Phillips and Gibson


End-Weight from the Speaker’s Perspective

Thomas Wasow 1,2

Explanations of the tendency to put long, complex constituents at the ends of sentences ("end-weight") usually take the listener’s perspective, claiming it facilitates parsing. I argue for a speaker-oriented explanation of end-weight, based on how it facilitates utterance planning. Parsing is facilitated when a more complete structure as possible can be determined early in the string, but production’s easiest when options for how to continue are kept open. That is, listeners should prefer early commitment and speakers should prefer late commitment. Corpus data show that different verbs exhibit different rates of word-order variation that are systematically related to differences in subcategorization possibilities in just the way predicted by a strategy of late commitment. Thus, a speaker-based account of lexical preferences vs. word ordering does a better job of explaining variation in weight effects than a listener-based account.

This is a longer version of a talk by the same name given at the Ninth Annual CUNY Conference on Human Sentence Processing, held in New York on March 21–23, 1986. Most of the results presented here are also reported in Wasow (in press). I am very grateful to Herb Clark for extensive advice and comments in the preparation of that paper. I also received valuable help from Jennifer Arnold, Bruce Bueno de Mesquita, Chris Culy, Adlele Goldberg, Georgia Green, Jack Hawkins, Maryellen MacDonald, Dics Osterle, John Perry, John Rickford, Ewart Thomas, and Arnold Zwicky. None of them is responsible for any of this paper’s shortcomings. Max Klee provided crucial assistance with the organization and analysis of the data. I have also benefited enormously from the use of the resources of the Center for the Study of Language and Information, in particular the on-line corpora from which most of the data in this paper were drawn. Research on this paper was supported in part by National Science Foundation (NSF) grant SBR-9109612.

1 Stanford University, Stanford, California 94305-2150.

2 Address all correspondence to Thomas Wasow, Department of Linguistics, Stanford University, Stanford, California 94305-2150.

347
INTRODUCTION

Consider 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.

In (1), each of the five bulleted phrases following the verb *document* is at least as long (measured in number of words) as the immediately preceding one, and the grammatical complexity seems intuitively to be increasing, as well. Notice, in particular, that the direct object noun phrase (NP), *the psychology of linguistic rules*, occurs after the prepositional phrase (PP) in great detail. This is an instance of what is sometimes called heavy NP shift (henceforth, HNPS). HNPS is one manifestation of a general tendency in many languages for constituents to occur in order of increasing size or complexity—what I call “grammatical weight.”

This observation is an old one, standardly attributed to Otto Behaghel (1909/10). Example (2) shows one of Behaghel's formulations of the generalization.

(2) *Das Gesetz der Wachsenden Glieder*: Von zwei Gliedern von verschiedenem Umfang steig der umfangreichere nach. (Behaghel, 1930, p. 85)

Following Quirk, Greenbaum, Leech, & Svartvik (1972), I will use the terms *end-weight* or *weight effects* to refer to this tendency.

In Wasow (in press), I considered at some length the question of how grammatical weight should be characterized. I examined three proposed structural weight measures in detail: number of words, number of nodes, and number of phrasal nodes. In the end, I concluded that these measures are indistinguishable as predictors of constituent order. This is not really surprising: In a large body of corpus data relating to three English constructions, I found the correlation coefficients for pairs of these measures ranged between .94 and .98. Moreover, all three measures are rather good predictors of constituent ordering, with about 90% of the examples in my collection conforming to end-weight.

The principal concern of the present paper is explaining why weight effects exist. That is, why should heavy constituents come late? One answer that has been offered repeatedly in the literature (e.g., Bever, 1970; Frazier & Fodor, 1978; Hawkins, 1990, 1994;Kimball, 1973) is that end-weight makes parsing easier. The intuitive idea is that saving complex elements for the end of the sentence permits the parser to determine the high-level structure of the sentence more quickly, which in turn aids in the analysis of the remainder of the sentence.

In this paper, I consider this idea and compare it to an alternative hypothesis, namely, that weight effects exist primarily to facilitate utterance planning and production.

