

Postverbal Behavior

Thomas Wasow

ISBN: 1-57586-402-9

Copyright notice: Excerpted from *Postverbal Behavior* by Thomas Wasow, published by CSLI Publications. ©2002 by CSLI Publications. All rights reserved. This text may be used and shared in accordance with the fair-use provisions of U.S. copyright law, and it may be archived and redistributed in electronic form, provided that this entire notice, including copyright information, is carried and provided that CSLI Publications is notified and no fee is charged for access. Archiving, redistribution, or republication of this text on other terms, in any medium, requires the consent of CSLI Publications.

Contents

Acknowledgments vii

Preface xi

- 1 Introduction 1
 - 1 An Example 1
 - 2 Objectives 3
 - 3 Some Phenomena Illustrating Order Variation 4
 - 4 Questions to be Addressed 12
 - 5 Why Only English Data? 14

- 2 Grammatical Weight 15
 - 1 Introduction 15
 - 2 Proposed Definitions of Weight 16
 - 3 Evaluating Characterizations of Weight 23
 - 4 Length and Complexity 32
 - 5 Explaining Weight Effects 42
 - 6 Conclusion 57
- Appendix A: Stimuli used in the questionnaire study 58
- Appendix B: Verb-Particle Frequency vs. HNPS Rates 60

3	Information Structure and Weight	61
1	Introduction	61
2	Characterizing the Role of Information Structure in Ordering	62
3	Information Structure vs. Grammatical Structure	67
4	Conclusion	81
4	Additional Factors	83
1	Introduction	83
2	Semantic Connectedness	83
3	Lexical Bias	87
4	Ambiguity Avoidance	88
5	Conclusion	109
	Appendix A: Stimuli from the questionnaire study	110
	Appendix B: Stimuli from the ambiguity experiment	113
5	Theoretical and Metatheoretical Implications	115
1	Introduction	115
2	Generative Analyses	116
3	Goals of Syntactic Research	131
4	Conclusion	147
6	Methodological Issues	149
1	Introduction	149
2	Simplicity	150
3	Generality	153
4	Coverage	156
5	Conclusion	164
	References	167
	Name Index	177
	Subject Index	181

Preface

The title of this book is obviously a play on the title of B.F. Skinner's *Verbal Behavior*. Chomsky's devastating 1959 review of that work in *Language* is generally regarded as a landmark in the history of linguistics—indeed of the cognitive sciences more generally. In alluding to Skinner's title, I do not mean to pay homage to his book (which I confess I have never read).

I do, however, have a point. The rejection of behaviorism in linguistics that was signalled most clearly by Chomsky's review was wholesale. Not only were unfounded theoretical strictures against mentalistic constructs overthrown; careful attention to the sources and quality of the data used to support theoretical claims were also abandoned. This doesn't exactly constitute throwing the baby out with the bathwater, but—extending the metaphor a bit—it is arguable that a washcloth and bar of soap got discarded.

As a student of Chomsky's thirty years ago, I learned the priorities and the methods of generative grammar. Elegant theories that 'captured generalizations' and clever arguments for such theories were what was most highly valued. Data were of interest only to the extent that they provided support for or evidence against some theoretical hypothesis. The only source of such data anyone employed was the introspective judgments of native speakers. Since almost all generative research at the time was about the native languages of the researchers, finding data to test hypotheses consisted, for the most part, of inventing critical example sentences and consulting one's own intuitions about their well-formedness (or, in some cases, their meanings).

The former dominance of behaviorist thinking was at that time a fresh memory for some of my teachers, and the shortcomings of behaviorism still occupied a significant place in the curriculum. This included not just demonstrations of the utility of non-observable and mentalistic theoretical entities, but also arguments that usage data (corpora) and statistical modeling were of no relevance to theoretical linguistics. Back then, I saw no reason to question the methodology we employed.

Not long after joining the Stanford faculty (in 1973), I began to interact with people working on natural language processing projects in the industrial research labs of Silicon Valley. This introduced me to a radically different perspective, where attention to how people actually used language was essential. I also found myself with academic colleagues—sociolinguists—whose research crucially involved attention to frequency information in usage data.

Through most of the 1980s I served as a consultant to a natural language processing project at Hewlett-Packard Laboratories. This heightened my awareness of the discrepancy between the kinds of data relevant to building useful language technologies and the sorts of examples employed by generative grammarians.

