Introduction

Hans-Jörg Schmid and Susanne Handl

1. Background

Significant innovations in theory-building tend to be accompanied by changes in methodology. For example, when generative grammar replaced American structuralism in the 1960s as the leading paradigm in linguistics, performance-based methods were abandoned in favour of introspection. Interestingly, a similar shift did not take place when cognitive linguists started to have a go at the basic assumptions of generativism. This is particularly remarkable in view of the catchword of the *usage-based* approach, which was introduced to cognitive linguistics by Ron Langacker (1987, 1988) to encapsulate the idea that knowledge of grammar is extracted from the actual use of linguistic structures (rather than implemented on the basis of an innate blueprint). While this would suggest that linguists pursuing a *usage-based* approach would actually look for relevant evidence in authentic language use, the introspective method continued to dominate cognitive linguistic research for a surprisingly long time. The required methodological changes were much slower in coming than the outpour of theoretical claims – and have in fact not been achieved in some quarters of the cognitive-linguistic community so far.

In recent years, however, the necessary methodological consequences resulting from a serious understanding of the *usage-based* programme have clearly been recognized. From this perspective, a linguistic approach qualifies as *usage-based* if, when formulating linguistic hypotheses, it takes a thorough look at the actual use of linguistic structures. In the most extreme versions of this view, usage-*based* theories of grammar have been replaced by distinctly inductive usage-*driven* ones. Both usage-based and usage-driven approaches are attractive for cognitive linguistics because they allow conclusions about how lexical, grammatical and pragmatic knowledge finds its way into the minds of the speaker-hearers of a language and comes to be stored there. Thus, by looking into real-life language, cognitive usage-based approaches expect to gain insights into cognitive foundations not only of language use, which has long been the main aim of psycholinguistic research, but also of language structure.

The recent trend towards a usage-based methodology, an important milestone of which is marked by the volume *Usage-based models of language* edited by Michael Barlow and Suzanne Kemmer (2000), manifests itself in a number of different strands. One important impetus has recently come from the empirical work on a usage-based theory of first language acquisition by Michael Tomasello and his team at Leipzig (cf., e.g., Tomasello 2000, 2003; Lieven et al. 2003). In the same period, historical linguistics has seen a move towards usage-based accounts of language change, e.g. in the work of Joan Bybee and Paul Hopper (cf., e.g., Bybee and Hopper 2001; Bybee 2006a, 2006b). Combining the corpus-linguistic methodology with cognitive-linguistic theorizing, linguists such as Dirk Geeraerts and his research group at Leuven (e.g. Tummers, Heylen and Geeraerts 2005), Stefan Gries (e.g. to appear), Hans-Jörg Schmid (2000) and Anatol Stefanowitsch (e.g. Stefanowitsch and Gries 2003, 2006) have tried to tap into the linguistic usage of large populations of speakers by investigating the material collected in computerized corpora.

2. Focus

Attempts to establish plausible links between linguistic data, on the one hand, and assumptions about their cognitive foundations, on the other, can only be convincing to the extent that they rely on observed recurrent linguistic behaviour, no matter whether it is recorded in the form of corpora of authentic language use, studied in linguistic experiments or simulated computationally on the basis of actual usage. For it is only for recurrent patterns of usage that it makes sense to assume that the underlying structure is intra-subjectively stable across time and intersubjectively similar across members of a speech community. Both characteristics are required if a given linguistic form is to be seen as manifesting a stored representation that is part of the 'grammar' of a language or variety of language.

Taking this obligation very seriously, the papers in the present volume all aim to bring together observed patterns of linguistic usage with cognitive-linguistic concepts and models. Equally importantly, all contributions have an empirical basis and show a high level of awareness of the potential and limits of the methodology applied. The methods used range from the investigation of corpora and tailor-made samples of authentic language use to linguistic and psycholinguistic experiments as well as computational simulations based on actual usage.

