SHORT COMMUNICATIONS

**Grammatical is easier: A review of
Peter W. Culicover & Andrzej Nowak, *Dynamical grammar. Minimalism, acquisition, and change*.
Oxford: Oxford University Press 2003
[Foundations of Syntax 2], 346 pp.**

The title of the book by Culicover and Nowak, *Dynamical grammar*, will sound as an oxymoron to a strict, traditional linguist. Grammar of a language is often associated with something static, a set of rules that describe a language or that are represented in a system of knowledge of a language user and allow him/her to produce acceptable sentences in this language. The term "dynamical", on the other hand, refers to processes or phenomena that unfold over time. Dynamical description of a phenomenon is thus one that tries to capture its change. What then may "grammar" have to do with "dynamics"?

The goal of the book, as the authors see it, is "to build a computational simulation of language acquisition and language change whose internal architecture does not represent knowledge of language in symbolic terms, as a grammar *per se*, but whose behaviour can be described using the devices of standard linguistic theory." "To learn a language" does not mean here the acquisition of formal rules of sound-meaning correspondences. Rules are not written, or even represented, in any explicit way in the system. Rather, the system's general properties together with specific input lead to its change in such a way as to behave as if "according to a rule".

It is something else to obey a rule and to behave as-if there was a rule (Wittgenstein, 1953). An example of this difference might be an attempt to explain behavior of a driver who stops at the crossroads at which there is a STOP sign. After observing several instances of such behavior we may conclude: a driver

has a rule represented in his mind: "when I see the STOP sign, I stop". This may describe the driver's behavior quite well, and may even be enough to predict it in most instances. There will be, however, situations in which the driver will not stop at the STOP sign, or will stop at crossroads without a STOP sign, and until we know the complex factors that really motivate the driver's behavior (avoiding being hit by another car, fear of police, being in a hurry etc, etc) we won't be able to really explain it. Rules simplify the description but not without costs.

According to the view presented in Culicover and Nowak's book, the architecture of a language faculty is a dynamical system. The main question of the book is: "What is the minimal machinery and prior knowledge" that must be assumed in such a system in order for it to learn a language. This situates their approach in the category of minimalist theories. The authors call it Concrete Minimalism – because the system is presented with minimal concrete manifestations of a language (i.e., information about strings of linguistic elements and their corresponding Conceptual Structure representations).

In this view, language is a dynamical phenomenon that can be viewed on four time scales: evolutionary (the time scale of the evolution of human cognition), historical (evolution of language), acquisition (the time scale of ontogeny), and the real time of communicative episodes. The book focuses on the middle two, and tries to show that grammar may be treated as an emergent property of the system – both at the level (or time-scale) of human development and at the level of language evolution.

In the case of acquisition, the goal is to seek such construction of a dynamical system for which it takes the smallest amount of energy to behave in a way that is best captured by the rules of grammar (grammatical is easier). Ungrammatical "behavior" of a system is thus not probable because of an energy barrier. The question – motivated by the explanatory minimalism mentioned above – is what are the minimal assumptions about the information that such a system must have, and about the properties of the system.

After a theoretical part, three chapters follow in which simulations are described. The first simulation deals with the distributional approaches to modeling language acquisition, which became quite popular types of explanation (e.g. recurrent neural networks by Elman, 1993, or Jordan, 1986). The conclusions from the analysis of behavior of the model Acqui are that "the distributional properties of words can provide at best a categorization that reflects semantic co-occurrence restrictions". Such models might be amended by some additional heuristics but they themselves do not seem to discover the "actual structure of the language".

Therefore one must presuppose something more in a learner, and the authors stipulate that it is the structure of meaning, represented in Conceptual Structure (CS). This is actually not a small or easy assumption, mainly because there is no single theory of meaning that most linguists would agree upon. Also, assuming that CS is a "hierarchically structured compositional semantic representation" seems to

take most of the "syntactic burden" off the shoulders of the strictly symbolic input, but the burden does not disappear: one still has to explain where the hierarchical and compositional properties of CS come from and what is its nature. The authors, however, are well aware of the problem. The meaning structure they accepted in their models is, as they say, one of many possible ones. They can do this because their main objective is not to solve the problem of meaning representation but rather to show how the properties of a meaning structure will manifest themselves in the course of acquiring semantics.

Thus the next model, CAMiLLe, takes into account information about the distributional properties of linguistic elements as well as about meanings corresponding to the symbolic input, presented to it during learning. The meanings are specified arbitrarily – they are not checked against their correctness or effectiveness nor it is explained what the mechanism of their coupling to the environment or even to the elements of language is. As said above, this is the conscious decision of the authors. However, one might easily imagine that e.g. the adequacy of categories in a given environment and their effectiveness as "meanings" could be verified in the model's "action" in the environment and its "communication" with the members of its community (see for example Steels and Belpaeme, 2005). The mechanisms of word/meaning coupling, on the other hand, could be, for example, those proposed in Bloom (2002), or in Tomasello (1999, 2000). The latter work could also help explain better, on the psychological level, the notion of "attentiveness" of CAMiLLe (the authors just assume that "object is being attended to when it is mentioned"). One aspect of the relation between syntax and conceptual structure that is harder to envision in this model is that the shaping relation between the two is *bi-directional*: not only CS influences syntax, but also syntax influences CS (see, for example Bowerman and Levinson, 2001).

CAMiLLe, trained on the real data from the CHILDES database, shows acquisition of many syntactic properties of natural language on the level of lexicon (e.g., nouns and verbs categorization), structure, word order. In some cases where CAMiLLe fails, it is suggested that the requisite information may come from the conceptual system, which even further strengthens its role in syntax acquisition. Access to meaning, capacity to form categories, and access to notions such as "phrase" and "head of the phrase", turned out to be prerequisites for language acquisition in CAMiLLe.

