2. GENERATIVE GRAMMAR AND FUNCTIONAL EXPLANATION

(1) **The Autonomy of Syntax:**
The syntactic rules and principles of a language are formulated without reference to meaning, discourse, or language use.

(2) A typical functionalist remark about the compatibility of generative grammar and functional explanation:
   The autonomy of syntax cuts off [sentence structure] from the pressures of communicative function. In the [formalist] vision, language is pure and autonomous, unconstrained and unshaped by purpose or function. (Bates and MacWhinney, 1989: 5).

(3) Chomsky does not exclude the idea that grammars are shaped in part by function:
   Surely there are significant connections between structure and function; this is not and has never been in doubt. ... Searle argues that ‘it is reasonable to suppose that the needs of communication influenced [language] structure’. I agree. (Chomsky, 1975: 56-58).
   See also his remarks in Chomsky, 1995 on how movement rules exist to aid processing and discourse.

(4) Other systems are both autonomous (in the relevant sense), yet explained functionally:
   a. Chess and other games.
   b. Our bodily organs.

(5) The structure-dependence of syntactic rules: Sentences are analyzed in terms of abstract phrases and rules/processes refer to these phrases

(6) A functional explanation for structure-dependence:
   … in general structure-independent operations are difficult for people to perform on temporally sequenced strings; for instance, it is difficult to give a fluent reading of the alphabet, or a well-known telephone number, backwards without practice. (Comrie, 1989: 23)

**To demonstrate the autonomy of syntax one must show that:**

(7) a. there exists an extensive set of purely formal generalizations orthogonal to generalizations governing meaning or discourse.
   b. these generalizations ‘interlock’ in a system

(8) a. A set of formal generalizations (FoGs):
   FoG₁ FoG₂ FoG₃ FoG₄ FoG₅ etc.
   b. that ‘interlock’ in a system:
Displaced *wh*-elements

(9)  a. Who did you see?
    b. The book *which* I read

(10) Some descriptive generalizations about displaced *wh*-elements:
    a. Displacement is almost always to the left rather than to the right.
    b. Displacement is far more common in VO than in OV languages:
       - 71% of V-final languages lack *Wh*-Movement
       - 42% of V-medial languages lack *Wh*-Movement
       - 16% of V-initial languages lack *Wh*-Movement
       (Dryer, 1991)
    c. Displacement is subject to locality conditions (e.g. Subjacency)
       i. What did you put on the table?
       ii. What did you ask John to put on the table?
       iii. Where did you ask John to put the book?
       iv. *What did you wonder where John put?

(10)’ These generalizations have functional explanations:
    a.'  i. *Wh*-expressions set up operator-variable structures. It is natural that operators should precede variables (since variables depend on operators — see Hawkins, 2001).
       ii. *Wh*-expressions function as scope markers. It is natural that a scope marker should precede the element(s) in its domain (Hawkins, 2001).
       iii. *Wh*-question elements focus on a single element demanding the hearer’s attention. It is natural that conversationally ‘urgent’ elements should appear early in the sentence (Givón, 1988).
       iv. *Wh*-relative elements are anaphoric to the head of the relative clause. It is natural for anaphoric elements to follow their antecedents (Hawkins, 2001) and to be close to them (Haiman, 1985). So movement will be to the left, not to the right.
    b.' Hawkins, 1995: 95 notes that heads, in general, are the best identifiers of their subcategorized arguments. If one hears the verb *give*, for example, one is primed to expect two associated internal arguments, one representing a recipient and the other an object undergoing transfer. On the other hand, a human NP might or might not be a recipient and an inanimate NP might or not be an object undergoing transfer. Hence, if arguments precede their heads, as they do in OV languages, extra cues are useful to identify their thematic status. Such can be accomplished by keeping them contiguous to the head (that is, by restricting their movement possibilities) and / or by endowing them with case marking that uniquely identifies their thematic role or helps to narrow down the possibilities.
c. Processing-based explanations for the existence of Subjacency go back at least to Erteschik-Shir, 1973 (for a discourse-processing account) and to Berwick and Weinberg, 1984 (for a purely syntactic parsing-based account).

So, given that *wh*-fronting has a functional explanation, do we still need the apparatus of formal grammar to characterize it? Yes, for several reasons:

(11) a. Functions help to restrict the positing of *wh*-elements, but they do not determine those positions. Hence we need formal statements characterizing those positions.
b. *Wh*-expressions in particular languages partake of a formal system with other (formal) elements — that is, there are purely formal generalizations in language (and individual languages) only remotely, if at all, related to functional generalizations.c. Functions can conflict with each other. Autonomy is a natural response to conflicting functions.

In the simplest cases, fronted *wh* is a scope marker and an operator:

(12) What did Mary eat? = for what *x*, Mary ate *x*

Appositive relatives are not operator-variable structures:

(13) Ralph Nader, who several million Americans voted for, lost the election.

Absence of weak crossover effects with appositive relatives (Lasnik and Stowell, 1991):

(14) a. Gerald, who his mother loves, is a nice guy.
b. This book, which its author wrote, last week, is a hit.

But ‘*wh*’ not always a scope marker:

(15) a. *niŋ-ne* aarø talli (Malayalam — Jayaseelan, 2002)
   you-acc. who beat (past)
   ‘Who beat you?’
b. *aarø* *niŋ-ne* talli
   who you-acc beat (past)

Partial *Wh*-Movement in German and Romani (McDaniel, 1989):

(16) Was *i* glaubt [Ip Hans] [Cp mit wem] [Ip Jakob jetzt t, spricht]?  
what think Hans with who Jacob now speak 
 ‘With whom does Hans believe Jacob is now talking?’
Multiple *wh* phrases, where one is left in situ

Who remembers where we bought which book (Baker, 1970)

Two interpretations of (18) (Chomsky, 1973: 283):

a. For which $x$, $x$ remembers for which $z$, for which $y$, we bought $y$ at $z$

b. For which $x$, for which $y$, $x$ remembers for which $z$ we bought $y$ at $z$

Fronted *wh*-phrases in different constructions play very different roles in terms of information structure:

Who did you see? (QUESTIONS: *wh*-phrase a focus)

The woman who I saw (RELATIVES: *wh*-phrase a topic)

I’ll buy whatever you are selling. (FREE RELATIVES: *wh*-phrase fills the semantic functions of the missing head noun.)

What John lost was his keys. (WH-CLEFTS: *wh*-phrase prepares the hearer for the focused (new) information in sentence-final position)

Subjacency accounts for ungrammaticality where there is no movement:

*Mary is taller than I believe the claim that Susan is.*

A simple (but incorrect) generalization about case assignment — displaced NPs bear the case of their surface position; displaced *wh*-phrases bear the case of their subcategorized (underlying) position:

a. She / *Her was seen by Mary.

b. She / *Her is easy to please.

A classical TG rule ordering account:

a. Passive

b. Case Assignment

c. *Wh*-Movement

But notice the case when the subject of a passive is *wh*-fronted:

Who / *Whom do you think was seen by John?

The cycle accounts for this automatically

Constraint violations that pose no parsing difficulties:
(30)  
  a. *What did you eat beans and ____?
  b. What did you eat beans with ____?

(31)  
  a. *Who were you hoping for ___ to win the game? (Fodor, 1984)
  b. *What did the baby play with ___ and the rattle?

**Equal problems for the parser, but differing grammaticality:** (Fodor, 1984)

(32)  
  a. *John tried for Mary to get along well with ___.
  b. John is too snobbish for Mary to get along well with ___.

(33)  
  a. *The second question, that he couldn’t answer ___ satisfactorily was obvious.
  b. The second question, it was obvious that he couldn’t answer ___ satisfactorily.

Lightfoot, 1999 on the incidentally dysfunctional consequences of a functionally-motivated constraint:

(34)  
Traces of movement must be lexically governed

**Condition (34) explains the following contrast:**

(35)  
  a. Who, was it apparent [ei that [Kay saw ei]]?
  b. *Who, was it apparent yesterday [ei that [Kay saw ei]]?

(36)  
  Jay’s picture
(37)  
  [Jay’s [picture ei]]
(38)  
  the picture of Jay’s
(39)  
  *the picture of [Jay’s [e ei]]

(40)  
Lightfoot notes that the condition might be functionally motivated:

... the general condition of movement traces ... may well be functionally motivated, possibly by parsing considerations. In parsing utterances, one needs to analyze the positions from which displaced elements have moved, traces. The UG condition discussed restricts traces to certain well-defined positions, and that presumably facilitates parsing. (Lightfoot, 1999: 249)

**Condition (34) blocks the straightforward extraction of subjects:**

(41)  
  a. *Who, do think [ei that e, saw Fay]?
  b. *Who, do you wonder [e, how [e, solved the problem]??]

**Strategies for undoing the damages of the lexical government condition:**

(42)  
  a. Adjust the complementizer to license the extraction.
  b. Use a resumptive pronoun in the extraction site.
  c. Move the subject first to a nonsubject position and then extract.
Strategy (42a) — English:
(43) Who do you think saw Fay?

Strategy (42b) — Swedish:
(44) Vilket ord visste ingen hur det/*e stavas? 
Which word knew no one how it/e is spelled? 
‘Which word did no one know how it is spelled?

Strategy (42c) — Italian:
(45) Chi credi che abbia telefonato e? 
who do-you-think that has telephoned? 
‘Who do you think has telephoned?

(46) NP
    NP S (The Keenan and Comrie 1977 Accessibility Hierarchy)
    NP1 VP
    V NP2 P-NP3

(47) Kirby, 1998 on parallel and non-parallel function relatives:
   a. NP = SUBJ VP
      NP S VP
      NP = SUBJ VP
      V NP
   b. NP = OBJ
      NP S VP
      NP V NP
      NP = OBJ
      V NP

Parallel function relatives (the relativized NP is circled)

   c. NP = SUBJ VP
      NP S VP
      NP = OBJ
      V NP
   d. NP = OBJ
      NP S VP
      NP = SUBJ VP
      V NP

Non-parallel function relatives (the relativized NP is circled)
The factor that best explains why a person’s grammar has the properties that it has is conventionality:

... a central aspect of a speaker’s use of language is convention. When I say *Who did you meet yesterday?*, I put the interrogative pronoun *Who* at the beginning of the sentence because that is the convention of my speech community. I know the conventions of my speech community, and my use of language will serve its purpose best most of the time if I conform to the conventions of my speech community. It may be that the initial position of *Who* is partly motivated by pragmatic universals of information structure, or partly specified by an innate Universal Grammar. In fact, one (or both) of those factors may be the motivation for the origin of the convention. But that is not why I have put it there in that utterance. (Croft, 2000: 7)

BASIC PROBLEM FOR A DIRECT LINKING BETWEEN FORM AND FUNCTION: the forces (functional or otherwise) that bring a construction into a language aren’t necessarily the same ones that keep it there.

The two genitives of Modern English:

a. GEN-N: Mary’s mother’s uncle’s lawyer
b. N-GEN: the leg of the table

The functional differentiation of the genitive in Early Modern English (Kroch, 1994; Rosenbach and Vezzosi, 2000)

a. GEN-N began to be favored for animate genitives
b. N-GEN began to be favored for inanimate genitives

a. the table’s leg (inanimates can occur in the GEN-N construction)
b. the treachery of the enemy (animates can occur in the N-GEN construction)

Pressures for maintaining both the GEN-N and N-GEN orders in English:

a. The pressure of conventionality.
b. The pressure to have animate specifiers and inanimate complements
c. Purely structural pressure, caused by the existence of noun phrases with the structure [NP’s N] and [N of NP] where there is no semantic possession at all:
   i. Tuesday’s lecture
   ii. the proof of the theorem

L₁ has property X, X motivated by external function FU-EX₁
L₂ has property Y, Y motivated by external function FU-EX₂

Can we say that FU-EX₁ explains why L₁ has X and FU-EX₂ explains why L₂ has Y?

No!
What functionalists have to say about ‘innateness’:

linguistic abilities is that they are part of a more general cognitive-learning schema, the avenue of ‘explaining’ some phenomenon by declaring the rule describing it to be innate is not available… (Foley and Van Valin, 1984: 13)

However, as an empirical claim innateness remains empty because it is … just a name given to the set of language universals, and using this name should not blind us to the fact that a name is not an explanation. (Comrie, 1989: 25).