I also began to have serious questions about the reliability of much of the standard judgment data. I published a ‘Topic-Comment’ column in *Natural Language and Linguistic Theory* (Wasow, 1986) poking fun at linguistics on a number of grounds, including the loose standards of evidence. To my surprise, this prompted a serious and angry response from Morris Halle (one of my favorite teachers in graduate school, whom I considered a role model) and James Higginbotham (Halle and Higginbotham, 1986).

Shortly thereafter, I was offered the position of Dean of Undergraduate Studies at Stanford. My doubts about the kind of work I had been taught to do were a contributing factor in my decision to accept. I spent the next four years as a full-time administrator.

When I returned to my normal professorial duties in 1991, I had no ongoing research program and began shopping around for new avenues to explore. One of several projects I got involved in at that point was what got me started on the line of research described in this monograph. My colleague John Rickford talked to me about a phenomenon he had been observing for several years, namely, the use of the string *as far as* to restrict the topic of a sentence. He had noticed an increasing tendency for speakers to use *as far as* NP without what we later started calling a verbal coda—that is, some form of *go* or *be concerned*. John had collected hundreds of examples of topic-restricting *as far as*, both with and without a coda—that is, utterances like *Now as far as misunderstanding goes, I’d just like to focus* and *As far as the white servants, it isn’t clear*. He

asked me if I could find any systematic syntactic properties that would correlate with the presence vs. absence of a coda.

A fact that jumped out from even a cursory examination of John's examples was that when long or complex NPs immediately followed *as far as*, the coda was less common. I proposed a scale of NP complexity based on the internal structure of the NP, and we found that this was a statistically significant factor in the presence or absence of a coda. More important in the present context is the fact that I began to think about the issue of structural complexity and wondering where it might manifest itself in syntax.

The phenomenon of heavy NP shift immediately came to mind. I began to look at the characterizations of 'heavy' in the literature on the subject, finding a great many proposals, almost none of which had ever been tested systematically.

The work on *as far as* was important to the present project in another way, for it was the first time that I was involved in a study of usage. I learned some important lessons about introspection from that. In particular, it quickly became evident to me that different people had radically different intuitions about *as far as* without a coda. A number of speakers categorically rejected all examples (in keeping with the admonitions of the usage handbooks), while others were much more permissive. Significantly, to my mind, some of those most adamant in insisting on the obligatoriness of the coda were linguists. One well-known linguist even told me he had never heard anyone use topic-restricting *as far as* without a coda. Corpus studies, on the other hand, showed that the phenomenon is extremely common. Indeed, when I inspected a random sample from the (spoken) Switchboard corpus, I found a majority of the instances of topic-restricting *as far as* had no coda. Another colleague (a philosopher of language this time) told me he found the variant without the coda totally unacceptable, but I subsequently heard him use it repeatedly in a formal presentation.

In short, both the substance of the present monograph and its methodology are an outgrowth of my participation in the *as far as* project (see Rickford, et al, 1995, for the results of that project). My interest in weight led me to look at Heavy NP Shift, which in turn got me interested in other alternations in the ordering of constituents after the verb in English, particularly the verb-Particle construction and the Dative Alternation. And I continued to gather data from a combination of usage examples (from corpora and casual observation) and systematic elicitation methods—including judgment questionnaires—rather than my own introspection.

My change in methodology was quite naturally accompanied by a change in the questions I tried to answer. In particular, instead of seeking to formulate and justify formal analyses (rules, lexical entries, parameters, or what have

you) within some generative theory, I asked, more straightforwardly, what led speakers (and writers) to select one postverbal constituent ordering rather than the other.

I quickly discovered that most generative linguists had little interest in the new question I was asking. Psycholinguists—especially those trained as psychologists—thought it was worth pursuing, so I began spending more time with psychologists. I observed that psychologists spend most of their time and effort on insuring the quality of their data—that is, on experimental design and analysis. That emphasis is strikingly different from generative linguists', and I have on occasion found myself wondering why psychologists give so little attention to developing high-level theories. It seems to me that both fields would benefit from acting a little more like the other.

Returning to the title of this monograph, the use of the word 'behavior' is meant to emphasize the fact that my attention is on the patterns of language as it is actually used. Insofar as I believe linguists should try to account for usage, I could be accused of empiricism. But it should be evident throughout that I am no Skinnerian.

