

Evam Kidd
ANU.

Converging Evidence

Methodological and theoretical issues
for linguistic research

Edited by

Doris Schönefeld
Universität Leipzig

John Benjamins Publishing Company
Amsterdam / Philadelphia

Human Cognitive Processing (HCP) Cognitive Foundations of Language Structure and Use

This book series is a forum for interdisciplinary research on the grammatical structure, semantic organization, and communicative function of language(s), and their anchoring in human cognitive faculties.

For an overview of all books published in this series, please see

<http://benjamins.com/catalog/hcp>

Editors

Klaus-Uwe Panther
University of Hamburg

Linda L. Thornburg

Editorial Board

Bogusław Bierwiaczonek
University of Economics and Humanities,
Poland

Elżbieta Górka
University of Warsaw

Mario Brdar
Josip Juraj Strossmayer University, Croatia

Martin Hilpert
Freiburg Institute for Advanced Studies

Barbara Dancygier
University of British Columbia

Zoltán Kövecses
Eötvös Loránd University, Hungary

N.J. Enfield
Max Planck Institute for Psycholinguistics,
Nijmegen & Radboud University Nijmegen

Teenie Matlock
University of California at Merced

Elisabeth Engberg-Pedersen
University of Copenhagen

Günter Radden
University of Hamburg

Ad Foolen
Radboud University Nijmegen

Francisco José Ruiz de Mendoza Ibáñez
University of La Rioja

Raymond W. Gibbs, Jr.
University of California at Santa Cruz

Doris Schönefeld
University of Leipzig

Rachel Giora
Tel Aviv University

Debra Ziegeler
Paul Valéry University, France

Volume 33

Converging Evidence. Methodological and theoretical issues
for linguistic research

Edited by Doris Schönefeld



The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences - Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Converging evidence : methodological and theoretical issues for linguistic research / edited by Doris Schönefeld.

p. cm. (Human Cognitive Processes, ISSN 1387-6724 : v. 33)

Includes bibliographical references and index.

1. Linguistics--Research--Methodology. I. Schönefeld, Doris, 1953-

P126.C676 2011

410.72--dc23

ISBN 978 90 272 2387 6 (Hb ; alk. paper)

ISBN 978 90 272 8451 8 (Eb)

2011029002

© 2011 - John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands

John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

Contributors	vii
Preface	ix
Introduction: On evidence and the convergence of evidence in linguistic research <i>Doris Schönefeld</i>	1
Issues in collecting converging evidence: Is metaphor always a matter of thought? <i>Gerard J. Steen</i>	33
PART 1. Multi-methodological approaches to constructional and idiomatic meaning	
1.1 <i>Cognition verb constructions</i>	
Perception and conception: The 'see x to be y' construction from a cognitive perspective <i>Thomas Egan</i>	57
Explaining diverging evidence: The case of clause-initial <i>I think</i> <i>Gunther Kaltenböck</i>	81
1.2 <i>Constructional alternatives</i>	
<i>I am about to die</i> vs. <i>I am going to die</i> : A usage-based comparison between two future-indicating constructions <i>Silke Höche</i>	115
Studying syntactic priming in corpora: Implications of different levels of granularity <i>Stefan Th. Gries</i>	143

Islands of (im)productivity in corpus data and acceptability judgments: Contrasting two potentiality constructions in Dutch <i>Ad Backus and Maria Mos</i>	165
1.3 <i>Idioms and creative language use</i>	
Compositional and embodied meanings of somatisms: A corpus-based approach to phraseologisms <i>Alexander Ziem and Sven Staffeldt</i>	195
Word-formation patterns in a cross-linguistic perspective: Testing predictions for novel object naming in Hungarian and German <i>Susanne R. Borgwaldt and Réka Benczes</i>	221
PART 2. Multi-methodological approaches to language acquisition	
The interaction of function and input frequency in L1-acquisition: The case of <i>was...für</i> 'what kind of...' questions in German <i>Rasmus Steinkrauss</i>	249
Relative clause acquisition and representation: Evidence from spontaneous speech, sentence repetition, and comprehension <i>Silke Brandt and Evan Kidd</i>	273
Converging evidence in the typology of motion events: A corpus-based approach to interlanguage <i>Nina Reshöft</i>	293
PART 3. Multi-methodological approaches to the study of discourse	
Differences in the use of emotion metaphors in expert-lay communication: Converging evidence from two complementary studies <i>Anke Beger</i>	319
Index	349

Contributors

Ad Backus Tilburg University A.M.Backus@uvt.nl Postal address: Faculty of Humanities P.O. Box 90153 5000 LE Tilburg, The Netherlands	Bienroder Weg 80 D-38106 Braunschweig Silke Brandt University of Basel silke.brandt@unibas.ch Postal address: Englisches Seminar Nadelberg 6 CH-4051 Basel, Switzerland
Anke Beger Flensburg University ankebege@web.de Postal address: Englisches Seminar Auf dem Campus 1 D-24943 Flensburg, Germany	Thomas Egan Hedmark University College Thomas.Egan@hihm.no Postal address: P.b. 4010 Bedriftssenteret 2306 Hamar, Norway
Réka Benczes Eötvös Loránd University rbenczes@gmail.com Postal address: School of English and American Studies H-1088 Budapest Rákóczi út 5, Hungary	Stefan Th. Gries University of California, Santa Barbara stgries@linguistics.ucsb.edu Postal address: South Hall 3431/3432 University of California, Santa Barbara Santa Barbara, CA 93106-3100, USA
Susanne Borgwaldt University of Braunschweig s.borgwaldt@tu-braunschweig.de Postal address: Institut für Germanistik Technische Universität Carolo Wilhelmina zu Braunschweig	Silke Höche University of Hannover silke.hoeche@engsem.uni-hannover.de Postal address: Englisches Seminar/Englische Sprachwissenschaft Königsworther Platz 1 D-30167 Hannover, Germany

- Gunther Kaltenböck**
University of Vienna
gunther.kaltenboeck@univie.ac.at
Postal address:
Institut für Anglistik
und Amerikanistik
Spitalgasse 2-4
A-1090 Wien, Austria
- Evan Kidd**
University of Manchester
evan.j.kidd@manchester.ac.uk
Postal address:
School of Psychological Sciences
Coupland 1
Oxford Road M13 9PL
Manchester, UK
- Maria Mos**
Tilburg University
maria.mos@uvt.nl
Postal address:
Department of Communication
and Information Sciences
Room D 411
P.O. Box 90153
5000 LE Tilburg, The Netherlands
- Nina Reshöft**
University of Bremen
n.reschoeft@uni-bremen.de
Postal address:
FB 10 Sprach- und
Literaturwissenschaften
GW2 Raum A3380
Bibliothekstr. 1
D-28359 Bremen, Germany
- Doris Schönefeld**
University of Leipzig
schoenefeld@uni-leipzig.de
- Postal address**
Institut für Anglistik
Beethovenstraße 15
D-04107 Leipzig, Germany
- Sven Staffeldt**
University of Würzburg
sven.staffeldt@uni-wuerzburg.de
Postal address:
Lehrstuhl für deutsche
Sprachwissenschaft
Am Hubland
D-97074 Würzburg, Germany
- Gerard J. Steen**
University of Amsterdam
gj.steen@let.vu.nl
Postal address:
Faculty of Arts, 11A-35
Department of Language
and Communication
VU University Amsterdam
De Boelelaan 1105
1081 HV Amsterdam, The Netherlands
- Rasmus Steinkrauss**
University of Groningen
E-mail: r.g.a.steinkrauss@rug.nl
Postal address:
Center for Language and Cognition
Oude Kijk in 't Jatstraat 26
9712 EK Groningen
- Alexander Ziem**
University of Düsseldorf
ziem@phil-fak.uni-duesseldorf.de
Postal address:
Heinrich-Heine-Universität Düsseldorf
Germanistische Sprachwissenschaft
Universitätsstr. 1
D-40225 Düsseldorf, Germany

Preface

This volume contains a selection of papers that are based on presentations given at the 3rd International Conference of the German Cognitive Linguistics Association held at Leipzig University, September 25–27, 2008. The conference was titled 'Converging evidence', and aimed at discussing and opening up research perspectives that can promote and produce new insights in the field of cognitive (usage-based) linguistics. The focus was on the presentation of (cognitively oriented) research resulting from the application of empirical methods, more specifically, from corpus-linguistic methods and experimental work.

The 13 papers collected in the volume demonstrate exactly this research agenda. They all report on findings obtained from corpus analyses and experimental tasks, and can be seen as textbook examples of how to conduct (empirical) research within the cognitive-linguistic framework. The papers have gone through a selective peer-review process intended to yield a high quality research product.

Credit for the quality of the papers goes first and foremost to the authors. They contributed a great deal of effort and creativity to produce this work, and I would very much like to thank them for their excellent cooperation in writing and re-writing their inspiring texts.

Credit also goes to the 30 reviewers, who donated time from busy schedules to carefully read and evaluate the submissions, each of which underwent two reviews. The appointed reviewers were (in alphabetical order): Kirsten Abboth-Smith, Michel Achard, Inbal Arnon, Johanna Barðdal, John Barnden, Silke Brandt, Cristiano Broccias, Paula Chesley, Alan Cienki, Timothy Colleman, Alexis Dimitriadis, Dagmar Divjak, Susanne Gahl, Volker Gast, Dirk Geeraerts, Gaetanella Gilquin, Stefan Gries, Beate Hampe, Martin Hilpert, Willem Hollmann, Florian Jaeger, Celeste Kidd, Andreas Langlotz, Cornelia Müller, Sae Oshima, Klaus-Uwe Panther, Neal Snider, Benedikt Szmeccanyi, Stefanie Wulff, and Arne Zeschel. I cannot thank them enough for writing thoughtful reviews on the contributions' strengths and weaknesses alike, thus stimulating a creative phase of revision.

Credit for the final polish and acceptance of the volume as a whole goes to the HCP series editors Linda Thornburg and Klaus-Uwe Panther. Without their constructive and helpful criticism the book would not have come into existence.

