The semantics of lexical typology

Maria Koptjevskaja-Tamm, Ekaterina Rakhilina, Martine Vanhove

The semantics of lexical typology
Maria Koptjevskaja-Tamm, Ekaterina Rakhilina, and Martine Vanhove

1. INTRODUCTION

Generally speaking, lexical typology is the systematic study of cross-linguistic variation in words and vocabularies (cf. Koptjevskaja-Tamm, 2008). Opinions differ, however, on what exactly can be meant by this relatively recent term. Many linguists will probably agree with Lehrer's (1992: 249) widely quoted definition that lexical typology is concerned with the "characteristic ways in which language […] packages semantic material into words" (cf. Koch, 2001; Brown, 2001). It is lexical typology in this understanding that will be our main interest here. Until recently lexical typology was both limited and marginal compared to the cross-linguistic research on grammar and on phonetics / phonology, apart from a few classical studies (e.g., Andersen, 1978; Berlin and Kay, 1969; Greenberg, 1980; Talmy, 1985; Viberg, 1984). The past two decades have, however, seen an increased interest in lexical-typological issues, both in the number and diversity of relevant studies and publications stemming from them, as well as in several publications suggesting various definitions of lexical typology, situating it within a broader research field and providing overviews over the field (e.g., Behrens and Sasse, 1997; Brown, 2001; Evans, 2011; Goddard, 2001; Koch, 2001; Koptjevskaja-Tamm, 2008; Lehmann, 1990).

The core concern of lexical typology, i.e. how languages express meanings by words, can be approached from slightly different perspectives. We can start from the *meanings,* or *concepts*, and ask how these are expressed in different languages – among other things, how semantic domains are distributed among the lexical items across languages. A simple example of this is given in Table 1, which shows how the inventories of body-part terms in six languages differ in the extent to which they distinguish between hand vs. arm, foot vs. leg, and finger vs. toe by conventionalised, lexicalized expressions.

Table 1: *Hand vs. arm, foot vs. leg, finger vs. toe in English, Italian, Rumanian, Estonian, Japanese and Khalkha Mongolian.*

English	Turkish	Rumanian	Estonian	Japanese	Khalkha Mongolian
hand	*el*	*mină*	*käsi*	*te*	*gar*
arm	*kol*	*brat,*	*käsi(vars)*	*ude*	
foot	*ayak*	*picior*	*jalg*	*ashi*	*höl*
leg	*bacak*				
finger	*parmak*	*deget*	*sõrm*	*yubi*	*huruu*
toe			*varvas*		

Lexical-typological research in this understanding typically asks questions such as how languages categorize particular domains (human body, kinship relations, colour, motion, perception, etc.) by means of lexical items, what parameters underlie categorization, whether languages are completely free to "carve up" the domains at an infinite and arbitrary number of places or whether there are limits on this, and whether any categories are universal (e.g., say 'relative', 'body', or 'red'). For many people lexical typology par excellence would primarily be associated with these kinds of issues, but lexical typology is broader than that, as will be shown immediately below.

Lexico-typological research can also start from the *expressions (lexemes)* and ask what different meanings can be expressed by them or by lexemes that are related to them synchronically and/or diachronically. The main focus here is on cross-linguistically recurrent patterns in the relations among the words and lexical items in the lexicon, e.g., semantic motivation (polysemy, semantic

1

associations / semantic shifts) and morphological motivation (derivational patterns, including compounding) (see Newman, this volume).

For some researchers, the main interest of lexical typology is different again. For instance, Lehmann (1990: 163) defines lexical typology as research which focuses on "typologically relevant features in the grammatical structure of the lexicon", rather than on "the semantics of individual lexical items, their configurations in lexical field or individual processes of word formation" (Lehmann (1990: 165). This view of lexical typology is also promoted in Behrens and Sasse (1997), in Nichols et al. (2004) and in the rich research on motion verbs stemming from Talmy's seminal chapter (1985).

In this chapter we will mainly focus on lexical typology understood as cross-linguistic research on domain categorization. We start by introducing some of the critical issues inherent in such research (Section 2), and then turn to the presentation and discussion of four different approaches to lexical typology in Sections 3–6. Section 7 is devoted to semantic maps as representations of meanings and generalizations in lexical typology, while Section 8 suggests a few future directions for lexical typology.

2. CRITICAL ISSUES

Lexical typology has to find its own way for balancing the methodological and theoretical ambitions of theoretical semantics, lexicography and general typology.

Typological research takes linguistic diversity as its point of departure, assumes that the variation across languages is restricted and aims at discovering systematicity behind it. Its central theoretical issues include the following: What are the parameters / dimensions along which a specific phenomenon varies across languages? In what patterns do these parameters (co-)occur? What generalisations can be made about attested vs. possible patterns? What is universal vs. language-particular in a given phenomenon? While cross-linguistic generalizations constitute a major result, many typological studies want to go further and find explanations for them. Possible explanations for the typological patterns and the generalizations over them may be environmental (rooted in the properties of the real world), biological (shaped by human perceptual and cognitive predispositions or simply innate), socio-historical or cultural. But also the distribution of the patterns across the languages, i.e. why language X has a certain pattern, calls for explanation. These are big and exciting questions, which all presuppose meticulous work, methodological awareness and, ideally, firm theoretical foundations, to which we now turn.

Typological research in general is dependent on comparable data coming from (many) different languages. Cross-linguistic identification of studied phenomena presupposes a procedure which ensures that we compare like with like. For data collection and cross-linguistic identification of phenomena, grammatical typology has historically been largely dependent on secondary data sources (such as reference grammars), with first-hand data sources gradually gaining more and more importance. However, for lexical typology first-hand sources of data are crucial, since the lexicon for most languages of the world is relatively poorly described. Although some lexico-typological studies use dictionaries as their main source (e.g., Andersen, 1978; Brown, 2005a; 2005b), most of the cross-linguistic research on the lexicon is based on elicited data – either by extra-linguistic stimuli (cf. Sections 5, 7) or by means of questionnaires, ranging from simple translational questionnaires to much more sophisticated 'frame-based' questionnaires, which elicit verbal descriptions of various situations (cf. Section 6). These are frequently complemented by corpus studies, with comparison of parallel texts (translations of one and the same text) as a relatively new but promising method for data collection (cf. Section 7). One particularly prominent feature of the lexicon as compared to grammar is its vastness and diversity, which creates additional challenges for both data collection and analysis (compare such culturally specific and subjective notions as emotions with the much more "visible" events of cutting and breaking). A lexico-typological study is typically restricted to one cognitive domain, often (although not necessarily)

with additional constraints on the part of speech of the studied words (e.g., verbs of perception); it typically requires a serious involvement of language experts, capable of discerning and discussing intricate meaning nuances, and usually operates with a much more limited language sample (normally between 10 and 50 languages) than what is the norm in grammatical and phonetic typology. There are notable exceptions here, such as Brown (2005a, 2005b) on body-part terminologies, Kay and Maffi (2005) on colour terms, Nerlove and Romney (1967) on sibling terminologies, and Wälchli and Cysouw (2011) on motion verbs, which all include between 100 and 245 languages.