While I began this project a decade ago, until this year, it always took a back seat to something else. I spent eight of those years directing Stanford's undergraduate cognitive science major, called the Symbolic Systems program. For four of those years, I also served as an associate dean. I taught throughout those years, and I coauthored an introductory syntax textbook with Ivan Sag. In addition, I did research on several unrelated—or only marginally related—topics, including negative inversion in African American vernacular English, idioms, and disfluencies in speech.

Happily, a year's fellowship at the Stanford Humanities Center has permitted me to pull together the pieces of research on constituent ordering that I and my collaborators have done over the past decade. Writing it up in one place made it clear where the holes were, and that has led to some additional studies. Much of the material in the first four chapters has appeared elsewhere, specifically, Wasow (1997a, b), Arnold, et al (2000), and Wasow and Arnold (in press), but all chapters contain some new material. Chapters 5 and 6 elaborate on the metatheoretical and methodological concerns I have described autobiographically in this preface. They attempt to articulate as clearly as possible how and why I have broken with the intellectual tradition in which I was raised. I hope they also convey my deep appreciation of the important contributions that have come out of that tradition.

Thomas Wasow
Stanford, California
April 2001

Introduction

1 An Example

Linguists generally classify English as a language with relatively fixed word order. Even a cursory examination of the sentences English speakers use, however, reveals that considerable variability in the ordering of constituents is possible, especially in the region following the main verb. Consider, for example, the following sentence from Steven Pinker's *The Language Instinct*, p. 131 (reformatted here for expository reasons):

1. In my laboratory we use it as an easily studied instance of mental grammar, allowing us to document
 - in great detail
 - the psychology of linguistic rules
 - from infancy to old age
 - in both normal and neurologically impaired people,
 - in much the same way that biologists focus on the fruit fly *Drosophila* to study the machinery of the genes.

The five bulleted phrases could, in principle, appear in 120 different possible orders. Pinker's choice of this particular ordering among so many possibilities was surely not arbitrary. As a good writer of English, he recognized

(presumably, without having to think about it consciously) that (1) sounds more natural than, e.g., (2):

2. In my laboratory we use it as an easily studied instance of mental grammar, allowing us to document, in both normal and neurologically impaired people, from infancy to old age, the psychology of linguistic rules, in great detail, in much the same way that biologists focus on the fruit fly *Drosophila* to study the machinery of the genes.

Other orderings sound still worse—to the point where some, such as (3), would probably be judged uninterpretable (or at least unacceptably awkward) by most readers.

3. In my laboratory we use it as an easily studied instance of mental grammar, allowing us to document, in both normal and neurologically impaired people, in much the same way that biologists focus on the fruit fly *Drosophila* to study the machinery of the genes, in great detail, from infancy to old age, the psychology of linguistic rules.

Whether distinctions among examples like (1)–(3) should be regarded as grammatical (that is, whether they are part of a competent speaker's knowledge of English) or whether they are a matter of style or performance (how a speaker puts grammatical knowledge to use) is a debatable question, to which I return in Chapter 5. But it is clear that speakers have robust preferences for certain orderings of constituents over others. The purpose of this monograph is to explore the linguistic and psycholinguistic bases for such preferences, and, in the process, to shed some light on the relationship between knowledge and use of language.

A first stab at an account of the preference for (1) might be to note that each bulleted phrase in (1) is at least as long (in words) as the preceding one. This is not true in (2) or (3). In fact, nondecreasing length is remarkably good predictor of the ordering of postverbal constituents in English: it holds in the overwhelming majority of the corpus materials I have studied—on the order of eighty to ninety percent. But it does not constitute an explanation; a full theory of ordering should provide reasons. A satisfactory explanation would relate the constraints on ordering to more basic factors, such as parsing strategies employed by listeners, innate principles that facilitate grammar acquisition, the informational structure of discourses, etc. Ideally, such an explanation

would also have something to say about the ten or twenty percent of the data not covered by the first stab.

2 Objectives

For expository purposes it will be convenient to adopt the following generalization, which (following Quirk, et al, 1972, 14.8) I will call ‘The Principle of End Weight’ (PEW):

4. Phrases are presented in order of increasing weight.

‘Weight’ is left undefined for now, though I am provisionally taking it to be a syntactically definable property. There are many definitions that make PEW generally true, but, as we shall see, probably none that can cover all cases.