Doris Schönefeld
Leipzig, Germany
May 2011

Introduction

On evidence and the convergence of evidence in linguistic research*

Doris Schönefeld
University of Leipzig

1. Introduction

The kind of evidence employed in linguistic research is determined by the researcher's linguistic orientation and adherence to a particular theory, along with its respective methodological commitments. This also holds for the requirement of converging evidence, that is, of evidence or findings that converge on one and the same conclusion. Commonly, converging evidence is assumed to emerge when a phenomenon is analyzed from more than one methodological perspective (cf. Gries, Hampe, and Schönefeld 2005, 2010). In a discussion of evidence arising from such multi-methodological research, Arppe et al. (2010) elaborate on the benefits and the problems of such a research strategy and make a strong case for the combination of different types of evidence in the analysis of language. At the same time they point at a major challenge in this respect: to ensure the comparability of the concepts and constructs under analysis as well as of aspects considered to be informative about these, such as various types of frequency, reaction times, or ratings. Steen (2007: 21, this volume) identifies a second type of converging evidence that results from 'phenomenological pluralism'. The second type encompasses evidence gained from analyses of an allegedly identical phenomenon in different domains and is problematic in that it needs to be ensured that the phenomena are the same indeed (cf. below and Steen, this volume).

The concern with methodological issues, in particular the question of what kinds of evidence can and should be used in linguistic research, has been an almost constant companion in discussions about the definition of linguistics and its

* I would like to thank the two anonymous referees and the two series editors for their extensive and stimulating feedback on earlier versions of this chapter, which made it possible to improve it in a substantial number of respects. Any remaining insufficiencies are entirely my own.

philosophical position. The notion of converging evidence is a fairly recent one, and the question of what it may add to linguistic research was raised as a result of the application of empirical research methods in linguistics where the formulation of testable hypotheses and their verification are of central concern.

This introductory chapter aims at specifying some of the major issues and problems connected with evidence drawn from various sources. They are discussed in connection with the linguistic frameworks preferably employing them, keeping a constant eye on situations in which the convergence of evidence may become an issue. As a necessary prerequisite for such an endeavor, Section 2 specifies the notion of converging evidence. Section 3 looks at introspective evidence, favored in *Generative Linguistics*. Included there is a discussion of empiricism in *Generative Linguistics*, which may come as a surprise, since this connection seems to be a contradiction in terms from the perspective of the paradigm's rationalist origin. Section 4 is concerned with empirical evidence, which is discussed in connection with the linguistic framework of cognitive linguistics, as an essential characteristic of corpus-linguistic work and for its role in truly usage-based cognitive linguistics. This Section is central to the book, with the chapters collected here focusing on and demonstrating the use of empirical data in (cognitively oriented) linguistic analyses. Section 5 proceeds to summarize what kinds of 'facts' and experience – data in the broadest sense – are at the linguist's disposal for making discoveries, before – as a kind of wrap up of the discussion – Section 6 motivates the need for evidence to converge. Section 7 gives a brief overview of the case studies assembled in this book in support of the argument about converging evidence. Finally, Section 8 considers some prospects for future research.

2. The potential of evidence to converge

In addition to the distinction of converging evidence according to its source (gained by different methods, or gained from different domains), it is necessary to discuss briefly and specify when the convergence of evidence is an issue.¹ The methods available in linguistic research are, roughly speaking, the employment of speaker intuitions and introspection, corpus data, experimentally elicited data, physiological data, and statistical methods (for more details see Section 5).

Looking at the first, 'non-empirical' method from the perspective of convergence of evidence, we can conceive of two situations where it matters. This is, firstly, when the analysis of a linguistic phenomenon does not rest on the linguist's

intuitions alone, but when other people's intuitions are taken into account as well, as is common practice when theorizing is tested against grammaticality judgments. Secondly, convergence will definitely become an aspect to consider when analysts are willing to test their intuitions against (empirical) facts of the language under analysis, usage data, and experimental facts.

As far as 'empirical' methods are concerned, e.g. the use of corpus data and the elicitation of experimental data, the question of their comparative value arises in a way that is similar to the way evidence arises from phenomenological pluralism: Just as we must be careful in deciding whether evidence from different domains can be understood to stand for the same phenomenon tested (cf. below and Steen, this volume), corpus evidence and experimental evidence found or elicited for 'the same' hypothesis are not automatically about the same phenomenon either. Thus, convergence is not necessarily the default situation when corpus and experimental data are drawn on in hypothesis testing. We may also be faced with non-convergence, which, nevertheless, need not always falsify the hypothesis at issue. This may have various reasons. Firstly, such diverging evidence may be due to the different nature of the two types of data: corpus data reflect language as product, whereas experimental data, tapping language processing online, more directly reflect language as process. Hence, if a hypothesis/prediction is concerned with the status as a mental construct of a particular (linguistic) unit/phenomenon, it is the latter kind of data that has more weight and can be taken as evidential in this respect, and corpus data need not necessarily converge on that. On the other hand, if corpus data allow for the verification of a hypothesis, experimental evidence may diverge, if, for example, the factors assumed to account for the data are not the ones actually involved. Beger, in the final chapter of this volume, reports on a case where the results obtained from a corpus analysis and experimentally elicited data diverge. In such a situation, the experiment can be considered as diagnostic and as a valuable corrective. It potentially allows for the discovery of more plausible explanations for the data observed, and thus serves its purpose as a hypothesis tester, too. Secondly, divergence may result from testing language use at two different levels of generality. One method, data analysis, usually looks at the usage in a corpus. This, if it is meant to reflect a speech community's usage at a particular time (or over time), contains data that represent the use of language by many speakers, thus revealing intersubjective linguistic behavior (at aggregate level). Experimental data do not necessarily generalize over speakers so that what is tested is individual usage. That is why experimental results may, of course, be different from what is found in more 'general' usage, where individual differences may just average out.

Thirdly, the fact that experimenting allows for the control of factors that go uncontrolled in observed usage data can also account for results that do not converge on exactly the same conclusion. In such cases, divergence does not necessarily

1. I am grateful to one of the reviewers for pointing out this requirement and for addressing the aspects that are being considered here.

falsify one's hypothesis either, but it may indicate the need for the hypothesis to be refined. A situation like this represents an almost classical case of an empirical research agenda, illustrating the idea of the empirical cycle (elaborated in Section 4). Conversely, experimental findings may also be tested against corpus data, and as the study by Gries (this volume) shows, corpus analyses – in combination with sophisticated statistical techniques – can add to what has been found on the basis of experiments, too. Hence, in such scenarios, what at first may seem to be cases of divergence turn out to be cases of complementation and mutual cross-fertilization.

Against the background of the above clarifications, the following sections discuss various types of evidence linguistic research may draw on, with reference to the linguistic theories and models that favor them.

3. Introspective evidence

Sometimes, introspection is considered equivalent to intuition, though both concepts differ from each other in that they focus on subtly different aspects of using one's knowledge: intuition is concerned with gaining rational insight, grasping or seeing what something is or that something is true, and, hence, is a matter of processing information. Introspection is a kind of self-observation and rather a matter of making oneself aware of one's activities when grasping something. It can be paraphrased as observing, examining and describing one's thoughts (for a brief argument about the distinction see e.g. Schütze (1996: 48–52)). In this sense of self-observation, introspection meets one criterion of empirical methods, where analyses are also based on observation, though admittedly on other-oriented observation. Itkonen (2003) sets introspection and intuition apart more clearly.² Understanding linguistics as a human science that is concerned with normative facts (cf. also López-Serena 2009), he attributes language to Popper's world 3 (cf. below) and postulates that investigations into language be based on intuition. He describes the essence of intuition as drawing on speakers' abilities to remember what they know of their language, what the norms of their language are (cf. Itkonen 2003: 41), and sets it apart from introspection and observation by linking each of them to one of the three Popperian ontological levels:

[...] i.e. ontology divided into the 'worlds' of physical states/events, of psychological states/events, and of social concepts and/or norms. The labels for these worlds are 'w-1', 'w-2', and 'w-3'. Each of these worlds requires its own characteristic act of knowledge: w-1 = observation, w-2 = introspection, w-3 = intuition. [...] All three types of acts are potentially conscious; and, as subjective psychological acts,

2. I thank one of the reviewers for bringing this publication to my attention.

they all emanate from w-2. (...) [...] There is an obvious conceptual interdependence between acts of knowledge and their objects. Observation is that which pertains to material or physical entities, and vice versa. [...] Similarly, intuition qua subjective act is that which pertains to social or intersubjective concepts/norms, and vice versa. (Itkonen 2003: 44–45)

As Itkonen emphasizes the social nature of language, it follows that intuition, "pertain[ing] to concepts or rules existing in an intersubjective normative reality" (Itkonen 1981: 128), suffices for the description of a speaker's knowledge of a language, whereas introspection, being concerned with subjective sensations mainly caused by physical events, does not seem to play a part in linguistic analysis. This is at least a possible conclusion from Itkonen's understanding of language and the absence of this cognitive act in Itkonen's further elaboration. Observation is the method of data analysis, since data as results of material events associate with the intersubjective physical world (cf. Itkonen 1981: 128). Considering the relation between intuition (of norms) and observation (of corpus occurrences), Itkonen (2005: 359–360) points out "observation is connected to the quantitative analysis of actual utterances whereas intuition is connected to the qualitative analysis of conceptual possibilities". He gives prominence to intuition, since data analysis does only work on the basis of a speaker's/analyst's intuitive knowledge of the rules of the language (cf. Itkonen 2005: 361). Moreover, regarding our knowledge of norms, observation (or quantitative analysis) does not tell us what the norms are, exactly because norms, the entities of normative reality, are not observable (cf. Itkonen 1981: 128).

Geeraerts (2005: 171–182) discusses Itkonen's concept of intuition critically, and notes that, in practice, it is almost identical to that of introspection (in the understanding of Generative Linguistics) in that the latter also draws on a speaker's knowledge of linguistic conventions.

In this chapter, I do not distinguish between intuition and introspection; both terms are used interchangeably for methods of analysis that draw on the (native) speaker's knowledge of her or his language.