A further issue is how the data may be analyzed and how the results of the analysis may be represented. The problem of a consistent meta-language for describing meaning, and in particular lexical meaning, cannot be overestimated. This is, in turn, related to the general enormous gap between theoretical semantics and theoretical lexicology, on the one hand, and actual lexicographic practices. Two of the major stumbling blocks are the problem of what can be meant by meaning – *denotation / extension* vs. *sense / descriptive meaning / intension* – and the problem of *polysemy / semantic generality / vagueness*. Although for many serious semanticists, lexicographers and lexicologists semantic analysis stands for understanding descriptive meanings, or senses, the enterprise gets easily insurmountable, especially when several languages are involved. As a consequence, much of cross-linguistic comparison is based on meanings defined as the denotational ranges of the expressions involved (cf. Sections 5, 7). Similarly, there are many different opinions on what counts as polysemy (see Riemer, 2005, for a recent overview of the problem), but distinguishing between several meanings of one and the same lexeme or one more general meaning typically requires sophisticated analyses and tests that are too difficult to carry out in cross-linguistic studies. As a rule of thumb, polysemy is acknowledged whenever a word may be used to denote entities, properties or situations that are assumed to belong to very different cognitive domains (e.g., 'foot of a person' vs. 'foot of a mountain', or 'to see a picture' vs. 'to see what somebody means'). Things become much more complicated when the two uses appear to belong to the same domain. Are the Khalkha Mongolian *gar* and the Russian *ruka* simply vague in not distinguishing between 'hand' and 'arm', or are they polysemous between the two meanings? How can we describe the fact that Russians can 'hear' smells, i.e. that the verb *slyšat'* 'hear' can combine with 'smell' for talking about smell perception (*slyšat zapax* 'feel (lit. hear) a smell')? Different researchers choose different strategies for dealing with such cases, stretching between explicit agnosticism (as François's 2008 notion of 'colexification' or Wälchli and Cysouw's 2012 view on situational contextually embedded meanings as analytic primitives, cf. Section 7) to pragmatic and partly ad hoc solutions to insistence on a strict demarcation between semantic generality and polysemy (as typical for Natural Semantic Metalanguage, cf. Section 4).

A further question is whether meanings should be analyzed and represented in terms of classical Aristotelian categories with necessary and sufficient meaning components or in terms of prototype categories (see Hampton, this volume). Related to this is the issue of compositionality, i.e., whether meanings can be decomposed and if so, to what extent, what the basic building blocks (e.g., whether there are universal or language-specific concepts, or a mixture of both, cf. Sections 3 and 4) are, and how the "composition" is carried out and represented (see Riemer, this volume). Logical formulas, explications (Section 4), semantic maps and diagrams (Section 7), pictures, tables (Section 3) have all been used for representing meaning, with relatively little communication among the different schools of semantics and lexicology and lexicographic practices.

3. COMPONENTIAL ANALYSIS

A once-influential trend in semantics, componential analysis was prevalent in the 1950–70s as a part of the structuralist paradigm. It was essentially an attempt to apply to the lexicon the concept of oppositions and binary features with roots in phonological and morphological analysis (see Newman, Riemer, chap. 14, in the present volume).

The main principle is to define each term in a domain by a set of distinctive semantic features,

e.g., *horse* [+adult, ±female] vs. *colt* [-adult, -female] vs. *mare* [+adult, +female] etc, and ultimately to use such oppositions to define the whole meaning of a term, e.g., *horse* [+animate, +quadruped, +mammal, +equine, +adult, ±female...].

The chief contribution of componential analysis to lexical typology concerns kinship terminologies, which have for a long time been a favourite semantic domain among anthropologists and anthropologically oriented linguists. Detailed and systematic descriptions of the domain are normally structured by a few relevant parameters, are available for many hundreds of languages, and there is a long tradition of classifying the resulting systems into a small number of types. As a rule, such classifications concentrate only on subparts of the kinship systems: Nerlove and Romney's (1967) study, for example, probably the most cross-linguistically systematic, focuses on sibling (i.e., children of the same biological parents) terminologies in 245 languages. These are based on eight logical KIN types as defined by three parameters (sex of ego, sex of relative, relative age) – e.g., whether one and the same term is used for all siblings, whether there are two separate terms for 'brother' and 'sister', whether there are four different terms ('younger brother', 'elder brother', 'younger sister', 'elder sister'). A simple example of this is given in Table 1, which shows the inventories of sibling terms in six languages – Russian (Indo-European: Slavic), Palula (Indo-European: Dardic), Turkish (Altaic: Turkic), Ewe (Niger-Kongo: Kwa), Jakarta Indonesian (Austronesian: Malayic), and Nukuoro (Austronesian: Polynesian).

Table 2: Sibling terms in six languages

Language	male sibling	female sibling	relative age
RUSSIAN	*brat*	*sestra*	age difference irrelevant
PALULA	*bhróo*	*bheén̠*	age difference irrelevant
	káaku	*kéeki*	elder
	kúuču	*kúuči*	younger
TURKISH	*aga*	*aba*	elder
	kardaş		younger
EWE	*efo*	*eda*	elder
	etse	*eʄoe*	younger
	nɔvi		age difference irrelevant
JAKARTA INDONESIAN	*abang*		elder
	kakak		
	adik		younger
	saudara		age difference irrelevant
NUKUORO	*teina*		age difference irrelevant

Restricting ourselves to sibling terms that are not hyperonyms of any others (e.g., excluding 'sibling' in languages distinguishing 'brother' and 'sister'), there are 4140 logically possible types, but only 12 of those are attested in more than one language from the 245 languages. The absolutely most dominant type (in its pure version attested in 78 languages) has four terms – elder brother, elder sister, younger brother, younger sister (cf. Palula and Ewe in Table 2); the next most frequent type (in its pure version attested in 38 languages) has three terms – elder brother, elder sister, and younger brother/sister (see Turkish in Table 2). The two next types occur in 21 languages each (i.e., in less than 10% of the sample) and include the average European type with a distinction between brother and sister, and another two-term type, with a distinction between elder and younger siblings (with Indonesian in Table 2 approaching it).

Importantly, not only do Nerlove and Romney present their classification and statistics, but they also suggest that the paucity of the attested systems may be explained by universal restrictions, such

as the cognitive restriction on avoidance of disjunctive categories <mark>(Bruner et al., 1956: 41-43).</mark> For instance, a category comprising male younger and elder siblings, i.e. 'brother', is conjunctive, since it contains the same component [male relative] combined with the two values of the component [relative age], as opposed to the disjunctive category comprising, say, male siblings and siblings of a male. Another universal constraint, this time of a more social character, is avoidance of categories in which sex of ego is a primary distinction. Nerlove and Romney go even further and try to link cross-linguistic variation to social variables, for example connecting the existence of sibling categories based on the relative sex of ego and relative in a language with brother-sister avoidance in the group speaking it.

Componential analysis of different denominations has been subject to massive criticism from various quarters, primarily for its categorical view of meaning (cf. Riemer, chap. 14, this volume). Its main appeal for lexical typology lies, however, in the promise of bringing out the structure behind a semantic domain and in this way facilitating systematic cross-linguistic comparison. Consider the above mentioned kinship asymmetry. It is an example of *linguistic universals*, i.e. generalizations on what is generally preferred/ dispreferred (or even possible / impossible) in human languages. Componential analysis, with its formalized and simplified meta-language, offers a seemingly smooth method for discovering and formulating generalizations based on combinations of co-occurring features. This is true in particular if the features are "easily" recognizable, not only within one language but also across languages, and if their values are clearly distinct from each other. However, even for these purposes componential analysis is of limited value, in that it may impose "artificial" structuring on a domain and obscure some of the important generalizations on the factors behind its categorization, as will be discussed below.

Apart from kinship terminology, there have hardly been any serious attempts to apply full-fledged componential analysis to cross-linguistic research on the lexicon whereby all the words in a domain will be decomposed into matrices of (binary) features, ideally independent of each other and relevant for the whole domain (a possible exception, although on a much smaller scale, is spatial dimension terms, cf. Lang 2001). Very few domains are in fact organized as paradigms. In practice, many features are dependent on each other, with many of them only applicable to a subset of the terms, often leading to taxonomies of various kinds. Already Lehrer in her comparison of cooking terms across several languages, which she couches in terms of components such as 'the use of water: [±water]', 'source of heat: radiated vs. conducted vs. hot surface', cooking time: long vs. short', 'cooking action: gentle vs. vigorous, etc., wrote the following:

> "A cursory examination of cooking words shows why many components are not combined into a single word: when cooking with water, it does not matter what the heat source is so long as it is sufficient to keep the water at a high enough temperature; when something cooked in liquid is cooked a long time, a gentle action is preferred so the liquid does not boil out" (Lehrer, 1974: 101).

The moral here is that the semantic properties of kinship terms, of cooking verbs and in fact of many (or perhaps even most?) other lexical systems are not randomly combined with each other, but form meaningful clusters that often lack symmetry. Even pairs of words that are generally considered to be antonyms and therefore only differ in the value of one component are normally asymmetrical – they differ in the scope of their meaning and in patterns of polysemy, so that one and the same word often has different antonyms in its different uses (cf. Paradis and Willners, 2011; see Storjohann, this volume for further details).

