Constructions at Work or at Rest?

Rens Bod*

"Constructions at Work" is an impressive piece of work which has already become a landmark in linguistics. In my commentary I will not go into the wealth of linguistic insights presented by Adele Goldberg, nor will I question the basic assumptions of Construction Grammar (CxG). Instead, I will deal with only one question: to what extent are we offered a model of language acquisition and/or language use in Goldberg's book?

At various places in the book we read that one of Goldberg's major goals is to address "How [do] learners acquire generalizations such that they can produce an open-ended number of novel utterances based on a finite amount of input" (e.g. p.11, p.227). But at the end of the day we learn that Goldberg proposes no model or theory that "can produce an open-ended number of novel utterances based on a finite amount of input". Sure enough, we get a lot of information about how constructions can be learned and generalized, and how a well-known machine learning algorithm, ADABoost, can be applied to combine multiple cues in generalizing constructions. But suprisingly enough, there is no exact model or theory in the book that describes how these constructions are used by learners to produce novel utterances.

The question that continuously came up while reading Goldberg's book, was: where is the model? I could find no description of an input-output procedure nor of a process-model that tells us how new utterances are produced on the basis of previous utterances. There is no precise definition of (i) the notion of a productive unit in CxG, (ii) the way productive units are acquired step by step from incoming input utterances, and (iii) the combination operations that combine constructions into (an open-ended number of) new utterances. At the end of the book, it even remained unclear to me what kind of an object a construction is according to Goldberg. Of course, I understand (and agree) that "any linguistic pattern is recognized as a construction […]" (p.5). Yet, the concept of "any linguistic pattern" is not well-defined. Linguistic patterns can be (partial) strings, (partial) phrase-structures, (partial) dependency structures, (partial) attribute-value matrices, (partial) typed-feature structures, etcetera. I doubt whether these are all taken as constructions. However, there is also a more precise definition of construction in the book: a "form-meaning pairing". But then, is a construction a pairing of strings, a function from a string to a predicate-argument structure, or yet something else?

And is the combination operation between constructions a concatenation operation, a substitution operation, a unification operation, some integration of these three, or something different? We learn on page 10 that "Constructions are combined freely to form actual expressions as long as they are not in conflict." As a first informal exposition this may do, but it is not a definition of how two or more constructions can be combined. I guess that at least a notion of substitution is involved if a construction with open slots or variables is combined with another construction. And most importantly

* University of St Andrews, Scotland, and University of Amsterdam, Netherlands, rens.bod@gmail.com. Many thanks to Willem Zuidema and an anonymous reviewer for comments on a previous version of this commentary.

perhaps, how are the productive units acquired from previous utterances in a step-by-step, incremental way?

My comments should not be seen as a pedantry: as long as we have no precise definition of the combination operation(s) between CxG's productive units, there is no way that CxG can be unequivocally tested as a linguistic model. I am not arguing for mathematical formalizations, only for precise definitions. Even when "definitive" definitions are out of reach at the moment, tentative definitions allow at least for testing CxG, after which these definitions can be adapted and improved.

One could claim that it should not be a theorist's goal to come up with precise definitions for combination operations let alone to construct a computational model. While I agree that the construction of a full-fledged computational model of language acquisition may be outside the realm of language theory, precise definitions of its productive units and how they are combined form a necessary condition for any linguistic theory.

Is this where Construction Grammar is moving to: a set of insightful but untestable ideas? Yet, it is not just Construction Grammar but almost any current linguistic theory that has given up on the construction of a precise, testable model of language use and language acquisition. Where are the good old days when every self-respecting linguist designed a precise model of language use and acquisition? But perhaps this is wishful thinking, and these days have never existed. Although I believe that Construction Grammar and Usage-Based Linguistics are by far the most promising approaches to language learning, they seem to underestimate the importance of exact definitions.

As far as I am aware, the only discipline that takes this enterprise seriously, is computational linguistics. However, during the last decade or so this discipline has become so impenetrable for non-computational linguists that it has lost almost all connections with the other fields of linguistics. The fact that computational linguists mostly publish in conference proceedings has not helped either, although all proceedings from 1965 onwards are freely available at the ACL-archives (http://acl.ldc.upenn.edu/). It is regrettable although perhaps understandable that Goldberg does not go into any of the work on statistical modeling in computational linguistics, which is so relevant for usage-based and constructionist models of language (given the importance of frequency in these models). Other linguists, such as Bybee (2006) and Bresnan and Hay (2006), have long recognized and emphasized the converging trends in theoretical and computational linguistics, suggesting a common development towards an integrated usage-based theory of language use and acquisition. In computational linguistics, there is nowadays a whole tradition that aims to combine construction/usage-based and computational approaches, such as Klein and Manning (2002), Steels (2004), Bod (2006), Zuidema (2006) and several others.