Chapter 6 deals with modeling language change, based on the assumption that "the same architecture and mechanisms of reorganization should suffice to account for language acquisition and language change". Even though this assumption seems debatable, (e.g., because language acquisition can be seen as just an element of a broader language change dynamics, which also has to encompass the processes of communication serving as efficacy criterion for the selection of linguistic structures), I understand that authors are still motivated by their concrete minimalism approach to ask questions such as: what is the minimal assumption

about the architecture of a system that produces results interesting from a language acquisition point of view. The same can be said about the next assumption in this section, namely, that change is entirely a product of the variability of linguistic input to a learner. One could ask: and what about changes driven from the level of the CS (conceptual structure)? What about the changes such as development of grammar being a consequence of semantic complexity, e.g., Batali, 1998; Schöne-mann, 1999? However again, given just basic mechanisms of self-organization the Culicover and Nowak's models were able to show interesting behavior, congruent with some diachronic linguistic data.

Finally Chapter 7, is a more complete explication of the Concrete Minimalism program, which, let's reiterate, has the goal to reconcile a dynamical perspective with the current syntactic theory by showing that the design features of language (as specified by this theory) may be seen as manifestations of a dynamical system.

Dynamical grammar is one of those rare books that – perhaps because of the difference in it's authors' background – raises hopes for bridging the gap between theoretical linguists and dynamical systems theorists. Of course, it is not for the first time that dynamics has been acknowledged as an element of linguistic explanation: we have such attempts at least since the second wave of neural network models in the eighties, and since the first theories of self-organization in phonology (Lindb-lom et al, 1984). But the novelty of this book, which makes it so important, lies in the combination of the scope of phenomena it aims at explaining, and at the same time attention to the constructs and details of linguistic theory. Even with such a background of dynamic modeling in language it is a courageous book, in that it has taken probably the most "rule oriented", most "symbolic" theory of grammar there is, in its entire complexity, and has shown, that the proposed rules might be seen as a description of the behavior of a dynamical system.

This is why the reconciling role of the book may be really significant. It may cause linguists to be more friendly to dynamical systems, and it may cause dynami-cal systems theorists to be more careful with modeling language without overly simplifying its nature.

The success of such an approach makes one hungry for more. Even though the ability of the authors to abstract from certain theoretical problems (while re-maining aware of them) is a definite asset in the enterprise of explanation of such complex phenomena as the structure of natural language, the dynamical systems theory seems to stand up well in the face of complexity. Thus, even though one may understand the authors' decision to limit the levels of dynamics taken into account (to two out of four time scales), including other time scales, especially the time scale of real time communication events, might lead to an interesting picture.

What would actually be fascinating, is to see, in the explanation of linguistic phenomena, the *intertwining of the time scales* at which the dynamical events pertaining to language happen, i.e., analyzing the mutual influence of those events. Thus, historical changes in language might be seen as (at least partially) stemming

from the functional sharpening of language as a social coordination tool (i.e., as not entirely a product of the variability of linguistic input to a learner). The scope of those changes is influenced by the learnability of structures on the time-scale of human development (see e.g. Smith, Brighton & Kirby, 2003) who see learnability not only as a limiting factor, but actually as a source of compositionality). During the course of development language might be seen as one of the factors that guide conceptual development, stabilizing some categories and not others. The effectiveness of language, on the other hand, is verified on the shortest scale of human communication episodes, which provides criteria for the process of selection of language structures and elements (for theoretical discussion see Rączaszek-Leonardi, 2003; for aspects of modeling such 3-level dynamics see Steels and Belpaeme, 2005).

The theoretical framework and modeling tools introduced in this book seem appropriate to address even such complex questions. An emphasis on syntax and attentiveness to the details of linguistic theory, which characterizes the collaboration of the two scientists, would – I believe – allow for doing it on a level more sophisticated than that of Steels and Belpaeme (2005) where only the single word-category couplings were considered.

References

Batali, J. (1998). Computational simulation of the emergence of grammar. In J.R. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the evolution of language* (pp.). Cambridge UK: Cambridge University Press.

Bloom, P. (2002). *How children learn the meanings of words*. MIT Press.

Bowerman, M. & Levinson, S.C. (Eds.) (2001). *Language acquisition and conceptual development*. Cambridge, UK: Cambridge University Press.

Jordan, M. I. (1986). *Serial order: A parallel distributed processing approach*. Technical Report ICS Report 8604, Institute for Cognitive Science, University of California.

Lindblom, B., MacNeilage, P., & Studdert-Kennedy, M. (1984). Self-organizing processes and the explanation of phonological universals. In B. Butterworth, B. Comrie, & Ö. Dahl (Eds.), *Explanations for language universals* (pp. 181-203). New York: Mouton.

Rączaszek-Leonardi, J. (2003). The interrelation of time scales in a description of language. *Views & Voices*, 1, 2, 93-108.

Schönemann, P.T. (1999). Syntax as emergent characteristic of the evolution of semantic complexity. *Minds and Machines*, 9, 309-346.

Smith, K., Brighton, H. Kirby, S. (2003). Complex systems in language evolution: the cultural emergence of compositional structure. *Advances in Complex Systems*, 6, 4, 537-558.

Steels, L. & Belpaeme, T. (2005). Coordinating Perceptually Grounded Categories

through Language. A Case Study for Colour. Accepted as target article for *Behavioral and Brain Sciences*.

Tomasello, M. (1999). *Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2000). The social-pragmatic theory of word learning. *Pragmatics*, 10, 4, 401-413.

Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: B. Blackwell.

Joanna Rączaszek-Leonardi
University of Warsaw