Lexical systems, with the lexical meanings and oppositions forming them, emerge from human experience of recurrent situations, including interaction with other people and various objects, and in this sense their asymmetries are inherent. Domains are often structured around several distinct prototypical situations or prototypical entities, and it may be more reasonable to take those as the main unit of description/analysis in a lexico-typological study rather than decomposing them into their particular distinctive features, the status of which is often quite questionable. Coming back to kinship terminologies, Wierzbicka (1987) points out the hidden circularity in defining 'mother' and 'father' via the component 'ascending generation', which can itself only be understood via

parenthood. 'Mother', as 'birth-giver', is, on the other hand, a highly important and universally lexicalized concept. Therefore, as Wierzbicka argues, it may be more fruitful to use the concept of 'mother' in the semantic analysis of kinship than to resort to such abstract components as 'ascending generation' and 'linearity'.

4. NATURAL SEMANTIC METALANGUAGE

Anna Wierzbicka, cited immediately above, together with Cliff Goddard are the portal figures for *Natural Semantic Metalanguage* (*NSM*), which shares the idea of meaning decomposition with componential analysis, but has a different philosophical foundation and embodies a completely different ideology. This approach is consistently anthropocentric, i.e. taking man as the measure of all things (cf. the notions of embodiment and perspectivization in Cognitive Linguistics). For instance, the seemingly random clusters of features that make up lexical meanings are described and explained as stemming from the functions of their referents in human life. Thus, the definition of 'cup' does not only mention a 'handle' as one of its components (as might be done in componential analysis), but also explicitly describes 'cups' as meant for hot liquids "to be drunk from while sitting at a table, lifting them to the mouth to drink a little and putting them down again" (Wierzbicka, 1984: 215). Definitions of colour terms explicitly refer to prototypical natural objects such as the sun for 'yellow' or living vegetation for 'green' (Wierzbicka, 2005), while temperature terms are described via fire (Goddard and Wierzbicka, 2007).

At the heart of this method lies the hypothesis that there is a small lexical-conceptual core shared by all languages, 'semantic primes', i.e., basic and universal meanings, which can be expressed by words or other linguistic expressions in all languages. The current list contains 64 items. Due to their universal "translatability", these can further be used as the *tertium comparationis* in the systematic investigation of lexical phenomena across languages, as the main building blocks in reductive paraphrases that aim at providing precise definitions, or "explications", of their meanings (e.g., Goddard and Wierzbicka 1994, Goddard, 2001; Wierzbicka, 2007).

According to the NSM homepage, NSM "can lay claim to being the most well-developed, comprehensive and practical approach to cross-linguistic and cross-cultural semantics on the contemporary scene" (http://www.une.edu.au/bcss/linguistics/nsm/), with a track record of thirty five years in empirical research and a wealth of publications. NSM has been applied to such traditionally central domains for lexical typology as kinship, colour and body (Wierzbicka, 1987; 2005; 2007). However the area where this method has led to particularly interesting results is cross-linguistic semantic comparison of complex, culturally relevant emotional and mental phenomena, where cross-linguistic diversity often arises from cultural stereotypes that profile different aspects of reality and speaker's attitudes (Wierzbicka, 1999; Harkins and Wierzbicka, 2001). A very fruitful idea for semantic comparison is to use a common structural pattern for words semantically related to each other (a "semantic template"). Similar ambitions are, of course, present in other semantic approaches, at least for the words belonging to one and the same language. However, NSM remains unrivalled in the extent to which this has been employed for cross-linguistic comparison, where semantic templates offer effective tools for analyzing, comparing and representing meanings across languages, in a way that enables capturing both similarities and fine-grained distinctions among them. To take an example, one of the semantic templates for emotion words builds on the idea that being happy, sad, pleased etc. means being in a particular state of mind (for instance, thinking that something good has happened) and because of that experiencing a particular kind of feeling (e.g., feeling good), which, in turn, is viewed as normally linked to this mental state. By a consistent application of such patterns, NSM researchers claim to demonstrate how words for seemingly similar emotions (say, 'anger' or 'happiness') differ across languages and in pinpointing what makes certain emotion terms (e.g., *amae* in Japanese) extremely culture-specific and seemingly untranslatable into other languages (see the contributions in Harkins and Wierzbicka, 2001).

Much of the recent work within NSM has been directed towards elaborating and testing semantic templates and other analytical constructs (primarily semantic molecules) for further semantic

domains across languages, such as various subclasses of physical actitivy verbs (cf. Goddard, 2012).

The theory has both positive and negative sides and its strong basic assumption on the existence of universally lexicalized semantic primes and on their manifestations in concrete languages is particularly debatable (cf. e.g., the discussion in Krifka (ed.) 2003, Riemer, 2006; Evans, 2011). To give just one example, it is far from clear whether any of the two Russian near-synonyms *xotet'* and *želat'* is a good candidate for the prime WANT, because both of them are influenced by being opposed to each other and develop connotations beyond the meaning of pure intent that the prime should signify (example from Apresjan, 2000).

5. DENOTATION-BASED, OR ETIC-GRID SEMANTICS

The denotation-based approach to cross-linguistic semantics was first launched in the second half of the 1960s by Brent Berlin and Paul Kay, who equipped their students with Munsell colour charts and sent them to see how these were mapped to the colour terms in the different languages spoken in the San Francisco Bay area. This was the birth of the groundbreaking systematic cross-linguistic research on colour often called "the Berlin-Kay paradigm". The Munsell colour charts provide a number of extra-linguistic contexts, or 'an etic grid' for capturing possible distinctions within the colour domain (organized according to the three dimensions of hue, value/lightness, and chroma/color purity), which enables comparison of denotational ranges of colour terms both within one and the same language and across languages. The various tasks that informants perform with the colour chips allow comparison with further cognitive findings. In the Berlin-Kay view, although languages are very different as to how many colour names they distinguish in their lexicon, as well as to exactly which portions of the colour spectrum each of them covers, this variation is severely restricted, primarily by the neurophysiology of vision (Berlin and Kay, 1969; Kay and McDaniel, 1978; Kay and Maffi, 2000). Colour remains the cross-linguistically most widely studied lexical domain in terms of the languages systematically covered by means of comparable and elaborated methodology (Kay and Maffi, 2005, lists 119 languages), and the intensity, diversity and depth of theoretical discussions (cf. e.g., Robertson and Hanley, 2010 for a recent overview; MacLaury et al., 2007, and the references there; and http://www.icsi.berkeley.edu/wcs/ for the World Color Survey Site).

In addition to colour, the denotation-based approach to cross-linguistic comparison has been actively promoted in the rich research at the Department of Language and Cognition at the Max-Planck Institute for Psycholinguistics in Nijmegen. The major results of this research have so far been achieved for 'body' (Majid et al., 2006), 'cut / break' (Majid and Bowerman, 2007), 'put / take' (Narasimhan and Kopecka, 2012), 'location' (Ameka and Levinson, 2007) and 'space' in general (Levinson and Wilkins, 2006) (http://www.mpi.nl/research/research-projects/categories). The "Nijmegen method" uses standardized stimuli, such as sets of pictures, videoclips and films for collecting data on a number of cognitive domains directly in the field (cf. http://fieldmanuals.mpi.nl/). Each set covers a shared denotational grid allowing systematic comparisons of semantic distinctions potentially relevant for the domain and may be used under different elicitation conditions, including games. To give a couple of examples, the data for the 24 languages in the 'cut / break' domain and for 19 languages in the 'put / take' project were collected by means of videoclips, such as 'X is tearing cloth into two pieces by hand', 'X is slicing a carrot across into multiple pieces with a knife', 'X is cutting hair with scissors', 'X is smashing a plate with a single blow of hammer', etc. (Majid et al., 2007), or 'X is putting a cup on the table / taking a cup off the table', 'X is pouring liquid into a container / pouring water out of a tin can', 'X is putting on a coat / taking off a coat' (Narasimhan et al., 2012). The data on the 'body' for the ten languages in Majid et al. (2006) involve, among others, drawing outlines of the various body-part terms on a picture of a human body.

