Typological approaches to lexical semantics

MARIA KOPTJEVSKAJA-TAMM, MARTINE VANHOVE, and PETER KOCH

1. Introduction

Although the term "lexical typology" is often used as if it were self-explanatory there is not much consensus on what exactly it can refer to, apart from the evident fact that it involves crosslinguistic research on the lexicon. 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. the overviews in Koch 2001 and Brown 2001). This would make lexical typology a sub-branch of semantic typology concerned with the lexicon, where semantic typology, in the definition of Evans (forthcoming), is "the systematic cross-linguistic study of how languages express meaning by way of signs". Other definitions of lexical typology, clinging to the apparently safer interface with grammar, focus on "typologically relevant features in the grammatical structure of the lexicon" (Lehmann 1990: 163) or on typologically relevant vs. language-specific patterns of the lexicon-grammar interaction (Behrens & Sasse 1997).

A reasonable way of defining lexical typology is to view it as the crosslinguistic and typological branch of lexicology (as itself delimited in Cruse et al. (eds.) 2002/05). In the same way as lexicology in general is not restricted to lexical semantics and is not completely opposed to either phonetics/phonology, morphology, or syntax, lexical typology can include phenomena that are not of primary interest for semantic typology and can raise questions that can be answered from different angles or within approaches that integrate several perspectives, goals, and methods. One such issue concerns possible vs. impossible words in different languages, different criteria for identifying words and interactions among them, universal vs. language-specific restrictions on possible and impossible, well- and ill-formed words. Traditional morphological typology, with its focus on how much and what kind of morphology is found across

languages (isolating vs. polysynthetic languages, etc.), represents one way of comparing possible words crosslinguistically. This huge and complex issue requires, however, a truly integrating approach, where morphological (and further grammatical), phonetic/phonological, and semantic criteria as well as psycholinguistic considerations (holistic storage and processing, acquisition) and sociolinguistic/pragmatic factors (such as the degree of conventionalization) are all relevant. Important recent contributions include Aikhenvald & Dixon (eds.) (2002) and the on-going project *Word Domains* within the research programme AUTOTYP (Bickel & Nichols 2007) which both focus on words as phonological and grammatical domains, as well as Corbett's (2007, forthcoming) "canonical approach" to inflection which evaluates the various formal ways in which the word forms of one and the same lexeme can be related to each other.

In addition to lexical-typological issues not covered by semantic typology, there are semantic-typological issues not covered by lexical typology. Semantic typology is, in fact, orthogonal to the distinction between grammar and lexicon (and even prosody). Studies in grammatical typology without any semantic orientation are relatively restricted; rather, they tend to differ in the degree to which semantic aspects of grammatical phenomena are foregrounded or backgrounded. In other words, lexical typology and semantic typology considerably overlap without being necessarily reducible to one another.

In this article we will mainly be concerned with semantically oriented lexical typology. In Section 2 we survey work on what we consider the three major, and often interrelated, focal questions that can be posed within lexical semantic typology. As we concentrate on research during the last ten years and have no ambitions of covering the whole field, our survey is far from exhaustive and the references provided are selective, with a bias towards our own work and interests. (For a more general overview see Koptjevskaja-Tamm forthcoming.) In *Linguistic Typology* lexical semantic typology has so far received relatively little attention, which the journal itself can of course hardly be blamed for. Section 3 deals with the relevance of lexical semantic typology for linguistic theory and vice versa. Section 4 discusses the central methodological problems facing lexical semantic typology and some perspectives for future research.

2.   Focal questions for lexical typology, and their reflection in *Linguistic Typology*

2.1.   *What meanings can or cannot be expressed by a single word in different languages?*

This question is the one that probably comes to mind first on hearing the term "lexical typology". Relevant issues concern lexicalizations and lexicalization

patterns, universal vs. language-specific lexicalizations, categorization within lexical fields or semantic domains by lexical items, and the architecture of lexical fields or semantic domains (e.g., basic vs. derived words).

The issues of categorization within lexical fields or conceptual domains have hardly been reflected at all in *LT*, apart from Davies et al. (1999), who discuss the colour system in Tsakhur, which includes an unusual basic term for turquoise (*a|nti:k'a*), and its theoretical implications. In general, systematic crosslinguistic research on semantic categorization has only been carried out for a handful of conceptual domains typically encoded by words (rather than by grammatical means): COLOUR, BODY, KINSHIP, PERCEPTION, MOTION, EVENTS OF BREAKING AND CUTTING, DIMENSION (for spatial dimension terms, see Lang 2001). The list can be made slightly longer if we include words and expressions with more grammatical meanings, such as INDEFINITE PRONOUNS (Haspelmath 1997), PHRASAL ADVERBIALS (Auwera 1998), SPATIAL ADPOSITIONS (Levinson & Meira 2003), and QUANTIFICATION (Bach et al. (eds.) 1995, Auwera 2001, Gil 2001).

The scarcity of these questions in *LT* is, thus, partly related to the relatively marginal place of crosslinguistic and typological research on lexicalization. This is only partly true, however, since the last decade has also seen several important collaborative crosslinguistic investigations devoted to categorization within lexical fields: e.g., studies on the BODY and on EVENTS OF BREAKING AND CUTTING (Majid et al. 2006 forthcoming) coordinated from the Max Planck Institute for Psycholinguistics in Nijmegen, and the project on AQUA-MOTION, the semantic field of moving and being in liquid medium, directed by Moscow linguists (Majsak & Rakhilina (eds.) 2006). The main publications for each of these projects, providing much information not previously available, are volumes with comparable papers on each of the languages investigated, ranging from 10 for BODY to 28 for BREAKING AND CUTTING and 43 for AQUA-MOTION. Although this book-type format is hardly suitable for *LT*, the journal will, hopefully, be used as a forum for spreading major generalizations deriving from such investigations.

The issue of universal vs. language-specific lexicalizations has been addressed in *LT* by Goddard (2001), in the probably most comprehensive overview of suggested lexico-semantic universals. Goddard aims at precise meaning definitions in the framework of Natural Semantic Metalanguage (NSM, see Goddard & Wierzbicka (eds.) 1994, among many other works). In accordance with these standards, as Goddard argues, SEE and HEAR seem to stand the proof of being universally lexicalized (at least as separate meanings within polysemous expressions), while such concepts as EAT and GIVE seem doubtful as lexical universals or, at least, can only be viewed as approximate ones. A few of the other surprises include the non-universal status of WATER and SUN, based on the fact that languages can have more than one word for each of those (cf.

'hot water' vs. 'non-hot water' in Japanese, 'sun low in the sky' vs. 'hot sun overhead' in Nyawaygi, Australia). However, opinions do vary strongly, also among the individual contributors to Goddard & Wierzbicka (eds.) (1994), both on the NSM approach in general and on what count as universal lexicalizations or what evidence there is for or against them (cf. Brown 2001).

2.2.   *What different meanings can be expressed by one and the same lexeme, by lexemes within one and the same synchronic word family, or by lexemes historically derived from each other?*

Here we are dealing with crosslinguistically recurrent patterns of motivation, i.e., in the relations among words and lexical items in the lexicon.

These questions can be approached from different angles. We can start with an individual LEXICAL ITEM, or with several items belonging to one and the same individual LEXICAL FIELD, and ask what other lexical or grammatical meanings can be expressed by the same form(s) or by forms derived from them. Hereby we focus on semantic relations, diachronic, synchronic, or both, between particular lexical units, or between particular meanings — i.e., on particular instances of SEMANTIC MOTIVATION (polysemy, semantic associations, semantic shifts). But we can also compare whole CLASSES or GROUPS OF WORDS where one of the classes contains words derived from, or formed on, words in the other one, and ask about the semantic relations associated with a particular word formation device (derivational patterns, compounding). The focus here is on the REGULARITIES in lexical motivation seen as an interaction of formal (morphological) and semantic aspects of motivation (cf. Koch & Marzo 2007). These will be dealt with separately in the following two subsections.

2.2.1.   *Particular instances of semantic motivation.*   The issues of categorization within lexical fields or conceptual domains and of semantic motivation are intimately related in being ultimately dependent on what counts as general meaning vs. polysemy. Thus, if the Russian noun *ruka* is seen as neutralizing the difference between 'hand' and 'arm', Russian differs from English in its categorization of the corresponding conceptual space; however, if 'hand' and 'arm' are recognized as two different senses of the polysemous lexeme *ruka*, Russian and English show the same categorization, but differ in their polysemy patterns, or semantic motivation. Current crosslinguistic research on the lexicon is, with very few exceptions (e.g., Riemer 2005), implicit and inconsistent in the question of polysemy or semantic generality, which complicates comparisons between different studies. By contrast, the method of semantic maps used in crosslinguistic comparisons of grammatical forms is explicitly agnostic about the distinction between polysemy and semantic generality (Haspelmath 2003: 231).

