

In: *Language Sciences* 34 (2012), 651-655.

## **Introduction: Converging data sources in cognitive linguistics**

András Kertész<sup>a\*</sup>, Monika Schwarz-Friesel<sup>b</sup>, Manfred Consten<sup>c</sup>

<sup>a</sup> *University of Debrecen, Institute of German Studies, Pf. 47, H-4010 Debrecen, Hungary*

<sup>b</sup> *Technical University of Berlin, Institute of Languages and Communication, Straße des 17. Juni 135, D-10623 Berlin, Germany*

<sup>c</sup> *Institute of German Linguistics, Fürstengraben 30, University of Jena, D-07737 Jena, Germany*

\* Corresponding author. E-mail address: [andras.kertesz@t-online.hu](mailto:andras.kertesz@t-online.hu)

### **Abstract**

The aim of the paper is to outline the topic of the present special issue. In the first section, the authors give a concise overview of current discussions on the structure and function of linguistic data and evidence in general. They argue that one of the main insights which the discussions led to is the need to integrate different data sources in linguistic theorising. The second section deals with the specific manifestation of these discussions in cognitive linguistics. The authors raise the questions which the papers in the special issue address with respect to cognitive linguistics: What data sources can be integrated?; How can different data sources be integrated?; How does the integration of different data sources affect the acceptability of hypotheses? Finally, in the last section, the structure of the special issue is summarised.

**Keywords:** linguistic data, linguistic evidence, cognitive linguistics, linguistic methodology

### **1. The current debate on linguistic data and evidence**

The history of modern linguistics has been inseparably associated with a series of methodological discussions on different aspects of the empiricalness of linguistic theories. For example, in the seventies of the twentieth century a heated discussion took place around the question of whether mainstream linguistic theories – in particular, generative linguistics – were empirical. The two philosophical positions confronted were that of Popperian critical rationalism and hermeneutics. In the past two decades, the problem was reduced to the relationship between corpus data and introspective data. Against the background of this distinction, one of the standpoints claimed that introspective data were not empirical and therefore not scientific. Generative linguists, on the other hand, regarded corpora on their own, without reference to linguistic intuitions and theories, as useless, and argued for the indispensability of introspective data. Somewhat later, besides introspective and corpus data, interest also turned to experimental data. Nevertheless, during the debate, each of the parties argued for their own data source and their own data type based on this source, excluding others. Arguments for one of the data sources – such as introspection or corpora or experiments in whatever sense – served as arguments for particular theories and against others. Thereby, ‘evidence’ was considered a designated data type stemming from a source whose reliability was not questioned by the proponents of the approach at issue. Therefore,

evidence was assumed to be the firm basis serving the acceptance or the rejection of particular hypotheses.

Currently, the nature of the problem seems to have changed again. The question is no longer whether linguistic theories should be empirical or not but rather, what kinds of data are permitted and how they function in different linguistic investigations (Penke and Rosenbach, 2007, p. 1). We highlight two important factors that triggered this change of perspective: one is rooted in cognitive linguistics, the other in generative linguistics.

Cognitive linguistics, as it developed in Germany in the early 1990s (Schwarz 1994, 1997, 2008), was directed towards interdisciplinary research and the acknowledgement of external data (namely, data from empirical and experimental research). It was recognised that introspection as the sole method in linguistics could not explain all cognitive phenomena concerning language and that some kind of data expansion was indispensable in order to fully understand and explain language as a mental system:

“Only through the employment of experimental methods and diagnostic procedures is it possible to get an insight into the representation, organisation and processing of linguistic units. Introspectively reached data, seen as internal evidence, must be complemented by external evidence resulting from empirical and experimental research. If linguistic models disregard external evidence delivered by empirical and experimental cognition research, there is a real danger that they will ‘theorize away’, and in the process lose all contact with cognitive reality. A theoretical approach in the framework of cognitive linguistics is empirically adequate on condition that it is compatible with psychological and neurophysiological results. All the data realms that may provide insight in the architecture and the dynamics of the language faculty deserve attention. Consequently, cognitive linguistics offers a broadened methodological approach than general, and above all, generative linguistics.” (Schwarz 1992, p. 47, <sup>3</sup>2008, pp. 47-48; translated by the author)

Another important factor was Schütze’s (1996) seminal work that fiercely criticised the data processing method of introspection in general and the way generative linguistics used introspective data to support its hypotheses in particular. Schütze proposed a considerably more refined treatment of linguistic data and the introduction of new data sources into generative linguistics such as experiments.

