Quantitative Methods in Cognitive Semantics: Introduction to the Volume

1. The Empirical Turn

Cognitive linguistics currently experiences a trend towards the use of empirical methods; it can be observed that conferences fill up increasingly with empirical studies supporting and extending the conceptual work in cognitive linguistics. At the same time, several volumes with the explicit aim to lay the foundations for empirical investigations have recently appeared (e.g. Stefanowitsch and Gries 2006; Kristiansen, Achard, Dirven and Ruiz de Mendoza Ibáñez 2006; González-Márquez, Mittelberg, Coulson and Spivey 2007). These new developments concern the relationship of cognitive linguistic findings to other empirical disciplines, such as psychology, corpus- and neurolinguistics, on the one hand and the attempt to achieve additional methodological rigour on the other. With respect to the former, Gibbs (2007), for instance, suggests cognitive linguistics to incorporate more psycholinguistic findings in order to live up to the requirement of cognitive reality. Stefanowitsch (2009) holds that in order for cognitive linguistics to become a cognitive science, it needs to pay more attention to its interfaces with other disciplines, which includes attention to methodology and to generally accepted criteria of scientific investigation, such as intersubjectivity, representativity, reliability, and validity. With regard to the latter, Tummers et al. (2005) suggest that more methodological rigour be needed in cognitive linguistics, and Geeraerts (2006), for instance, explicitly argues for an 'empirical turn' in cognitive linguistics. He evaluates the current situation as a 'theoretical chaos', albeit a creative one, which needs to give room to a situation in which competing approaches can be objectively compared (Geeraerts 2006: 21). The use of quantitative methodologies in cognitive semantics presented i this volume is similarly an attempt at studying cognitive semantics empirically,[1] at connecting cognitive semantics with research in other disciplines, especially corpus linguistics and usage-based approaches to language, and at introducing more methodological rigour into the discipline.

The question now is: How can meaning in cognitive semantics be investigated by means of quantitative methods? Or, as Geeraerts puts it in this volume: "how can meaning, the most

---

1 Some scholars understand this criterion to be defining for scientific studies in general (see, for instance, Stefanowitsch, this volume).

qualitative of all linguistic features, be expressed in numbers?"

In fact there are many who consider the use of quantitative methodologies to be too problematic to produce valid results. Already the definition of the numerator, the item under consideration, may not be trivial; for instance, it may be polysemous, its interpretation may be context-dependent (cf. Bednarek 2008), and even different word forms of the item in question may have different uses (cf. Newman, this volume). The definition of the denominator, the actual and potential contexts in which the item under consideration may occur (cf. Schmid, this volume), and the domain, the context in which the data were observed, i.e. the corpus under consideration, may be similarly problematic (Fischer, this volume). Especially in a constructivist approach such as cognitive linguistics, categories cannot be assumed to be objectively given. Heritage (1995: 400-402) illustrates the problem for the occurrence of the interjection 'oh' as the third turn after responses to questions. While "the third turn after replies to questions" may appear as a useful operationalisation, in practice it is often not trivial to identify objectively which utterances constitute questions and which responses.

Moreover, the problems increase in the realm of meaning; given the cognitive nature of meanings in cognitive semantics, the use of quantitative methodology is not obvious. In general, quantitative methods introduce an external, objectifying perspective on the object of study (Heritage 1995: 406). While this ensures intersubjectivity in the scientific endeavour, it may be problematic for the study of meaning as "the most subjective of all linguistic phenomena" (Geeraerts, this volume).

The current volume illustrates some ways in which applying corpus linguistic methods and statistical analysis to the study of meaning may be fruitful; it presents practical solutions for the methodological problems outlined above and demonstrates procedures and methods for the empirical study of meaning. However, it also addresses the problems arising in a quantitative approach to meaning and the implications the use of quantitative methodologies may have for the development of the discipline. Thus, the volume reflects the struggle of a relatively young discipline determining its future course of action.