3.1 Introspective evidence – the (sole) source of insight in Generative Linguistics

In Generative Linguistics, the linguistic paradigm dominant in the second half of the twentieth century, studying language and modeling it is always equivalent to studying and modeling linguistic competence, i.e. native speakers' tacit knowledge of the grammar of their language. The generally accepted method to make this knowledge accessible and explicit is to consult the native speakers' intuitions about expressions or strings of words of their language. These intuitions are about

the well-formedness of sentences and about sentence structure, and it is assumed that they can be unearthed by means of grammaticality judgments, where native speakers are asked to say whether a given sentence is well-formed or not and whether it has a particular structure or not (cf. also Radford 1988: 9). The judgments are used as evidence in the process of testing the hypotheses formulated, the theory or model constructed by the linguist. Once researchers have gained some affirmative evidence for the phenomena they postulate, they may proceed deductively to develop hypotheses of what it is to know a language into a more complex model or theory of language. For testing such (deduced) hypotheses, e.g. about a particular structure or structural principle of the language under analysis, expressions that either exhibit or deviate from the hypothesized principle can be constructed and again be tested for their grammaticality. As soon as these expressions are judged to be well-formed, this is taken as sufficient evidence for the correctness of the extrapolated structure or principle, and the judgment represents evidence in support of the theory from which the predictions have been generated. If they are judged to be ill-formed, the hypothesis must be revised or even discarded. From such a perspective, the ideal person for linguistic theory construction is a linguist who is also a native speaker of the language s/he studies. This is because the study of language is theory-driven and intuition-based, and the linguist would then profitably and conveniently represent the expert with respect to both (theoretical) linguistic knowledge and speaker intuition. When constructing a model/theory of some linguistic phenomenon and elaborating the predictions it generates, the linguist can simultaneously draw on intuitions of well-formedness.

Hence, hypothesis formation and testing in this framework draws on introspective evidence alone, it is the sole kind of evidence used in the process of constructing linguistic theory and is – for the sake of its acknowledgement and general acceptance – not required to converge with any other kind of evidence. By implying that competence can be known on the basis of intuition alone, or by deduction from intuited notions, the generative framework places itself in the philosophical tradition of rationalism, where, as Lehmann (2004: 188) puts it,

[...] linguistic data play no role whatsoever, since the linguist only externalizes what is already in his mind. Representations of speech events are neither needed as a basis of inductive generalizations nor as touchstones of empirical verification or falsification of hypotheses deduced from the theory, since it is not an empirical theory.

A second explanation, suggested by one of the reviewers, for the generative linguists' lack of interest in performance data could be their understanding of grammar as a model generating all *possible sentences* of a language. These can indeed be inferred from grammaticality judgments by extrapolating the linguistic knowledge

underlying them. Such a procedure is, however, in stark contrast with empirical approaches to linguistics, which are concerned with *actually occurring sentences*, from which grammatical knowledge is assumed to arise.

Hence, it does not come as a surprise that criticism of the restriction to, and the confidence in, introspection alone has been leveled from other, more empirically oriented, theoretical positions, such as functional and cognitive linguistics. Concerning the method of extracting the intuitions of native speakers the problem arises that grammaticality judgments are notorious for not being uniform.³ Different speakers may disagree on the grammaticality of a particular structure just as well as one and the same speaker may feel different about the grammaticality of one structure at different times or in different situations, which indeed questions their diagnostic value. It is far from clear whether such judgments can be taken as direct measures of linguistic competence, because they seem to be influenced by a considerable number of other factors, such as the subjects' metalinguistic skills, which are not controlled for in the usual practice of eliciting the judgments. Hence, they need to be handled with care as regards their diagnostic significance (for a convincing demonstration thereof see Schütze 1996 (Chapters 3–5 in particular) and Keller 1998, for example).

The rejection of performance data for linguistic analyses within this framework follows from the assumption that they reflect competence imperfectly. They may be defective or ungrammatical, due to attention or memory lapses or other psychological factors such as tiredness or boredom, and must therefore be understood as contaminated by non-linguistic factors. Such 'flawed' performance phenomena as slips of the tongue, false starts or restarts, anacolutha, incomplete utterances, repetitions, and the like are not revealing about a speaker's linguistic knowledge: "Linguistic theory is concerned primarily with an ideal speaker-listener, ... who knows its [sic] language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations ... in applying his knowledge of the language in actual performance" (Chomsky 1965: 3).

Moreover, from the assumption that the set of sentences of a language is infinite, Chomsky (1957: 15) concluded that it is "obvious that the set of grammatical sentences cannot be identified with any particular corpus of utterances obtained by the linguist in his field work". A third argument against the suitability of corpora for linguistic description is that

3. This problem will not be discussed in detail. Suffice it to say that such concepts and factors as acceptability, context-sensitivity, and the understanding of grammaticality as a dichotomy (with the values + or -) or a continuum add to the complexity of grammaticality judgments and the feasibility of their 'translation' into a speaker's tacit knowledge of his language (for a discussion of the concept of grammaticality see e.g. Schütze 1996: 19–53, Fetzter 2004).

[t]he problem for the linguist [...] is to determine from the data of performance the underlying system of rules that must have been mastered by the speaker-listener and that he puts to use in actual performance. [...] [T]his [information about the speaker-hearer's competence, D.S.] is neither presented for direct observation nor extractable from data by inductive procedures of any known sort. (Chomsky 1965: 4, 18; see also Itkonen's argument rendered in Section 2)

It is a logical consequence of the assumed impossibility of extracting the underlying rule system from real performance data that introspective judgments are required in generative theorizing (for a similar argument, cf. Riemer 2009: 661).

3.2 Generative Linguistics in more recent times – broadening the methodological spectrum

Even in Generative Linguistics, the discussion about research procedures and the value of particular types of evidence to be produced in support of one's ideas and claims, empirical data in particular, is enjoying a revival. In 2009, Volume 28.1 of the *Zeitschrift für Sprachwissenschaft* provided a platform for a debate on the issue of evidence from within the generative framework and from an external perspective. The general debate is about the relationship between empirical work and theory formation in linguistics, more particularly, the focus is on the kind of data employed in linguistic research, their nature, elicitation and interpretation.

First of all (and not surprisingly from a generative perspective), in most of the (framework-internal) papers, the relation between theory and empirical work is elaborated in one direction: it is a matter of the empirical foundation of (generative) grammar theory (cf. Fanselow 2009: 133). As a consequence of the priority given to deductive theory formation, the discussion about empiricism is qualified as not being of central concern, and if data are to be employed, this is only fruitful if the analysis is theory-driven, so that appropriate questions can be asked (cf. Fanselow 2009: 134).

A second point made is that the empirical methods currently available are insufficient and inadequate for the main goal of linguistic research (i.e. the formation of a theory modeling competence), and the argument is basically in favor of employing acceptability judgments. The judgments themselves would generate sufficient data for theory testing, the more so, when the data show variation which would need to be accounted for by the theory under construction. This procedure would represent a good method of empirical foundation, and the 'new' empirical methods could be employed in the clarification of the nature of differences in acceptability judgments in problematic cases (cf. Fanselow 2009: 136–137). Hence, it is not surprising that acceptability judgments and the concept of well-formedness are placed center stage. Meyer (2009: 141–150) elaborates on the latter and

emphasizes its multi-dimensional nature, including such aspects as a linguistic unit's occurrence (in language use, incorporating the dimension of language variation), the frequency of occurrence, and its correspondence to a linguistic rule or regularity. He argues that this multi-dimensional approach is required if empirical facts, such as different data types, scalar values, sample effects, and the like are supposed to be accounted for, simultaneously pointing to the empirical challenge of appropriate operationalizations.

Meyer's article shows that the methodological discussion within the generative framework is no longer restricted to introspection, as, for example, information on occurrence and frequency of a linguistic unit can only be drawn from usage data. As far as their role in linguistic research is concerned, Schütze (2009) argues that data collected from corpora (the Web in this case) should supplement judgments. Just briefly commenting on 'idea-sparking' uses, he is more concerned with an 'evidentiary scenario', corresponding to the theory-driven procedure when corpus data are used as evidence for particular theoretical claims. He qualifies this procedure by saying that "merely having found instances of a construction of interest should not be construed as evidence of anything ipso facto" (Schütze 2009: 152), a qualification with which we could not agree more (cf. also Section 4.2). In addition, Schütze illustrates that the sole reliance on corpus examples has its drawbacks, concluding that search data cannot stand as evidence on their own:

My view is that evidentiary uses of Web data require checking those data against the intuitions of live human speakers of the language. In case such verification fails, considerations should be given to how the exemplars might have arisen *other than* as genuine instances of use of the language in question, and how their frequency compares to relevant base rates. (Schütze 2009: 155)

This emphasizes and specifies a point also made by Sinclair (1991: 39) and McEnery and Wilson (1996, cf. also Section 4.2), that data need to be interpreted and evaluated, and that the rightful place of introspection is just there.

4. Empirical evidence

Before going into detail regarding the employment of empirical data in specific linguistic theories, we will look at what is common practice in empirical linguistic research in general.⁴ With the proclamation of doing empirical research, linguists commit themselves to the 'empirical cycle', a methodological manual or programme first propagated for doing empirical research in the field of psychology by

4. For a complex argument on empirical linguistics see Sampson (2001).

Adriaan D. de Groot (1969). Following the steps suggested there, the linguist inductively produces generalizations and formulates hypotheses regarding particular phenomena on the basis of what he has observed and collected. These hypotheses are meant to explain the phenomena at issue and are simultaneously the ground from which to deduce consequences in the form of further (more general) testable predictions. Hypothesis formulation always needs to be accompanied with hypothesis testing – either by further observations, or by simulation or testing for specific factors in experiments and evaluating the results. The practice to be applied here is to look for data and/or design experiments for refuting the null hypotheses H_0 (rather than for verifying one's own, or alternative, hypotheses (H_1)). Such a procedure presupposes that the hypotheses formulated must be testable.

In short, the research agenda can be summarized to (ideally) comprise the following phases:

- observation
- descriptive research of what has been observed
- exploratory research (using inductive and deductive reasoning) resulting in the formation of (testable!) hypotheses
- analytical research of empirically testing these hypotheses and
- evaluation.