A PARsING ACCOUNT OF WEIGHT EFFECTS

**Hawkins's Idea**

The intuition that the primary function of end-weight is to simplify parsing has been developed in considerable detail by Hawkins (1994). He wrote:

[C]onstituents occur in the orders they do so that syntactic groupings and their immediate constituents (ICs) can be recognized (and produced) as rapidly and efficiently as possible in language performance. Different orderings of elements result in more or less rapid IC recognition (Hawkins, 1994, p. 57)

Despite the parenthetical allusion to production, Hawkins's elaboration of this idea focused entirely on parsing. This is clear from his discussion of HNPS, using the following examples:

(5) a. I gave Mary the valuable book that was extremely difficult to find.

```
1 2 3 4 5 6 7 8 9 10 11
```

I employ the following standard abbreviations: “VP” for Verb Phrase; “NP” for Noun Phrase; “PP” for Prepositional Phrase; and “V” for Verb.

---

3 Translation: "The law of the growing constituents." Of two constituents of different size, the larger one follows the smaller one. (Thanks to Hinrich Schütze for help with the translation—TW.)
End-Weight from the Speaker's Perspective

Consider, for example, the following (reformatted) sentence, taken from the Aligned-Hansard corpus, a collection of transcripts from the Canadian Parliament:

(5) I share
- with that part of the population of which I am a member, and
- those others who are identifiable by colour, sex or disability,
- the feeling that we are experiencing today a day of tragedy which I would attribute, as my colleague did previously, to the inability of Hcn. Members opposite to put themselves in the shoes of those who have suffered from discrimination in the past and who will suffer from the Government’s failure to attack systemic discrimination today.

Hawkins’s theory would seem to require that the speaker have both phrases fully planned at the point of uttering with, so that they could be produced in order of increasing weight. It is far more likely, however, that the actual formulation of the phrases took place during the course of uttering the sentence, much of it long after with was uttered.

In fact, speakers often begin uttering sentences before they have decided on the formulation of all the constituents. This is evident from the pervasiveness of disfluencies such as uh, um, repetitions, false starts, and tie like. Many of these disfluencies occur in mid-sentence, indicating that sentence planning is not complete when utterance of the sentence begins. Moreover, an analysis of the distribution of different types of disfluencies suggests that it is not only word selection, but also the choice of grammatical construction, that is ongoing during speech. See Clark (1994), Clark (1996), Clark and Wasow (1996), Fox Tree and Clark (1994), and Levelt (1989) for further discussion of disfluencies and utterance planning.

The complexity of the structures needed to express different components of a thought may well influence constituent ordering, but the determination of order need not wait until the constituents are fully formulated. I contend that speakers select the ordering they do primarily because it facilitates utterance planning and production, not because it will help the listener in parsing. To test this claim, I needed to find examples of ordering preferences where the interests of the speaker diverge from those of the listener. These are relatively rare, since what makes things easier to produce usually makes them easier to analyze, as well.\textsuperscript{5} I have, however, found some such cases, and they support a production-based account of weight phenomena.

\textsuperscript{5} See Gibson and Pearlmutter (1994) for some discussion of this issue, and of the relationship between corpus frequency and psycholinguistic complexity.
EVIDENCE FOR THE ROLE OF PLANNING IN WEIGHT EFFECTS

Collocations and Heavy NP Shift:

First, an anecdote: San Francisco Giants' announcer Hank Greenwald regularly uses (6) (and the equivalent sentences with other players' names in final position).

(6) That will bring to the plate Barry Bonds

To the plate is undeniably longer and more complex than Barry Bonds. There is no apparent reason why it should be easier to parse (6) than it is to parse That will bring Barry Bonds to the plate. But the noncanonical "shifted" order is clearly useful to Greenwald, since it gives him an extra second or so to check his scoreboard and see who the next batter is. The string bring to the plate is a conventionalized collocation in Greenwald's speech, which he naturally produces without discontinuities. Yet it remains semantically transparent, in the sense that an English-speaking listener familiar with baseball would have no trouble interpreting it on first encounter. Consequently, the shifted constituent order is of use on y to Greenwald, not to the listener.

Many English collocations consist of a transitive verb plus a PP, with an open slot for the direct object: NP. Some of these (e.g., bring to an end) are merely conventionalized ways of saying something that is semantically transparent, whereas others (e.g., bring to bear) are semantically more opaque.

I analyzed 827 sentences from the Aligned-Hansard corpus containing the following word sequences: attribute...to, bring...to, obtain...from, share...with, and take...into. Each sentence was coded according to whether it exhibited HNPS, whether the verb was part of a transitive V-PP collocation, and, if so, whether that collocation was semantically transparent.