Due to the high profile of the issue of grammaticalization in recent research, there are many examples of crosslinguistically recurrent (diachronic) semantic shifts from lexical to grammatical items. Somewhat surprisingly, *LT* has no articles specifically devoted to such cases, even though they are occasionally mentioned in some publications (the clearest case is the brief note on the lexical sources of modal meaning in Auwera & Plungjan 1998). The recent literature includes Enfield (2003) for ACQUIRE, Newman (1997, 2002) for GIVE and verbs of posture, Majsak (2006) for motion verbs, Heine & Kuteva (2002) for scores of items, and many other studies too numerous to be done justice to here.

Apart from that, crosslinguistic research on semantic motivation has so far been very restricted. Brown (2001) reports on several crosslinguistically recurrent connections (WOOD – TREE, SEED – FRUIT, WIND – AIR), for which social and cultural factors have been suggested. The few other notable exceptions include research on semantic change and extensions related to BODY and on PERCEPTION verbs developing COGNITIVE meanings. Here we will briefly report on three recent and on-going projects that are specifically oriented towards the systematic crosslinguistic study of recurrent semantic motivation patterns in the lexicon. The aims include (i) distinguishing universal, areal, genetic, and cultural preferences and (ii) understanding the cognitive mechanisms at work there.

The project "Typology of semantic associations" lead by one of us (MV) within the *Fédération Typologie et Universaux Linguistiques* at the CNRS (Paris) has a panchronic orientation. "Semantic associations" (also named semantic extensions, derivations, parallels, affinities, or connections) is a cover term for three different but closely related semantic properties of the lexicon: polysemy, heterosemy (i.e., semantic extensions through morphosyntactic derivation), and semantic change. The basic assumption is that it is legitimate to have a notion that can be studied both from a synchronic and a diachronic perspective, since synchronic facts (polysemies, heterosemies, and contextual and pragmatic uses) underlie semantic change, and vice versa, synchronic polysemies are often a consequence of semantic innovation. Or, to quote Sweetser (1990: 45–46), "[t]hrough a historical analysis of 'routes' of semantic change, it is possible to elucidate synchronic semantic connections between lexical domains; similarly, synchronic connections may help clarify reasons for shifts of meaning in past linguistic history".

The project involves a group of field linguists, semanticists, cognitivists, and a specialist in Natural Language Processing (NLP), and relies mainly on first hand data collected by the members of the group, in order to maximize the advantages of the deep insight that field linguists have into the languages and cultures they study. The empirical research was conducted on a set of 45 languages belonging to eight genetic stocks, plus one pidgin and one creole, with

an average of 25 languages for each semantic association. The latter include (i) PERCEPTION, COGNITION (see, hear, take, know, understand . . .); (ii) OFFSPRING, SMALL, PART of a whole; (iii) FRIEND, OTHER, DUALITY, COMPLEMENTARITY, etc.; (iv) BREATH, SMELL, LIFE, SOUL, etc.; (v) ROOTS, ORIGIN, BASIS, CHIEF, etc.; (vi) EAT, TRY, DESTROY, etc.; (vii) DRINK, SMOKE, CATCH, SUFFER, etc.; (viii) MEAT, ANIMAL; (ix) EYE, EYEBALL, EYELASH, EYEBROW, EYELID; (x) CAT, MONKEY, MASK, CHILD, STUPID, MISFORTUNE, etc. (Vanhove (ed.) forthcoming). The choice of semantic notions was guided by the necessity of exploring different reasons for semantic associations, be they cultural or cognitive. It turns out that very few semantic associations can be considered universal (the semantic extension of auditory perception to cognition and internal reception might be one of them; cf. Vanhove forthcoming; also Evans & Wilkins 2000); and even if some of them are to be found in unrelated genetic groups for which no contact can be suspected, such parallelisms can very often be accounted for in terms of similar cultural environments. Thus, the crosslinguistically recurrent associations between cats or monkeys and small insects appear to stem from the cross-culturally recurrent conception of these living beings as connected to supernatural powers (Masson forthcoming), whereas the frequent connection between meat and animals is most certainly due to the importance of hunting and of game as a source of food (Boyeldieu forthcoming). Other theoretically and methodologically important contributions consist in applying the methodology of semantic maps to crosslinguistically recurrent semantic associations (for BREATH, SMELL, LIFE, SOUL etc.: François forthcoming) and testing the possibility of "exploiting lexical data bases obtained by field linguists in order to study a given corpus in a unified manner, to measure the semantic proximity between lexical terms and to compare the semantic networks in languages" (Gaume et al. forthcoming). Recent work on French dictionaries by Bruno Gaume and the research group *DiLan* (cf. Duvignau & Gaume 2004, also http://dilan.irit.fr/) provides a meeting place for typological and comparative issues. Dictionaries, once coded as graphs, show remarkable structural properties that do not seem to be language-dependent and support interactive linguistic and psycholinguistic models based on dynamic networks and the dynamic construction of meaning.

It is within this framework that Gaume proposes a method (already applied to French dictionaries, see http://prox.irit.fr/) for the automatic analysis of semantic associations crosslinguistically. The aim is to make an inventory of these semantic associations, to analyze their structures, to categorize them, and to measure their linguistic distribution. As he states, "to organize a cartography of all natural languages according to their semantic associations by hand, would be a gigantic task" (Gaume et al. forthcoming). Having a robust automatic method capable of capturing and measuring the confluences present in a paradigmatic network makes it possible to achieve at least two main tasks:

constructing the data and making the systematic and quantitative inventory of the semantic associations present in the data. The NLP tool *Prox* that has been used has proved to be a powerful automated tool for systematic searches and measurements of semantic associations. It can use existing data even if they show certain weaknesses as compared to linguistic reality.

Crosslinguistically recurrent panchronic semantic associations are also investigated in the research project *The Catalogue of Semantic Shifts*, lead by Anna Zalizniak at the Institute of Linguistics of the Russian Academy at Moscow (Zalizniak forthcoming); already approximately eight hundred semantic associations have been identified, each of them having from two to seven realizations in languages belonging to different linguistic families.

The strictly diachronic research project "Lexical Change – Polygenesis – Cognitive Constants: The Human Body" (*LexiType$_{Dia}$*, 1999–2004), run by one of us (PK) within the Collaborative Research Centre 441 *Linguistic Data Structures* at Tübingen University, focuses on lexical change leading to the emergence of body part terms within the domain of HEAD in well-defined language samples (cf. earlier important research in Matisoff 1978, Wilkins 1996, and Brown 2001). The goals of the project include suggesting a typology of innovative denomination strategies for a given set of body part concepts and investigating to what extent these can be explained by universal, genetic, areal, and cultural factors. The central tool here is a rigorous analytical framework for the threefold cross-classification of any type of lexical change, namely (i) cognitive relations (contiguity, metaphorical similarity, different taxonomic relations, etc.); (ii) formal devices (no formal change, i.e., only SEMANTIC change!, gender change, conversion, suffixation, compounding, etc.); (iii) "stratificational" properties (autochthonous processes vs. borrowing vs. calques, etc.) (cf. Blank 2003, Gévaudan forthcoming). A key issue of this approach, as opposed to the one taken in Wilkins (1996), is the systematic possibility of identifying one and the same cognitive relation as a constant behind an unlimited variety of formal devices (cf. "heterosemy", above) and/or various stratificational properties in different languages. For instance, the change EYEBROW ↔ EYELID, which is an example of contiguity (the relationships between a frame and its elements, or between different elements of one and the same frame, where the frame in this case is EYE), emerges from a simple semantic change in Hopi *puvùpwpi('at)*, from compounding in Latin *supercilium < cilium* (EYELID) + *super* (ABOVE), and from a blend in Romanian *sprânceană < * Latin *gena* (EYELID) × *supercilia* (EYEBROW). Ultimately, every target-source pair of concepts in the data is explicitly linked by a cognitive relation. *LexiType$_{Dia}$* makes use of diachronic data whenever available and, in certain cases, also relies on intragenetic comparison, i.e., comparative data in the same language or language family that are synchronic "traces" of diachronic processes (polysemy of the same lexical item, cognates, members of the same word family, etc.).

Starting with TARGET rather than with SOURCE concepts provides methodological advantages in testing hypotheses about possible source concepts from which a possible lexical change (of a particular semantic type) will be likely to take place. In addition, this perspective is closer to the language user's point of view and has therefore a firmer basis in the pragmatics of language change. In real discourse, speakers create semantic innovations not IN ORDER TO CHANGE an existing lexical item of their language, but IN ORDER TO EXPRESS a given (target) concept. Accordingly the project *LexiType_{Dia}* starts from specific body part target concepts expressed by lexical items in order to find out the relevant diachronic source concepts and does not consider lexical or grammatical items deriving from body part nouns (which is, in itself, another interesting issue).

