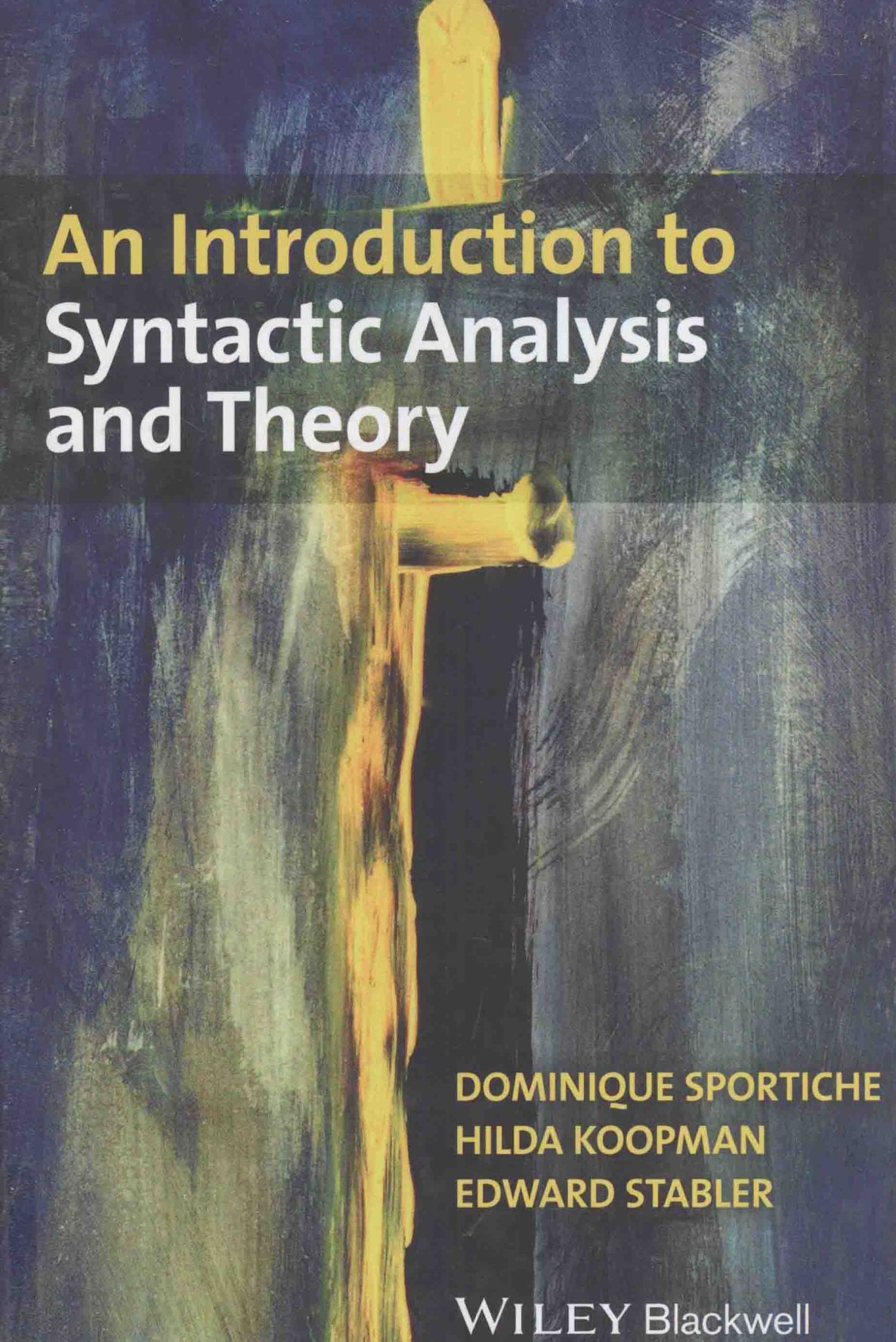
An Introduction to
Syntactic Analysis
and Theory
DOMINIQUE SPORTICHE
HILDA KOOPMAN
EDWARD STABLER
WILEY Blackwell

An Introduction to Syntactic Analysis and Theory

DOMINIQUE SPORTICHE
HILDA KOOPMAN
EDWARD STABLER

WILEY Blackwell

This edition first published 2014

Registered Office
John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial Offices
350 Main Street, Malden, MA 02148-5020, USA
9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

For details of our global editorial offices, for customer services, and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

Library of Congress Cataloging-in-Publication Data

Sportiche, Dominique.
An introduction to syntactic analysis and theory / Dominique Sportiche, Hilda Koopman, Edward Stabler.
pages cm
Includes index.
ISBN 978-1-4051-0016-8 (cloth) – ISBN 978-1-4051-0017-5 (pbk.)
1. Grammar, Comparative and general–Syntax. 2. Generative grammar. I. Sportiche, Dominique, author.
II. Koopman, Hilda Judith, author. III. Stabler, Edward P., author.
P291.S57 2013
415–dc23
2013015034

A catalogue record for this book is available from the British Library.