On the basis of such a strategy, the model envisaged by the researcher for the description and explanation of a particular linguistic phenomenon can be put on firm ground. When further elaborating and extending the model to cover more data, the researcher will traverse the cycle again, collecting more data, forming more general hypotheses, running experiments to test them, etc.

Empirical research in linguistics is most obvious in usage-based approaches, where, by definition, usage data are the prime source of linguistic discovery: they represent the observed facts from which hypotheses are inductively constructed and generalizations made which allow for predictions about further data. At the same time, data are an important source of evidence in that they are employed as the test case for the hypotheses made: New data must be in accordance with one's claims, and as soon as some data disconfirm them, the claims must be reconsidered and adapted so that they cover also the new data, or alternative explanations must be looked for, covering both the original and the new 'facts'. In some studies collected in this volume the usage data employed converge on the hypotheses to be tested (cf. Backus and Mos, Egan, Höche, Reshöft). In others, the data analyses result in a modification or refinement of the hypotheses made (cf. Beger, Borgwaldt and Benczes, Brandt and Kidd; Kaltenböck; and Steinkrauss). This situation is indicative of the potential of data analyses in general and for hypothesis testing in particular.

A second source of evidence associated with empirical research is elicited data, that is data available as a result of employing questionnaires and running experiments, where the elicitation may also be triggered by non-linguistic stimuli. The former data can either be categorized with introspective data (though in a more objective form, since they represent the intuitions of many speakers), or with empirical data. This depends on what the questions are intended to elicit: grammaticality judgments or the subjects' 'linguistic behavior' (e.g. as in translation or gap-filling tasks).⁵

Experimentation, which allows for the controlled elicitation of data, has its place in the phase of empirical hypothesis testing, where predictions following from intuition or the generalizations over usage data are put to the test. The study by Borgwaldt and Benczes is a case in point. Against the backdrop of the standard practice to formulate testable hypotheses, these may (but need not) also provide a pointer to the design of experimental studies. If an experiment can produce results refuting H_0 , this adds to the validity of the alternative hypothesis. If the results do not corroborate the model envisaged, an explanation must be looked for, for example by checking all the (other) factors potentially involved in the experimental activity and controlling for them accordingly. If nothing can be identified, this should feedback to theory construction and give rise to re-thinking one's model.

In other words, traversing the empirical cycle is actually concerned with collecting evidence that converges on the same assumption(s). That is, a procedure in which a hypothesis was formulated on the basis of usage data and is tested in an experiment represents a showcase of converging evidence in that results of two methodologically different analyses point to the same explanation of the observed phenomenon. The studies by Backus and Mos, Beger, and Brandt and Kidd pretty much follow such a strategy and reveal its usefulness and potential. Working through the phases of the empirical cycle, it is obvious that researchers will – deliberately or unconsciously – also consult their own knowledge of the language, their intuitions. The generalizations made and the interpretations given to the data do not come about independently of a researcher's theoretical orientation and presuppositions; hypothesis construction does not work without logical deduction from the facts AND predetermined conceptions. Even 'corpus-driven grammar', a radically empiricist approach to language description mainly developed by Hunston and Francis (2000, for example), where corpus data are analyzed for distributional patterns with a minimum of theoretical presuppositions, is not 'theory-free'. For the identification of these patterns rests on the recognition of such notions as word categories, and, although these seem to be pre-theoretical at first sight, it must be noted that in the categorization process different criteria are susceptible to

5. This difference was brought to my attention by one of the reviewers.

different weightings – depending on the linguist's conviction: does one take notional, morphological and/or a distributional aspects as decisive for a word's category? That is the linguist's intuitions and theoretical predilections cannot be blocked completely. The important point to make from the perspective of empiricism is that the generalizations and hypotheses made by the linguist should indeed be data-based rather than theory-driven (which is in contrast to the practice in Generative Linguistics (cf. Section 3.2)) and that they must have been ascertained by a rigorous search for all available evidence and a judicious assessment of its probative value (for a more detailed discussion about different types of evidence to be used in linguistic theorizing see Schönefeld 2001: 110–113).

4.1 Empirical evidence – evidence naturally employed within Cognitive Linguistics

As has been shown, Generative Linguistics places itself in the logical tradition of linguistic research, mainly relying on deductive reasoning and introspection. Since the late 1980s, a competing linguistic paradigm has been emerging, cognitive linguistics, which, exactly as elaborated by Kuhn (1970) for the development of science on the basis of his studies of theories in physics and astronomy, clashes with much of the 'received wisdom' of Generative Linguistics.

Cognitive linguistics sees itself as a usage-based approach to the analysis of language. Hence, it is only natural that the research questions asked and tackled within its framework should also be approached from the perspective of language use or performance (from which also the grammatical knowledge (the competence) of a speaker is assumed to arise). Unlike the generative practice of eliciting intuitions by means of grammaticality judgments in order to test or validate their theoretical claims, usage-based research is empirical. This implies that theory construction starts out from the description and analysis of observable facts in the first place. The observable facts for the construction of a linguistic model are what we hear people say and see them write, i.e., it is performance data from which any generalization should be made. It is, however, crucial to note here again that also observational data are not 'pure' in the sense of theory-free. For, as soon as observations are processed and described, the observer needs to categorize and to make judgments as to what it is that has been observed, and this is always shaped by preceding experience, such as the observer's linguistic knowledge, the ideas he has about language (that is, the model he adheres to), his research tradition, other observations, and the like. We need to be aware of these influences that might skew observation (cf. Section 6), and one way to counteract this danger of 'seeing what one expects to see' or 'overlooking what one does not expect' is to look for converging evidence. Moreover, saying that theory formation in cognitive

linguistics is data-based does not exclude introspection/intuition at all, but it will certainly have its place in the interpretation of the data and in the determination of their theoretical significance. Intuition is also important to cognitive linguistics for another reason, which follows from its central concern with meaning (understood as conceptualization). It is widely acknowledged that the analysis of a linguistic unit's conceptual content is most naturally and successfully accomplished by employing one's intuition. In Talmy's (2007: xiii) view, introspection is the only methodology allowing for direct access to meaning: "Meaning is a consciousness phenomenon and, if it is to be taken on as a target of research, introspection – itself a process occurring in consciousness – is the relevant instrumentality able to reach its venue". Harder (2007: 1248) argues along similar lines "since mental content is only accessible to human subjects, there is no way of getting at the central data without them [intuitions D.S.]". On the other hand, if meaning is understood as an effect of contextualized use, as function in context, it is mandatory to consult the corpus. Such a view of meaning was held by Firth (1957: 11, 14), for example, who claims that a word's meaning is known by the company it keeps, and it eventually reflects in collocational analyses (e.g. Sinclair 1991), one of the first forms of corpus-linguistic work.⁶ In cognitive-linguistic research, the interaction between a word's meaning and context is explicitly drawn on when the meanings of constructions are investigated by applying collostructional analyses (cf. Gries and Stefanowitsch (2004), Stefanowitsch and Gries (2003), (2005)). The focus on use is also present in modern cognitive-linguistic theorizing about the concept of meaning and is pushed quite far by Croft and Cruse (2004: 97), who propagate an approach "where neither meanings nor structural relations are specified in the lexicon, but are construed 'on-line', in actual situations of use," with the actual expressions figuring as just one of several components contributing partial clues to the construal (cf. Croft and Cruse 2004: 98).

Hence, unlike the introspective evidence considered sufficient in Generative Linguistics, drawing on (converging) evidence of different types suggests itself as essential for doing research within a data- or usage-based framework, which Cognitive Linguistics evidently adheres to.

In addition to the implicit methodological anchoring provided by the label 'usage-based', cognitive linguists have explicitly discussed methodological issues from the start. Lakoff (1990: 46) summarizes and generalizes for the methodology

6. Such an understanding of 'meaning' can be traced back to Malinowski, who – according to Steiner (1983: 57–58) – defines 'meaning' by referring to both 'function' and 'context': 'Meaning is function in context, it is use of language and not concept. The same concept of meaning is taken up and further elaborated by Halliday.

to be applied in cognitive linguistics that it needs to reflect the *generalization commitment* and the *cognitive commitment*:

From my perspective, the generalization commitment is a commitment to linguistics as a scientific endeavor; a commitment to seek general principles. The cognitive commitment is a commitment not to isolate linguistics from the study of the mind, but to take seriously the widest range of other data about the mind. Neither of these commitments, in themselves, imposes a particular form on the answer. As such, they are methodological, not substantive commitments.

It is the latter commitment that hints at the requirement for evidence to converge on one conclusion, though the evidence Lakoff had in mind is of a different type from what has been discussed thus far. He was more concerned with the compatibility of claims made for linguistic phenomena with findings, or evidence, produced in neighbouring fields, such as psychology, neurology and biology; that is evidence as a product of 'phenomenological pluralism'.

Assessing the cognitive-linguistic enterprise, Langacker (1999: 23–37) makes a statement on 'some methodological principles', too. He argues for such basic principles (shared with functional linguists) as '... concern with the integrity of one's data, the need for precise formulations leading to testable hypotheses, the desirability of accounting for a wide range of phenomena with a limited set of constructs' (Langacker 1999: 24). The latter, corresponding to Lakoff's generalization commitment, he sees embodied in Cognitive Grammar's assumption that any linguistic unit is symbolic, i.e. a form-meaning pairing. Following from this, grammar and lexicon are 'the same' in that they represent symbols. They are nevertheless distinguishable in that they exhibit different degrees of symbolic specificity, complexity and entrenchment (cf. Langacker 1991: 44).

Regarding the formulation and testing of hypotheses, Langacker specifies as the most fundamental methodological principle "to look for converging evidence from multiple sources" (Langacker 1999: 26). He illustrates this commitment for particular linguistic constructs (such as profiling and search domain, for example) by drawing on evidence from three different sources. The constructs

are shown to be necessary for the adequate semantic description of multiple phenomena in various languages. ... [to be] commensurate with (if not identical to) independently observable cognitive abilities. ... [to be] critical for the explicit characterization of varied grammatical phenomena. (Langacker 1999: 27)

This, again, is an appeal to look for evidence for allegedly the same constructs from different domains of investigation. Or, to put it differently, the researcher is invited to consult evidence from different domains and test whether it can be assumed to converge on the same conclusion. Again, this is the type of evidence described and discussed as a product of 'phenomenological pluralism'.