2. Quantitative Methods in Cognitive Semantics

In cognitive semantics, meaning is understood as a cognitive phenomenon; Lakoff (1987), for instance, outlines the cognitive linguistic program in contrast to so-called objectivist approaches to language (see also Langacker 1987: 5; Fillmore 1975). Thus, meaning in cognitive linguistics is taken to correspond to dynamic, context-sensitive cognitive construal, for instance:

> "The term conceptualization is interpreted broadly as embracing any kind of mental experience. It subsumes (a) both established and novel conceptions; (b) not only abstract or intellectual 'concepts' but also sensory, motor, and emotive experience; (c) conceptions that are not instantaneous but change or unfold through processing time; and (d) full apprehension of the physical, social, cultural, and linguistic context. Thus, far from being either static or solipsistic, conceptualization is viewed as the dynamic activity of embodied minds interacting with their environments." (Langacker 1998: 3)

Crucial aspects of such construal are, according to Langacker (1998), granularity, metaphor, perspective, and prominence, in particular profiling. These processes guide the way in which meaning is cognitively construed in different linguistic structures.

Langacker's definition of meaning in cognitive semantics also means that linguistic meaning and world knowledge cannot be separated categorically. This view is generally shared in cognitive linguistics; for instance, in various seminal papers (e.g. 1975, 1982, 1985; Fillmore and Atkins 1992) Fillmore outlines the cognitive semantic approach to meaning that he calls a *semantics of understanding*. This comprises "what one knows by virtue of being a speaker of the language" (Fillmore 1985: 252). Fillmore argues that a semantic theory should aim to account for "the relationship between linguistic texts, the context in which they are instanced, and the process and products of their interpretation" (1985: 222). That is, "in determining the situated meanings of uses of the sentence, one interprets the sentence's conventional meaning with its linguistic and extralinguistic context" (1985: 233).

Cognitive semantics thus focuses on conceptualization and understanding; yet, it also makes use of the notion of reference, although this notion is heavily relied on in formal semantic approaches, which are generally rejected from a cognitive linguistic perspective. Reference may provide the central anchor point to compare, for instance, different construals (Berlin and Kay 1969, Geeraerts, Grondelaers and Bakema 1994; Langacker 2008: ch.9).

Finally, usage plays an important role in cognitive semantics. Especially Langacker has been advocating the usage-based approach (e.g. 1987, 1988, 2000, 2008) that holds that meaning emerges from use. Given Fillmore's above definition of the semantics of understanding, expressions are taken to receive their meanings in contexts of use. Such a perspective entails different degrees of schematization, the study of creativity and analogical reasoning, and opens up for the possibility that even different word forms may be entrenched differently. In the same way as meaning emerges from use, language use is taken to be determined by the conceptual meaning by an item, and, in the same way, its reference is determined by its conceptual content.

To sum up, cognitive semantics involves four different aspects of meaning (see Figure 1): conceptualisation, usage, world knowledge and reference. These are taken to be highly interrelated and in many ways inseparable, such that context, meaning, reference and use are taken to co-determine each other.

conceptualisation ——— reference
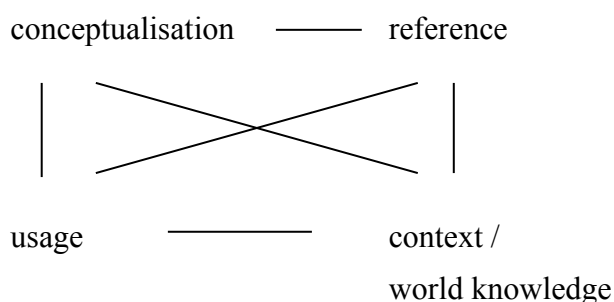

usage ——— context / world knowledge

Figure 1

These different aspects of meaning lend themselves to study by means of quantitative methodologies to different degrees. While conceptualisation is generally inaccessible to direct scientific study, usage can be investigated intersubjectively in collections of usage events, such as linguistic corpora.

Correspondingly, several papers in this volume approach cognitive semantics with corpus linguistic methods; many researchers investigate language usage therefore not in its own right but because they take it as an indicators for something else (more interesting) (cf. Dirk Geeraerts, this volume, and Anatol Stefanowitsch, this volume), namely the conceptual meanings of the items and structures under consideration. Since language use is objectively accessible, it can serve as operationalisation of more subjective factors if the link between the two factors can be unambiguously established. Furthermore, in contrast to many experimental techniques, which study

conceptualisation in off-line experiments, corpora represent the 'results' of on-line language production (cf. Gries and Divjak 2009).