The Baldwin Effect: A mechanism by which advantageous traits can be incorporated into the genome even if they don’t directly increase the survival and reproductive possibilities of those that have them.


Three classes of Arguments from the Poverty of the Stimulus for innateness:

a. The abstractness and/or complexity of the principle is so great that no conceivable mechanism of inductive learning could have resulted in its acquisition.

b. The principle appears at such an early age and so rapidly that, again, no theory other than one that posits its innate origins can reasonably explain its possession by the child.

c. Some aspect of verbal behavior (normally judgments of sentential acceptability) is at great remove from anything likely to have been experienced. Therefore an innate principle must underlie that behavior (even though we may not know the precise nature of the principle).

Two examples of type (58a):

I. Chomsky, 1975 on structure-dependence: Chomsky’s well-known and often repeated argument for the innateness of the structure-dependent nature of grammatical rules.

Chomsky reasons as follows: Suppose that a child learning English always forms questions such as (59b) from declaratives such as (59a):

a. The man is tall.

b. Is the man tall?

In principle, the child could be working with one of two hypotheses:

a. To form questions one simply inverts the first occurrence of is (or an auxiliary).

b. More complexly, the child analyzes the sentence into abstract phrases and preposes the occurrence of is (or an auxiliary) that occurs after the first noun phrase.

Questions formed from declaratives such as (61a) decide the matter — the child assumes the second hypothesis. Given the first hypothesis, the child would incorrectly produce (61b). However, following the second, the child correctly produces (61c):

a. The man who is tall is in the room.

b. *Is the man who tall is in the room?
c. Is the man who is tall in the room?

How did the child come to have the second hypothesis? Chomsky argues that:

It is certainly absurd to argue that children are trained to use the structure-dependent rule, ...A person may go through a considerable portion of his life without ever facing relevant
evidence, but he will have no hesitation in using the structure-dependent rule, even if his
experience is consistent with hypothesis 1. ... The principle of structure-dependence is not
learned, but forms part of the conditions for language learning. (Chomsky, 1975: 32-33)

II. Hoekstra and Kooij, 1988 on Subjacency. Subjacency prohibits the formation of *wh-
questions if a *wh-phrase intervenes between the filler and the gap. Thus it predicts correctly
that (62a) is ambiguous as to the scope of where, while (62b) is not.

In (62a), where can refer both to the place of John’s saying and the place of getting off the bus,
while in (62b), where can refer only to the place of John’s asking:

(62) a. Where, did John say ___ that we had to get off the bus ___? (ambiguous)
    b. Where, did John ask ___ whether we had to get off the bus *___? (not ambiguous)

H&K suggest that there is no way that the child could have learned this inductively.

An example of type (58b):

The discussion of the acquisition of anaphoric binding in Crain and McKee, 1986. Consider
(63a-d):

(63) a. The Ninja Turtle danced while he ate pizza.
    b. While he danced, the Ninja Turtle ate pizza.
    c. His archrival danced while the Ninja Turtle ate pizza.
    d. He danced while the Ninja Turtle ate pizza.

In (63a-b) he can be coreferential with the Ninja Turtle,
In (63c) his can be coreferential with the Ninja Turtle,
However, in (63d) he and Ninja Turtle cannot corefer.

Even though the relevant principles are quite complex and abstract, Crain and McKee, 1986
have shown that even children as young as two years old accept (63a-b) and reject (63d).

They conclude that the principles of anaphoric binding must be innate.

An example of type (58c):

English speakers consistently judge (64) as well-formed and (65a-b) as ill-formed. Yet
‘parasitic gap’ sentences are extremely rare in actual discourse. Therefore the principles
guiding these judgments must be innate.

(64) This is the paper that I filed before reading.
(65) a. *I filed the paper before reading.
    b. *This is the paper that I filed the notes before reading.
Language serves many functions, which pull on it in many different directions (thought / communication).

For this reason, virtually all linguists agree that there can be no *simple* relationship between form and function.

However, two functional forces do seem powerful enough to have ‘left their mark’ on grammar — the force pushing form and meaning into alignment (pressure for iconicity) and the force favoring the identification of the structure of the sentence as rapidly as possible (parsing pressure).

Even these two pressures can conflict with each other, however — in some cases dramatically (for ex, any case where parsing pressure to postpone proper subpart of some semantic unit) (or where preference for topic before predication conflicts with pressure to have long-before-short, as in Japanese)

The problem, then, is to provide grammar with the degree of stability rendering it immune from the constant push-pull of conflicting forces.

A natural solution to the problem is to provide language with a relatively stable core *immune* to the immanent pressure coming from all sides.

That is, a natural solution is to embody language with a *structural system* at its core.

Put another way, an autonomous syntax as an intermediate system between form and function is a clever design solution to the problem of how to make language both learnable and usable:

This system allows language to be

- nonarbitrary enough to facilitate acquisition and use

and yet

- stable enough not be pushed this way and that by the functional force of the moment

*So autonomous syntax is functionally motivated!*

REFERENCES


4b. ECONOMY AND SIMPLICITY IN LANGUAGE STRUCTURE

(1) Ways in which ‘economy’, ‘simplicity’, least-effort’, etc. have been used:
   a. In the sense of ‘Occam’s Razor’, driving all empirical investigation.
   b. As a distinct functional force, corresponding to ‘iconicity’, ‘parsing’, and so on.
   c. As a construct internal to generative-grammar.

(2) The two major criticisms of ‘economy’, as it is used within generative grammar:
   a. Generative grammarians are interested only in a formal ‘playing with symbols’ economy — a sense of ‘economy’ that excludes a host of psychologically-relevant facts about grammars.
   b. Generative grammarians focus too narrowly on economy within one component of the grammar (in particular, syntax), without consideration of the consequences for the ‘economy’ of the system as a whole.

On (1a) — Occam’s razor:

(3) A sometimes-heard argument that would appear to challenge Occam’s razor:
    Language is a biological system and, like all such systems, is full of redundancy. Therefore, it is a mistake to search for the simplest (i.e. most non redundant) analysis of any phenomenon.

The fallacy here is ignoring the fact that one searches for the simplest analysis consistent with the facts

(4) Functionalists appeal to Occam’s razor as frequently as formalists do:
    If grammar is autonomous … then any attempt to relate grammatical theory to broader cognitive capacities is doomed to failure. … [But in fact] syntactic structures are grounded in ordinary spatial thought via conceptual metaphor. … It will be argued at length that basic grammatical properties emerge as metaphoric inferences out of basic inferential patterns applicable to ordinary physical objects. (Deane 1992: 2-3).

Deane never argues that it is impossible to formulate autonomous rules to capture island constraints — His argument is based on an Occam’s razor appeal to elegance, given a set of facts.

On (1b) — Economy as a functional force:

(5) All external explanations are appeals to ‘economy’ / ‘least effort’
   a. ‘Deletions’ under formal or discourse identity allow the speaker to economize production effort.
b. An iconic relation between form and meaning is economical for both speaker and hearer as it leads to less ‘mental computation’ for each

c. Redundancy (as in grammatical agreement) is economical for the hearer, as it facilitates the recovery of relations among grammatical elements

d. Appeals to ‘metaphorical extension’ as part of an explanation for grammatical form presuppose that the organization of grammatical structure is simpler / more economical when there are established cognitive links among related elements.

e. Etc.

Better, therefore, not to use the term ‘economy’ at all in this sense

On (1c) — Simplicity within generative grammar:

A legacy of Prague School structuralism is that the correct grammar is the most economical in terms of number of symbols employed, rules written, etc.

A problem: noted in Chomsky and Halle 1968: (6a-b) and (7a-b) are equal in their number of distinctive feature specifications. But the (a) rules are more natural / common than the (b) rules:

(6) a. \(\phi \rightarrow \emptyset / C \_\_ [+\text{son}]\#\)
   b. \(\phi \rightarrow k / C \_\_ [+\text{son}]\#\)

(7) a. \([k, g] \rightarrow [\check{c}, j] / \_\_ [i, e]\]
   b. \([p, b] \rightarrow [t, d] / \_\_ [i, e]\]

(8) Some principles of GB, very informally (and rather inaccurately) stated:
   a. Binding. An anaphor is coreferential to its closest possible antecedent.
   b. Bounding. An element can move only to its closest potential landing site.
   d. Case. Verbs and prepositions case-mark elements in their local domain.
   e. Theta. A head theta-marks elements in its local domain.
   f. Control. PRO takes its closest possible antecedent as a controller.

The leading idea of the minimalist program is to distill core of each principle into one overriding ‘economy’ principle

(9) Chomsky on the economy of language design and language use:
Language design appears to be problematic from a parsing-theoretic perspective, though elegant regarded in isolation from considerations of use. (Chomsky 1991: 448)

On criticism (2a) — Generative grammarians are interested only in a formal ‘playing with symbols’ economy — a sense of ‘economy’ that excludes a host of psychologically-relevant facts about grammars.
(10) Givón on the ‘gutting of the data base’

The pseudotheoretical baggage which transformational-generative linguistics carried into structuralism was coupled with another — equally ominous — development: the gutting, beyond recognition, of the concept of relevant data in linguistics. … Thus by dissociating itself from the consideration of communicative function, speech processing, cognitive-perceptual structure, diachronic, ontogenetic and phylogenetic evolution, and world-view pragmatics and ontology, transformational linguistics had already restricted itself to the narrow band of language-internal data covered by the Bloomfieldians. (Givón 1979: 23)

(11) Statements like Givon’s raise two questions:

a. What data are relevant for the construction of a grammatical theory?

b. Are we in a position to use all forms of potentially relevant data?

Short answer to (11a) — all data are potentially relevant
Short answer to (11b) — ‘No’

(12) A fallacious argument:

Language users attend closely to frequency, as is evidenced by the relationship between frequency of use of a form and facts about grammars and how they change. Therefore grammars have to incorporate facts about relative frequency.

This argument is fallacious because X can affect Y without being part of a system that includes Y.

(13) Paraphrase of an argument by Sandra Thompson:

Chomsky says that phrases like the shooting of the hunters are ambiguous. But I have scanned thousands of pages of text and found hundreds of examples of structures of the form ‘the N of the NP’ and have discovered that in 96% of the cases, the NP is interpreted as the object of the gerund, not as the subject. Therefore, Chomsky is wrong.

But grammars encode structural possibilities — not propensities for taking advantage of these possibilities

On criticism (2b) — Generative grammarians focus too narrowly on economy within one component of the grammar (in particular, syntax), without consideration of the consequences for the ‘economy’ of the system as a whole.

This criticism has a lot to recommend it.

(14) The ‘Definiteness Effect’:

a. ?There is the meeting in room 205 today.

b. ?There’s the dog running loose somewhere in the neighborhood.

(15) Syntactic accounts of the definiteness effect: Milsark 1977; Reuland 1985; Safir 1985
(16) Ward and Birner 1995 — The correct account is a pragmatic one — in such sentences are required to represent a hearer-new entity. Hence (17b) is fine.

(17) There were the same people at both conferences.

**Croft 1995 on the three components of the autonomy of syntax:**

(18) a. At least some elements of syntax are arbitrary (ARBITRARINESS);
   b. The arbitrary elements participate in a system (SYSTEMATICITY);
   c. That system is self-contained (SELF-CONTAINEDNESS).