Cover image: © Fotosearch
Cover design by E&P Design

Set in 9.5/11.5pt Galliard by SPi Publisher Services, Pondicherry, India
Printed in Malaysia by Ho Printing (M) Sdn Bhd

1 2014

An Introduction to Syntactic Analysis and Theory

To Noam with gratitude beyond words, whose influence is found on every page, and whose charismatic ideas have made intellectual life exciting, Chomsky-style.

Acknowledgments

This book was written more slowly than we anticipated, over years of teaching syntax to advanced undergraduate students and graduate students at UCLA. This means many people have contributed to it in many different ways. We would like to thank:

- Robin Clark who inspired a very first attempt to write an uncompromising introduction to contemporary syntactic theory. We hope you will like the result.
- The generations of students at UCLA and elsewhere, whose reactions, positive and negative, have forced us to work harder.
- The colleagues who have used earlier versions of this book and gave us very valuable feedback: Ora Matushansky, Léa Nash, and Elena Soare at Université Paris 8, Vincent Homer at École normale supérieure in Paris, Maire Noonan at McGill University, Chris Collins at NYU, Greg Kobele at the University of Chicago, and Benjamin George, Peter Hallman, Anoop Mahajan, Keir Moulton, Robyn Orfitelli, and Martin Prinzhorn at UCLA.
- Three colleagues who gave us extremely extensive comments: Joseph Emonds, Chris Collins, and Leston Buell. We tried our best to incorporate your many suggestions.
- The anonymous reviewers who provided encouragement, and criticisms that we tried to address.
- Our teaching assistants who have experienced the effects of earlier versions from the trenches and our graduate students, whose comments made this book much better than it could ever have been without them: Natasha Abner, Byron Ahn, Melanie Bervoets, Ivano Caponigro, Isabelle Charnavel, Vincent Homer, Tomoko Ishizuka, Ananda Lima, Robyn Orfitelli, Dave Schueller, Sarah van Wagenen, Jun Yashima.
- The many colleagues whose work and taste have influenced us in too many ways to list and who are also part of this book: Alec Marantz, Anna Szabolcsi, Barry Schein, Benjamin Spector, Danny Fox, Gennaro Chierchia, Guglielmo Cinque, Hagit Borer, Henk van Riemsdijk, Isabelle Charnavel, Jean Roger Vergnaud, Jim Huang, Joseph Aoun, Luigi Rizzi, Maria Luisa Zubizaretta, Maria Polinsky, Martin Prinzhorn, Michal Starke, Misi Brody, Morris Halle, Norbert Hornstein, Philippe Schlenker, Richard Carter, Richard Kayne, Tim Stowell, Vincent Homer, Viola Schmitt.
- Last but not least, Noémie and Sophie who, as often the sole native speakers of English close by, have suffered endless torment over acceptability judgments. If you read this book, you will perhaps feel that the end justified the means.

Contents

Additional and updated materials are available at www.wiley.com/go/syntacticanalysis

1
Introduction

Linguistics is a domain in which language is studied. The notion of language is a common sense notion. In general, a common sense name is not sufficient to characterize a field of research, as there may be many different fields studying more or less the object referred to by the same common-sense name. For example, one can study the oceans from the point of view of a marine biologist, a climate oceanographer, a plate tectonics physicist, a zoologist, a botanist, or a chemist. To get a more precise idea of the field of linguistics, it is necessary to define the type of questions that one is asking about the object of study. Most of what modern linguistics studies falls under the heading of the study of the "language faculty" that human beings possess as part of their human biology. The capacity for language shared by all normal members of our species includes, among other things, the ability to physically manifest thoughts, to linguistically express diverse and original ideas, in diverse and original ways. This faculty also underlies the ability to understand others, to have coherent conversations, and to deduce other people's intentions from their utterances. These properties of language point to a field where the object of study is some mental property of individuals. So linguistics is a part of individual psychology, viewed in contemporary research as part of the study of the human brain.

Investigating this complex and often mysterious faculty appears a daunting task. One way to start is to try to formulate sensible questions about the output of our linguistic capacity – the sounds, the words, the sentences – so as to provide a framework within which incremental knowledge about this faculty can be gained.

