Syntactic Analysis

An HPSG-Based Approach

In syntactic analysis, as in linguistics generally, the skills required to first identify, and then make sense of, complex patterns in linguistic data involve a certain specific kind of reasoning, where various alternatives are entertained and modified in light of progressively broader empirical coverage. Rather than focus on transmitting the details of complex theoretical superstructures, this textbook takes a practical, analytical approach, starting from a small set of powerful analytic tools, applied first to simple phenomena and then to the passive, complement, and raising/control constructions. The analytic tools are then applied to unbounded dependences, via detailed argumentation. What emerges is that syntactic structure and intricate networks of dependencies linking different parts of those structures are straightforward projections of lexical valence, in tandem with very general rules regulating the sharing of feature values. Featuring integrated exercises and problems throughout each chapter, this book equips students with the analytical tools for recognizing and assessing linguistic patterns.

ROBERT D. LEVINE is Professor of Linguistics at Ohio State University. He is the co-author of *The Unity of Unbounded Dependency Constructions* (with Thomas E. Hukari, 2015) and the editor of *Formal Grammar* (1992) and *Studies in Contemporary Phrase Structure Grammar* (with Georgia M. Green, Cambridge, 2010). He has also published many articles in journals such as *Natural Language and Linguistic Theory*, *Language*, *Linguistic Inquiry*, *Linguistics and Philosophy* and *Journal of Linguistics*.

CAMBRIDGE TEXTBOOKS IN LINGUISTICS

Syntactic Analysis

An HPSG-Based Approach

In this series:

R. CANN *Formal Semantics*
J. LAVER *Principles of Phonetics*
F. R. PALMER *Grammatical Roles and Relations*
M. A. JONES *Foundations of French Syntax*
A. RADFORD *Syntactic Theory and the Structure of English: A Minimalist Approach*
R. D. VAN VALIN, JR, AND R. J. LAPOLLA *Syntax: Structure, Meaning and Function*
A. DURANTI *Linguistic Anthropology*
A. CRUTTENDEN *Intonation* Second edition
J. K. CHAMBERS AND P. TRUDGILL *Dialectology* Second edition
C. LYONS *Definiteness*
R. KAGER *Optimality Theory*
J. A. HOLM *An Introduction to Pidgins and Creoles*
G. G. CORBETT *Number*
C. J. EWEN AND H. VAN DER HULST *The Phonological Structure of Words*
F. R. PALMER *Mood and Modality* Second edition
B. J. BLAKE *Case* Second edition
E. GUSSMAN *Phonology: Analysis and Theory*
M. YIP *Tone*
W. CROFT *Typology and Universals* Second edition
F. COULMAS *Writing Systems: An Introduction to their Linguistic Analysis*
P. J. HOPPER AND E. C. TRAUGOTT *Grammaticalization* Second edition
L. WHITE *Second Language Acquisition and Universal Grammar*
I. PLAG *Word-Formation in English*
W. CROFT AND A. CRUSE *Cognitive Linguistics*
A. SIEWIERSKA *Person*
A. RADFORD *Minimalist Syntax: Exploring the structure of English*
D. BÜ RING *Binding Theory*
M. BUTT *Theories of Case*
N. HORNSTEIN, J. NUÑES AND K. GROHMANN *Understanding Minimalism*
B. C. LUST *Child Language: Acquisition and Growth*
G. G. CORBETT *Agreement*
J. C. L. INGRAM *Neurolinguistics: An Introduction to Spoken Language Processing and its Disorders*
J. CLACKSON *Indo-European Linguistics: An Introduction*
M. ARIEL *Pragmatics and Grammar*
R. CANN, R. KEMPSON AND E. GREGOROMICHELAKI *Semantics: An Introduction to Meaning in Language*
Y. MATRAS *Language Contact*
D. BIBER AND S. CONRAD *Register, Genre and Style*
L. JEFFRIES AND D. MCINTYRE *Stylistics*
R. HUDSON *An Introduction to Word Grammar*
M. L. MURPHY *Lexical Meaning*
J. M. MEISEL *First and Second Language Acquisition*
T. MCENERY AND A. HARDIE *Corpus Linguistics: Method, Language and Practice*
J. SAKEL AND D. L. EVERETT *Linguistic Fieldwork: A Student Guide*
A. SPENCER AND A. LUIS *Clitics: An Introduction*
G. CORBETT *Features*
A. MCMAHON AND R. MCMAHON *Evolutionary Linguistics*
B. CLARK *Relevance Theory*
B. LONG PENG *Analyzing Sound Patterns*
B. DANCYGIER AND E. SWEETSER *Figurative Language*
J. BYBEE *Language Change*
S. G. THOMASON *Endangered Languages: An Introduction*
A. RADFORD *Analysing English Sentences* Second edition
R. CLIFT *Conversation Analysis*