Following these initiatives, the new millennium has seen a series of works centring on the problem of linguistic data in this sense (see for example Borsley (Ed.), 2005, Kepser and Reis (Eds.), 2005, Penke and Rosenbach (Eds.), 2007, Schwarz, <sup>3</sup>2008, Schwarz, (Ed.) 2004, Kertész and Rákosi, 2012 etc.). The following quotations illustrate the contemporary significance of the problem of linguistic data and evidence:

“When we posted our Call for Papers, we were surprised by the response we received. Obviously, we had hit some nerve, and the field was *ripe for* a discussion of linguistic evidence.” (Penke and Rosenbach, 2007, p. vii; emphasis added)

“It seems likely, however, that questions about data will loom somewhat *larger* in linguistics in the future than they have in the past.” (Borsley, 2005, p. 1475; emphasis added)

“Above all it is clear that questions about data are *more important* than is sometimes assumed.” (Borsley, 2005, p. 1479; emphasis added)

“In short, linguistic evidence is an *extremely important* topic as well as a challenging problem for linguists of all persuasions.” (Kepser and Reis, 2005, pp. 1-2; emphasis added)

“Given the fundamental nature of the problem, linguistic evidence is *a remarkably new topic* of linguistic discussion.” (Kepser and Reis, 2005, p. 2; emphasis added)

As the literature cited above has shown, there is now a tendency pointing towards the acceptance of various tenets, which include the following (see Kertész and Rákosi, 2012 for a detailed overview of the state of the art):

- There is no single data source that is to be preferred against others at the outset. The great diversity of data sources has been acknowledged by metalinguistic research, none of which is *a priori* better or worse than another (Kepser and Reis, 2005, p. 1; Borsley, 2005, 1476 ff., Penke and Rosenbach, 2004, p. 514).
- Thereby, all data sources – among them, introspective, corpus and experimental data alike – are considered to be problematic in the sense that their structure, function and effectiveness is made subject to explicit metatheoretical reflection.
- There is a need to combine several data types rooted in different sources in order to enhance the acceptability of the hypotheses the data are expected to support.
- There are no general criteria for defining evidence; rather, evidence has to be relativised to a couple of factors such as problems, theoretical frameworks and the ‘open-mindedness’ of linguists (Borsley, 2005, p. 1476, Penke and Rosenbach, 2004, p. 493, Kepser and Reis, 2005, p. 3).

## **2. The problem of data and evidence in cognitive linguistics**

The papers in the present special issue will focus on particular aspects of the problem of data and evidence in *cognitive linguistics*. With respect to the state of the art outlined above, this reduction is motivated by the following considerations.

First, in cognitive linguistics a methodological discussion has been emerging which is closely related to the general problem of linguistic data/evidence and the trend summarised above (see e.g. Kristiansen et al. (Eds.), 2006, Gonzales-Marquez et al. (Eds), 2007). The discussion focuses primarily on the relationship between introspective data, corpus data and experimental data. On the one hand, the most widely applied classic approaches to cognitive linguistics (associated with names like Langacker, Lakoff, Talmy, Fillmore and Fauconnier) make primarily use of introspective data. Therefore, just like generative linguistics, the latter are exposed to the criticism that their data processing does not rest on an empirically acceptable method (see e.g. Geeraerts, 2006, etc.). They are contrasted with the perspective which bases cognitive linguistic research on corpus linguistic (Stefanowitch and Gries, 2006a, b) or experimental data (see e.g. Gentner and Bowdle, 2008, Gentner and Wolff, 1997 etc.). On the other hand, instead of contrasting introspection with corpus linguistic or experimental data processing, a series of approaches try to combine introspection with these data sources. Just to mention one example, a significant part of cognitive metaphor research focuses on the experimental control of introspectively obtained hypotheses which maintain that there are metaphorical concepts underlying metaphorical expressions (see e.g. Gibbs, 2006, Gibbs et al., 2004 etc.). These approaches are motivated by the background assumption that the more data sources support a particular hypothesis, the more acceptable this hypothesis is.

Second, the data problem in cognitive linguistics has not only been raised in analogy to the general problem, but from specific points of view as well. For example, the applicability of neurolinguistic methods (Dodge and Lakoff, 2005) or specific data processing techniques such as Steen's (2007) 'metaphorical propositional analysis' have been considered.

Third, raising the data problem has also been related to the fierce criticism of classic approaches to cognitive linguistics (Rakova, 2002, Haser, 2005, Geeraerts, 2006) and through this, it affects the foundational issues of theory construction in cognitive linguistics.