Accordingly, Stefan Gries and Dagmar Divjak argue in this volume that the 'behavioural profile' of a word, i.e. a quantified description of its distribution in large collections of data regarding a set of linguistic variables, reflects its semantics. The underlying assumption is that the exact properties of a concept determine the word's usage as it is manifest in corpora. In their approach, the quantitative behavioural profile may then serve as the basis for further quantitative investigation.

Unlike Stefan Gries and Dagmar Divjak, other researchers do not operate on the quantified results of the quantitative methods directly but rather investigate those linguistic items further that the quantitative analysis identified as significantly attracted to the linguistic structure under consideration (see Timothy Coleman, Martin Hilpert and Stefan Fuhs, this volume, as well as Stefanowitsch and Gries to appear). Similarly, Kerstin Fischer (this volume) uses correlating features of the linguistic contexts in comparable corpora as indicators for speakers' conceptualisations of the communication situation, which in turn provide evidence for the functions of the constructions under consideration. A fourth example of quantitative methods focusing on language use is Hans-Joerg Schmid's investigation of the relationship between objectively observable frequency and the cognitive linguistic concept of entrenchment. Entrenchment, like conceptionalization, is not objectively accessible, and thus needs to be studied indirectly. Schmid's investigation concerns the question whether frequency, as an observable feature of language use, can be interpreted as an indicator of the mental phenomenon entrenchment.

However, to employ quantitative methods to investigate language use in order to study meaning is only one of several uses of quantitative methods in cognitive semantics illustrated in this volume. Dirk Geeraerts, for instance, argues that quantifying, and even statistical, procedures are merely a consequence of using empirical methods in general (this volume).

A very different use of quantitative methodology is illustrated in Stefanie Wulff's paper; she uses quantitative methods to develop a mathematical model that describes and predicts compositionality in V-NP constructions. That is, she uses results from corpus linguistic investigations to quantify the contributions of the words inside a construction and its degree of compositionality and then describes their relationship in a mathematical model.

Dagmar Divjak's paper illustrates another use of quantitative methods in cognitive semantics; the author uses logistic regression to measure the influence of particular factors in

linguistic choice. That is, the quantitative procedure she chooses allows her to identify those variables that predict a particular linguistic form best, enabling her to distinguish between competing linguistic analyses of the phenomenon under consideration.

Finally, quantitative methods may also be helpful for investigating general linguistic issues such as the determinants of meaning extension. Arne Zeschel illustrates how quantitative methods can be applied to address the relationship between established and novel semantic types in order to explain the productivity of certain word senses, i.e. the degree with which they allow semantic extensions and analogical formations. By showing that there is a high statistical correlation between the numbers of established semantic types and novel instantiations, he provides empirical support for an important theoretical concept.

To conclude, quantitative methodology may play several different roles in cognitive semantics, supporting the scientific process in numerous different ways. While the most prominent use is the investigation of language usage to identify conceptual meanings or entrenchment, quantitative methods may also serve to model findings and to contribute to theory development by providing means to test hypotheses, to evaluate different theoretical models and to identify those factors that describe the observable patterning of the data best.

3. Quantitative Methods for Cognitive Semantics

Both John Newman and Hans-Joerg Schmid discuss (in this volume) the problems of introducing methods into cognitive linguistics which originate from another discipline and which may be too complicated for cognitive linguists at large or which may not be psycholinguistically plausible. This issue is similar to the role of psycholinguistic methods in cognitive linguistics discussed by Gibbs (2007). Gibbs holds that even though cognitive linguistics should incorporate psycholinguistic findings more, cognitive linguists should not all try to become psychologists now; instead, the different competences should remain in the different disciplines. Applying this proposal to quantitative cognitive semantics would mean that cognitive linguists do what they have been doing for the last thirty years (which is very much based on linguistic intuitions, cf. Geeraerts and Stefanowitsch in this volume) and leave it to a statistician to test cognitive linguistic hypotheses in quantitative investigations.