The development of the terms for the target concepts EYELASH, EYELID, EYEBROW, and EYEBALL, all pertaining to the frame EYE (Koch forthcoming), will illustrate some of the generalizations achieved in *LexiType_{Dia}*. Here the languages in the global sample show a few recurrent denomination strategies. One, illustrated by Indonesian *bulu mata*, lit. 'hair eye', is to conceptualize EYELASH via taxonomic subordination, i.e., as a special kind of hair pertaining to the eye. The others are to conceptualize EYELASH via metaphorical similarity to a hairy object, e.g., BURNT END OF A WICK (*ukope* in Swahili) or PANICLE (*ripse* in Estonian), or to FRINGE (e.g., *eyelash* in English or *ögon-frans* in Swedish), in many cases within a compound where the modifying part designates the frame, EYE. Comparable strategies are also attested for EYELID and EYEBROW, e.g., 'eye flesh/skin' (taxonomic subordination) vs. 'eye lid' (metaphorical similarity) for EYELID, and 'eye hair' (taxonomic subordination) vs. 'eye edge/fringe' (metaphorical similarity) for EYEBROW, whereas EYEBALL is overwhelmingly based on metaphorical similarity ('eye ball/apple/fruit/nut, etc.'). Mihatsch (2005) suggests that the choice among these strategies is to a certain degree dependent on the grammatical properties of the nouns in the languages (which, in turn, may reflect major conceptualization preferences in nouns). Thus, languages in which nouns are obligatorily marked for number tend to develop terms for EYELASH, EYELID, and EYEBROW on the basis of metaphorical similarity, while those without obligatory number marking (e.g., Chinese, Tzeltal, Japanese) tend to resort to taxonomic subordination. These results fit in quite nicely with recent work on linguistic relativity, in particular, with psycholinguistic experiments on object sorting or classification carried out with speakers of languages with obligatory nominal plural marking (e.g., English) vs. those lacking this property and, in addition, resorting to obligatory numeral classifiers (e.g., Yucatec Maya and Japanese; cf. Lucy & Gaskins 2001, Imai & Gentner 1997). While the former tend to classify simple objects on the basis of their shape (e.g., considering a plastic comb with a handle as similar to a wooden comb of the same shape), the latter tend to rely on substance (e.g., considering a plastic comb with a handle as more similar to an-
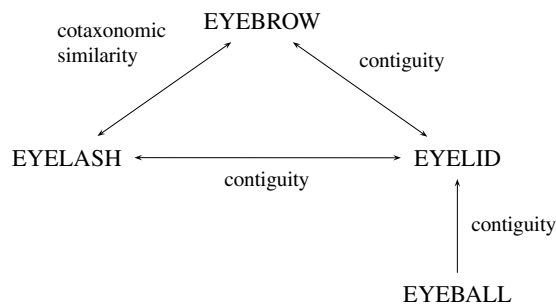
EYEBROW

cotaxonomic
similarity

contiguity

EYELASH ⟷ EYELID

contiguity

contiguity

EYEBALL

Figure 1. *Attested semantic shifts within the frame* EYE *(Koch forthcoming)*

other plastic object than to a wooden object of the same shape). In the same vein, metaphorical similarity in body part denomination is more sensitive to the shape of the compared entities, while taxonomic subordination is in these cases more sensitive to the substance.

Another interesting generalization concerns the frequently attested shifts in the denomination of these body parts summarized in Figure 1. While terms for many visible body parts seem to follow the "unidirectional law of synecdochic change" of Wilkins (1996) (i.e., visible part > visible whole, to the exclusion of visible whole > visible part, e.g., NAIL > FINGER > HAND > ARM), the changes involving the body part terms "around the eye" are bidirectional (apart from the apparently unidirectional development EYEBALL ⇒ EYELID).

In general, the terms for the body parts "around the eye" overwhelmingly develop via dependent conceptualization, i.e., via reference to other body parts (either to the frame, EYE, or to subparts within it). This is a cognitive constant, which, together with the frequent bidirectional semantic shifts among them cries, of course, for explanation.

The main results of *LexiType_{Dia}* will appear in Steinberg (in preparation), which systematizes and represents on maps the denomination strategies with respect to 26 target concepts in the conceptual domain HEAD within two language samples, a 14-language Indo-European sample and a representative world-wide 22-language sample. Thanks to the threefold, i.e., cognitive, formal, and stratificational, analytical framework described above, the world-wide sample yields insights which go far beyond (and partially correct) Wilkins (1996), while the Indo-European sample provides the possibility of studying genetic issues and contact phenomena (cf. also Mihatsch & Dvořák 2004 on the denomination of the concept FACE in a global 50-language sample).

Questions of polysemy, and in particular universal metaphoric and metonymic processes, and of the ways of representing it are a central concern of cog-

nitive semantics. The relevant knowledge accumulated within this theoretical framework is impressive; however, cognitive semantics has on the whole operated with a limited number of languages and has relatively modestly empirically-founded insights with regard to crosslinguistic semantic comparison. Systematic crosslinguistic research on semantic motivation along the lines reported in this section has therefore an important theoretical dimension.

2.2.2.   *Regularities in lexical motivation.*   Turning to the issue of regularities in lexical motivation, seen as an interaction of formal (morphological) and semantic aspects of motivation, the relevant questions include these: What meanings correspond to "more basic" vs. "regularly derived" words? What formal (primarily morphological) devices are there in a language for forming words from other words or lexical units from other lexical units (e.g., derivation, compounding)? What meaning relations can be expressed by these devices?

A large-scale *LT* publication addressing the issue of more basic vs. regularly derived words and how it interacts with the formal word-forming devices in a language is Nichols et al. (2004). It presents a typology of 80 languages based on their treatment of what the authors view as semantically basic and almost universally intransitive verbs such as SIT, FEAR, LAUGH, FALL and their transitive counterparts (all in all 18 pairs). The main question is whether the two sets of words are formally related to each other and if yes, how. Languages turn out to be fairly consistent in whether they treat intransitives as basic and transitives as derived by means of causative morphology (TRANSITIVIZING languages), whether they derive intransitives by means of anti-causative morphology (DETRANSITIVIZING languages), whether both intransitives and transitives are encoded by the same labile verb (NEUTRAL languages), or whether both intransitives and transitives have the same status (INDETERMINATE languages). There are also various further statistically significant generalizations on the "inner logics" of the types themselves, on their correlation to other linguistic phenomena (grammatical and lexical, e.g., alignment and voice alternations, complexity, aspect and Aktionsart), and on their distribution across the languages of the world – in the robust tradition of the standard modern large-scale typological research.

Another *LT* publication, Plank (2005), explores the formal devices for forming delocutive verbs across languages, i.e., "verbs derived from a base X which mean 'by saying or uttering "X" (to someone) to perform an act which is culturally associated with the meaning or force of X', where X is a variable ranging over types of things that can be said or uttered" (Plank 2005: 459).

2.3.   *What crosslinguistic patterns are there in lexicon-grammar interaction?*

Many crosslinguistic studies can be seen as contributions to lexical typology, understood as the search for typologically relevant features in the grammatical structure of the lexicon, or as typologically significant correlations between lexicon and grammar. They vary in how and to what extent they fit into the typological research framework and tradition(s), and in how and to what extent they consider the lexicon. Some are restricted to the lexicon-grammar interaction for a particular conceptual domain or lexical field or even for a particular lexical meaning, e.g., body part terms in adnominal possession and in special syntactic constructions (Chappell & McGregor (eds.) 1996; the literature is too extensive to be listed), kin terms in grammar (Dahl & Koptjevskaja-Tamm 2001), GIVE and argument linking (Haspelmath 2005a, Kittilä 2006), different classes of complement-taking verbs and the structure of complementation (Cristofaro 2003), WANT and the structure of desiderative clauses (Haspelmath 2005b, Khanina 2005), and many more: we have merely mentioned a few relevant things, chosen more or less arbitrarily. Curiously, although these issues are clearly anchored in the lexicon, this fact is not normally acknowledged and there is little awareness that the relevant studies focus on lexical phenomena and are rooted in particular lexical meanings. Veselinova's (2006) large-scale study of suppletion in verb paradigms is an interesting example of lexicon-grammar interaction: it shows that suppletion tends to be linked to verbs with particular lexical meanings (e.g., MOTION), with different meanings picked up by suppletion according to different grammatical categories (e.g., tense-aspect-mood, or imperative).

But many other traditional grammatical phenomena can also be viewed as lexical.

Consider, first of all, the issue of word classes or parts of speech across languages, which has figured prominently in *LT*. Although word classes are an example par excellence of the interaction – and significant correlation – between lexicon and grammar, the jump from individual language descriptions to large-scale crosslinguistic research tends to reduce lexical information to very few representatives for each presumptive word class (like BIG and GOOD for potential adjectives), not always systematically checked and/or completely comparable across the languages in the sample. In a number of crosslinguistic works word class behaviour is, fortunately, studied with more attention to lexical semantics and against the background of relatively fine-grained lexical distinctions, e.g., Dixon (1977) and later Dixon & Aikhenvald (eds.) (2004).

