Phrase Structure Generalizations

Jim Blevins

University of Cambridge/Collegium de Lyon

`jpb39@cam.ac.uk`

Structure and Evidence in Linguistics:

A Workshop in Honour of Ivan A. Sag

Stanford University, 2013.04.30

Syntactic representations

- Modern syntactic theories are largely theories of representations and of the devices that produce those representations.
- This is a recent conception, even in the Americanist school:
  - There is no theory of syntactic representations in Bloomfield.
  - The descriptivist 'procedures' of IC analysis consist mainly of 'recipes' for segmenting expressions into smaller parts.
- Moreover, the simple representational model that ultimately emerged in the Post-Bloomfieldian tradition was constrained by unargued assumptions about 'economy' or 'scientific compactness'.

Lexical economy

The inflectional forms are relatively easy to describe, since they occur in parallel paradigmatic sets; the traditional grammar of familiar languages gives us a picture of their inflectional systems. It may be worth noticing, however, that our traditional grammars fall short of scientific compactness by dealing with an identical feature over and over again as it occurs in different paradigmatic types. Thus, in a Latin grammar, we find the nominative-singular sign -s noted separately for each of the types amīcus 'friend', lapis 'stone', dux 'leader', tussis 'cough', manus 'hand', faciēs 'face', when, of course, it should be noted only once, with a full statement as to where it is and where it is not used. (Bloomfield 1933: 238)

▶ Scientific compactness is sensitive to the redundancy of elements, but not to the complexity of the statement of their distribution.

Syntactic economy

- A Bloomfieldian lexicon consists of maximally simple units ('morphemes') that are combined by arbitrarily complex rules.

- Transformational analysis projects this conception onto syntax:

  [T]he sentences of English can be characterized by a small family of elementary sentence structures and a few small families of elementary transformations on these. (Harris 1964: 115)

  [W]e can apparently define a grammatical transformation in terms of ... a sequence of elementary transformations drawn from a base set including substitutions, deletions and adjunctions. It seems also that these form larger repeated units ... (Chomsky 1965: 147)

- Scientific compactness is achieved within a transformational model by applying elementary operations to simple initial representations.

The representational alternative

- The phrase structure tradition that runs from early GPSG through current models of SBCG was instrumental in articulating an alternative conception of syntax, one in which richer (and more precisely characterized) representations play a far greater role.
- Yet this tradition faced a fundamental analytical challenge: Which aspects of a phrase structure grammar should be generalized?
  - PSGs provide no systematic representation of feature information; even abandoning the distribution class features of Harris (1951).
  - PSGs admit only restricted types of constituent structures, excluding entire classes of traditional analyses with non-contiguous units.

Consolidation and extension

- ▶ We can look back to the descriptive coverage and formal clarity achieved by extending the treatment of feature information.
- ▶ We can also look forward to further advances obtainable by generalizing feature and constituent structures in parallel.

Subject-auxiliary 'inversion' (Gazdar et al. 1982)

▶ A refined feature classification of verbal elements provides a
description of the form and distribution of auxiliary constructions:
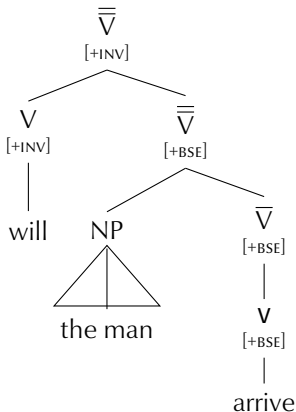
Symbiotic generalization

- This analysis improved on coverage and precision of 'affix hopping'.
- Yet the use of features such as [±ɪɴᴠ] was independent of the assumptions about constituent structure in this representation:
  - It is critical is that *will* functions as the head of the inverted clause.
  - But the analysis does not depend on the assumption that *will* is an IC of the clause or that the other IC is a 'clause-like' unit *the man arrive*.
- Hence the central insight of the GPSG analysis carries over to an analysis in which the immediate constituents of the construction are *the man* and *will arrive* (Hockett 1958,Gleason 1961, etc.).
- A contemporary version of this traditional analysis is expressed within the Head Grammar formalism of Pollard (1984), reflecting the brief convergence of phrase structure and categorial approaches.