The linguistic phenomena investigated in the contributions run the gamut from the lexico-conceptual and collocational level to morphological and grammatical categories, constructions and pragmatic functions. Cutting across the grouping of the papers into lexical and grammatical studies that divide the volume into two parts (see below, Section 3), two complementary perspectives of language and cognition are represented: in one set of papers, the established methods of psycholinguistic experimentation, quantitative corpus analysis and computational simulation are exploited to demonstrate the viability and increase the plausibility and force of cognitive-linguistic thinking. The papers in the second group test well-known cognitive-linguistic approaches such as conceptual metaphor theory, the theory of idealized cognitive models and construction grammar against authentic data demonstrating their applicability and explanatory potential, but also their limitations. Both groups include papers reaching beyond the scope of traditional cognitive-linguistic topics, e.g. by taking a critical stance of reductionist cognitive thinking.

3. The contributions

As mentioned above, the volume is divided into two parts, each comprising five papers. The papers in the first part focus on lexical patterns and their relations to cognitive processes and cognitive-linguistic concepts. They are ordered according to the complexity of the linguistic elements studied, from individual lexical items to concepts and collocations.

The same principle underlies the arrangement of the papers in the second part. Starting with inflectional morphemes and grammatical categories, the grammatical patterns investigated include argument-structure constructions and valency patterns as well as the pragmatic functions of sentence mood.

In the first contribution on the lexicon, **George Dunbar** addresses a problem that has a long history in cognitive semantics, viz. the distinction between ambiguous and vague lexemes. While ambiguous lexemes have traditionally been considered to have a number of distinct senses, vague ones are seen to carry one meaning that is interpreted in different ways depending on actual usage contexts. Taking up a proposal by Tuggy (1993), who pleaded for a scalar approach that treats ambiguity and vagueness as two poles of a continuum, Dunbar describes a computational model implementing this continuum, which is based on a connectionist network

and validated against the lexicographic decisions taken by the corpus-based COBUILD dictionary. Dunbar closes his paper by arguing that the general mechanism underlying his model gives a good account of a number of general cognitive and perceptual phenomena.

The focus of **Dylan Glynn's** paper is also a semantic relation, viz. synonymy. Studying the three near-synonyms *annoy, bother* and *hassle* denoting slightly different aspects of the concept BOTHER, Glynn emphasizes that semantic investigations must take into account not only the lexico-grammatical frames providing patterns for occurrences of individual lexemes, but also use-related and user-related aspects like registers and regional varieties. His approach is corpus-driven and quantitative, and highly sensitive to the power and limitations of the methods applied. In order to come to grips with the highly multivariate data situation, Glynn uses advanced statistical methods such as correspondence analysis and hierarchical cluster analysis. These multidimensional techniques allow him to map usage patterns that arguably correspond to ways of carving up conceptual space as suggested to speakers of English by the grammar and lexicon of that language.

Olaf Jäkel applies the theory of idealized cognitive models (cf. Lakoff 1987: 113–114 *et passim*) to the study of public boundary disputes concerning the highly controversial concepts of LIFE and DEATH. His investigation focuses on the entrance boundary of LIFE, with linguistic material taken from the public discourse on embryonic stem cell research going on in both English (United States) and German (Germany) in the years 2000 to 2002. By close scrutiny of the data collected, Jäkel manages to show how scientists and politicians involved in the stem cell debate quarrel over denotational incongruencies, each party trying to dislocate or relocate denotational boundaries to suit their aims. The conceptual basis of this dispute is provided by diverging cognitive models of LIFE, including the conservative model, which sees life as beginning with conception, and the biotechnical model, according to which human life proper does not begin before *nidation*, a term introduced fairly recently to denote the settling of the foetus in the female womb.

Like Jäkel, **Brigitte Nerlich** studies usage-patterns in public discourse with the aim of unravelling the conceptual framing of public events. Nerlich looks into press releases and interviews published by scientists as well as the press coverage of key events in science and presents two case studies, one on the alleged breakthrough towards the possibility of 'cloning' the first human being in the laboratories of South Korean scientist Woo-Suk

Hwang, and one on the emergence of so-called *superbugs* heralding the *post-antibiotic apocalypse*. In her analyses, Nerlich extends conceptual metaphor theory in order to study the politics and ethics of *discourse metaphors* in authentic contexts. She manages to show how usage-patterns that rely on entrenched conceptual metaphors are deliberately launched and exploited by scientists themselves and by the press to influence public opinion, for example, with the ultimate aim of creating the public hysteria that will force politicians to provide more funding.