A huge appeal of this approach is its objectivity. With a standard set of stimuli, the researcher's influence on the results is minimal, and it is easy to compare data from different speakers and

different languages, no matter when or where it has been gathered. This makes the method particularly suited for studies of underdescribed languages.

However, it is not free from problems. Some domains do not lend themselves to representations by stimuli sets – complex mental and social phenomena such as emotions, unpleasant physical sensations such as pain, or abstract notions, such as possession, are obvious examples. These require other methods of data collection and analysis, for instance, verbal descriptions. A deeper problem lies within the stimuli themselves: the set needs to be complete and cover the domain exhaustively, since anything not featured in the experiment is likely to be overlooked. However, there may also be various interesting and unexpected sides of the phenomena that the researcher has not been aware of while preparing the stimuli and that may turn out to be important for their understanding – these may be disclosed by non-elicited data. Also the fact that the stimuli are artificial and largely divorced from the context, may affect the results, as debated by people such as Levinson (2000), and Wierzbicka (2005), who have questioned the validity of the decontextualized denotation-based methodology underlying a lion's share of the studies in the Berlin-Kay paradigm.

For example, colour terms are often chosen depending on the type of the object. *Red* as a hair colour is very different from the prototypical red and is, in fact, close to orange, but calling hair *orange* strongly implies that it is dyed. Colours related to humans and animals tend to have very specific terms, such as *blond* (only human hair) or *bay* (only horses). When new colour terms emerge, often inspired by new objects, old terms shift their meanings to accommodate the change and gradually fall out of use. For instance, the old word for brown in Russian, *buryj*, has been almost completely displaced by *koričnevyj*, which is derived from the word for the exotic spice cinnamon. Among the last objects for which *buryj* is still used are dirt and animal fur. The rule behind this process is that new and exotic terms are used to describe artificial colours, while old terms stick with natural objects. This means that by disregarding the names of natural colours, with their limited scope and declining usage, the researcher might miss what once used to be the core colour vocabulary of the language and will not be able to fully understand its evolution, which has been one of the major goals in the Berlin and Kay enterprise.

To be fair to the denotation-based approach, the degree of decontextualization varies significantly between different techniques and studies. Both retelling a film for someone who has not seen it or exchanging verbal instructions during a game come closer to language in normal use than describing a series of disconnected videoclips or naming colour chips. And in fact, most of the data acquired by stimuli-based elicitation in the research at the Max-Planck institute are complemented by data coming from other sources (e.g., texts in the relevant languages). A successful study in lexical typology obviously benefits from a combination of different types of data.

6. COMBINATORIAL LEXICAL TYPOLOGY

Combinatorial lexical typology, primarily developed by the Moscow Lexico-Typological group (directed by Ekaterina Rakhilina), is the most recent contribution to cross-linguistic semantic comparison in lexicon (e.g., Rakhilina, 2010; Reznikova et al., 2012; Koptjevskaja-Tamm et al., 2010). This method can be traced back to the early claim of the Moscow school of semantics (with Apresjan as the main figure, e.g., Apresjan, 2000) that even the closest synonyms have different collocations, and that every word should be described with particular attention to its surroundings. The idea, also shared by Wierzbicka and Firth and strongly argued for by cognitive linguistics, assumes that the combinability of a lexeme in different contexts, its "linguistic behaviour" (Wierzbicka), or "semantic range" (Firth, 1957) is normally motivated by its semantic properties and can provide cues for the understanding of its meaning, where the meaning of a word is its intension, rather than denotation, or extension. The semantic properties of every single lexeme are therefore to be studied individually, by a systematic analysis of its combinability. A semantic description of a lexeme includes both a description of its linguistic behaviour and a suggestion as to what semantic properties might motivate it. In Russia, this idea gave rise to a rich lexicographic

8

tradition (e.g., NOS, 2004) whereas the development of corpus linguistics has enabled its serious application to lexical typology. Corpora, available for an increasing number of languages, give a relatively full and reliable picture of words in context, and, most importantly, of the contexts preferred by a given lexeme.

A key notion in the combinatorial approach is a "frame", which is understood somewhat differently than in the mainstream cognitive linguistics (e.g., Fillmore 1982; FrameNet https://framenet.icsi.berkeley.edu/fndrupal/home). Whereas for Fillmore a frame is defined as a set of participants with their assigned syntactic roles, the combinatorial lexical typology views a frame as a group of prototypical situations (and contexts that represent these situations in texts) which have functional similarity and are relevant for lexicalization. Similarly to NSM, this approach is anthropocentric and assumes that words partition semantic domains by breaking up reality into cognitively relevant fragments, rather than randomly. In particular, the prototypical situations are so closely related to the predicate's argument types that a description of its arguments can serve as a sufficient representation of a situation. For example, 'oldness' is a small domain that consists of four main frames that can roughly be described as 'worn' (*old shoes, rags...*), 'aged' (*old people, horses, trees...*), 'former' (*old address, government, procedure...*) and 'ancient' (*old myths, music, grammatical construction...*) (Taylor, 1992). The English adjective *old* covers all of them, even though it can have specific antonyms corresponding to some of the frames – *young*, *current* and *modern* respectively. Other languages divide the domain between two or more lexemes, and the richest systems (found in Nanai, Yakut, Komi, Kurdish, Malagasy, Quechua and Ossetic) have separate adjectives for all four situations.

Combinatorial lexical typology is similar to the denotation-based approach in working with 'etic grids' of contexts, enabling systematic comparison of how languages group these contexts within their own concepts. However, whereas in the denotation-based approach the grid consists of extra-linguistic stimuli, the lexical-combinatorial method uses a list of frames, or rather a list of diagnostic contexts that represent them. For instance, for 'old' it would be 'shoes', 'people' etc, for the verbs of rotation such contexts as 'the Earth around the Sun', 'barrel on the ground' and 'hawk over its prey' and some others (Krugljakova, 2010), for the verbs of oscillation – 'trees in the wind', 'a broken chair', 'a long skirt when walking' etc. Searches for the key words and their collocations are made in dictionaries and corpora, and the list is also used as a questionnaire for native speakers. The core of the method consists in finding contexts in which a word cannot be replaced by a particular near-synonym, and determining which of their properties determine such restrictions.

In this approach each domain presents a landscape of frames that languages carve up by their lexemes. Lexemes can both overlap in their ranges and leave parts of the landscape unlabelled. Frames are believed to constitute the analytic primitives in a lexico-typological comparison. They have been extensively discussed for several domains such as aquamotion, swaying and shaking, rotation, sharpness and bluntness, pain, etc. (for details see Rakhilina and Reznikova, 2014).

7. SEMANTIC MAPS

As noted at the end of the preceding section, much of the cross-linguistic research on the lexicon (e.g., both combinatorial lexical typology and denotation-based semantic typology) departs from 'etic grids' of contexts and asks the question of how languages group these contexts within their own concepts. In sections 3 and 4 we showed how two different decompositional approaches to semantics (componential analysis and Natural Semantic Metalanguage) formulate the results of cross-linguistic semantic comparison in terms of their specific metalanguages. This section deals with a family of representational techniques that are particularly suitable for 'etic-based' semantic comparison, namely semantic maps.

A semantic map is a tool for visualizing (representing) crosslinguistic commonalities in the multifunctionality patterns manifested by semantically / functionally "comparable" linguistic expressions (e.g., morphemes, words, constructions) of particular languages. The idea is that one and the same linguistic expression normally has a range of functions (uses, meanings, contexts), and

these ranges usually show considerable overlaps across languages, without necessarily being completely identical. The main guiding principle for constructing a semantic map may be called the "contiguity / connectivity requirement". I.e., functions (uses, meanings, contexts) that are often associated with one and the same linguistic expression are represented as nodes adjacent to each other, or as a contiguous region in a semantic map (cf. Haspelmath, 2003, for an illuminating presentation of the semantic map methodology). However, and importantly, the fact that one and the same linguistic expression is represented as associated with several nodes does not necessarily mean that it is analyzed as polysemic. In other words, "multifunctionality", inherent in the method, "does not imply a commitment to a particular choice among monosemic and polysemic analyses" (Haspelmath, 2003: 213).

