Volume 8
Issue 1
2010

Commentary on Cristofaro - Cognitive Mechanisms Need to Be Operationalized

Remi van Trijp

Sony Computer Science Laboratory, Paris

Cognitive Mechanisms Need to Be Operationalized

Comment on 'Semantic Maps and Mental Representation'
by Sonia Cristofaro (2010)

Remi van Trijp
Sony Computer Science Laboratory, Paris

Cristofaro (2010) voices a growing trend within cognitive-functional linguistics to explain language universals in terms of productive processes in language use rather than in terms of universal conceptual representations. Such mechanisms include metonymy and generalization. Even though I fully agree with this general approach, the proposed mechanisms remain too fuzzy and incomplete to shed new light on the matter of mental representations. This is not so much a criticism directed at Cristofaro as it is a plea for coupling empirical observations to computational models: formal modeling is the only way to *demonstrate* the validity and completeness of any linguistic theory, and hence forms a crucial part of empirical science.

Let me illustrate the limits of verbal theorizing through "generalization", a mechanism proposed by various cognitive linguists and explained by Cristofaro (2010) in section 2.2. Cristofaro writes that generalization "is the loss of some of the meaning features associated with a grammatical form, with the consequent expansion of the range of appropriate contexts of use for that form." She argues that, for example, *be going to* could generalize into a future marker because a motion is always directed towards a point in the future, so the expression already contained futurity in its meaning. However, from what is known through computational linguistics, machine learning, and multi-agent experiments in artificial intelligence, this definition of "generalization" is problematic for various reasons.

The first problem is that ambiguous contexts do not require such a meaning shift if the target meaning "is already there". Concrete computational models, such as Memory-Based Language Processing (Daelemans and van den Bosch 2005), have shown that generalization is possible *without* abstraction: novel instances can be correctly classified and processed through similarity-based reasoning over stored exemplars. In fact, exemplar-based models turn out to be more robust because making abstractions can be harmful to a language user's performance. So the proposal that ambiguous contexts invite language users to perform meaning shifts does not explain *why* speakers would prefer this "risky" strategy over robust classifications which remain faithful to previous *successful* uses of a word or construction.

The mechanism of "generalization as a loss of meaning" also does not explain how such innovations could ever be propagated in a linguistic community and hence become the new convention in a language. In fact, the issue of propagation poses severe constraints on plausible innovation processes, as demonstrated by multi-agent experiments on the evolution of language (e.g. Steels 1996). These experiments show that language users can never innovate with the certainty that other language users will somehow make the same innovation. Rather than shifting the meaning of a specific construction (or word), a language user must maintain several exemplars in his memory at the same time with various degrees of entrenchment in order to stay in sync with the rest of the population, and to achieve a robust linguistic performance. This would be in line with the observation of "synchronic layering" (Hopper and Traugott 1993: section 5.5), i.e., that speakers keep on using both the "old" and the "new" use of a construction for a long time.

Finally, loss of meaning does not lead to "more grammaticalized" behavior or (cross-linguistically recurrent) patterns of multifunctionality. This is clearly demonstrated in experiments on flexible word meaning by Wellens (2008) and Wellens, Loetzsch and Steels (2008), in which artificial agents adapt the meaning features of words if this is required for reaching communicative success in certain contexts. In the experiments, some words indeed lose specific lexical features, but they still remain lexical items. Other words, however, become richer in meaning features if this is required by the communicative contexts in which they occur. The experiments thus show that the ability to shift the meaning of a word (a) does not automatically imply that it takes on additional grammatical properties, (b) cannot explain the directionality of grammaticalization, and (c) does not automatically lead to an expansion of the contexts in which a word can occur (but rather to a shift of contexts). What it ultimately shows is that this particular mechanism of generalization implicitly assumes a so-called "Gesamtbedeutung" for categories, whereas grammar requires a "polysemy" or "usage type" approach (Croft 1991:1). Neither Cristofaro (2010) nor any of her cited sources would ever describe grammar in terms of "Gesamtbedeutung definitions", but this is an overlooked consequence of the generalization mechanism they propose.

As already mentioned at the beginning of this commentary, I strongly agree with the general hypothesis that patterns of multifunctionality are the result of dynamic processes in communicative interactions. However, as the above discussion indicates, empirical observations only show the *result* (e.g. generalization) of certain processes, but not the actual processes themselves. Linguists therefore need to combine their empirical observations with computational or mathematical models and robotic experiments which can demonstrate these processes in action. A particularly interesting approach has been pioneered by Steels (1996), who argues that language should be seen as a complex adaptive system (see Steels 2000 for a summary of this view) in which a community of language users have to collectively solve the problem of developing a shared communication system. Steels proposes concrete operationalizations of abstract mechanisms (usually inspired by biological systems) such as self-organization, selectionism, co-evolution through structural coupling, reinforcement learning, and level formation. Experiments on the evolution of grammar have already demonstrated that for example generalization can be a *side-effect* of such dynamic processes in locally distributed communicative interactions rather than the mechanism responsible for multifunctionality (Steels 2004; van Trijp 2010). So in order to devise a truly explanatory theory of linguistic universals, we need to unravel the abstract processes that only indirectly manifest themselves in the observed facts.

References

Cristofaro, Sonia. 2010. Semantic maps and mental representation. Linguistic Discovery, this issue.

Croft, William. 1991. Syntactic categories and grammatical relations: The cognitive organization of information. Chicago: Chicago University Press.

Daelemans, Walter and Antal van den Bosch. 2005. Memory-based language processing. Cambridge: Cambridge University Press.

Hopper, Paul J. and Elizabeth Closs Traugott. 1993. Grammaticalization. Cambridge: Cambridge University Press.

Steels, Luc. 1996. A self-organizing spatial vocabulary. Artificial Life Journal 2/3.319-332.

-----. 2000. Language as a complex adaptive system. Proceedings of PPSN VI, ed. by M. Schoenauer, 17-26. Berlin: Springer.

-----. 2004. The constructivist development of grounded construction grammars. Proceedings of the Annual Meeting of the Association for Computational Linguistics. ACL, ed. by W. Daelemans.

van Trijp, Remi. 2010. Grammaticalization and semantic maps: Evidence from artificial language evolution. Linguistic Discovery, this issue.

Wellens, Pieter. 2008. Coping with combinatorial uncertainty in word learning: A flexible usage-based model. The evolution of language. Proceedings of the 7th International Conference on the Evolution of Language, ed. by A.D.M. Smith, K. Smith and R.F. i Cancho, 370-377. Singapore: World Scientific.

Wellens, Pieter, Martin Loetzsch and Luc Steels. 2008. Flexible word meaning in embodied agents. Connection Science 20/2.173-191.

Author's contact information:
Remi van Trijp
Sony Computer Science Laboratory Paris
6 Rue Amyot
75005 Paris (France)
remi@csl.sony.fr