

Quantifying semantic regularity across languages

Asifa Majid & Stephen C. Levinson

Semantic maps in typological work are often produced on the basis of underlying conceptual spaces constructed by intuition and inspection. Here we argue that if the underlying conceptual spaces are thought of as multi-dimensional spaces, structured as similarity spaces, then this allows the application of sophisticated quantitative methods. For the typologist interested in quantifying how similar semantic categorization is across languages, these methods offer exciting new possibilities. Focussing on specific domains, they can be used to study different - although interrelated - questions, such as:

(1) What are the semantic distinctions being made within a domain? And where languages make different distinctions, do they respect the coherence of the same underlying conceptual space?

(2) How similar (or different) are languages in their overall pattern of categorization? Can we measure degrees of convergence?

(3) Are children faster at learning semantic distinctions that are typologically frequent in comparison to rare distinctions?

Drawing on findings from several collaborative cross-linguistic projects based at the Max Planck Institute for Psycholinguistics, we examine the semantic categorization of events as reflected in verbs and constructions, and highlight techniques for analyzing large cross-linguistic data sets that can capture both shared category structure and language variability.

In the case of event domains, the projects begin with an etic grid of event types - a set of videoclips - which vary along a number of parameters pertinent to the domain of study. Speaker descriptions are then elicited from a range of geographically, genetically and typologically diverse languages. The descriptions are analyzed using multivariate statistics, such as factor analysis, correspondence analysis and multidimensional scaling. These techniques not only visually represent cross-linguistic regularity, but also quantify precisely how much structure is shared, and what constitutes an unusual pattern of categorization.

The picture emerging is one of robust regularity. For instance, for "cutting and breaking" events all languages recognize a dimension having to do with how predictable the location of separation in an entity will be. Categorization of reciprocal events shows much more variation, with some quite different solutions to the problem of how to encode such events, but nevertheless recurrent categorization patterns emerge. Finally, we show how the multivariate conceptual spaces extracted from cross-linguistic studies can be used to investigate language acquisition too.

Semantic map geometry: two approaches

Joost Zwarts (Radboud University Nijmegen & Utrecht University)

Common to all semantic map approaches is the idea of a 'geometric' layout of meanings, which represents graphically how meanings (or functions) of words (or grams) are related to each other. Where does this geometry come from? In most semantic map applications, the geometry emerges *a posteriori* from the linguistic data, in an inductive way, either by constructing the smallest graph of meanings in which every word covers a connected subgraph, or by applying statistical scaling techniques. However, it is also possible to work in the opposite direction, from an *a priori* geometry or grid of meanings, deducing relations that can be tested against linguistic data. The colour space offers the classical example of such a language-independent geometry of meanings (Gärdenfors 2000). The prepositional network of Lakoff (1987) and the reciprocal lattice of Dalrymple et al. (1998) can also be interpreted as conceptual spaces of this type.

The point of this paper is that we need *both* approaches, complementing each other. Often, a data-driven approach is the only way to get some idea about how a set of meanings hangs together. It is both a powerful heuristic and an important check on misguided *a priori* assumptions about a particular meaning space. However, the approach also has its limitations.

1 A semantic map should not be the theoretical endpoint. We want to know why the meanings are distributed in a particular way, but it actually turns out to be difficult to make the step from a data-driven semantic map to a semantic model of the underlying conceptual space. This is even harder when statistical mapping methods are applied. By using an exclusively inductive approach, the semantic map approach runs the risk of broadening the gap with semantic theories, both from the formal and cognitive paradigm. We therefore need to work from the other end too: define a geometry on the basis of particular semantic assumptions and study the cross-linguistic mapping of such a geometry.

2 One of the exciting things about semantic maps is that they could embody a non-classical, but constrained theory of categorization, thanks to the connectivity (convexity, contiguity) property. However, some small-scale maps show a distribution of data that can easily be captured in terms of necessary and sufficient features (as I will show for the modality map of Van der Auwera and Plungian 1998 and the A – S – P map of Croft 2001). If we want to show that regions on a semantic map are really more than classical feature bundles, a model of the underlying semantic geometry is inevitable.

3 In the data-driven mapping approach the important connectivity hypothesis is part of the methodology itself and as a result its validity can only be studied indirectly. There is no room for principled exceptions to connectivity, unless we already have some idea about what meanings are non-adjacent on independent semantic grounds (as I will illustrate with the modality map). A purely data-driven approach can not recognize the individual exceptions and working in the opposite direction is more fruitful here.

4 The *a priori* approach allows us to separate two roles of non-discreteness in mapping, which can be obscured in scaling methods. There can be non-discreteness in the conceptual space itself (the famous cups and saucers of Labov), but this should be distinguished from the non-discreteness that results from the way linguistic data distribute over a discrete geometry of meanings. I will show that a 'three-dimensional' map, in which words are not regions but *hills*, helps us to give a proper place to this distinction.

At the moment there are no good examples (apart from colour terminology) where large amounts of linguistic data are tested against an extensive '*a priori*' conceptual space, but I will present a range of examples of smaller scale that suggest the direction in which this work might go, involving prepositions, clothing items and birds' names.

A data-driven, inductive approach to semantic maps has serious limitations, but, at the same time, a purely theory-driven, *a priori* approach to semantic maps does not work either. It is only when we are willing to go back and forth between semantic modeling and linguistic data that we can hope to gain insight in the way languages divide up spaces of meanings into words and grammatical markers.

Polysemous Qualities and Universal Networks

The topic of this talk is a reflection about the conceptual organization of qualities involved in polysemous associations and about their universal nature. This analysis follows on from a study carried out by a Franco-German working group on polysemous qualities as expressed in twenty-two African languages. In this frame, I proposed a model of the semantic networks built by the polysemous qualities following the method of semantic maps (Haspelmath, 2003).

The results showed that what is common between each particular network of a specific African language is not exactly the high number of recurring cross-linguistic polysemous associations but rather several semantic networks made up by qualities involved in recurring polysemous associations (see the annex). Such networks seem to be shared by each individual as an idealized cognitive model. That is why I called them “universal networks” - if we consider that the term “universal” does not refer to a systematic rule but a tendency (high or not).

The aim of this presentation so is to go deeper into the above-described results by the way of a confrontation with a sample of Indo-European languages in order to further justify the existence of these universal networks.

The study is based on a cognitive approach as developed by linguists and psycholinguists (Lakoff, 1987; Langacker, 1993; Lazard, 1992; Koch, 2004) as well as philosophers and psychologists (Proust, 1997; Searle, 1985). It will be shown how it is possible to observe cognitive correspondences between (a) the different universal networks, (b) some particular semantic domains - e.g. acrid taste, important dimension, small dimension, strong resistance, weak resistance... - which characterize these networks and (c) some linguistic and cognitive processes involved in the construction of meaning, i.e. inferential processes which are relative to symbolism, iconicity, pragmatism or conceptualization, on a larger cross-linguistic study.

BLANK Andreas, 2000. « Pour une approche cognitive du changement sémantique lexical : aspect sémasiologique », In *Mémoires de la société linguistique de Paris*, Tome IX, Peeters : Paris, p. 59-74

FUCHS Catherine, 1999. “Diversity in linguistic representations: a challenge for cognitive science”, in C. FUCHS & S. ROBERT (eds), *Language diversity and cognitive representations*, Amsterdam, Benjamins, pp. 3-19

HASPELMATH Martin, 2003. « The geometry of grammatical meaning : semantic maps and cross-linguistic comparison ». In TOMASELLO Michael (ed.), *The new psychology of language*, vol. 2, Mahwah, NJ: Erlbaum, pp. 211-242

KOCH Peter, 2000. « Pour une approche cognitive du changement sémantique lexical : Aspect onomasiologique ». In *Mémoires de la société linguistique de Paris*, Tome IX, Peeters : Paris.

KOCH Peter, 2004. “Diachronic onomasiology and semantic reconstruction”. In Wiltrud Mihatsch/Reinhild Steinberg (éds.), *Lexical Data and Universals of Semantic Change*, Tübingen: Stauffenburg, 79-106.

LAKOFF George, 1987. *Women, fire and dangerous things*. The University of Chicago Press : Chicago.