In *LT*, an instructive example of how the close attention to the behaviour of the items in a large lexical sample can lead to a revaluation of earlier analyses is offered in Evans & Osada's papers on Mundari (2005a, b) and in the commentary (Peterson 2005, Hengeveld & Rijkhoff 2005, Croft 2005). A radically

lexicon-based stance is also taken in Pustet's (2000) study of verb-adjective synonymies of the type to RESEMBLE vs. SIMILAR in seven languages (cf. also Pustet 2003). Pustet combines the generalizations and the various parameters suggested in earlier research on copulas and word class distinctions (primarily in Croft 1991 and Stassen 1997) and tests their compatibility by using the behaviour of 164 verb-adjective minimal pairs. The lexical items fall into a number of lexical classes based on various combinations of four different parameters, which together underlie lexical categorization. Pustet's principled sample (extended to 530–850 items in Pustet 2003) can most probably be used for further work on word classes and on their interaction with various grammatical categories. Taking a closer look at the lexical semantics and working with more fine-grained lexical classes will certainly give many more exciting and deeper insights into word class classification.

Crosslinguistic variation in categorization within major word classes also offers many challenges for research on significant patterns in the lexicon-grammar interaction. Consider the category of Aktionsart. Most modern research on Aktionsart has its roots in Vendler's (1967) verb classes (states, activities, accomplishments, achievements), whereas "[a]n urgent desideratum is the investigation of the role of lexicon, in particular the subcategorization of situation types", as Sasse puts it in his *LT* overview (2002: 263). Two later publications in *LT* offer promising steps in this direction. Tatevosov's (2002) study is based on a principled list of 100 "predicative meanings" (normally expressed by verbs or verb-based expressions) coming from several cognitive domains and covering the "basic" verbal lexicon; these are checked for all possible combinations with the verbal tense-aspect categories and their resulting meanings in four genetically unrelated languages: Bagwalal (Daghestanian, North East Caucasian), Mari (Finno-Ugric, Uralic), Tatar (Turkic), and Russian (Slavic, Indo-European). Already this comparison falsifies the common assumptions "that notions on which Vendlerian classes are based are logically universal, hence are not subject to crosslinguistic variation" and that verbs or verb phrases with "similar meanings" in different languages (i.e., translational equivalents) will belong to the same verbal class as their English equivalents (Tatevosov 2002: 322). Thus, "actionality" (used by Tatevosov instead of Aktionsart) turns out to be a parameter based on a universal set of elementary semantic distinctions, but allowing for different settings in different languages. Different languages show their own language specific sub-categorizations of the verb lexicon that can only be discovered via empirical investigations rather than taken for granted. A good example of the latter is found in Botne's (2003) crosslinguistic study of *die* and its correspondences in 18 languages. DIE is an example par excellence of Vendler's achievement verbs (telic, or bounded, and punctual) in that it refers to the acute point demarcating life and death. Botne shows that languages can differ in their lexicalization of the different stages in

the process leading to death, which in turn has important consequences for the Aktionsart categorization of the corresponding verb in a particular language. (We will return to the consequences of this insight in Section 4.)

It is with respect to lexicon-grammar interaction that lexical-typological research is of particular relevance for formal semantics. Formal semantics has always been concerned with compositionality, i.e., with how the meaning of a complex expression can be derived from the meanings of its parts. In accordance with this and following its own progress, formal semantics has been paying more and more attention to particular aspects of lexical meaning, e.g., the semantics of verbs in connection with the study of verbal aspect or of argument structure, semantics of nouns in connection with the count-mass distinction, quantification, possession, etc., which has resulted in an constantly increasing number of studies on diverse languages (to a large extent inspired by Bach et al. (eds.) 1995) and appearing in the major semantics outlets (*Journal of Semantics*, *Natural Language Semantics*, *Linguistics and Philosophy*, etc.).

It is now widely acknowledged in various linguistic theories that many grammatical phenomena are rooted in the lexicon. Lexicon-grammar interaction will surely provide lots of challenges for future lexical-typological research.

3.   The relevance of typological research into lexical semantics

Lexical semantic typology has different facets, not necessarily all captured by the questions posed in the preceding section. For instance, since the lexicon of a language is a dynamic and constantly changing complex structure, lexical-typological research has both synchronic and diachronic dimensions. Some of the questions relevant for historically oriented lexical typology have been mentioned in Section 2.2 (crosslinguistically recurrent patterns in semantic change, grammaticalization, and lexicalization processes). In addition, since the lexicons of most languages show different layers of origin with many words coming from "outside" (direct loans, loan translations, etc.), an interesting aspect of historical lexical typology is the search for crosslinguistically recurrent patterns in contact-induced lexicalization and lexical change – e.g., differences in borrowability among the different parts of the lexicon and the corresponding processes in the integration of new words, or patterns of lexical acculturation (i.e., how lexicons adjust to new objects and concepts). Important contributions here include Brown (1999) and the ongoing project on *Loanword Typology* at the Max Planck Institute for Evolutionary Anthropology (coordinated by Martin Haspelmath and Uri Tadmor, http://www.eva.mpg.de/lingua/files/lwt.html), none of which has so far been reflected in *LT*. Research on crosslinguistically recurrent patterns of semantic motivation and on their likelihood to arise independently in different languages, to be inherited, and to diffuse via language

contact (Section 2.2) is also highly relevant for historical linguistics and, in particular, for areal linguistics, as has already been demonstrated, for instance, in Enfield's (2003) study of ACQUIRE in mainland Southeast Asia and in Evans & Wilkins' (2001) study of verbs for hearing vs. seeing in Australian languages.

In addition to the issues we have discussed here there are also interesting questions on the interaction between lexicon and phonology, on overall principles of taxonomic categorization, or the organization of the lexicon, and many others. A particularly important question is whether it is possible to discern central, essential characteristics of the lexicon that would have important repercussions for the language as a whole, would provide a good basis for crosslinguistic comparison, and might even have important implications outside the language. The most influential candidate for this role has undoubtedly been Talmy's "lexicalization-pattern" theory for motion and causation events (1985, 1991, 2000). In general, Talmy's studies of space, causality, force dynamics, differential semantic packaging in closed-class and open-class items have had a profound impact on the semantic research of particular languages and crosslinguistically. Especially the binary distinction between VERB-FRAMED and SATELLITE-FRAMED languages (Talmy 1991), further building on Talmy's original three-fold classification of languages with respect to the encoding of spatial events, has had wide-ranging effects, both on semantic issues in general and outside linguistics. In particular, the rich research by Dan Slobin and his associates has demonstrated multiple connections between the lexicalization type of a certain language and its discourse organization, child language acquisition of the domain, gestures and "thinking for speaking", i.e., mild cognitive effects of linguistic relativity (e.g., Slobin 2003, Slobin & Bowerman 2007, Kita & Özyürek 2003). In short, a language's lexicalization pattern in the motion domain is believed to belong to its typologically relevant properties, even though there remain many questions for further research on the lexicalization-pattern typology itself and on its empirical underpinnings (e.g., cf. Wälchli in preparation a).

A number of other crosslinguistic studies investigate the "lexical profile" of a language, aiming at essential characteristics of the lexicon and its implications. As mentioned in Section 2.2, Nichols et al. (2004) are explicit about the relevance of their classification into transitivizing, detransitivizing, neutral, and indeterminate languages for other linguistic phenomena such as alignment and voice alternations, complexity, aspect and Aktionsart. Kibrik (2003) suggests to approach lexico-semantic profiles of individual languages and, further, lexical typology by identifying types of meanings that are rendered in a non-derived way (by bare verb roots) and further exploring individual derivation processes and the processes of meaning construction. These together can further account for various grammatical peculiarities of particular languages, which Kibrik demonstrates using the example of the striking grammatical prop-

erties of Athabascan languages. Other suggestions include Viberg's (2006) method to use frequencies of occurrence to identify the basic or nuclear verbs that together make up the lexical typological profile of the verb lexicon in a particular language (in this case, one of several European languages), as well as the matrix for measuring the semantic and formal motivation profiles of individual languages exploited in the Tübingen project *LexiType_{Syn}* on the empirical basis of speaker judgments (Koch & Marzo 2007). Each of the latter three suggestions has at present only been applied to one or two languages and is still awaiting further testing. In any case, without systematic crosslinguistic research on the lexicon we will not be able to tackle what Evans (forthcoming) calls the issue of "semiotic ecology", i.e., of how semantic choices made in one linguistic subsystem affect those in others.

This brings us to the core concern in this stocktaking issue of *LT* – the relation between typology, in this case lexical semantic typology, and theoretical linguistics. As in most of the other contributions to this issue we would like to conclude that there does not have to be any antagonism between the two: good lexical semantic typology is good theoretical semantics, and good theoretical semantics is good semantic typology. So what is the theoretical contribution of lexical semantic typology? In a nutshell, it is primarily the omnipresent issue of universality vs. language-specificity, or variability, in this case in lexical semantic phenomena, and of possible explanations for the attested facts. This applies to different kinds of phenomena. Thus, for instance, to what extent are languages free in carving the world (at an infinite and arbitrary number of places) by lexical items? Are there limits to the meanings that can be expressed within a word? To what extent are languages free to combine different meanings within one and the same lexeme?

