INSTITUT FÜR SPRACHWISSENSCHAFT
UNIVERSITÄT ZU KÖLN

ARBEITSPAPIER NR. 30 (Neue Folge)

Lexical Typology: A Programmatic Sketch

Leila Behrens
Hans-Jürgen Sasse

August 1997

Lexical Typology: A Programmatic Sketch

0   Preliminaries

The present paper is an attempt to lay the foundation for Lexical Typology as a new kind of linguistic typology.[1] The goal of Lexical Typology is to investigate cross-linguistically significant patterns of interaction between lexicon and grammar.

To avoid misinterpretations from the outset, we will begin with a lexical semantic exercise clarifying the sense of the adjective "lexical" and the compositional nature of the term "Lexical Typology" as understood here. The typological approach advocated in this paper relies, in the first place, on the development of interdisciplinary lexicon research during the past 20 years, in which the modeling of the **interaction** between lexicon and grammar has become a central point of interest. A basic consensus emerging from these studies is that lexicon and grammar do not constitute components or subject matters of linguistics which would lend themselves to independent investigation. Rather, lexicon and grammar represent different **perspectives** in the scientific study of a composite "lexico-grammar", which forms an organic whole (cf. Halliday 1992, Gross 1994).

Proceeding from this empirically well-founded hypothesis, we assume that significant typological insights can only be gained through simultaneous consideration of lexicon (lexical semantics, categorial structuring) and grammar (semantics and structure of grammatical categories). The adjective "lexical" in the expression "Lexical Typology" is therefore intended to refer to the "lexical perspective" and the entire expression is to be read as "taking a lexical perspective in the typological investigation of lexico-grammar". From this it follows that we do not intend a restrictive reading of "Lexical Typology" in the sense of a cross-linguistic study of the structure of vocabulary ignoring issues of grammatical organization. The typological approach propagated here does not directly relate to the tradition of the well-known "subsystem typologies" such as "morphological typology", "syntactic typology", etc. The adjectives "morphological", "syntactic", etc. in these expressions are conceived of as referring to distinct linguistic components such as morphology, syntax, etc. based either on the assumption that these are autonomous mental modules or on the weak heuristic assumption that they can be investigated independently of each other. For example, a "syntactic typology" is usually understood as a "typology restricted to the domain of syntax" abstracting from regularities in all other components. Thus, a word order typology abstracts from cross-linguistic differences in morphological structure or compositional principles. It is not in this sense that we want to interpret the term "Lexical Typology" here, that is, the adjective "lexical" is not to be understood as referring to a distinct linguistic component named "lexicon". Rather, "lexical" refers to

---

the perspective from which we attempt to investigate typological regularities of the interaction of lexicon and grammar.

The concept of lexico-grammar is based on the observation that in the languages of the world, the organization of the lexicon (i.e. its semantic and formal structuring) is to a large extent dependent on the morphological and syntactic make-up (i.e. on constructional resources and the semantics and formal behavior of morphosyntactic categories). Such grammatical issues directly control compositionality, categorization of lexical items, inherent lexical semantics as opposed to phrasal or sentential semantics, and sense relations within lexical items (for instance, ambiguity) and between lexical items of the whole lexicon. In turn, grammatical entities (either classical grammatical categories such as subject or constructions such as diatheses) can only be understood by recourse to language-specific lexical, especially lexical-semantic, regularities. Linguistic research operating with the concept of lexico-grammar tries to account for this interweavement of lexical and grammatical structuring. Viewed in the context of comparative linguistic research, the concept of lexico-grammar leads to the assumption that we can expect, in different languages, quite divergent patterns of interactions between lexicon and grammar, and that these divergences are of great typological significance. It is therefore proposed that lexical semantics and its repercussions on grammar be assigned a central role in typological investigations. To this end, we will lay much emphasis on the discovery of principles of ambiguity and compositionality. These principles are presumably universal on a higher level of abstraction but typologically variable in their concrete individual manifestations. They therefore strongly influence the make-up of an individual language's grammar and lexicon.

The paper is organized as follows. We will first sketch some of the fundamental assumptions essential for an adequate comparative approach to language (section 1). We will then proceed to an outline of the role of recent lexicological research as a pacemaker for Lexical Typology (section 2). Section 3 will present a few ideas for a generalized approach based on these developments. In section 4, we will briefly discuss some methodological prerequisites we consider necessary for the intended kind of research.

1   Basic Assumptions and Problems

Traditional typology has been dominated by an approach which presupposes, as a basis for classification, certain substantive entities and structural principles. These entities and principles may be valid for certain individual languages from whom they have been abstracted and generalized but have empirically been found to not constitute absolute universals. Moreover, all prominent typologies of the past are grammatical typologies in which lexical-semantic variability does not play an important role. A well-known example of the kind of typology we have in mind here is Greenberg's word order typology (Greenberg 1963/1966). It is based on the following two assumptions: (a) for

any language it is possible to identify homogeneous formal entities corresponding to "subject", "object" and "verb"; (b) any language possesses a basic word order which can be described in terms of these categories. Both assumptions have turned out to be empirically false. An obvious conclusion that has been drawn from this state of affairs was the attempt to restrict the validity of word order typology to precisely that subset of languages which fulfill the two above-mentioned conditions (Comrie 1981/1989: 35-36). Another attempt was to preserve the universal validity of the classificatory principle by successively replacing "subject", etc. with related but slightly different categories (e.g. "subject" with "topic"). According to this modification, basic word order could still be conceived of as a universal principle of languages, but it was now allowed that the languages display variation with respect to the categories ordered in a canonical fashion.

One can find several parallels to these repair strategies - restriction of the typological domain to a subset of languages and generalization of the typological parameters - in the recent history of linguistics. Another typical example is the treatment of configurationality. On the one hand, scholars have opted for a restriction of the structural principle of "configurationality" to those languages which meet certain conditions such as fixed constituent order, lack of discontinuous expressions, lack of "pro-drop", presence of overt "expletive" elements such as English *it, there*, French *il* (cf. Hale 1982 and subsequent publications). On the other, attempts are made to define configurationality at a more abstract level (cf. É. Kiss 1987, 1995). According to such an approach, configurationality can be defined not only with respect to syntactic relations but also with respect to discourse entities, etc.

We think that any kind of typological research today has to cope from the start with the following empirical facts:

(1) a. Substantive grammatical categories, i.e. categories which constitute complex configurations of formal and semantic or pragmatic properties in well-investigated standard languages, do not enjoy a universal status if taken as holistic entities. This has been amply demonstrated for categories such as "subject" (cf. Keenan 1976) or for distinctions such as that between "mass nouns" and "count nouns" (cf. Behrens 1995); in principle, it would be demonstrable for **any** traditional grammatical category.

b. Structural principles, which in turn are based on substantive and complex grammatical categories, can likewise not be postulated as linguistic universals. It is clear, for example, that "configurationality" (in the narrow sense) cannot be taken as a universal structural principle of natural language (cf. the long dispute initiated by Ken Hale). Likewise, GB's principle that "all clauses have a subject" (the "predication principle" inspired by Williams 1980) has been invalidated on empirical grounds in a long chain of works over the past decades and cannot even be rescued by the pro-drop parameter as Haberland and Heltoft (1992) have shown. The same holds true of the "projection principle", and so forth.

c. Semantic "ingredients" of allegedly universal categories are not distributed in a universally predictable way among the lexicon and the grammar or among specific parts of the grammar. Individual languages can be expected to exhibit different kinds of "division of labor" among the components of a lexico-grammar (compare, for example, the rudimentary typology of the distribution of "aspect-like" and "aktionsart-like" phenomena presented in Sasse 1991). Consequently, there is no universal schema for the correlation between lexical information and grammatical patterns.

d. Languages display partial semantic or constructional similarities and similarities with respect to meaning variations and structural variations. From this it does not follow that it is possible to cross-linguistically identify primary senses (basic senses) or primary structures (basic structures) for lexical or phrasal units. For instance, we cannot assume that languages which possess an "active-like" and a "passive-like" voice always assign primary status to the former: In Tagalog, the active-like voice ("actor focus") is clearly not "basic" vis-à-vis other voices (cf. Schachter 1977). Likewise, the primary status of a linguistic unit's use vis-à-vis other uses in one language does not allow conclusions for similar constellations in other languages. For example, the spatio-temporally concrete use of a noun is not necessarily more "basic" than, say, its use to denote a kind in a taxonomic hierarchy: In Hungarian, for instance, a plural referring to sorts ("kinds of...") does not have a secondary status vis-à-vis other types of plural (cf. Behrens 1995).