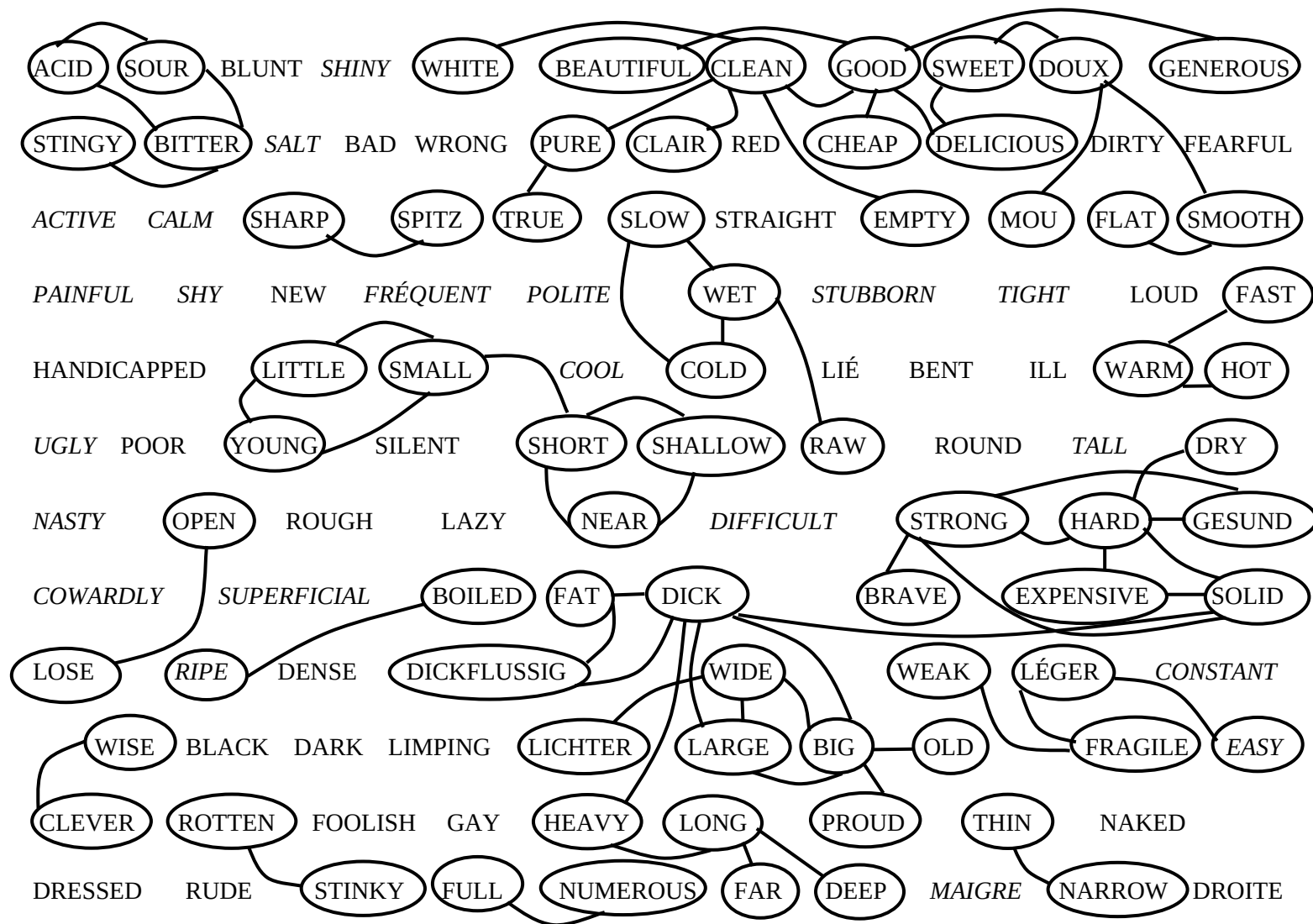
LANGACKER, Ronald W., 1990 *Concept, image and symbol : the cognitive basis of grammar*. Berlin, Mouton de Gruyter.

LANGACKER Ronald W., 1993. « Reference-point constructions ». *Cognitive linguistics* 4-1, pp. 1-38.

LAZARD Gilbert, 1992. « Y a-t-il des catégories interlangagières ? », in S. Anschutz (ed), *Texte, Sätze, Wörter und Moneme*, Festschrift K. Heger, Heidelberg, Heidelberg Orientverlag : 427-434.

PROUST Joëlle (dir.), 1997. « Présentation ». In *Perception et intermodalité, approches actuelles de la question de Molyneux*, Paris, P.U.F., pp. 1-18

SEARLE John, 1985, *Du cerveau au savoir*. Conférences Reith 1984 de la BBC. Hermann, Paris



The big and the little: on the difference between domains and functions in creating semantic maps

This paper argues for a rigorous distinction between functions and domains when it comes to the geometry of semantic maps. We define a domain as consisting of two or more functions that have to be primitive (i.e., that do not themselves consist of functions). It is argued that very often linguistic analyses take a top-down approach where a bottom-up approach would be more fruitful. A top-down approach starts with the domain (e.g., epistemic modality) and asks whether a given morpheme in language X is an instantiation of that domain. This can lead to serious problems if the domain is overly broad or if the domain is too poorly defined to yield reliable cross-linguistic results. An example of an overly broad domain is epistemic modality. While it would appear to be not hard to place a given morpheme in this domain, it is still too broad. It has been argued that English *must* and Swedish *lär* are both instances of epistemic modality, yet in actuality their interpretations do not show any overlap. Other examples of domains that are too broad are various temporal domains, such as past tense. An example of a domain that is too poorly defined to be of any real use in typological studies is that of reality status (realis / irrealis). The various sub-domains that make up this domain vary greatly from language to language (for instance, in some languages, Imperative and Prohibitive are part of realis, in others part of irrealis, in yet others one is part of realis, the other of irrealis). Hence, stating that morpheme X in language Y is an Irrealis morpheme is not very helpful, as they may have a vastly different semantic range.

In a bottom-up approach, we start by examining the semantic range of an individual morpheme and compare that with morphemes from other languages that have a similar range. That way, the emphasis is on functions and more precise comparisons can be made. Later on, one can decide on such issues as which functions comprise a given domain, or where a given domain ends. This approach is especially helpful in deciding on whether a given morpheme belongs to one domain or another. One such example is the current debate about the difference between the domains of epistemic modality and evidentiality which will be used as main example in the paper. It turns out that English *must* has a completely different range than Swedish *lär*, for instance while Dutch *moeten* overlaps with both *must* and *lär*. This only shows up if we consider their semantics on the function level, not on the domain level.

It will be shown that by taking a bottom-up approach we can get a clearer picture of the nature of both domains which in turn can help us to draw the boundary between both domains. Other domains that will benefit from a bottom-up approach are such areas as the perfect (is it a tense or an aspect?) and, to take a slightly different example, possession.

How variation in sampling changes semantic maps built on a comparison of parallel text data in the domain of motion events (verb stems and case/adpositions)

This paper deals with non-implicational semantic maps, built automatically using classical multi-dimensional scaling from a direct comparison of parallel text data (the Gospel according to Mark) in the domain of motion events (verb stems and case/adpositions) in more than 100 languages from all continents in more than 300 parallel clauses, and investigates how robust the result is if different biased smaller samples of (a) selected languages or (b) selected clauses are taken as a basis for the analysis. For implicational semantic maps, Haspelmath (2003: 217) claims that “[e]xperience shows that it is generally sufficient to look at a dozen genealogically diverse languages to arrive at a stable map that does not undergo significant changes as more languages are considered.” This paper explores what are the actual differences in the results if the method is applied to subsamples of a dozen or more languages of particular continents (Africa, Eurasia, Oceania, the Americas) or particular language families (Indo-European, Austronesian, Niger-Congo). In the same way it is investigated what differences can result if the sample of clauses is manipulated. It is found that the selection of analytical primitives (the objects represented on the semantic map representing the functions compared cross-linguistically) is at least equally important as the sampling of languages.

The theoretical background of the approach taken here is briefly summarized below: Both implicational and non-implicational semantic maps (for the latter cf. Cysouw submitted, Wälchli submitted) rely on the single principle that cross-linguistically recurrent identity in form reflects similarity in meaning. This principle rests—implicitly or explicitly—on a theory of similarity semantics. Similarity semantics is concerned with the network of similarity and dissimilarity relationships emerging from the cumulative pairwise comparison of meanings and is particularly appropriate for comparing the meanings of contextually-embedded concrete utterances without having to assume an arbitrary set of primitive semantic units, since it operates without any reference to the notion of semantic identity (identical meanings, if there is such a thing in context, are treated as maximally similar meanings). Even if radically different, similarity semantics shares with Fregean truth semantics its indirect approach to meaning. It is not argued that similarity semantics is the only way how humans process meaning; rather the various forms of semantic map approaches are empirical methods to investigate how far we can go with similarity semantics alone, based on a minimal set of basic entities, similarly, e.g., as those assumed in Locke’s (1690-1714, book iv, ch. i) theory of knowledge. Similarity semantics is also compatible with the structuralists’ (de Saussure, Hjelmslev) concept of meaning as a continuous mass, analogous to the continuous phonetic space. Like phonemes, semantic categories of particular languages categorize particular areas of the continuous semantic space and typology is an indirect method to reconstruct the underlying semantic space, which cannot be measured directly unlike the articulatory-acoustic space in phonetics.

In semantic map approaches one has to distinguish strictly between underlying theoretical assumptions, such as outlined above, and aspects of the practical method applied. Most practical approaches to semantic maps (including the one applied here) assume that cross-linguistically identified “translation-equivalent” functions are identical, where they are in fact only similar. The semantic map approach works in practice to the extent that cross-linguistically identified functions are more similar than the functions of particular languages to be compared. Since cross-linguistically different categorization patterns can be distinguished in semantic maps only to the extent that there is a sufficiently high resolution of analytic primitives, it is crucial to match functions to be identified cross-linguistically as sharply as possible, which is best done by matching concrete examples. This is also necessary because concepts are at least partly based on exemplary knowledge in natural languages (Goldberg 2006: 229; Marty 1908: 530).