One major objective of this monograph is to characterize grammatical weight so as to maximize the generality of PEW and provide the basis for an explanation of why PEW holds. To a large extent, this has already been done by Hawkins (1994). Most of the studies described here are consistent with Hawkins’s ‘performance theory of order’, but they also raise some questions about it and suggest ways in which it might be refined. The issue of the definition of weight is the topic of Chapter 2.

Another central goal of the present work is to examine factors that can explain why postverbal constituents in English sometimes occur in orders violating PEW. One that has received a great deal of attention over a period of many decades is what might be called information status: what role do the interpretations of different phrases play in the information the speaker wishes to convey to the listener? The literature on this subject is vast and confusing. I will not attempt to summarize it, nor to resolve most of the issues it raises. Rather, in Chapter 3, I consider the question of whether the syntactic and informational properties that have been identified as influencing ordering might simply be different ways of looking at a single factor.

While weight and information status are well-known factors influencing ordering, I have come across a number of others that have received little or no attention in the literature. Chapter 4 describes these and considers the reasons they might have this influence.

Chapters 5 and 6 examine the implications for syntactic research of the work described in Chapter 2–4. These implications concern both theory and methodology. What kinds of syntactic theories seem best suited to expressing the sorts of findings described in the earlier chapters? Do they, in fact, have any bearing on questions of syntactic theory? Why do I employ methods so differ-

ent from what generative grammarians standardly do? Chapter 5 focuses on questions of theory, especially whether probabilistic constraints have a place in a grammatical theory. In the process, I revisit some old arguments against the usefulness of statistical modeling in linguistics, and I consider why, after several decades of acceptance, these are now being ignored by so many researchers. Chapter 6 is concerned with methodology, addressing the relevance of usage data and psycholinguistic experiments to competence theories, as well as the ubiquitous use of judgment data in generative syntax.

3 Some Phenomena Illustrating Order Variation

The canonical ordering of elements within the English VP is summarized by Gazdar and Pullum (1981; 120) as (5):

5. $H < N'' < P'' < V''$

In their notation, this means that the head verb comes first, followed by any noun phrases, followed by any prepositional phrases, followed by any verb phrases or clauses.

Note that if weight is given almost any natural structural definition, (5) conforms to PEW. This is no coincidence. Since the verb is a single word, it is shorter and less complex than any phrasal constituents. Prepositional phrases and clauses normally contain NPs as proper parts, so it is natural to regard them as heavier than NPs. Finally, since any postverbal PP is necessarily contained in a clause but relatively few PPs contain clauses, it seems plausible to consider clauses heavier than PPs. In short, (5) is arguably an instance of PEW.

(6a) is an example (from the Brown corpus) of a sentence whose VP contains all four types of elements ordered as in (5); (6b–f) are the sentences resulting from reordering the postverbal constituents. While they clearly sound worse, to varying degrees, at least the first two would probably be judged as acceptable by most speakers.

- 6 a. told [the Rotary Club of Providence] [at its luncheon at the Sheraton-Biltmore Hotel] [that about half the people in the country want the 'welfare' type of government...]
 b. told [the Rotary Club of Providence][that about half the people in the country want the 'welfare' type of government...][at its luncheon at the Sheraton-Biltmore Hotel]

- c. told [at its luncheon at the Sheraton-Biltmore Hotel] [the Rotary Club of Providence] [that about half the people in the country want the ‘welfare’ type of government...]
- d. told [at its luncheon at the Sheraton-Biltmore Hotel] [that about half the people in the country want the ‘welfare’ type of government...] [the Rotary Club of Providence]
- e. told [that about half the people in the country want the ‘welfare’ type of government...] [the Rotary Club of Providence] [at its luncheon at the Sheraton-Biltmore Hotel]
- f. told [that about half the people in the country want the ‘welfare’ type of government...] [at its luncheon at the Sheraton-Biltmore Hotel] [the Rotary Club of Providence]

It seems, then, that English permits more freedom of ordering than is sometimes recognized. The remainder of this section lists some constructions that illustrate this point more clearly.