None of the authors of this volume suggest this, and on the contrary, many claim that cognitive semantics needs to rely on empirical methods more, while others demonstrate the usefulness of such methods for various problem areas in cognitive semantics. However, it is important to bear in mind that the methodological discussion needs to be understandable to the 'end user', i.e. the cognitive linguist, as John Newman argues. The current volume tries to bridge this gap by making a set of quantitative methods more accessible to the reader and by illustrating their advantages and uses. Moreover, most of the methods employed in this volume are rather 'standard' procedures such that they are available with most office software or free to download; in any case, the introduction provided here is intended to promote at least a passive understanding of the methods for the interested reader.

The most commonly used method in this volume is the collostructional method. This method covers in fact three different procedures, collexeme analysis (Stefanowitsch and Gries 2003), distinctive-collexeme analysis (Gries and Stefanowitsch 2004) and covarying collexeme analysis (Stefanowitsch and Gries 2005). The method used in the volume by Timothy Coleman, Martin Hilpert and Stefan Fuhs is the collexeme analysis. In this method, the frequency of a word in a given construction is compared to its occurrence in other constructions and to the frequency of the construction in the whole corpus. The resulting measure, the collostructional strength, describes the association of a word in a construction and thus can serve as starting point for further quantitative and qualitative investigations. This method is also discussed in Hans-Joerg Schmid's article in detail in comparison with some other measures of frequency in a corpus.

In addition to collexeme analysis, other statistical procedures are employed in the volume. For instance, Stefan Gries and Dagmar Divjak apply a method from unsupervised learning to the quantitative distributions they identify, cluster analysis. In this method, all instances in a data set are compared to all others. The algorithm then determines a distance between each data point and sorts the instances into a hierarchical set of classes, i.e. in a tree-like representation, the dendrogram, where spatial proximity represents similarity. This method can, for instance, be used to organise results from a collexeme analysis into semantic classes (cf. Gries and Stefanowitsch 2010) or to detect prototypical word classes (Gries and Divjak, this volume).

Another statistical method used in this volume is correlation analysis; correlation describes the strength of the relationship between two variables, suggesting possible causal relationships. The correlation coefficient used in the papers by Arne Zeschel and Kerstin Fischer in this volume is the most common one, the Pearson product moment coefficient, which describes the strength of a linear

relationship between two data sets. For instance, in a strong positive correlation, if x is high, y will also be high; alternatively, in a strong negative correlation, if x is high, y will be low. The Pearson product moment correlation coefficient is part of most standard statistical analysis software.

Moreover, Dagmar Divjak uses logistic regression in her study. Regression analysis measures the influence of a particular variable in a set of independent variables. That is, it analyses the contribution of a particular variable to the behaviour of an independent variable if all other variables are kept constant. It can thus be used as an exploratory method, identifying which independent variable influences the results of the dependent variable the most, but it is also a powerful technique from confirmatory statistics. Dagmar Divjak uses this method to identify the contribution of a set of hypothesized factors for the absence or presence of imperfective versus perfective aspect in Russian.

The methods used and discussed in this volume do not constitute an exhaustive list of quantitative methods that may be usefully applied to cognitive semantics; many other statistical (e.g. Glynn 2009 and this volume) and computational (e.g. Fischer 2000; Elman 2006) methods have been used in the area of cognitive semantics, and so this volume may only illustrate a commonly used subset of quantitative methods for cognitive semantics. The focus of the volume is rather on determining the potential of quantitative methodologies for the description of meaning and to determine the path cognitive semantics is going to take in the near future.

4. Structure of the Volume

The volume addresses the use of quantitative corpus methodologies for the study of cognitive semantics. It intends to combine high quality, up-to-date research on a broad range of cognitive semantic questions.

The volume starts with an outline of the history and principles of corpus cognitive semantics. Following this overview, it presents a discussion of those assumptions that may so far have prevented cognitive semanticists from engaging in quantitative methodologies. Particularly for the domain of semantics, the use of quantitative methodologies is not obvious, and especially the first two and the last two papers address the particular methodological problems arising. The first article sets the scene for rigorous hypothesis formation and testing, which is put to use in the

following articles. The second article concerns some detailed methodological issues a cognitive semantician using quantitative methodologies should bear in mind.