The examples given in (1) a. through (1) d. are characteristic instances of a specific sort of typological variation which has for a long time raised questions of cross-linguistic comparability and universal generalizability of morphosyntactic and lexical categories, structural principles, so-called "markedness" relations and the like. However, these variations are not accidental deviations from universal or prototypical grammatical categories permitting the conclusion that cross-linguistic identification and typological evaluation of grammatical patterns could be done directly on a semantic and "intuitive basis" (cf. Croft (1990: 12-13) and the critique by Matthews (1993: 48)). Rather, these variations frequently relate to typologically significant differences in the interlocking principles of lexical and grammatical organization.

A brief discussion of one of the examples alluded to above will suffice to make the issue clear. In European languages there is a strong association between thematic roles and their syntactic realization. Diathesis is based on this association combined with a hierarchy of syntactic relations with the subject at the top. The primary status of actives via-à-vis passives in these languages corresponds to such a hierarchy, which is in turn manifested in specific patterns of valency and argument frames being central lexical properties of situation expressions (verbs). Passive is thus a method of reorganizing argument relations whose basic structure is lexically fixed.

In Tagalog, there is no association between thematic roles and syntactic relations and, consequently, no hierarchy of syntactic relations prototypically related to thematic

roles. The "active-like" voice and the "passive-like" voice are two members of a larger set of role-marking devices, which are all of equal status and not involved in any kind of derivational unidirectionality. Function words marking syntactic relations in Tagalog (*ang*, *ng*) are therefore not multiply ambiguous with respect to thematic roles. However, they express, in addition to their function as syntactic relation markers, certain aspects of determination (definiteness, specificity). Thus, Tagalog differs from European languages not only in its complete disentanglement of syntactic relations and thematic roles but also in the fact that it conflates determination and syntactic relations. This conflation has repeatedly led to classifications of these particles as either articles or case markers, depending on what the various authors considered as their primary function.[2] Such a situation may become the source of serious misunderstandings when elements with conflated functions are assigned to a "universal" category associated with one of the functions and then compared with language-specific "instances" of this category. For example, Kornai/Pullum (1990: 34) take for granted that the particle *ang* in Tagalog is an article and present it as counterevidence against cross-linguistic restrictions on the syntactic behavior of articles. Those features of *ang* which make up the counterevidence are, of course, due to the "case" function of this particle.

Moreover, recent approaches to "factorize" primitives of traditional grammar such as "syntactic relations" or the "mass/count" distinction have shown that the scientific concepts of such categories are often based on the language-specific clustering of components in a limited number of well-studied languages which historically constituted the empirical input of linguistic theories. For example, the customary theoretical treatment of "count nouns" and "mass nouns" is doubtlessly modeled on a situation such as found in English with its specific correspondence of the lexical semantic properties of nouns, the principles of number marking, the semantics of morphological number categories, and the usage of quantifiers and determiners. This orientation towards the specific form-meaning correlations in one model language (or a few of them) often generates a background of expectations as to the cross-linguistic behavior of linguistic units in a certain domain, eventually resulting in a grid which sets a frame for questionnaires, tests, etc.

However, when investigating the equivalents of such categories in other languages, one finds that their ingredients may be distributed in quite a different way across the different levels of analysis, often resulting in formidable dissimilarities in the overall lexico-grammatical organization. To elaborate further on the mass/count example, it is well known that the situation in numeral classifier languages is quite dissimilar from that of English in that the behavior of determiners and quantifiers, number marking and the lexical semantics of nouns do not at all cluster in the same way as in English. This has given rise to extensive discussion in the literature, but the different solutions proposed are again typically formulated against the background of an

---

[2] For a discussion of various interpretations of these particles as articles or case markers see Himmelmann (1983:85ff). The conflation of the two functions was pointed out as early as in Blake (1925:206).

English-type lexicon-grammar interaction (for instance statements to the effect that there are languages having "only mass nouns", or which "underspecify the distinction between count nouns and mass nouns", or "express the distinction syntactically", and so forth).

What is the proper remedy? As we have seen, the common denominator of variations of the kind depicted above is the fact that, in different languages, different form-function correlations in the lexicon are systematically connected with different form-function correlations in the grammar. The variability found in these connections may be subject to cross-linguistic analysis and classification. We find here a range of typological patterns as well as universal principles. In other words, specific constellations in the organization of the grammar (grammatical patterns) correspond, in a typologically significant way, to specific constellations in the organization of the lexicon (lexicalization patterns).

It is clear that the systematization of such correspondences cannot be achieved as long as typological investigation confines itself to overt morphological phenomena. Given that the whole enterprise centrally involves semantics, it also involves polysemy as an important principle of human language. Largely neglected so far in typological studies, regular polysemy, i.e. recurrent patterns of ambiguity, will be of central relevance to Lexical Typology. Systematic ambiguities often correlate with grammatical contexts (cf. section 2). The study of Lexical Typology has to encompass both cases of systematic ambiguities usually restricted to smaller or larger lexical-semantic classes within major lexical categories as well as those cases which affect entire lexical classes and result in "ambiguities of sentence semantics" such as the generic vs non-generic use of nouns. Moreover, a comprehensive typology of language must also deal with the fact that lexical ambiguity in one case may correspond to overt grammatical markings in another, both within single languages and across languages.

The considerations presented so far can be summarized in the following three basic research goals of Lexical Typology:

- A typological systematization of the interaction between lexicalization patterns and grammatical patterns;
- A cross-linguistic investigation of the distribution of ambiguities across lexicon and grammar;
- An investigation of the types of correspondence between systematic ambiguity patterns and morphological operations.

For obvious reasons it is difficult to find theoretical concepts for dealing with these issues in a traditional typological framework. First and foremost, traditional typology has a strong morphosyntactic bias with a poor morpheme-based concept of lexicon. Second, the investigation of cross-linguistic divergencies and similarities in the correspondences between lexicon and grammar presupposes detailed multi-factor analyses in individual languages, which must precede language comparison. Traditional

typology has largely been based on coarse-grained large-sample analysis. Furthermore, the patterns and principles we are looking for cannot be discovered by simply combining some of the single independent subsystem typologies hitherto proposed. An essential drawback of these typological approaches resides in the fact that they are restricted to certain aspects of grammatical structure (word structure, linear structure of major constituents (word order), general syntacto-semantic pattern of syntactic relations, etc.), which do not lend themselves easily to unification into an integrated macro-typology.