As everywhere in good typology, the sine qua non of lexical semantic typology is trying to answer such questions as according to what parameters a specific phenomenon can vary across languages, in what patterns these parameters (co-)occur, and what generalizations can be made about attested vs. possible patterns. Already this step requires the solid integration of theoretical and factual knowledge (or, if you like, good interaction between theoretical and descriptive brains): crosslinguistic facts and crosslinguistic generalizations can win or lose depending on the assumptions of what counts as meaning and/or the metalanguage adopted in a particular theory, and, conversely, crosslinguistic facts and crosslinguistic generalizations can guide theoretical semanticists in their search for an appropriate metalanguage and for the essence of meaning in general.

Let's mention just a few choices and keep in mind that there are many more. For instance, which option is more appropriate for capturing crosslinguistic generalizations in a particular case or more generally – to analyze meanings in terms of classical Aristotelian categories with necessary and sufficient mean-

ing components, as has been practiced in different versions of componential analysis and/or in formal semantics, or in terms of prototype categories, in the tradition of the mainstream cognitively oriented approaches to semantics? Is there some common "alphabet of human thought" in terms of which all semantic categories in all languages can be composed (the position taken in the Natural Semantic Metalanguage approach), or do different languages build up their semantic categories from different basic stocks, or is there a golden mean, so that there are both universal semantic concepts and language-particular basic stocks? If there are universal semantic concepts, what are they and how elementary should they be? Cf. the various ontologies suggested in formal semantics (e.g., Bittner 2006), including the different treatments of mass vs. count nouns (e.g., Chierchia 1998; cf. also a further "decomposition" of this distinction in Behrens & Sasse 2003). How should meaning be represented to start with and what is the appropriate or allowed level of generality or vagueness here? There are lots and lots of options here, with relatively little communication among the different schools of semantics.

The next step in good typological work, including lexical semantic typology, is to advance explanations for the attested patterns – and they can be environmental (rooted in the properties of the real world), biological (shaped by human perceptual and cognitive predispositions or simply innate), social, or cultural. This is, again, an enterprise that requires a high level of theoretical awareness. How do we go about to recognize a crosslinguistically recurrent polysemy pattern as an instance of a particular conceptual metaphor or a particular image schema transformation in the tradition of the cognitive semantics? How does the crosslinguistically attested variation in the observed lexicon-grammar interaction agree with the compositionality considerations that are central for formal semantics? Again, there are lots and lots of theoretical decisions to be made and theoretical insights to be gained.

Theoretical linguistics in general and theoretical semantics in particular often talks about universality. For many such claims, robust empirical crosslinguistic research on the lexicon is what is badly needed.

4.   Urgent methodological problems for lexical typology

On the methodological side, the basic and most urgent desiderata for lexical typology are: (i) to refine the existent and to develop new methods of data collection; (ii) to improve standards in the crosslinguistic identification of the phenomena under study and in their (semantic) analysis; (iii) to achieve a reasonable consensus on the meta-language for semantic explications and on the ways of representing meanings.

As to the methodology of data collection, there are striking differences here between "traditional" mainstream typology and lexical typology. Morphosyn-

tactic typology has been largely dependent on secondary data sources, with reference grammars as the typological data source par excellence, in many cases complemented by sporadic consultations with native speakers and/or language experts. Recent decades have seen an elaboration of other techniques for typological data collection, but reference grammars still keep their leading position in the field. Studies in morphosyntactic typology are typically a "one researcher's job": even when data collection involves, say, filling in questionnaires, the people doing that part of the job normally count as consultants rather than co-authors. (There are, of course, exceptions here: e.g., the tradition of the Leningrad/St.Peterburg Typological School, or the numerous collections edited by Aikhenvald and Dixon.)

By contrast, secondary sources are of marginal importance for lexical typology in most of its aspects. Consider the three groups of questions that have primarily figured in the present article – categorization within conceptual domains, semantic and formal motivation, and lexicon-grammar interaction. Relevant data are normally scattered across different kinds of secondary sources: a typical thesaurus dictionary might provide information on categorizations within conceptual domains, but hardly anything on the polysemy patterns and on other formal-semantic relations within word families – information that can sometimes be found in a "normal" dictionary. Some information on lexicon-grammar interaction might find its way into a reference grammar (which seldom lists all the words showing a particular grammatical behaviour), some might be given in a dictionary. In addition, dictionaries are a favourite object of ridicule in theoretical work on semantics and lexicography, providing vague and circular definitions. Ideally, a desideratum would be to have a source that for every word in a language would give a precise meaning definition, show both its exact relations to other words, and define its grammatical properties. There are a few attempts on the market towards this desideratum – e.g., the long tradition of the "interpretational-combinatorial dictionary" with roots in the Moscow School of Semantics (cf. Iordanskaja & Paperno 1996 for an excellent treatment of Russian body part terms in this tradition), the Berkeley FrameNet project based on Frame Semantics (http://framenet.icsi.berkeley.edu/index.php?option=com_frontpage&Itemid=1), or the expanding enterprise of Word Net for several European languages (http://www.globalwordnet.org/). However, the lexicon for most languages of the world is – and will remain – relatively poorly described, at least for the purposes of consistent crosslinguistic research.

Most lexical-typological research is therefore in need of constantly inventing, testing, and elaborating its methods of data collection. Two of the major stumbling stones for any semantic work are the problem of what can be meant by meaning – denotation/extension vs. sense/descriptive meaning/intension – and the problem of polysemy/semantic generality. Although for many seri-

ous semanticists, lexicographers, and lexicologists semantic analysis stands for coming to grips with descriptive meanings, or senses, the enterprise is far from obvious even for the researcher's native tongue; for other languages it easily gets insurmountable. As a consequence, much crosslinguistic comparison is based on meanings defined as denotations. In other words, the question of "What meanings can be or cannot be expressed by single words in a language?" often amounts to "What are possible/impossible denotational ranges of single words in a language?". There are various methods for eliciting, defining, and evaluating expression-denotation couples (pictures, videoclips, Munsell colour chips, etc.), but these are not easily transmittable among different research areas, even among those that ask comparable questions. To take a simple case, visual stimuli for eliciting words referring to CUTTING AND BREAKING EVENTS can certainly serve as a model for research on some other conceptual domains involving dynamic situations with clearly visible actions and results (say, DRESSING/UNDRESSING or PUTTING). But already moving to domains based on other perceptual modalities is far from trivial: SOUNDS, TEMPERATURE, TASTE are still awaiting good data collection techniques and guidelines. There are various reasons for why such denotation-based approaches are not always sufficient, including Quine's (1960: 29) famous "gavagai" problem: If a person whose language you don't know says "gavagai" when a white rabbit appears in front of you both, how can you be sure about what (s)he really means? Another big problem is that many meanings – or, rather, many conceptual domains – hardly lend themselves to being investigated via denotation-based techniques: for instance, how do you get at the meaning of *think* or *love*? (Cf. also Evans & Sasse 2007, but also Pavlenko 2002 for a comparison of emotional descriptions in Russian and English narratives elicited through the same short film.)

The comparison of parallel texts (translations of one and the same text) is a relatively new and expanding method for data collection in crosslinguistic work. The extent to which parallel texts can be used in lexical-typological research is, of course, extremely dependent on the object of study and on the genre of the parallel texts and is best suited for frequent phenomena. Thus, most available texts in many languages are the various versions of the New Testament and the Universal Declaration of Human Rights. While motion verbs frequently occur in the former and generic statements in the latter, which has been used by Wälchli in his large-scale investigations of motion verbs and co-compounding (Wälchli in preparation a, b; 2005), these sources will be of restricted value for the study of, say, PAIN expressions, even though such examples do occasionally occur in the New Testament.

Denotation-based techniques for data collection, parallel corpora, and questionnaires (where "meaning" amounts to "translational equivalence") effectively neglect the issue of semantic generality/polysemy. They often provide a number of contexts , or an "etic grid" for capturing (logically) possible dis-

tinctions within a domain, with the result that the meaning of a word can easily become reduced to the set of its uses (an "etic definition"). The logical step from an ETIC definition to an EMIC one (i.e., finding out the commonalities behind the different uses and, ideally, arriving at a reasonable characterization of the descriptive meaning) goes hand in hand with deciding what constitutes one meaning, i.e., distinguishing between semantic generality and polysemy (cf. Evans forthcoming for the discussion of etic vs. emic definitions in semantic typology).