References

- Cysouw, Michael (submitted). Building semantic maps: the case of person marking. In M. Miestamo, & B. Wälchli (eds.), *New Challenges in Typology: Broadening the Horizons and Redefining the Foundations*.
- Goldberg, Adele E. (2006). *Constructions at work. The nature of generalization in language*. Oxford: Oxford University Press.
- Haspelmath, Martin (2003). The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In Michael Tomasello (ed.), *The New Psychology of Language 2*: 211-242. Mahwah, NJ: Lawrence Erlbaum.
- Locke, John (1975 [1690-1714]). *An essay concerning human understanding*. edited by P.H. Nidditch. Oxford: Clarendon.
- Marty, Anton (1908). *Untersuchungen zur Grundlegung der allgemeinen Grammatik und Sprachphilosophie*. Erster Band. Halle: Niemeyer.
- Wälchli, Bernhard (submitted). Constructing semantic maps from etic parallel text data.

Abstract for ALT VII

What do "do" verbs do? Towards a typology of generalised action verbs

All languages appear to have one or more ‘generalised action verbs’ (Van Valin & LaPolla 1997), which, like English *do* in *Who **did** this?*, are used as ‘pro-verbs’ in contexts where the nature of an event is unknown or left unspecified.

It has been claimed in the literature that the concept ‘DO’ is universal, and moreover, that it is universally linked to the notion of agency (e.g. Goddard & Wierzbicka 1994: 42-3). This assumption is also at the heart of proposals to use DO in the semantic decomposition of verbs, pioneered by Dowty (1979: 110-125). Foley & Van Valin (1984: 47-53) and Van Valin & LaPolla (1997: 102-129) even use two kinds of ‘do’ operators, one to represent activities, and one to represent agency.

Through a cross-linguistic study of generalised action verbs, it will be demonstrated that they are not necessarily agentive in nature, but may cover a wide range of functions, including the following:

- Verb of manufacturing, as e.g. in German (*machen*) and French (*faire*)
- Causative verb, as e.g. in French (*faire*) (cf. Moreno 1993)
- Grammaticalised auxiliary, as e.g. in English (*do*)
- Marker of quotations, as e.g. in many Northern Australian languages (Rumsey 1994; McGregor 1994), Papuan languages (Foley 1986: 119), and African languages (Güldemann 2001: 237-245)
- Verbaliser with non-verbal predicates or onomatopoeia, as in the German example *und auf einmal machte es “platsch”* ‘and suddenly it went (lit. ‘made’) “splash”’
- Inchoative verb, as e.g. in Wintu (Pitkin 1985: xii-xix), Yimas (Foley 1991: 293-301, 334-336), and Samoan (Mosel & Hovdhaugen 1992: 113)
- Eventive Verb, i.e. a verb translated as ‘happen’, ‘occur’, as e.g. in Hopi (*hìnti* in the Hopi Dictionary Project, 1998) and Yimas (Foley 1991: 293-301, 334-336)
- A verb used to render a feeling or emotional reaction, as e.g. in Hopi (*hìnti*, see above) and Yimas (Foley 1991: 293-301, 334-336), and Kalam (Pawley 1994: 407-8)
- A verb used to predicate a quality of an entity with a nominal or adverbial complement (e.g. Ewe *wɔ* in *é-wɔ ké* ‘it is sandy’, Ameka 1994: 71).

Note that the semantic range of these verbs includes a number of concepts for which primitives that are supposedly semantically distinct from DO have been introduced in the literature, e.g. CAUSE and BECOME in the Foley/Van Valin/LaPolla framework, and HAPPEN, SAY, and FEEL, which are claimed to be semantic primitives in Wierzbicka’s ‘Natural Semantic Metalanguage’ framework (cf. Wierzbicka 1994). The cross-linguistic data suggest, however, that the range of functions of generalised action verbs is by no means random, and that similar functions are found in numerous unrelated languages. I will propose a semantic map accounting for the most frequent functions and their formal and semantic relationships, as well as linking these to the paths of grammaticalization that are attested for generalised action verbs.

References

- Ameka, Felix K., 1994. Ewe. In Cliff Goddard & Anna Wierzbicka (eds.), *Semantic and Lexical Universals*. Amsterdam: John Benjamins. 57-86.
- Dowty, David R., 1979. *Word meaning and Montague Grammar. The semantics of verbs and times in Generative Semantics and in Montague's PTQ*. Dordrecht: Reidel.
- Foley, William A., 1986. *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.
- Foley, William A., 1991. *The Yimas language of New Guinea*. Stanford: Stanford University Press.
- Foley, William A. & Robert D. Van Valin Jr., 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Goddard, Cliff & Anna Wierzbicka, 1994. Introducing lexical primitives. In Cliff Goddard & Anna Wierzbicka (eds.), *Semantic and Lexical Universals*. Amsterdam: John Benjamins. 31-45.
- Güldemann, Tom, 2001. *Quotative constructions in African languages: a synchronic and diachronic survey*. Leipzig: Habilitationsschrift, Fakultät für Geschichte, Kunst- und Orientalwissenschaften der Univ. Leipzig.
- Hopi Dictionary Project, 1998. *Hopi Dictionary - Hopìikwa Lavytutuveni*. Tucson: The University of Arizona Press.
- McGregor, William, 1994. The grammar of reported speech and thought in Gooniyandi. *Australian Journal of Linguistics* 14,1: 63-92.
- Moreno, Juan Carlos, 1993. 'Make' and the semantic origins of causativity: a typological study. In Bernard Comrie & Maria Polinsky (eds.), *Causatives and Transitivity*. Amsterdam: John Benjamins. 155-164.
- Mosel, Ulrike & Even Hovdhaugen, 1992. *Samoan Reference Grammar*. Oslo: Scandinavian University Press.
- Pawley, Andrew, 1994. Kalam Exponents of lexical and semantic primitives. In Cliff Goddard & Anna Wierzbicka (eds.), *Semantic and Lexical Universals*. Amsterdam: John Benjamins. 387-422.
- Pitkin, Harvey, 1985. *Wintu Dictionary*. Berkeley and Los Angeles: University of California Press.
- Rumsey, Alan, 1994. On the transitivity of 'say' constructions in Bunuba. *Australian Journal of Linguistics* 14, 2: 137-153.
- Van Valin, Robert D. & Randy J. LaPolla, 1997. *Syntax. Structure, meaning and function*. Cambridge: Cambridge University Press.
- Wierzbicka, Anna, 1994. Semantic Primitives across languages: a critical review. In Cliff Goddard & Anna Wierzbicka (eds.), *Semantic and Lexical Universals*. Amsterdam: John Benjamins. 445-500.

Analytical dimensions and the functional map of Parts-of-Speech

Kees Hengeveld & Eva van Lier, University of Amsterdam

The meaning of linguistic units can be analyzed at two different levels. In Functional Discourse Grammar (Hengeveld & Mackenzie, forthcoming), the *Interpersonal* level gives a formal representation of linguistic units in terms of two basic communicative functions, called ascriptive acts and referential acts. At the *Representational* level, linguistic units are described in terms of their semantic designation. Units at the Representational level may correspond to different Interpersonal functions. Furthermore, each linguistic unit consists of an obligatory part, its head, and an optional modifier.

In this paper (cf. Hengeveld & Van Lier, submitted), we use the distinction between the Interpersonal and the Representational levels of analysis on the one hand, and the head-modifier distinction on the other hand, to define the following four possible functions of lexical items:

- (i) The head of a representational unit that is used as an ascriptive act.
- (ii) A representational unit that is used as a modifier of the head of an ascriptive act.
- (iii) The head of a representational unit that is used as a referential act.
- (iv) A representational unit that is used as a modifier of the head of a referential act.

Cross-linguistically, there is considerable variation in terms of the freedom of lexeme classes to express one or more of these four functions (Hengeveld 1992 and Hengeveld et al. 2004). Originally, the constraints on this variation were formulated in terms of the implicational hierarchy below, where the four functions are ordered according to the likelihood that a language would have a specialized lexeme class for the expression of that function (with the chance of specialization increasing to the left).

(i) > (iii) > (iv) > (ii)

In the present paper, however, we argue that this hierarchy is in fact the superficial reflection of a two-dimensional functional map, based on the analytical primitives Head-Modifier and Ascription-Reference, as shown in the figure below:

	Head	Modifier
Ascription	(i)	(ii)
Reference	(iii)	(iv)

Each of the two dimensions independently reflects a predominance relation: (Head > Modifier) and (Ascription > Reference). In addition, the two relations are hierarchically ordered with respect to one another: ((Ascription > Reference) > (Head > Modifier)).

On the basis of data from a 50-language variety sample, we show that this two-dimensional map approach yields a higher coverage than the original proposal, while still enabling a clear-cut and cross-linguistically comparable description of the mapping of groups of lexemes onto a functional space.

The map that we propose is not a semantic map in the strict sense, because it does not take into account the entity types of the units that express the various functions (cf. Croft 2001, Van Lier 2006). Taking up this point, we will explore the possibilities of enriching our functional map with this third dimension.