Semantic maps have become a popular method in grammatical typology of the last two decades, where they have been used for capturing both synchronic facts and patterns of development (cf. Cysouw et al., 2010, for a recent contribution). Most attempts to apply semantic maps to lexical typology are more recent and date back to the second half of the 2000s. These can be grouped into two main types – semantic maps per se, or implicational semantic maps (see 7.1), and probabilistic semantic maps (7.2).

The explicit non-commitment of semantic maps methodology to the issues of monosemy / polysemy makes it particularly suitable for lexical typology, where it has been applied both to the uses of one and the same lexeme within one domain (i.e., to the issues of domain categorization), as well as to the uses of one and the same lexeme or of lexemes related to it in different domains (normally seen as patterns of polysemy and semantic shifts). In particular the latter applications share certain similarities with conceptual / semantic networks as these are practiced in Cognitive Linguistics (cf. Newman, Geeraerts ?, Lemmens, this volume), but are theoretically much less sophisticated. Keeping in line with the main interest of this chapter, we will predominantly focus on the use of semantic maps in the research on domain categorization; however, we will start with two examples of semantic maps that aim at capturing patterns of polysemy and semantic shifts.

7.1. Implicational semantic maps in lexical typology

The generalization (prediction) inherent in a semantic map is that if a linguistic expression is associated with two non-adjacent nodes *a* and *c*, it will also be associated with the nod *b* connecting the two. The "contiguity / connectivity requirement" underlying the methodology of semantic maps is therefore reminiscent of – but also simultaneously weaker than – the traditional typological tool of implicational generalizations (if *a* then *b*). Early examples of graphical representations of implicational lexico-typological generalizations are found in Viberg's (1984) influential paper on perception verbs in about fifty languages. As Viberg demonstrates, "[o]ne of the most striking characteristics of the lexicalization patterns of the verbs of perception is the large amount of polysemy with respect to the sense modalities" (Viberg, 1984: 136). He shows that lexicalization of perception by verbs across languages and the patterns of sense conflation ("patterns of polysemy") follow the sense-modality hierarchy, from sight followed by hearing, as the highest sense modalities, to smell, taste and touch as the lowest ones. To give an example, if a language has a dedicated verb for touching, tasting or smelling, it will distinguish between hearing and seeing verbs.

One of the first attempts to fully elaborate the methodology of implicational semantic maps for lexical typology is found in François (2008). François' semantic maps target cross-linguistically recurrent patterns of semantic shifts as manifested by polysemous words and words related to them by derivation and compounding and illustrate the methodology by means of the pivot concept 'breath' (cf. Newman, this volume, for details on François' methods and an illustration).

Semantic maps are currently widely used by the Moscow Lexico-Typological Group (cf. Section 6 and Rakhilina and Reznikova, 2014), where the nodes correspond to "frames", i.e. the typologically relevant situations. Fig. 1 shows the semantic map for the domain of EMPTINESS and its lexicalization in Chinese and Serbian, where the crosslinguistically relevant frames are defined as 'hollow shape' (e.g., "a hollow cylinder") 'empty container' (e.g., "an empty bottle"), 'location

10

empty of people' (e.g., "an empty house), 'large space without objects in it' (e.g., "an empty field"), 'small flat surface without things on it' (e.g., "an empty table"), and 'empty hanger'.
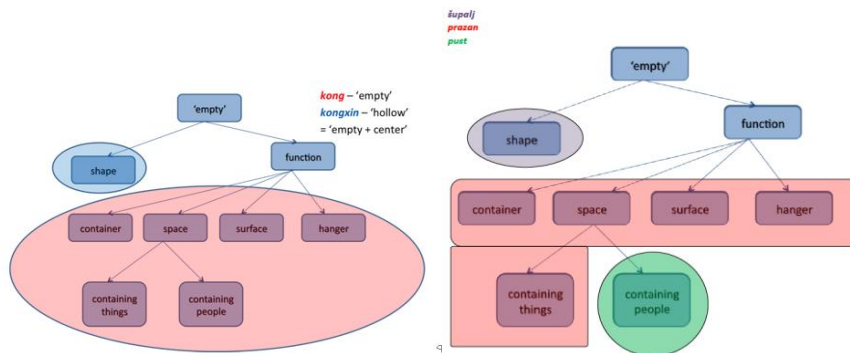

Fig. 1: Semantic maps of EMPTINESS *in Mandarin and in Serbian (in Rakhilina and Reznikova 2014)*

7.2. Probabilistic semantic maps

Probabilistic semantic maps make use of statistical methods based on correspondence analysis of similarity matrices, where matrices contain occurrences of particular linguistic expressions (e.g., lexemes) in a number of different contexts across languages. Although also anchored in the notions of semantic spaces and semantic maps, the resulting representations differ significantly from those considered in 7.1. The data themselves may come from very different sources – e.g., informants' replies to a particular set of stimuli, or particular sentences in parallel texts, i.e. translations of one and the same text into different languages. Below we will illustrate both approaches.

The earliest example of a cross-linguistic probabilistic semantic map comes from the denotation-based Nijmegen project on "cutting" and "breaking" events (Majid et al., 2007), where the data consist of descriptions of events represented in videoclips by speakers of 24 languages (cf. Section 5 for the details). These are arranged, for each of the languages separately, in a similarity matrix which is built on the assumption that clips that are often described with the same predicate are more similar semantically than clips described with different predicates. Correspondence analysis was then jointly applied to these matrices to extract the main dimensions along which the languages (or rather the informants) grouped or differentiated events. All in all, four dimensions suffice to categorize the replies cross-linguistically, with the main dimension being the relative predictability of the locus of separation in the affected object: cf. high predictability for slicing carrots with a small knife, poor predictability for smashing a plate with a hammer, and in-between predictability for chopping off a branch with an axe. Within events with a poorly predictable locus of separation some languages make a further distinction between smashing a rigid object like a plate, pot, or carrot with a sharp blow and snapping a long object like a stick or a carrot into two pieces between the hands or over the knee. Finally, tearing events (e.g., a hand action on a two-dimensional flexible object) are very often endowed with a verb of their own. The resulting semantic space and the first three dimensions underlying it are shown in Fig. 2.
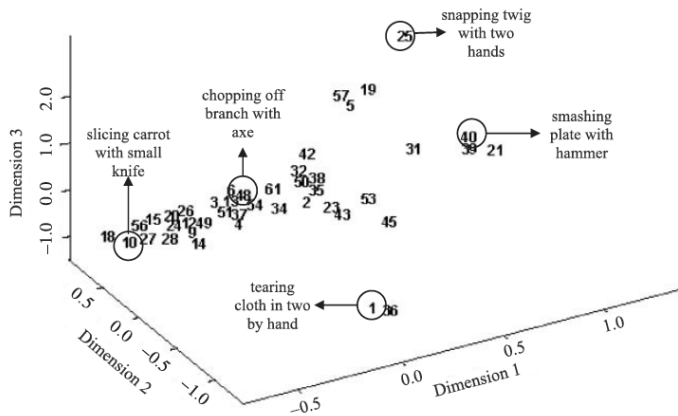
Fig. 2: A three-dimensional plot of cutting and breaking events (Majid et al., 2007: 143)

The most recent example of a cross-linguistic probabilistic semantic map comes from Wälchli and Cysouw's (2012) study of motion events in parallel texts, i.e. in translations of the Gospel according to Mark into 100 languages from all continents. The method uses Multidimensional Scaling (MDS) for visualization and the Hamming distance <mark>(here the fraction of languages for which the relevant predicates are lexicalized differently)</mark> for the statistics, and the idea is that

> "[t]he closer two contextually embedded situations are represented in a semantic map the more likely it is that they are represented by the same category in any language in the database. A probability space is accurate to the extent that it predicts crosslinguistically recurrent tendencies in the categorization of form classes" (Wälchli and Cysouw, 2012: 679).