Finally, word lists, as we have seen, may well be used for some purposes (e.g., for checking the word class categorization of "property" words or the Aktionsart categorization of verbs), but are of marginal value when too little is known about the lexical meaning of phenomena under consideration or when the phenomena involve too many language-specific lexical idiosyncrasies. Consider pluralia tantum, nouns that only occur in the plural form, like *scissors*. These are very unevenly distributed across languages: some have none (e.g., Hungarian), others just a few (e.g., Daghestanian languages), while still others have dozens or even hundreds (e.g., Baltic languages). In Koptjevskaja-Tamm & Wälchli (2001) two principled samples of lexemes that are encoded by pluralia tantum nouns in Lithuanian vs. Russian are used for collecting comparable data across 40 European languages. The study was part of a larger investigation with the goal of establishing contact phenomena in the Northeastern part of Europe and was based on the hypothesis that the degree of overlapping in the distribution of pluralia tantum nouns across languages could be used as a measure for their contacts. The two samples turned out to be useful for this particular end, due to the fact that the distribution of pluralia tantum in a language is highly idiosyncratic. But the same fact causes difficulties for crosslinguistic studies of pluralia tantum in general: they often occur in comparable domains (e.g., heterogeneous substances like *leftovers*, diseases like *measles*, festivities like *Weihnachten* 'Christmas' in German), but are very language-specific when it comes to the lexical meanings, which precludes the use of a consistent word list for crosslinguistic data collection.

We think that successful lexical-typological research can only rarely be a "one researcher's job", but should in most cases build on collaborative work involving language experts, semanticists, typologists (and, possibly, other specialists as well – as an example of interaction of cognitive science and lexical typology cf. Gaume et al. forthcoming). The work on lexical universals in the Natural Semantic Metalanguage tradition (Goddard & Wierzbicka 1994), the different domain-categorization studies co-ordinated from the Max Planck Institute in Nijmegen (Levinson & Meira 2003; Majid et al. 2006, forthcoming), the project on AQUA-MOTION verbs directed by Moscow linguists (Majsak & Rakhilina 2006) are all examples of semantic-typological research based on the methodology that had been elaborated, tested, and improved by the group

of language experts, who have further collected and analyzed data in close collaboration with native speakers.

The issue of data collection is intimately related to the issue of crosslinguistic identification, which is a key concern for typological research in general. We have to be sure that we compare like with like, rather than apples with pears. However, another key concern for typological research is to find a reasonable level of abstraction, at which the richness of language-specific details can be reduced to manageable patterns. The two concerns are interrelated in various ways; most importantly, what counts as "like and like" is often dependent on the research object and goal. We will illustrate this by several examples from the studies on motivation.

Recall the above-mentioned discussion of universal lexicalization in Goddard (2001) that amounted to there being very few meanings that can easily translate among languages, in particular if precise semantic identity is required. Crosslinguistic research on motivation considers (or explicitly discusses) this fact very seldom; the different meanings expressed by one and the same lexeme or by derivationally related lexemes are normally taken for granted, as self-evident, and easily identified across languages and in a particular language. Consider the grammaticalization paths HAND ⇒ FIVE, attested in various languages, including Samoan (Polynesian, Austronesian) and Turkana (Nilotic, Nilo-Saharan), and HAND ⇒ POSSESSION, attested, among other languages, in Kono (Mande, Niger-Congo), Zande (Ubangian, Niger-Congo) (Heine & Kuteva 2002: 166–167), and Estonian (Finno-Ugric, Uralic). Since all these languages use the same lexeme for HAND and ARM, how would we know that it is HAND that has been grammaticalized, rather than, say, ARM (EXCLUDING HAND) or ARM (INCLUDING HAND)? A strict proof for the case would include, first, arguments in favour of polysemy HAND/ARM rather than semantic generality in all these languages and, second, evidence for the grammaticalized meanings being based on HAND to the exclusion of ARM. We are not aware of serious attempts to do anything along these lines. The interpretation of these particular examples and the postulation of these particular semantic links are, most probably, founded on common sense and intuition rather than on strict argumentation, and on parallels with other languages which clearly distinguish between HAND and ARM.

Likewise, almost none of the 18 verb pairs used in Nichols et al. (2004) and viewed by the authors as semantically basic and almost universal, belongs to Goddard's (2001) list of lexico-semantic universals (with some, like SIT and FEAR, being explicitly excluded from it). The exact semantics and precise semantic identity of the verbs on the list is, however, not a point here: the 18 verb pairs have been chosen on pragmatic grounds, as representing certain combinations of general parameters, corresponding to frequently encoded situations, and having approximate translational equivalents in many languages.

Stricter requirements on semantic comparability would in fact create obstacles for achieving the principal objective of the study. Obtaining one-word expressions with the same semantics for 18 events (and, in addition, representing the various combinations of interesting parameters) in 80 languages is hardly conceivable, while achieving the right semantics by means of word combinations would most probably conceal the basic derivational relations.

In other cases the relatively low degree of semantic precision in the definitions is less justified and can be an obstacle for deeper insights and effective crosslinguistic comparison. Among the various grammaticalization paths for motion verbs several are often defined as starting with COME and GO (cf. Heine & Kuteva 2002). The English verbs *come* and *go* as semantic metalabels are not totally felicitous; among other things, they encode the deictic distinction between centripetal and centrifugal motion, absent from many languages of the world (see Ricca 1993), and neutralize the distinction between motion on foot vs. in a vehicle (cf. also Goddard 2001: 28). Descriptions like COME ⇒ CONTINUOUS or GO ⇒ HABITUAL are therefore too vague for understanding the underlying logic of the development; they do, however, serve as preliminary crude classifications and as guidelines for future research.

An issue that is rarely explicitly acknowledged and discussed in crosslinguistic studies on lexicon-grammar interactions concerns possible implications of such variation for the lexical semantics of the items under consideration. Consider Botne's (2003: 276) conclusions of his crosslinguistic study on the Aktionsart categorization of the correspondences to *die*:

> This small, exploratory study has shown that [...] achievement verbs, though unified by the punctual, culminative nature of their nucleus, may be conceptualized in different languages as encoding durative preliminary (onset) or postliminary (coda) phases in addition to the punctual nucleus. Consequently, die verbs frequently have a complex temporal structure and do not simply encode a point of transition [...] [T]he same 'concept' will not necessarily be encoded with the same phases in every language. Consequently, appropriate crosslinguistic comparison and analysis of these kinds of verbs will perforce require a close analysis of a particular verb in each language.

But if languages differ as to which of the phases leading to death they encode in their DIE verbs, it is rather doubtful that we can view them as encoding "the same concept"! The crosslinguistic identification of phenomena based on approximate rather than precise semantic identity can be justified when the primary focus of crosslinguistic research is not on the lexical semantics per se. However, it is also reasonable to take the next step and use crosslinguistic variation in grammatical behaviour as evidence for lexical-semantic differences.

It should be mentioned here that grammatical typology on the whole hardly ever cares about precise semantics: the only prerequisite is that we can roughly

identify linguistic phenomena across languages via certain conditions that they
have to meet, e.g., via a certain function that has to be expressed by a construc-
tion. Thus, for instance, a possessive NP is recognized by its ability to refer to
legal ownership (*Peter's bag*), to kin relations (*Peter's son*), or to relations be-
tween a person and his body parts (*Peter's leg*). The fact that the same construc-
tion in English can occasionally refer to temporal and local relations (*yester-
day's magazine, London's museums*), whereas many other languages are much
more restrictive in this respect, is of marginal interest for the crosslinguistic
identification of possessive NPs themselves. There are, of course, certain limits
to the semantic vagueness that can underlie systematic crosslinguistic identifi-
cation. At the opposite end from the careful lexical-typological investigations
in the domains of COLOUR, BODY, SPACE, AQUA-MOTION etc. (cf. above) are, for
instance, some studies arguing for the universality of various metaphors, like,
for instance, ANGER IS HEAT (Kövecses 1996). What can probably be done is
to test some of their concrete manifestations, e.g., whether the words for anger
(and other emotions) can be described by temperature terms.

Finally, a crucial complication for crosslinguistic studies on the lexicon –
and, further, in most crosslinguistic research where meaning is involved – is
created by the problem of a consistent meta-language for representing mean-
ings within and across languages, which, in turn, is related to the general enor-
mous gap between theoretical semantics and theoretical lexicology, on the one
hand, and actual lexicographic practice, on the other.

The interaction between descriptive practice and grammatical typology has
been profitable for both. On the one hand, linguistic typology has to a large
extent been focusing on issues that are traditionally found in grammars. On
the other hand, the cumulative typological experience has had a noticeable
positive impact on the general level and quality of modern grammatical de-
scriptions of less well known languages, as compared to earlier descriptions.
This is to a large extent due to the "guidelines" for uncovering various phe-
nomena in new languages that have explicitly or implicitly been formulated
in typological research (as, e.g., in Shopen (ed.) 1985, a collection of typo-
logical surveys specifically designed for these purposes). These descriptions,
in turn, provide a better empirical basis for new typological generalizations.
Let's hope for a similar profitable interaction between lexical-typological re-
search and descriptive practice in the near future, where initial steps have
already been taken in the shape of elicitation guidelines as developed at the
Max Planck Institute in Nijmegen and by the AQUA-MOTION project in Moscow
(http://aquamotion.narod.ru).