3.1 *Heavy NP Shift*

Clear exceptions to (5) are not hard to find. For example, (1) contains a PP (*in great detail*) that precedes an NP (*the psychology of linguistic rules*), violating one part of (5). In the generative literature, the occurrence of a nonNP constituent between a verb and a following NP is known variously as ‘Complex NP Shift’ (Ross 1967), ‘Heavy NP Shift’ (Kimball 1973), or simply ‘NP Shift’ (Larson 1988); sticking with the weight metaphor, I will adopt Kimball’s terminology (HNPS, for short). I hasten to add that the name should not be taken as an endorsement of the view that the phenomenon should be analyzed using a movement transformation (see Chapter 5 for relevant discussion).

As the name implies, the NPs in HNPS examples tend to be heavy, in the intuitive sense of long and complex. The bracketed NP in (7) is an example familiar to linguists:

7. The problem for the linguist, as well as for the child learning the language, is to determine from the data of performance [the underlying system of rules that has been mastered by the speaker-hearer and that he puts to use in actual performance].
[Chomsky, 1965; 4]

Putting the postverbal constituents in their canonical order, as in (8a), or replacing the heavy NP by a light one, as in (8b), produces a far less natural sounding sentence.

- 8 a. The problem for the linguist, as well as for the child learning the language, is to determine the underlying system of rules that has been mastered by the speaker-hearer and that he puts to use in actual performance from the data of performance.
 b. The problem for the linguist, as well as for the child learning the language, is to determine from the data of performance the rules.

As the name implies, HNPS usually exemplifies PEW; this is documented in the next chapter by means of corpus studies.

3.2 *Extraposition from NP*

Ross (1967, Chapter 1) discussed a putative transformation he calls 'Extraposition from NP', which could extract an S from inside an NP and move it to the end of the clause. (9) is one of his examples, in which, under his analysis, the relative clause, *which I had cleaned*, has been extraposed from its position within the subject NP.

9. A gun went off *which I had cleaned*.

A somewhat more general characterization is given by Quirk, et al (1985; 18.39):

Sometimes only part of an element is postponed. The most commonly affected part is the postmodification of a noun phrase....The postponement...results in a 'discontinuous' noun phrase...

(10) lists some of Quirk, et al's examples (italicization in original):

- 10 a. *A rumour circulated widely that he was secretly engaged to the Marchioness.*
 b. *A steering committee had been formed, consisting of Messrs Ogawa, Schultz, and Robinson.*
 c. We heard *the story* from his own lips *of how he was stranded for days without food.*
 d. ...is usually accompanied by *the assignment* to it *of a marked...focus*

In all such cases, a heavy constituent (subordinate clause, PP, or VP) is positioned at the end of its clause, and the nonfinal NP of which it is a modifier or complement is consequently lighter than it would be without the discontinuity. Both the final position of the (usually heavy) extraposed element and the lightening of the NP serve to increase the probability of satisfying PEW. Since discontinuous elements presumably add to the processing complexity of the language, the fact that such discontinuities are tolerated in this construction suggests that PEW plays a powerful role in constituent ordering.

3.3 *The Verb-Particle Construction*

Many verbs combine with a preposition or adverb to form what the generative literature generally refers to as a verb-particle combination. When such complex verbs are transitive, they characteristically allow the particle to occur either before or after the object NP, as in (11):

- 11 a. Pat picked up a book. ~ Pat picked a book up.
 b. Chris took away the gun. ~ Chris took the gun away.

The forms in which the particle is separated from the verb (what I will be referring to as the ‘split’ ordering) are *prima facie* counterexamples to PEW, since the particles are manifestly short, simple constituents. This is one reason to look for other factors influencing ordering.

It is, however, well-known that the weight of the object NP influences the preferred position of the particle. When the object is a personal pronoun, it must precede the particle, as illustrated in (12):

- 12 a. Pat picked it up.
 b. *Pat picked up it.

And when the object is heavy (by virtue of length and/or some other criterion of complexity), the postparticle position is preferred:

- 13 a. Pat picked up a very large mint-green hardcover book.
 b. ?Pat picked a very large mint-green hardcover book up.
 c. I figured out who died.
 d. ?I figured who died out.

The factors influencing particle position will be discussed at length in Chapters 2 and 4.