To sum up it can be said that an extension of traditional typological approaches does not appear to lead to fruitful directions for Lexical Typology. These approaches merely allow statements of partial regularities situated on a single level of analysis, rely on morpheme glossing and ignore entire lexicalization patterns, which results in serious misinterpretations of the constructional apparatus of a language and hence constitutes a considerable source of mistakes. In particular, when it comes to less well-documented languages a "double standard" is often employed: The degree of methodological sophistication standard for languages attested by an enormous amount of material such as English or French is hardly ever fulfilled in the investigation of poorly attested languages.

We therefore have to transgress the boundaries of traditional typology and resort to work done in other linguistic disciplines. Productive theoretical directions of research along these lines have been developed in the framework of recent lexicological work. We will now turn to a brief review of some of the issues that have been raised in this context.

2 Current Lexicon Research as a Pacemaker

2.1 Interdisciplinary Convergence in Lexicon Research

Increasing interest in lexical subregularities and productive lexical processes during the seventies (cf. Jackendoff 1975) marks the beginning of a very fruitful development resulting from the cooperation of theoretical linguistics, lexicography (in particular computational lexicography), computational linguistics and psycholinguistics. The following factors have played a prominent part in this converging development:

(2) a. In the eighties, lexical semantics acquired a central position as an independent subdiscipline (cf. Testen et al. 1984; Cruse 1986; Kuczaj/Barrett 1986; Hüllen/Schulze 1988; Gorfein 1989; Pustejovsky 1989; Boguraev/Pustejovsky 1990; Lehrer/Kittay 1992). Attention was directed to different types of lexical ambiguity differences, i.e. to different types of interaction between sense variation and context-dependence. Above all, a considerable interest in

systematic polysemy was arising (Apresjan 1974/1992; Nunberg 1978, 1979; Norrick 1981; Deane 1984; Pustejovsky 1989, 1995; Lehrer 1990).

b. In syntax-oriented research, the following insight was generally accepted: There can be no successful study of syntax without recourse to lexical-semantic properties. At the same time, it was recognized that the opposite is equally true: Without "an appreciation of the syntactic structure of a language, the study of lexical semantics is bound to fail" (cf. Pustejovsky 1991: 410). The investigation of regular correspondences between lexical-semantic and syntactic variation thus became an important research object (cf. Levin 1993).

c. Cognitive linguistics has made considerable headway in the cognitive foundation of semantic dimensions and contributed significantly to stimulate research interest in the systematic character and the derivability of metonymic and metaphoric relations (cf. Lakoff/Johnson 1980; Fauconnier 1985; Dirven 1985; Rudzka-Ostyn 1985; Lakoff 1987; Fass 1988, 1993; Martin 1991).

d. The development in computational lexicography and corpus research has created fundamentally new conditions for the investigation of lexical microstructure and paradigmatic subregularities. It has developed methods and tools for the extraction and systematization of lexical information (i.e. both lexical-semantic information and information about syntactic behavior) from large data corpora (i.e. text corpora and dictionaries) (cf. Zernik 1991a; Boguraev/Briscoe 1989; Wilks et al. 1988; Krovetz 1991; Slator 1991; Zernik 1991b; Atkins 1987; Justeson/Katz 1991, 1993; Smajda 1991; Anick/Pustejovsky 1990; Pustejovsky et al. 1993; Calzolari/Bindi 1990; Byrd 1989; Rizk 1989; Ide/Véronis 1990; Boguraev/Pustejovsky 1990; Klavans/ Chodorow/Wacholder 1990; Calzolari 1991; Klavans/Tzoukerman 1990; Kilgariff 1991; Bindi et al. 1991; Ahlswede 1993). Of particular relevance for the present approach is the systematization of machine-readable dictionaries and the extraction of systematic patterns of ambiguity and lexical-syntactic correspondences (cf. Atkins/Kegl/Levin 1986, 1988) which have led to insights of entirely new quality. This research is based on language-independent methods which render their application to typologically dissimilar languages possible (cf. Behrens 1994 on the results of Tagalog dictionary comparison).

Recent interdisciplinary lexicon research thus offers a host of new concepts and methods for the analysis of the interaction between lexicon and grammar. In particular, there is common agreement on a number of principles which have become standard in both theoretical and application-oriented lexicon research and which can be fruitfully exploited in the typological comparison of languages (cf. Zampolli 1994; Walker et al. 1995b):

(3) a. There is a great number of lexical-semantic and lexical-grammatical aspects (dimensions), according to which lexical units can and must be cross-classified (cf. Walker et al. 1995b).

b. Lexical decomposition is feasible. It cannot be expected, however, that certain types of features (for instance aspectual features for verbs) covering certain lexical-semantic properties will express all aspects of word meaning in terms

of language-independent semantic representation. Nowadays, semantic decomposition is no longer strongly connected with the aim of differentiating all lexical elements in a language. Rather, it is often only partially executed for controlling some lexical-grammatical regularities (cf. Sanfilippo 1992). Employed in this way, it is a very useful tool. Furthermore, it is generally accepted that we must distinguish between language-independent ("interlingual") features of lexical decomposition and language-specific features. Nevertheless, there is general agreement that rich multidimensional representations of subtle distinctions are desirable both in the context of single language studies and cross-linguistic studies. Multidimensional and highly structured representations have turned out to be particularly suitable in the multilingual environment, i.e. in the context of typological studies and in the context of computational linguistic applications such as machine translation or multilingual databases (cf. Talmy 1975, 1985; Pustejovsky 1989; Walker et al. 1995a).

c. There are regular correlations (or "alternations") between differing semantic interpretations of single lexical forms and their occurring in differing syntactic environments, which are central to the understanding of ambiguity, compositionality, and the interaction between syntax and lexicon (cf. Pustejovsky 1991).

d. The translation equivalents of members of a lexical class established in the source language on the basis of certain specific linguistic properties do not necessarily constitute a uniform class in the target language as well but may be distributed across different classes (cf. Zampolli 1994: 4).

e. Translation equivalence can often be achieved only at the phrasal or sentential level. The "lexical and structural mismatches" or "translation mismatches", which are observed in such cases, can frequently be attributed to divergent lexicalization patterns, which, in combination with divergent grammatical patterns, result in differences in compositionality (cf. Sanfilippo 1992; Dorr 1992/1993).

2.2   Talmy's Approach

Of particular interest in the present connection is recent work done by Leonard Talmy (1985, 1988a, 1988b, 1991a, 1991b). In this work three linguistic concepts play a key role: the concept of conflation of semantic components; the concept of multiple usage regularly corresponding with different environments and constituting lexicalization patterns; and the concept of semantic domains within which varying lexicalization patterns may be found.

