

Syntax

A Generative Introduction

Third Edition



Andrew Carnie

 **WILEY-BLACKWELL**

Syntax

Introducing Linguistics

This outstanding series is an indispensable resource for students and teachers – a concise and engaging introduction to the central subjects of contemporary linguistics. Presupposing no prior knowledge on the part of the reader, each volume sets out the fundamental skills and knowledge of the field, and so provides the ideal educational platform for further study in linguistics.

- 1 Andrew Spencer *Phonology*
- 2 John Saeed *Semantics*, 3rd Edition
- 3 Barbara Johnstone *Discourse Analysis*, 2nd Edition
- 4 Andrew Carnie *Syntax*, 3rd Edition
5. Anne Baker and Kees Hengevelt *Linguistics*

Syntax

A Generative Introduction
Third Edition

Andrew Carnie



WILEY-BLACKWELL

A John Wiley & Sons, Ltd., Publication

This third edition first published 2013
© 2013 Andrew Carnie

Edition History: Blackwell Publishing, Ltd (1e, 2002 and 2e, 2007)

Wiley-Blackwell is an imprint of John Wiley & Sons, formed by the merger of Wiley's global Scientific, Technical and Medical business with Blackwell Publishing.

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.

The right of Andrew Carnie to be identified as the author of this work has been asserted in accordance with the UK Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data

Carnie, Andrew, 1969-

Syntax : a generative introduction / Andrew Carnie. -- Third Edition.

pages cm. -- (Introducing linguistics ; 16)

Includes bibliographical references and index.

ISBN 978-0-470-65531-3 (pbk.)

1. Grammar, Comparative and general--Syntax. 2. Generative grammar. I. Title.

P295.C37 2012

415--dc23

2012009763

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

Cover Image: Vassily Kandinsky, Sweet Pink, ink with watercolour, 1929, Roethel 925. Sotheby's, London / Superstock. © ADAGP, Paris and DACS, London 2012.

Cover design by Nicki Averill

Set in 10 point Palatino 10 point Arial by Language Technologies Incorporated

Dedicated with love to my parents, Robert and Jean
and in memory of my teacher and mentor, Ken Hale

This book is accompanied by a workbook: *The Syntax Workbook: A Companion to Carnie's Syntax*, available for purchase from <http://www.wiley.com>. The workbook is optional and serves as a supplement for students who want additional practice in syntactic analysis.

Additional online material including bonus chapters can be found on the book's website:

<http://www.wiley.com/go/carnie>

The text formatting in this book has been adapted and modified by ReadSmart®, an algorithm which imposes almost undetectable changes in presentation. This formatting is based on simple psycholinguistic principles, and has been shown in studies to enhance readers' reading speed, comprehension and enjoyment and to make the text more persuasive. The ReadSmart® treatment of the text was provided by Language Technologies Inc., which offers this service for all forms of publication.

www.readsmart.com

Contents

Preface and Acknowledgments	xiii
Part 1 Preliminaries	1
1 Generative Grammar	3
0. Preliminaries	3
1. Syntax as a Cognitive Science	5
2. Modeling Syntax	6
3. Syntax as Science – the Scientific Method	7
4. Where Do the Rules Come From?	18
5. Choosing among Theories about Syntax	29
6. The Scientific Method and the Structure of this Textbook	29
7. Conclusion	31
Ideas, Rules, and Constraints Introduced in this Chapter	31
General Problem Sets	33
Challenge Problem Sets	36
2 Parts of Speech	43
0. Words and Why They Matter to Syntax	44
1. Determining Part of Speech	45
2. The Major Parts of Speech: N, V, Adj, and Adv	48
3. Open vs. Closed; Lexical vs. Functional	51
4. Subcategories and Features	54
5. Conclusion	61
Ideas, Rules, and Constraints Introduced in this Chapter	61
General Problem Sets	62
Challenge Problem Sets	66
3 Constituency, Trees, and Rules	71
0. Introduction	71
1. Rules and Trees	74
2. How to Draw a Tree	90
3. Modification and Ambiguity	96
4. Constituency Tests	98
5. Constituency in Other Languages	100
6. Summary and Conclusion	106
Ideas, Rules, and Constraints Introduced in this Chapter	106
General Problem Sets	107
Challenge Problem Sets	114