From a more external perspective, the one of a psychologist, Gibbs (2007) critically evaluates cognitive-linguistic research with a particular focus on the way the researchers arrive at their theories and hypotheses. In the (cognitive) psychologists' view, cognitive linguistics is not yet really empirical, because objective scientific experiments are not common practice. They criticize cognitive linguistic work for a lack of objective evidence in the form of objective, replicable data gained from experimental tests:

What is needed [...] is for cognitive linguists to be more sensitive to some of the important properties of framing experimental hypotheses (e.g. constructing falsifiable hypotheses, considering alternative hypotheses), and trying to articulate their ideas, and empirical findings in ways that may be tested by scholars in other disciplines. (cf. Gibbs 2007: 17)

Even if we take this to be a rather pessimistic description of the situation, it seems to hint at a gap between aspirations and reality in cognitive-linguistic research, and Gibbs's criticism needs to be taken seriously. So it does not come as a surprise that criticism is leveled at a quantitatively (and at times also qualitatively) unsatisfactory use of empirical methods also from within the cognitive-linguistic paradigm. Geeraerts (2006) elaborates on the tensions between the inclination towards empirical methods (almost 'inherent' to a usage-based form of linguistic investigation) and the reality of the still dominant traditional introspective methodology of conceptual analysis. He argues against this suboptimal situation (cf. Geeraerts 2006: 39–44) and encourages the application of empirical methods, also in the field of semantic analysis. Successful attempts at analyzing such outspoken semantic-conceptual phenomena as metaphor on the basis of corpus data (cf. Stefanowitsch and Gries 2006) make it obvious that this is a promising way to go. However, for the current situation, it must be admitted that neither empirical methods nor the consultation of different types of evidence can be rated as the default situation in cognitive-linguistic research.

All in all, the work done by linguists who feel affiliated to the cognitive framework, shows no restriction in their methodology to the sole employment of just one type of evidence, namely that of introspective reasoning. As proponents of a usage-based approach to language, cognitive linguists draw on the analysis of performance data, and should primarily do so: they search data for regularities and more general patterns, and draw conclusions as to their motivation(s) and what they mean with respect to our linguistic and cognitive functioning. They also take data as the test case for their theoretical claims. It has also been shown that there is considerable demand for improvement both in the phases of the formation of (testable) hypotheses and (empirical) testing. This needs to be added to the list of

desiderata in cognitive-linguistic research. Accordingly, testable hypotheses and their testing are in the focus of most of the contributions to this volume.

4.2 Empirical evidence – an essential characteristic of corpus linguistics

In parallel with, but independently of, the developments in the cognitive-linguistic paradigm, the discussion of methodological issues was also provoked by corpus-linguistic work, which defines itself on the basis of the methodology employed: the observation of naturally occurring data. As the label makes obvious, corpus linguistics is usage-based by definition. It takes its discoveries and findings from the observation of actual usage and introspection – the collected data need to be interpreted. The use of corpora as the principal source for describing language marks corpus linguistics as a field adhering to an empiricist view of scientific inquiry (cf. also Leech 1992: 107–111), placing it in the philosophical framework of empiricism. Knowledge is drawn from observable data, the data are replicable and allow for verification/falsification on the basis of further data. In the wake of this empirical orientation, Stubbs (1993: 8) requires that “[l]anguage should be studied in actual, attested, authentic instances of use, not as intuitive, invented, isolated sentences”. In accordance with Schütze (cf. Section 3.2), it needs to be emphasized that it does not suffice to rely for testing one’s hypotheses on a small number of authentic instances only, but that a representative inventory of data needs to be employed.⁷ McEnery and Wilson (1996: 1–19) likewise characterize corpus linguistics as resting on (computer-aided) empiricism, which makes it a systematic and rigorous approach, but they also make a strong point of its affinity to cognitive rationalism. They suggest that the best possible research strategy is to view corpus- and introspection-based approaches to linguistics as complementary:

7. I am aware of the problem of the concept of ‘representativeness’ of corpora and corpus data. The former is topicalized in almost any publication about corpus linguistics (e.g. Sinclair 1987; Biber 1992 (2005); McEnery and Wilson 1996; Biber, Conrad, and Reppen 1998), and will not be discussed here. Suffice it to say that representativeness as such is so multi-faceted a phenomenon that, though it can be checked or achieved for particular parameters, it is extraordinarily difficult to achieve in full. In this respect, the concept compares to that of ‘equivalence’ in translation studies. The consequence arising from this was spelled out by one of the reviewers, commenting that the desideratum of a representative corpus leaves us with two options, the first of which is more preferable: (i) either representativeness is regarded as a challenge to be faced, or (ii) linguists should restrict themselves to making claims about corpora rather than about mental representations of linguistic data. The question of ‘representativeness’ (or rather significance) of corpus data is taken up in Section 4.3.

[...] artificial data can have a place in modern corpus linguistics. Yet it should always be used with naturally occurring data which can act as a control, a yardstick if you will. Corpus linguistics is, and should be, a synthesis of introspective and observational techniques, relying on a mix of artificial and natural observation. (McEnery and Wilson 1996: 16)

In practice this means that text corpora serve as the empirical foundation on which the analysis of linguistic units, their properties and relations rests. The automatic, systematic and quantitative analysis of corpora is seen as the prerequisite for hypothesis formation and hypothesis testing (though not to the exclusion of introspection and logical extrapolation), since they represent both the linguistic material over which to generalize and an objective test case for the researcher’s hypotheses. In the first case, hypothesis formation, corpus analyses aim at the identification and exploration of linguistic entities and regularities. In this phase, the analyst’s theoretical background assumptions will, as has already been hinted at in Section 4, play a role and it is important that they be made explicit. In particular, when the corpora used are tagged, parsed and/or semantically annotated, the information that can be gathered from them will be affected by theoretical preconceptions and decisions linked with the tagging, parsing and annotation procedures respectively. Theory and introspection will have to enter the scene in exploratory research, when assumptions are made as to which theoretical constructs the identified linguistic facts are indicative of.

In the second case, hypotheses – no matter if formed on the basis of corpus-linguistic analyses or deduced by extrapolation – are tested against corpus data and, if confirmed, can be made the basis for making further testable hypotheses. Two things require special care in this phase: It is important to formulate potentially falsifiable hypotheses (cf. the argument in 4.1), and it is decisive for the proper functioning of the test that the variables of one’s hypothesis are operationalized in a way that what is measured or counted actually represents a factor involved in the theoretical construct being tested. For the decision of how the variables in the researcher’s hypothesis can or need to be investigated in corpus-based research, Gries (2009: 177–178) suggests, “you answer two interrelated questions. Firstly, ‘what will I perceive when I perform our study and observe the values/levels of the variable involved?’ Secondly, ‘which mathematical concept will I use – counts/frequencies, averages, dispersions, or correlations?’” The second question makes it obvious that in corpus linguistics the corpus is not just used for the identification of interesting linguistic phenomena, but is systematically and exhaustively searched so that frequencies of occurrence can (easily) be calculated and be analyzed for what they could mean for the construction of a linguistic model. A careful quantitative evaluation of the data elicited from a representative corpus opens the analysis to the incorporation of probabilistic aspects, if data variation can be shown to

significance testing presupposes quantification. They are required in corpus-based studies to evaluate the position and relevance of the discovered phenomena in the language under analysis and a model thereof. Raw frequencies and percentages, however, are difficult to evaluate and interpret. For example, for an interpretation of the frequency of two co-occurring words, we would need to know what the expectation for their co-occurrence is on the basis of the total occurrences of both words at issue. In technical terms, we need to have a baseline frequency against which the observed frequency can be evaluated and tested for its significance. In Tummers's et al. (2005: 235) words, this is a general requirement in corpus-based research: "In order to identify the relevant, i.e. significant, features of language use, descriptive and inferential statistical techniques have to be used". Significance tests tell us whether it is justified to extend one's findings from the sample analysis to the total population, i.e. to the language under analysis. In this respect, they are indispensable in order to be able to tell apart pseudo-findings from real findings, with the former turning out to be accidental coincidences in a corpus and the latter representing the meaningful relations or phenomena which are important for the theoretical modeling of the language under analysis.

As Tummers et al. (2005: 242) point out, frequency counts have another shortcoming, namely that of not allowing for the control of other factors potentially involved in the distribution of the phenomenon analyzed. Though it is conceivable to consider all the potentially relevant factors in separate analyses, this would not help in the identification of possible interactions between them. That is why multifactorial analyses have been introduced to usage-based analyses, allowing for the simultaneous analysis of the impact on the dependent phenomenon of several factors as well as of the interaction between them (as shown by Gries 2003, this volume, and publications of the research team *Quantitative Lexicology and Variation Linguistics* at Leuven University, for example).⁸ Heylen et al. (2008) elaborate at length and in great detail and discuss critically the use of several types of multivariate statistics in cognitive-linguistic research. For details, the interested reader is referred to this text.

To conclude, the current situation in usage-based cognitive-linguistic analysis can be described as gradually extending its methodology from frequency-oriented systematic and exhaustive corpus analyses (for the recognition of probabilities and trends in language use) to the more sophisticated statistical testing of explanations for the findings made.⁹ The more advanced studies also take into account the

8. The people most prominently involved in the group's research activities are Dirk Geeraerts, Dirk Speelman, and Stefan Grondelaers.

9. 'Exhaustive' is meant in the sense of searching the complete corpus for the phenomenon at issue and drawing on the full list of extracted data.

relate to particular contextual and co-textual factors, for example. These are aspects notoriously difficult to capture on the basis of introspection alone, as was pointed out by Sinclair (1991: 39) from early on:

The problem about all kinds of introspection is that it does not give evidence about usage. The informant will not be able to distinguish among various kinds of language patterning – psychological associations, semantic groupings, and so on. Actual usage plays a very minor role in one's consciousness of language and one would be recording largely ideas about language rather than facts of it.

He concludes from this that the rightful place of introspection is in the evaluation of evidence rather than in its creation.