Fourth, during the past decade, attempts have also been made to integrate cognitive linguistic research with data processing methods stemming from quite different traditions such as, for example, conversation analysis (Deppermann, 2007, Liebert et al. (Eds.), 1997) or grammaticalisation research (Bybee, 2007).

Against this background, the present special issue focuses on one of the basic aspects of current discussions on data in cognitive linguistics, namely *the convergence of data sources*. The contributions consider the following questions:

- What data sources can be integrated?
- How can different data sources be integrated?
- How does the integration of different data sources affect the acceptability of hypotheses?

### **3. The structure of the special issue**

The contributions consider the above questions from different points of view. They both tackle the foundational issues of cognitive linguistics related to them and illustrate the possible answers suggested by examples and case studies taken from the practice of cognitive linguistics research.

The first paper by Monika Schwarz-Friesel entitled *On the status of external evidence in the theories of cognitive linguistics: compatibility problems or signs of stagnation in the field? or: Why do some linguists behave like Fodor's input systems?* gives a general overview of the state of the art, criticises current positions and suggests a fruitful new perspective. Her starting point is that in recent publications on linguistic data and evidence the ways in which linguists use data have been ignored, however fundamental this aspect is. She maintains that many linguists insist on the methods and ideas they traditionally use and are not aware of the need for a radical change in linguistic work. The author discusses three antagonistic approaches to cognitive linguistics. Two of them have in common that they do not reflect on their basic assumptions and lack methodological and theoretical openness. Therefore, the current debate about data is a symptom of a fundamental crisis in cognitive linguistics. Moreover, in the practice of research, questions about data are inseparable from questions about the people who handle these data. Although this partly sociological, partly scientific problem is of great relevance for the progress of linguistic inquiry, it has been neglected so far.

Klaas Willems's article *Intuition, introspection and observation in linguistic inquiry* explores the relationship between observation of naturally occurring utterances and intuition in linguistic inquiry. In the first part of the paper, the author starts his line of argumentation with the insight that the reason why cognitive linguistics in general and Cognitive Grammar in particular have not succeeded in developing a usage-based model is that two problematic assumptions underpin the cognitive paradigm: the way meaning is defined with special respect to the encyclopaedic conception of linguistic semantics; and the failure to distinguish between intuition and introspection as methods of inquiry. The second part of the paper

discusses a case study on the cognitive account of the meaning of the accusative and dative case with two-way prepositions in German. The case study serves to exemplify how questions concerning the empiricalness of linguistic inquiry emerge, what the role of introspective judgments is and what status linguistic intuition has. The author also shows that quantitative data are not sufficient for obtaining a fully coherent analysis of case meanings. Rather, to establish the intersubjective, language-specific meaning potentials of cases on the basis of intuition – as distinguished from introspection – is important as well.

In her paper *The fabulous engine: Strengths and flaws of psycholinguistic experiments*, Csilla Rákosi analyses the nature and limits of psycholinguistic experiments which are related to problems raised in cognitive linguistics. According to a widely held view among linguists, experiments can be regarded as completely reliable data sources that make it possible to turn linguistics into a mature empirical discipline meeting the standards of the natural sciences. However, the author argues that this requirement is rooted in obsolete and unworkable tenets of the standard view of the analytical philosophy of science. In order to provide a viable methodology for psycholinguistic experiments motivated by problems discussed in cognitive linguistics, she suggests a strategy consisting of three components. First, metascientific reflection on the nature of experiments as data sources in linguistics has to consider the research activities of linguists conducting experiments. Second, insights gained by philosophers of science studying experiments in the natural sciences have to serve as possible points of departure for analogies between the structure and function of experiments carried out in physics, chemistry, biology, etc. on the one hand, and in linguistics on the other hand. Third, the methodological guidelines have to be compatible with a general account of linguistic theorising. The workability of this strategy is tested with the help of a case study on psycholinguistic experiments in metaphor research.

In their contribution *Circularity effects in corpus studies – why annotations sometimes go round in circles* Manfred Consten and Annegret Loll investigate the status of annotated corpus data in empirical linguistics. Their central claim is that annotators should be regarded as co-producers of data. Accordingly, annotations depend on certain theoretical categories and are therefore theory-laden. As regards the minimal requirements which annotated data are expected to meet, first they have to allow the reconstruction of the original raw data, and second, annotations must follow guidelines in order to avoid that the annotator's decisions are arbitrary. The paper exemplifies the problems which result from the close relation between annotation categories and their theoretical prerequisites by two quantitative corpus studies on anaphora as a discourse-functional phenomenon. One is a newspaper corpus study, the other a study on a second-language acquisition corpus. The studies yield the finding that it is two circles that are responsible for the problems which the authors discuss. The first circle stems from the interplay of deductive and inductive procedures. This interplay causes an impact of theory on annotation. The second circle is rooted in the relations between language structures and their discourse functions. The latter are not observable independently of the structural features of the utterance.