Susanne Handl and **Eva-Maria Graf** introduce an acquisitional aspect into the pattern discussion, relating the contextualist notions of idiom principle and open-choice principle (cf. Sinclair 1991) to the cognitive notions of holistic and analytic language processing (cf. Wray 2002). Drawing on the hypothesis that the quality and evolution of recurring word combinations in different stages of linguistic development provide insights into the anchoring and processing of language in the mind, they classify two essential types of word co-occurrences, i.e. lexical collocations and patterns. Their analysis of these types in children's and adolescents' corpora shows that in a phase of predominantly holistic language processing the percentage of lexical collocations is higher, whereas in an analytical phase, speakers produce more patterns, as they have become aware of the separability and combinability of previously unanalyzed linguistic chunks.

The first paper of the second part, authored by **Ewa Dąbrowska**, moves the interest in language acquisition from the lexicon to grammar. Dąbrowska presents two empirical studies which show that children rely on low-level generalizations when acquiring their first language. One study deals with the inflectional marking of the dative singular in Polish, the other with questions with long-distance dependencies in English (e.g. *what do you think you're doing* or *who do you think you are*). In both cases Dąbrowska reports experimental evidence suggesting that low-level schemas are psychologically more basic and often preferred to the higher-level generalizations proposed in the form of 'rules' by generative grammar. What is also striking is that these low-level schemata tend to hinge on prototypical lexical realizations of constructions.

Klaus-Michael Köpcke, **Klaus-Uwe Panther** and **David Zubin** argue for a conceptual-pragmatic approach to explaining gender agreement in German. Providing a wealth of attested examples from various sources they adapt Corbett's (2003) gender agreement hierarchy by replacing Corbett's formal categories with the conceptual-pragmatic functions of specifying, modifying, predicating and reference-tracking. While grammatical agree-

ment dominates in specifying and modifying contexts, conceptual agreement tends to prevail in uses with reference-tracking function, especially when other syntactic factors (like high degree of syntactic embeddedness) and discourse factors (such as narrative concerns) support this choice.

Ulrich Detges tackles a grammatical problem similar to the lexical one dealt with by Dunbar. The French-language phenomenon which he studies from both a diachronic and a synchronic point of view traditionally goes by the name of *imparfait de politesse* and has often been considered a mere usage variant of the 'normal' *imparfait* by many researchers. By means of an in-depth quantitative and qualitative corpus study, Detges is able to show that the so-called *imparfait de politesse* actually encompasses two types of phenomena that should be distinguished, namely one more variable pattern manifesting a range of verbs that invite a metonymic inference yielding a down-toning effect, and another more specific one consisting of *je voulais* ('I wanted') and a *verbum dicendi* such as *dire* 'to say', *parler* 'to speak', *demander* 'to ask', *proposer* 'to propose' etc. While the first type retains many aspects of 'normal' uses of the *imparfait*, the second one has become entrenched as a discourse marker with a present-tense meaning serving a range of specific textual functions. Detges concludes that the second pattern offers a case of a polysemous meaning of a grammatical construction, since *je voulais* + *verbum dicendi* is still motivated by the meaning of the *imparfait*, but too removed for it to be experienced as being derived from the latter by present-day native speakers of French.

Like Ewa Dąbrowska, **Thomas Herbst** addresses the nature and degree of generalizations stored in the minds of native speakers of a language. His focus lies on a comparison of the predictions made by construction grammar, as represented by Goldberg's (1995, 2006) argument-structure constructions, with those (implicitly made) by European valency models of grammar. While the former postulates fairly high-level generalizations assisted by lower-level schemata, descriptions of verbs in terms of their valency patterns have typically been item-specific, as they often defy generalizations based on shared meanings. Herbst's rich data come from diverse corpora as well as the *Valency Dictionary of English* (Herbst et al. 2004), which identifies several hundred valency patterns of English verbs, nouns and adjectives. Herbst concludes by stating that construction grammar is probably better equipped than valency grammar to account for grammatical creativity (cf., e.g., Goldberg's by now notorious … *sneezed the tissue off the table*, 1995: 152). On the other hand, construction grammar still has to find a way of adequately accounting for how the wealth of

item-specific knowledge of grammatical patterns is stored in long-term memory.