4	Structural Relations	117
0.	Introduction	117
1.	The Parts of a Tree	118
2.	Domination	120
3.	Precedence	125
4.	C-command	127
5.	Grammatical Relations	132
6.	Summary and Conclusions	135
	Ideas, Rules, and Constraints Introduced in this Chapter	135
	General Problem Sets	137
	Challenge Problem Sets	144
5	Binding Theory	147
0.	Introduction	147
1.	The Notions <i>Coindex</i> and <i>Antecedent</i>	149
2.	Binding	151
3.	Locality Conditions on the Binding of Anaphors	153
4.	The Distribution of Pronouns	155
5.	The Distribution of R-expressions	156
6.	Conclusion	156
	Ideas, Rules, and Constraints Introduced in this Chapter	157
	General Problem Sets	158
	Challenge Problem Sets	159
Part 2	The Base	163
6	X-bar Theory	165
0.	Introduction	165
1.	Bar-level Projections	167
2.	Generalizing the Rules: The X-bar Schema	172
3.	Complements, Adjuncts, and Specifiers	175
4.	Some Definitional Housekeeping	185
5.	Parameters of Word Order	186
6.	Drawing Trees in X-bar Notation	189
7.	X-bar Theory: A Summary	198
	Ideas, Rules, and Constraints Introduced in this Chapter	199
	General Problem Sets	200
	Challenge Problem Sets	204

7	Extending X-bar Theory to Functional Categories	207
0.	Introduction	207
1.	Determiner Phrases (DPs)	208
2.	A Descriptive Tangent into Clause Types	211
3.	Complementizer Phrases (CPs)	216
4.	Tense Phrases (TPs)	219
	Ideas, Rules, and Constraints Introduced in this Chapter	221
	General Problem Sets	222
	Challenge Problem Sets	224
8	Constraining X-bar: Theta Theory	227
0.	Introduction	227
1.	Some Basic Terminology	228
2.	Thematic Relations and Theta Roles	229
3.	The Lexicon	236
4.	Expletives and the Extended Projection Principle	237
5.	Conclusion	239
	Ideas, Rules, and Constraints Introduced in this Chapter	240
	General Problem Sets	241
	Challenge Problem Sets	245
9	Auxiliaries and Functional Categories	249
0.	Introduction	250
1.	Complementizers	250
2.	Determiners	254
3.	Understanding Tense, Aspect, Voice, and Mood	258
4.	Auxiliaries	263
5.	Conclusion	278
	Ideas, Rules, and Constraints Introduced in this Chapter	279
	General Problem Sets	281
	Challenge Problem Sets	284
Part 3	Movement	287
10	Head-to-Head Movement	289
0.	Introduction	289
1.	Verb Movement ($V \rightarrow T$)	292
2.	T Movement ($T \rightarrow C$)	308
3.	Do-support	311
	Ideas, Rules, and Constraints Introduced in this Chapter	312
	General Problem Sets	313
	Challenge Problem Sets	318

11 DP Movement	323
0. Introduction	323
1. A Puzzle for the Theory of Theta Roles	324
2. Passives	330
3. Case	335
4. Raising: Reprise	339
5. Passives: Reprise	342
6. Tying Up a Loose End	344
7. Conclusion	346
Ideas, Rules, and Constraints Introduced in this Chapter	346
General Problem Sets	347
Challenge Problem Sets	352
12 <i>Wh</i> -movement and Locality Constraints	357
0. Introduction	357
1. Movement in <i>Wh</i> -questions	359
2. Relative Clauses	369
3. Islands	374
4. The Minimal Link Condition	377
5. Echo Questions (<i>Wh</i> -in-situ) in English	382
6. Conclusion	383
Ideas, Rules, and Constraints Introduced in this Chapter	384
General Problem Sets	385
Challenge Problem Sets	387
13 A Unified Theory of Movement	391
0. Introduction	391
1. Move	393
2. Explaining Cross-linguistic Differences	396
3. Scope, Covert Movement, and the MLC	401
4. Conclusion	405
Ideas, Rules, and Constraints Introduced in this Chapter	405
General Problem Sets	406
Challenge Problem Sets	406

Part 4	Advanced Topics	409
14	Expanded VPs	411
0.	Introduction	411
1.	The Problem of Ditransitive Verbs	412
2.	The Active Voice Head and Other Light Verbs	413
3.	Object Shift	416
4.	Ditransitives: Reprise	421
	Ideas, Rules, and Constraints Introduced in this Chapter	425
	General Problem Sets	425
	Challenge Problem Sets	427
15	Raising, Control, and Empty Categories	429
0.	Introduction	430
1.	Raising vs. Control	431
2.	Two Kinds of Raising, Two Kinds of Control	439
3.	Control Theory	445
4.	Another Kind of Null Subject: “Little” <i>pro</i>	449
5.	Conclusion	450
	Ideas, Rules, and Constraints Introduced in this Chapter	451
	General Problem Sets	452
	Challenge Problem Sets	454
16	Ellipsis	457
0.	Ellipsis	457
1.	LF-copying or PF-deletion	460
2.	Antecedent-Contained Deletion and Pseudogapping	466
3.	Conclusion	470
	Ideas, Rules, and Constraints Introduced in this Chapter	471
	General Problem Sets	472
	Challenge Problem Sets	475
17	Advanced Topics in Binding Theory	477
0.	Introduction	477
1.	Levels of Representation	478
2.	The Definition of Binding Domain	480
	Ideas, Rules, and Constraints Introduced in this Chapter	486
	General Problem Sets	487
	Challenge Problem Sets	489