Within cognitive metaphor research, there are rival approaches which fiercely criticise each other's hypotheses and methods. Despite their conflicts, cognitive theories of metaphor have a lot in common. This follows from the circumstance that they belong to cognitive science and share its basic assumptions and goals. The paper by András Kertész, Csilla Rákosi and Péter Csatár – entitled *Data, problems, heuristics and results in cognitive metaphor research* – deals with the problem of how this complicated relationship can be accounted for. The first part of the paper presents a metascientific approach based on B. von Eckardt's notion of 'research framework'. Within this framework, the second part of the paper compares Lakoff and Johnson's, Glucksberg's, and Gentner's cognitive theories of metaphor, and analyses their influence on each other. The authors show that due to the criticism and

interaction among the three rival approaches, they continuously revise their research problems, data processing methods, theoretical model and heuristics.

Rita Brdar-Szabó and Mario Brdar devote their paper to *The problem of data in the cognitive linguistic research on metonymy: a cross-linguistic perspective*. The paper has two goals. Its general goal is to tackle a series of questions which have been raised by cross-linguistic (contrastive) data and which call for a revision of some widely accepted views. Its special methodological goal is to demonstrate the cyclic interaction of the introspection-driven research and the authentic-data driven research. The authors exemplify this interaction by two case studies. The first centres on the referential metonymy of the CAPITAL-FOR-GOVERNMENT type, while the second examines the illocutionary metonymy that motivates constructions that realise the instructional speech act in cooking recipes. Analysing the common features of the two case studies and extending them to a number of languages, the authors conclude that it is structural factors that motivate the contrastive facts. Raising the question of what the functional-cognitive background of the structural facts is, reveals a deeper sort of motivation.

Arnulf Deppermann's essay asks the question: *How does 'cognition' matter to the analysis of talk-in-interaction?* Both conversation analysis (CA) and discursive psychology (DP) maintain that cognitive issues are respecified as discursive phenomena. However, they do this differently. The author exemplifies the approach of discursive psychologists who investigate discursive practices of talking about mental phenomena and use mental predicates by a case study on the use of constructions with German *verstehen* ('to understand') in conversation. The approach of some conversation analysts who examine the way participants display mental states in talk-in-interaction is illustrated by a case study on how grammatical constructions are used to display different types of inferences drawn from a partner's prior turn. On the one hand, the author argues for the constructivist and anti-essentialist stance of the two approaches from a methodological point of view. On the other hand, he also shows that there is no avoiding tacit assumptions about cognitive processes in the practice of CA and DP. He concludes that research on talk-in-interaction would be well advised to integrate explicitly a cognitive perspective.

Finally, in his paper entitled *Pragmatic evidence, context, and story design: An essay on recent developments in experimental pragmatics* Jörg Meibauer raises the question of what pragmatic evidence is. Thereby he focuses on cognitive pragmatics which in his view deals with the acquisition, production and interpretation of utterances. He compares two prominent accounts of cognitive pragmatics, namely, Levinson's Presumptive Meanings approach and Sperber and Wilson's Relevance Theory. In discussing the way the two approaches handle Grice's notions of Generalised Conversational Implicature and Particularised Conversational Implicature, he argues that stories play an important role both in traditional philosophical approaches to pragmatics and in experimental pragmatics. Thus, pragmatic evidence is to be conceived of as a story that reflects the pragmatic intuitions of a speaker or hearer related to a given context. The author concludes that experimental pragmatics would profit from the systematic handling of stories representing contexts.

Since the special issue narrows down the general problem of linguistic data and evidence to the questions (a)-(c) with respect to cognitive linguistics, it tackles basic methodological problems of linguistic research in general and of cognitive linguistics in particular. It addresses cognitive scientists, cognitive linguists, philosophers of science and theoretical linguists alike. Above all, the special issue is meant to contribute to the clarification of a foundational problem of linguistic research which will stay in the foreground of interest in the coming years.

## Acknowledgements

Work on the present special issue as well as the present paper was supported by the Research Group for Theoretical Linguistics of the Hungarian Academy of Sciences and the projects OTKA NI 68436 and TÁMOP 4.2.1./B-09/1/KONV-2010-0007.