Following these stage-setting articles is a set of articles that address central concepts in semantic research: semantic creativity, situation-dependency, force-dynamics and coercion. These articles show how the discussion of these central semantic concepts, which also lie at the core of cognitive linguistics, can profit from a quantitative approach.

The following four articles provide in depth studies of selected methodological issues in cognitive semantics, combining detailed empirical investigation with methodological discussion. They concern the study of compositionality, the relationship between aspect and modality, lexical semantics and the multifaceted semantics of an argument structure construction respectively, which are discussed with respect to a variety of different languages.

Finally, the volume is concluded by two theoretical articles; the first one addresses several reservations previously expressed against the use of quantitative methods in cognitive semantics and proposes the notion of behavioural profile to frame theoretically how quantitative approaches to experiential categories, such as meaning, may be investigated. The second one outlines a general approach to empirical cognitive semantics.

In particular, the volume starts out with an overview over the development of the scientific discipline of corpus cognitive semantics by **Dylan Glynn**. It is followed by a stage-setting article by **Dirk Geeraerts**, entitled "The Doctor and the Semantician" in which he first addresses the question whether meaning can be studied empirically at all and then discusses false conceptions why scholars may think they need not address semantics using empirical methodologies. He continues by suggesting to identify indices for the semantic concept in question that are easier to establish than the concept under consideration itself. From there, he outlines a scientific procedure at which scholars in empirical cognitive semantics may orient. He suggests to proceed in cycles of hypotheses and testing, developing the theory in several steps.

John Newman's "Balancing Acts: Empirical Pursuits in Cognitive Linguistics" addresses some of the issues that scholars intending to use quantitative methods need to consider. In particular, he considers those assumptions often not addressed in corpus-linguistic research. For instance, while the usage-based approach is spreading in cognitive semantics, the concept of 'situated instances of language use' is generally restricted to linguistic cotext and excludes paralinguistic and suprasegmental linguistic features, non-verbal behaviours and aspects of the context of situation. Moreover, large-scale corpus-linguistic investigations usually concern written

rather than conversational language. John Newman continues by considering the problems of ambiguous tags and of lemma-based corpus queries which often neglect that different word forms may be entrenched differently. In all, he argues for a balance between different kinds of methodologies.

Some of these considerations recur in **Hans-Joerg Schmid**'s "Does Frequency in Text Instantiate Entrenchment in the Cognitive System?"; he discusses different measures for frequency and compares calculations of collostructional strength (Stefanowitsch and Gries 2003) with the possibly more intuitive measures *attraction* and *reliance*; he concludes that subjective judgements may enter the calculations of the seemingly objective measure *frequency*. Furthermore, he argues that the relationship between frequency and entrenchment is more complex than usually assumed and still in need of further research that integrates both corpus-linguistic and psycholinguistic findings.

The theoretical section on corpus methods in cognitive semantics is followed by a section in which the authors address specific theoretical issues. **Stefan Fuhs** discusses in his article "Aspectual Types across Predicates in the English Durative Construction" the relationship between aspectual composition and coercion. He uses collostructional analysis (Stefanowitsch and Gries 2003) to investigate the English durative construction. While the results show that the construction is significantly associated with atelic verbs, as many as 25% of the data exhibit coercion effects, where the construction causes generally telic verbs to become atelic. His study therefore supports a constructional perspective on meaning construal in which both the verb and the construction contribute meaning aspects and illustrates how quantitative methods can contribute to shedding light into compositional semantics.

In "The Force Dynamics of English Complement Clauses: A Usage-based Account", **Martin Hilpert** investigates English gerund classes with infinitive complements using a collexeme (Stefanowitsch and Gries 2003) analysis to identify the lexical material associated with this construction. The results from the quantitative study are then investigated in terms of Talmy's (2000) force dynamics. The quantitative evidence supports the qualitative analysis showing that the construction under consideration is primarily used to express force dynamic meanings and that judgements on the well-formedness of novel examples can be predicted on the basis of the model developed.

