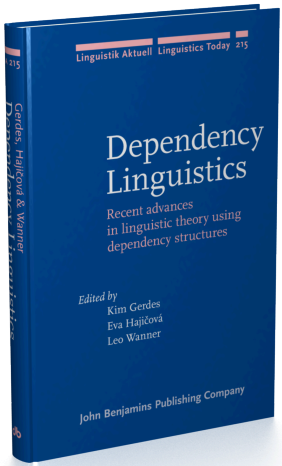
Linguistik Aktuell/Linguistics Today

215

Foreword

Kim Gerdes | Sorbonne Nouvelle
Eva Hajičová | Charles University
Leo Wanner | ICREA and Pompeu Fabra University

https://doi.org/10.1075/la.215.002for



Pages ix–xii of
Dependency Linguistics: Recent advances in linguistic theory using dependency structures
Edited by Kim Gerdes, Eva Hajičová and Leo Wanner
[**Linguistik Aktuell/Linguistics Today**, 215]
2014. xi, 355 pp.

© John Benjamins Publishing Company

This electronic file may not be altered in any way. For any reuse of this material written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

For further information, please contact rights@benjamins.nl or consult our website at benjamins.com/rights

John Benjamins Publishing Company

Foreword

Kim Gerdes, Eva Hajičová & Leo Wanner
Sorbonne Nouvelle / Charles University / ICREA and Pompeu Fabra University

1. Is dependency a linguistic domain?

Are you a dependency linguist? Is Dependency Linguistics a field of linguistics at all? It certainly is a field that is different from syntax or language acquisition in that it does not have a delimited part of language that it tries to shed light on. It rather resembles Computational Linguistics that is defined by its primary tool, computers, because Dependency Linguistics is defined by the primary linguistic representation it uses: dependency structures, i.e. hierarchical relations primarily between the basic units of language (words, morphemes, semantemes, illocutionary units, etc.). The links are often labeled and categorized, and the units and links together commonly form a graph, usually a directed acyclic graph, sometimes even a tree or a chain. You are thus a dependency linguist if you assign a prominent role to these kinds of structures in your analyses – be it in syntax, semantics, discourse analysis, or any other field of language sciences.

The choice of dependency graphs as the primary representation opposes Dependency Linguistics directly to constituent analyses, where the main goal is to define groupings, commonly continuous chunks of text that result in constituent or even phrase structure trees, if the groupings can be hierarchically organized. Just like dependency, the constituent approach stems from but is not limited to syntax. The idea of phrase structures influenced many other fields of linguistics as it was, and sometimes still is, the only formal representation taught in the ordinary linguistics curriculum.

The generative grammatical tradition that, in its origins, solely attempts to construct a system that distinguishes grammatical from ungrammatical sentences, left linguistics in a state where the result of the grammatical analysis, namely phrase structure, was difficult to connect to deeper (semantic, conceptual) structures. The result was a far-reaching separation between, on the one side Natural Language Processing (NLP) that needed deeper analyses for parsing, translation, classification, generation etc., and, on the other side, generative linguistics that

built complex structures with the declared goal to model Language as a whole, where the structures got more and more complicated the further the described language diverged from English. In the second half of the 20th century, only a few linguists, often referring themselves to Tesnière, continued to describe language in terms of dependency, mainly because they were working on free word order languages, where the use of phrase structure is more clearly absurd.

Since the 1990s, NLP is turning towards dependency analysis and in the past five years, dependency has become hegemonic. Thus, it is illustrative that the very large majority of parsers presented at recent NLP conferences are explicitly dependency based; Machine Translation is also moving more and more towards dependency, etc. It seems, however, that the connection between computational linguists and dependency linguists remains sporadic: What happens commonly is that someone transfers an existing treebank into a dependency format that fits his or her needs, and other NLP researchers attempt to reproduce this annotation, with statistical or rule-based grammars. Not that the situation was better when parsers still automated phrase structure construction and linguistics discussed "move alpha". Yet, we believe that the state of affairs is different today and dependency linguists and computational linguists have a lot to share.

2. This volume

The Dependency Linguistics conference Depling 2011 brought together a number of scholars from theoretical and applied linguistics as well as from the domain of Natural Language Processing. All the submissions to the conference were critically reviewed and commented upon by renowned specialists in the field, three to four for each paper. Their comments were an important contribution to the final versions of the papers. This volume unites some of the linguistically oriented articles from the conference (in their revised and extended forms) and gives a general overview over the current state of the art in dependency centered linguistics.

The volume starts out by delimiting the domain of research: Igor Meľčuk, arguably one of the most prominent dependency linguists of our days, gives in his paper an introduction to dependency and states criteria that allow for the distinction of different levels of dependency – from semantics, over syntax and morphology. In a similar vein, the two following papers in the spirit of the Prague school of dependency linguistics delimit dependency grammar: Jarmila Panevová and Magda Ševčíková pick up the classical question of the separation between grammar and lexicon, and Pavlína Jinová, Lucie Mladová and Jiří Mírovský tackle the question whether dependency can be extended beyond sentence boundaries. Then, Henrik Høeg Müller and Iørn Korzen demonstrate how the Copenhagen

Dependency Treebank has been extended from syntax to morphology and semantics. Dependency treebank creation is also the subject of Kristiina Muhonen and Tanja Purtonen's paper: They exemplify how to create satisfactory annotation for controversial structures such as, for instance, ellipsis using the help of the future users of the treebank.

The other papers in the volume address more specific phenomena in the light of the dependency paradigm. The paper by Orsolya Vincze and Margarita Alonso Ramos explores the problem of the representation of Spanish person names at different levels of linguistic dependency representation.

The next two papers tackle the problem of the representation of coordination in a dependency framework from different angles: Nicolas Mazziotta proposes a new analysis based on data from Old French, and Markus Dickinson and Marwa Ragheb tackle the annotation of coordination on a corpus of learner language.

Eva Maria Duran Eppler analyses exciting data from bilingual code-switching by taking into account a measure that is specific to dependency grammar: The distance between words that are connected by a dependency link. Vered Silber-Varod shows that in some cases those dependency links can go beyond prosodic boundaries of spoken Hebrew. Continuous segments of the dependency tree, called Catenae, are shown by Thomas Groß to be instrumental not only for modeling syntax, but also for the analysis of morphological phenomena.

The concern of Kateřina Rysová's paper is word order in Czech taking into account information structure and semantics. Timothy Osborne addresses the problem of rising, illustrating it mainly on English and German.

Andreas Pankau tackles the astonishing fact of Wh-Copying in German, and Dina El Kassas, finally, looks at the contrary phenomenon in Arabic: Pronoun dropping.

Overall, we can state that this comprehensive and coherent collection of papers covers all fields of contemporary dependency linguistics, ranging from definitional challenges of dependency to concrete analyses of various cross-linguistic phenomena of syntax in its interplay with phonetics, morphology, and semantics, including phenomena for which classical simple phrase-structure based models have proven to be unsatisfactory.