The domain of motion verbs turns out to require ten to twenty dimensions for capturing the crosslinguistically most frequent lexicalization patterns (cf. the four dimensions for breaking and cutting events), with several crosslinguistically recurrent lexicalization clusters for which the authors use descriptive labels such as 'go', 'come' or 'arrive'. These three serve as the main illustration in the paper and are represented below by the semantic maps for two unrelated languages – Classical Greek and Sora (Austroasiatic: Munda).
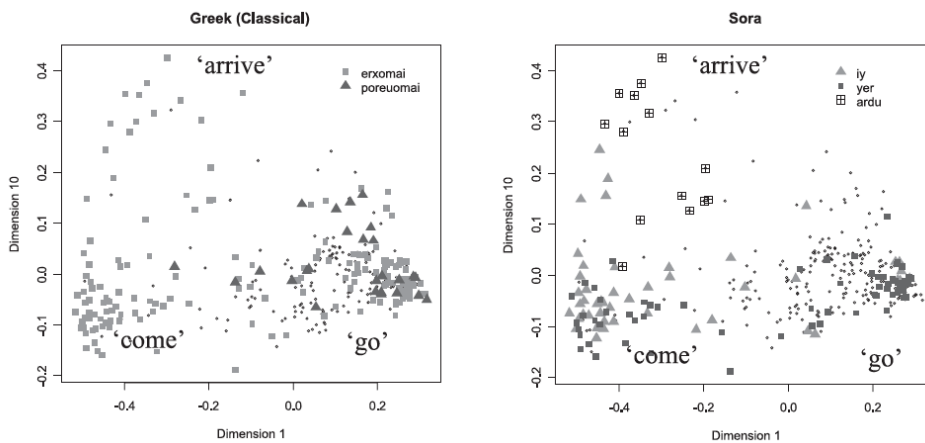

Fig. 3: Probabilistic semantic maps of the 'go', 'come' and 'arrive' domain in Classical Greek and Sora in the Gospel after Mark (Wälchli and Cysouw, 2012: 701).

Semantic maps are particularly useful when the arrangement of the nods and/or their general geometry receives a further explanation and/or can be used for various predictions. For instance, due to the intermediate level of predictability of the locus of separation for chopping off a branch with an axe, some languages group this event together with "precise control" events (such as slicing carrots with a small knife), while others group it with "imprecise control" events (such as smashing a plate with a hammer; Majid et al., 2007). Another example concerns the temperature domain (Koptjevskaja-Tamm, forthcoming), where one of the main dimensions underlying the crosslinguistic variation concerns the distinctions among three frames of temperature evaluation – tactile temperature (*The stones are cold*), ambient temperature (*It is cold here*), and personal-feeling temperature (*He feels cold*). While English 'cold' can be used for all the three frames (competing with *freeze* for personal-feeling temperatures), other languages may require several different lexemes here. Armenian distinguishes between the 'tactile cold' *saŕn*, the 'ambient cold' *c'urt*, and the 'personal-feeling (to be) cold' *mrsum*; Kamang (Timor-Alor-Pantar) distinguishes between the 'non-personal cold' (i.e., tactile + ambient) *kamal* and the 'personal-feeling cold' *faatei*, while Palula (Indo-European: Dardic), finally, distinguishes between the 'tactile cold' *taátu* and the 'non-tactile cold' (i.e., ambient + personal-feeling) *húluk*. The recurrent pattern, visualized in the semantic maps in Fig. 4, is that personal-feeling temperatures are often singled out by languages in various ways, whereas the linguistic encoding of ambient temperature may share properties with those of either tactile or personal-feeling temperature.
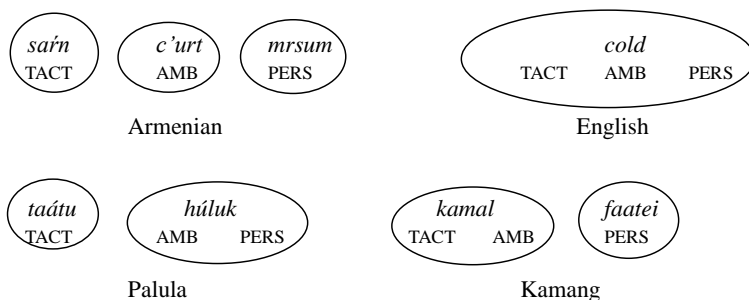

Fig. 4: 'Cold' in Armenian, English, Palula and Kamang

The motivation for this lies in the conceptual and perceptual affinities of ambient temperature with both other frames of temperature evaluation. On the one hand, ambient and personal-feeling temperatures are rooted in the same type of experience, thermal comfort, whereas tactile temperature relates to evaluation of the temperature of other entities based on perception received by the skin. On the other hand, tactile and ambient temperatures are about temperatures that can be verified from "outside", whereas personal-feeling temperature is about a subjective "inner" experience of a living being (for details cf. Kopjevskaja-Tamm, forthcoming).

Now, while semantic maps in general offer a practical method for representing cross-linguistic generalizations, they do not embody any stance on the relation between an 'etic' concept (i.e., the range of the uses of a particular linguistic expression) to the corresponding 'emic' concept (i.e., its descriptive meaning / sense / intension). And in fact most of the researchers do not even touch upon this issue. The question that is, however, sometimes discussed is whether semantic maps necessarily correspond to a speaker's mental representations, with some voices arguing for, or at least assuming this position (e.g., Haspelmath, 2003) and some against (Wälchli and Cysouw, 2012).

An interesting and so far unexploited aspect of semantic maps is their potential applicability for combining cross-linguistic generalizations on domain categorization with those on cross-domain semantic shifts. First of all, semantic extensions normally originate not from the lexeme as a whole,

but from its particular use, or 'frame', and the same frame, or the same use, may lead to similar semantic shifts across languages. For instance adjectives denoting functional sharpness (as in *sharp knife*) frequently evolve into intensifiers (as in *sharp contrast*). This shift, however, is not observed for a similar class of adjectives denoting pointed shape (as in *pointed shoes*). Specialized adjectives like *pointed* are not likely to develop intensifier meaning, although both types of "sharpness" are semantically very close, and are typically expressed by one and the same polysemous lexical item cross-linguistically.

In addition, there may be interesting regularities in mapping across whole domains. As an example, take the extensive research on the semantic extensions from perception to cognition, cf. "to see" as 'understand' in *Do you see what I mean* (Sweetser, 1990; Evans and Wilkins, 2000), further enriched with prehension predicates, such as 'seize', in Vanhove (2008). Vanhove proposes an implicational hierarchy to describe the predicates' potential for semantic extension into cognition: hearing > vision > prehension. In other words, if a language has a prehension word which maps onto the domain of mental perception, it also has another lexical item with a similar semantic association for vision and the auditory sense, but the reverse is not true.

A possible solution for capturing such generalizations might be three-dimensional semantic maps, with maps of the source and target domains aligned as layers and shifts connecting the nodes on one layer to the nodes on the other.

8. FUTURE PROSPECTS

Systematic research in lexical typology has so far been carried out on rather limited language samples. These samples are often quite sufficient for falsifying some assumptions on the universality of a particular phenomenon and for unveiling major patterns in its cross-linguistic variation, but are hardly adequate for drawing safe conclusions on the interplay among the various factors behind it or for clearly distinguishing between universal determinants and those due to historical relations among the languages. Systematic research in lexical typology needs therefore to be extended to more linguistic phenomena and to more languages. In particular, sign languages have been largely missing in most lexico-typological studies.

An important task facing lexical typology is to focus more on historical processes, and especially on cross-domain semantic shifts, i.e. on the emergence and functioning of metaphoric and metonymical word-senses (see also Newman, Lemmens, this volume). While Cognitive Linguistics emphasizes the universality of many cognitive metaphors, the research on regular semantic shifts across languages is so far very limited. There is, however, ample evidence of cross-linguistically recurrent patterns here. A few examples, in addition to the development of cognitive verbs mentioned in 7.2., include the research on the sources of body-part terms (e.g., Wilkins 1996, Koch, 2008), Urban's (2012) large-scale study of motivational patterns underlying many different kinds of referring expressions, and Reznikova et al.'s (2012) study of metaphorical sources of pain expressions. Of the approaches discussed above, the classic structuralist method is inherently static, psycholinguistic experiments lack the tools to capture lexical metaphors, and the NSM school with its interest in abstract domains and culture-specific concepts is less concerned with the typology of shifts from physical domains. But even for the purposes of domain-categorization research, semantic shifts are worthy of special attention in providing additional clues for understanding the structure of their source domain. Just as a word having several antonyms is usually an indication of polysemy, with antonyms corresponding to distinct word-senses that can be represented by separate lexemes in some other language, the source of a semantic shift may also point out a word-sense, even though this does not necessarily hold in all cases (cf. the example of 'sharp contrast' mentioned in 7.2).