## References

- Borsley, R.D. (Eds.), 2005. Data in Theoretical Linguistics. Special issue of *Lingua* 115/11.
- Bybee, J., 2007. Diachronic Linguistics. In: Geeraerts, D., Cuyckens, H. (Eds.), *The Oxford Handbook of Cognitive Linguistics*. Oxford University Press, Oxford, pp.945-987.
- Deppermann, A., 2007. *Grammatik und Semantik aus gesprächsanalytischer Sicht*. De Gruyter, Berlin/New York.
- Dodge, E., Lakoff, G., 2005. Image Schemas: From Linguistic Analysis to Neural Grounding. In: Hampe, B. (Ed.), *From Perception to Meaning: Image Schemas in Cognitive Linguistics*. De Gruyter, Berlin/New York, pp. 57-91.
- Haser, V., 2005. Metaphor, Metonymy, and Experientialist Philosophy. *Challenging Cognitive Semantics*. De Gruyter, Berlin/New York.
- Geeraerts, D., 2006. Methodology in Cognitive Linguistics. In: Kristiansen et al. (Eds.), pp. 21-49.
- Gentner, D., Bowdle, B., 2008. Metaphor as Structure-Mapping. In: Gibbs, R.W. (Eds.), *The Cambridge Handbook of Metaphor and Thought*. Cambridge University Press, Cambridge, pp. 109-128.
- Gentner, D., Wolff, P., 1997. Alignment in the Processing of Metaphor. *Journal of Memory and Language* 37, 331-355.
- Gibbs, Raymond W., 2006. Introspection and Cognitive Linguistics: Should we Trust our Own Intuition? *Annual Review of Cognitive Linguistics* 4, 135-151.
- Gibbs, R. W., Lima, P. L. C., Francozo, E., 2004. Metaphor is grounded in embodied experience. *Journal of Pragmatics* 36, 1189-1210.
- Gonzales-Marquez, M., Mittelberg, I., Coulson, S. and Spivey, M. J. (Eds.), 2007. *Methods in Cognitive Linguistics*. Benjamins, Amsterdam/Philadelphia.
- Kepler, S., Reis, M. (Eds.), 2005. *Linguistic Evidence: Empirical, Theoretical and Computational Perspectives*. De Gruyter, Berlin/New York.
- Kertész, A., Rákosi, Cs., 2012. *Data and Evidence in Linguistics: A Plausible Argumentation Model*. Cambridge University Press, Cambridge.
- Liebert, W.-A., Redeker, G., Waugh, L. (Eds.), 1997. *Discourse and Perspective in Cognitive Linguistics*. Benjamins, Amsterdam/Philadelphia.
- Kristiansen, G., Michel, A., Dirven, R., Ruiz de Mendoza Ibáñez, F. (Eds.), 2006. *Cognitive Linguistics: Current Applications and Future Perspectives*. Mouton de Gruyter, Berlin/New York.
- Penke, M., Rosenbach, A., 2007. Preface. In: Penke and Rosenbach (Eds.), pp. vii-ix.
- Penke, M., Rosenbach, A. (Eds.), 2007. *What Counts as Evidence in Linguistics?* Benjamins, Amsterdam/Philadelphia.
- Rakova, M. 2002. The Philosophy of Embodied Realism: A High Price to Pay? *Cognitive Linguistics* 13, 215-244.
- Schütze, C. T., 1996. *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*. The University of Chicago Press, Chicago.
- Schwarz, M., 1992 (<sup>3</sup>2008). *Einführung in die Kognitive Linguistik*. Francke, Tübingen.

- Schwarz, M. (Ed.) 1994. Kognitive Semantik/Cognitive Semantics. Ergebnisse, Probleme, Perspektiven. Narr (= TBL 395), Tübingen.
- Schwarz, M., 1997. Kognitive Linguistik? Eine Straße in den Geist! In: Kertész, A. (Ed.), 1997. Metalinguistik im Wandel. Lang, Frankfurt a. M., 19–29
- Steen, G. J., 2007. Finding Metaphor in Grammar and Usage: A Methodological Analysis of Theory and Research. Benjamins, Amsterdam/Philadelphia.
- Stefanowitsch, A., Gries, S. (Eds.), 2006a. Corpus-based Approaches to Metaphor and Metonymy. De Gruyter, Berlin/New York.
- Stefanowitsch, A., Gries, S. (Eds.), 2006b. Corpora in Cognitive Linguistics: Corpus-based Approaches to Syntax and Lexis. De Gruyter, Berlin/New York.