

p 250435306

Studies in Generative Grammar 85

Linguistic Evidence

Empirical, Theoretical
and Computational Perspectives

Editors

Henk van Riemsdijk
Harry van der Hulst
Jan Koster

Edited by

Stephan Kepser
Marga Reis

GG 430.402

Mouton de Gruyter
Berlin · New York

Mouton de Gruyter
Berlin · New York

Evidence in Linguistics

Stephan Kepser and Marga Reis

As is well known, the central objects of linguistic enquiry – language, languages, and the factors/mechanisms systematically (co-)governing language acquisition, language processing, language use, and language change – cannot be directly accessed; they must be reconstructed from the accessible manifestations of linguistic behaviour. These manifestations constitute the realm of possibly usable linguistic data. Since they fall into many types – introspective data, corpus data, data from (psycho-)linguistic experiments, synchronic vs. diachronic data, typological data, neurolinguistic data, data from first and second language learning, from language disorders, etc. –, and since each type, apart from historical data, can be instantiated by infinitely many tokens, the linguist's central task of building theories about the above-mentioned linguistic objects is invariably bound up with several empirical tasks as well: (i) collecting/selecting a representative as well as reliable database from one or more data types, (ii) evaluating the various data types as to how they reflect linguistic competence (recall that even so-called primary data from introspection as well as authentic language production are complex performance data involving different nonlinguistic factors), (iii) assessing the relationship between the various data types such that comparison between studies of the same issue based on different data types is possible, and potential conflicts in results can in principle be resolved.

As will be obvious, the three empirical tasks are largely interdependent. However, they are to a considerable degree dependent on linguistic theorising as well: Task (i) must typically be solved for specific linguistic problems, the specific shape of which is determined by linguistic theory proper. Tasks (ii) and (iii) must be related to theories about the interaction of linguistic competence with nonlinguistic faculties and factors in performance. Thus, gaining relevant linguistic evidence from the mass of potentially available data is neither a trivial matter nor a purely methodical one that can be pursued in isolation from concrete linguistic enquiry and their theoretical concerns. Moreover, providing useful data collections (be it appropriately annotated corpora, collections of controlled speaker judgements, experimentally elicited data, etc.) is also a linguistically challenging 'practical' task. In short, *linguistic*

evidence is an extremely important topic as well as a challenging problem for linguists of all persuasions.

Given the fundamental nature of the problem, *linguistic evidence* is a remarkably new topic of linguistic discussion. Traditionally, concrete speech events, i.e. naturally occurring written or spoken utterances, were taken without further ado as the only relevant source of linguistic data, although the need for 'abstracting' the linguistically relevant traits from these data was by no means unknown (cf. Bühler 1932: 97, 1934: 14–15). Within structuralism, this tradition gained explicit methodological and theoretical status ('distributionalism'). Thus the explicit mentalistic turn of generative grammar which claimed the priority of explanatory over descriptive goals and introspective over corpus data was bound to inspire a heated debate concerning the status of linguistics as an empirical science in general and the nature of proper linguistic evidence in particular. This debate, however, died down after the seventies without virtually any consequences on linguistic practice: Generative linguists continued relying more or less on introspective data gained in rather informal ways, non-generative linguists continued relying more or less on corpus data that were often just as informally obtained.

In recent times, this has begun to change. Regarding the use of introspective data, an important turning point was the book by Schütze (1996), who was the first to argue forcefully for a systematic approach to the collection of speaker judgements. Since then, many authors have followed his lead and shown in various ways the necessity of controlling the many factors that influence speaker judgements in order to obtain more reliable data. As a consequence, there is a growing awareness among generative linguists that it is imperative to collect introspective data in systematically controlled ways, and moreover useful to complement them by data from other sources, both of which increasingly influences their linguistic practice. Regarding corpus data, the importance of this source of evidence has grown significantly since about the mid nineties, when when really large amounts of language data of many types became electronically available and easily accessible for the first time. Frequently, these data were annotated in linguistically relevant ways which made these sources even more valuable. At the same time, computational linguists developed methods of accessing and evaluating these corpora. Consequently, linguists have now access to corpora that are several orders of magnitude larger than they were before. And the size and number of such corpora is still rapidly growing. Hence the renaissance of corpus linguistics to be observed since the nineties is by no means a coincidence.

Both developments, by voiding mutual reservations concerning solidity

and practicability of method, have also paved the way for a rapprochement between introspective and corpus linguists, as evidenced by several recent publications in which the question of what should count as linguistic evidence is discussed from either perspective, on the whole opting for using corpus as well as introspective evidence (see, e.g., the recent special issues of *Lingua* and *Studies in Language*). But an astonishing number of participants in the discussion are still trying to argue that one of these types of linguistic evidence is generally significantly superior to the other (see, e.g., Lehmann (2004) and Borsley (2005b)).

It is one of the main aims of this volume to overcome the corpus data versus introspective data opposition and to argue for a view that values and employs different types of linguistic evidence each in their own right. Evidence involving different domains of data will shed different, but altogether more, light on the issues under investigation, be it that the various findings support each other, help with the correct interpretation, or by contradicting each other, lead to factors of influence so far overlooked. This ties in naturally with the fact we started out with that there are more domains and sources of evidence that should be taken into account than just corpus data and introspective data. These insights may sound simple, but, unfortunately, a look into the discussion on evidence in linguistics shows that they are not generally accepted.