References:

- Croft, William 2001: *Radical Construction Grammar*. Oxford: Oxford University Press.
- Hengeveld, Kees 1992: Parts of Speech. In: Michael Fortescue, Peter Harder & Lars Kristoffersen (eds.): *Layered structure and reference in a functional perspective*. Amsterdam/Philadelphia: Benjamins. 29—55.
- Hengeveld, Kees, Jan Rijkhoff & Anna Siewierska. 2004: Parts-of-speech systems and word order. *Journal of Linguistics* 40: 527—570.
- Hengeveld, Kees & Eva van Lier, submitted: 'Lexical and complex heads in Functional Discourse Grammar.'
- Hengeveld, Kees & J. Lachlan Mackenzie, forthcoming: *Functional Discourse Grammar*. Oxford: Oxford University Press.
- Van Lier, Eva 2006: Parts-of-Speech systems and dependent clauses. *Folia Linguistica* 40, 3-4, 239—304.

Modality's semantic map revisited

Semantic maps essentially account for the synchronic polyfunctionality of linguistic constructions. This polyfunctionality is taken to result from diachronic evolution. Maps may or may not represent claims about the directionality of the presupposed evolution. If they do, the lines that connect the contiguous meanings are arrows. Often, though not always, developments are strictly unidirectional. The issue that the paper will illustrate is this: what do we do when we have a map with a well argued directionality hypothesis that certain data appear to violate? Just like for data that appear to flout the contiguity requirement (meanings covered by a marker have to be contiguous or go back to a common ancestor), one can either give up part of the semantic map or look for a non-semantic motivation. One such non-semantic motivation appeals to language contact, the idea being that language contact may steer constructions in directions not allowed by the semantics of the map.

This general problem will be illustrated with the semantic map of modality, as proposed by van der Auwera & Plungian (1998). This map, as well as some of the work on which it is based (esp. Bybee et al 1994), describes a directionality from participant-internal possibility (also 'ability'), as in (1), to participant-external possibility (also 'circumstantial possibility') as in (2).

- (1) I *can* swim.
- (2) To reach the station you *can* take bus 66.

The recalcitrant data involve modals that derive from the lexical item 'get'; the resulting modals can be called 'acquisitive modals'. The two hotbeds of acquisitive modality are the Baltic area and Southeast Asia (on the latter see Enfield 2002). In both there is at least indirect evidence for a development from participant-external to participant-internal modality: languages in these areas may employ an acquisitive lexeme for only participant-external possibility, for both participant-internal and participant-external possibility, for neither of the two; but never for only participant-internal possibility. The relevant languages belong to different families, e.g. Chinese, Thai, Vietnamese, Hmong in Southeast Asia or Swedish, Finnish, and Latvian in the Baltic. Both areas also testify strong contact inference, and there are claims in the literature that specifically point to the relevance of contact interference for the fate of acquisitive modals. Although we give due consideration to language contact, we will nevertheless show that there is enough direct diachronic evidence (from Chinese) and that there is a sufficiently plausible semantic scenario for the development of acquisitive modals for us to revise the relevant part of the original semantic map. It will also be shown that the revised idea about the relation between participant-internal and participant-external possibility carries over to necessity. The mistake of 8 years ago will be argued to stem from a Standard Average European bias.

Bybee, J., Perkins, R. and Pagliuca, W. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.

Enfield, N.J. 2002. *Linguistic Epidemiology: Semantics and Grammar of Language Contact in Mainland Southeast Asia* [Curzon Asiatic Linguistics]. London and New York: Routledge.

van der Auwera, J. and Plungian, V. 1998. "Modality's semantic map". *Linguistic Typology* 2:79-124.

A semantic map of epistemic expressions – Abstract

Epistemic expressions are defined here as linguistic items and constructions that express either degree of certainty (e.g. certainty, doubt, probability, epistemic necessity, or epistemic possibility) or source of information (e.g. direct, indirect-inferential, or indirect-reportive evidence), or both.

For some time now, epistemic expressions have been intensively studied, and the semantic-map approach has been applied to them. Anderson 1986 provides a semantic map of grammaticalized expressions of source of information (“evidentials”), and van der Auwera & Plungian 1998 provides a semantic map of expressions of epistemic as well as non-epistemic necessity and possibility. However, while it is undisputed that epistemic expressions are semantically closely related to each other, no one has so far provided a semantic map that takes into account both expressions of degree of certainty and expressions of source of information. In a couple of studies (notably, Givón 1982 and Akatsuka 1985) degree of certainty and source of information are related to each other in terms of a scale (an “epistemic scale”), which is not far from being a genuine semantic map. But the relevant studies draw upon data from only three to four languages.

This paper presents a unified semantic map of epistemic expressions – that is, a map which covers both different degrees of certainty and different types of information source. The map is based on a survey of epistemic expressions from more than 50 languages representing geographical as well as genetic diversity. It is compatible moreover with data from an additional great number of languages discussed in Givón 1982, Akatsuka 1985, Bybee, Perkins & Pagliuca 1994, and Aikhenvald 2004. The main features of the semantic map are as follows: 1) The meanings (or functions) of epistemic expressions constitute a continuous region – that is, each of the epistemic meanings distinguished is connected to at least one other epistemic meaning by what Haspelmath 2003 refers to as a “connecting line”. 2) Within the overall continuous region, degree-of-certainty meanings make up one continuous subregion, while source-of-information meanings make up another one – that is, it holds for both types of meaning that each meaning distinguished is connected to at least one other meaning of the same type by a connecting line. 3) The two subregions are connected to each other in a systematic way: while high degree of certainty is connected to highly reliable source of information (i.e. direct evidence) by a connecting line, less degree of certainty is connected to less reliable source of information (i.e. indirect evidence).

With these features the semantic map of epistemic expressions has important implications for the discussion of the relationship between ‘epistemic modality’ and ‘evidentiality’. However, the map has implications for ‘semantic-mapping theory’ as well. In a discussion of the different connecting lines of the map the paper argues that a distinction should be made between essentially conceptual and essentially functional connecting lines – thus, one might prefer to talk about ‘functional-conceptual space’ rather than about “conceptual space” (e.g. Croft 2003). Subsequently, in an outline of general properties of epistemic expressions the paper argues that a distinction should be made between connecting lines internal and connecting lines external to a semantic domain.

References

- Aikhenvald, A. Y. 2004. *Evidentiality*. Oxford: Oxford University Press.
- Akatsuka, N. 1985. “Conditionals and the epistemic scale”. *Language* 61.3: 625-639.
- Anderson, L. B. 1986. “Evidentials, paths of change, and mental maps: Typologically regular asymmetries”. Chafe & Nichols 1986 (eds.). *Evidentiality: The Linguistic Coding of Epistemology*. Norwood: Ablex. 273-312.
- Bybee, J., Perkins, R. D. & Pagliuca, W. 1994. *The Evolution of Grammar*. Chicago: University of Chicago Press.
- Croft, W. 2003. *Typology and Universals* [2nd ed]. Cambridge: Cambridge University Press.
- Givón, T. 1982. “Evidentiality and epistemic space”. *Studies in Language* 6.1. 23-49.
- Haspelmath, M. 2003. “The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison”. Tomasello, M. (ed.). 2003. *The New Psychology of Language* 2. London: Lawrence Erlbaum. 211-242.
- van der Auwera, J. & Plungian, V. A. 1998. “Modality’s semantic map”. *Linguistic Typology* 2. 79-124.

A semantic map for epistemic and inferential functions

The purpose of this presentation is to consider the issue of representing various epistemic and inferential functions in the form of a semantic map. Inferential functions may be interpreted as purely evidential, or they may combine both evidential and epistemic notions. From a typological perspective, epistemic and inferential functions have very intricate relations to each other and to some other functions, too. Many of these relations are not represented in the map, proposed by van der Auwera and Plungian (1998). It is important for the development of the semantic map methodology to find an adequate way of representing these kinds of relations.

The presentation is based on the obtained results of my typological study of epistemic modality and its relation to evidentiality. In this study, I have selected a genealogically stratified sample of the languages of the world (60 languages). On the selected languages, reference material, consisting of descriptions of epistemic modality and evidentiality, has been used. The data consists of various kinds of grammatical expressions, considered in these descriptions. By means of a detailed analysis of the data, it is argued in this study that epistemic modality and evidentiality are closely related categories. Both of these categories delineate the speaker's attitude to the truth-value or factual status of the proposition. An epistemic modality marker indicates the speaker's attitude towards the proposition in terms of the degrees of certainty. An evidential marker indicates the source of the speaker's attitude. Especially various types of inferential functions often combine epistemic and evidential notions. In these functions, either epistemic or evidential notions are predominant. These notions can often be described by means of two semantic parameters and their values (cf. Vilkki 2006a). For example, the English modal *must* has many functions, and one of them can be described as "the degree of the speaker's certainty: certain" and "the source of the speaker's attitude: the speaker infers P on the basis of general knowledge". In addition, the value of the first parameter is predominant. For the description of the inferential functions of some languages, a third parameter, indicating "degrees of the reliability of evidence", is needed. The values of this parameter can combine with the values of the second parameter or with the values of both the first and the second parameter. They are rarely predominant.