3.4 *The Dative Alternation*

Most English verbs that occur with two NP objects also occur in a construction with one of the two arguments realized as the object of a preposition (*to* or *for*, depending on the role of the NP). Paradigm examples of verbs exhibiting this alternation are *give* and *buy*, as illustrated in (14) and (15):

- 14 a. Chris gave the boy a book.
 b. Chris gave a book to the boy.
 15 a. Pat bought the boy a book.
 b. Pat bought a book for the boy.

Considerable attention has been devoted to the role of semantic and pragmatic factors in the choice between the double object and prepositional constructions in such examples (see, e.g., Green 1974, Oehrle 1976, Ertshik-Shir 1979, Dryer 1986, and Goldberg 1995). But there has been very little discussion of the structural factors typically identified with weight (that is, length and syntactic complexity) in connection with this choice (which I will call ‘the dative alternation’—DA for short). Aside from the prohibition against unstressed pronouns as the second object in the double object construction, the interaction of weight with DA seems to have gone unmentioned in the literature before Hawkins (1994: 212–214).

Despite this oversight, DA provides another example of PEW. This can be seen anecdotally by considering the verb *begrudge*, which is standardly cited as one that does not participate in the alternation in question, occurring only in the double object construction (Benson, et al, 1986: 22, and Levin, 1993: 47); however, when the double object construction would strongly violate PEW, *begrudge* does appear in the prepositional alternative.

- 16 a. But no one could begrudge its splendid facilities to a city
 which lost 16,000 of Armenia’s 25,000 dead on December 7,
 1988, and was half-ruined by the earthquake. (Hector Corpus)
 b. We don’t have to begrudge it to our children. (*Ah Wilderness!*,
 by Eugene O’Neill)

The effects of weight and other factors on DA is considered at length in the next three chapters.

3.5 *Multiple PPs*

When English verb phrases contain multiple PPs—as in (1)—there is characteristically a good deal of freedom in how to order them. This is illustrated in (17), in which both orders are fully acceptable.

- 17 a. Pat talked to many people about the movie.
 b. Pat talked about the movie to many people.

Hawkins (2000) investigated the relative ordering of postverbal PPs in a sample of English texts he selected. Although traditional handbooks of English grammar claim that the preferred ordering among PPs depends on their semantic function (viz., manner before place before time), Hawkins found no evidence in support of this claim. Instead, he found a strong weight effect (in keeping with PEW) and a weaker but significant effect of semantic dependencies between verbs and the PPs that follow them. Hawkins's findings are discussed at greater length in Chapter 4.

3.6 *The Locative Alternation*

Another, slightly more complex case is what Pinker (1989) calls the locative alternation, exemplified in (18) and (19):

- 18 a. Pat sprayed the wall with red paint.
 b. Pat sprayed red paint on the wall.
 19 a. Pat drained the radiator of the coolant.
 b. Pat drained the coolant from the radiator.

In this case, there are meaning differences between the variants that might dictate the choice of one over the other in a given context (see, e.g., Anderson 1971, Talmy 1976, Pinker 1989). Hence, while this might arguably be a case of postverbal constituent order variation in English, it is not part of the present study.

3.7 *Some Other Constructions*

Although the discussion in this monograph concentrates on the relative positions of elements following the verb (as the title implies), there are ordering alternations in English involving preverbal elements. One clear example is extraposition. In this construction, exemplified in the following examples taken from Erdmann (1988), the subject slot is filled by the dummy *it*, with a subordinate clause or VP appearing later in the sentence:

- 20 a. It is evident that one reform calls for another...

[*Daily Telegraph*, 3/13/75, 16: 7]

- b. It is dangerous to trust the word of even a scrupulous diarist...

[*Spectator*, 1/25/75, 93: 3]

Most analyses relate such examples to forms with clausal subjects, such as (21):

- 21 a. That one reform calls for another is evident.

- b. To trust the word of even a scrupulous diarist is dangerous.

Like the postverbal alternations, extraposition evidently conforms to PEW. Erdmann (1988) did a corpus study of adjectival predicates taking sentential arguments, either as subjects or in the extraposition construction. Of the 1101 examples he found, 1033 (93.8%) exhibited extraposition. Erdmann also divided the adjectival predicates into light and heavy, using his definition of 'heavy' (namely, not head-final—one of several definitions I examine in the next chapter), and looked for a relationship between predicate heaviness and extraposition. When the predicate was light, 95.6% of the examples had the subordinate clause in extraposed position; when the predicate was heavy, the corresponding figure was only 82.1%. This difference is highly significant ($\chi^2 = 39.819, p < .001$)¹.