**Some (likely) arbitrary facts about English syntax:**

(19) a. He is likely to be late.
   b. *He is probable to be late. (likely, but not probable, allows raising)

(20) a. He allowed the rope to go slack.
   b. *He let the rope to go slack. (let doesn’t take infinitive marker)

(21) a. He isn’t sufficiently tall.
   b. *He isn’t enough tall. / He isn’t tall enough. (enough is only degree modifier that occurs postadjectivally)

(22) The weakening of ‘self-contained’ syntax in some recent work:

   a. The incorporation of theta-roles into grammatical description, where theta-roles are ‘semantic properties assigned by heads’ (Chomsky 1986: 93). For example, consider the idea that syntactic structure is a projection of lexical structure, where lexical structure is stated in terms of ‘theta-grids’. In its extreme form, this involves trying to derive ‘C-Selection’ from ‘S-Selection’, utilizing lexical representations like:

   persuade _____ goal proposition
   hit _____ patient goal

   b. The idea that movement is triggered by features of functional categories with semantic content:

   Syntactic movement … must be triggered by the satisfaction of certain quasi-morphological requirements of heads. … [S]uch features have an interpretive import (Wh, Neg, Top, Foc, …): they determine the interpretation of the category bearing them and of its immediate constituents …, function as scope markers for phrases with the relevant quantificational force in a local configuration, etc. (Rizzi 1997: 282)

   (23) **Culicover and Jackendoff 2001 on control:**

   a. Mary attempted to take a vacation. (Agent control)
   b. Bill told Mary to come visit him. (Addressee control)
   c. I persuaded John to see the doctor. (Patient control)
   d. We taught the dog to roll over. (Recipient control)
   e. Sam promised Mary to write more often. (Source control)
f. Tom required Alice to hand in the assignment. (Holder control)

(24) **Coindexing, but no coreference:**
   a. [no qualified applicants] presented themselves
   b. it, seems, to be raining
   c. John is easy [O, [to please t]]

(25) **The coindexing - coreference distinction parallels the distinction between**
   a. Grammatical gender and ‘natural’ gender
   b. Formal features involved phenomena such as classifiers and
      their (approximate) semantic counterparts
   c. Etc.

The Grammar-internal Structure-Preserving Constraint of Emonds 1976 prohibits:
(26) a. *John regretted that liver he would never eat.
    b. *Mary feared that speaking at today’s lunch would be our
       congressman.

Hooper and Thompson 1973 noted that the following sentences seem fully
acceptable:
(27) a. John said that liver he would never eat.
    b. Mary repeated that speaking at today’s lunch would be our
       congressman.

If Hooper and Thompson are right, then syntax is/can still be autonomous,
just ‘smaller’

MORAL: Grammar-internal economy cannot focus just on syntax

REFERENCES

York: Praeger.
Chomsky, N. 1991. Some Notes on Economy of Derivation and
Representation, ed. R. Freidin, *Principles and Parameters in Comparative
and Row.
Oxford: Oxford University Press.
Inquiry 32.*
Emonds, J. E. 1976. *A Transformational Approach to English Syntax.* New York:
Academic Press.


5a. Prototypes and Syntactic Categories

(1) The standard generative view of categories: The grammar manipulates discrete formal objects drawn from a universal vocabulary.

(2) The challenge from prototype theory: Categories have a ‘best case’ and deviations from that best case, where the best case is semantically and/or pragmatically characterizable:

```
CAT
```

(3) Categorial Prototypicality:
   a. Grammatical categories have ‘best case’ members and members that systematically depart from the ‘best case’.
   b. The optimal grammatical description of morphosyntactic processes involves reference to degree of categorial deviation from the ‘best case’.

(4) Fuzzy Categories
   The boundaries between categories are nondistinct.

(5) The Fake NP Squish (Ross 1973a)
   a. Animates
   b. Events
   c. Abstracts
   d. Expletive *it
   e. Expletive *there
   f. Opaque idiom chunks

(6) a. John is likely to be shown to have cheated.
   b. ?The performance is likely to be shown to have begun late.
   c. *No headway is likely to have been shown to have been made.

(7) The Nouniness Squish (Ross 1973b: 141)
   that clauses > for to clauses > embedded questions > Acc ing complements > Poss ing complements > action nominals > derived nominals > underived nominals

(8) Preposition deletion
   — must apply before that and for-to complements (9a),
   — may optionally apply before embedded questions (9b), and
   — may not apply before more nouny elements (9c)

(9) a. I was surprised (*at) that you had measles.
   b. I was surprised (at) how far you could throw the ball.
c. I was surprised *(at) Jim’s victory.

<table>
<thead>
<tr>
<th>Syntactic Category</th>
<th>Noun</th>
<th>Adjective</th>
<th>Verb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semantic class</td>
<td>Object</td>
<td>Property</td>
<td>Action</td>
</tr>
<tr>
<td>Valency</td>
<td>0</td>
<td>1</td>
<td>≥1</td>
</tr>
<tr>
<td>Stativity</td>
<td>state</td>
<td>state</td>
<td>process</td>
</tr>
<tr>
<td>Persistence</td>
<td>persistent</td>
<td>persistent</td>
<td>transitory</td>
</tr>
<tr>
<td>Gradability</td>
<td>nongradable</td>
<td>gradable</td>
<td>nongradable</td>
</tr>
<tr>
<td>Pragmatic function</td>
<td>Reference</td>
<td>Modification</td>
<td>Predication</td>
</tr>
</tbody>
</table>

Table 1
*Prototypical correlations of syntactic categories* (Croft 1991: 55, 65)

(10) **A prototypicality hierarchy for nouns**
   a. 1st and 2nd person pronouns
   b. 3rd person pronouns
   c. proper names
   d. human common nouns
   e. nonhuman animate common nouns
   f. inanimate concrete common nouns
   g. abstract nouns
   
   Decreasing prototypicality

(11) For adjectives: age, dimension, value, color are said to be most prototypical
(12) For verbs: causative action verbs are said to be most prototypical

The two central claims of prototype theory as far as morphosyntax is concerned go back to old work in markedness:

(13) **Syntagmatic simplicity** — prototypical member of a category requires no more morphemes (or are in general shorter) than a nonprototypical member

(14) **Paradigmatic complexity** — prototypical members manifest more
   a. structural elaboration
   b. inflectional possibilities
   than nonprototypical members
(15) English pronouns (more prototypical than nouns):
   a. SS — they are shorter than nouns
   b. PC — they distinguish between nominative and accusative case and nouns
   do not

(16) German singulars (more prototypical than plurals)
   a. SS — nominative sing. doesn’t have suffix; nom. pl. does
   b. PC — nominative singular can have definite article *der, die, das
       nominative plurals only *die

(17) Special morphemes to indicate nonprototypical
   a. happi+ness
   b. the fly+ing object

(18) Predication Hierarchy
    Action word (V) < Property word (A) < Object word (N)
    No language will require a special morpheme for verbal predication if
    it does not require one for adjectival and nominal predication.
    No language will require a special morpheme for adjectival predication
    if it does not require one for nominal.

(19) More illustrations of Paradigmatic Complexity:
    a. Structural elaboration:
       — many more rules can apply to animates than to inanimates
       (in many languages, only animates can passivize)
    b. Inflectional possibilities:
       — infinitives & participles are nonprototypical verbs (restricted in
         tense, aspect & agreement)
       — stative verbs don’t occur in progressive
       — predicate nouns & adjectives are nonprototypical (often don’t
         inflect, eg German predicate adjectives don’t show agreement)
         das rote Buch; *das Buch ist rote/rotes

(20) The strongest claim about prototypicality and grammatical
description:
    These [markedness, hierarchy, and prototype] patterns are universal, and are
    therefore part of the grammatical description of any language. Language-
specific facts involve the degree to which typological universals are
    conventionalized in a particular language; e.g. what cut-off point in the
    animacy hierarchy is used to structurally and behaviorally mark direct objects
    (Croft 1990: 154)

The following examples show that there is no cut-off point for progressive
morphology in English on an absolute scale pf prototypicality:

(21) a. Mary was throwing the ball.
    b. *Mary was containing 10 billion DNA molecules.

(22) a. The portrait is hanging on the wall of the bedroom.
    b. I’m enjoying my sabbatical year.

(23) *I’m noticing a diesel fuel truck passing by my window.

(24) a. Tomorrow, the Mariners are playing the Yankees.
    b. *Tomorrow, the Mariners are playing well.
Likewise for a variety of other processes (25-29):

Various subject alternations:
(25)  a. The world saw the beginning of a new era in 1492.
    b. 1492 saw the beginning of a new era.
(26)  a. We sleep five people in each room.
    b. Each room sleeps five people.
(27)  a. The middle class will benefit from the new tax laws.
    b. The new tax laws will benefit the middle class.

*There*-Insertion:
(28)  a. A ship appeared on the horizon.
    b. There appeared a ship on the horizon.

Locative inversion:
(29)  a. A flowering plant is on the window sill.
    b. On the window sill is a flowering plant.

*Measure verbs and passive:*
(30)  a. The book cost a lot of money.
    b. John weighed 180 pounds.
(31)  a. *A lot of money was cost by the book.
    b. *180 pounds was weighed by John.
(32)  a. The boys make good cakes.  (Bresnan 1978)
    b. The boys make good cooks.
(33)  a. Good cakes are made by the boys.
    b. *Good cooks are made by the boys.

Passives in (31a-b) and (33b) are impossible because the post-verbal phrases are not arguments of the verb. The ‘prototypicality’ of the verb does not enter directly into the explanation.

*There as a nonprototypical NP.*

(34)  The Fake NP Squish
    a.  Animates
    b.  Events
    c.  Abstracts
    d.  Expletive *it
    e.  Expletive *there
    f.  Opaque idiom chunks

Why call it expletive *there* a NP?
— it raises over the verb *seem* and others of its class (35a);
— it occurs as a passive subject (35b);
— it inverts over auxiliaries (35c);
— it can be coindexed with tags (35d):
(35)  

a. There seems to be a problem.
b. There was believed to be a problem.
c. Is there a problem?
d. There is a problem, isn’t there?

How does expletive there behave like a ‘non-prototypical NP? According to Ross:
— it doesn’t undergo the rule of ‘promotion’ (36a-b);
— it doesn’t allow raising to reapply (37a-b);
— it doesn’t occur in the ‘think of ... as X’ construction (38a-b) or the ‘what’s ...
  doing X’ construction (39a-b);
— it doesn’t allow ‘being-deletion’ (40a-b);
— it doesn’t occur in dislocation constructions (41a-b);
— it doesn’t undergo tough-movement (42a-b),
— topicalization (43a-b),
— ‘swooping’ (44a-b),
— ‘equi’ (45a-b),
— or conjunction reduction (46a-b).

Promotion:  
(36)  
a. Harpo’s being willing to return surprised me. / Harpo surprised me by
  being willing to return.
b. There being heat in the furnace surprised me. / *There surprised me by
  being heat in the furnace.

Double Raising:  
(37)  
a. John is likely ___ to be shown ___ to have cheated.
b. ?*There is likely ___ to be shown ___ to be no way out of this shoe.

Think of ... as NP:  
(38)  
a. I thought of Freud as being wiggy.
b. *I thought of there as being too much homework.

What’s ... doing X?:  
(39)  
a. What’s he doing in jail?
b. *What’s there doing being no mistrial?

Being Deletion:  
(40)  
a. Hinswood (being) in the tub is a funny thought.
b. There *(being) no more Schlitz is a funny thought.

Left Dislocation:  
(41)  
a. Those guys, they’re smuggling my armadillo to Helen.
b. *There, there are three armadillos in the road.

Tough Movement:  
(42)  
a. John will be difficult to prove to be likely to win.
b. *There will be difficult to prove likely to be enough to eat.

Topicalization:  
(43)  
a. John, I don’t consider very intelligent.
b. *There, I don’t consider to be enough booze in the eggnog.

Swooping:  
(44)  
a. I gave Sandra my zwieback, and she didn’t want any. / I gave Sandra, and
  she didn’t want any, my zwieback.
b. I find there to be no grounds for contempt proceedings, and there may
  have been previously. / *I find there, which may have been previously, to be no
  grounds for contempt proceedings.

Equi:
(45)  a. After he laughed politely, Oliver wiped his mustache. / After laughing politely, Oliver wiped his mustache.
     b. After there is a confrontation, there’s always some good old-time head-busting. / *After being a confrontation, there’s always some good old-time head-busting.

Conjunction Reduction:
(46)  a. Manny wept and Sheila wept. / Manny and Sheila wept.
     b. There were diplodocuses, there are platypuses, and there may well also be diplatocodypuses. / *There were diplodocuses, are platypuses, and may well also be diplatocodypuses.

The ungrammaticality of the starred sentences follows from the meaning of there
(abstract setting, conceptual space) and pragmatic conditions on the use of an
element with that meaning. There is no reason for the grammar to make reference
to the lack of prototypicality of there.

(47)  The square circle elapsed the dragon.