In the presentation, it will be argued that the parameter "the degree of the speaker's certainty" should be used as a vertical dimension and the parameter "the source of the speaker's attitude" should be used as a horizontal dimension on the semantic map for epistemic and inferential functions. The parameter "degrees of the reliability of evidence" is, however, difficult to represent as a third dimension, because the combinations of the values of this parameter and the values of the other two parameters are quite variable. Therefore, an alternative way of representing this parameter, using different shades, will be discussed. The names of the specific functions of epistemic and inferential expressions are based on the values of the three parameters. Selection of the relevant functions and arranging the functions on the map follows the principles presented by, for example, Haspelmath (2003). It will be proposed that the predominant values of the functions could be indicated by using bold lines and shades. Some implicational universals are also presented. Finally, the possibility of using semantic maps for the representation of secondary functions of epistemic and inferential expressions will also be briefly discussed (cf. Vilkki2006b).

References:

- Haspelmath, M. (2003) The geometry of grammatical meaning: semantic maps and cross-linguistic comparison. In M. Tomasello (ed.) *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, Volume 2:211-242. Mahwah, NJ: Erlbaum.
- van der Auwera, J. and Plungian, V.A. (1998) Modality's semantic map. *Linguistic Typology* 2: 79-124.
- Vilki, L. (2006a) A network-based description of the polysemy of Russian epistemic parentheticals. Paper presented at Perspectives on Slavistics 2006 conference, Regensburg, Germany.
- (2006b) Politeness, face and facework: current issues. In M. Suominen & al. (eds.) *A Man of Measure: Festschrift in Honour of Fred Karlsson on his 60th Birthday*. A Special Supplement to SKY Journal of Linguistics, Volume 19:322-332. Turku: The Linguistic Association of Finland
(<http://www.ling.helsinki.fi/sky/julkaisut/sky2006special.shtml>).

The twofold conceptual space of coordination relations

Caterina Mauri (University of Pavia)

The aim of this paper is to depict the conceptual space within which the three basic coordination relations of *combination* ('and'), *contrast* ('but') and *alternative* ('or') are located (Croft's distinction between 'semantic map' and 'conceptual space' will be followed here, cf. Croft 2003: 144-52). The notion of coordination relation is defined in purely functional terms as a relation established between functionally parallel states of affairs (henceforth SoAs), i.e. each having an autonomous cognitive profile and the same illocutionary force (see Mauri 2007: chapter 2). Every construction used to establish one or more coordination relations is considered a coordinating construction, regardless of its morphosyntactic properties.

As pointed out, among others, by Dik (1968) and Haspelmath (2004), further subtypes may be identified within each coordination relation. Combination may be TEMPORAL (simultaneous vs. sequential) or ATEMPORAL, depending on the location of the SoAs on the temporal axis. Contrast may be OPPOSITIVE, CORRECTIVE or COUNTEREXPECTATIVE, depending on the origin of the conflict (cf. Haspelmath, to appear). Alternative may be SIMPLE or CHOICE-AIMED, depending on the necessity to make a choice between the available possibilities (cf. 'standard' vs. 'interrogative' disjunction, Haspelmath (to appear)). This research, based on a 74 language sample, examines the cross-linguistic coding of the three basic coordination relations and their subtypes with respect to two parameters: (i) the presence and morphophonological complexity of overt coordinating markers (mono-/polymorphemic, mono-/polysyllabic markers), and (ii) the semantic domain of each attested marker, that is, the set of relations it may be used for (general vs. dedicated markers).

Two main results have been achieved in this survey. First of all, the semantic domains of the attested markers have revealed a neat bipartition within the coordination conceptual space, which relates combination to contrast on the one hand and combination to alternative on the other hand. As exemplified in Fig. 1, combination and contrast markers show recurrent overlapping polysemy patterns across languages, pointing to the following combination-contrast conceptual space: [*sequential comb* - *simultaneous comb.* - *atemporal comb.* - *oppositive contrast* - *corrective contrast* - *counterexpectative contrast*] (see Malchukov 2004 for a slightly different assessment). To the contrary, combination and alternative relations tend to be coded by means of completely different markers, thus showing a reduced semantic overlap. However, in languages with no overt marker for alternative, the two relations are expressed by means of the same construction, namely alternative is systematically conveyed through the combination of possibilities. In such cases, the potential status of each combined SoA is obligatorily marked by means of some irrealis markers (like *maṇaya* in example (1), cf. Mauri, forthcoming). No polysemy pattern is attested between the coding of contrast and alternative.

Secondly, the exam of the morphophonological complexity of the attested markers highlights the hierarchical structure characterizing the twofold coordination conceptual space. As highlighted by Kortmann (1997: 78) for subordinators, a simple morphophonological structure tends to correlate with a basic and general semantics, mainly because markers expressing basic and general relations have a high frequency of use and consequently undergo a high morphophonological erosion (Croft 2003: 110-16). This form-function asymmetry is mirrored by data in the sample. Combination markers, which express the most basic and unspecified relation, are structurally simpler than both contrast and alternative markers, and general markers are structurally simpler than dedicated ones. In particular: (i) if a language has one of the markers indicated on the following hierarchy, it will be at least as morphophonologically complex as the markers to its left: [*dedicated marker for sequential combination*, *general marker expressing at least one combination relation* > *general marker only expressing contrast relations* > *dedicated marker for a contrast relation*]; (ii) in a language, markers used to express alternative relations, either general or dedicated, are at least as morphophonologically complex as the markers used to express at least one combination relation. The comparison of contrast and alternative markers, instead, does not reveal any regular cross-linguistic pattern.

To conclude, I will argue that combination, contrast and alternative do not stand on the same level, but combination is more basic and is implied by the other two relations. Based on the attested polysemy patterns and on the morphophonological complexity of the coordinating markers, I propose a twofold, hierarchical conceptual space, structured along two perpendicular axes of increasing semantic specificity having their origin in the combination relation (Fig. 2). On the one hand, a combination of SoAs may be specified in terms of some discontinuity (Givón 1990: 849) originating a contrast. On the other hand, a combination may be specified in terms of the irrealis of the SoAs it links, creating a set of alternative possibilities. Along the two axes, the more a coordination relation is semantically specified, the more complex will be the marker expressing it.

References

- Croft, W. (2003). *Typology and universals* (2nd ed.). Cambridge: Cambridge University Press.
- Dik, S. (1968). *Coordination. Its implications for the theory of general linguistics*. Amsterdam: North-Holland Publishing Company.
- Givón, T. (1990). *Syntax: a functional-typological introduction*, Volume II. Amsterdam; Philadelphia: John Benjamins.
- Haspelmath, M. (2004). Coordinating constructions: an overview. In M. Haspelmath (ed.), *Coordinating constructions*, pp. 3-39. Amsterdam: John Benjamins.
- Haspelmath, M. (to appear 2007). Coordination. In T. Shopen (ed.), *Language typology and linguistic description* (2nd ed.). Cambridge: CUP.

- Kortmann, B. (1997). *Adverbial subordination. A typology and history of adverbial subordinators based on European languages*. Berlin/New York: Mouton de Gruyter.
- Malchukov, A. L. (2004). Towards a Semantic Typology of Adversative and Contrast Marking. *Journal of Semantics* 21, 177–198.
- Mauri, C. (2007). *Coordination relations: a cross-linguistic study*. University of Pavia: PhD Dissertation.
- Mauri, C. (forthcoming). The irreality of alternatives: towards a typology of disjunction. *Studies in Language* 31.
- Merlan, F. (1982). *Mangarayî*, Volume 4 of *Lingua Descriptive Studies*. Amsterdam: North Holland Publishing Company.

Figures and examples

- (1) Mangarayi, Gunwingguan, Australian (Merlan 1982: 39)

<i>manjaya</i>	<i>ja-Ø-niya-n</i>	<i>manjaya</i>	<i>dayi</i>
perhaps	3-3sg-come-PRES	perhaps	NEG

‘Perhaps he’ll come, perhaps not.’, i.e ‘it is possible that he may or may not come’

	SEQ	SIM	ATEMP	OPPOSIT	CORRECT	COUNTEREX
Mangarayi	—					gana
Tuvaluani	—	kae				
Hdi	lá	—				ám á
Maori	aa	hoki	—		engari	
Hausa	kuma					amman
Russian	i		a			no
S-Croatian	i		a		věč	ali
Polish	i		a		ale	
German	und				sondern	aber
Italian	e				ma	
	mentre			benà	però	

Figure 1: *The combination-contrast conceptual space: some attested semantic maps.*

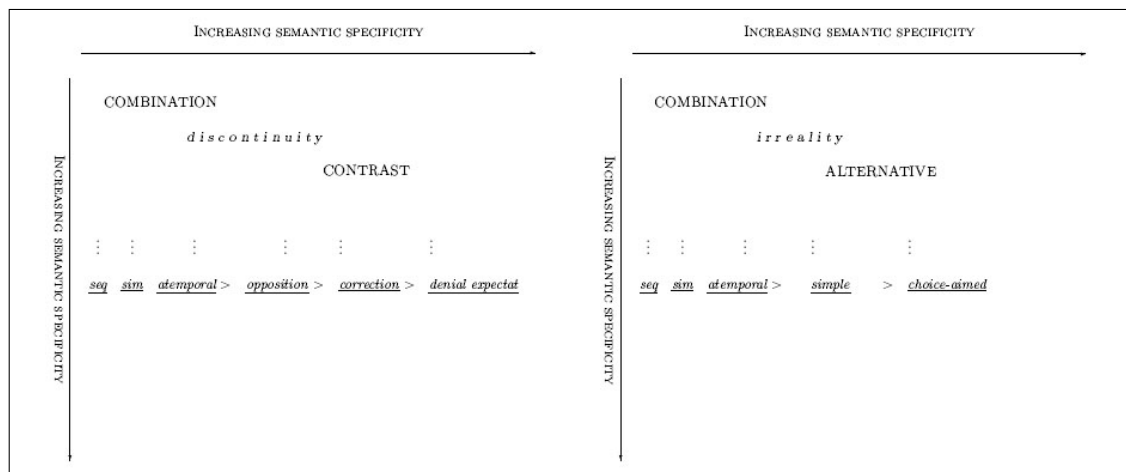


Figure 2: *The conceptual space of coordination relations: two dimensions of increasing semantic specificity.*