In much the same way corpus analyses are also carried out when results of earlier studies need to be confirmed or replicated, or when the results of studies having drawn on data other than those collected from corpora are to be validated.

Since almost all the case studies collected in this volume incorporate corpus analyses (in the sense of systematic and exhaustive searches allowing for quantitative evaluation), they exactly fit this methodological pattern: the authors' own (or other) hypotheses are tested against (new) corpus data. The (more general and deeper) insights following from such a strategy may trigger suggestions and questions for further experimental investigation, as is shown by Backus and Mos; Berger; and Brandt and Kidd (cf. also Section 4).

4.3 Empirical evidence – towards a truly usage-based approach in cognitive linguistics

Since the methodological aspirations of cognitive linguistics obviously coincide with what has been received practice in (computer) corpus linguistics for quite some time, it is not surprising that, quite recently, there have been successful attempts at an explicit unification of cognitive-linguistic theorizing and corpus-linguistic methodology. Studies exemplifying this procedure were encouraged, promoted and collected by Gries and Stefanowitsch in two volumes (2006), and they make a convincing case for the fruitfulness of corpus-linguistic techniques for the analysis and explanation of phenomena that are of central interest within the cognitive-linguistic framework, namely the lexicon-grammar continuum and metaphor and metonymy. The explicit turn to corpus-linguistic methodology (as the most likely form of usage-based approaches applied there) in cognitive linguistics is also discussed by Tummers, Heylen, and Geeraerts (in different orders, 2005, 2008). It is basically three points that need to be taken up here to complete the picture of the types of evidence designed so far: quantification, significance testing, and multifactorial analyses. The first two activities are associated naturally in that

complexity of language by looking for – and also developing – multifactorial methods so that as many variables as have been found to be influential on the phenomenon under analysis can be incorporated into the analysis. The studies in this volume can be situated in this area: usage-based cognitive linguistics. The most advanced study in this respect is the one by Gries, who discusses and demonstrates more sophisticated statistical testing and the employment of multifactorial analyses in linguistic research.

5. Data in linguistic research – a summary and overview

In the previous sections, various types of evidence employed in linguistic research were introduced and discussed, introspective data and corpus data in particular. Experimentally elicited data were also mentioned, though in passing, and statistical methods for determining the significance of empirically triggered findings.

Haspelmath (2009: 158) takes an even broader perspective, listing and discussing eight different types of data systematically in relation to the questions they could help to answer. He notes that all data having been used in linguistic analyses have shown to be relevant in principle. We will add to our data survey those that have been referred to only briefly or not at all. Experimental data (elicited in psycholinguistic tests) were already mentioned as evidence used when testing hypotheses following from generalizations over usage data (see Section 4) and they are shown to be informative by the studies of Backus and Mos; Beger; and Brandt and Kidd. Haspelmath points to another function experimental data may have in linguistic research: they are required for research into the regularities of a speaker's mental grammar, because the mere encounter in corpora of a particular linguistic unit or construction does not say anything about its mental representation (cf. the argument in Section 2). Another type of data adds to exactly this kind of research question: physiological data, such as fMRI and ERP, resulting from neuro-linguistic investigations. However, the problem with physiological and experimentally elicited data alike consists again in the potential interference of other factors. Since both in experiments and in the observation and measurement of people's physiological reactions to particular stimuli many factors must be assumed to be involved in the cause and effect relation to be tested, the results produced and observed can only be interpreted with great caution. Effects (the dependent variable) attributed to the factor assumed to be tested (the independent variable) may just as well be caused by one or more factors not even known to be involved in the experimental task. It is exactly here that the need for converging evidence becomes straightforwardly obvious: if the results of several

experiments and/or measurements allow for the same conclusions, this adds to their explanatory power.

Schlesewsky (2009: 177) argues in a similar vein, focusing on the discrepancy between multidimensional data and unidimensional interpretations in current psycholinguistic research. Interestingly, he sees in this discrepancy a potential reason for the felt gradient character of many data, which may simply follow from a negligence of factors that are also part and parcel of the issue under analysis. The solution he envisages is methodological pluralism. The development and installation of multifactorial analyses also in cognitive linguistics (cf. Section 5 and the study by Gries) is a promising way to go and improve the quality of empirical research in this field. Notable examples include Gries (2003: blurb), whose aim is to "demonstrate the superiority of corpus-based, multifactorial and probabilistic approaches towards grammatical phenomena over traditional analyses based on acceptability judgments and minimal pair tests", and the publications by the QLVL research team. In Heylen et al. (2008), the authors broadly and critically discuss the approaches by both Gries and colleagues and the QLVL group, exploring "how they can complement each other", and arguing that "these two approaches have the potential to converge into a single promising methodology for empirical usage-based research" (Gries 2003: 93).

A second point raised (and bemoaned) by Schlesewsky is an asymmetrical interaction between theory and empirical work. The latter would always be aimed at experimentally validating theoretical assumptions, without theory-incompatible data having repercussions in theory formation (cf. Schlesewsky 2009: 175–176). However, this asymmetry is naturally counterbalanced in usage-based accounts, where data are not only the test case for hypotheses but also the basis from which these are formulated in the first place. Moreover, the studies by Kaltenböck and Brandt and Kidd also serve as examples of how 'theory-incompatible' data can and should be handled in an approach that is responsive to the data and takes them seriously.

From all this and from what has been found in the more general discussion of introspective and empirical evidence in linguistic research, it turns out that neither introspection nor empirical data can be used to the exclusion of the other. Rather, linguists seem to use the two types of data at different stages of their research agendas and, hence, for different purposes, depending on the research tradition they work in. This is what makes Gast (2006: 114) conclude that "empiricism" and 'rationalism' should be regarded as scalar rather than complementary opposites, and what we should really talk about is the way responsibility can be apportioned between observation and reasoning for an accurate description and analysis of language".

6. Conclusion to the discussion

6.1 Why evidence is needed to converge

Concluding the discussion of the types of evidence linguistic theorizing should draw on and the merits and problems inherent in the individual types, I argue lastly – from a usage-based perspective – that all types are required: usage data, introspection, and experiment, and that a converging evidence approach suggests itself as the order of the day, for various reasons.

First of all, as has been made explicit, each of the types is particularly suitable for particular phases in the empirical cycle. Typically, the analysis of usage data is prominent in descriptive and exploratory research, where it may be complemented by (the analyst's and/or his subjects') introspection. Experimentation is prominent in analytical research, where also multifactorial analyses may be employed and significance testing has its place. It is, however, also not uncommon to use further usage data here, for the replication of previous findings or their corroboration in different conditions, for example. A research algorithm relying on just one sort of evidence and neglecting the others will always imply the risk of being insufficient in that certain aspects or factors might be wrongly interpreted or totally overlooked.

This already implies that, secondly, there is no restriction to one particular type of evidence to be drawn on in the respective phases of the research cycle. It is especially useful to employ different types of evidence produced by different methodologies in the phase of hypothesis testing. If in this phase, (comparable) evidence resulting from methodological pluralism can be shown to converge on one and the same conclusion, this adds to the probability of the tested hypothesis being correct or true. Diverging evidence should start an investigation into potential causes, before actually refuting the hypothesis tested (see Section 2). The usefulness and potential of such a procedure becomes apparent in the majority of the studies collected here: Backus and Mos; Beger; Brandt and Kidd; Borgwaldt and Benczes; Egan; Höche; Kaltenböck; Reshöft; and Steinkrauss. These studies cover a wide variety of linguistic phenomena, above all from the fields of semantics and language acquisition, and prove both the practicality and the effectiveness of this research strategy.

Evidence arising from phenomenological pluralism may also be employed in the verification process of a claim, though with due care only, namely if the phenomena brought into a relationship can be shown to actually relate in the way assumed in independent analyses.

In addition to this practice, which can be considered as a general recommendation for usage-based research, a number of concrete situations and examples

have been mentioned in the course of the argument in which the search for converging evidence is indicated:

- Evidence converging on the same conclusion would reduce the risk of generalizations and hypotheses being skewed by the researcher's linguistic experience and predilections.
- Searching for converging evidence is appropriate for the analysis and explanation of such a complex and multi-faceted phenomenon as language. The many factors being involved in any verbal activity need to be recognized and controlled for, and there are at least two strategies the researcher can adopt: multifactorial analyses and checking the effect of (at least some of) the factors assumed to be responsible for the verbal behavior under analysis in a number of different experiments. Compatible results would definitely buttress the researcher's hypotheses and conclusions respectively. Gries' study (this volume) puts forward a convincing case for the importance of and the gain to be achieved from deciding on the appropriate type of multi-factorial analysis that fits the research question best.
- Along similar lines, converging evidence adds to the weight of a hypothesis, i.e. to the probability of a hypothesis to be true. This would be the case if, for example, a data-based hypothesis can be corroborated by results elicited in experiments probing the same phenomenon in a more controlled way.
- Converging evidence is needed to determine the significance of a finding: Since all usage-based research works on the basis of samples only, inferential statistics is required in addition to (simple) frequency counts to separate accidental findings from (theoretically) informative ones.

Furthermore, there are also other scenarios in which the search for a convergence of evidence should be put on the research agenda:

- Converging evidence is desirable to avoid unjustified generalizations: individual observations or analyses do not allow for generalizations or the postulation of a theoretical construct, if they are not backed up by at least further observations of the same type (replication studies). They would have much more explanatory power, if they could be corroborated as a result of employing other methodologies.
- Converging evidence is desirable to guard the researcher against what is known as the confirmation bias. It is natural that observations that coincide with the researcher's theoretical assumptions are taken as confirming evidence and are given great mental weight. On the other hand, observations that contradict these assumptions are taken as exceptions and hardly leave a mental trace. In an even stronger form, such a bias can result in a selection

bias, so that the researcher dismisses evidence that disconfirms his hypothesis (or rather confirms the null hypothesis). In such a situation, the theory becomes more important than the question being investigated, so that the evidence is made to fit, instead of being used to test, the theory. In the 'worst-case scenario', the neglect of contradictory observations may cause untenable ideas to become 'received wisdom' and lead to the construction of false theories. Schlesewsky (2009: 173) reports such a case from the realms of neurolinguistics, where the initial interpretation of neuro-physiological data was maintained and drawn on in later studies, although meanwhile data had been produced which disprove it. A more widely known example to emphasize the reality of this danger is the adherence of generative grammar theory to transformations (up to the development of the Minimalist Program), although experiments designed in the 1960s to produce evidence for the cognitive costs of such operations could not corroborate their psychological reality.

