatory Categorial Grammar.

3. Colin Phillips (personal communication) points out that feature checking could also capitalize on the fact that VP adjuncts express a restricted set of relations (time, location, instrument). NP adjuncts form a looser modifying relationship to the NP head that they adjoin to. Assuming this, we could construe VP adjunction as potentially checking a feature that NP adjunction would not. This would allow us to predict a preference for VP adjuncts over NP adjuncts.

4. This point will become important in the discussion of Phillips’s proposals in section 11.5.

5. Stevenson and Merlo explain this contrast using a lexical theory devised by Hale and Keyser (1993). This theory is embedded in the processing model proposed by Stevenson (1993). Stevenson and Merlo’s central claim is that the lexical representations needed to process unergative reduced-relative clauses cannot be computed by a model like Stevenson’s, given lexical representations of the type Hale and Keyser propose. The account is extremely interesting; however, as shown in (i), it is too strong in that even unergatives can be improved in the appropriate discourse context.

(i) a. only horses raced at good tracks are eligible for the Triple Crown
   b. I saw a horse raced with a good jockey, and a horse raced with a bad one
   c. a horse raced with a good jockey will learn to go for the inside

This type of improvement due to different types of quantification was first pointed out by Crain and Steedman (1985) and has been elaborated on in much subsequent work by these authors. I will have more to say about these cases in section 11.6.

6. This is not to say that the initial preference as either a main or a relative clause cannot be shifted. It is merely to say that the initial preference is difficult to change once it is set.


8. Space considerations preclude fully contrasting a grammatically based account with the constraint-based theory. Important topics that I will not review or contrast here involve the importance of pragmatic and semantic factors. Do pragmatic factors interact with frequency or syntactically derived constraints in a modular or nonmodular way? My goal in this section is simply to discuss the role of grammatical principles in a theory where it is assumed that frequency plays some role.
9. The verbs *fought, studied, lectured, watched,* and *surrendered* were tested as biased intransitive verbs in MacDonald's set. She also tested the verbs *raced, moved,* and *marched,* which Stevenson and Merlo treat as unergative.

10. Clearly, these results are only suggestive since the class of verbs that yielded difficulty in MacDonald's experiments was primarily but not completely composed of Merlo and Stevenson's "ordinary" verbs. I am planning to retest MacDonald's stimuli with a class of nonmixed purely "ordinary" verbs. My intuitions suggest, however, that the "ordinary" verbs pattern with the unergative verbs.

References