Grammaticalisation and Semantic Maps

Luc Steels^{1,2} and Remi van Trijp²

¹ VUB AI Lab, Vrije Universiteit Brussels – Pleinlaan 2 – 1050 Brussels (Belgium)

² Sony Computer Science Laboratory, Paris – 6, Rue Amyot – 75005 Paris (France)

Semantic Maps have offered linguists an appealing and empirically rooted methodology for describing recurrent structural patterns in how languages categorize conceptual space. They have also been argued to provide a route through which grammaticalisation processes operate. Although some researchers argue that semantic maps are universal and given (Haspelmath (2003) and Croft and Poole (forthcoming)) others provide evidence that there are no fixed nor universal maps (Cysouw (forthcoming)).

Here we take the position that semantic maps are a useful way to map out the grammatical evolution of a language (particularly the evolution of semantic structuring) but that this grammatical evolution is a consequence of distributed processes whereby agents shape and reshape their language. So it is a challenge to find out what these processes are and whether they indeed generate the kind of semantic maps observed for human languages. Semantic maps of different languages will be similar because the same evolutionary pathways are followed.

In our work, we have taken a design stance towards the question of the origins of linguistic structure. In our experiments we investigate the emergence of grammatical features in populations of autonomous artificial agents that play language games about situations they perceive through a sensori-motor embodiment. In past experiments, we have already shown how such a population could self-organize and develop a lexicon and a shared conceptual space for objects by playing locally situated “language games” (Steels, 2003). In this talk, we will present new experiments in which we investigate whether semantic maps for case markers could emerge through grammaticalisation processes without the need for an underlying universal conceptual space.

References:

- Croft, William, and Keith T. Poole (forthcoming). *Inferring universals from grammatical variation: multidimensional scaling for typological analysis*.
- Cysouw, Michael (forthcoming). Building semantic maps: the case of person marking. In: Bernard Wälchli & Matti Miestamo (eds.), Berlin: Mouton.
- Haspelmath, Martin (2003). The geometry of grammatical meaning: semantic maps and cross-linguistic comparison. In: M. Tomasello (ed.), *The new psychology of language*, vol. 2, New York: Erlbaum, 211-243.
- Levinson, Stephen C. (2003). *Space in language and cognition: Explorations in cognitive diversity*, Cambridge: Cambridge UP.
- Lucy, John (1992). *Grammatical categories and cognition: a case study of the Linguistic Relativity Hypothesis*, Cambridge: Cambridge UP.
- Steels, Luc (2003). Evolving grounded communication for robots. *Trends in Cognitive Science*. 7(7), July 2003, pp. 308-312.
<<http://www.csl.sony.fr/downloads/papers/2003/steels-03c.pdf>>

Diachronic issues in a map of case functions

The semantic map approach has recently gained a wider currency in typological studies, but in many areas substantial research that established the basis for semantic maps preceded its emergence. One of these areas is case functions.

Currently, there are two major manners of representation in semantic maps. One, as favored for example by Croft (2001 etc.), puts the emphasize on degree of similarity represented through degree of spatial adjacency. The logical continuation of such an approach is the conceptualization of the relationship between two meanings or functions on the basis of their statistical frequency of co-occurrence in the same linguistic form. The other approach, as favored saliently by Haspelmath (2003 etc.), pursues the possibility of specific connections between individual meanings to the exclusion of other connections which are in principle possible in terms of similarity as well, but supposedly do not actually occur, for cognitive-conceptual or other reasons. In other words, the latter approach posits the existence of various constraints on configurations on a semantic map, while the former in principle does not. While it is not at all clear yet which approach comes closer to linguistic 'reality', the author of this abstract assumes it to be more profitable to pursue the latter approach (the 'individual connection approach') as far as it can be supported by the data, simply because more constraints also mean more informativeness. Especially, individual connections, if dynamicized, can also be related relatively easily to a diachronic dimension and to grammaticalization research.

Thus, this paper seeks to explore the interrelation between the connection between case functions on a semantic map of this area, and the diachronic relationships between case functions that have been posited in grammaticalization research already relatively early (Heine et al. 1991, Lehmann 1983; 2002). The focus is on the area of instrument-comitative, already treated earlier by the author (Narrog&Ito (to appear)) and the related area of agent-ablative. It is shown that some connections in this area are almost universally acknowledged and relatively unproblematic (e.g. comitative > instrument) while others are controversial and even have the potential to contradict seemingly universal tendencies of grammaticalization (e.g. instrument <> agent). The primary goal of this presentation will be to clarify the directionality of controversial connections in this area on the basis of a 200-languages sample.

How “semantic” are semantic maps?

A pilot study of passive and impersonal constructions in European languages

Andrea Sansò
Università di Pavia

1. Introduction and aims. Semantic maps are often defined as *multi-level* representations of linguistic meaning/function in which each point represents a *semantic structure* associated with one or more grammatical entities (or grams), and the connections between points represent relations between the functions/meanings of grams. How these “semantic structures” should look like is largely an individual choice of the creator of the map, and often it is not easy to tell if we are dealing with different usages or with different meanings/senses of grams. As a result, *function*, *meaning*, *sense*, and *usage* are used by practitioners of semantic maps as if they were interchangeable, and the claim underlying this method, be it explicitly stated or not, is that different contextual meanings (=usages/functions) of a given grammatical entity directly reflect its conventional meanings (=senses/meanings), both being part of the semantic characterization of that entity (for an exemplar discussion, see Haspelmath 2003: 212-213). Moreover, although in principle the semantic-map approach to cross-linguistic diversity is able to transcend the boundary between sentences and discourse (see, e.g., Croft [2001: 93, adapted]: “conceptual spaces also represent conventional *pragmatic* or *discourse-functional* or *informational-structural* or even *stylistic* and *social* dimensions of the use of a grammatical form or construction”), semantic maps have rarely been used in the realm of discourse in a systematic way. This paper is an attempt at making the semantic structures that form semantic maps more suitable to deal with phenomena traditionally falling within the realm of discourse (such as, e.g., voice phenomena, anaphoric relations, topic/focus constructions, etc.). The purpose of this paper is thus twofold. First and foremost, I will use discourse micro-structures as a diagnostics for building a semantic map of *agent defocusing* (Myhill 1997, Sansò 2006), a general function that is manifested in a variety of ways in the languages of the world, and that appears to be preferentially associated with passive and impersonal constructions across languages. The second aim is more general: I will illustrate how *discourse-functional* or *informational-structural* dimensions of the use of a grammatical form may be captured by making use of semantic maps. Passive and impersonal constructions, being highly sensitive to discourse conditions, are an ideal domain for this purpose.

2. Corpus and data. The corpus used in this pilot study consists of Umberto Eco’s novel *Il nome della rosa* along with its translations in 9 European languages (Spanish, Romanian, French, German, Dutch, Danish, Modern Greek, Polish, Czech). The construction types analyzed in this study include: (i) so-called *periphrastic passives*, in which the verb phrase consists of an auxiliary plus the past participle of the verb; (ii) *inflectional passive/medial paradigms*; (iii) passive and impersonal constructions in which a reflexive marker is used (labelled as *middle constructions*, following Abraham 1995, Steinbach 2002, among others); (iv) so-called *impersonal passives*, i.e. constructions in which the predicate is associated with passive morphology, but either there is no patient (i.e. the corresponding active clause is intransitive), or the patient is marked in the same way in which it is marked in the active sentence; (v) so-called *man-constructions*, i.e. constructions having some general noun (“man”, “people”) as subject; (vi) constructions involving the impersonal or vague use of a personal pronoun, or the corresponding inflected form of the verb (so-called “vague *you*” and “vague *they*” constructions).