1.1 Where to Start

We can start by looking at a simple case of basic linguistic behavior in an individual. When I hear language, my ears are sensing small, rapid variations in air pressure. This vibration, sound, is transformed into nerve impulses that travel to my brain, where the signal is somehow decoded and transformed into an idea, a thought. This is speech perception, or recognition. A similar process occurs in users of sign language: a visual signal is somehow decomposed and converted into an idea. Inversely, when I speak, an idea in my mind is physically manifested through speech or visual signs: this is speech production. These simple observations raise many questions. What exactly goes on when we produce or recognize speech? How does perception or production unfold in real time and how is this coded in brain tissue?

An Introduction to Syntactic Analysis and Theory, First Edition.
Dominique Sportiche, Hilda Koopman, and Edward Stabler.

To understand language processing in the brain, we aim to understand how it rapidly changes state in ways that we can interpret as analytical steps in decoding the linguistic signal. How does the brain do this, exactly? In engineering jargon, we are faced with a problem of "reverse engineering," a common problem for industrial spies. We have a machine – a body, particularly a brain – capable of accomplishing a certain task and we try to understand how it works and how it could be built. We have similar questions in vision, acoustic processing, and other domains of cognitive science. Note that our question is not simply how a particular ability could be produced or imitated in principle. Rather, we aim to identify how it is really produced in the human language user.

This "reverse engineering" problem is very difficult to approach directly. First, to find the physical mechanisms responsible for some ability, we need to have a good understanding of the basic properties of that ability. For example, studies show that certain areas of the brain (Broca's and Wernicke's areas, etc.) are active when we perform certain kinds of linguistic tasks, but to decipher what exactly they are doing, what computations are being performed, we need to know a lot about what these tasks involve. To obtain this kind of understanding, we must initially approach the problems of production and perception abstractly.

We start with the hypothesis that, stored in our mind somehow, we possess some information about certain basic things, some atoms of information that we can deploy in given contexts. As a starting point, we could assume that these atoms are "words" or something similar, such as *table* and *book*, and *eat* and *curious*, and *the* and *before*, and further that these elements have phonetic properties, or instructions on how to physically realize the atoms through speech, as well as meanings, a pairing of the elements to bits of thoughts.

So when it comes to expressing a thought to someone else, a simple story is that I search in my mind for the words corresponding to what I want to talk about and string them together in a particular way to express the thought. Of course, this string of words has to be physically manifested, as in speech or signing, which means that a signal has to be sent to the motor centers that coordinate physical gestures of the mouth or the hand, as the case may be.

Similarly, successful perception of a sentence might be achieved by an electrical signal sent to my brain by the hearing system, which is then segmented by my brain into words, each corresponding to a bit of thought. As my brain receives the particular sequential arrangement of these words, it can also calculate what the combination of the words in the sentence means as a whole.

We can schematize these processes with the following flow chart:

sounds ↔ ordered sets of words ↔ meanings

If we go from left to right, we have a crude model of spoken language perception, from right to left a crude model of spoken language production. If we replaced "sounds" with "signs," we would have crude models of sign language perception and production.

Understanding how this could work, even at an abstract level, is not easy. For example, it is imaginable that the rules governing how sounds are composed into ordered sets of words in a particular situation depend on what the hearer is looking at when perceiving the sounds. This would mean that we have to worry the visual system and how it relates to the linguistic system. This is why we decide to concern ourselves with an even more abstract problem, and will not investigate how the flow chart above really works, how it unfolds in real time. Instead, we will ask: What **necessary linguistic** properties does the language processing task have? By **linguistic**, we mean that we are going to abstract away from the influence of visual clues or background knowledge about the world, and (initially at least) focus on grammatical properties (akin to those found in traditional grammar books). By **necessary**, we mean that we will concentrate on properties that hold across all normal uses of the language, properties that linguistic computations must respect in order to yield the linguistic phenomena we observe.

Here is a sample of more precise questions about necessary linguistic properties:

- Do all orders of words yield meaningful expressions (in the way that all orders of decimal digits represent numbers)? If not, why not?
- Do meaningful word sequences have any structure beyond their linear, temporal ordering? If so, what kind of structure, and why would this structure exist?
- How are the meanings of a sentence determined by (or restricted by) the meanings of its component parts?

The discovery that speech can be symbolically transcribed, can be written down, is certainly among the most momentous human discoveries ever. It allowed the transmission of information across distances and across generations in ways that were never before possible.