Given the various methodological problems and our general poor knowledge
of lexicon-rooted phenomena, a large part of lexical-typological research in
the nearest future will, most probably, have to be restricted to small numbers
of languages. This might be a disappointment for some typologists, but it is

undoubtedly better to have high-quality research on just a few languages than poor investigations on scores. There is also ample evidence that even closely related languages can manifest striking differences in their lexical organization, which also has far-reaching consequences for sampling decisions in lexical typology.

There are many difficult issues in crosslinguistic research into the lexicon. What seems to be the best strategy for the moment is to explore various ways to go. Let us hope that *Linguistic Typology* will be a reliable travel companion in these explorations.

Received: 15 March 2007   *Stockholms universitet*
Revised: 5 April 2007   *Llacan, Inalco, CNRS*
                                                            Universität Tübingen

Correspondence addresses: (Koptjevskaja-Tamm) Department of Linguistics, Stockholm University, 106 91, Stockholm, Sweden; e-mail: tamm@ling.su.se; (Vanhove) Llacan, CNRS, B.P. 8, 7, rue Guy Môquet, 94801 Villejuif Cedex, France; e-mail: vanhove@vjf.cnrs.fr; (Koch) Romanisches Seminar, Universität Tübingen, Wilhelmstr. 50, 72074 Tübingen, Germany; e-mail peter.koch@uni-tuebingen.de

References

Aikhenvald, Alexandra & R. M. W. Dixon (eds.) (2002). *Word: A Cross-linguistic Typology.* Oxford: Oxford University Press.

Auwera, Johan van der (1998). Phrasal adverbials in the languages of Europe. In Johan van der Auwera (ed.), *Adverbial Constructions in the Languages of Europe*, 25–145. Berlin: Mouton de Gruyter.

— (2001). On the lexical typology of modals, quantifiers, and connectives. In István Kenesei & Robert M. Harnish (eds.), *Perspectives on Semantics, Pragmatics, and Discourse: A Festschrift for Ferenc Kiefer*, 173–186. Amsterdam: Benjamins.

Auwera, Johan van der & Vladimir Plungian (1998). Modality's semantic map. *Linguistic Typology* 2: 79–124.

Bach, Emmon, Eloise Jelinek, Angelika Kratzer, & Barbara H. Partee (eds.) (1995). *Quantification in Natural Languages*. Dordrecht: Kluwer.

Behrens, Leila & Hans-Jürgen Sasse (2003). *The Microstructure of Lexicon-Grammar-Interaction: A Study of 'Gold' in English and Arabic*. München: Lincom Europa.

Bickel, Balthasar & Johanna Nichols (2007). AUTOTYP. http://www.uni-leipzig.de/~autotyp/projects/wd_dom/wd_dom.html

Bittner, Maria (2006). Ontology for human talk and thought (not robotics). *Theoretical Linguistics* 32: 45–56.

Blank, Andreas (2003). Words and concepts in time: Towards diachronic cognitive onomasiology. In Regine Eckardt, Klaus von Heusinger, & Christoph Schwarze (eds.), *Words in Time: Diachronic Semantics from Different Points of View*, 37–65. Berlin: Mouton de Gruyter.

Botne, Robert (2003). To die across languages. *Linguistic Typology* 2: 233–278.

Boyeldieu, Pascal (forthcoming). From semantic change to polysemy: The cases of 'meat/animal' and 'drink'. In Vanhove (ed.) forthcoming.

Brown, Cecil H. (1999). *Lexical Acculturation in Native American Languages*. New York: Oxford University Press.

— (2001). Lexical typology from an anthropological point of view. In Haspelmath et al. (eds.) 2001, Volume 2, 1178–1190.

Chappell, Hilary & William McGregor (eds.) (1996). *The Grammar of Inalienability: A Typological Perspective on Body Part Terms and the Part-Whole Relation*. Berlin: Mouton de Gruyter.

Chierchia, Gennaro (1998). Reference to kinds across languages. *Natural Language Semantics* 6: 339–405.

Corbett, Greville G. (2007). Canonical typology, suppletion and possible words. *Language* 83.

— (forthcoming). Higher order exceptionality in inflectional morphology. In Horst J. Simon & Heike Wiese (eds.), *Expecting the Unexpected: Exceptions in Grammar*. Berlin: Mouton de Gruyter.

Cristofaro, Sonia (2003). *Complementation*. Oxford: Oxford University Press.

Croft, William (1991). *Syntactic Categories and Grammatical Relations*. Chicago: University of Chicago Press.

— (2005). Word classes, parts of speech, and syntactic argumentation. *Linguistic Typology* 9: 431–441.

Cruse, Alan D., Franz Hundsnurscher, Michael Job, & Peter Rolf Lutzeier (eds.) (2002/05). *Lexicology: An International Handbook on the Nature and Structure of Words and Vocabularies*. 2 volumes. Berlin: Walter de Gruyter.

Dahl, Östen & Maria Koptjevskaja-Tamm (2001). Kinship in grammar. In Irène Baron, Michael Herslund, & Finn Sørensen (eds.), *Dimensions of Possession*, 201–225. Amsterdam: Benjamins.

Davies, Ian, Tat'jana Sosenskaja, & Greville G. Corbett (1999). Colours in Tsakhur: First account of the basic colour terms of a Nakh-Daghestanian language. *Linguistic Typology* 3: 179–207.

Dixon, R. M. W. (1977). Where have all the adjectives gone? *Studies in Language* 1: 19–80.

Dixon, R. M. W. & Alexandra Aikhenvald (eds.) (2004*). Adjective Classes*. Oxford: Oxford University Press.

Duvignau, Karine & Bruno Gaume (2004). Linguistic, psycholinguistic and computational approaches to the lexicon: For early verb-learning. *ESSCS Journal, Journal of the European Society for the Study of Cognitive Systems* 6: 255–269.

Enfield, Nick J. (2003). *Linguistic Epidemiology*. London: Routledge, Curzon.

Evans, Nicholas (forthcoming). Semantic typology. In Jae Jung Sung (ed.), *The Oxford Handbook of Typology*. Oxford: Oxford University Press.

Evans, Nicholas & Toshiki Osada (2005a). Mundari: The myth of a language without word classes. *Linguistic Typology* 9: 351–390.

— (2005b). Mundari and argumentation in word class analysis. *Linguistic Typology* 9: 442–457.

Evans, Nicholas & Hans-Jürgen Sasse (2007). Searching for meaning in the Library of Babel: Field semantics and problems of digital archiving. *Archives and Social Studies: A Journal of Interdisciplinary Research* 1.

Evans, Nicholas & David P. Wilkins (2000). In the mind's ear: The semantic extensions of perception verbs in Australian languages. *Language* 76: 546–592.

François, Alexandre (forthcoming). Semantic maps and the typology of colexification: Intertwining polysemous networks across languages. In Vanhove (ed.) forthcoming.

Gaume, Bruno, Karine Duvignau, & Martine Vanhove (forthcoming). Semantic associations and confluences in paradigmatic networks. In Vanhove (ed.) forthcoming.

Gévaudan, Paul (forthcoming). *Klassifikation lexikalischer Entwicklungen: Semantische, morphologische und stratische Filiation.* Tübingen: Niemeyer.

Gil, David (2001). Quantifiers. In Haspelmath et al. (eds.) 2001, Volume 2, 1275–1294.

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

Goddard, Cliff & Anna Wierzbicka (eds.) (1994). *Semantic and Lexical Universals.* Amsterdam: Benjamins.

Haspelmath, Martin (1997). *Indefinite Pronouns*. Oxford: Oxford University Press.

— (2003). The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In Michael Tomasello (ed.), *The New Psychology of Language*, Volume 2, 211–242. Mahwah, N.J.: Erlbaum.

— (2005a). Ditransitive constructions: The verb 'give'. In Haspelmath et al. (eds.) 2005, 426–429.

— (2005b). 'Want' complement clauses. In Haspelmath et al. (eds.) 2005, 502–505.

—

Haspelmath, Martin, Matthew Dryer, David Gil, & Bernard Comrie (eds.) (2005). *The World Atlas of Language Structures*. Oxford: Oxford University Press.

Haspelmath, Martin, Ekkehard König, Wulf Oesterreicher, & Wolfgang Raible (eds.) (2001). *Language Typology and Language Universals: An International Handbook*. 2 volumes. Berlin: Walter de Gruyter.

Haspelmath, Martin & Uri Tadmor (eds.) (in preparation). *Lexical Borrowability.* 2 volumes. Berlin: Mouton de Gruyter.

Heine, Bernd & Tanja Kuteva (2002). *World Lexicon of Grammaticalization.* Cambridge: Cambridge University Press.

Hengeveld, Kees & Jan Rijkhoff (2005). Mundari as a flexible language. *Linguistic Typology* 9: 406–431.