To give a concrete example let me briefly point out some connections between CxG and a computational model I have worked on myself: data-oriented parsing (DOP). I will only go into the more recent versions of DOP such as LFG-DOP (Bod and Kaplan 1998) and HPSG-DOP (Arnold and Linardaki 2007). In DOP, new utterances are produced and understood by statistically generalizing over a corpus of previous utterances. Informally, the productive units in DOP are, like in CxG, taken as "any linguistic pattern" be they morphological, syntactic, semantic or pragmatic in nature. But in concrete instantiations of DOP the productive units (or constructions) are usually taken as syntactic-semantic patterns. For example, in LFG-DOP the productive units are defined as any connected part of constituent structure and semantic structure (see Bod 2006). These parts are learned from exemplars of previous utterance-analyses and are formally equivalent to a functional mapping from subtrees (representing syntactic surface form) to attribute-value matrices (representing functional and semantic structure). The combination operation between the productive units consists of simple node substitution between subtrees combined with unification between corresponding attribute-value matrices. A statistical model is used to rank alternative utterances for a given meaning or alternative meanings for a given utterance. All these formal details of DOP may turn out to be inadequate in the long run, but they are precise enough to be turned into a testable computational model, which has indeed been carried out (e.g. Bod 2008; Zuidema 2006).

The acquisition of the productive units is dealt with by an unsupervised generalization of DOP known as U-DOP (Bod 2007). U-DOP initially allows any, possibly discontinuous fragment of previous utterances to form a productive unit, and lets the statistics decide which form-meaning pairs are actually learned and stored in the corpus (to which next DOP can be applied). This integrated U-DOP/DOP model has been tested against a wide variety of child language phenomena from the CHILDES database, ranging from learning separable particle verbs to learning auxiliary fronting (Bod 2008). It offers to linguists and cognitive scientists a precise model for testing the usage-based approach against child and adult data.

It thus seems that the DOP approach is not only consonant with CxG but that it actually stands as a testable realization of CxG: it specifies how constructions are acquired, how they are combined and how they can undergo changes in the light of new input. Goldberg's notion of "exemplar-based abstraction" (p.48) is furthermore congenial to U-DOP's acquisition procedure that abstracts over different utterances by statistical induction (Bod 2008). Yet U-DOP/DOP is not the only possible computational realization of CxG. There are other computational approaches to language learning, based on somewhat different notions of productive unit and different combination operations, e.g. van Zaanen (2000), Clark (2001), Klein and Manning (2002), Dennis (2005), Zuidema (2006), Seginer (2007) to name a few.

Despite the various, sometimes subtle differences between these computational learning approaches, they share one very important goal, which is, as far as I understand, the same goal as in CxG: to develop a usage-based theory of language learning that acquires generalizations which can produce an unlimited number of new utterances based on a finite number of previous utterances. I am convinced that this goal can be achieved, if only the two – linguistic and computational – strands take each other's work into account.

References

Arnold, Doug, and Evita Linardaki

2007 HPSG-DOP: Towards Exemplar-based HPSG. *Proceedings of the Workshop on Exemplar Based Models of Language Acquisition and Use*, Dublin, 42–51.

Bod, Rens

2006 Exemplar-Based Syntax: How to Get Productivity from Examples. *The Linguistic Review* 23, 291–320.

2007 Is the End of Supervised Parsing in Sight? *Proceedings ACL 2007*. Prague, 400–407.

2008 *Explorations in Language Learning by Integrating Analogy and Probability*. Submitted for publication.

Bod, Rens, and Ronald Kaplan
1998 A probabilistic corpus-driven model for lexical-functional analysis. *Proceedings COLING-AC 1998*, 145–151.

Bybee, Joan
2006 From Usage to Grammar: The Mind's Response to Repetition. *Language* 82(4), 711–733.

Clark, Alexander
2001 Unsupervised induction of stochastic context-free grammars using distributional clustering. *Proceedings CoNLL 2001*, 105–112.

Dennis, Simon
2005 An Exemplar-Based Approach to Unsupervised Parsing. *Proceedings CogSci 2005*, Stresa, Italy.

Hay, Jennifer, and Joan Bresnan
2006 Spoken syntax: The phonetics of giving a hand in new zealand english. *The Linguistic Review* 23, 321–349.

Klein, Dan, and Chris Manning
2002 A general constituent-context model for improved grammar induction. *Proceedings ACL 2002*, Philadelphia, 128–135.

Seginer, Yoav
2007 Fast Unsupervised Incremental Parsing. *Proceedings ACL 2007*, 384–391.

Steels, Luc
2004 Constructivist Development of Grounded Construction Grammar. *Proceedings ACL 2004*, 9–16.

van Zaanen, Menno
2000 ABL: Alignment-Based Learning. *Proceedings COLING 2000*, 961–967.

Zuidema, Willem
2006 What are the Productive Units of Natural Language Grammar? A DOP Approach to the Automatic Identification of Constructions. *Proceedings CoNLL 2006*, 29–36.