> In the famous 1632 *Dialogue Concerning the Two Chief World Systems* written by Galileo Galilei one of the characters, Sir Giovanni Francesco Sagredo, "a man of noble extraction and trenchant wit," so marvels about this invention: "But surpassing all stupendous inventions, what sublimity of mind was his who dreamed of finding means to communicate his deepest thoughts to any other person, though distant by mighty intervals of place and time! Of talking with those who are in India; of speaking to those who are not yet born and will not he born for a thousand or ten thousand years; and with what facility, by the different arrangements of twenty characters upon a page!"

What is less obvious is that this invention is also a theoretical discovery about human psychology: the structure of language, the nature of this system, allows thoughts to be transmitted in this way. One fundamental aspect of this discovery can be stated as follows: the speech signal – even though it is a continuous physical phenomenon – namely a continuous variation of air pressure for normal speech, or continuous motion for sign language – can be represented with a finite (sometimes small) number of discrete units. A particularly striking example of this property is illustrated by alphabetic writing systems, as the quote in the box opposite emphasizes: with a small number of symbols – letters – they can very effectively (partially) code speech. Informally speaking, it is clear that this segmentation of the speech signal occurs at various levels of "graininess." Alphabetic writing systems segment the speech signal in very small units of writing, while Chinese characters or Egyptian hieroglyphic writing systems segment it in somewhat larger units. That this segmentation can occur at various degrees of "magnification" can be illustrated with the following string:

these books burned

Evidence confirms that English speakers segment and analyze this expression at some levels roughly like these, where the number in the leftmost column indicates the number of units in its line (and the first line is written in phonetic alphabet and is meant to represent the sequence of sounds found in this string):

12	ð	i:	z	b	ʊ	k	s	b	ə	r	n	d		
6	Demonstrative		Plural	Noun			Plural	Verb				Past		
3	Demonstrative+Plural			Noun+Plural				Verb+Past						
2	Subject							Predicate						
1	Sentence													

Linguists have extensively documented the relevance of such segmentations for our understanding of language structure. Hypothesizing that these modes of segmentation, these different "levels of

magnification," correspond to real psychological properties of the speech signal, we will need to at least answer the following questions about each level as part of understanding how our flow chart above really works:

- What is the inventory of the smallest pieces, the atomic elements, that are assembled at each level?
- What are the rules or principles that govern how these units can be assembled?

Traditionally, linguists have postulated the following divisions: Phonology studies the atoms and combinations of sounds; Morphology considers the atoms and how words are built; and Syntax considers how words are put together to form phrases. Of course, this preliminary division into such subdomains may not be correct. It could be that the atoms of syntax are not words but morphemes, and that the rules of combination for morphemes are the same as the rules of combination for words. If this were the case, there would really be no distinction between morphology and syntax. Or morphology might be part of phonology. (These kinds of proposals have been seriously explored.) But we will start with this traditional picture of the components, modifying it as necessary. In fact, where syntax books start with words, this book will start with morphology, i.e. the structure of words.

These two questions – what are the atoms of language, and how are they combined – characterize very well what this book is about. We view language as a system of symbols (e.g. sounds, which are paired up with meanings), and a combinatory system, where symbols combine to yield more complex objects, themselves associated with meanings. Here we concentrate on how aspects of this particular symbolic combinatory system are put together: this is a technical manual that almost exclusively worries about how to investigate and characterize the combinatory structure of morphological and syntactic units.

Even though the scope of this book is relatively narrow, the research perspective on language it embodies is part of a much broader research program that tries to characterize cognitive functions. The questions we address here are limited to the structure of the syntactic and morphological combinatory systems but, as will become clear, a substantial amount is known about these systems that suggests that non-trivial principles regulate how it works. This in turn raises all sorts of questions which we will not address but that are central research questions: how is this system put to use when we speak or understand? How much computational power is needed to master such a system? Is language fundamentally shaped by our communicative intentions? Where do the fundamental principles of language structure come from? Are they innate? Are they learned? Are they completely specific to language, or only partially so, or not at all? Are they specific to humans? Or only partially so? How did they appear in our species? Suddenly? Progressively? By themselves?

The approach to the study of language described above took off in the mid 20th century and is now a dynamic field incorporating an increasing panoply of methods or tools, from the methods used by traditional grammarians and language fieldworkers, to laboratory methods originating in experimental psychology, to neuro-imagery, to mathematical methods imported from pure and applied mathematics, to statistical tools and the tools of modern genetics. Noam Chomsky, pictured here on the left, presently (2013) Institute Professor at the Massachusetts Institute of Technology is the most influential pioneer of this research perspective and agenda and the research methods that carry it out.