English idiom chunks

Ross 1981 and Lakoff 1987 argue that idioms can be ranked internally in terms of
their degree of prototypicality:

(48)  a. to stub one’s toe
     b. to hold one’s breath
     c. to lose one’s way
     d. to take one’s time

Only to stub one’s toe can be converted into a past participle - noun sequence.
(49)  a. A stubbed toe can be very painful.
     b. *Held breath is usually fetid when released.
     c. *A lost way has been the cause of many a missed appointment.
     d. *Taken time might tend to irritate your boss.

To stub one’s toe and to hold one’s breath allow gapping in their conjuncts:
(50)  a. I stubbed my toe, and she hers.
     b. I held my breath, and she hers.
     c. *I lost my way, and she hers.
     d. *I took my time, and she hers.

Pluralization with conjoined subjects distinguishes to stub one’s toe from to hold
one’s breath and both of these from to lose one’s way and to take one’s time:
(51)  a. Betty and Sue stubbed their toes.
     b. *Betty and Sue stubbed their toe.
     c. Betty and Sue held their breaths.
     d. Betty and Sue held their breath.
     e. *Betty and Sue lost their ways.
     f. Betty and Sue lost their way.
     g. *Betty and Sue took their times.
     h. Betty and Sue took their time.
All but to take one’s time allow pronominalization:

(52)  
  a. I stubbed my toe, but didn’t hurt it.
  b. Sam held his breath for a few seconds, and then released it.
  c. Harry lost his way, but found it again.
  d. *Harry took his time, but wasted it.

The following examples show that the above facts have nothing to do with the ‘prototypicality’ of the idiom:

(53)  
  a. *Held cats often try to jump out of your arms.
  b. *The taken jewels were never returned.
(54)  
  A lost child is a pathetic sight.
(55)  
  a. I lost my way, and she her way.
  b. I took my time, and she her time.
  c. ?I ate my ice cream and she hers.
  d. In the race to get to the airport, Mary and John lost their way, but we didn’t lose ours (and so we won).
(56)  
  a. *Betty and Sue broke their leg.
  b. *My bowl is full of rices.
(57)  
  a. Harry took his time, and wasted it.
  b. Harry took his time, which doesn’t mean that he didn’t find a way to waste it.

Event structure and inflectional possibilities.

(58)  
  The event structure representations in Pustejovsky 1991: 
  
a. States:

\[
\begin{array}{c}
S \\
| \\
e
\end{array}
\]

b. Processes:

\[
\begin{array}{c}
P \\
\triangledown \\
e_1 \ldots e_2
\end{array}
\]

c. Achievements and Accomplishments have the same schematic structure (both are called ‘transitions’), though the former are non-agentive and the latter agentive:

\[
\begin{array}{c}
T \\
\quad E_1 \\
\quad \neg E_2
\end{array}
\]

The more complex the Event Structure, the more aspectual possibilities — and hence the richer the Functional Structure — and hence the more inflectional possibilities

In other words, one can derive a large part of paradigmatic complexity without any reference to prototypes.
Two possible objections to an exclusively semantic account of restrictions on inflection:

(59) In some languages, inflections are restricted to some forms, but not others, even though there is no semantic incompatibility between the form and the inflection.

(60) Some inflections seem to be semantically empty (e.g. agreement)

**Prototypes and markedness**

Some markedness hierarchies:

(61) Number (Greenberg 1963): singular < plural < dual < trial/paucal

(62) Animacy (Dixon 1979: 85): 1st, 2nd person pronouns < 3rd person pronouns < proper names < human common nouns < nonhuman animate common nouns < inanimate common nouns

(63) Definiteness (Croft 1990): definite < referential indefinite < nonreferential indefinite

What is the difference between prototypicality and (un)markedness?

Markedness is relative; prototypicality is absolute

(64) a. kē-φ ‘to’ / dari-φ ‘from’ + noun phrase denoting a place (Malay)

   b. kē-pada ‘to’ / dari-pada ‘from’ + noun phrase denoting a person

**Fuzzy categories**

**Ross 1972 on near**

Near seems to be something between an A and a P:

(65) a. The shed is near to the barn. (like an A)

   b. The shed is near the barn. (like a P)

But it meets all tests for A:

(66) a. The gas station is near to the supermarket.

   b. Near enough to the supermarket

   c. Nearer to the supermarket

   d. The near shore

And it meets all tests for P:

(67) a. The gas station is near the supermarket.

   b. *The gas station is near enough the supermarket.

   c. The gas station is right near (*to) the supermarket.

(68) The gas station is nearer the supermarket than the bank.

(69) The seaplane right now is more over the lake than over the mountain.

**The nouniness squish:**

(70) The ‘Nouniness Squish’:

that clauses > for to clauses > embedded questions > Acc ing complements > Poss ing complements > action nominals > derived nominals > underived nominals
Ross says that the nounier the complement, the harder to extract from it:

(71)  
  a. I wonder who he resented (it) that I went steady with.  
  b. I wonder who he would resent (it) for me to go steady with.  
  c. *I wonder who he resented how long I went steady with.  
  d. ?I wonder who he resented me going out with.  
  e. ??I wonder who he resented my going out with.  
  f. ?*I wonder who he resented my careless examining of.  
  g. ?*I wonder who he resented my careless examination of.  
  h. ??I wonder who he resented the daughter of.  

Ross says that pied-piping is sensitive to the degree of nouniness; the more nouny, the easier to pied-pipe:

(72)  
  a. *Eloise, [for us to love [whom]] they liked, is an accomplished washboardiste.  
  b. *Eloise, [us loving [whom]] they liked, is an accomplished washboardiste.  
  c. *Eloise, [our loving [whom]] they liked, is an accomplished washboardiste.  
  d. ?*Eloise, [our loving of [whom]] they liked, is an accomplished washboardiste.  
  e. ?Eloise, [our love for [whom]] they liked, is an accomplished washboardiste.  
  f. Eloise, [a part of [whom]] they liked, is an accomplished washboardiste.  

Both can be accounted for in ‘discrete’ terms

REFERENCES

OVERVIEW OF PAPER: H sets out to explain the common prohibition against definite articles cooccurring with possessive phrases. He argues that appealing to a single structural position for ‘determiners’ will not work. Instead he puts forward a functional explanation: since possessed NPs are usually definite, ‘economic motivation’ dictates that there is no need for definite articles and possessors to cooccur.

SUMMARY EVALUATION: H’s arguments against the structural explanation are quite convincing. However, issues not discussed by H undercut to a certain degree the functional explanation that he puts forward.

A. ECONOMY VERSUS REDUNDANCY

Pressure for economy, to be sure, is an important motivating force in language. But grammars are also filled with redundancy. Charles Hockett has offered the opinion that:

In every language, redundancy, measured in phonological terms, hovers near 50%. ... It may be that the redundancy figure would be about the same were it measured in grammatico-lexical terms ... (Hockett 1963: 24)

Here is a more recent statement, from the point of view of functional-cognitive linguistics:

[G]rammar is massive and highly redundant, rather than stripped down and economical. (Kemmer and Barlow 2000: ix)

Presumably, redundancy is also motivated functionally, as repetition of some grammatical feature is an aid to hearer comprehension. The question, then, is when economy-effects tend to be dominant and when redundancy-effects tend to be dominant (and why). In the absence of an answer to this question, any appeal to economy (or redundancy) has the danger of vacuity.

Economy effects are most manifest in ‘non-local’ contexts, that is where structure/meaning is recoverable from earlier stretches of structure or from the discourse itself. Typical examples are ‘deletions under identity’ (1) and null arguments (2):

(1) John promised to help Mary on Tuesday and he did [help Mary on Tuesday].

(2) Chinese (Huang 1984: 533)
Redundancy effects, on the other hand, tend to be local (for discussion, see Lapointe 1988). That is, they show up in specifier-head configurations and affect features such as number, gender, case, person, declension class, tense, aspect, mood, voice, and definiteness. Some examples of definiteness agreement (all taken from Ferguson and Barlow 1988):

(3) ha-isha ha-tov-a (Hebrew)
    the-woman the-good-FEM.SG
    ‘the good woman’

(4) al-mu‘allim-u al-jadiid-u (Classical Arabic)
    DEF-teacher(MASC.SG)-NOM DEF-new(MASC.SG)-NOM
    ‘the new teacher’

(5) he gune he agath-e (Classical Greek)
    the.FEM.SG woman the.FEM.SG good-FEM.SG
    ‘the good woman’

If it is the case that definiteness is a feature more likely to be repeated than economized on, then H’s explanation for definite article - possessor complementarity does not go through.

B. DEFINITE AND INDEFINITE ARTICLES

H’s argument is based on the idea that since possessed phrases are usually definite, economy dictates that the definite article is dispensable. The prediction would seem to follow, then, that when the possessed phrase is indefinite, the indefinite article (if there is one in the language) would be more likely to occur. That is, we should find languages where indefinite articles occur with possessive NPs, but not definite articles, but no languages in which the reverse situation holds. (This prediction is parallel to H’s explanation (p. 231) for why we find special case marking on animate patients and overt coding of possession of alienable nouns.)

Is the prediction fulfilled? From H’s few remarks on indefinite articles, it would appear not to be. So for example, H tells us (p. 240) that in Romanian the definite article cooccurs with the possessive, but the indefinite article does not:

(6) a. profesorul meu
teacher-the my
b. *un professormeu
   a   teacher       my

What is needed for H’s argument, then, is a survey of both kinds of articles and a demonstration that a significant number of languages allow structures like (7a), but not like (7b):

(7) a. a my book
    b. the my book

C. ON H’S UNIVERSAL 1

Universal 1: If possessed NPs show the definite article, then so do nonpossessed NPs.

This universal allows economical inexplicit languages (e.g., English), explicit uneconomical languages (e.g., Italian), and languages that completely lack definite articles (e.g., Russian), but it excludes uneconomical inexplicit languages that (uneconomically) have the definite article in possessed NPs but (inexplicitly) lack the definite article in nonpossessed NPs. (H, p. 234)

Things are more complicated than that, however. Let us look at the four language types, as far as the ‘features’ ±ECONOMICAL and ±EXPLICIT are concerned:

<table>
<thead>
<tr>
<th>Language type</th>
<th>Instantiated by</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ECO +EXP</td>
<td>Brazilian Portuguese</td>
<td>the my book</td>
</tr>
<tr>
<td></td>
<td></td>
<td>my book</td>
</tr>
<tr>
<td>+ECO -EXP</td>
<td>English</td>
<td>*the my book</td>
</tr>
<tr>
<td></td>
<td></td>
<td>my book</td>
</tr>
<tr>
<td>-ECO +EXP</td>
<td>Italian</td>
<td>the my book</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*my book</td>
</tr>
<tr>
<td>-ECO -EXP</td>
<td>NONE</td>
<td>the my book</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*the (old) book</td>
</tr>
</tbody>
</table>

As H notes, within individual NPs, economy conflicts with explicitness. However, looking at languages as a whole, we see that both economical and explicit constructions are possible. Brazilian Portuguese is a testament to that. Presumably, when economy is called for by the discourse situation, then a BP speaker will be more likely to say my book and when explicitness is called for, then the speaker will be more likely to say the my book. The question, then, is why more languages are not like Brazilian Portuguese? That is, if H’s theory predicts why there are no uneconomical inexplicit languages, then it would also seem to predict (incorrectly) that there should be lots of language in which both economy and explicitness are possible.
Furthermore, there is probably an explanation (quasi-structural, quasi-semantic) for the nonexistence of the fourth type of language that does not appeal directly to either economy or explicitness. Note that in such languages, the equivalent of (8a) is possible, but not that of (8b):

\[(8)\]
\[
\begin{align*}
&\text{a. } \text{DEF ART - POSS PRON - N} \\
&\text{b. } \text{*DEF ART - N}
\end{align*}
\]

In such a language, the full expansion of NP would be possible, but not a reduced expansion in which a semantically optional element is missing (‘semantically optional’, since not all nouns are possessed).

We would not expect such a state of affairs to exist, both on structural and on semantic grounds, In other words, H’s fourth type of language should be rare for the same reason that we would not expect to find languages in which (9a) is possible, but not (9b); or (10a), but not (10b):

\[(9)\]
\[
\begin{align*}
&\text{a. } \text{the old book} \\
&\text{b. } \text{*the book}
\end{align*}
\]

\[(10)\]
\[
\begin{align*}
&\text{a. } \text{an old book} \\
&\text{b. } \text{*a book}
\end{align*}
\]

The explanation of (9) and (10) is surely parallel to that of (8), even though neither ‘economy’ not ‘explicitness’ appears to be playing a role.

\[\text{D. ON H’S UNIVERSAL 2}\]

Universal 2: If possessed NPs with a preposed possessor show the definite article, then so do possessed NPs with a postposed possessor.