Out of those sentences, 439 (53%) of the V-PP pairings were collocations, of which 237 (54%) exhibited HNPS. In contrast, among the noncollocational VPs, only 59 (15%) exhibited HNPS, χ²(1) = 137.33, p < .001.

The classification of the collocational examples as either transparent or opaque was based on the judgment that a transparent collocation should be comprehensible to a speaker who knows the literal meaning of each of the words in it, but who has never before encountered them in this combination.

The lion's share of the examples involved the collocations bring X's attention, bring to an end/phase, and take into account/consideration; after some deliberation, I coded bring X's attention and bring to an end/phase as transparent and take into account/consideration as opaque. Of the 192 transparent collocations, 90 (47%) exhibited HNPS, and of the 247 opaque collocations, 147 (60%) exhibited HNPS.

When the verb and PP form an opaque collocation, HNPS is useful to both speaker and listener. The speaker benefits because, having decided on the collocation, uttering all of it immediately provides extra time and reduces memory load. The listener benefits because the correct meaning can only be assigned to the verb once the PP is uttered, so early occurrence of the PP facilitates incremental interpretation. When the verb and PP form a transparent collocation, the advantage to the speaker is maintained. That is, once a collocation has been selected by the speaker, producing it immediately buys time to plan the direct object: NP, irrespective of whether the collocation is transparent or opaque. For the listener, on the other hand, the tasks of parsing and interpretation are the same in a transparent collocation as in a noncollocational VP. From the listener's perspective, there would be no reason to expect that transparent collocations would have a higher rate of HNPS than noncollocational VPs.

In short, in deciding between a speaker-based and listener-based account of weight effects, one crucial test is the behavior of transparent collocations. The speaker-based account predicts that they should behave like opaque collocations; the listener-based account predicts that they should behave like non-collocational VPs.

Figure 1 suggests that the behavior of transparent collocations is intermediate, but much closer to that of opaque collocations than noncollocations. Two-way comparisons support this impression. The difference between HNPS rates in transparent collocations and noncollocational VPs is highly significant, χ²(1) = 68.03, p < .001. The difference in HNPS rates between transparent and opaque collocations is also significant, χ²(1) = 6.95, p < .01, suggesting that the listener's perspective is not irrelevant to constituent ordering. However, the relative magnitude of the effects lends support to the idea that utterance planning is the more important factor in weight effects.

* I am grateful to John Perry for this observation.

† The sentences were found by searching for utterances containing any of these verb-preposition sequences. In the course of coding, spurious examples (such as those with bring...to where to was an infinitive marker) were collected. While this is a sample selected to have a high rate of HNPS, there is no reason to believe that it is unrepresentative with respect to the difference in shifting rates between collocational and non-collocational VPs.

‡ There were few cases where this classification was unclear. In those cases, I consulted Spears (1993).
End-Weight from the Speaker's Perspective

The verbs listed are the ones which appear in the relevant constructions most frequently in the parsed version of the Brown corpus, available as part of the Treebank (see Marcus, Santorini, & Marcinkiewicz, 1993).

The difference between these two verb classes is illustrated in (8) and (9):

(8) $V_i$
   a. *Pat brought to the party.
   b. Pat brought a box with a ribbon around it to the party.
   c. Pat brought to the party a box with a ribbon around it.

(9) $V_s$
   a. Pat wrote on the blackboard.
   b. Pat wrote something about Chris on the blackboard.
   c. Pat wrote on the blackboard something about Chris.

The difference in grammaticality between (8a) and (9a) illustrates the defining difference between the two verb classes. The arrows in the (8b) and (9b) and (8c) and (9c) sentences indicate the point of commitment. That is, these are the points at which the speaker has committed herself to producing a VP containing both NP and PP as immediate constituents. They are also the points at which the listener can be certain that the VP will contain both an NP and a PP.

In (8), it is clear that HNPS (in 8c) makes the commitment earlier. Hence, from the listener's perspective, it would be helpful (for the reasons

---

Lexical Dispositions

The most direct conflict I have found between interests of speaker and listener occurred with respect to what I will call point of commitment. For the listener, the more predictable the remainder of the sentence, the better, for fewer possible continuations compatible with the string at any given point entail less load on memory and less work for the parser later on. For the speaker, on the other hand, keeping options open postpones decision making, which reduces the amount of planning needed and gives the speaker more time to formulate and articulate thoughts. This, in turn, should minimize the chances of having to correct or abort an utterance.