In "Accounting for the Role of Situation in Language Use in a Cognitive Semantic Representation of Sentence Mood", **Kerstin Fischer** addresses how both quantitative and

qualitative methods can be used to establish speakers' different construals of an objectively identical situation. Situation is a notoriously difficult concept in cognitive semantics because cognitive semantics assumes categories to be subjectively construed by language; a quantitative approach may provide the missing link since quantitative methods may correlate the differences in subjective construal with different sets of linguistic choices. Fischer illustrates her proposal on the choice of grammatical mood in task-oriented interactions. The interdependencies between situational construal and grammatical choice are then modelled in a computational cognitive semantic representation of grammatical mood.

Arne Zeschel's article "Exemplars and Analogy: Semantic Extension in Constructional Networks" addresses the question of how speakers generalise from conventional, learned expressions to a productive constructional schema that allows the creation of new, unconventional instances. He investigates the question by considering the relationship between the number of already established types per semantic reading and novel instances in an increasingly more fine-grained semantic analysis. He demonstrates a procedure that takes great care to avoid pitfalls like those discussed in the papers by Newman and Schmid. He finds for the three semantic levels of analysis large positive correlations between the semantic type frequency of established uses and the number of creative extensions, allowing conclusions regarding schematization based on semantic types.

The next section concerns specific methodological questions. It is lead by **Stefanie Wulff**'s paper "Marrying cognitive-linguistic and corpus-based methods: On the compositionality of English V NP-idioms" which focuses on a core semantic issue, compositionality. Cognitive linguistics suggests that syntactic structures be signs similar to lexical items, combining form and meaning directly. In such an approach, both lexical items and constructions contribute to the meaning of the whole construct, and they may do so to different degrees. Stefanie Wulff now develops a quantitative measure based on a large scale corpus linguistic investigation that takes all of the component words of a construction into consideration. This measure reflects differences is entrenchement of instantiations of the V NP-construction (e.g. *make the headway, call the police, tell the story*), using the number of shared collocates as a measure for semantic similarity.

The following paper illustrates the use of quantitative methods for the study of lexical semantics. **Dylan Glynn** examines the polysemy of the English verb *bother* and offers an example of multifactorial research coupled with fine-grained semantic analysis. He focuses on one of the basic hurdles of corpus-driven cognitive semantics – the subjective nature of manual semantic

coding. He argues that even highly subjective semantic coding can be verified using confirmatory statistical analysis. Although objectivity may be impossible with this approach, verification may be the key to enabling the field to advance in this direction. Glynn argues that statistical modelling may further verify the results.

In "Beyond the Dative Alternation: The Semantics of the Dutch *aan*-Dative", **Timothy Colleman** addresses the Dutch prepositional dative construction from a cognitive semantic perspective, thus focusing not on its distinctive features that distinguish its meaning from the near-synonymous double object construction, but instead providing a semantic description of the conceptual meaning of the construction under consideration. He uses a collexeme analysis (Stefanowitsch and Gries 2003) to identify the verb attracted to the construction under consideration. The ranking reveals a broad semantic range in which possessional transfer plays a major role – however, by far not the only one. He therefore introduces a multidimensional analysis (cf. Geeraerts 1998) which starts with the definition of the semantic core of the construction and proceeds by identifying metaphorically motivated sense extensions along certain dimensions, in this case, the two dimensions *contact* and *direction*, resulting in a family of 'caused contact' senses.

In her article "Corpus-based Evidence for an Idiosyncratic Aspect-Modality Relation in Russian", **Dagmar Divjak** discusses the relationship between aspect and modality in Russian, evaluating two competing theoretical proposals; her corpus linguistic analysis not only supports one of the competing models empirically, but is also in accordance with cognitive semantic analyses of the Russian aspectual system. She thus shows how quantitative analyses outperform intuition-based studies and provides the necessary correction of previous hypotheses in linguistic typology, supporting careful hand-crafted studies and tying in with cognitive models of aspect in Russian.