The fact that articles are more likely to be omitted with preposed possessors finds a natural explanation within my economy-based approach: if the possessor preceded the head noun, then at the time the hearer encounters the head noun he or she already has the information about its probable definiteness. Thus, overt indication of definiteness is still more redundant under these circumstances. (H, p. 235)

In other words, we should not find language in which (11a) is possible, but not (11b):

\[(11)\]
\[
\begin{align*}
&\text{a. } \text{that John’s book} \\
&\text{b. } \text{that book of John’s}
\end{align*}
\]

And H gives, on p. 234, examples of a number of languages to instantiate that point.

The question, though, is whether H’s explanation is general enough. Emonds 1976 observed that English in general does not allow full phrases between the determiner and the head of an NP, even though they are allowed in post-head position:
Hawkins 1994 provides a parsing-based explanation for this tendency that generalizes to all VO languages (I believe that H’s illustrative languages on p. 234 are all VO). Hawkins explanation would appear to subsume H’s. (It would be interesting to see what the situation is for OV languages, where H’s and Hawkins’s predictions might differ.)

E. ON H’S UNIVERSAL 3

Universal 3: If possessed NPs with a kinship term as head noun show the definite article, then so do possessed NPs with a nonkinship term as head noun.

I propose that articles are more likely to be omitted with kinship nouns because the possessive relation inherent in them … (H, p. 235)

This universal is illustrated by H (p. 236) for Italian (16) and Brazilian Portuguese (17), as well as a few other languages:

(16) a. la mia casa ‘my house’
    b. *la mia madre ‘my mother’
(17) a. a minha casa ‘my house’
    b. *a minha mãe ‘my mother’

I don’t question the basic universal, but the explanation might be more complex than the one suggested by H in terms of ‘relationality’. In Italian (Cinzia Russi, p. c.) the definite article is possible with plural kinship terms, even though one’s sisters are as ‘relational’ as one’s mother:

(18) le mie sorelle

Furthermore, the definite article is required with singular kinship terms, if the term is a diminutive or term of endearment:

(19) a. la mia mammina
    b. *mia mammina
(20) a. la mia sorellina
    b. *mia sorellina
REFERENCES


7. Processing explanations and discourse-based explanations

(1) Two important external (functional) explanations:
   a. Parsing: There is pressure to shape grammar so the hearer can
determine the structure of the sentence as rapidly as possible.
   b. Information flow in discourse: There is pressure for the syntactic
structure of a sentence to mirror the flow of information in discourse

(2) The most comprehensive attempt to demonstrate how parsing pressure
has influenced grammatical structure is Hawkins 1994

Central insight — it is in the interest of the hearer to recognize the syntactic
groupings in a sentence as rapidly as possible.

(3) This preference is realized both in language use and in the grammar
itself.
   a. Use: when speakers have the choice, they will follow the
parser’s preference
   b. Grammar: facts about grammars will reflect parsing preferences

Central principle: EIC:

(4) Early Immediate Constituents (EIC)
The hearer (and therefore the parsing mechanism) prefers orderings of
elements that lead to the most rapid recognition possible of the structure of
the sentence

EIC Explains why long (or heavy) elements tend to come after short (or light)
one in VO languages

Why, for example, (5b) sounds better than (5a) and why (6b) sounds better
than (6a)

Examples of long (or ‘heavy’) elements (in boldface) tending to follow short
(or ‘light’) elements (in italics):
(5) a. ?I consider everybody who agrees with me and my disciples about the
nature of the cosmos smart.
   b. I consider smart everybody who agrees with me and my disciples
about the nature of the cosmos.
(6) a. ?I met the twenty three people who I had taken Astronomy 201 with
last semester in the park.
   b. I met in the park the twenty three people who I had taken Astronomy
201 with last semester.
Notice that the distance between V and P in (7a) is much longer than the distance between V and D in (7b). So (6b) sounds much better than (6a).

(8) Typological predictions of the EIC
   a. VO langs tend to extrapose S' subjects
   b. OV languages, like Japanese, tend to prepose S' objects.

(9) a. That John will leave is likely.
    b. It is likely that John will leave.

(10) a. Mary-ga [kinoo John-ga kekkonsi-ta to] it-ta
     Mary yesterday John married that said
    'Mary said that John got married yesterday'
    b. [kinoo John-ga kekkonsi-ta to] Mary-ga it-ta

EIC provides an explanation: By extraposing the sentential subject in English, the space in which the main clause constituents can be recognized is dramatically shortened, as illustrated in (11a-b).

   By preposing the sentential object in Japanese, we find an analogous recognition domain shortening, as in (12a-b):

(11) a. \[s_s[\text{that} s_s[\text{John will leave}]]_{vp}[\text{is likely}]\]

b. \[s_{np}[\text{it}]_{vp}[\text{is likely} s_{s}[\text{that} s_s[\text{John will leave}]]]\]

(12) a. \[s_{s}[\text{Mary-ga} s_{s}[s_s[\text{kinoo John-ga kekkonsi-ta} to] it-ta]]\]

b. \[s_{s}[s_{s}[\text{kinoo John-ga kekkonsi-ta} to] s_{s}[\text{Mary-ga} s_{s}[\text{it-ta}]]\]

EIC explains why VO languages have prepositions and OV languages have postpositions.
There are four logical possibilities, illustrated in (13a-d): VO and prepositional (13a); OV and postpositional (13b); VO and postpositional (13c); and OV and prepositional (13d):

(13) a.      VP b.                      VP
             V      NP     PP         PP        NP       V
             P            NP           NP      P
VO and prepositional (common) OV and postpositional (common)

In (13a) and (13b), the two common structures, the recognition domain for the VP is just the distance betw V and P, crossing over the object NP

But in (13c) and (13d), the uncommon structures, the recognition domain is longer — it involves the object of the preposition as well.

Another important processing principle proposed in Hawkins 1997:

(14) **Dependent Nodes Later**: If node B is dependent on node A for a property assignment, the processor prefers B to follow A.

(15) Examples of Dependent Nodes Later:

a. Fillers tend to precede gaps
   i. Wh-questions
   ii. Relative clauses
   iii. Control structures
   iv. A wide variety of ‘deletion’ constructions

b. Antecedents tend to precede anaphors

c. Topics tend to precede predications (cf. Japanese *wa*)

d. Restrictive relative clauses tend to precede appositives

e. Agents tend to precede patients

f. Quantifiers/operators tend to precede elements within their scope
Explanations based on information flow in discourse

(16) ‘Free word order’ languages, where all orders of S, V, and O are possible (Example from Russian):
   a. Lenin citiruet Marksa.
   b. Lenin Marksa citiruet.
   c. Citiruet Lenin Marksa.
   d. Citiruet Marksa Lenin.
   e. Marksa Lenin citiruet.
   f. Marksa citiruet Lenin.

(17) The principle of ‘Communicative Dynamism’ (Mathesius 1929; Firbas 1964; Firbas 1966; Vachek 1966): Information moves from ‘more thematic’ to ‘less thematic’ (essentially, from old information to new information) as the sentence progresses.

(18) The principle of ‘Communicative Task Urgency’ (Givón 1988): Information moves from new to old; from more important to less important.

(19) John is the person that I talked to.

(20) Three criteria for identifying convincing functional explanations:
   a. precise formulation
   b. demonstrable linkage between cause and effect
   c. measurable typological consequences

By these criteria, processing explanations do very well and information flow explanations very poorly.

The two competing theories about how information flow in discourse is supposed to influence grammar:

1. Communicative Dynamism (old information precedes new information)
2. Communicative Task Urgency (new information precedes old information)

Both can’t be right at the same time!

old-before new is generally true for VO languages
new-before-old is generally true for OV languages

Jack Hawkins’s parsing theory (EIC) predicts:
   — short-before-long for VO languages (e.g. in English, post-verbal PP’s tend to be ordered in terms of increasing length)
   — long-before-short for OV languages (e.g. in Japanese, -ga, -o, and –ni phrases tend to be ordered in terms of decreasing length)

In other words, where EIC predicts short-before-long, you get old-before new
   where EIC predicts long-before-short, you get new-before-old

But old information is shorter than new information —

So both ‘Communicative Dynamism’ and ‘Communicative Task Urgency’ are parsing effects
   — they have nothing to do with discourse principles affecting grammatical structure!
The fact that (short) topic constituents tend to precede (long) comment constituents, even in OV languages, might seem to support Communicative Dynamism

But this is also a processing effect:

a. New information is generally predicated of old information
   (Dependent Nodes Later principle at work — it is easier for the hearer if a ‘dependent’ node (the comment) follows an ‘independent’ node (the topic)

b. There are processing advantages for the speaker too (Arnold, Wasow, Losongco, and Ginstrom 2000)

General difficulties with information flow-based arguments designed to explain the nature of syntactic structure:

a. Word order facts can be reduced to the effects of parsing pressure to a great extent.

b. The desire to maintain structural parallelism is more important than the desire to model information flow.

c. There is little reason to believe that the conveying of information is a central ‘function’ of language.

d. Information flow-based explanations attribute to speakers and hearers more knowledge than they actually are likely to have.

e. The recognition of form takes precedence over the recognition of the information conveyed by that form.

I counted on your son.
I played in the rain.
I counted on your son yesterday.
I played in the rain yesterday.
I counted [on your good-for-nothing son] yesterday.
I played [in the hard driving rain] yesterday.

(26a) involves breaking up a syntactic/semantic dependency, (26b) does not.

REFERENCES
9b. Discussion of Bernard Comrie (C) ‘Reference-tracking: Description and explanation. (Sprachtypol Univ. Forsch. 51: 335-346; 1998).

OVERVIEW OF PAPER: C sets out to explain a set of generalizations about reference-tracking:

- If a language has 1st and 2nd person reflexives, it will also have 3rd person reflexives. Explanation: 1st and 2nd person referents are unique. But 3rd person referents are open-ended. It is therefore more ‘useful’ to have 3rd person reflexives, since they serve to narrow down the class of possible referents. Within a given domain, if a language makes a less-than-useful distinction, it will also make a useful one.¹
- Explicit coreference marking is more common in a local domain than in a nonlocal domain. Explanation: Such marking serves a useful function, since ‘the normal situation … is for an action that includes more than one participant to have an agent acting on a patient that is a distinct entity from the agent’. Noncoreference marking is more common in an extended domain than in a local domain. Explanation: The unmarked situation is to have referential continuity, so special marking is used to indicate deviation from this pattern.

SUMMARY EVALUATION: C’s first explanation seems reasonable, but a better functional explanation for the facts is possible that does not demand appeal to the notion of ‘usefulness’. There are serious problems with C’s second and third explanations.

ON GENERAL APPEALS TO ‘USEFULNESS’ IN FUNCTIONAL EXPLANATION

Unlike appeals to parsing-efficiency as a functional explanation, which are predictive, general appeals to ‘usefulness’ have a post-facto feel to them. One observes a generalization and comes up with a plausible story to account for it. But there is nothing predictive about accounts like C’s. One can come up with any number of things that it would be useful for a language to do (or for language in general to do), but which never occur. One needs a theory — which is now lacking — of why some useful features result in grammatical coding, and some do not.

Jack Hawkins has made the following perceptive observation on C’s first argument:

This explanation appeals, ultimately, to the number of occasions on which a formal distinction is useful in actual communication. I find it highly plausible. And yet it raises a fundamental question. When can we refer to such discourse considerations in general, and when not? After all, there are lots of useful real-world distinctions that could potentially be grammaticalized in languages which either never receive their own separate morphemes or structures, or which do so only rarely. Speakers regularly have occasion to refer deictically to objects

¹ This argument first appeared in print in Faltz 1977/1985: 120.
on their left as opposed to their right side, or vice-versa. Yet deictic systems across languages are not organized on this basis. The height of an object in relation to the speaker is another potentially useful distinguishing criterion, and here there are some languages that incorporate a height dimension into systems of spatial deixis (e.g. Daga and Dyirbal, cf. Anderson and Keenan 1985), but the majority operate only with a system of relative distance from the speaker (and in some cases from the hearer as well, cf. Anderson and Keenan 1985). Hence, although Comrie’s explanation for his reflexive universal is very convincing, it is somewhat after the fact. If other similar motivations for the grammaticalization of real-world distinctions do not result in distinctive grammatical morphemes or structures in this way, when can we appeal to a discourse-pragmatic explanation and when not? And can we assume that there will be a simple correlation between the degree of usefulness of some real-world distinction, and the frequency of its grammaticalization across languages? (Hawkins 1988: 12).