Hence, speakers benefit from keeping options open, but the listener's task is simplified if sentences are highly predictable. In other words, listeners should prefer early commitment, and speakers should prefer late commitment.

To test the choice between early and late commitment, I examined two verb classes, described and exemplified in (7).

(7) • $V_i$ ("transitive verbs") require NP objects in all their subcategorizations.
   Examples: bring, carry, make, place, put, set, take

• $V_s$ ("prepositional verbs") can occur with NP objects but also have uses with an immediately following PP and no NP object.
   Examples: add, build, call, draw, give, hold, leave, see, show, write

---

* More specifically, I identified all verbs that exhibited HNPS at least four times and occurred with both an NP and a PP as sisters (in either order) a total of at least 100 times. Excluding be and save (which tend to be highly atypical in their behavior), there are 21 such verbs: add, become, bring, build, call, carry, consider, draw, find, give, hold, keep, leave, make, place, put, see, sex, show, take, and write. With help from Benson, Benson, & Isar (1986), a dictionary with extensive subcategorization information, I classified these 21 verbs as follows:

$V_i$: bring, carry, make, place, put, set, take

$V_s$: add, build, call, draw, give, hold, leave, see, show, write

Others: become, consider, find, keep

Some aspects of this classification might be questioned. For example, some verbs in $V_i$ occur in idioms with no following NP, as in I really took to Sandy and Pat made me a monkey; likewise, keep can occur in idioms with an immediately following PP, as in You always keep to yourself. The classification was based on nonidiomatic uses. Another potential objection is the assignment of give to $V_s$ rather than to $V_i$; this was on the basis of examples like I gave at the office and Give the United Way.
the grammaticality of (9a) together with the relatively marked character of HNPS would lead the parser to temporarily misanalyze the occurrence of \( \text{wrote} \) in (9c) as intransitive.

Summing up, then, the listener’s perspective predicts an elevated HNPS rate for \( V_t \) and a possibly depressed HNPS rate for \( V_p \). The speaker’s perspective predicts a depressed HNPS rate for \( V_t \) and a normal HNPS rate for \( V_p \).

The actual rates of occurrence of HNPS for \( V_t \) and \( V_p \) are depicted in Fig. 2,\(^{10}\) based on a total of 2,025 sentences of the relevant types involving these verbs.\(^{11}\) The predictions made from the speaker’s perspective are supported, and those made from the listener’s perspective are not.\(^{12}\) The HNPS rate for \( V_t \) is very close to that of the corpus overall (6.4%), and the rate for \( V_p \) is significantly higher.

Since my explanation of the difference between HNPS rates in \( V_t \) and \( V_p \) is based on the planning process during speech, the predictions should be tested against spoken data. With this in mind, I analyzed analogous data from the Switchboard corpus, a collection of telephone conversations recorded and transcribed by Texas Instruments.\(^{13}\) As in the

---

\(^{10}\) The HNPS rate for the Others verb class is 10.8%, which is higher than either \( V_t \) or \( V_p \), though the difference between \( V_t \) and Others is not statistically significant.

\(^{11}\) If give were to be reclassified as a member of \( V_t \) (see footnote 8), the results would be much stronger: \( \chi^2(1) = 21.05, p < .001 \).

\(^{12}\) Maryellen MacDonald (personal communication) has suggested that the verbs I examined should be divided up somewhat differently. Following Stalings, MacDonald & O’Searghda (1995), she proposed that verbs that can take either NP objects or sentential complements (together with a PP) will have an enhanced disposition to shift. Note that most of the verbs in my \( V_p \) class are of this type, along with \text{find}, \text{keep}, and \text{consider} from the Others class. Reclassifying as she has suggested, however, yields no statistically significant correlations between verb class and HNPS. Nevertheless, for present purposes, it would not matter even if her hypothesis proved superior, because both explanations rely on a speaker-oriented account of why heavy NPs tend to occur late.

\(^{13}\) Specifically, utterances were extracted in which one of the verbs in question occurred, immediately followed by a preposition, a determiner, or a pronoun. Clearly, this procedure failed to capture many immediately postverbal NPs, such as bare plurals or proper names. Since there is no reason to believe that the \( vO \) verb classes would differ in their disposition to take such objects, this should not affect the results. All examples with prepositions adjacent to the verb were inspected, and those exhibiting HNPS were retained. There were only 42 of these. Of the almost 6,000 utterances with determiners or pronouns immediately following the verb, 400 randomly chosen examples (200 from each verb class) were inspected, and the number of genuine \( V \) NP PP sequences in each class was extrapolated from these samples.
written data, HNPS occurred approximately twice as frequently in \( V_s \) as in \( V_n \). See Fig. 3.