Apparently, it is not so much the origin of evidence that counts. What is more important is adequacy and the status of the data as true 'evidence'. Adequacy means that the data put forward to support a certain claim actually do so. This can only be decided on an individual level, i.e., for the particular linguistic problem in question. It is therefore of no concern to us here. Whether certain data can be regarded as true evidence touches the key questions of reliability and reproducibility of data. Reproducibility of data is a base demand in all areas of science for these data to be considered true evidence for something. Typical counterexamples are example sentences held to be (un)acceptable by virtue of the linguist's own judgement only (especially if fortified by the belief in individual 'dialects'), or quoting a single occurrence of a construction found in the world wide web, which is by some regarded as the largest accessible corpus as support for this construction's grammatical existence. Reliability encompasses reproducibility, but requires more. A proper analysis and control of the factors that influence the constitution of the data are necessary as well. With reproducibility and reliability secured, data can be fruitfully used as evidence for strengthening or refuting hypotheses. The contributions to the present book are examples of how this can be done

in linguistic practice.

An important aspect of this book, and a consequence of what we pointed out at the outset about the theoretical underpinnings of issues of linguistic evidence, is the absence of purely abstract discussions of methodologies. Rather, all issues concerning linguistic evidence taken up in the various contributions are addressed in relation to specific linguistic research problems. The main reason for this is our belief that it is only with respect to concrete problems that the quality of the method and of the various types of evidence brought to bear on them can be evaluated. Apart from that it is just more convincing to see how using different types of evidence and different methods of obtaining it may in fact further our understanding of such concrete problems.

It stands to reason then that a volume on 'Linguistic Evidence' should cover a wide range of data types (and methods for turning data into evidence) to be applied to an equally wide range of linguistic phenomena. The present volume does: As for data types, many sources of evidence come into play: corpus data, introspective data, psycholinguistic data, data from computational linguistics, language acquisition data, data from historical linguistics, and sign language data. In several contributions, different data types are comparatively evaluated, which yields particularly insightful results. What is remarkably absent is quarrel about the status of introspective vs. corpus data; both are recognised throughout as equally valid sources of evidence. We take this as a hopeful sign that the longstanding but fruitless either-or confrontation of these data types will finally be overcome. Different ways for gaining linguistic evidence are also well represented in this volume, papers applying/exploring psycholinguistic methods forming perhaps the largest group. A good part of them is concerned with experimental data from language processing, exploring systematic ways for measuring and interpreting these data. But there are also papers exploring methods for collecting reliable as well as reproducible grammaticality judgements. These data types and methods are applied insightfully to phenomena from such diverse areas as syntax, semantics, phonology, morphology, psycholinguistics, historical linguistics, language acquisition, corpus linguistics, computational linguistics, and patholinguistics. For books, such diversity of topics is not always a virtue. But in this case, it serves to underline the fundamental importance issues of linguistic evidence have for all fields of linguistics. It also indicates that awareness of these issues has by now reached almost all these fields.

The present book is based on the conference on *Linguistic Evidence. Empirical, Theoretical, and Computational Perspectives* that took place in Tübingen, January 29 – February 1, 2004. It was organised by the Collaborative

Research Centre (SFB) 441 on "Linguistic Data Structures. On the Relation between Data and Theory in Linguistics" at the University of Tübingen, which has supported in-depth studies of linguistic evidence in all its aspects since 1999. The contributions to this volume are elaborated versions of the conference presentations, plus a paper by H. Weiß designed to complement the historical section. Unfortunately, four papers presented at the conference were not submitted for publication.

The editors of this volume wish to express their gratitude to the members of the collaborative research centre (SFB) 441 on *Linguistic Data Structures* at the University of Tübingen for many interesting discussions on key issues of evidence in linguistics, and for their vigorous support when organising the above-mentioned conference. In this regard we owe particular thanks to Sam Featherston, Beate Starke, and Dirk Wiebel. We also want to thank the members of the conference programme committee for their excellent work.

When preparing the present volume we received again generous support by many, to whom we are very grateful. In particular, we wish to thank the colleagues who reviewed the papers for publication, for their extremely useful comments and criticisms, and the group of helpers without whom editing this volume might have become a mission impossible: Iris Banholzer, Ansgar Höckh, Chris Sapp, and Bettina Zeisler. We are also grateful to the German Science Foundation (DFG) for their generous support of the collaborative research centre 441 and of the conference on *Linguistic Evidence*.

Stephan Kepser and Marga Reis

November 2005

References

- Borsley, Robert D., (ed.)
2005a *Data in Theoretical Linguistics*, volume 115(11) of *Lingua*.
- Borsley, Robert D.
2005b Introduction. *Lingua*, 115(11): 1475–1480.
- Bühler, Karl
1932 *Das Ganze der Sprachtheorie, ihr Aufbau und ihre Teile*. In *Bericht über den XII. Kongreß der Deutschen Gesellschaft für Psychologie*, pp. 95–122. Fischer, Jena.

6 *Stephan Kepser and Marga Reis*

- 1934 *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Fischer, Jena.
2nd edition Stuttgart, 1965.
- Lehmann, Christian
- 2004 Data in linguistics. *The Linguistic Review*, 21: 175–210.
- Penka, Martina and Anette Rosenbach
- 2004 What counts as evidence in linguistics. *Studies in Language*, 28(3):
480–526.
- Schütze, Carson
- 1996 *The Empirical Base of Linguistics: Grammaticality Judgments and
Linguistic Methodology*. University of Chicago Press.