Patric Bach and **Dietmar Zaefferer** investigate the pragmatic functions of declarative and interrogative sentences. They compare two languages: German, where interrogatives are marked by subject-operator inversion at the beginning of sentences, and Japanese, where interrogatives are marked by a sentence-final interrogative particle (*ka*). Their main concern is how the difference between assertives and interrogatives is processed cognitively and whether it has an effect on the cognitive representations of the contents of the corresponding sentences. These research questions are investigated with original experiments exploiting the so-called Simon effect, i.e., the observation that ipsilateral responses are faster and more accurate than contralateral ones. By systematically varying the place where visual and verbal information was displayed on the computer screen and the side of the keys that informants had to press on the computer keyboard, Bach and Zaefferer were able to isolate the effect of the assertion-question distinction and the effect of the forward-typing (German) and backward-typing (Japanese) language. They present evidence from their tests suggesting that declarative sentences are processed in a richer and more fleshed-out propositional form than interrogatives and that the types of cognitive representations constructed depend on the position of the interrogative marker in the sentence.

References

Barlow, Michael, and Suzanne Kemmer
    2000   *Usage-based Models of Grammar*. Cambridge: Cambridge University Press.
Bybee, Joan
    2006a   *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.
    2006b   From usage to grammar: the mind's response to repetition. *Language* 82 (4): 711–733.
Bybee, Joan, and Paul Hopper (eds.)
    2001   *Frequency and the Emergence of Linguistic Structure*. Amsterdam – Philadelphia: John Benjamins.
Corbett, Greville
    2003   Agreement: Terms and boundaries. In *Role of Agreement in Natural Language: Proceedings of the Fifth Annual Texas Linguistics Soci-*

ety, W. E. Griffin (ed.), 109–122. Texas Linguistic Forum 53: http://uts.cc.utexas.edu/~tls/2001tls/2001proceeds.html.

Goldberg, Adele

    1995   *Constructions: a Construction Grammar Approach to Argument Structure*. Chicago: Chicago University Press.

    2006   *Constructions at Work. The Nature of Generalization inLlanguage*. Oxford: Oxford University Press.

Gries, Stefan Th.

    t.a.   Corpus data in usage-based linguistics: What's the right degree of granularity for the analysis of argument structure constructions? In *Expanding Cognitive Linguistic Horizons*; Mario Brdar, Milena Žic Fuchs, and Stefan Th. Gries (eds.), Amsterdam – Philadelphia: John Benjamins.

Lakoff, George

    1987   *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. Chicago – London: University of Chicago Press.

Langacker, Ronald

    1987   *Foundations of Cognitive Grammar, Vol. I: Theoretical Prerequisites*, Stanford,CA: Stanford University Press.

    1988   A usage-based model. In *Topics in Cognitive Linguistics*; Brygida Rudzka-Ostyn (ed.), 127–161. Amsterdam – Philadelphia: John Benjamins.

Lieven, Elena, Heike Behrens, Jennifer Speares, and Michael Tomasello

    2003   Early syntactic creativity: a usage-based approach. *Journal of Child Language* 30: 333–370.

Schmid, Hans-Jörg

    2000   *English Abstract Nouns as Conceptual Shells. From Corpus to Cognition*. Berlin – New York: Mouton de Gruyter.

Sinclair, John M.

    1991   *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Stefanowitsch, Anatol, and Stefan Th. Gries

    2003   Collostructions: investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8 (2): 209–243.

Stefanowitsch, Anatol, and Stefan Th. Gries (eds.)

    2006   *Corpus-based Approaches to Metaphor and Metonymy*. Berlin – New York: Mouton de Gruyter.

Tomasello, Michael

    2000   First steps in a usage based theory of first language acquisition. *Cognitive Linguistics* 11: 61–82.

    2003   *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.

Tuggy, David
    1993   Ambiguity, polysemy and vagueness. *Cognitive Linguistics* 4: 273–290.

Tummers, José, Kris Heylen, and Dirk Geeraerts
    2005   Usage-based approaches in Cognitive Linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory* 1: 225–261.

Wray, Alison
    2002   *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.