Subject-auxiliary 'wrapping'

- ▶ The IC analysis can be exhibited as a derivation tree (though the same structure can also be represented as a graph or as a DOM list):

will the man arrive, $_{S[+INV]}$

the man, $_{NP}$ will arrive, $_{VP[+INV]}$

will, $_{V[+INV]}$ arrive, $_{V[+BSE]}$
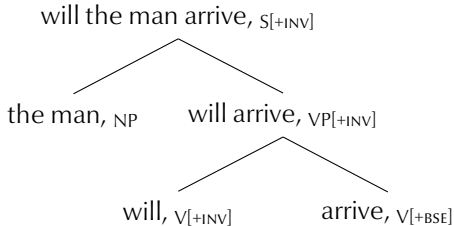
Uninverted (subject-predicate) order

▶ A [–INV] auxiliary will occur with a basic (non-inverted order) in a
clause that exhibits an identical immediate constituent structure:

the man will arrive, S[–INV]

the man, NP   will arrive, VP[–INV]

will, V[–INV]   arrive, V[+BSE]
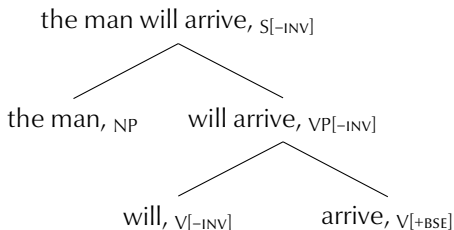
Phrasal verbs as complex transitives I

► When we turn our attention to other low-hanging fruit in the
traditional literature, we notice that a similar generalization permits
a phrase structure analysis of phrasal verbs (Wells 1947):

pick the cat up, $_{VP}$

the cat, $_{NP}$ pick up, $_{V}$

pick, $_{V}$ up, $_{PRT}$
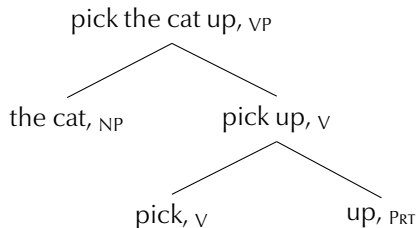
Phrasal verbs as complex transitives II

- 'Particle shift' can again be treated as structurally neutral:

pick up the cat, $_{VP}$

the cat, $_{NP}$ pick up, $_{V}$

pick, $_{V}$ up, $_{PRT}$
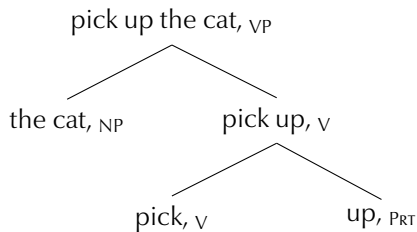
Structural ambiguity as a source of variation

- In this case, the order is not conditioned by feature variation, [+ɪɴᴠ] vs [–ɪɴᴠ], but is attributable, at least in part, to structural ambiguity.
- Direct objects in English are constrained to follow governing verbs.
- Constructions like [[*pick* ᵥ] *up* ᵥ] thus permit two solutions:
  - Placing the object after the simple verb yields *pick the cat up*
  - Placing the object after the phrasal verb yields *pick up the cat*
- A similar ambiguity may underlie the pattern [[*pick* ᵥ]-*er up* ᵥ]-*er*.

Unlike constituent coordination (Sag et al. 1985)

- ▶ Conjuncts are traditionally required to be 'alike': $X \rightarrow X\ conj\ X$
- ▶ Yet heterocategorial conjunctions appear to be possible in English.
- ▶ In *Pat is a Republican and proud of it* an NP is conjoined with an AP.
- ▶ The GPSG analysis posits partially unspecified coordinate nodes.
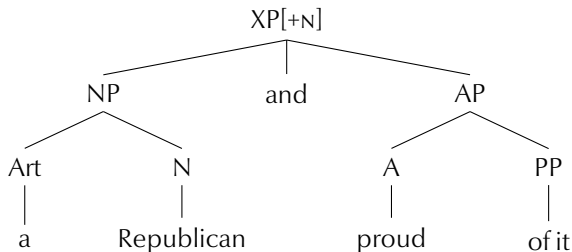
UCC as Left Node Raising?

- ▶ But if reentrant structures are allowed, these examples can be treated as 'like' coordinate VPs in which the conjuncts share a verb:
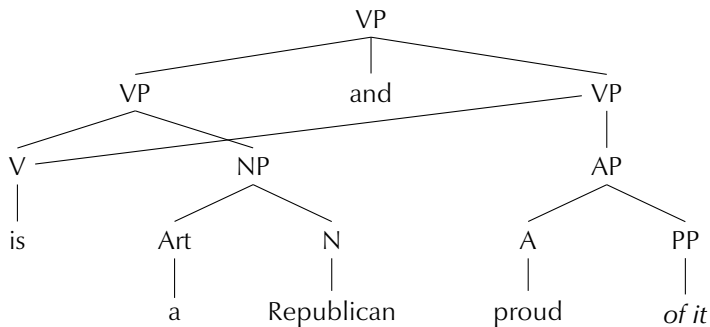
NCC as Left Node Raising?