Conflation of Semantic Components

The basic units of Talmy's typology are form-meaning relations. He proceeds from the assumption that the relationship between meaning and form in natural language is

largely not one-to-one. We share this assumption and regard it as a fundamental methodological principle: A combination of semantic components can be expressed by a single form, and a single semantic component (or type of component) can be expressed by a combination of forms. This may be called the "principle of conflation", i.e. the lexical integration of several semantic components. The phenomenon of conflation is well-known from grammatical morphemes (cf. the notion of "portmanteau morphemes"); it has long been recognized as a characteristic of the fusional technique of inflectional morphology as evident in languages such as Indo-European. This principle is now employed by Talmy in the description of lexical units and their typological comparison. A classical example of typological variation in conflation patterns is the difference between Germanic and Romance languages in the treatment of semantic components in verbs of motion. Whereas Germanic languages usually conflate the semantic component of MANNER with the component of MOTION (cf. Engl. *float*) and express the semantic component PATH as a locative preposition (*into, out of*), Romance languages usually integrate the PATH rather than the MANNER component in the verbs (Span. *entrar, salir*). With such PATH verbs, the MANNER component can be optionally realized by a gerund form from a special MANNER verb (*flotando*):

(4)   a. ENG: *The bottle floated **into** the cave.*
         SPAN: *La botella entró a la cueva (flotando).*
      b. ENG: *The bottle floated **out** of the cave.*
         SPAN: *La botella salió de la cueva (flotando).*

Lexicalization Patterns with Systematically Related Uses

Moreover, the English verb *float* has two systematically related uses: in the first one (*float₁*) the PATH component is not expressed in the sentence at all; in the second one (*float₂*) it is expressed by a directional phrase (cf. (5)). Thus, the English sentences under (4) contrasted with the Spanish sentences illustrate only the second use of the lexical form *float*. In contrast to this use (*float₂*), the semantic component of MOTION is not incorporated in the verb form when it is used without a directional PP (*float₁*), or, to put it in Talmy's words, in this case "the verb *float* refers to the buoyancy relation between an object and a medium" (1985: 64). These two different uses can be proved and demonstrated by two different sentence paraphrases, where (a) the grammatical instance of *float₁* can be substituted by and is connected through "usage equivalence" (in Talmy's terminology) to a grammatical instance of a stative expression (like *be afloat*) and (b) the grammatical instance of *float₂* is "usage equivalent" to a grammatical instance of a construction consisting of the verb *move* together with an infinite form of *float* used in the first sense. Note that the differing paraphrases in (5) make visible whether or not the relevant semantic component (MOTION) is incorporated in the verb: The paraphrase for the second use contains the verb *move* and is comparable to the preferred Spanish construction in expressing a MANNER verb as an infinite form (*move floating*).

(5) a. *The craft floated₁ on a cushion of air.*
     Paraphrase: *The craft was afloat on a cushion of air.*
     Usage equivalence: *float₁ = be afloat*
   b. *The craft floated₂ into the hangar on a cushion of air.*
     Paraphrase: *The craft moved into the hangar floating₁ on a cushion of air.*
     Usage equivalence: *float₂ = move floating₁*

Talmy calls such multiple uses "lexicalization doublets", emphasizing the fact that "lexicalization doublets" occurring in a recurrent fashion in the lexicon of a single language may constitute typologically significant lexicalization patterns. Thus discovering multiple uses by a controlled method of paraphrasing is an essential step in determining the semantic components of lexical units in a single language and investigating recurrent patterns of multiple uses is an essential step in the typological comparison of lexicalization patterns.

Systematically related uses constituting lexicalization patterns are often referred to as "systematic ambiguities" in the literature and are presently called "alternations", so far as the diverging semantic interpretations systematically correlate with different syntactic environments. "Alternations" have also been the subject of extensive research in syntactically oriented approaches in recent years. On a heuristic basis similar to Talmy's, alternations and semantic components have been connected in comprehensive studies by scholars such as Beth Levin (1993).

Since each language possesses its own complex system of alternations, there are significant differences among languages with respect to the different alternation and conflation patterns which may prevail in a given semantic domain.

Semantic Domains

A significant feature of Talmy's approach is the replacement of universal "primitive" grammatical categories by "semantic domains" such as, for example, the domain of "aspect" or the domain of "causation". In the languages of the world, these domains may interact in manifold ways and may thus form superordinate combinations of domains. For instance, typologically significant patterns in the combination of aspect and causativity, which depend on language-specific conditions of ambiguity and the language-specific morphological and syntactic potential, can be discovered on the basis of a universal schema operating with the semantic distinctions among "stative", "inchoative" and "causative".

3   A Generalized Approach

Current interdisciplinary lexicon research has provided innovative ideas in at least three areas essential to the typological study of lexicon-grammar interaction. These have led to the three principles extracted from Talmy's work in the preceding section:

- the **principle of the one-to-many relationship between form and meaning** which replaces the assumption of a basic one-to-one relationship;
- the **principle of systematic correspondences between semantic interpretation and grammatical environment** (in short, the principle of alternations) based on the investigation of systematic ambiguities and their grammatical contexts;
- the **principle of comparability on the basis of semantic domains**, i.e. the establishment of universal semantic domains as a standard of comparison instead of grammatical categories conceived of as universal "primitives" for which uniform expression forms and uniform superordinate meanings were to be expected in every language.

We postulate that these three principles be the basic theoretical assumptions underlying Lexical Typology. In the following, we will add a few remarks on these principles in order to touch on some points of possible further refinement.

The concept of alternations is a fruitful one since it paves the way for an adequate description of correlations of systematic ambiguities with grammatical contexts. However, it has to be expanded in order to fit the requirements of lexical typological research described in section 1. Given that the notion of alternation has been developed on the basis of English, alternations in the narrow sense have been understood as only those in which the grammatical context is a syntactic one. In the context of cross-linguistic research, however, it is clear that alternations associated with different types of grammatical context must be made comparable. For example, **all** transitivity alternations have to be seen in a uniform connection, irrespective of whether they come about by means of variation of syntactic environment alone (as in English) or by affixation (e.g. prefixation of *be-* as in German). Lexical Typology therefore requires a generalized concept of alternations, in which all kinds of systematically varying grammatical environments are regarded as correlates of a lexically determined semantic distinction (i.e. not only syntactic correlates, but also morphological and phonological ones). Different types of correlates often coexist in one single language. An example of this can be found in English noun-verb conversions: in some of them, the semantic distinction correlates with a syntactic distinction alone (as in *focus* (N) vs *focus* (V)), in others it correlates with a syntactic and a phonological one (as in *cónvict* (N) vs *convíct* (V)).

Semantic domains can be conceived of as cognitively based interconnected semantic landscapes (cf. Anderson 1982). They mirror current empirical knowledge gained from language-specific work and demarcate functional-semantic areas in which

our linguistic experience with the grammatical and lexical structure of many languages leads us to expect that certain semantic contrasts will be conventionalized, i.e. lexicalized and/or grammaticalized. Semantic domains are apt to serve as heuristic basis for cross-linguistic comparability: Semantic contrasts in individual languages are not necessarily expected to be realized by a uniform linguistic entity, but may be distributed across several (morphological/phonological/syntactic) expression forms, which nevertheless form a coherent common system. In addition, they may be simultaneously effective both in lexicalization patterns and in grammatical paradigms.

The assumption of **combined** semantic domains in Talmy's sense is of utmost importance for cross-linguistic comparison. It enables us to take account of the conflation of semantic contrasts such as found, for example, in the phrase-marking particles of Tagalog, which simultaneously indicate determination and - by agreement with the verb - participant roles. Given that the picture of the correlation between form and meaning is rather complex even within a single language, the existence of many-to-many-relationships is all the more likely from a cross-linguistic perspective.

The objects of typological comparison across languages are lexicalized patterns of systematic ambiguities or alternations and patterns of semantic conflation in a certain semantic domain or several interconnected domains. This has important consequences for the method of language comparison. The usual method is based on a direct comparison of sentences or texts which are translation-equivalent to each other. However, the comparison of random occurrences of single uses or members of alternations in sentences normally elicited or in multilingual texts is certainly not sufficient for the goal pursued here, since this kind of data does not cover the entire "usage-range" or alternation system. Thus it is necessary to systematically take the whole system of alternation and conflation patterns in the compared languages into account.