A related issue concerns syntactic and morphosyntactic properties of words. It is well-established in lexicology and lexicography that collocation, valency, case marking and other patterns of argument realization, as well as many morphological characteristics such as countability of nouns or

verbal aspectual classes, can be motivated by semantics. However, there is a long tradition of distinguishing between grammatical typology, focusing on the grammatical behaviour of words and on morphosyntactic patterns as encoding meanings, and lexical typology, that has to a large extent been restricted to domain-categorization by lexical means, without further considering their grammatical behaviour. Fortunately, the recent developments within the crosslinguistic research on domain-categorization are witness of an ambition to reconcile lexical and grammatical interests and to engage in a dialogue with linguistic grammatical theory. For instance, one of the leading issues in the project on cutting and breaking events (Majid et al., 2007) has been the interface between syntax and lexical semantics, i.e. to what extent and how the argument structure properties of a verb are predictable from its meaning. Construction Grammar is gaining popularity as an appropriate framework for lexico-typological research, e.g., on PAIN predicates (Reznikova et al., 2012), TEMPERATURE (Koptjevskaja-Tamm, 2011), and LOCATION–EXISTENCE–POSSESSION (Koch, 2012). Construction-Grammar inspired schemas are capable of covering linguistic phenomena on different levels (lexicon and grammar) and their interaction.

An in-depth study of lexical items that would take into account polysemy and formal properties requires exponentially more time and labour than simply identifying the core vocabulary in a domain. To make it feasible lexical typology needs to look for methods that would automate the collection and preliminary analysis of data as much as possible. Computational linguistics, in particular machine translation, has made significant progress in word-sense disambiguation, identifying and grouping contexts, aligning parallel texts and other related tasks. Implementing computational methods in a lexico-typological study would alleviate some of the manual work involved, and increased use of statistics, in particular, would give more weight to typological generalizations.

Of course, only a handful of languages at present have corpora of sufficient size and representativeness to make them fully useful to lexical typology, but as corpus linguistics continues growing, more corpora in more languages become available, with new tools developing for extracting a wealth of relevant information from them.

RELATED TOPICS

Chap. 6. Cognitive semantics: Maarten Lemmens
Chap. 8. Categories, prototypes and exemplars: James Hampton
Chap. 14. Lexical decomposition: Nick Riemer
Chap. 15. Sense individuation: Dirk Geeraerts
Chap. 16. Sense relations: Petra Storjohann
Chap. 17. Semantic extension : John Newman

> **Commenté [M2]:** Should we add Peter Koch's if he had time to send it before he passed away ?

REFERENCES

Ameka, F.K. and Levinson S.C. eds., 2007. *The typology and semantics of locative predication: Posturals, positionals and other beasts*. [Special Issue]. *Linguistics,* 45(5).

Andersen, E., 1978. Lexical universals of body-part terminology. In: Joseph H. Greenberg, ed., 1978. *Universals of human language.* Stanford: Stanford University Press, pp. 335–368.

Apresjan, J.D., 2000. *Systematic lexicography*. Translated from Russian by K. Windle. Oxford: Oxford University Press.

Behrens, L. and Sasse, H.-J., 1997. *Lexical typology: A programmatic sketch. Arbeitspapier* Nr. 30 (Neue Folge). Cologne: University of Cologne, Department of linguistics.

Berlin, B. and Kay, P. 1969. *Basic color terms: their universality and evolution.* Berkeley, CA: University of California Press.

Brown, C.H., 2001. Lexical typology from an anthropological point of view. In: M. Haspelmath et al., eds., 2001. pp. 1178–1190.

Brown, C.H., 2005a. Hand and arm. In: M. Haspelmath et al., eds., 2001. pp. 522-525.

Brown, C.H., 2005b. Finger and hand. In: M. Haspelmath et al., eds., 2001. pp. 526-529.

Bruner, J.S., Goodnow, J.J. and Austin, G.A. eds., 1956. *A Study of Thinking*. New York: John Wiley.

Cysouw, M., Haspelmath, M. and Malchukov A. eds., 2010. "Semantic maps: methods and application". [Special issue] *Linguistic discovery*, 8(1), [online] Available at http://journals.dartmouth.edu/cgi-bin/WebObjects/Journals.woa/1/xmlpage/1/issue/34 [accessed 21 May 2014].

Evans, N., 2011. Semantic typology. In: Jae Jung Song, ed., 2011. *The Oxford Handbook of Typology*. Oxford: Oxford University Press, pp. 504–533

Evans, N. and Wilkins, D.P., 2000. In the mind's ear: The semantic extensions of perception verbs in Australian languages. *Language* 76, pp. 546-592.

Fillmore, C.J., 1982. *Frame semantics. Linguistics in the Morning Calm*. Seoul, South Korea: Hanshin Publishing Co, pp. 111-137.

Firth, J.R., 1957. *Studies in linguistic analysis*. Oxford: Blackwell.

François, A., 2008. Semantic maps and the typology of colexification: Intertwining polysemous networks across languages. In: M.Vanhove, ed., 2008. pp. 163-215.

Goddard, C. and Wierzbicka, A., 2007. NSM analyses of the semantics of physical qualities: sweet, hot, hard, heavy, rough, sharp in cross-linguistic perspective. *Studies in Language*, 31(4), pp. 765–800.

Goddard, C., 2001. Lexico-semantic universals: A critical overview. *Linguistic Typology*, 5(1), pp. 1–65.

Goddard, C., 2012. Semantic primes, semantic molecules, semantic templates: Key concepts in the NSM approach to lexical typology. In: M. Koptjevskaja-Tamm and M. Vanhove, eds., 2012. pp. 711–743

Greenberg, J.H., 1980. Universals of kinship terminology: their nature and the problem of their explanation. In: Jacques Maquet, ed., *On Linguistic Anthropology: essays in honor of Harry Hoijer*, 9-32. Malibu: Undena Publications [Reprinted in Keith Denning and Suzanne Kemmer, Stanford: Stanford University Press, 1990, *Selected Writings: Selected Writings of Joseph Greenberg H.*, pp. 310-327].

Harkins, J.and Wierzbicka, A. eds., 2001. *Emotions in cross-linguistic perspective*. Berlin: Mouton de Gruyter.

Haspelmath, M., 2003. The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In: M. Tomasello, ed., 2003. *The New Psychology of Language*, 2, Mahwah, NJ: Lawrence Erlbaum, pp. 211–242.

Haspelmath, M., Dryer, M., Gil, D. and Comrie B. eds., 2005. *The World Atlas of Language Structures (WALS)*. Oxford: Oxford University Press.

Haspelmath, M., König, E., Oesterreicher, W. and Raible W. eds., 2001. *Language Typology and Language Universals*, v.1-2. Berlin: Walter de Gruyter.

Kay, P. and Maffi L., 2005. Colour terms. In: M. Haspelmath et al., eds., 2005. pp. 534-545.

Koch, P., 2001. Lexical typology from a cognitive and linguistic point of view. In: M. Haspelmath et al., eds., 2001. pp. 1142–1178.

Koch, P., 2008. Cognitive onomasiology and lexical change: around the eye. In: M. Vanhove, ed., 2008. pp. 107–137.

Koch, P., 2012. Location, existence, and possession: A constructional-typological exploration. In: M. Koptjevskaja-Tamm and M. Vanhove, eds., 2012. pp. 533–604.

Koptjevskaja-Tamm, M., 2008. Approaching lexical typology. In: M. Vanhove, ed., 2008. pp. 3–52.

Koptjevskaja.Tamm, M., Divjak, D. and Rakhilina, E., 2010. Aquamotion verbs in Slavic and Germanic: A case study in lexical typology. In: V. Driagina-Hasko and R. Perelmutter, eds., 2010. *New approaches to Slavic verbs of* motion. Amsterdam, Philadelphia: John Benjamins, pp. 315–341.