A similar argument for explaining weight effects from the speaker’s perspective can be constructed on the basis of the dative alternation (DA). Verbs that undergo DA can be divided into the following two classes:

10. \( V_s \) (for “sentential verbs”): Dative alternation verbs which also have subcategorizations in which they are followed by an NP and a that-clause or an infinitival VP.
Examples: offer, show, teach, sell, write

11. \( V_n \) (for “nonsentential verbs”): Dative alternation verbs without such subcategorizations.
Examples: assign, bring, give, hand, pay, send, take

Again, the verbs listed represent those that occur most frequently in the constructions in question in the Brown corpus.

The NP that occurs between \( V_s \) verbs and the following clause or VP has the same semantic role (namely, goal) as the first object in the double-object construction. Hence, with \( V_s \) verbs, the production of an immediately postverbal goal NP leaves open two options for continuation: an NP object or verbal complement (S or VP). With \( V_n \) verbs, on the other hand, an immediately postverbal goal NP must be followed by a second (theme) object NP. This is illustrated in (11) and (12).

11 \( V_s \)

a. Pat told a story to the children.
b. Pat told the children a story.
c. Pat told the children that it was bedtime.

12 \( V_n \)

a. Pat assigned a problem to the students.
b. Pat assigned the students a problem.
c. Pat assigned the students that they do a problem.

For verbs in \( V_s \), production of the goal NP leaves open two possible continuations, illustrated in (11b) and (11c). Hence, the double-object construction (11b) constitutes late commitment. For verbs in \( V_n \), on the other hand,

It is natural to wonder why the HNPS rates in the Switchboard corpus are so much lower than those in the Brown corpus. I have not investigated this question systematically, so I can only speculate. One possibility has to do with the circumstances of the conversations in the Switchboard corpus; the conversations were between people who did not know each other and were on topics from an assigned list; further, the interlocutors knew they were being recorded. Consequently, it seems likely that the sort of informational factors that might play a role in HNPS (postposing an NP to highlight its discourse significance) may have been reduced.

the goal NP must be followed by another NP and the double-object construction (12b) is hence an early commitment. Consequently, the speaker’s perspective predicts a higher rate of double-object construction for \( V_s \), whereas the listener’s perspective predicts a higher rate of the double-object construction for \( V_n \).

Testing these predictions against the twelve DA verbs listed in (10) again strongly supports the speaker’s perspective. The results from both corpora, based on 1,046 examples from Brown and 2,799 from Switchboard, are exhibited in Fig. 4.

CONCLUSIONS

The data presented in this paper provide evidence for the following two novel claims:

• A production-based explanation for weight effects (specifically, in terms of utterance planning) can account for data that would be mysterious under a parsing-based explanation.

• There is considerable, often systematic, lexical variation in weight effects, related to the subcategorization possibilities of the verbs being used. This argues against trying to explain grammatical weight solely in terms of tree configurations.

![Fig. 4. Dative alternation (DA) rates by verb class \( V_n \) = nonsentential verbs; \( V_s \) = sentential verbs. \( \chi^2(1) = 52.14, p < .001 \) for Brown corpus; \( \chi^2(1) = 180.91, p < .001 \) for Switchboard corpus.](image-url)
The evidence presented for the latter claim is in direct conflict with the following claim by Hawkins (1994, p. 68):

It does not appear as if any contextual biases or predictions are of much use, or even desirable, in anticipating forthcoming structure. ... In head-initial languages, context effects could only remove the motive for rearranging [immediate] constituents: if the PP could be anticipated in advance of its appearance, there would be less reason to postpone the heavy NP. ... Since the preferences and dispreferences to be documented in this book seem to follow the dictates of structural length regardless of the contextual predictability of the relevant ICs ... this suggests that the human parser is informationally encapsulated, in the sense of Jerry Fodor.

I close with two methodological observations. The first is that usage data can be a useful source of insight into language processing. The second is that psycholinguistic explanations of grammatical phenomena should not focus one-sidedly on the listener. Language use is characteristically a joint activity. If we wish to explain aspects of language structure in terms of performance, we should look at the tasks facing both participants, and at how the participants carry out those tasks.

REFERENCES