Earlier issues not listed are also available

Syntactic Analysis

An HPSG-Based Approach

ROBERT D. LEVINE

Ohio State University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

484324, 2nd Floor, Ansari Road, Daryaganj, Delhi – 110002, India

79 Anson Road, #06–0406, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107018884
10.10179781139093521

First published 2017

Printed in the United Kingdom by Clays, St Ives plc

A catalogue record for this publication is available from the British Library

Library of Congress Cataloging-in-Publication data
Names: Levine, Robert, 1947– author.
Title: Syntactic analysis : an HPSG-based approach / Robert D. Levine.
Description: Cambridge ; New York, NY : Cambridge University Press, [2016] |
Series: Cambridge textbooks in linguistics | Includes bibliographical references and index.
Identifiers: LCCN 2016022602 | ISBN 9781107018884 (Hardback)
Subjects: LCSH: Grammar, Comparative and general–Syntax. |
Head-driven phrase structure grammar. | Generative grammar. | Linguistic analysis (Linguistics) |
BISAC: LANGUAGE ARTS & DISCIPLINES / General.
Classification: LCC P158.4 .L384 2016 | DDC 415.01/823–dc23 LC record
available at https://lccn.loc.gov/2016022602

ISBN 978-1-107-01888-4 Hardback
ISBN 978-1-107-61412-3 Paperback

Contents

Preface

There is a long tradition in syntax textbooks whereby the early sections (or chapters) present a particular 'take' on the key properties of human language (e.g., its unboundedness, the fact that speakers can pass intuitive and more or less definitive judgments on the well-formedness as stand-alone sentences of arbitrary strings of words, and so on), followed by an overview of fundamental data structure that are said to best capture these properties, followed by a series of chapters applying those data structures to a range of natural language data. Along the way, technical refinements are introduced, attempts are made to formalize the often informal statements of descriptive machinery given at the outset to jump-start the discussion, and the discussion increasingly becomes focused on the content of the theoretical framework, with natural language data used as points of departure for exploring that content. Every such framework seems to have its own 'set piece' examples which demonstrate its explanatory reach, its ability to capture apparently profound generalizations with a minimal number of additional assumptions. The virtues of this narrative organization are obvious: the point of any science is to capture the behavior of its objects of interest as parsimoniously as possible, which in turn requires an analytic toolkit to capture the generalizations that represent that science's *discoveries* about its domain of inquiry. In order to say anything useful about that domain, students must first acquire a basic working familiarity with those conceptual tools by applying them to simpler phenomena (typically using optimally simple or idealized data sets) and progressively refine and expand their mastery of the framework by tackling increasingly challenging or even open-ended problems. So far, so good.

But this kind of storyline faces a certain kind of risk: the text becomes in effect a kind of recipe book of stock analyses, with large chunks of the thinking that went into these analyses presented as faits accomplis, which students are expected to internalize and extend to new data. The result is a heavily 'top-down' presentation of syntax, making it a matter of mastering a typically complex set of technical tools, specialized notations, and axioms. But anyone who has spent much time doing syntax is well aware that, in practice, research in the field has more the quality of a series of difficult challenges, counterexamples, analytic directions that go nowhere, and novel data that force investigators to backtrack and modify their thinking continuously. What makes syntax so attractive and exciting as a field of study are the demands it places on its practitioners for flexibility, lateral thinking, and persistence. These aspects of the field are, I think,

inherent in the nature of the kind of data that syntacticians try to make sense of, and in order to teach syntax, we need to focus on eliciting from students just this kind of flexible thinking and skill in finding analyses that work. Many modern textbooks, regardless of theoretical stance, seem to me to favor detailed presentations of a preferred system, but what students need to master in order to pursue syntax successfully is the ability to construct well-thought-out arguments for specific proposals. Ideally, the details of a particular framework emerge incrementally as a body of successful analyses is built up, with new problems leading to extensions or revisions of the original proposals.