ON AMBIGUITY-REDUCTION AS A MOTIVATING FORCE

All three of C’s arguments are based, essentially, on ambiguity-reduction as a motivating factor in language. It is well known that language tolerates a tremendous amount of (formal) ambiguity, particularly in the reference system. Consider the inclusive/exclusive pronoun distinction. According to Nichols 1992, a minority of languages (about 42%) make this distinction. And the distinction is heavily areal — only in Australia, Oceania, and South America do more than 60% of languages manifest it. Yet the distinction is quite ‘useful’ from an ambiguity-reduction perspective. We all have been in many situations where we or someone else have said “We are going to do X, Y, and Z” and it has not been clear whether the person addressed was included in the “we” or not.

One could make the same point about the falling together of 2nd person singular and plural pronouns in English (where it is total) and in other European languages (where it affects the polite form).

Consider also the fact that a majority of the world’s languages are null subject. The point is that overwhelmingly (I think) argument reference can assigned from the discourse. Why should there be functional pressure at all to mark 3rd person coreference? In real life, how much more uncertain is the reference of a 3rd person object than that of a 1st or 2nd?

To be convincing, C should take a language like Old English, which does not have a special 3rd person reflexive form, and show that a significant amount of confusion arises from that fact.

A BETTER FUNCTIONAL EXPLANATION OF THE SAME FACTS

There is an explanation of the fact that the existence of 1st and 2nd person reflexive pronouns in a language implies the existence of 3rd person reflexive pronouns that does not involve problematic appeals to ‘usefulness’. In
languages that have reflexive pronouns in all three persons, 3\textsuperscript{rd} person reflexives are \textit{used more frequently} than 1\textsuperscript{st} and 2\textsuperscript{nd}. Consider English. In a million-word collection of British English texts (Johansson and Hofland 1989), 3\textsuperscript{rd} person singular reflexives were 5.8 times more likely to occur than 1\textsuperscript{st} person and 10.5 times more likely to occur than 2\textsuperscript{nd} person (see Table I).

<table>
<thead>
<tr>
<th>Reflexive pronoun</th>
<th>Number of occurrences in corpus</th>
</tr>
</thead>
<tbody>
<tr>
<td>myself</td>
<td>169</td>
</tr>
<tr>
<td>yourself</td>
<td>94</td>
</tr>
<tr>
<td>himself</td>
<td>511</td>
</tr>
<tr>
<td>herself</td>
<td>203</td>
</tr>
<tr>
<td>itself</td>
<td>272</td>
</tr>
<tr>
<td>TOTAL 3\textsuperscript{rd} PERS. SG.</td>
<td>986</td>
</tr>
</tbody>
</table>

Table I
Reflexive Pronoun Occurrence in English

Language users (for whatever reason) more frequently use identical subjects and objects in the 3\textsuperscript{rd} person than in the 1\textsuperscript{st} or 2\textsuperscript{nd}. Given the well-accepted dictum that more frequently used expressions are more likely to be lexicalized / grammaticalized than less frequently used ones, the implicational relationship among reflexive pronouns follows automatically. There is no need to appeal to ambiguity-reducing ‘usefulness’.

\textbf{SUPPOSE THAT THE FACTS WERE THE OPPOSITE …}

A tell-tale sign of a weak argument is that a reversal of the facts also lends itself to a ‘neat’ explanation. Let’s make the counterfactual assumption that many languages have reflexive pronouns in the 1\textsuperscript{st} and 2\textsuperscript{nd} person, but not in the 3\textsuperscript{rd} person. Can we ‘explain’ why such should be true?

The answer appears to be ‘yes’. First note that a long-standing assumption is that 1\textsuperscript{st} and 2\textsuperscript{nd} person is more ‘prototypical’ than 3\textsuperscript{rd} person. In particular, consider the Animacy Hierarchy (Silverstein 1976; Dixon 1979):

\begin{enumerate}
\item Animacy Hierarchy (Dixon 1979: 85): 1\textsuperscript{st} Person Pronoun > 2\textsuperscript{nd} Person Pronoun > 3\textsuperscript{rd} Person Pronoun > Proper Noun > Human Common Noun > Animate Common Noun > Inanimate Common Noun
\end{enumerate}

Many grammatical processes appeal to degree of animacy, the most prototypically animate nouns being 1\textsuperscript{st} and 2\textsuperscript{nd} person. This hierarchy appears to be implicated in phenomena as diverse as case-marking and the number of honorific pronoun forms (for the latter, see Brown and Levinson 1978 /1987). Now, it is a staple of markedness / prototype theory that the more prototypical /less marked member of an opposition shows more morphosyntactic complexity than the less prototypical / more marked member. Consider German gender distinctions, for example. In the nominative case, German makes three overt distinctions in the (unmarked) singular, while there is only one form for the (marked) plural. It ‘stands to reason’, then, that more overt distinctions (e.g. \textit{me} vs. \textit{myself}) would be made for pronouns in the 1\textsuperscript{st} and 2\textsuperscript{nd} person than in the 3\textsuperscript{rd}.
This bad (but why bad?) argument might be supported by the observation that in informal conversation (but not in more formal genres) 1st and 2nd person reference is more common than 3rd:

Egophoric reference is defined as reference to speech act participants and generic reference. As shown by adult conversational data from Swedish, English, and Spanish, and longitudinal data from one Swedish child, the majority of all animate arguments of verbs in conversation are egophoric. (Dahl 2000: 37).

For other studies that confirm this claim, see Ariel 2000.

Given that frequency correlates with both prototypicality and morphosyntactic elaboration, it ‘follows’ that languages would be more likely to have 1st and 2nd person reflexives than 3rd.

C’S CORE EXPLANATION AND EXPANDED BINDING DOMAINS

In C’s analysis, non-coreference is marked in local domains because ‘the normal situation … is for an action that includes more than one participant to have an agent acting on a patient that is a distinct entity from the agent’. As C observes, many languages allow an extended binding domain for reflexives. C implies that what we have here is an innocuous extension of a functionally-motivated formal generalization to a situation where the function is not directly served. (An analogy might be drawn from case-marking. It is functionally useful for a language to distinguish nominative from accusative case. But no harm is done for a case-marking language to mark nominative case on NP subjects of intransitive verbs, even though the function of distinguishing subjects from objects does not apply.)

But extending the domain of reflexivization is downright counter-functional, given C’s theory. C claims that the function of the reflexive pronoun is to signal a non-prototypical situation. More often than not, however long-distance reflexives will be used in prototypical situations. So consider the following sentence from Korean (from Moon 1995):

(2) John-i-un [Tomj-i caki,i/l-lul pipanhayss-ul ttay] amwu-malto an hayssta

John-top Tom-nom self-acc criticize-comp when anything did not say 'John, did not say anything when Tom, criticized him,/himself/

John is not only a possible antecedent for the reflexive caki, but is actually the preferred antecedent. This fact is quite troublesome for C’s account. The sentence in which caki occurs is ‘normal’ according to C, since agent and patient have different reference. Yet the reflexive pronoun occurs in it.

For Korean (and possibly for other languages with long-distance anaphors) the farthest antecedent is the most preferred — that fact runs counter to the explanation that C offers for the role of anaphors.
So the question is why languages would develop extended binding domains (long-distance anaphors) in the first place, if the motivation for reflexive pronouns is what C says it is.

**TOWARDS A BETTER FORMAL (AND THEREFORE FUNCTIONAL?) ACCOUNT OF THE BINDING FACTS THAN THAT PROVIDED BY C**

In many ways, the account of anaphor binding in C is remarkably similar to the now ‘classic’ binding theory of Chomsky 1981. Both assume that the normal ‘unmarked’ situation for anaphor binding is that it take place in a very restricted local domain. The evident fact that many, probably most, languages, have a much more expanded domain is an embarrassment for both approaches.

A more adequate approach to binding is provided in Burzio 1996. For Burzio, reflexives are *always* possible, no matter how far from the antecedent, unless blocked by some relevant intervening category. There is an implicational relationship among blockers, going from strongest (on the left) to weakest (on the right):

(3) **Blocking hierarchy for anaphor binding:** tensed indicative clauses > tensed subjunctive clauses > infinitival clauses > noun phrases / perception verb complements

Some other facts that are built into Burzio’s theory:

(4) **Languages that do not have verb agreement, such as Chinese, Japanese, and Korean, permit anaphors as subjects of tensed clauses:**

(5) a. *John said that himself would come.*

   b. Zhangsan, shuo [ziji, hui lai]

   ‘Zhangsan said [self will come]’

(6) **Long distance anaphors are consistently monomorphemic, while those that do have strict locality conditions are bimorphemic:**

   a. English has only one (inflected) anaphor, *him/her/itself*, which must be local.

   b. Korean has monomorphemic anaphors, *caki* and *casin*, which can be long-distance and an compound anaphor, *cakicasin*, which must be local.

(7) **There is a strong correlation between the class of elements that can serve as antecedents for anaphors and those elements that block anaphor binding:**

(8) Experiencers in Chinese and other languages can serve as antecedents for anaphor binding, but when they occur between a potential antecedent and an anaphor, they block binding.
Possessive anaphora are a subcase of long-distance anaphora.

Hence in English there is no form like (10):

(10) *Mary read herself’s book.

Burzio ties all of these observations/claims together in a formal analysis that need not concern us here. But the point is, if he is (largely) correct, then the ultimate functional explanation of the generalization(s) governing the distribution of reflexives would have to be very different from that envisioned by C.

ON THE INDEPENDENT NEED FOR FORMAL ANALYSIS

C never dismisses the need for formal analysis of reference-tracking systems. He merely states that ‘an explanatory account … will often have to go beyond the purely formal statement of the generalizations …’ (p. 346). But since a commonly-heard view is that, given a functional explanation, one doesn’t need formal principles, it is worth pointing out why such principles are indeed needed.

Take English, which comes as close as any language to instantiating C’s prototypical agent-action-patient situation where we find reflexives. Interestingly, however, we find reflexives possible where the reflexive is in an oblique phrase (11a-b) and even (in 11b), where the antecedent is a direct object:

(11) a. John smeared the peanut butter on himself.
    b. I talked to Mary about herself.

The reflexive need not even have a thematic relation to the verb, as ‘exceptional case marking’ sentences show:

(12) Sally believes herself to have won.

Most interestingly of all, English allows a reflexive which is a (logical) subject of a lower clause and where there is not even a surface grammatical relation with the verb whose subject is antecedent of the reflexive:

(13) Tom would prefer [CP for himself [IP to be the one chosen]]

All of these cases follow from the formal account presented in the classic binding theory. To the extent that this account gets the facts right, it would appear that such an account is needed, since C’s functional explanation is much too broad-brush to capture the details of anaphor binding in English.

Another reason that a formal account is needed is that there is not necessarily a one-to-one relationship between formal generalizations and functional explanations. A centerpiece of the classic binding theory is the (a priori unexpected) parallel between constraints on anaphora and those on ‘NP
movements’ (passive, raising, etc.). Both are constrained by both c-command and locality:

(14) a. S
    NP    S
    NP    NP

   legitimate movement operations

b. S
    NP    S
    NP    NP

   legitimate anaphoric relations

Assuming that this generalization is a real one, it would never have been discovered by a search for purely ‘functional’ principles, in the absence of a theory that searches for formal generalizations independently of functional explanations.