Another type of example Erdmann (1988) studied was *it*-clefts. These are examples like the following (taken from Erdmann's article):

- 22 a. It is, of course, unemployment and its attendant consequences that politicians fear most.

[*Spectator*, 2/15/75, 74: 13]

- b. It is the colleges which are the outstanding characteristics of Oxford and Cambridge.

[Burgess, *A Guide to English Schools*: 186]

Erdmann (1988) also discusses what he calls '*what*-clefts' (often referred to in the literature as 'pseudoclefts' or '*wh*-clefts'). Two of his examples are given in (23):

1. Erdmann organized his data somewhat differently from the way I am presenting them. Hence, the numbers that appear in this paragraph, while trivially computed from what he wrote, do not actually appear in his paper.

- 23 a. What was needed was a political breakthrough.

[*Daily Telegraph*, 3/21/75, 32: 4]

- b. What we want is the chance to earn a living, as we know we can.

[*Daily Telegraph*, 3/25/75, 2: 7]

Interestingly, pseudoclefts look at first like a construction that runs counter to PEW. The characteristic pseudocleft structure is: *wh*-clause-copula-XP, where XP can be any major phrasal category, and the *wh*-clause contains a gap that could be filled by XP. Assuming that clauses are normally heavier than other types of phrases, this sequence has its heaviest constituent at the beginning, contrary to PEW.

However, when the XP is heavy and corresponds to a subject gap in the *wh*-clause, the pseudocleft construction can be used as a way of placing a heavy subject at the end of the sentence, in conformity with PEW. (24) is an example, again taken from Erdmann (1988), with bracketing of the heavy NP added:

24. What was to Henry's credit was [his whole state of mind, which encompassed the welfare of Ireland, as something quite separate from that of England, and at the same time perfectly desirable]. (Fraser, *Cromwell*: 579)

The alternative formulation in (25) would begin with a very long and complex NP, in strong violation of PEW.

25. His whole state of mind, which encompassed the welfare of Ireland, as something quite separate from that of England, and at the same time perfectly desirable was to Henry's credit.

Erdmann examined 251 examples of pseudoclefts with NPs in the postcopular position and found that 89.0% of the NPs corresponding to subject gaps were heavy (by his definition), whereas only 78.4% of the NPs corresponding to object gaps were heavy. This result is statistically significant ($p < .05$). Thus, it seems that the pseudocleft construction, though superficially a counterexample to PEW, provides some evidence in its support, as well.

In short, English exhibits numerous constructions in which constituent ordering can be varied. Where the alternations in form correlate with an obvious difference in interpretation, the explanation for the choice of one form over another seems obvious: the speaker/writer chooses the form that conveys the intended meaning most accurately. But in cases like those listed above

(with the possible exception of the locative alternation), whatever meaning differences exist between the two orderings are quite subtle. Minimally, the truth conditions are not affected by the choice of order. In such cases, we have already seen that a structural notion of ‘weight’ probably influences order. But this notion needs to be characterized more precisely, and the roles (if any) of other factors need to be investigated.

4 Questions to be Addressed

There are three main questions I address in this monograph. The first is the subject of the next three chapters, which present the findings of various empirical studies. The second is addressed in those chapters as well, but receives more direct attention in the final two chapters, which are also the locus of discussion of the third question.

4.1 *What factors influence postverbal ordering in English?*

I take it as established that PEW is correct. However, making that a truly substantive claim requires that the notion of ‘weight’ be given more content. Once that is done, one can ask what, if any, other factors play a role. Hawkins (1994) suggests that structural factors will suffice to account for ordering, but his later work (especially Hawkins (2001)) clearly recognizes a role for nonstructural factors. My empirical investigations lead me to conclude that there are indeed a number of distinct factors that influence constituent ordering.

This conclusion leads naturally to another question: Is there an overarching generalization that links the various factors that influence ordering? It would be preferable on methodological grounds to find a single property that can account for all the factors I have found that influence ordering. But the viability of such an account is clearly dependent on the evidence, and to date I have not been able to find a unifying property.