A near-perfect example of what I have in mind here is the great Soames/Perlmutter textbook *Syntactic Argumentation and the Structure of English*, and while the following chapters are arranged in a quite different, more traditional format, I've tried to follow as much as possible Soames and Perlmutter's expository practice in two respects:

- emphasizing the burden on the investigator not only to propose an analysis which accounts for the data in question, but also to experiment with a few plausible alternatives and push them as far as possible before pronouncing them untenable;
- relying on a small, powerful set of analytic tools, and trying to get as much mileage out of their use as possible before going on to introduce elaborations, and then only the most conservative innovations that are needed to capture the observed regularities.

To equip students with the argumentation skills necessary for making sense of a range of syntactic dependencies, Soames and Perlmutter took the ordered transformational rule cycle as their conceptual workhorse. In the following chapters, I use instead the satisfaction of lexically specified valence specifications, with the formal platform of Head-driven Phrase Structure Grammar's (HPSG) feature logic, as per Pollard and Sag (1994), assumed in the background, as the central explanatory key to a wide range of local dependencies, incorporating analytic insights that originated in the paradigm-transforming work of the pioneers of Generalized Phrase Structure Grammar, HPSG's direct ancestor. From a pedagogical standpoint, the lexically based version of HPSG employed here (as vs the version of HPSG which has evolved over the past two decades, via increasing reliance on hierarchies of often very elaborate syntactic types, into a version of construction grammar) is particularly user-friendly; phenomena which were treated by cyclic 'relation-changing' rules in the 'Standard Theory' era – passive, extraposition, *there* insertion, Raising and similar operations – are all simply statable in terms of valence specifications, and identity between the values of valence features, using the logic of path identities that the 1994 avatar of HPSG was based on.

But I want to emphasize that this book is *not* intended to be an introduction to strongly lexicalist HPSG; rather, as its title suggests, it's an introduction to the art of reasoning about the phrase structure of natural language sentences

and their subparts, *grounded in* the formally explicit and consistent technology of feature structures and their description logics (with these formalisms hidden from the user), and guided by what I take to be the fundamental insight of HPSG: that syntactic structure is nothing but a reflection of the steps mandated by the grammars of individual languages through which the valence requirements of lexical items are satisfied. It's important that the conceptual resources of a framework have explicit denotations – something which is unfortunately not the case in a good deal of contemporary syntactic theorizing – but formally secure foundations are nothing more than a necessary condition on saying something that actually corresponds to a well-defined object. Arguing for the superiority of a particular analysis requires that one not only recognizes the pattern or regularity that needs to be captured, and how a given proposal duly captures that pattern, but also that other possibly quite plausible alternatives be ruled out by confronting them with contradictory data. And there is no simple procedure for determining what such data might be. The narrative thread in the rest of this book therefore consistently follows the mindset of the analyst confronting a range of possible solutions to an increasingly wide-ranging set of interlocking problems. This is, in fact, pretty much the permanent situation of syntacticians. As Paul Postal put it, in perhaps the most eloquent passage in the whole literature of grammatical theory:

> we remain with hardly a single reasonably articulated analysis of any component of the grammar which even approaches relative stability or unchallengeability. Proposal after proposal … has collapsed or slid into the murk of competing claims and contradictory evidence … as the documentation of the factual enormity of, for example, English grammar continues and expands … we can … see with increasing clarity the quite incredible scope of human grammar and correspondingly the limitations of any set of principles that can be posited today to reconstruct this system. Even the best theory is none too good. One must, I think, be led to an increased respect for the subject under study and, ultimately, for the unknown forces which brought it into existence. (Postal 1972: 160–162)