REFERENCES


10a. Typology and generative grammar

Typological facts are irrelevant to generative grammar

(1) This claim does not mean to imply:
   (a) that one does not need to study as many languages as possible
   (b) that typology lacks theoretical interest

<table>
<thead>
<tr>
<th></th>
<th>VSO</th>
<th>SVO</th>
<th>SOV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prep</td>
<td>6</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Postp</td>
<td>0</td>
<td>3</td>
<td>11</td>
</tr>
</tbody>
</table>

Correlations between word order and adposition order (Greenberg 1963)

Table 1

(2) UG characterizes the notion ‘possible human language’, not the notion ‘probable human language’.

Chomsky’s early view of Greenberg’s work:
Insofar as attention is restricted to surface structures, the most that can be expected is the discovery of statistical tendencies, such as those presented by Greenberg (1963). (Chomsky 1965: 118; emphasis added)

Chomsky’s more recent view:
There has also been very productive study of generalizations that are more directly observable: generalizations about the word orders we actually see, for example. The work of Joseph Greenberg has been particularly instructive and influential in this regard. These universals are probably descriptive generalizations that should be derived from principles of UG. (Chomsky 1998: 33; emphasis added)

<table>
<thead>
<tr>
<th>VO correlate</th>
<th>OV correlate</th>
</tr>
</thead>
<tbody>
<tr>
<td>adposition - NP</td>
<td>NP - adposition</td>
</tr>
<tr>
<td>copula verb - predicate</td>
<td>predicate - copula verb</td>
</tr>
<tr>
<td>‘want’ - VP</td>
<td>VP - ‘want’</td>
</tr>
<tr>
<td>tense/aspect auxiliary verb - VP</td>
<td>VP - tense/aspect auxiliary verb</td>
</tr>
<tr>
<td>negative auxiliary - VP</td>
<td>VP - negative auxiliary</td>
</tr>
<tr>
<td>complementizer - S</td>
<td>S - complementizer</td>
</tr>
<tr>
<td>question particle - S</td>
<td>S - question particle</td>
</tr>
<tr>
<td>adverbial subordinator - S</td>
<td>S - adverbial subordinator</td>
</tr>
<tr>
<td>article - N’</td>
<td>N’ - article</td>
</tr>
<tr>
<td>plural word - N’</td>
<td>N’ - plural word</td>
</tr>
<tr>
<td>noun - genitive</td>
<td>genitive - noun</td>
</tr>
<tr>
<td>noun - relative clause</td>
<td>relative clause - noun</td>
</tr>
<tr>
<td>adjective - standard of comparison</td>
<td>standard of comparison - adjective</td>
</tr>
<tr>
<td>verb - PP</td>
<td>PP - verb</td>
</tr>
<tr>
<td>verb - manner adverb</td>
<td>manner adverb - verb</td>
</tr>
</tbody>
</table>

Correlation pairs reported in Dryer (1992)

Table 2

(3) $W_1 - W_2 - W_3 - ... - W_n \rightarrow W_n - ... - W_3 - W_2 - W_1$
(4) Guiding assumption in mainstream generative work: there is no disparity between the notions ‘typologically significant generalization’ and ‘linguistically (i.e. grammatically) significant generalization’

Chinese is consistently head final except in rule expanding X’ to X⁰ (if the head is verbal it precedes the complement)

(5) (Huang 1994):
   a. XP —> YP  X’
   b. X’ —> YP  X’
   c. X’ —> c’.  X⁰ YP iff X = [+v]
      c''. YP X⁰ otherwise

(6) (Kayne 1994)

\[
\begin{align*}
&\text{CP} \\
&\quad [\text{spec, CP}] \\
&\quad \text{IP} \\
&\quad \text{C} \\
&\quad \text{TP}_1
\end{align*}
\]

(7) *NDem & NumN predicted by Kayne
But Berber, Welsh, Hebrew, Zapotec, etc. have this correlation
These languages have and extra movement of demonstratives (Cinque 1996)

(8) Possible orders of V, direct object, complement PP₁, and adjunct PP₂:
   a. PP₂ PP₁ NP V  
   b. PP₂ PP₁ V NP  
   c. PP₂ NP V PP₁  
   d. PP₂ V NP PP₁  
   e. PP₁ NP V PP₂  
   f. PP₁ V NP PP₂  
   g. NP V PP₁ PP₂  
   h. V NP PP₁ PP₂

<table>
<thead>
<tr>
<th>HEADEDNESS</th>
<th>THETA</th>
<th>CASE</th>
<th>LANGUAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. final</td>
<td>left</td>
<td>left</td>
<td>Japanese</td>
</tr>
<tr>
<td>b. final</td>
<td>left</td>
<td>right</td>
<td>Chinese (future)</td>
</tr>
<tr>
<td>c. final</td>
<td>right</td>
<td>left</td>
<td>*</td>
</tr>
<tr>
<td>d. final</td>
<td>right</td>
<td>right</td>
<td>Chinese (pres)</td>
</tr>
<tr>
<td>e. initial</td>
<td>left</td>
<td>left</td>
<td>Kpelle (past)</td>
</tr>
<tr>
<td>f. initial</td>
<td>left</td>
<td>right</td>
<td>*</td>
</tr>
<tr>
<td>g. initial</td>
<td>right</td>
<td>left</td>
<td>Kpelle (present)</td>
</tr>
<tr>
<td>h. initial</td>
<td>right</td>
<td>right</td>
<td>English</td>
</tr>
</tbody>
</table>

Combinations of the headedness, direction of theta-role assignment, and direction of case assignment parameters (Travis 1989)
Table 3

For specific proposals concerning marked values to entail testable claims, these claims will have to hold in an ‘external’ domain, a domain other than that of the distribution of morphemes or grammatical well-formedness. Claims to explanatory adequacy will have to be grounded in such domains. Natural candidates for such a
domain wherein markedness proposals make empirically testable claims are
language change and acquisition. (Lightfoot 1979: 76-77; emphasis added)

[In] determining which notions are encoded in a language’s
morphology, the child is faced with a formidable search problem ...
[B]y imposing a weighting on the child’s hypotheses, one could
account for the large disparities in the prevalence of various
grammatical encodings in the world’s languages, and in the speed of
acquisition of various encodings by children. (Pinker 1984: 168-171)

One intriguing possibility is that the relative accessibility for children
of alternative schemes for partitioning meaning in a given conceptual
domain is correlated with the frequency with which these schemes are
instantiated in the languages of the world. ... It is plausible that relative
frequency is correlated with ‘ease’ or ‘naturalness’ for the human
mind. (Bowerman 1985: 1306).

Three hypotheses linking typological generalizations to aspects of I-language:

(9) a. Cross-linguistically frequent properties of language are reflected by
correspondingly simple (unmarked) properties of grammars.
   b. Cross-linguistically frequent properties of language are acquired
ey early by the child.
   c. Cross-linguistically frequent properties of language are
diachronically stable.

If (9a-c) were right, then typology would be relevant to gram theory in two
complementary ways:
   1. One could appeal to grammatical theory to explain typological
distribution
   2. Typological distribution would be a reliable heuristic for
grammatical analysis
But they are not right!

Emonds 1980: Irish and other VSO languages are underlingly SVO: they have rule
that fronts the verb. Their grammars are ‘more complicated, therefore rarer’ (p. 44).

Preposition-Stranding:
(10) a. Who did you talk to?
    b. Mary was spoken to.

Stranding grammars are simpler than non-stranding grammars:
(11) a. NON-STRANDING LANGUAGES: The lexical categories N, V, and
A are proper governors. The lexical category P is not a proper governor.
    b. STRANDING LANGUAGES: All four lexical categories are proper
governors.

A popular reanalysis approach to stranding:
(12) a. You talked \text{PP}[to who] > You \text{V}[talk to] who > Who did you \text{V}[talk to] ei?
    b. e was spoken \text{PP}[to Mary] > e was \text{V}[spoken to] Mary > Mary was
\text{V}[spoken to] ei

But reanalysis demands implausible lexical items:
(13) a. Which shoes did you [walk across Europe in] ___? (Jones 1987)
b. Which of the two knives did you [pay twice for] ___? (Inada 1981)

Gapping doesn’t treat the V-P complex as a verb:
(14) a. *John looked at Mary and Bill ___ Sue.
b. John looked at Mary and Bill ___ at Sue.

Nor does Heavy NP-shift:
(15) a. John looked at [the woman he loved] very often.
b. John looked very often [at the woman he loved]
c. *John looked at very often [the woman he loved].

Extraposition out of a lexical item would be necessary:
(16) What did you [talk to that guy ___ about] ___ who was here yesterday?

Mutually incompatible analyses, where wh-movement and passive both apply:
(17) a. Which problems has Harry been [[talked to] e about] e?
b. Who would you like to be [[sung to] e by] e?

If P is a proper governor, then V and P don’t have to be adjacent:
(18) a. Who did you give all those books about golf to?
b. Which burner did you leave the pot on?

Most interesting prediction of my analysis: stranding is possible with extraction of NP from PP adjuncts to VP:

(19)

```
VP
  /\           /
 V'           PP
     /\       /\       /
  V     (NP)  P     NP
       [+wh]
```

This is correct:
(20) a. Which shoes did you walk across Europe in?
b. Which ball park did Ruth hit the most home runs in?
c. Which knife shall we use to cut the turkey with?
d. Which red-headed man is Mary standing beside?
There is little or no correlation between age of acquisition and typological frequency:

a. American children acquire P-stranding before pied-piping (Karin Stromswold, personal communication)

b. French-speaking children have the verb-raising parameter from the earliest multi-word utterances (Pierce 1992)

c. there is no period during which German-speaking children fail to set the V2 parameter (Poeppel and Wexler 1993)

d. the null subject parameter is set very early, regardless of whether the language is null subject or not (Valian 1991)

e. children acquiring English, German, and French evidence strong knowledge of locality in wh-extraction domains at early ages (Roeper and De Villiers 1994)

The Strong Continuity Hypothesis: The parameters of UG are set early and are not subject to change in the course of development (see especially Lust, Suñer, and Whitman 1994)

Early Immediate Constituents (EIC) (Hawkins 1994): The parser prefers orderings of elements that lead to rapid recognition of the immediate constituents of the phrase being processed.

The Greenbergian correlations hold better at surface level than at deep level.
(25) German and Dutch:
a. main clause declaratives: SVO
b. subordinate clauses: SOV
Virtually all generative grammarians have argued underlyingly they are underlyingly SOV (Bach 1962; Bierwisch 1963; 1966; Koster 1975; Bennis and Hoekstra 1984)

(26) **Structure Preserving Constraint** (Emonds 1976): Rules that distort phrase structure configurations apply only in main clauses

If this constraint is right, German and Dutch *have to be* SOV.

(27) a. 

```
  IP
   \   
  /   
 subj V obj
```

SVO underlying order — SOV underlying order

Movement violates SPC

b. 

```
  IP
   \   
  /   
 subj V obj
```

Movement obeys SPC

But German and Dutch are not *typologically* SOV

(28) **SURFACE ORDER AND TYPOLOGICAL GENERALIZATIONS**

a. Dutch and German are more ‘consistent’ at the surface than underlyingly.

b. Despite Bach (1970a), Amharic is typologically more like an OV than a VO language.

c. Languages with discourse-governed word order (where it is hard to make a case for one or another deep order of elements) manifest the correlations associated with that surface order having the highest text frequency (see Table 5).

d. Some languages have been argued by generativists to have *no* underlying order of subject, object, and verb (Kiss 1987 for Hungarian, Hale 1992 for Warlpiri, etc.) Clearly, in such languages the Greenbergian correlations cannot be captured at D-structure.
<table>
<thead>
<tr>
<th>Language</th>
<th>Percentage</th>
<th>Gen, Po, Rel, Clause-final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ute</td>
<td>72%</td>
<td>GenN, Po</td>
</tr>
<tr>
<td>Tlingit</td>
<td>67%</td>
<td>GenN, Po, RelN</td>
</tr>
<tr>
<td>Huallaga Quechua</td>
<td>69%</td>
<td>GenN, Po, RelN</td>
</tr>
<tr>
<td>Trumai</td>
<td>65%</td>
<td>GenN, Po</td>
</tr>
<tr>
<td>Koryak</td>
<td>66%</td>
<td>GenN, Po</td>
</tr>
<tr>
<td>Tacana</td>
<td>86%</td>
<td>GenN, Po, Clause-final subordinator</td>
</tr>
<tr>
<td>Takelma</td>
<td>85%</td>
<td>GenN, Po</td>
</tr>
<tr>
<td>Hupa</td>
<td>53%</td>
<td>GenN, Po</td>
</tr>
<tr>
<td>Cherokee</td>
<td>(more frequent)</td>
<td>GenN, Po</td>
</tr>
<tr>
<td>Korana</td>
<td>89%</td>
<td>GenN, Po, Clause-final Q, etc.</td>
</tr>
</tbody>
</table>

Frequency of OV order and OV characteristics among languages with purely discourse-governed orders (Dryer 1989)

Table 5

(29) Transformational theories in which ordering of elements is a low-level phenomenon: Sanders 1970; Staal 1967; Peterson 1971; Hudson 1972;

... and Chomsky 1995.

