Formalizing Construction Grammar in TAG

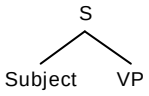
In this paper, we show how the balancing advantages of Tree Adjoining Grammar (TAG) and a Zwickyan Construction Grammar (CG) may be unified into a single approach, capturing both the details of complex clausal expressions and the relationships holding in the grammar between the smaller pieces of syntactic structure licensing these expressions. We envision a hybrid grammar in which a partial CG is formalized as a "metagrammar" for a TAG (in the spirit of Xia 2001, if not in every detail), defining the set of elementary trees on which the TAG operates and the all-important relationships between them. The TAG will then take this metagrammar's output (i.e. a set of elementary trees anchored to lexical items, simple-clause-sized or smaller) and license (potentially multi-clausal) sentence-level objects using its normal combinatorics of substitution and adjunction.

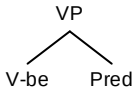
It is the considered opinion of a number of researchers (cf. Joseph 1997, Ginzburg & Sag 2000:5 and references therein, and Fillmore & Kay 1999 inter alia) that the distinction between "core" and "periphery" linguistic phenomena (as in, e.g. Chomsky 1981:8) is, at best, poorly motivated and, at worst, methodologically dangerous. The CG framework proposed in Zwicky (1994) is designed to achieve the goal of a syntactic framework which integrates, in a single theory, simpler, more general constructions (e.g. Subject-VP or Subject-Auxiliary Inversion) with constructions of increasing complexity (such as the Inverted WH-Cleft exemplified in 3 below) and constructions of increasing lexical specificity ("by and large"). However, for all its benefits, the CG in Zwicky (1994) is unformalized and its mechanism for combining constructions appears to be too powerful, allowing everything in principle. TAG (Joshi et al. 1975, Joshi 1985, inter alia), on the other hand, constitutes a highly constrained formalism for building (and thus encoding generalizations about) clausal and multi-clausal syntax, whose mathematical properties are well-known (Vijay-Shanker and Joshi 1985). A TAG does not, however, contain any mechanism for encoding relationships between the elementary trees it combines. TAGs do not capture generalizations at the level of the basic clause and below-clause levels (Frank and Kroch 1995:12 also make this point) in the way that Zwickyian CG does. In fact, we believe that the strengths and weaknesses of CG and TAG are entirely complementary.

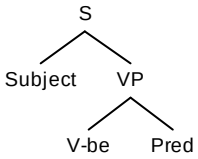
In order to illustrate our approach, take the set-piece constructions of Zwicky (1994): the WH-Cleft and Inverted WH-Cleft (illustrated in 2 and 3, respectively). While the TAG itself licenses the clausal expressions found in examples 1-3, it is the CG-based metagrammar that will state the fact that, for instance, 1, 2, and 3 are related in having a common basic constituent structure and ordering of elements: Subject+VP. Furthermore, the metagrammar captures the generalization that the set of objects licensed as the subject of the WH-Cleft (e.g. "what we saw") is precisely the set of elements licensed as the predicate of the V in the Inverted WH-Cleft. Figures 4-5 show a graphic representation of the basic constructions (Subj-VP) and (VP-Pred). Constructions are implemented as tree descriptions, i.e. underspecified trees. Such descriptions combine into bigger descriptions (such as the one for the copula in 6). Combination consists in straightforward addition of dominance and precedence relations. The construction 6 then combines with the one in 7 (a valence set for "be"), producing a description with exactly two minimal instantiations: the elementary trees 8 and 9, that are used in the derivation of sentences 3 and 2, respectively. Now, these trees are directly structurally related, because they make use of the same smaller tree descriptions.

Thus we see a twofold benefit: TAG can be viewed as a viable formalization of a CG, while also contributing a formal clausal and multi-clausal syntax for free.

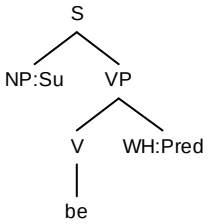
1. We will see a flying pig. (Zwicky 1994:612)
2. What we saw was a flying pig. (Zwicky 1994: example 9.6)
3. A flying pig was what we saw (Zwicky 1994: example 9.7)

4. 

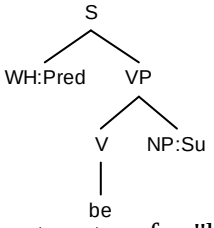
5. 

6. 

7. { NP, be, Wh-Phrase }

8. 

elementary tree for "be" in 3.

9. 

elementary tree for "be" in 2.

References:

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Ginzburg, Jonathan and Ivan A. Sag. 2000. *Interrogative Investigations: The Form, Meaning, and Use of English Interrogatives.* Stanford: CSLI Publications.

Joseph, Brian D. 1997. On the linguistics of marginality: The centrality of the periphery. In G. Anderson et al., eds., *Papers from the 33rd Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, pp. 197–213.

Joshi, Aravind K. 1985. How much context-sensitivity is required to provide reasonable structural descriptions: Tree Adjoining Grammars. In David Dowty, Lauri Kartunnen, and Arnold Zwicky (eds.), *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives*. Cambridge: Cambridge University Press.

Joshi, A. K., L. S. Levy and M. Takahashi. 1975. Tree Adjunct Grammars. *Journal of Computer and System Science,* 10:136–163.

Fillmore, Charles J. and Paul Kay. 1999. Grammatical constructions and linguistic generalizations: the *What's X doing Y?* construction. *Language* 75:1–33.

Frank, Robert and Anthony Kroch. 1995. Generalized transformations and the theory of grammar. *Studia Linguistica* 49:103–151.

Kroch, A. and A. K. Joshi. 1985. The Linguistic Relevance of Tree Adjoining Grammar. University of Pennsylvania Department of Computer and Information Sciences Technical Report MS-CIS-85-16.

Kroch, Anthony. 1989. Asymmetries in long distance extraction in a Tree Adjoining Grammar. In Mark Baltin and Anthony Kroch (eds.), *Alternative Conceptions of Phrase Structure,* 66–98. Chicago: University of Chicago Press.

Vijay-Shanker, K. and A. K. Joshi. 1985. Some computational properties of Tree Adjoining Grammars. In *Proceedings of ACL*.

Xia, Fei. 2001. *Automatic Grammar Generation From Two Different Perspectives*. Doctoral Dissertation, University of Pennsylvania.

Zwicky, Arnold M. 1994. Dealing out meaning: Fundamentals of syntactic constructions. *BLS* 20:611–25