What Postal wrote in 1972 is as true today as ever. And it holds even with respect to the framework that this volume is couched in. HPSG, like all other phrase-structure based approaches, has significant lacunæ in its coverage, particularly in the areas of coordination and ellipsis, two vast domains of phenomena which raise serious questions about the ultimate viability of phrase structure configuration as the appropriate representational language for expressions in natural languages, and which I therefore have had nothing to say about in this book. So far as semantics is concerned, the somewhat different problem is one of intrinsic difficulty: the only formally explicit framework mapping HPSG syntactic representations to actual truth-conditional representations, Lexical Resource Semantics, is technically far too difficult to include in a source with this book's intended audience. Rather than substituting one of the various markup languages which have, for most of HPSG's evolution, been used as a substitute

for a genuine semantics, I've used very informal characterizations of relevant truth conditions where necessary, in the expectation that instructors using this text will supplement the text with their own favorite syntax/semantics translation mechanisms.

Despite these shortcomings, the approach taken in HPSG to both local and nonlocal dependencies gives students the critical analytic means not only to understand individual syntactic dependencies but also to see how they can be chained together in arbitrarily complex ways. Thus, the lexicon and a very small number of lexical and structural combinatory rules will license sentences such as *It seems to have already been decided that John was guilty*, where the independent analyses of extraposition, passive, and 'raising' properties of both *seems* and the auxiliaries jointly entail the well-formedness of such sentences as a straightforward consequence of the feature path identities imposed by the HPSG constraint system. Following the first chapter laying the basic arguments for, and the most perspicuous representation of, the internal hierarchical structure of sentences, the next four chapters develop a set of interconnected arguments about the organization of the auxiliary complex in English sentences and the consequent basic picture of clause structure, relatedness between constructions as an epiphenomenon of systematic relations between the valence properties of morphologically related lexical items, and the propagation of information in nonfinite clauses. The addition of nonlocal dependencies requires an additional mechanism beyond the simple valence saturation which is sufficient to handle the example just given, for reasons that I've taken some pains to lay out in detail in Chapter 6.

But I think it's also important for students studying syntax at even a fairly basic level to appreciate that much of the data which we've taken to reflect something about syntactic combinatorics is in fact a reflection of interacting functional factors – working memory, processing routines and their inherent bottlenecks, pragmatic coherence (or lack of coherence), prosodic requirements, and so on. For this reason, I've included a final chapter, part of which looks at what were regarded for a long time as syntactic island phenomena, and summarizes some of what has been learned over the past extremely productive two decades about the extragrammatical sources of island effects. These findings are of compelling interest in their own right, I think, but for students there is a particularly important message: on the one hand, syntax may actually be, not easier, but *simpler* than we thought it was – but on the other, the lesson of this research is that one can't do syntax in isolation from psycholinguistics, pragmatics and semantics, and prosody, at least if one wishes to avoid what I believe are the problematic assumptions of the past.

I am enormously indebted to a number of people for the conversations, arguments, and collaborations which have shaped my thinking about the content of framework which serves as the platform for the material in this book: Carl Pollard, Robert Borsley, Detmar Meurers, Manfred Sailer, Frank Richter, Gert Webelhuth, Stefan Müller, Tibor Kiss, Erhard Hinrichs, Geoff Pullum,

the late Ivan Sag and, most especially, my long-time friend and collaborator Thomas Hukari, none of whom bear the slightest responsibility for any shortcomings in what I've written. Thanks also to Hoskuldur Thrainsson for some very useful information about Icelandic, and to Bob Borsley for discussing with me data from Welsh which show that it belongs to the class of languages which mark extraction pathways. A special debt to the almost thirty years of my graduate and undergraduate syntax students at Ohio State University, who shaped my thinking about both syntax and the teaching of syntax.

Finally, my sincere thanks to Gert Webelhuth, who gave me invaluable detailed comments on an early draft, to Helen Barton, my unfailingly patient and encouraging editor at CUP, and to Daphne and Adrian.