3. Results. Even in a typologically and genetically homogeneous language sample, structurally similar constructions show considerable differences in use (see, e.g., Figures 2-4): these differences are not chaotic, but systematic to a certain extent, and can be captured through a careful inspection of texts, which alone can shed light on semantic nuances that would otherwise be downplayed or ignored. These differences can be formalized by means of a conceptual space whose nodes are not atomic meanings/functions, but **clusters of discourse properties of the event and its main participants** (A[gent] and P[atient]; see Figure 1): the discourse status of A and P, and their degree of individuation (in the sense of Hopper and Thompson 1980) are in a direct, positive relationship with the overall degree of elaboration of the event, i.e. the degree at which an event is conceptually distinguished into separate participants and sub-events. To be more precise, I will argue for the existence of an array of *situation types* which have agent defocusing as their basic component but show some crucial differences that can result in their being coded in different ways both within a single language and across languages. *Situation types* are defined, following Kemmer (1993: 7), as “sets of situational or semantic/pragmatic contexts that are systematically associated with a particular form of expression”. ‘Semantic/pragmatic contexts’ are not simply ‘real world contexts’ existing independently of the language-user, but include ‘real world’ information filtered through the conceptual apparatus of the speaker. Every language has a large inventory of lexico-grammatical devices that allow a given real-world situation to be portrayed in different ways, *under any conceivable set of discourse conditions*. The constructions examined in this paper are precisely among those lexico-grammatical devices that allow different conceptualizations of the same states of affairs: they *share the basic component of agent defocusing*, but encode different situation types, and their semantic contribution to the discourse in which they are embodied crucially depends on the way they conceptualize the event denoted by the verb.

Situation type	Features of A, P, and the event
Patient-oriented process	A is less discourse-central than P; P is highly topical; medium/high degree of elaboration of the event: the state of affairs is represented from the point of view of the patient.
Bare happening	A is de-emphasized, but corresponds to some specific individual in the world; P is not particularly topical; the event is a past, realis one, but is conceptualized as a naked fact, in summary fashion
Agentless generic event	A is generically identifiable as a subgroup of humanity (e.g., people in a given location) or represents virtually all humanity; P is not particularly topical; the event is a generic (or irrealis) one, which either did not occur, or which is presented as occurring in a non-real (contingent) world

Figure 1. *A conceptual space of agent defocusing.*

	Patient	highly topical/more discourse-central than the agent	not particularly topical
Agent			
less discourse-central than P			
de-emphasised, but specific			
generic (identifiable as a subgroup of humanity)			
generic (representing virtually all humanity)			

Figure 2. *A semantic map for passive and impersonal constructions in Italian*

	Patient	highly topical/more discourse-central than the agent	not particularly topical
Agent			
less discourse-central than P			
de-emphasised, but specific			
generic (identifiable as a subgroup of humanity)			
generic (representing virtually all humanity)			

Figure 3. *A semantic map for passive and impersonal constructions in Spanish*

	Patient	highly topical/more discourse-central than the agent	not particularly topical
Agent			
less discourse-central than P			
de-emphasised, but specific			
generic (identifiable as a subgroup of humanity)			
generic (representing virtually all humanity)			

Figure 4. *A semantic map for passive and impersonal constructions in Polish*

References

- Abraham, W. 1995. Diathesis: The middle, particularly in West-Germanic. What does reflexivization have to do with valency reduction? In: W. Abraham, T. Givón and S.A. Thompson (eds.), *Discourse grammar and typology. Papers in honor of John W.M. Verhaar*, 3-47. Amsterdam-Philadelphia: Benjamins.
- Croft, W. 2001. *Radical Construction Grammar*. Oxford: Oxford University Press.
- Haspelmath, M. 2003. The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In: M. Tomasello (ed.), *The New Psychology of Language*, vol. II, 211-242. Mahwah, NJ: Erlbaum.
- Hopper, P. J. & S. A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56: 251-299.
- Kemmer, S. 1993. *The middle voice*. Amsterdam-Philadelphia: Benjamins.
- Myhill, J. 1997. Towards a functional typology of agent defocusing. *Linguistics* 35: 799-844.
- Sansò, A. 2006. 'Agent defocusing' revisited. Passive and impersonal constructions in some European languages. In: W. Abraham, L. Leisiö (eds.), *Passivization and Typology. Form and Function*, 229-270. Amsterdam-Philadelphia: Benjamins.
- Steinbach, M. 2002. *Middle Voice. A comparative study in the syntax-semantics interface of German*. Amsterdam-Philadelphia: Benjamins.

Semantic maps, conceptual spaces, and mental representation

The typological literature generally assumes that semantic maps and the underlying conceptual spaces have mental reality, that is, they correspond to a universal arrangement of the relevant conceptual situations in a speaker's mind, based on perceived relations of similarity between these conceptual situations (Croft 2001 and 2003, Haspelmath 2003, among others). Croft (2003) goes far as arguing that the universal distributions found for particular morphosyntactic patterns (e.g. presence vs. absence of number inflection), not just the multifunctionality patterns found for individual morphemes, are the manifestation of a universal conceptual space that has mental reality.

The paper provides a number of arguments that challenge this view, based on various types of evidence from grammaticalization processes and synchronic implicational universals. In particular:

(i) There appear to be two types of connections between the conceptual situations involved in a multifunctionality pattern. Multifunctionality patterns that originate from metaphorical transfer are indeed based on some perceived similarity between the relevant conceptual situations, so these situations are arguably associated in terms of mental representation. As has been increasingly emphasized in the literature on grammaticalization (e.g. Bybee 2003, Heine 2003), however, several multifunctionality patterns originate from metonymic extensions at a construction-based or discourse-based level. In such cases, the multifunctionality pattern originates from the cooccurrence of the relevant conceptual situations in some particular construction or discourse context, not any specific similarity between these situations. There is therefore no evidence that these situations are associated in terms of mental representation.

(ii) Some multifunctionality patterns do not appear to be based on any specific connection between the relevant conceptual situations. For example, the same relative element may be used to relativize a variety of syntactic roles depending on accessibility to relativization, or, more generally, identifiability of these roles, not any connection between the roles as such. Similarly, inflectional markedness patterns, as described in Croft 2003, reveal that use of the same morpheme to encode different values of a particular inflectional parameter (such as different number values) depend on the relative frequency of the values for a cross-cutting parameter (e.g. different animacy values), not any relationship between the values encoded by the morpheme.

All this challenges the idea that cross-linguistic patterns of multifunctionality as such, as described by semantic maps and conceptual spaces, can be regarded as evidence for a specific universal arrangement of the relevant conceptual situations in a speaker's mind. These patterns reveal a number of mechanisms of form-function correspondence that are arguably valid for all speakers. This does not imply, however, that any specific association exists between the relevant conceptual situations in terms of mental representation.

Mapping Case and Agency: A Lattice-Based Approach

Scott Grimm : Department of Linguistics, Stanford University : sgrimm@stanford.edu

The typological literature has demonstrated that parameters such as agency and affectedness influence the realization of case-marking; yet, explicitly connecting individual parameters with the semantics of case-marking patterns has largely proven elusive. Here a feature-based representation of agency properties is proposed, loosely based on Dowty (1991), but reformulated in terms of privative opposition and hierarchically organized via a lattice. This approach generates a structure which can account for individual case systems as well as deliver predictions about typological generalizations. As such, this system complements the work on the semantic maps of case markers, while building upon insights accrued from work in formal and lexical semantics. For instance, one of the aims of this lattice structure is to illuminate the correspondence between the multi-functionality of a given case marker with the semantic similarity among its multiple functions.

I assume a set of event-based properties entailed by the verb referring to modes of participation in events: *instigation*, *motion*, *sentience*, *volition*, corresponding to the active ingredients of agency, and degrees of *persistence*, corresponding to affectedness. *Persistence* is a two-tiered notion, for something can persist existentially, its essence remains the same throughout the event, or it can persist qualitatively, it persists in all its particulars. Either of these can obtain at the beginning and/or the end of the event, in terms of features: *existential persistence (beginning)*, *existential persistence (end)*, *qualitative persistence (beginning)*, and *qualitative persistence (end)*. Establishing agency properties in this manner leads to two diametrically opposed classes in privative opposition: one a maximal agent possessing all the properties and the other not entailing any. Ordering these properties and their combinations by inclusion (modulo impossibilities, e.g., *volition* must co-occur with *sentience*) yields a partial order, which can be structured as a lattice. This lattice structure provides a space upon which argument structures can be mapped.

The agency features above are responsible for argument realization, i.e., which arguments are selected as subject, object, etc. Inasmuch as governed cases make reference to argument structure properties, a case-marker can be represented as ranging over one or more (connected) node(s) in the lattice. Once a region is established for the core use of a case-marker, it is then incumbent on the semantic features of that region to explain the more peripheral uses of that case. For example, that the dative, which prototypically marks the recipient, often marks experiencers can be grounded in the fact that both functions map to the same node (they are qualitatively affected and +*sentience*, but -*volition*). Similarly, the association of the instrumental case with the comitative function is expected in that the region of the instrumental (+*total persistence*, +*motion*) differs from that of the comitative by one feature (+*sentience*). This structure makes the further prediction that grammaticalization should only proceed through connected nodes.

While this framework is related to the account of case in Jakobson (1984), it is not limited to one language-particular case system, rather it shares with semantic maps that the general space of the structure corresponds to the typological space, of which any given language-particular system is one particular subspace. In sum, a comparison between the inductive method of semantic maps and the deductive method put forth here promises to be instructive concerning how the findings of formal and lexical semantics may contribute to the work on semantic maps and vice versa.

References

David Dowty. Thematic proto-roles and argument selection. *Language*, 67:547–619, 1991.

Roman Jakobson. Contribution to the general theory of case: General meanings of the russian cases. In Linda R. Waugh, editor, *Russian and Slavic Grammar: Studies 1931-1981*, pages 59–103. Mouton, 1984.