► A parallel analysis applies to examples of nonconstituent
coordination, to which no feature generalization strategy applies:
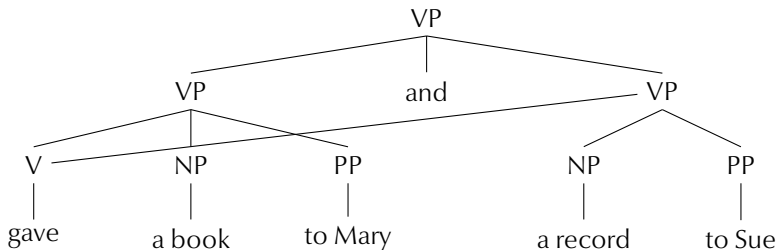
Reentrancy and order

- On this generalization, UCC and NCC are the directional duals of Right Node Raising constructions, as analyzed by McCawley (1982).
- The contrasts between RNR and these putative cases of LNR are attributable in large part to independent word order patterns:
  - Verbs are serialized in the leftmost conjunct in LNR constructions because they must precede their complements in all conjuncts.
  - Complements are serialized in the rightmost conjunct in RNR constructions because they must follow their verbs in all conjuncts.

The locus of reentrancy

- ▶ Thus the notion of structural reentrancy, introduced in HPSG feature structures, has a direct analogue in constituent structures.
- ▶ Reentrant structures simplify the feature analysis of coordinate nodes in UCC, the constituent structure of individual conjuncts in NCC, and extend the scope of a standard coordinate schema.
- ▶ It appears that reentrancies are structures whose time has come:
  - ▶ Reentrancies are assigned to noun phrase constructions in the Cambridge Grammar of English (Huddleston & Pullum 2002).
  - ▶ Reentrancies are even incorporated in recent Minimalist accounts of coordinate questions (Citko & Gračanin-Yuksek 2013).

What can we conclude?

- The research tradition that spans GPSG (Gazdar et al. 1985), HPSG (Pollard & Sag 1987, 1994), and models of SBCG (Sag 2010, 2012) has established a representational approach to a range of phenomena.
- This tradition continues to generate research questions, identify and confront empirical challenges and throw up analytic choices.
- The continued relevance of this tradition derives from two defining characteristics of Ivan's work: a dedication to comprehensive coverage, and a commitment to formulating proposals with sufficient precision to determine when they are or might be wrong.

▶ What principle guides the work that we are celebrating here?

"the method of rigorously stating a proposed theory and applying it strictly to linguistic material with no attempt to avoid unacceptable conclusions by ad hoc adjustments or loose formulation"

Bloomfield, L. (1933). *Language*. Chicago: University of Chicago Press.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Citko, B. & Gračanin-Yuksek, M. (2013). Towards a new typology of coordinated *wh*-questions. *Journal of Linguistics* **49**, 1–32.

Gazdar, G., Klein, E., Pullum, G. K. & Sag, I. A. (1985). *Generalized Phrase Structure Grammar*. Cambridge: Harvard University Press.

Gazdar, G., Pullum, G. K. & Sag, I. A. (1982). Auxiliaries and related phenomena in a restrictive theory of grammar. *Language* **58**, 591–638.

Gleason, H. A. (1961). *An Introduction to Descriptive Linguistics*. New York: Holt, Rinehart and Winston, 2nd edition.

Harris, Z. S. (1951). *Methods in Structural Linguistics*. Chicago: University of Chicago Press.

Harris, Z. S. (1964). The elementary transformations. *Transformations and discourse analysis papers, no. 54*, University of Pennsylvania.

Hockett, C. F. (1958). *A Course in Modern Linguistics*. New York: MacMillan.

Huddleston, R. & Pullum, G. K. (2002). *The Cambridge Grammar of English*. Cambridge: Cambridge University Press.

McCawley, J. D. (1982). Parentheticals and discontinuous constituent structure. *Linguistic Inquiry* **13**, 91–106.

Pollard, C. (1984). *Generalized Phrase Structure Grammars, Head Grammars and Natural Language*. Ph.D. thesis, Stanford.

Pollard, C. & Sag, I. A. (1987). *Information-Based Syntax and Semantics*. Stanford: CSLI.

Pollard, C. & Sag, I. A. (1994). *Head-driven Phrase Structure Grammar*. Stanford: University of Chicago Press.

Sag, I. A. (2010). English filler gap constructions. *Language* **86**, 486–545.

Sag, I. A. (2012). Sign-based Construction Grammar: An informal synopsis. In Sag, I. A. & Boas, H. (eds.), *Sign-Based Construction Grammar*, Stanford: CSLI Publications.

Sag, I. A., Gazdar, G., Wasow, T. & Weisler, S. (1985). Coordination and how to distinguish categories. *Natural Language and Linguistic Theory* **3**, 117–71.

Wells, R. S. (1947). Immediate constituents. *Language* **23**, 81–117.