The volume is concluded with two papers that address the role of quantitative methods in cognitive semantics in general. **Stefan Gries and Dagmar Divjak** take up criticisms and reservations against the use of quantitative methods in cognitive linguistics. After addressing each of these reservation, they outline the behavioural profile approach in their chapter "Quantitative Approaches in Usage-based Cognitive Semantics: Myths, Erroneous Assumptions, and a Proposal" to illustrate its use for the investigation of two of the core problems in cognitive semantics: polysemy and the identification of prototypical word senses. A behavioural profile consists of the percentages of (co-)occurrence of a word or sense with respect to ID tags, that is, certain morphological, syntactic and semantic properties of a word and a context. The authors continue by showing how such a profile can be used by means of a cluster analytic approach to address

polysemy, near-synonymy and contrastive semantics. The results of a behavioural profile approach can moreover be compared with experimental evidence.

Finally, **Anatol Stefanowitsch** discusses in "Empirical Cognitive Semantics: Some Thoughts" the development of cognitive semantics 'from an art to a science', in which the use of empirical methodologies plays a key role. He introduces the concept of operationalization and illustrates how this concept can be applied to research in cognitive semantics focusing on one of four kinds of meaning: meaning as concept, proposition, reference and use. He then considers what consequences the empirical approach resting on the operationalization of meaning may have on the development of cognitive semantics and concludes that this move will change the discipline radically while providing cognitive semantics with an empirically grounded, scientific basis.

The research combined in this volume presents innovative approaches to genuinely semantic problems by means of quantitative methods, some of which developed especially for the current purposes, others applied from other research areas to the study of semantic phenomena. The volume also provides an overview of different quantitative methodologies for the study of cognitive semantics (for example, collexeme, correlation and regression analysis) and a discussion of the theoretical and methodological consequences of the quantitative approaches taken.

References

Bednarek, Monika   2008   Semantic Preference and Semantic Prosody Re-Examined. *Corpus Linguistics and Linguistic Theory* 4, 2: 119-139.

Berlin, Brent and Paul Kay   1969   *Basic Color Terms. Their Universality and Evolution.* Berkeley, CA: University of California Press.

Elman, Jeff   2006   Computational approaches to language acquisition. In: K. Brown (ed.), *Encyclopedia of Language & Linguistics*, pp. 726-732. Second Edition, Volume 2, Oxford: Elsevier.

Fillmore, Charles J.   1975   An Alternative to Checklist Theories of Meaning. *Proceedings of BLS* 1: 123-131.

Fillmore, Charles J.   1976   Frame semantics and the nature of language. In: *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and*

Speech 280: 20-32.

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

Fillmore, Charles J.   1988   The mechanisms of 'construction grammar'. *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistic Society* 14: 35-33.

Fillmore, Charles J., Paul Kay and Mary O'Connor   1988   Regularity and idiomaticity in grammatical constructions: the case of let alone. *Language* 64, 3: 501–538.

Fillmore, Charles J. and Beryl T. Atkins   1992   Towards a frame-based lexicon: the case of RISK. In: Adrienne Lehrer and Eva Kittay (eds.), *Frames, Fields, and Contrasts*, Erlbaum Publishers, pp. 75–102.

Fischer, Kerstin   2000   *From Cognitive Semantics to Lexical Pragmatics: The Functional Polysemy of Discourse Particles*. Berlin/New York: Mouton de Gruyter.

Geeraerts, Dirk   1998   The semantic structure of the indirect object in Dutch. In: WillyVan Langendonck and William Van Belle (eds.), *The Dative II: Theoretical Approaches*, pp. 185-210. Amsterdam: John Benjamins.

Geeraerts, Dirk   2007   Methodology in Cognitive Linguistics. In: Mónica González-Márquez, Irene Mittelberg, Seana Coulson and Michael J. Spivey (eds.), *Methods in Cognitive Linguistics*. Amsterdam: Benjamins.

Geeraerts, Dirk, Stefan Grondelaers and Peter Bakema   1994   The structure of lexical variation. Meaning, naming, and context. Berlin/New York: Mouton de Gruyter.

Gibbs, Raymond W. Jr.   2007   Why Cognitive Linguistics Should Care More about Empirical Methods. In: González-Márquez, Mónica, Irene Mittelberg, Seana Coulson and Michael J. Spivey (eds.), *Methods in Cognitive Linguistics*. Amsterdam: Benjamins.