Mapping the semantics of pronominal clitics in Iranian

Abstract submitted for the Workshop on Semantic Maps, *ALT VII*

Semantic maps have proved highly applicable for representing the functions typically associated with dative-type grams. They also have considerable explanatory power as a model for understanding diachronic developments (Haspelmath, 2003). In this contribution I will combine these two aspects by mapping the functional shifts in the clitic pronouns of Iranian, which have undergone various radial extensions from a benefactive/external possessor core over a two-and-a-half thousand year time frame (Haig, Forthcoming). Functional expansion occurred through massive syncretism, as the erstwhile distinct Old Iranian Dative and Accusative clitics disappeared and their functions were absorbed by what was, etymologically, the old Iranian Genitive. The result was a single set of ‘oblique’ clitic pronouns, found in an extremely wide range of Iranian languages past and present. But unlike the dative clitics familiar from Romance and other languages, these pronouns also extended functionally to express transitive subjects in the past tenses.

Although the resulting systems exhibit rich cross-language variation, the distribution of functions can be shown to be tightly constrained, hence allowing the formulation of a number of hypotheses regarding possible paths of historical development. It will be argued that the two-dimensional approach to case functions advocated by Lehmann et al. (2004), distinguishing direct (control/affectedness) from indirect involvement, provides the most apt framework for capturing the essence of the Iranian system, and for tracing its development in a comparative perspective. The broader issues to be raised in the paper concern the efficacy of semantic maps for modelling patterns of syncretism, and the relative amount of construction-specific meaning which needs to be included in the functional categories adopted for a particular semantic map.

References

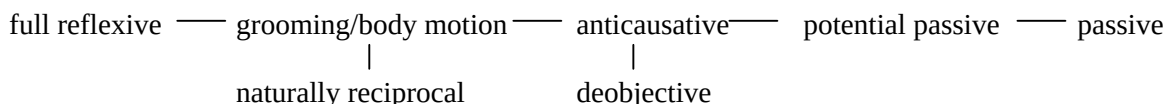
- Haig, Geoffrey. Forthcoming. *Alignment shifts in Iranian languages*. Berlin: Mouton de Gruyter.
- Haspelmath, Martin. 2003. The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In *The new psychology of language, Vol. 1*, ed. Michael Tomasello, 211–242. Mahwah, N.J.: Erlbaum.
- Lehmann, Christian, Yong-Min Shin, and Elisabeth Verhoeven. 2004. Person prominence and relational prominence. On the typology of syntactic rela-

tions with particular reference to Yucatec Maya [Second revised version].
Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt
No. 12, available online from the Institute's website.

Semantic maps and implicational universals in diachronic studies: the case of reflexives

Nicoletta Puddu
University of Cagliari

I will analyse the diachronic development of the reflexive marker (henceforth RM) in ancient Indo-European languages using the semantic map proposed by Haspelmath (2003). He suggests to represent the polifunctionality of RMs by using the following map:



According to Haspelmath (2003) semantic maps are also particularly useful in diachronic studies, since they show a clear directionality in semantic change. We would expect then that the more a RM is grammaticalized, the more it extends to the right in the semantic map.

Ancient Indo-European languages do not show a common reflexive strategy. Rather, it seems that the creation of a dedicated reflexive marker was a later development in Indo-European languages (see Puddu 2005, *forthcoming*). Eastern IE languages used a nominal strategy as a RM, which still had a clear referential meaning. Vedic, for instance, used the word for ‘body’ *tanū-*. In these languages, as we would expect, the RM is used only as a full reflexive.

In Ancient Greek the RM *hè autón* was already fully grammaticalized, and it extended only to grooming verbs (e.g. Hom. Il. 14, 162 *entunasan hè autén* ‘adorning herself’). Also in Latin the RM *se* was restricted to grooming or body motion (Pl. *Am.* 273 *se commovent in caelo* ‘they move in the sky’), while in Gothic it was extended to anticausative uses (e.g. *ushaffan* ‘raise’ vs. *ushaffan sik* ‘rise’). In all these languages the RM could also be used with reference to the subject of the main clause. However, while in Archaic Latin and in Ancient Greek the dependent clause could be in the indicative, subjunctive, infinitival or participial form, in Gothic it could only be in the infinitival or participial form. Huang (2000) proposed the following universal for long distance reflexivization at the sentence level:

NPs > small clauses > infinitivals > subjunctives > indicatives

In Puddu *forthcoming* I have used Huang’s (2000) hierarchy to demonstrate the original anaphoric value of **se-*. Here, I will argue that this hierarchy is “complementary” to the semantic map proposed by Haspelmath (2003). On the basis of a corpus study, I will argue that the extension of the RM in the middle domain is linked to the contemporary restriction of its uses outside the clause. In other words, I will show that the polifunctionality of the RM in IE language is:

- directly proportional to its grammaticalization;
- inversely proportional to the possibility of its use with reference outside of the sentence.

References

Haspelmath, Martin (2003) ‘The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison’. In *The New Psychology of Language. Cognitive and Functional Approaches to Language Structure*, vol. 2, Michael Tomasello (ed.), 211–242. Mahwah/London: Lawrence Erlbaum Associates.

Huang, Yan (2000) *Anaphora. A Cross-linguistic Approach*. Oxford: Oxford University Press.

Puddu, Nicoletta (2005) *Riflessivi e intensificatori: greco, latino e le altre lingue indoeuropee*. Pisa: ETS

Puddu, Nicoletta (*forthcoming*) ‘Typology and Historical Linguistics: Some remarks on reflexives in ancient Indo-European languages’. In Miestamo, M. e Wälchli, B. (eds.) *New Trends in Typology: Young Typologists’ Contributions to Linguistic Theory*. Berlin/New York: Mouton de Gruyter.

**Semantic maps and word formation:
Agents, Instruments and related semantic functions**

Eugenio R. Luján

Depto. de Filología Griega y Lingüística Indoeuropea
Universidad Complutense de Madrid

[\(erlujan@filol.ucm.es\)](mailto:erlujan@filol.ucm.es)

The semantic map methodology has been applied mainly to the analysis of grammatical morphemes (affixes and adpositions); see, e.g., Haspelmath 1999 for 'Dative' or Haspelmath 2003: 226-229 for Instrumental and related semantic roles. Although this methodology is still in need of some refinement, it has already proved as a very useful tool for the study of the type of structured polysemy that is characteristic of grammatical morphemes. Semantic maps have also proved to be extremely useful for the analysis of grammaticalization paths between synchronically linked semantic functions.

The semantic map methodology can be further extended to the analysis of procedures of word formation, including both derivational and compositional procedures. I have been working lately on the use of semantic maps for the analysis of Agent and Instrument nouns and the grammaticalization processes associated with them. This poses some interesting theoretical and practical problems. For instance, it has been stressed (e.g., Haspelmath 1997: 10-13) how difficult it is to isolate semantic functions when dealing with grammatical morphemes cross-linguistically, if the possible proliferation of functions identified only on semantic criteria is to be avoided. That difficulty will increase when dealing with affixes and compounds, given that we cannot apply the standard syntactic procedures used to isolate semantic functions in functional-typological linguistics. We are thus driven to explore the bases on which different semantic functions can be isolated when dealing with procedures of word formation.

Semantic maps based on the analysis of procedures of word formation also allow for interesting comparisons with semantic maps elaborated on the basis of grammatical morphemes. Some general remarks can be made. First, languages tend to grammaticalize a much lesser number of procedures of word formation than grammatical morphemes for the expression of semantic functions – e.g., procedures of word formation of Agents and Instruments are frequently found but languages do not usually have procedures of word formation for Beneficiaries. Second, linguistic categorization (and its reflection in semantic maps) may be different when dealing with grammatical morphemes and procedures of word formation – semantic functions that are directly linked to each other in semantic maps based on grammatical morphemes may or may not be linked in semantic maps based on procedures of word formation, and the other way round. Third, even if certain grammatical functions are organized in the same way in semantic maps based on grammatical morphemes and procedures of word formation and, accordingly, meanings are extended following the linking lines between them (Croft – Shyldkrot – Kemmer 1987, Haspelmath 2003: 233-237), grammaticalization processes may follow opposite directions – e.g., Instrument markers frequently evolve into Agent markers while it is Agent nouns that usually evolve into Instrument nouns.

REFERENCES

- Croft, W. – H. B.-Z. Shyldkrot – S. Kemmer 1987: "Diachronic semantic processes in the middle voice", in: A. Giacalone Ramat – O. Carruba – G. Bernini (eds.), *Papers from the 7th International Conference on Historical Linguistics*, Amsterdam – Philadelphia, John Benjamins, 179-192.
- Haspelmath, M. 1997: *From Space to Time: temporal adverbials in the world's languages*, Munich, Lincom.
- 1999: Haspelmath, M. 1999: "External possession in a European areal perspective", in: D. Payne – I. Barshi (eds.), *External Possession*, Amsterdam - Philadelphia, John Benjamins, 109-135.
- 2003: "The geometry of grammatical meanings: semantic maps and cross-linguistic comparison", in: M. Tomasello (ed.), *The New Psychology of Language*, Mahwah, Erlbaum, 211-242.