Imai, Mutsumi & Dedre Gentner (1997). A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition* 62: 179–200.

Iordanskaja, Lidija & Slava Paperno (1996). *A Russian-English Collocational Dictionary of the Human Body*. Columbus, Ohio: Slavica. Available at http://LexiconBrige.com/

Khanina, Olesya (2005). Jazykovoe oformlenie situacii želanija: opyt tipologičeskogo issledovanija. [Linguistic encoding of 'wanting': A typological approach.] Doctoral dissertation, Moscow State University.

Kibrik, Andrej A. (2003). Lexical semantics as a key to typological peculiarities. Paper at the *Workshop on Lexical/Semantic Typology* at the ALT conference in Cagliari, September 2003.

Kita, Sotaro & Aslı Özyürek (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* 48: 16–32.

Kittilä, Seppo (2006). Anomaly of the verb 'give' explained by its high (formal and semantic) transitivity. *Linguistics* 44: 569–609.

Koch, Peter (2001). Lexical typology from a cognitive and linguistic point of view. In Haspelmath et al. (eds.) 2001, Volume 2, 1142–1178.

— (forthcoming). Cognitive onomasiology and lexical change: around the eye. In Vanhove (ed.) forthcoming.

Koch, Peter & Marzo, Daniela (2007). A two-dimensional approach to the study of motivation in lexical typology and its first application to French high-frequency vocabulary. *Studies in Language* 31: 259–291.

Koptjevskaja-Tamm, Maria (forthcoming). Approaching lexical typology. In Vanhove (ed.) forthcoming.

Koptjevskaja-Tamm, Maria & Bernhard Wälchli (2001). The Circum-Baltic languages: An areal-typological approach. In Östen Dahl & Maria Koptjevskaja-Tamm (eds.), *Circum-Baltic Languages: Their Typology and Contacts*, Volume 2, 615–750. Amsterdam: Benjamins.

Kövecses, Zoltán (1996). Anger: Its language, conceptualization, and physiology in the light of cross-cultural evidence. In John Taylor & Robert MacLaury (eds.), *Language and the Cognitive Construal of the World*, 181–196. Berlin: Mouton de Gruyter.

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

Lehmann, Christian (1990). Towards lexical typology. In William Croft, Keith Denning, & Suzanne Kemmer (eds.), *Studies in Typology and Diachrony: Papers Presented to Joseph H. Greenberg on his 75th Birthday*, 161–185. Amsterdam: Benjamins.

Lehrer, Adrienne (1992). A theory of vocabulary structure: Retrospectives and prospectives. In Manfred Pütz (ed.), *Thirty Years of Linguistic Evolution: Studies in Honour of René Dirven on the Occasion of his Sixtieth Birthday*, 243–256. Amsterdam: Benjamins.

Levinson, Stephen & Sergio Meira (2003). 'Natural concepts' in the spatial topological domain – adpositional meanings in cross-linguistic perspective: An exercise in semantic typology. *Language* 79: 485–516.

Lucy, John & Suzanne Gaskins (2001). Grammatical categories and the development of classification preferences: A comparative approach. In Stephen Levinson & Melissa Bowerman (eds.), *Language Acquisition and Conceptual Development*, 257–283. Cambridge: Cambridge University Press.

Majid, Asifa, Melissa Bowerman, Miriam van Staden, & James S. Boster (forthcoming). The semantic categories of 'cutting and breaking' events across languages. *Cognitive Linguistics*.

Majid, Asifa, Nick Enfield, & Miriam van Staden (2006). Parts of the body: Cross-linguistic categorization. *Language Sciences* 28: 137–360.

Majsak, Timur A. (2005). *Tipologija grammatikalizacii konstrukcij s glagolami dviženija i glagolami pozicii.* [A typology of grammaticalization of constructions with motion and position verbs.] Moskva: Jazyki slavjanskix kul'tur.

Majsak, Timur & Ekaterina Rakhilina (eds.) (2006). *Glagoly dviženija v vode: Leksičeskaja tipologija.* [Aqua-motion verbs: A lexical typology.] Moskva: Indrik.

Masson, Michel (forthcoming). Cats and bugs: Some remarks about semantic parallelisms. In Vanhove (ed.) forthcoming.

Matisoff, James (1978). *Variational Semantics in Tibeto-Burman: The 'Organic' Approach to Linguistic Comparison*. Philadelphia: Institute for the Study of Human Issues.

Mihatsch, Wiltrud (2005). Experimental data vs. diachronic typological data: Two types of evidence for linguistic relativity. In Stephan Kepser & Marga Reis (eds.), *Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives*, 371–392. Berlin: Mouton de Gruyter.

Mihatsch, Wiltrud & Boštjan Dvořák (2004). The concept FACE: Paths of lexical change. In Wiltrud Mihatsch & Reinhild Steinberg (eds.), *Lexical Data and Universals of Semantic Change*, 231–254. Tübingen: Stauffenburg.

Newman, John (ed.) (1997). *The Linguistics of Giving*. Amsterdam: Benjamins.

— (ed.) (2002). *The Linguistics of Sitting, Standing and Lying.* Amsterdam: Benjamins.

Nichols, Johanna, David A. Peterson, & Jonathan Barnes (2004). Transitivizing and detransitivizing languages. *Linguistic Typology* 8: 149–211.

Pavlenko, Galina (2002). Emotions and the body in Russian and English. *Pragmatics and Cognition* 10: 207–241.

Peterson, John (2005). There's a grain of truth in every 'myth', or, Why the discussion of lexical classes in Mundari isn't quite over yet. *Linguistic Typology,* 9: 391–405.

Plank, Frans (2005). Delocutive verbs, crosslinguistically. *Linguistic Typology* 9: 459–491.

Pustet, Regina (2000). How arbitrary is lexical categorization? Verbs vs. adjectives. *Linguistic Typology* 4: 175–212.

— (2003). *Copulas: Universals in the Categorization of the Lexicon.* Oxford: Oxford University Press.

Quine, W. V. O. (1960). *Word and Object*. Cambridge, Mass.: MIT Press.

Ricca, Davide (1993). *I verbi deittici di movimento in Europa: Una ricerca interlinguistica.* Firenze: La Nuova Italia Editrice.

Riemer, Nicholas (2005). *The Semantics of Polysemy: Reading Meaning in English and Warlpiri.* Berlin: Mouton de Gruyter.

Sasse, Hans-Jürgen (2002). Recent activity in the theory of aspect: Accomplishments, achievements, or just non-progressive state? *Linguistic Typology* 6: 199–271.

Shopen, Timothy (ed.) (1985). *Language Typology and Syntactic Description.* 3 volumes. Cambridge: Cambridge University Press.

Slobin, Dan I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In Dedre Gentner & Susan Goldin-Meadow (eds.), *Language in Mind: Advances in the Study of Language and Thought*, 157–191. Cambridge, Mass.: MIT Press.

Slobin, Dan I. & Melissa Bowerman (2007). Interfaces between linguistic typology and child language research. *Linguistic Typology* 11: 213–226.

Stassen, Leon (1997). *Intransitive Predication.* Oxford: Oxford University Press.

Steinberg, Reinhild (in preparation). *Lexikalische Polygenese im Konzeptbereich KOPF.*

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

Talmy, Leonard (1985). Lexicalization patterns. In Shopen (ed.) 1985, Volume 3, 57–149.

— (1991). Path to realization – via aspect and result: A typology of event conflation. *Berkeley Linguistic Society* 17: 480–519.

— (2000). *Toward a Cognitive Semantics.* 2 volumes. Cambridge, Mass.: MIT Press.

Tatevosov, Sergej (2002). The parameter of actionality. *Linguistic Typology* 6: 317–401.

Vanhove, Martine (forthcoming). Semantic associations between sensory modalities, prehension and mental perceptions: A cross-linguistic perspective. In Vanhove (ed.) forthcoming.

Vanhove, Martine (ed.) (forthcoming). *From Polysemy to Semantic Change: Towards a Typology of Lexical Semantic Associations.* Amsterdam: Benjamins.

Veselinova, Ljuba (2006). *Suppletion in Verb Paradigms.* Amsterdam: Benjamins.

Viberg, Åke (2006). Towards a lexical profile of the Swedish verb lexicon. *Sprachtypologie und Universalienforschung* 59: 103–129.

Wälchli, Bernhard (2005). *Co-compounds and Natural Coordination.* Oxford: Oxford University Press.

— (in preparation a). Lexicalization patterns in motion events revisited. Available at http://ling.uni-konstanz.de/pages/home/a20_11/waelchli/waelchli-lexpatt.pdf

— (in preparation b.). Constructing semantic maps from parallel text data. Available at http://ling.uni-konstanz.de/pages/home/a20_11/waelchli/publications.htm

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

Zalizniak, Anna (forthcoming). A catalogue of semantic shifts: Towards a typology of semantic derivation. In Vanhove (ed.) forthcoming.