4.2 *Why do the factors influencing ordering have the effects they do?*

The question of why certain factors affect constituent ordering immediately raises the fundamental question of what would count as an answer. Many sorts of accounts are regarded as explanations in linguistics, including histories, formal simplifications (sometimes augmented with learnability arguments), and processing models.

In the case of PEW, it seems most natural to seek psycholinguistic explanations. The typical reaction to sentences that strongly violate PEW (such as (8a) or (8b) above) is that they are awkward and hard to analyze. Hence, it is not surprising that a number of investigators (e.g., Bever 1970, Kimball 1973, Fra-

zier & Fodor 1978, Hawkins 1990, 1994) have proposed that PEW should be explained in terms of the architecture of or strategies employed by the human parser.

Assuming that PEW does generally facilitate parsing, one can still ask why speakers put heavy elements later. Where the canonical ordering (that is, the one in (5)) conforms to PEW, speakers are simply following the rules of their grammar; arguably, in this case the ordering that is easiest to parse has become grammaticalized. For noncanonical orderings, such as heavy NP shift, however, a parsing explanation requires that we ascribe a good deal of charity to the speaker. It assumes that the speaker goes to the extra trouble of selecting a marked construction in order to make life easier for the addressee. In Chapter 2, I argue that saving heavy elements for the ends of sentences facilitates utterance planning and production. Hence, the explanation for PEW should not rest entirely on parsing considerations.

For the other factors influencing ordering, similar functional explanations should be sought. Alas, my success in finding such explanations for all the factors I know of has been only very partial.

4.3 *What are the implications of these findings for theoretical syntax?*

Some of the phenomena listed in section 3 above have been discussed at length in the generative literature of the past four decades. But generative analyses have made little or no reference to the factors discussed in Chapters 2–4. This is because generative grammarians have traditionally been concerned only with what forms are possible, not with the reasons for choosing among various grammatically well-formed alternatives.

In Chapter 5, I suggest that this may be a mistake—that categorical constraints on possible forms may simply be the limiting case of preferences. Moreover, I argue that certain architectures for a theory of grammar are better suited than others to deal with noncategorical preferences.

This, in turn, raises methodological issues. It is customary in the generative literature to assume that native speakers can provide categorical introspective judgments of (absolute or relative) acceptability that are sufficient to justify an analysis. I employ such judgments occasionally in this monograph, but most of my argumentation is based on data from usage—that is, from examples that happened to catch my eye (or ear), from on-line corpora, and from production experiments conducted in conjunction with Jennifer Arnold and others. Such data are rarely categorical, and often provide counterexamples to categorical claims based on introspection. In addition, where acceptability judgments strike me as too delicate for the intuitions of one speaker to be reliable, I employ questionnaires administered to multiple speakers. In Chapter 6, I dis-

cuss methodological issues underlying my decision to rely so heavily on quantitative data, and consider the reasons that generative grammarians have traditionally avoided them.

5 Why Only English Data?

Weight effects have been documented for a substantial number of languages (see Behaghel 1909/10 and Hawkins 1994). Yet the present work is based entirely on English evidence. It is natural to ask why I did not investigate other languages. There are several reasons, based largely on practical considerations.

As mentioned above, my arguments below are supported largely by corpus data. In particular, I made use of the parsed corpora in the Penn Treebank (see Marcus, et al, 1993). This made it possible to do quantitative analyses of substantial numbers of examples based on details of tree structures. Comparable parsed corpora (especially, parsed corpora of speech) are not available for any other language. Even if they were, many of the analyses of the sort I present in chapters 2–4 can only be carried out by someone with an excellent command of the languages in question.

Moreover, the wealth of information that is contained in these corpora is enormous. The studies reported on in this monograph only scratch the surface of what could be extracted. I consequently found it most productive to focus on English, leaving analogous studies of other languages for future research.

Another reason for restricting attention to English has to do with cross-linguistic variation in weight effects. Hawkins (1994; 66–67) discovered that some left-branching languages, such as Japanese and Korean, exhibit a tendency to put heavy constituents at the beginnings of sentences, rather than at their ends. He formulated his parsing-based theory of weight effects so that it predicts this difference, but a relatively minor change in the theory could make it predict the opposite. I remain skeptical of his account of the difference, though I confess that I have nothing better to offer. Under these circumstances, it seems most prudent for me to restrict my attention to English.