Commentary on Malchukov - Three Questions about Analyzing Semantic Maps

Bernhard Wälchli

University of Bern

Three Questions about Analyzing Semantic Maps

Comment on 'Analyzing Semantic Maps: A Multifactorial Approach' by Andrej L. Malchukov (2010)

Bernhard Wälchli
University of Bern

1. Introduction

Malchukov (2010) argues that probabilistic semantic maps ("similarity maps") can be turned into implicational maps ("semantic maps") if conflicting factors are "featured out". As far as conflicting factors are concerned, this contribution to the semantic map discussion is highly important because it is a further step towards a better understanding of the limitations of the isomorphism principle (recurrent identity in form is similarity in meaning), which is the basis for both implicational and probabilistic semantic maps. However, whether Malchukov's claim is true is an empirical question, of course. Personally, I am convinced that he is mistaken. Implicational maps require a very strong preselection of data, first of all on the level of granularity. The suggestion that linguists have a choice between implicational maps and probabilistic maps is an illusion—that choice is made by the datasets. Most datasets do not lend themselves to implicational maps, and the question is simply whether the semantic map approach should be restricted to such datasets where an implicational analysis is possible or whether we should be allowed to study messier datasets with more robust methods of analysis. Here, I would like to raise three questions about Malchukov's approach.

My first question concerns the *treatment of conflicting evidence*. Malchukov takes it for granted that all cases of "formal similarity" which do not reflect "semantic similarity" must be excluded. But wouldn't it be better to apply more robust methods for building semantic maps, where a little noise does not do any harm?

The second question concerns the *definition of functions*. According to Malchukov, these are preselected. However, preselection remains largely implicit. So, how are functions preselected?

The third question concerns Malchukov's *underlying concept of similarity* centering around the idea of adjacency, but not made fully explicit in the paper. What is the underlying theory of similarity? I will argue that the notion of adjacent functions is a matter of convenience without any theoretical foundation.

2. Wouldn't It Be Better to Control for Noise Rather Than Exclude It?

Semantic maps rest on the assumption that *recurrent* formal identity reflects semantic similarity. Important is the word *recurrent*, which means that some cases of formal identity may be accidental, but all cases taken together reflect systematic relationships.

I agree with Malchukov that there may be certain systematic exceptions to this assumption— zero marking, for instance. However, Malchukov argues that it is possible to control for interfering factors by excluding them in building semantic maps. With interfering factors included, we get raw "similarity maps", with them excluded, we get refined "semantic maps", and only the latter are amenable to semantic analysis, he argues. What I would like to know is

how this is done in practice in a concrete investigation resting on a large typological database. What Malchukov does in practice is to consider conflicting factors only where they do harm to the expected semantic map. He does not eliminate the factor; he eliminates only its undesired effect. This is illicit. If "heavier" markers can disturb the adjacency of lighter markers, then heavy and light markers should not be combined at all in building semantic maps. Moreover, all markers having both lexical and grammatical functions must be excluded (this, obviously, requiring that lexical and grammatical functions must always be sharply distinguished). Furthermore, all zero-marking categories and all unmarked members of oppositions must be excluded. The unanswered question is: how much will be left if we start excluding everything that might do some harm? It is important to base semantic maps on databases. Only if this is done can we quantify the amount of data that needs to be excluded in data-exclusion approaches.

I am inclined to believe that exclusion is the wrong approach. If the signal underlying semantic maps is strong enough and an appropriate method is used to extract it, a little "noise" will not do any harm. However, if the conflicting factors are the major trend in the data and if the same-form-ergo-similar-meaning principle is only scarcely relevant, this is bad news for the semantic map method altogether. If this should be the case, the method must simply be abandoned. Malchukov's approach is important because we are badly in need of a better understanding of possible violations of the same-form-ergo-similar-meaning principle, which are rarely discussed in the literature on semantic maps. However, identifying conflicting factors does not automatically entail that all cases affected by conflicting factors must be excluded. In multifactorial analyses in statistics, the aim is to control for factors, but this does not require excluding all cases where multiple factors might be involved.

Malchukov seems to assume that once all noise is excluded, the remaining data will support an implicational map without exceptions. This is an empirical question. Let us try with a large database and a large set of functions! The emerging picture in probabilistic maps is not "messy" because the method is messy, but rather because this method can be applied to datasets which come closer to reflecting the real amount of diversity in discourse. I would love to use pure implicational semantic maps if my databases would allow for them.