Converging evidence is desirable in situations in which the data at issue may be interpreted by several hypotheses. The decision between such competing hypotheses can be promoted by further, independent evidence. Such a strategy was shown to be effective by Gries, Hampe, and Schönefeld (2005, 2010) for example, who approached the question of what is the better predictor for the association between a word and a construction, frequency or collocation strength, from both a corpus-based and an experimental perspective.

Converging evidence is desirable when testing hypotheses resorting to (conditioning) factors from the field of language processing. To be sure, corpus data can suggest assumptions regarding language-processing procedures, but they should not be made the only evidence in the verification process. For, the corpus does not give first-hand information on what was actually going on in the formulation process of the written and spoken text respectively. In order to tap the procedures going on in language production and comprehension, experimentation should follow in the research agenda to test whether the hypotheses formulated on the basis of corpus findings can be corroborated (cf. Section 2).

6.2 Converging evidence in practice

In the contributions to the volume at hand, converging evidence is lived in practice. From a (predominantly) cognitive-linguistic perspective, all of them illustrate a research strategy that draws on at least two different methodologies. Thus, they demonstrate a pluralism of methods in one project and study respectively rather than giving a general survey of the methodological potential available to (cognitive) linguists. This feature sets the present volume apart from another book-length

study explicitly concerned with methods in cognitive linguistics, González-Márquez et al. (2007: xxii), a volume "intended as a handbook to exploring the empirical dimension of the theoretical questions raised by [Cognitive L]inguistics". That book may be seen as a prerequisite to the book at hand.¹⁰ Whereas the former discusses the broad spectrum of empirical methods available to cognitively oriented linguistic research, the focus in the present volume is on the use of corpus data, not least because corpus-based approaches are currently enjoying a heyday, also in (cognitive) linguistics. Most of the contributions collected in this volume can be located within this context. They report on studies in which corpus data were employed in combination with the various other methodological tools discussed above. The corpus data are processed and evaluated either in more 'traditional' qualitative analyses based on raw frequencies or in quantitative corpus analyses. This way the studies show how corpus data can be systematically searched, analyzed, and interpreted for the elucidation of linguistic problems, also using statistical methods and techniques. In addition, the present volume also illustrates how experimentally elicited findings and those from other domains can be employed as evidence for the claims one makes with respect to a particular linguistic phenomenon. A brief summary of each contribution is presented in Section 7.

7. The contributions to this volume

As has been seen, this introductory chapter has taken a theoretical stance on the issue of evidence and its convergence. It has considered what counts as evidence in linguistic research in general and in usage-based (cognitive) research in particular and has located the chapters collected in this volume within this debate.

In the lead chapter that follows, Gerard Steen takes a reverse perspective, starting out from a central research question raised in cognitive linguistics (the nature of metaphor) and developing an argument about the value of the method of converging evidence in cognitive linguistics. Steen argues that metaphor constitutes a showcase for the method, but that it equally constitutes a showcase for the difficulties involved in handling converging evidence. In Steen's view, the notion of converging evidence typically displays a conflation between two phenomena, which he labels 'phenomenological pluralism' and 'methodological pluralism' (cf. Steen 2007). The latter is acknowledged for its potential to produce converging

10. The foreword, in which Talmy (2007) discusses different types of evidence for their pros and cons and for their suitability to the investigation into particular research phenomena or questions, is of special importance also to the concern here.

evidence, the former is characterized as exciting and attractive, but problematic, which is explicated in detail.

The body of this volume is subdivided into three parts that focus on investigations into different linguistic phenomena: meaning, language acquisition, and discourse. As specified by the book's theme, the studies are concerned with evidence from more than one source, seeking to test their hypotheses against at least two types of evidence: introspectively gained evidence and evidence gained from the analysis of empirical data in the form of corpus data or experimentally elicited data. When a third methodology is incorporated into a study, it is the experimental elicitation of data in order to test individual factors assumed to be involved in the phenomenon under analysis in a controlled way.

7.1 Part 1: Multi-methodological approaches to constructional and idiomatic meaning

Chapters 3–9, constituting Part 1 of the volume, examine the role of corpus-based and experimental evidence in the analysis of constructional and idiomatic meanings. Two are concerned with cognition verb constructions.

In the first, Thomas Egan uses corpus data to test introspective interpretations of the 'see x to be y' construction proposed in the 1970s, and finds his data to converge on these intuitive hypotheses, thus contradicting earlier findings based on work with smaller corpora.

Gunther Kaltenböck's chapter investigates the syntactic status of clause-initial *I think* followed by a so-called complement clause and in doing so illustrates how quantitative and qualitative corpus analysis can be used to gain insights into syntactic structure. The analysis of syntactic and prosodic aspects of a construction yields various types of evidence, which at first glance seem incompatible but can be accounted for in a grammar that allows for a dynamic relationship between usage and structure and incorporates a diachronic dimension.

The second sub-group in Part 1 contains three chapters that focus on constructional alternatives.

Silke Höche looks at similarities and differences of the 'be about to V' and the 'be going to V' patterns from a cognitive-linguistic, usage-based perspective, incorporating several sources of insight. Her comparative investigation of the two patterns demonstrates the successful integration of two sources of evidence: introspection in conjunction with prior attempts at spelling out formal and functional features of the constructions, and corpus investigation in conjunction with inferential statistical analysis.

Using the dative alternation as an example, Stefan Th. Gries investigates syntactic priming on the basis of corpus data from the British component of the

International Corpus of English. The methodological point of this chapter is the use of an 'unusual' source of evidence, as syntactic priming is a phenomenon almost exclusively investigated through experimenting. He finds that the results of his analyses do not only converge on those gained from experimental data, but also go beyond them, making additional factors visible that are significantly involved in priming.

Ad Backus and Mario Mos analyze two Dutch constructions expressing likelihood and possibility for their productivity and semantic overlap. The analysis of the 'is tee V' construction and the 'V-baar' construction exploits two sources: it draws on both corpus data in conjunction with inferential statistical analysis and experimental data in the form of acceptability judgments. The results of both types of analysis converge on the identification of the same similarities and differences of the constructions and speak for the existence of the constructions as entrenched patterns.

The final two chapters in Part 1 use the respective methodologies for the investigation of idioms and creative language use.

Alexander Ziem and Sven Staffeldt's chapter is concerned with the analysis of idiomatic uses of a body-part term, *finger*, on the basis of corpus data and extrapolations from a basic tenet of cognitive linguistics, that meaning is grounded in experience. On the basis of evidence from corpus data, they suggest that the hypotheses following from the 'embodiment' position need to be refined.

Susanne Borgwaldt and Réka Benczes study word formation strategies in two typologically different languages (German and Hungarian) by eliciting data from native speakers in a novel-object naming task. Apart from replicating results from studies of German they carried out earlier, their findings require the refinement of hypotheses made for Hungarian.

7.2 Part 2: Multi-methodological approaches to language acquisition

The three chapters in Part 2 focus on the collection and use of empirical evidence in the investigation into language acquisition, another field that is prominently represented in usage-based research. Again relying on at least two types of evidence, two studies enquire into the L1 acquisition of particular constructions.

Rasmus Steinkraus uses longitudinal data from a German child's acquisition of a group of wh-questions in order to test the assumption that acquisition is largely determined by the input frequency with which the child experiences the respective constructions. His results show that further factors need to be considered to account for what happens in acquisition, namely the function(s) of the constructions at issue and what has already been acquired before.

Silke Brandt and Evan Kidd provide evidence for truly usage-based strategies in children's acquisition and use of relative clauses. Comparing corpus data of such

constructions with the results from two experimental tasks, they find that children perform best with those types of relative clauses that are the most frequent exemplars in usage data, reflecting that children give careful consideration to subtle factors of the input they are exposed to.

Nina Reshöft seeks to identify typological differences in the expression of motion events by Romance learners of English. The analysis of corpus data reveals that these learners show a tendency to transfer their L1-specific lexicalization patterns for the expression of motion events to their interlanguage, thus producing evidence for the assumed distinction between verb-framed and satellite-framed languages.

7.3 Part 3: Multi-methodological approaches to the study of discourse

Part 3 is represented by one study. Anke Beger turns to an analysis of experts and laymen's use of emotion metaphors in the discourse of counseling. What she finds from the analysis of corpus data is then tested against data elicited in a guided interview. The latter data turn out to be complementary to the former in that they reveal a new factor to be causally related to the use of the respective metaphors. This speaks for the efficiency of a research strategy that uses evidence from multiple sources.

8. Convergence of evidence: Prospects for future research

The studies collected in this book give proof of the feasibility and usefulness of multi-methodological analyses of linguistic phenomena, recommending them as an endeavor worth pursuing. In the conclusion to the discussion (Section 6.1), the increased power and efficiency of using multi-source evidence was pointed out, but it also needs to be kept in mind that such a research strategy is not without its pitfalls and challenges, one of which is the broad methodological expertise required for pursuing it. For example, the researcher must be aware of the exact nature of the different types of evidence obtained, so that the comparability of the evidence can be assumed if not ensured (cf. also Arppe et al. 2010: 4–5). As far as evidence resulting from phenomenological pluralism is concerned, it is strongly suggested that the respective expert knowledge is pooled and shared. That is, future research will, in all probability, benefit from collaborative and interdisciplinary work. For usage-based research into language, potential collaborators are – besides psycholinguists – psychologists and sociologists, for example, who could contribute both their expertise in their fields and their methodological know-how. For, in social research, the (multi-methodological) sociological approach, also known as the triangulation (of methods) has been practiced for some time (cf. Olsen 2004: 103), and in psychology it is also gaining ground (cf. Flick 2004):

A second aspect to guide future research, addressed in Section 4.3, is elaborated and discussed by e.g. Gries 2003, Tummers et al. (2005), and Gries (2006) and concerns quantification in empirical/usage-based research. The former two argue for the use of quantitative techniques in usage-based linguistics. More particularly, they advocate the use of multivariate statistics for the explanation of language phenomena, because such statistics make it possible to consider the manifold factors involved in language use in their complex interaction (cf. Tummers et al. 2005: 243). The latter (cf. Gries 2006: 198, 200) emphasizes the necessity for (corpus) linguists to use more powerful statistical tools (than raw frequencies), to do rigorous statistical testing, and to employ such measures as effect size and confidence intervals in the validation of their empirical results.

