LINGUISTIC DISCOVERY
DARTMOUTH COLLEGE

Volume 8
Issue 1
2010

Commentary on Cysouw - What Should Be on a Map?

Heiko Narrog

Tohoku University

doi: 10.1349/PS1.1537-0852.A.365

url: http://journals.dartmouth.edu/cgi-bin/WebObjects/Journals.woa/1/xmlpage/1/article/365

Linguistic Discovery
Published by the Dartmouth College Library

ISSN 1537-0852
linguistic-discovery.dartmouth.edu

What Should Be on a Map?

Comment on 'Semantic Maps as Metrics on Meaning' by Michael Cysouw (2010)

Heiko Narrog
Tohoku University

Cysouw's paper is one of the most radical and most theoretically intriguing contributions in this issue, both with respect to his definition of meaning and its operationalization for the construction of semantic maps, which are conceived as a 'cross-linguistic metric on meaning' (section 1). On the basis of this definition, Cysouw offers a refinement of the terminology used in semantic maps (section 2) and a refinement of the terminology concerning linguistic expressions which serve as building blocks for the maps (sections 5, 6). Section 7 exemplifies the approach, making use of pairs of constructions (i.e. 'constructional behavior'), that is, a type of data rarely used for constructing semantic maps.

Although Cysouw's approach is radical, it is logically stringent, and there is not much to disagree with once one has accepted its premises. On the other hand, this paper is not easy on its readers, as it is not very explicit. Metaphorically speaking, it is painted in bold strokes, and one would wish that the author, like the old Flemish masters, had some disciple in his workshop who would fill in the details, that is, someone who would elaborate on both contents and technical details, as for example the relevant page numbers of the works cited. My comment is therefore less of a critique rather than a peer review which rates the paper under consideration as A, but nonetheless requests some clarifications.

The largest potential clarification would be one which is in fact impossible to provide within the scope of a paper, namely, the nature of meaning in linguistic expressions. Its definition as "the collection of all contexts in which the expression can be used" is just one possible solution. In the context of Cysouw's research, I understand it as a working definition which is necessary as a basis for building up what follows, but its validity can certainly be questioned. The justification given in Cysouw (2010) is rather short, and it seems to me that this paper has to be read together with Wälchli's (this issue) in order to gain a fuller understanding of its background. In their view, semantic maps are defined in terms of the representation of measurable distances. This definition practically excludes the so-called 'classical' or 'traditional' maps, at least in the fashion in which we know them, since distances between meanings or functions on these maps are not conceived as the result of a calculation of similarity. Some of the papers in this issue which explicitly discuss the pros and cons of classical vs. statistical maps (Malchukov 2010; Mauri 2010; Narrog 2010) highlight the advantages of classical maps, a discussion which I will not repeat here, while other authors (Boye 2010; Cristofaro 2010; de Haan 2010; Hengeveld and van Lier 2010) straightforwardly take the makeup of classical maps as a given premise on which they base their respective inquiries. This indicates that classical maps still present a viable approach to the study of semantic structures. It thus appears to me that the new terminology offered in Cysouw's (2010) table 1, rather than replacing Haspelmath's (2003) terminology, serves as a proposal for a specific subset of semantic maps, namely those plotted mathematically on the basis of statistical data. As such, it is more explicit than what has heretofore been offered.

Another important presupposition which is not explicitly discussed but taken for granted by Cysouw (2010) and encountered again in more detail in Wälchli (2010) is that this statistical approach, like other approaches to semantic maps as well, requires some notion of iconicity between form and meaning.

In my view, sections 5 and 6 would have profited from additional examples and perhaps a

table summarizing the terminology. However, I wish to move on to the case study (section 7). As already mentioned (and also stated by the author himself), working with pairs of linguistic expressions instead of single terms is quite unique. This method is therefore also able to deal with data for which it is difficult to imagine how they could be represented as a 'classical' map. Classical maps, though they too require some notion of similarity, essentially highlight relationships of polysemy/polyfunctionality and thus meaning extension. No such polysemy can reasonably be assumed for lexical expressions like *wake up* and *teach/learn*, which are placed adjacently in figure 2. If the author had focused on comparing statistical to classical maps, he might have presented this as a case where statistical maps offer the opportunity to represent a larger range (i.e. more types) of data than classical maps. The main point of this case study, however, is that there is more than one way to construct a map, depending on how one analyzes the data (or more explicitly, depending on the aspect of the data which is quantified, and on the way in which it is quantified), and this point is illustrated quite convincingly. Subject to the properties of a concrete data set in question, there is usually a variety of possibilities among which the researcher can choose, and semantic maps, even if calculated automatically, do not automatically emanate from the data. Also, there are solutions which make more sense (sections 7.2, 7.3) and such that make less sense (section 7.4). This is again reminiscent of Wälchli's (2010) stance that semantic maps are "dynamic". Both papers demonstrate, from different perspectives, the extent to which semantic maps are "relative", as it were, and dependent on decisions made by the researcher rather than being absolute representations of some mental reality. Still, from the point of view of a researcher seeking some kind of absolute map as a tool of semantic analysis, it could be argued that maps (and metrics) which are apparently less informative should be dismissed, and that there will always exist one map, based on the most comprehensive and most representative data set to which the most appropriate criteria of measurement have been applied, which is optimal, that is, which is closest to representing some universal cognitive reality, for instance, and which is therefore the ultimate goal to achieve.