2. How Are Preselected Functions Defined?

One of the biggest problems of functional linguistics is that functional domains are not usually clearly defined. According to Givón (1981:167), it is sufficient to "define a functional domain in a broad, lax fashion." This may be enough for a typological study focusing on one domain, but not for a study investigating the relationship of many of them. Malchukov seems to feel that assigning the name of a case function, such as "Possessor", is not enough. He provides examples to illustrate what is meant. I would like to know what exactly the status of these examples is in defining semantic maps. I would claim that they do more than exemplifying. They narrow down (sharpen) the domain at issue. Much depends on which examples are chosen. If we choose for Material *The house is made of wood* and for Possessor *I saw the house of the Smiths* instead of those given, we should add a link between MAT and POS in Figure 8. If we choose *This house belongs to John* for Possessor, we have to add a link between POS and G(oal). In a methodologically well-founded approach, it should be made clear on what grounds functions are preselected, and especially what the status of examples is. If functions can be defined sharply only by exemplification, then building semantic maps requires exemplary semantics. If examples are irrelevant, it should be demonstrated how functions can be defined without examples.

3. What Is the Underlying Theory of Similarity?

Malchukov discusses cases where "formal similarity" is not a reflex of "semantic similarity". In order to understand what this really means, we must know what he understands by similarity. This is not made explicit in the paper, which is why I try to infer his point of view from the few indications given in the text. The key passage is "semantic similarities support polysemies of adjacent (connected) categories on the map, while non-adjacent categories need not share any common semantic features" (Malchukov: 184). This suggests that, on the one hand, *adjacency* is an essential condition and that, on the other hand, similarity is understood as *partial identity*. I will argue here that basing similarity on adjacency is possible only if resolution is unalterable, and that understanding similarity as partial identity is not the only option.

   Adjacency depends on resolution. On a rainbow color scale, red and yellow may be adjacent if only four colors distinguished, but they will not be adjacent if a hundred different shades are distinguished. If adjacency is indispensable for defining similarity, this would mean that the degree of resolution cannot be altered. However, I would argue that the resolution of functions (the analytic primitives of the map) are not given, but chosen by the linguist building the map. If we take semantic roles, Van Valin (2001:30-31) argues convincingly for a continuum from verb-specific semantic roles to grammatical relations. As the level of generalization increases, the contrast between verb-specific semantic roles (such as "Thinker", "Believer") is neutralized via a more general set of verb-specific semantic roles (such as "Cognizer") to thematic roles (such as "Experiencer"), semantic macroroles (such as "Actor" and "Undergoer"), and finally grammatical relations (such as "Subject").[1] Linguists choose the degree of resolution in sampling analytic primitives (functions) for semantic maps. As a consequence, defining similarity as adjacency in semantic space is not possible. The term "adjacent" only makes sense if understood as "significantly closer than any other preselected function". As soon as the set of preselected functions changes, adjacency relationships change, and the more entities there are, the more difficult it is to define adjacency. Adjacency is therefore an artifact of the choice of functions and has no theoretical foundation.

   In the semantic map approach, it is common to speak of formal identity and semantic similarity (e.g., Haspelmath 2001:215). Instead, Malchukov speaks of formal similarity. This suggests that he views similarity as partial identity, since the formal similarity we find in languages is often a partial identity of forms. In discussing semantic maps, we should be aware that there are two philosophical possibilities to view similarity according to whether it is a notion more basic or less basic than identity. Of course, we can follow such philosophers as Anton Marty (1908:407), according to whom there are two kinds of similarities, both derived from identity: partial identity of complexes and close species of the same genus—but there is also the opposite point of view of such philosophers as Fritz Mauthner, for whom similarity is the more basic concept. The point of view one adopts has important consequences. Prioritizing identity entails essentialism: it forces us to decompose the complexes into underlying units and/or to delimit the genera. If identity is more basic than similarity in meaning, the semantic space cannot be an amorphous mass as claimed by the structuralists, but will, at some level, consist of underlying discrete semantic features or primes which allow us to identify partial identity in meanings. The possibility that all meanings can be neatly decomposed into a basic set of

---

[1]Of course, further differentiation is possible, and there are other possibilities to slice the cake.

semantic primes cannot be denied. However, if somebody holds this view, I would expect an explication as to the nature of the presumed semantic features. To summarize, more effort should be made in the semantic map approach to clarify its theoretical foundations.

References

Givón, Talmy. 1981 Typology and functional domains. Studies in Language 5.163-193.

Haspelmath, Martin. 2003. The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison: The new psychology of language, ed. by Michael Tomasello, vol. 2, 211-242. Mahwah, NJ: Erlbaum.

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

Marty, Anton. 1908. Untersuchungen zur Grundlegung der allgemeinen Grammatik und Sprachphilosophie. Erster Band. Halle: Niemeyer.

Van Valin, Robert D., Jr. 2001. An introduction to syntax. Cambridge: Cambridge University Press.

Author's contact information:
Bernhard Wälchli
Institut für Sprachwissenschaft
Universität Bern
Länggassstr. 49
3000 Bern 9, Switzerland
waelchli@isw.unibe.ch