4   Methods and Resources

The linguistic facts relevant to the study of lexicon-grammar interaction can only be obtained by multidimensional microstructural analysis within the various domains, whereby language-specific complex categories are disentangled and made comparable. Fine-grained analyses will therefore constitute the chief methodological procedure in Lexical Typology. We will now briefly address the following basic problems arising in this context: selection of sources of information, differences among languages with respect to the status of documentation, and language sampling.

4.1   Sources of Information

One of the main problems which pose themselves for an integrated typology of lexicon-grammar interaction is that of data acquisition. As long as typological research is more

or less confined to straightforward formal ("surface") phenomena in one restricted linguistic area, say, syntactic constructions, it is comparably easy to describe and classify significant cross-linguistic patterns in that area on the basis of a small data sample extracted from reference grammars and texts. For Lexical Typology, however, exhaustive investigations into semantic and grammatical aspects of the usage of lexical elements are essential. This is a far more complicated task: The combinatorics of lexical elements are subject to much larger variation than that of grammatical constructions and can hardly be discovered on the basis of a single source of data alone. Each of the usual methods of data acquisition will be incomplete when taken individually because different sources will yield different types of information. Several sources have thus to be combined to complement each other.

For the empirical aspect of the kind of research we are proposing here, this means that the investigation must be based on a multitude of data acquisition processes to achieve an optimal exploitation of the various defective methods. Our proposal is a combination of at least the following four methods and data sources:

- Extraction of information, in particular of recurrent lexical patterns, from dictionaries and comparison of monolingual and multilingual dictionaries.
- Tests with informants.
- Systematic search in large machine-readable corpora if available.
- Systematic comparison of texts available in several languages (translations).

A few comments on these four data sources may be in order here.

Systematic exploitation of dictionaries has an object-linguistic and a metatheoretical aspect. Its object-linguistic benefit lies in the fact that dictionaries provide implicit information about recurrent paradigmatic patterns such as lexical-categorial ambiguities, systematic lexical-semantic ambiguities (for instance systematic metonymies like the INSTITUTION/BUILDING metonymy) or systematic collocational patterns. These types of information are more or less systematically captured in the macrostructure and microstructure of dictionaries and can be extracted relatively easily. This makes dictionaries a very important data source for determining the entire range of lexical subregularities. Such lexical subregularities are usually not exhaustively covered in grammatical descriptions, cannot be obtained by mere introspection, and cannot be comprehensively extracted even from larger text corpora due to their high degree of randomness. Thus, dictionaries assume an important heuristic function in the discovery of undetected patterns.

The other kind of information obtainable from dictionaries is more of a metalinguistic nature. Heavy representational inconsistencies within and between dictionaries often point to problems that have remained unsolved in theoretical linguistic work. Strong inconherence in the application of grammatical features, for instance, may suggest that the system of grammatical categories used as grammatical codes is not adequately defined for the language in question. For example, a comparison

of Tagalog dictionaries carried out in Behrens (1994) has revealed that the application of traditional lexical categories ("word classes") in Tagalog present a much more serious problem than commonly believed.

Work with native speakers is an indispensable complementation to the dictionary method. Dictionaries throw light on the types of lexical patterns existing in a certain language whereas they are not very informative about the boundaries of the input domain of semi-productive regularities, i.e., whether or not a specific lexical form which meets the relevant semantic and formal conditions for being subject to a lexical process actually undergoes this process. Failure to list senses or uses which could be generated on the basis of ambiguity/alternation patterns raise the same problem as non-listed morphological formations. Such "gaps" can indicate (a) that the expected use is not possible at all (i.e. entirely "blocked") or (b) that it is fully transparent and thus redundant or (c) that it is not very frequent and/or restricted to specific contexts, and so on. We need the informant with his creative capacity and his knowledge about usualized form-meaning pairs to supplement such missing information. His judgments on "normality conditions" (cf. Cruse 1986) and contextual restrictions are very useful for finer determination of the limits of semi-productive lexical patterns.

According to our experience, work with large (machine-readable) text corpora has the following merits in the context of lexicon research as proposed here. First, spontaneous connected text can serve as a control device for testing hypotheses about contextual restrictions gained by consulting dictionaries and/or by elicitation/introspection. Second, it permits the frequency analysis of lexical patterns. Third, it can reveal innovative patterns not yet captured by dictionaries or it can demonstrate uses which are overwhelmingly not yet accepted by informants.

Systematic comparison of multilingual texts is the only method that allows a simultaneous comparison of all languages of which a typology is to be established, i.e., it is the only method by which the overall (semantic and constructional) variation range in the investigated domain in the respective languages can be followed. Research based on dictionaries and large text corpora is mostly semasiologically oriented. It is usually carried out in a monolingual context or in a bilingual context via a selected mediator language, rather than in a multilingual context. However, we are also interested in how expression modes (systematic ambiguity in the lexicon, morphological derivation, syntactic constructions) for certain concepts or concept types (for instance, "individuated" pieces of material, objects seen as "sorts" in a taxonomy, etc.) are distributed in a particular language and in different languages (onomasiological perspective). Multilingual text comparison is a fruitful research tool for uncovering significant divergencies in the compositional make-up of languages. Furthermore, this method permits a contrastive testing of the interaction between lexicon and discourse.

4.2   Differences in Documentation and the Sample Problem

Differences in the documentation depth of languages render the undiscriminating employment of data acquisition methods difficult. There are statements in the literature to the effect that it is possible to base a typology of grammatical categories on information abstracted from standard reference grammars alone (Bybee et al. 1994: xv). This may indeed suffice for a coarse-grained large-sample assessment in a certain grammatical area, if the expected results do not go beyond a general picture of the type of categories frequently grammaticalized in the relevant domain.

For the kind of work we have in mind here, the situation is different. We have stressed repeatedly that fine-grained analyses of lexical microstructure presuppose access to a specific type of detailed lexical semantic information, which is not available for most of the languages of the world. This means that the data acquisition methods and the research tools enumerated above have to be handled in a variable way depending on the comprehensiveness of the documentation. It is trivial that in the investigation of well-known European standard languages with centuries of lexicographical tradition, work with informants will be of a quite different status than in the investigation of less well-described languages. In the former, work with native speakers will be chiefly restricted to testing the range of application of linguistic units, while in the latter it is more often than not necessary to elicit basic information on lexical semantics and on the use of central grammatical categories. This even pertains to relatively well-studied non-European languages such as Tagalog or Swahili, languages quite frequently used in theoretical linguistic argumentation.

It is therefore recommended to start from small samples of languages for which information on lexical microstructure can be easily obtained. This does not necessarily mean that small and poorly described "exotic" languages are a priori excluded from Lexical Typological investigation. On the contrary, we want to suggest that this kind of information be given prominence in future activities of language documentation. Detailed studies on lexical semantics, in particular polysemy and systematic patterns of ambiguity, in as many languages as possible are imperative for a large-scale study of Lexical Typology.

References

Ahlswede, Thomas E. (1993), Sense Disambiguation for Humans and Machines. In: *Making Sense of Words. Proceedings of the Ninth Annual Conference of the UW Centre for the New Oxford English Dictionary, September 27-28, 1993, Oxford, England*, Oxford, 75-88.

Anderson, Lloyd (1982), The 'Perfect' as a Universal and as a Language Particular Category. In: Hopper, Paul J. (ed.), *Tense and Aspect: Between Semantics and*

Pragmatics (= Typological Studies in Language 1). Amsterdam: John Benjamins, 227-264.