From these considerations, the cornerstones for future research in empirical linguistics are evident: more multi-methodological research and more quantification.

References

- Arppe, Antti et al. 2010. Cognitive corpus linguistics: Five points of debate on current theory and methodology. *Corpora* 5: 1–27.
- Biber, Douglas. 2005. Representativeness in corpus design. In Geoffrey Sampson & Diana McCarthy, eds., *Corpus Linguistics. Readings in a Widening Discipline*, 17–197. New York: Continuum.
- Biber, Douglas, Susan Conrad, & Randi Reppen. 1998. *Corpus Linguistics. Investigating Language Structure and Use*. Cambridge: CUP.
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Croft, William & D. Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: CUP.
- De Groot, Adriaan D. 1969. *Methodology: Foundations of Inference and Research in the Behavioural Sciences*. The Hague: Mouton.
- Fanselow, Gisbert. 2009. Die (generative) Syntax in den Zeiten der Empiriediskussion. *Zeitschrift für Sprachwissenschaft* 28: 133–139.
- Fetzer, Anita. 2004. *Recontextualising Context: Grammaticality Meets Appropriateness*. Amsterdam & Philadelphia: Benjamins.
- Firth, John R. 1957. A synopsis of linguistic theory, 1930–1955. *Studies in Linguistic Analysis* Special Volume. Philological Society: 1–32.
- Flick, Uwe. 2004. *Triangulation. Eine Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gast, Volker. 2006. Introduction. *ZAA* 54.2: 113–120.
- Geeraerts, Dirk. 2005. Lectal variation and empirical data in Cognitive Linguistics. In Francesco J. Ruiz de Mendoza & Sandra Peña Cervel, eds., *Cognitive Linguistics: Internal Dynamics and Interdisciplinary Interaction*, 163–189. Berlin: Mouton de Gruyter.
- Geeraerts, Dirk. 2006. Methodology in Cognitive Linguistics. In Gitte Kristiansen, Michel Achard, René Driven, & Francisco J. Ruiz de Mendoza Ibáñez, eds., *Cognitive Linguistics: Current Applications and Future Perspectives*, 21–49. Berlin: Mouton de Gruyter.

- Gibbs, Raymond W. 2007. Why cognitive linguists should care more about empirical methods. In Monica González-Márquez, Irene Mittelberg, Seana Coulson, & Michael J. Spivey, eds., *Methods in Cognitive Linguistics*, 2–18. Amsterdam & Philadelphia: Benjamins.
- González-Márquez, Monica, Irene Mittelberg, Seana Coulson & Michael J. Spivey, eds. 2007. *Methods in Cognitive Linguistics*. Amsterdam & Philadelphia: Benjamins.
- Gries, Stefan Th. 2003. *Multifactorial Analysis in Corpus Linguistics: A Study of Particle Placement*. New York: Continuum.
- Gries, Stefan Th. 2006. Some Proposals towards a More Rigorous Corpus Linguistics, ZAA 54 (2): 191–202.
- Gries, Stefan Th. 2009. *Quantitative Corpus Linguistics with R: A Practical Introduction*. New York: Routledge.
- Gries, Stefan Th, Beate Hampe & Doris Schönefeld. 2005. Converging evidence: Bringing together experimental and corpus data on the association of verbs and constructions. *Cognitive Linguistics* 16 (4): 635–676.
- Gries, Stefan Th, Beate Hampe & Doris Schönefeld. 2010. Converging evidence II: More on the association of verbs and constructions. In John Newman & Sally Rice, eds., *Experimental and Empirical Methods in the Study of Conceptual Structure, Discourse, and Language*, 59–90. Stanford, CA: CSLI.
- Gries, Stefan Th. & Anatol Stefanowitsch. 2004. Covarying collexemes in the into-causative. In Michel Achard & Suzanne Kemmer, eds., *Language, Culture and Mind*, 225–236. Stanford: CSLI Publications.
- Gries, Stefan Th. & Anatol Stefanowitsch. 2006. *Corpora in Cognitive Linguistics: Corpus-Based Approaches to Syntax and Lexis*. Berlin: Mouton de Gruyter.
- Harder, Peter. 2007. Cognitive Linguistics and Philosophy. In Dirk Geeraerts & Hubert Cuyckens, eds., *The Oxford Handbook of Cognitive Linguistics*, 1241–1265. Oxford: OUP.
- Häselmath, Martin. 2009. Welche Fragen können wir mit herkömmlichen Daten beantworten? *Zeitschrift für Sprachwissenschaft* 28: 157–162.
- Heylen, Kris, José Tummers & Dirk Geeraerts. 2008. Methodological issues in corpus-based Cognitive Linguistics. In Gille Kristiansen & René Dirven, eds., *Cognitive Sociolinguistics: Language Variation, Cultural Models, Social Systems*, 91–128. Berlin: Mouton de Gruyter.
- Hunston, Susan & Gill Francis. 2000. *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English*. Amsterdam: Benjamins.
- Itkonen, Esa. 1981. The concept of linguistic intuition. In Florian Coulmas, ed., *A Festschrift for Native Speaker*, 127–140. Mouton: The Hague.
- Itkonen, Esa. 2003. *What is Language? A Study in the Philosophy of Linguistics*. Turku: Abo Akademi tryckeri.
- Itkonen, Esa. 2005. Concerning the synthesis between intuition-based study of norms and observation-based study of corpora. *SKY Journal of Linguistics* 18: 357–377.
- Keller, Frank. 1998. *Grammaticality Judgments and Linguistic Methodology*. Retrieved on 18 May 2009 from <http://biblioteca.universia.net/ficha.do?id=41569708&http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.52.2559>.
- Kuhn, Thomas. 1970. *The Structure of Scientific Revolutions*. Chicago: The University of Chicago Press.
- Lakoff, George. 1990. The invariance hypothesis: Is abstract reason based on image-schemas? *Cognitive Linguistics* 1: 39–74.
- Langacker, Ronald W. 1991. *Foundations of Cognitive Grammar* 2. Stanford, CA: Stanford University Press.
- Langacker, Ronald W. 1999. Assessing the cognitive linguistic enterprise. In Theo Janssen & Gisela Redeker, eds., *Cognitive Linguistics: Foundations, Scope and Methodology*, 13–59. Berlin: Mouton de Gruyter.
- Leech, Geoffrey. 1992. Corpora and theories of linguistic performance. In Jan Svartvik, ed., *Directions in Corpus Linguistics*, 105–122. Berlin: Mouton de Gruyter.
- Lehmann, Christian. 2004. Data in linguistics. *The Linguistic Review* 21 (3–4): 275–310.
- López-Serena, Araceli. 2009. Intuition, acceptability and grammaticality: A reply to Riemer. *Language Sciences* 31: 634–648.
- McEnery, Tony & Andrew Wilson. 1996. *Corpus Linguistics*. Edinburgh: EUP.
- Meyer, Markus. 2009. Sprachliche Wohlgeformtheit – eine kritische Bestandsaufnahme. *Zeitschrift für Sprachwissenschaft* 28: 141–150.
- Olsen, Wendy. 2004. Triangulation in Social Research: Qualitative and Quantitative Methods Can Really Be Mixed. *Developments in Sociology* 20: 103–121.
- Radford, Andrew. 1988. *Transformational Grammar: A First Course*. Cambridge: CUP.
- Riemer, Nick. 2009. On not having read Itkonen: Empiricism and intuitions in the generative data debate. *Language Sciences* 31: 649–662.
- Sampson, Geoffrey. 2001. *Empirical Linguistics*. London: Continuum.
- Schlesewsky, Matthias. 2009. Linguistische Daten aus experimentellen Umgebungen: Eine multi-experimentelle und multimodale Perspektive. *Zeitschrift für Sprachwissenschaft* 28: 169–178.
- Schönefeld, Doris. 2001. *Where Lexicon and Syntax Meet*. Berlin: Mouton de Gruyter.
- Schütze, Carson T. (1996) *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*. Chicago: The University of Chicago Press.
- Schütze Carson T. 2009. Web searches should supplement judgments, not supplant them. *Zeitschrift für Sprachwissenschaft* 28: 151–156.
- Sinclair, John McH. 1987. *Looking up: An Account of the Cobuild Project in Lexical Computing*. London: Collins ELT.
- Sinclair, John McH. 1991. *Corpus, Concordances, Collocations*. Oxford: OUP.
- Steen, Gerard. 2007. *Finding Metaphor in Grammar and Usage*. Amsterdam & Philadelphia: Benjamins.
- Steiner, Erich. 1983. *Die Entwicklung des Britischen Kontextualismus*. Heidelberg: Julius Groos.
- Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collocations: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8: 209–243.
- Stefanowitsch, Anatol & Stefan Th. Gries. 2005. Covarying collexemes. *Corpus Linguistics and Linguistic Theory* 1.1: 1–43.
- Stefanowitsch, Anatol & Stefan T. Gries. 2006. *Corpus-based Approaches to Metaphor and Metonymy*. Berlin: Mouton de Gruyter.
- Stubbs, Michael. 1993. British traditions in text analysis: From Firth to Sinclair. In Mona Baker, Gill Francis, & Elena Tognini-Bonelli, eds., *Text and Technology*, 1–33. Amsterdam & Philadelphia: Benjamins.
- Tahay, Leonard. 2007. Foreword. In Monica González-Márquez, Irene Mittelberg, Seana Coulson & Michael J. Spivey, eds., *Methods in Cognitive Linguistics*, xi–xxi. Amsterdam & Philadelphia: Benjamins.
- Tummers, Jose, Kris Heylen & Dirk Geeraerts. 2005. Usage-based approaches in Cognitive Linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory* 1–2: 225–261.