Glynn, Dylan 2009   Polysemy, syntax and variation. In: Vyvyan Evans and Stephanie S. Pourcel (eds.), *New Directions in Cognitive Linguistics*, pp. 77-104. Amsterdam/Philadelphia: John Benjamins.

González-Márquez, Mónica, Irene Mittelberg, Seana Coulson and Michael J. Spivey (eds.) 2007   *Methods in Cognitive Linguistics*. Amsterdam: Benjamins.

Gries, Stefan and Anatol Stefanowitsch   2004   Extending collostructional analysis: A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics* 8: 31-61.

Gries, Stefan and Dagmar Divjak 2009 Behavioral Profiles: A Corpus-based Approach to Cognitive Semantic Analysis. In: Vyvyan Evans and Stephanie S. Pourcel (eds.), *New Directions in Cognitive Linguistics*, pp. 57-75. Amsterdam/Philadelphia: John Benjamins.

Gries, Stefan Th. and Anatol Stefanowitsch 2010 Cluster Analysis and the Identification of Collexeme Classes. In: John Newman and Sally Rice (eds.): *Empirical and Experimental Methods in Cognitive/Functional Research,* pp. 73-90. Stanford: CSLI.

Heritage, John 1995 Conversation Analysis: Methodological Aspects. In: Uta M. Quasthoff (ed.), *Aspects of Oral Communication*, pp.391-418. Berlin: De Gruyter.

Heylen, Kris, Yves Peirsman and Dirk Geeraerts 2008 Automatic synonymy extraction. A Comparison of Syntactic Context Models. LOT Computational Linguistics in the Netherlands 2007, pp.101-116

Kristiansen, Gitte, Michel Achard, René Dirven and Francisco José Ruiz de Mendoza Ibáñez (Eds.) 2006 *Cognitive Linguistics: Current Applications and Future Perspectives*. Berlin: Mouton de Gruyter.

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

Langacker, Ronald W. 1987 *Foundations of Cognitive Grammar, Volume 1: Theoretical Prerequisites*, Stanford: Stanford University Press.

Langacker, Ronald W. 1991 *Concept, Image, Symbol*. Berlin/New York: Mouton de Gruyter.

Langacker, Robald W. 2000 A dynamic usage-based model. In: Michael Barlow and Suzanne Kemmer (eds.), *Usage-Based Models of Language*, pp. 1-63. Stanford: CSLI.

Langacker, Ronald 2008 *Cognitive Grammar: A Basic Introduction*. New York: Oxford University Press.

Stefanowitsch, Anatol 2009 Cognitive Linguistics as a Cognitive Science. Keynote Lecture at the Conference on Bi-directional Perspectives in the Cognitive Sciences in Marburg, Germany, March 2nd, 2009, available at http://www-user.uni-bremen.de/~anatol/.

Stefanowitsch, Anatol and Stefan Th. Gries 2003 Collostructions: Investigating the interaction of words and constructions. Proceedings of the 18th International Conference on Computational Linguistics, pp. 747-753.

Stefanowitsch, Anatol and Stefan Th. Gries 2005 Co-varying Collexemes. *Corpus Linguistics and Linguistic Theory* 1, 1: 1-46.

Stefanowitsch, Anatol and Stefan Th. Gries. 2008 Channel and constructional meaning: A collostructional case study. In: Gitte Kristiansen and René Dirven (eds), *Cognitive Sociolinguistics: Language Variation, Cultural Models, Social Systems*, pp. 129–152. Berlin and New York: Mouton de Gruyter.

Stefanowitsch, Anatol and Stefan Th. Gries (eds.) 2006 *Corpora in Cognitive Linguistics: Corpus-based Approaches to Syntax and Lexis* (Trends in Linguistics: Studies and Monographs), Berlin: Mouton de Gruyter.

Tummers, Jose, Kris Heylen and Dirk Geeraerts 2005 Usage-based Approaches in Cognitive Linguistics: A Technical State of the Art. *Corpus Linguistics and Linguistic Theory* 1, 2: 225-261.