Koptjevskaja-Tamm, M. and Vanhove, M. eds., 2012. *New Directions in Lexical Typology*. [special issue] *Linguistics* 50(3).

Koptjevskaja-Tamm, M. ed., forth. *The linguistics of temperature*. Amsterdam, Philadelphia: John Benjamins.

Krugljakova, V.A., 2010. *Semantika glagolov vraščenija v tipologičeskoj perspektive* [Semantics of rotation verbs in typological perspective]. Ph.D. thesis, Russian State University for Humanities.

Lang, Ewald. 2001. Spatial dimension terms. In Haspelmath et al., eds., v. 2, 1251–1275.

Lehmann, C., 1990. Towards lexical typology. In: W. Croft, K. Denning and S. Kemmer, eds., 1990. *Studies in typology and diachrony*: *Papers presented to Joseph H. Greenberg on his 75th birthday*. Amsterdam, Philadelphia: John Benjamins, pp. 161–185

Lehrer, A., 1992. A theory of vocabulary structure: Retrospectives and prospectives. In: Pütz, Manfred, ed., 1992. *Thirty years of linguistic evolution. Studies in honour of René Dirven on the occasion of his sixtieth birthday*. Amsterdam, Philadelphia: John Benjamins, pp. 243–256.

Lehrer, Adrienne. 1974. Semantic fields and lexical structure. Amsterdam, London: North-Holland Publishing Company.

Levinson, Steven, 2000. Yélî Dnye and the theory of basic color terms. *Journal of Linguistic Anthropology*, 10(1), pp. 3–55.

Levinson, S. and Wilkins, D., 2006. *Grammars of space*. Cambridge: Cambridge University Press.

MacLaury, R.E., Paramei, G.V. and Dedrick, D. eds., 2007. *Anthropology of Color: Interdisciplinary multilevel modeling*. Amsterdam, Philadelphia: John Benjamins.

Majid, A. and Bowerman, M. eds., 2007. *"Cutting and breaking" events: A crosslinguistic perspective.* [special issue] *Cognitive Linguistics*, 18(2).

Majid, A., Enfield, N.J. and van Staden, M. eds., 2006. *Parts of the body: Cross-linguistic categorisation.* [Special issue]. *Language Sciences*, 28(2-3).

Narasimhan, B. and Kopecka, A. eds. 2012. *Events of 'putting' and 'taking': A crosslinguistic perspective.* Amsterdam, Philadelphia: John Benjamins.

Nerlove, S. and Romney, A.K., 1967. Sibling terminology and cross-sex behavior. *American Anthropologist*, 74, pp 1249-1253.

Nichols, J., Peterson, D. A. and Barnes, J., 2004. Transitivizing and detransitivizing languages. *Linguistic Typology*, 8(2), pp. 149-211.

NOS = *Novyj objasnitel'nyj slovar' sinonimov russkogo jazyka* [A new explanatory dictionary of synonyms in the Russian language]. 2004. Moskva; Vena: Jazyki slavjanskoj kul'tury: Venskij slavističeskij al'manax.

Paradis, C. and Willners, C., 2011. Antonymy: from conventionalization to meaning-making. *Review of Cognitive Linguistics*, 9(2), pp. 367–391.

Rakhilina, E., 2010. Verbs of rotation in Russian and Polish. In: V. Hasko and R. Perelmutter, eds., 2010. *New Approaches to Slavic Verbs of Motion*. Amsterdam, Philadelphia: John Benjamins, pp. 291-314.

Rakhilina, E. and Reznikova, T., 2014. Doing lexical typology with frames and semantic maps. *Higher School of Economics Working Papers in Linguistics* (WP BRP 10/LNG/2014).

Reznikova, T., Rakhilina, E. and Bonch-Osmolovskaya, A., 2012. Towards a typology of pain predicates. In: M. Koptjevskaja-Tamm and M. Vanhove, eds., 2012. pp. 421–466.

Riemer, N., 2005. *The semantics of polysemy*. Berlin, New York: Mouton de Gruyter.

Riemer, N., 2006. Reductive Paraphrase and Meaning: A Critique of Wierzbickian Semantics. *Linguistics and Philosophy*, 29(3), pp. 347-379.

Robertson, D. and Hanley, J.R., 2010. Relatively speaking: an account of the relationship between language and thought in the color domain. In: B.C. Malt and P. Wolff, eds., 2010. *Words and the mind. How words capture human experience*. Oxford: Oxford University Press, pp. 183–198.

Sweetser, E., 1990. *From Etymology to Pragmatics. Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge: Cambridge University Press.

Talmy, L., 1985. Lexicalization patterns. In: Timothy Shopen, ed., *Language typology and synchrohic description, vol. 3.* Cambridge: Cambridge University Press, pp. 57–149.

Taylor, J.R., 1992. Old problems: adjectives in Cognitive Grammar. *Cognitive Linguistics* 3(1), pp. 1-35.

Urban, M., 2012. *Analyzibility and semantic associations in referring expressions. A study in comparative lexicology*. PhD diss., Leiden University.

Vanhove, M., 2008. Semantic associations between sensory modalities, prehension and mental perceptions: A crosslinguistic perspective. In: M. Vanhove, ed., 2008. pp. 342-370.

Vanhove, M. ed., 2008. *From Polysemy to Semantic Change*. Amsterdam, Philadelphia: John Benjamins.

Viberg, Å., 1984. The verbs of perception: A typological study. In: B. Butterworth, B. Comrie and Ö. Dahl, eds., 1984. *Explanations for Language Universals*, Berlin, New York, Amsterdam: Mouton, pp. 123-162.

Wälchli, B. and Cysouw, M., 2012. Lexical typology through similarity semantics: Toward a semantic map of motion verbs. In: M. Koptjevskaja-Tamm and M. Vanhove, eds., 2012. pp. 671–710.

Wierzbicka, A., 1984. Cups and mugs: Lexicography and conceptual analysis. *Australian Journal of Linguistics*, 4(2), pp. 205-255.

Wierzbicka, A., 1987. Kinship Semantics: Lexical Universals as a Key to Psychological Reality. *Anthropological Linguistics*, 29(2), pp. 131-156.

Wierzbicka, A., 1999. *Emotions across languages and cultures: diversity and universals*. Cambridge: Cambridge University Press.

Wierzbicka, A., 2005. There are no "color universals". But there are universals of visual semantics. *Anthropological Linguistics*, 47(2), pp. 217 – 244.

Wierzbicka, A., 2007. Bodies and their parts: an NSM approach to semantic typology. *Language Sciences*, 29, pp. 14-65.

Wilkins, David P. 1996. Natural tendencies of semantic change and the search for cognates. in: Durie, Mark and Malcolm Ross (eds.), *The Comparative Method Reviewed. Regularity and Irregularity in Language Change*. New York/Oxford: Oxford University Press, 264–304.

FURTHER READING

Koptjevskaja-Tamm, M. and Vanhove, M. eds., 2012. *New Directions in Lexical Typology*. [special issue] *Linguistics* 50(3).
This volume is the most recent update, representing a whole range of approaches to lexical typology.

Majid, A. and Bowerman, M. eds., 2007. *"Cutting and breaking" events: A crosslinguistic perspective*. [special issue] *Cognitive Linguistics*, 18(2).
A detailed research on crosslinguistic similarities and differences in semantic categorization, applying statistical modeling to the descriptions of cutting and breaking events.

Narasimhan, B. and Kopecka, A. eds. 2012. *Events of 'putting' and 'taking': A crosslinguistic perspective*. Amsterdam, Philadelphia: John Benjamins.
An indepth crosslinguistic study of placement and removal event descriptions describing their lexical and grammatical means, syntax-semantics mappings, lexical semantics, and asymmetries in the encoding of the two domains.

Wierzbicka, A., 1999. *Emotions across languages and cultures: diversity and universals*. Cambridge: Cambridge University Press.
A classic of the Natural Semantic Metalanguage approach: an introduction to the theory as a tool for cross-cultural analysis and the discovery of semantic primes and universals, clearly illustrated by the domain of emotions.