Anick, Peter / Pustejovsky, James (1990), An Application of Lexical Semantics to Knowledge Acquisition from Corpora. In: *Proceedings of COLING-90, Vol. 2, Helsinki*. Helsinki: Universitas Helsingiensis, 7-12.

Apresjan, J.D. (1974/1992), *Lexical Semantics. User's Guide to Contemporary Russian Vocabulary*. Ann Arbor: Karoma Publishers (Original Russian Edition: 1974, Moskau: Nauka).

Atkins, Beryl T. (1987), Semantic ID Tags: Corpus Evidence for Dictionary Senses. In: *The Uses of Large Text Databases. Proceedings of theThird Annual Conference of the UW Centre for the New Oxford English Dictionary, November 9-10, 1987, Waterloo, Canada*, Waterloo, 17-36.

Atkins, Beryl T. / Kegl, Judy / Levin, Beth (1986), Explicit and Implicit Information in Dictionaries. In: *Advances in Lexicology. Proceedings of the Second Annual Conference of the UW Centre for the New Oxford English Dictionary, November 9-11, 1986, Waterloo, Canada*, Waterloo, 45-63.

Atkins, Beryl T. / Kegl, Judy / Levin, Beth (1988), Anatomy of a Verb Entry: from Linguistic Theory to Lexicographic Practice. In: *International Journal of Lexicography 1(2)*, 84-126.

Atkins, Beryl T. / Zampolli, Antonio (eds.) (1994), *Computational Approaches to the Lexicon*. New York: Oxford University Press.

Behrens, Leila (1994), Alternationen - ein Schlüssel für die Universalienforschung? In: *Zeitschrift für Sprachwissenschaft 13(2)*, 149-200.

Behrens, Leila (1995), Categorizing Between Lexicon and Grammar.The MASS/COUNT Distinction In a Cross-Linguistic Perspective. In: *Lexicology 1*, 1-112.

Bindi, Remo / Calzolari, Nicoletta / Monachini, Monica / Pirrelli, Vito (1991), Lexical Knowledge Acquisition from Textual Corpora: A Multivariate Statistic Approach as an Integration to Traditional Methodologies. In: *Using Corpora. Proceedings of the Seventh Annual Conference of the UW Centre for the New Oxford English Dictionary, September 29 - October 1, 1991, Oxford, England*, Oxford, 170-196.

Blake, Frank R. (1925), *A Grammar of the Tagalog Language*. New Haven, Connecticut: American Oriental Society.

Boguraev, Branimir K. / Briscoe, Ted (eds.) (1989), *Computational Lexicography for Natural Language Processing*. London / New York: Longman.

Boguraev, Branimir K. / Pustejovsky, James (1990), Lexical Ambiguity and The Role of Knowledge Representation in Lexicon Design. In: *Proceedings of COLING-90, Vol. 2*, 36-41.

Bybee, Joan / Perkins, Revere / Pagliuca, William (1994), *The Evolution of Grammar. Tense, Aspect, and Modality in the Languages of the World*. Chicago and London: The University of Chicago Press.

Byrd, Roy J. (1989), Discovering Relationships Among Word Senses. In: *Dictionaries in the Electronic Age. Proceedings of the Fifth Annual Conference of the UW Centre for the New Oxford English Dictionary, September 18-19, 1989, Oxford, England*. Oxford, England, 67-79.

Calzolari, Nicoletta / Bindi, Remo (1990), Acquisition of Lexical Information from A Large Textual Italian Corpus. In: *Proceedings of COLING-90, Vol. 3, Helsinki.* Helsinki: Universitas Helsingiensis, 54-59.

Comrie, Bernard (1981/1989), *Language Universals and Linguistic Typology. Syntax and Morphology.* Oxford: Basil Blackwell.

Croft, William (1990), *Typology and Universals.* Cambridge: Cambridge University Press.

Cruse, D. Alan (1986), *Lexical Semantics.* Cambridge: Cambridge University Press.

Deane, Paul D. (1984), Three Types of Polysemy. In: David, Rood (ed.), *Papers from the 18th Mid-America Linguistics Conference at Boulder.* Boulder, CO: Department of Linguistics, University of Colorado at Boulder, 136-143.

Dirven, René (1985), Metaphor as a Basic Means for Extending the Lexicon. In: Paprotté, Wolf / Dirven, René (eds.), *The Ubiquity of Metaphor. Metaphor in Language and Thought.* Amsterdam / Philadelphia: John Benjamins, 86-119

Dorr, Bonnie J. (1992 / 1993), The Use of Lexical Semantics in Interlingual Machine Translation. In: *Machine Translation 7(3)*, 135-193.

É. Kiss, Katalin (1987), *Configurationality in Hungarian.* Dordrecht: Reidel.

É. Kiss, Katalin (ed.) (1995), *Discourse Configurational Languages.* New York / Oxford: Oxford University Press.

Fass, Dan C. (1988), Metonymy and Metaphor: What's the Difference? In: *Proceedings of COLING 88*, 177-181.

Fass, Dan C. (1993), Lexical Semantic Constraints. In: Pustejovsky, James (ed.), *Semantics and the Lexicon.* Dordrecht: Kluwer, 263-289.

Fauconnier, Gilles (1985), *Mental Spaces: Aspects of Meaning Construction in Natural Language.* Cambridge, MA: The MIT Press.

Gorfein, David S. (ed.) (1989), *Resolving Semantic Ambiguity.* New York et al.: Springer-Verlag.

Greenberg, Joseph H. (1963/1966), Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In: Greenberg, Joseph H. (ed.), *Universals of Language.* Cambridge, MA: The MIT Press, 73-113.

Gross, Maurice (1994), Constructing Lexicon-Grammars. In: Atkins, B.T. Sue / Zampolli, Antonio (eds.) (1994), 213-263.

Haberland, Hartmut / Heltoft, Lars (1992), Universals, Explanations and Pragmatics. In: Kefer, Michel / van der Auwera, Johan (eds.), *Meaning and Grammar. Cross-Linguistic Perspectives.* Berlin: Mouton de Gruyter, 17-26.

Hale, Kenneth (1982), Preliminary Remarks on Configurationality. In: Pustejovsky, J. / Sells, P. (eds.), *Proceedings of NELS 12.* Amherst, MA.

Halliday, Michael A.K. (1992), *Language as System and Language as Instance: The Corpus as a Theoretical Construct.* In: Svartvik, Jan (ed.), *Directions in Corpus Linguistics. Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991.* Berlin / New York: Mouton / Walter de Gruyter, 61-77.

Himmelmann, Nikolaus (1983), *Linking im Tagalog.* Munich: M.A.Thesis.

Hüllen, Werner / Schulze, Rainer (eds.) (1988), *Understanding the Lexicon: Meaning, Sense and World Knowledge in Lexical Semantics.* Tübingen: Niemeyer.

Ide, Nancy M. / Véronis, Jean (1990), Mapping Dictionaries: A Spreading Activation Approach. In: *Electronic Text Research. Proceedings of the Sixth Annual*

Conference of the UW Centre for the New Oxford English Dictionary, October 28-30, 1991, Waterloo, Canada, Waterloo, 52-64.

Jackendoff, Ray S. (1975), Morphological and Semantic Regularities in the Lexicon. In: *Language 51(3)*, 639-671.

Justeson, John S. / Katz, Slava M. (1991), Redefining Antonymy: The Textual Structure of a Semantic Relation. In: *Using Corpora. Proceedings of the Seventh Annual Conference of the UW Centre for the New Oxford English Dictionary, September 29 - October 1, 1991, Oxford, England*, Oxford, 138-153.