Language design appears to be problematic from a parsing-theoretic perspective, though elegant regarded in isolation from considerations of use. (Chomsky 1991: 448)

REFERENCES


11a. OPTIMALITY THEORY AND FUNCTIONAL EXPLANATION

<table>
<thead>
<tr>
<th>name</th>
<th>grammatical constraint</th>
<th>corresponding user constraint</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAY (Grimshaw 1997, Speas 1997)</td>
<td>“Do not move”</td>
<td>Leaving material in canonical position helps the hearer to identify grammatical relationships and reduces processing costs for the speaker</td>
</tr>
<tr>
<td>TELEGRAPH (Pesetsky 1998)</td>
<td>“Do not pronounce function words”</td>
<td>Leaving out function words reduces pronunciation costs for the speaker in a way that is minimally disruptive for understanding by the hearer</td>
</tr>
<tr>
<td>RECOVERABILITY (Pesetsky 1998)</td>
<td>“A syntactic unit with semantic content must be pronounced unless it has a sufficiently local antecedent”</td>
<td>Omitting a meaning-bearing element in pronunciation makes the hearer’s task of extracting the intended meaning from the speech signal very difficult unless it can be inferred from the context</td>
</tr>
</tbody>
</table>

Table I
Some OT constraints and their possible functions (Haspelmath 1999)

(1) **Thematic Hierarchy** Agent > Beneficiary > Recipient/Experiencer > Instrumental > Theme/Patient > Location
(2) **Relational Hierarchy** Subject > Direct Object > Indirect Object > Oblique > Genitive > Object of Comparative
(3) **Animacy Hierarchy** 1st Person Pronoun > 2nd Person Pronoun > 3rd Person Pronoun > Proper Noun > Human Common Noun > Animate Common Noun > Inanimate Common Noun
(4) **Person Hierarchy** Local (1st, 2nd) > 3rd
(5) **Prominence Hierarchy** Discourse Prominent > Not Discourse Prominent
(6) **Relational**: Subject > Object > Oblique
(7) **Person**: Local (1st, 2nd) > 3rd
(8) **Thematic Role**: Agent > Patient
(9) **Prominence**: Discourse Prominent (X) > Not Discourse Prominent (x)

(10) a. *Su/Pat
    b. *Su/x
    c. *Obl/Local

(11) a. *Su/Pat » *Su/Ag
    b. *Su/Ag » *Su/Pat
Tableau I

*Su/x *Su/Pat *GR/Pers

ACTIVE
Agt/Su/3/x-Pat/Oj/1/X *!

PASSIVE
Pat/Su/1/X-Agt/Obl/3/x *

Table I

English (prominent patient) (Aissen 1999: 689)

(12) ** Animacy: Human > Animate > Inanimate
(13) ** Definiteness: Personal Pronoun > Proper Noun > Definite NP > Indefinite Specific NP > Non-specific NP

(14) a. *Oj/Hum » *Oj/Anim » *Oj/Inam
   b. *Oj/Pro » *Oj/PN » *Oj/Def » *Oj/Spec » *Oj/NSpec

(15) *φc ‘Star Zero’: Penalizes the absence of a value for the feature CASE.

Local conjunction of *φc with the
subhierarchy on object animacy (14a)

*Oj/Hum & *φc »
*Oj/Anim & *φc »
*Oj/Inam & *φc »

Local conjunction of *φc with the
subhierarchy on object definiteness
(14b)

*Oj/Pro & *φc »
*Oj/PN & *φc »
*Oj/Def & *φc »
*Oj/Spec & *φc »
*Oj/NSpec & *φc »

Table II

Local conjunction of *φc with object-oriented subhierarchies (Aissen 2000: 9)

(16) **STRUCc: penalizes a value for the morphological category CASE

Hebrew specific, indefinite patient

Patient: specific indefinite

<table>
<thead>
<tr>
<th>*Oj/Def &amp; *φc</th>
<th>*STRUCc</th>
<th>*Oj/Spec &amp; *φc</th>
<th>*Oj/NSpec &amp; *φc</th>
</tr>
</thead>
</table>
| Oj: specific indefinite CASE: ACC
| *Oj/Def & *φc | *STRUCc | *Oj/Spec & *φc | *Oj/NSpec & *φc |
| Oj: specific indefinite CASE: none |

Tableau II

Hebrew specific indefinite patients (Aissen 2000: 14)

<table>
<thead>
<tr>
<th>Patient: specific indefinite</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Oj/Def &amp; *φc</td>
</tr>
<tr>
<td>*STRUCc</td>
</tr>
<tr>
<td>*Oj/Spec &amp; *φc</td>
</tr>
<tr>
<td>*Oj/NSpec &amp; *φc</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Oj: specific indefinite CASE: ACC</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Oj/Def &amp; *φc</td>
</tr>
<tr>
<td>*STRUCc</td>
</tr>
<tr>
<td>*Oj/Spec &amp; *φc</td>
</tr>
<tr>
<td>*Oj/NSpec &amp; *φc</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Oj: specific indefinite CASE: none</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Oj/Def &amp; *φc</td>
</tr>
<tr>
<td>*STRUCc</td>
</tr>
<tr>
<td>*Oj/Spec &amp; *φc</td>
</tr>
<tr>
<td>*Oj/NSpec &amp; *φc</td>
</tr>
</tbody>
</table>

Tableau III

Turkish specific indefinite patients (Aissen 2000: 15)

(17) ** NO LEXICAL HEAD MOVEMENT: A lexical head cannot move
(18) LEFT EDGE (CP): The first pronounced word in CP is a function word related to the main verb of that CP

(19) a. Je crois que Pierre a faim.
    b. *Je crois Pierre a faim.

(20) a. I found a book for you to think about.
    b. *I found a book you to think about.

<table>
<thead>
<tr>
<th>name</th>
<th>grammatical constraint</th>
<th>corresponding user constraint</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO LEXICAL HEAD MOVEMENT (Grimshaw 1997)</td>
<td>&quot;A lexical head cannot move&quot;</td>
<td>Moving a lexical head away from its dependents makes it difficult for the hearer to pair heads and dependents</td>
</tr>
<tr>
<td>LEFT EDGE (CP) (Pesetsky 1998)</td>
<td>&quot;The first pronounced word in CP is a function word related to the main verb of that CP&quot;</td>
<td>Not explicitly marking embedded clauses in terms of their boundaries and finiteness impedes rapid recognition of their meaning and role</td>
</tr>
</tbody>
</table>

Table Ia
Possible functions for NO LEXICAL HEAD MOVEMENT and LEFT EDGE (CP)

(21) OP-SPEC: Syntactic operators must be in specifier position

(22) HEAD-RIGHT: Heads uniformly follow their complements and adjuncts

(23) Versions of the Thematic Hierarchy:
   a. Fillmore 1968: 33
      Agent > Instrumental > Objective
   b. Jackendoff 1972
      Agent > Location/ Source/ Goal > Theme
   c. Ostler 1980
      (relational predicates) Theme > Goal > Source > Path
      (actional predicates) Source > Path > Theme > Goal
   d. Givón 1984: 139
      Agent > Dative/ Beneficiary > Patient > Locative > Instrument /
      Associative > Manner
   e. Kiparsky 1985: 20
      Agent > Source > Goal > Instrument > Theme > Locative
   f. Carrier-Duncan 1985: 7
      Agent > Theme > Goal/ Source/ Location
   g. Larson 1988: 382-383
      Agent > Theme > Goal > Location (and other obliques)
   h. Wilkins 1988: 211
      Agent > Patient > Location/ Source/ Goal > Theme
i. Randall 1988: 138 (for effects of lexical rules on argument structure)
Theme > Agent > Instrument/ Source/ Path/ Goal/ Location/ …

Agent > Beneficiary > Recipient/ Experiencer > Instrumental > Theme/ Patient > Location

k. Baker 1989: 544
Agent > Instrument > Patient/ Theme > Goal/ Location

l. Grimshaw 1990: 8
Agent > Experiencer > Goal/ Source/ Location > Theme

m. Jackendoff 1990: 261
Agent > Patient/ Beneficiary > Theme > Source/ Goal/ Reference Object > Identificational Goal/ Reference Object

definition

n. Langacker 1990
Agent > Instrument > Patient/ Mover/ Experiencer (‘Energy Flow Hierarchy’ for subject choice)
Agent > Experiencer > Other (‘Initiative Hierarchy’)

o. Speas 1990: 16
Agent > Experiencer > Theme > Goal/ Source/ Location > Manner/ Time

p. Dowty 1991; Rugemalira 1994
Proto-Agent > Proto-Patient

q. Kiefer 1995
Actor > Agent > Beneficiary > Theme/ Patient > Instrument

r. Van Valin and Lapolla 1997: 127 (continuum in terms of LS (=logical structure) argument positions) — this is an elaboration of Foley and Van Valin 1984:59:
Agent > Effector/ Mover/ User/ etc. > Location/ Perceiver/ Cognizer/ Experiencer/ etc. > Theme/ Stimulus/ Implement/ etc. > Patient/ Entity

(24) a. Mary attempted to take a vacation.
b. Bill told Mary to come visit him.
c. I persuaded John to see the doctor.
d. We taught the dog to roll over.
e. Sam promised Mary to write more often.
f. Tom required Alice to hand in the assignment.

(25) Relational Hierarchy: Subject > Direct Object > Indirect Object > Oblique > Genitive > Object of Comparative

(26) a. John kills the ducklings.
b. The ducklings are killed by John.

(27) The higher in prominence a direct object, the more likely it is to be overtly case marked.

(28) Animacy Hierarchy: 1st Person Pronoun > 2nd Person Pronoun > 3rd 
Person Pronoun > Proper Noun > Human Common Noun > 
Animate Common Noun > Inanimate Common Noun

(29) Definiteness Hierarchy: Personal Pronoun > Proper Noun > Definite 
NP > Indefinite Specific NP > Non-specific NP

(30) Grammars in which object case is overtly marked choose a point on the 
Animacy and/or Definiteness Hierarchies and mark case at that point and at 
every point of higher prominence.

b. Turkish: Object case-mark point d on the Definiteness Hierarchy.

(32) If an individual can lift a weight of \( x \) pounds, then he or she can lift a lighter weight.

(33) a. Smith can lift 200 pounds.
b. Jones can lift 300 pounds.

REFERENCES

Aissen, Judith. 1999. 'Markedness and Subject Choice in Optimality Theory', 
*Natural Language and Linguistic Theory* 17, 673-711.
Unpublished ms., UCSC.
Alsina, Alex. 1996. 'Passive Types and the Theory of Object Asymmetries', 
*Natural Language and Linguistic Theory* 14, 673-723.
Baker, Mark C. 1989. 'Object Sharing and Projection in Serial Verb 
Constructions', *Linguistic Inquiry* 20, 513-553.
Bresnan, Joan W. and Jonni Kanerva. 1989. 'Locative Inversion in Chichewa: A 
Case Study of Factorization in Grammar', *Linguistic Inquiry* 20, 1-50.
Bresnan, Joan W. and Liobi Moshi. 1990. 'Object Asymmetries in Comparative 
Carrier-Duncan, Jill. 1985. 'Linking of Thematic Roles in Derivational Word 
Formation', *Linguistic Inquiry* 16, 1-34.
Dowty, David R. 1991. 'Thematic Proto-Roles and Argument Selection', 
*Language* 67, 547-619.
*Universals in Linguistic Theory*, Holt, Rinehart and Winston, New York, 
pp. 1-90.
Benjamins, Amsterdam.