The last comment that I wish to offer here concerns figures 2 and 3 and the text leading to their presentation. Again, I must admit that I had some difficulties following the discussion and interpreting the maps here. Figure 2 as such is not easy to grasp if one is not familiar with the methodology. What strikes me as a layman with respect to statistical maps is that the map appears to be built on two different sets of data: firstly, the first dimension of an MDS resulting from a previous separate calculation, and secondly, the proportion of anticausative strategies as already provided by Haspelmath (1993), but calculated in a different, mathematically more appropriate manner. By contrast, the first thing that I would have expected was a map constructed from a single set of data, presenting the two (or more) dimensions calculated on the basis of these data. Therefore, I somehow missed an explanation concerning the lack of such a map. I was also struck by the fact that 'die/kill' ended up in the immediate vicinity of 'boil' and 'freeze'. If semantic maps are to feed into semantic analysis, or should reflect semantic analysis, as Malchukov (2010) suggests in this issue, this finding is hard to swallow. The pairs 'freeze' (tr.) – 'freeze' (itr.) and 'die' – 'kill' do not seem to be similar with respect to their difference in transitivity, nor are they similar with respect to the way in which the difference between the two concepts is cross-linguistically encoded−the only thing they do have in common is the relative lack of anticausative strategies. In Haspelmath's (1993:104) table for anticausative vs. causative type of expression, the pair 'die' – 'kill' is effectively excluded from the anticausative/causative calculation, for lack of a language which construes this pair anticausatively. On a MDS map, the peculiarity of this pair can probably only be captured through an additional dimension (e.g. a dimension representing degree of suppletion), or it cannot be properly captured statistically at all for lack of pairs that behave similarly, and thus for lack of a valid set of data serving as a basis to calculate this additional dimension. More generally speaking, this raises the question whether in

the end the traditional table as presented by Haspelmath (despite mathematical deficiencies) is not in some important respect more informative, and thus (at least in that same respect) ironically superior to the semantic map in Cysouw's (2010) figure 2 as a visualization of differences between causative/inchoative verb pairs. This leads back to the question what should be represented on a map in order to make it an adequate tool for capturing meaning, a question that can be tricky in constructing both classical and automatically calculated maps alike (see also Malchukov 2010).

References

Boye, Kasper. 2010. Semantic maps and the identification of cross-linguistic generic categories: Evidentiality and its relation to Epistemic Modality. Linguistic Discovery, this issue.

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

Cysouw, Michael. 2010. Semantic maps as metrics on meaning. Linguistic Discovery, this issue.

de Haan, Ferdinand. 2010. Building a semantic map: Top-down versus bottom-up approaches. Linguistic Discovery, this issue.

Haspelmath, Martin. 1993. More on the typology of inchoative/causative verb alternations. Causatives and transitivity, ed. by Bernard Comrie and Maria Polinsky, 87-120. Amsterdam: Benjamins.

Hengeveld, Kees and Eva van Lier. 2010. An implicational map of parts of speech. Linguistic Discovery, this issue.

Malchukov, Andrej L. 2010. Analyzing semantic maps: A multifactorial approach. Linguistic Discovery, this issue.

Mauri, Caterina. 2010. Semantic maps or coding maps? Towards a unified account of the coding degree, coding complexity and coding distance of coordination relations. Linguistic Discovery, this issue.

Narrog, Heiko. 2010. A diachronic dimension in maps of case functions. Linguistic Discovery, this issue.

Wälchli, Bernhard. 2010. Similarity semantics and building probabilistic semantic maps from parallel texts. Linguistic Discovery, this issue.

Author's contact information:
Heiko Narrog
Graduate School of Information Sciences
Tohoku University
Kawauchi 41
Aoba-ku
Sendai-shi, 980-8576
Japan
narrog@gmail.com