Justeson, John S. / Katz, Slava M. (1993), Principled Disambiguation: Discriminating Adjective Senses with Modified Nouns. In: *Making Sense of Words. Proceedings of the Ninth Annual Conference of the UW Centre for the New Oxford English Dictionary, September 27-28, 1993, Oxford, England*, Oxford, 57-74.

Keenan, Edward L. (1976), Towards a Universal Definition of "Subject". In: Li, Charles N. (ed.), *Subject and Topic*. New York: Academic Press, 303-333.

Kilgariff, Adam (1991) Corpus Word Usages and Dictionary Word Senses: What is the Match? An Empirical Study. In: *Using Corpora. Proceedings of the Seventh Annual Conference of the UW Centre for the New Oxford English Dictionary, September 29 - October 1, 1991, Oxford, England*, Oxford, 23-39.

Klavans, Judith L. / Chodorow, Martin S. / Wacholder, Nina (1990), From Dictionary to Knowledge Base via Taxonomy. In: *Electronic Text Research. Proceedings of the Sixth Annual Conference of the UW Centre for the New Oxford English Dictionary, October 28-30, 1991, Waterloo, Canada*, Waterloo, 110-131.

Klavans, Judith L. / Tzoukermann, Evelyne (1990), Linking Bilingual Corpora and Machine Readable Dictionaries with the BICORD System. In: *Electronic Text Research. Proceedings of the Sixth Annual Conference of the UW Centre for the New Oxford English Dictionary, October 28-30, 1991, Waterloo, Canada*, Waterloo, 19-30.

Kornai, András / Pullum, Geoffrey K. (1990), The X-bar Theory of Phrase Structure. In: *Language 66(1)*, 24-50.

Krovetz, Robert (1991), Lexical Acquisition and Information Retrieval. In: Zernik, Uri (ed.), 45-64.

Kuczaj, Stan A. / Barrett, Martyn D. (1985), *The Development of Word Meaning*. Progress in Cognitive Development Research. New York et al.: Springer-Verlag.

Lakoff, George (1987), *Women, Fire and Dangerous Things*. Chicago: Chicago University Press.

Lakoff, George / Johnson, Mark (1980), *Metaphors We Live By*. Chicago / London: Chicago University Press.

Lehrer, Adrienne (1990), Polysemy, Conventionality, and the Structure of the Lexicon. In: *Cognitive Linguistics 1(2)*, 207-246.

Lehrer, Adrienne / Kittay, Eva F. (eds.) (1992), *Frames, Fields, and Contrasts. New Essays in Semantic and Lexical Organization*. Hillsdale, N.J. et al.: Lawrence Erlbaum.

Levin, Beth (1993), *English Verb Classes and Alternations. A Preliminary Investigation*. Chicago/London: The University of Chicago Press.

Martin, James H. (1991), Representing and Acquiring Metaphor-Based Polysemy. In: Zernik, Uri (ed.), 389-415.

Matthews, Peter H. (1993), *Grammatical Theory in the United States from Bloomfield to Chomsky.* Cambridge et al.: Cambridge University Press.

Norrick, Neal R. (1981), *Semiotic Principles in Semantic Theory.* Amsterdam: John Benjamins.

Nunberg, Geoffrey (1978), *The Pragmatics of Reference.* Bloomington, Indiana: Indiana University Linguistics Club.

Nunberg, Geoffrey (1979), The Non-Uniqueness of Semantic Solution: Polysemy. In: *Linguistics and Philosophy 3*, 143-184.

Pustejovsky, James (1989), Current Issues in Computational Lexical Semantics. In: *Proceedings of the 4th European ACL, Manchester 1989*, XVII-XXV. Manchester.

Pustejovsky, James (1991), The Generative Lexicon. In: *Computational Linguistics 17(3)*, 409-441.

Pustejovsky, James (1995), *The Generative Lexicon.* Cambridge, MA / London, England: The MIT Press.

Pustejovsky, James / Bergler, Sabine / Anick, Peter (1993), Lexical Semantic Techniques for Corpus Analysis. *Computational Linguistics 19 (2)*, 331-358.

Rizk, Omneya A. (1989), *Sense Disambiguation of Word Translations in Bilingual Dictionaries: Trying to Solve the Mapping Problem Automatically.* In: IBM Research Report RC 14666 (#65812). Yorktown Heights, N.Y.: T.J. Watson Research Center.

Rudzka-Ostyn, Brygida (1985), Metaphoric Processes in Word Formation: The Case of Prefixed Verbs. In: Paprotté, Wolf / Dirven, René (eds.), *The Ubiquity of Metaphor. Metaphor in Language and Thought.* Amsterdam / Philadelphia: John Benjamins, 209-241.

Sanfilippo, Antonio / Briscoe, Ted / Copestake, Ann / Marti, Maria Antonia / Taulé, Mariona / Alonge, Antonietta (1992), Translation Equivalence and Lexicalization in the ACQUILEX LKB. In: *Proceedings of the Fourth International Conference on Theoretical and Methodological Issues in Machine Translation, Montreal.* Montreal, 1-11.

Sasse, Hans-Jürgen (1991), Aspect and Aktionsart: A Reconciliation. In: *Belgian Journal of Linguistics 6*, 31-45.

Schachter, Paul (1977), Reference Related and Role Related Properties of Subjects. In: Cole, P./Sadock, J.M. (eds.): *Syntax and Semantics 8: Grammatical relations.* New York: Academic Press 1977, 297-306.

Slator, Brian M. (1991), Using Context for Sense Preference. In: Zernik, Uri (ed.), 65-90.

Smajda, Frank (1991), Macrocoding the Lexicon with Co-occurrence Knowledge. In: Zernik, Uri (ed.), 165-189.

Talmy, Leonard (1975), Semantics and Syntax of Motion. In: Kimball, J. (ed.), *Syntax and Semantics 4.* New York et al.: Academic Press, 181-238.

Talmy, Leonard (1985), Lexicalization Patterns: Semantic Structure in Lexical Forms. In: Shopen, Timothy (ed.), *Language Typology and Syntactic Description, III: Grammatical Categories and the Lexicon.* Cambridge: Cambridge University Press, 57-149.

Talmy, Leonard (1988a), The Relation of Grammar to Cognition. In: Rudzka-Ostyn, Brygida (ed.), *Topics in Cognitive Linguistics.* Amsterdam / Philadelphia: John Benjamins, 165-205.

Talmy, Leonard (1988b), Force Dynamics in Language and Cognition. In: *Cognitive Science 12,* 49-100.

Talmy, Leonard (1991a), Path to Realization: A Typology of Event Conflation. In: *Proceedings of the Berkeley Linguistic Society 17.* Berkeley: Berkeley Linguistic Society, 480-519.

Talmy, Leonard (1991b), Path to Realization: A Typology of Event Integration. In: *Buffalo Papers in Linguistics 1.*

Testen, David / Mishra, Veena / Drogo, Joseph (eds.) (1984), *Papers from the Parasession on Lexical Semantics.* Chicago: Chicago Linguistic Society.

Walker, Donald E. / Zampolli, Antonio / Calzolari, Nicoletta (1995b), Introduction. In: Walker, Donald E. et al. (eds.), 1-19.

Walker, Donald E. / Zampolli, Antonio / Calzolari, Nicoletta (eds.) (1995a), *Automating the Lexicon. Research and Practice in a Multilingual Environment.* Oxford: Oxford University Press.

Wilks, Yorick / Fass, Dan / Guo, Cheng-Ming / Mcdonald, James / Plate, Tony / Slator, Brian (1988), Machine Tractable Dictionaries as Tools and Resources for Natural Language Processing. In: *Proceedings of COLING 88,* 750-755.

Williams, Edwin (1980), Predication. In: *Linguistic Inquiry 11,* 203-238

Zampolli, Antonio (1994), Introduction. In: Atkins, Beryl T. / Zampolli, Antonio (eds.), *Computational Approaches to the Lexicon.* New York: Oxford University Press, 3-15.

Zernik, Uri (1991b), Train1 vs. Train2: Tagging Word Senses in Corpus. In: Zernik, Uri (ed.), 91-112.

Zernik, Uri (ed.) (1991a), *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon.* Hillsdale, N.J. et al.: Lawrence Erlbaum.

Von 1968 an erschienen die von Prof. Dr. Hansjakob Seiler herausgegebenen Arbeitspapiere des Instituts für Sprachwissenschaft. Nach der Emeritierung von Prof. Dr. Seiler im März 1986 wurde eine neue Folge mit neuer Zählung und dem Zusatz "Neue Folge" (N. F.) begonnen. Herausgeber ist das Institut für Sprachwissenschaft.

Arbeitspapiere Köln (Liste noch vorrätiger Arbeitspapiere)

3. SEILER, H. & SCHEFFCYK, A. 1969. Die Sprechsituation in Linguistik und Kommunikationswissenschaft. Referat einer Diskussion.

5. BRETTSCHNEIDER, G. 1969. Das Aufstellen einer morphophonemischen Kartei (illustriert an der Morphophonemik des japanischen Verbs).

6. PENČEV, J. 1969. Einige semantische Besonderheiten der bulgarischen Geschmacksadjektive.

14. ROSENKRANZ, B. 1970. Georg von der Gabelentz und die Junggrammatische Schule.

18. SEILER, H. 1971. Possessivität und Universalien. Zwei Vorträge gehalten im Dezember 1971: I. Zum Problem der Possessivität im Cahuilla (Uto-Aztekisch,Südkalifornien) II. Possessivität und Universalien.

23. BRETTSCHNEIDER, G. & LEHMANN, Ch. 1974. Der Schlagwortkatalog des Instituts für Sprachwissenschaft der Universität Köln.

24. WIESEMANN, U. 1974. Time Distinctions in Kaingang.

26. SEILER, H. u.a. 1975. Deskriptive und etikettierende Benennung; Relativkonstruktionen, (Becker, Katz, Walter, Habel, Schwendy, Kirsch, Clasen, Seip).

29. VAN DEN BOOM, H. & SAMUELSDORFF, P. 1976. "Aspects"-Kommentar. Protokolle eines Seminars aus dem WS 1975/76.

36. STEPHANY, U. 1978. The Modality Constituent - A Neglected Area in the Study of First Language Acquisition.

37. LEHMANN, Ch. 1980. Guidelines for Interlinear Morphemic Translation. A proposal for a standardization.

40. PAUL, W. 1982. Die Koverben im Chinesischen (with an English summary).

41. SCHLÖGEL, S. 1983. Zum Passiv im Türkischen.

42. BREIDBACH, W. 1983. Zur Possession im Samoanischen.

43. STEPHANY, U. 1983. The development of modality in language acquisition.

44. SEILER, H. Die Indianersprachen Nordamerikas. Ausarbeitung der Vorlesung SS 1980.

45. KUKUČZKA, E. 1984. Lokalrelationen und Postpositionen im Tamil.

49. PREMPER, W. 1986. Kollektion im Arabischen.

50. FACHNER, R. 1986. Der Relativsatz im Bambara.

51. PUSTET, R. 1986. Zur Frage der Universalität des "Subjekts": Das Ayacucho-Quechua.

52. REICHERT, Ch. 1986. Verteilung und Leistung der Personalaffixe im Ungarischen.

Neue Folge (Die fettgedruckten Nummern der Arbeitspapiere sind vorrätig.)

1. HOFMANN, Gudrun 1986. Zum Verständnis epistemischer Modalausdrücke des Deutschen im Kindergartenalter.

2. BREIDBACH, Winfried 1986. Die Verben mit der Bedeutung 'weggehen' im Althochdeutschen.

3. HASPELMATH, Martin 1987. Verbal nouns or verbal adjectives? The case of the Latin gerundive and gerund.

4. MOSEL, Ulrike 1987. Inhalt und Aufbau deskriptiver Grammatiken (How to write a grammar)

5. HASPELMATH, Martin 1987. Transitivity alternations of the anticausative type.

6. BREIDBACH, Winfried 1988. Die Schiffsbezeichnungen des Alt- und Mittelhochdeutschen. Ein onomasiologisches und etymologisches Glossar.

7. HAASE, Martin 1988. Der baskische Relativsatz auf dem Kontinuum der Nominalisierung.

8. THOMADAKI, Evangelia 1988. Neugriechische Wortbildung.

9. SASSE, Hans-Jürgen 1988. Der irokesische Sprachtyp.

10. HAASE, Martin 1989. Komposition und Derivation: Ein Kontinuum der Grammatikalisierung.

11. BAUMANN, Monika 1989. Die Integration englischer Lehnwörter in die samoanische Sprache. Das Verhalten der englischen Konsonantenphoneme.

12. SASSE, Hans-Jürgen 1990. Theory of Language Death; Language Decay and Contact-Induced Change: Similarities and Differences.

13. SCHULTZE-BERNDT, Eva 1991. Phonetische Substanz und phonologische Theorie. Eine Fallstudie zum Erstspracherwerb des Deutschen.

14. SASSE, Hans-Jürgen (Hg.) 1991. Aspektsysteme.

15. HIMMELMANN, Nikolaus P. 1991. The Philippine Challenge to Universal Grammar.

16. HIMMELMANN, Nikolaus P. 1992. Grammar and Grammaticalization.

17. COMPES, Isabel & KUTSCHER, Silvia & RUDORF, Carmen 1993. Pfade der Grammatikalisierung: Ein systematisierter Überblick.

18. COMPES, Isabel & OTTO, Barbara 1994. Nicht-morphologische Nominalinkorporation - etwas ganz anderes?

19. DROSSARD, Werner 1994. The Systematization of Tagalog Morphosyntax.

20. ÁGEL, Vilmos 1994. Valenzrealisierung, Grammatik und Valenz.

21. KEUSEN, Anna 1994. Analysis of a Cayuga Particle: *ne:'* as a Focus Marker.

22. STEPHANY, Ursula 1995. The Acquisition of Greek.

23. SASSE, Hans-Jürgen 1995. A Preliminary Bibliography on Focus.

24. KUTSCHER, Silvia & MATTISSEN, Johanna & WODARG, Anke (Hrsg.) 1995. Das Mutafi-Lazische.

25. GARCÍA CID, Aranzazu 1995. Parenthesen, Einschübe und Kommentare: Zur Klassifikation von Nebenprädikationen in gesprochenen spanischen Texten.

26. JOSWIG, Andreas 1996. Die grammatischen Rollen des Objekts im Swahili.

27. SASSE, Hans-Jürgen 1996. Theticity.

28. SASSE, Hans-Jürgen 1997. Aspektsemantik und Lexikonorganisation: Beobachtungen zum Cayuga (Nordirokesisch)

29. COMPES, Isabel 1997. Die *ona-lea* Konstruktion im Samoanischen. Eine Untersuchung zur Struktur narrativer Texte.