18 Polysynthesis, Incorporation, and Non-configurationality	491
0. Introduction	491
1. Polysynthesis	492
2. Incorporation	494
3. Scrambling and Non-configurationality	496
4. Conclusions	505
Ideas, Rules, and Constraints Introduced in this Chapter	505
General Problem Sets	506
Challenge Problem Sets	507
Conclusions and Directions for Further Study	509
References	511
Index	519

Go to

www.wiley.com/go/carnie

*for bonus chapters on Lexical-Functional Grammar (LFG) and Head-driven
Phrase Structure Grammar (HPSG), as well as a glossary and other
supplementary materials.*

Preface and Acknowledgments

Almost every preface to every syntax textbook out there starts out by telling the reader how different this book is from every other syntax textbook. On one hand, this is often the truth: each author shows their own particular spin or emphasis. This is certainly true of this textbook. For example, you'll be hard-pressed to find another textbook on Principles and Parameters syntax that uses as many Irish examples as this one does. Nor will you find another P&P textbook with a supplementary discussion of alternative theoretical approaches like LFG or HPSG. On the other hand, let's face facts. The basic material to be covered in an introductory textbook doesn't really vary much. One linguist may prefer a little more on binding theory, and a little less on control, etc. In this text, I've attempted to provide a relatively balanced presentation of most of the major issues and I've tried to do this in a student-friendly way. I've occasionally abstracted away from some of the thornier controversies, when I felt they weren't crucial to a student understanding the basics. This may make the professional syntactician feel that I've cut corners or laid out too rosy a picture. I did this on purpose, however, to give students a chance to absorb the fundamentals before challenging the issues. This was a deliberate pedagogical choice. I'm well aware that sometimes I've glossed over controversies, but I think a student has to learn the basics of how the system works before they can seriously critique and evaluate the model. This is a textbook, not a scholarly tome, so its aim is to reach as many students as possible. The style is deliberately low-key and friendly. This doesn't mean I don't want the students to challenge the material I've presented here. Throughout the book, you'll find grey "textboxes" that contain issues for further discussion or interesting tidbits. Many of the problem sets also invite the student to challenge the black and white presentation I've given in the text. I encourage instructors to assign these, and students to do them, as they form an important part of the textbook. Instructors may note that if a favorite topic is not dealt with in the body of the text, a problem set may very well treat the question.

A quick word on the level of this textbook: This book is intended as an introduction to syntactic theory. It takes the student through most of the major issues in Principles and Parameters, from tree drawing to constraints on movement. While this book is written as an introduction, some students have reported it to be challenging. I use this text in my upper-division undergraduate introduction to syntax course with success, but I can certainly see it being used in more advanced classes. I hope instructors will flesh out the book, and walk their students through some of the thornier issues.

This textbook has grown out of my lecture notes for my own classes. Needless to say, the form and shape of these notes have been influenced in terms of choice of material and presentation by the textbooks my own students have used. While

the book you are reading is entirely my fault, it does owe a particular intellectual debt to the following three textbooks, which I have used in teaching at various times:

Cowper, Elizabeth (1992) *A Concise Introduction to Syntactic Theory: The Government and Binding Approach*. Chicago: Chicago University Press.

Haegeman, Liliane (1994) *Introduction to Government and Binding Theory* (2nd edition). Oxford: Blackwell.

Radford, Andrew (1988) *Transformational Grammar: A First Course*. Cambridge: Cambridge University Press.

I'd like to thank the authors of these books for breaking ground in presenting a complicated and integrated theory to the beginner. Writing this book has given me new appreciation for the difficulty of this task and their presentation of the material has undoubtedly influenced mine.

Sadly, during the final stages of putting the first edition of this text together, my dissertation director, teacher, mentor, and academic hero, Ken Hale, passed away after a long illness. Ken always pushed the idea that theoretical syntax is best informed by cross-linguistic research, while at the same time the accurate documentation of languages requires a sophisticated understanding of grammatical theory. These were important lessons that I learned from Ken and I hope students will glean the significance of both by reading this text. While I was writing this book (and much other work) Ken gave me many comments and his unfettered support. He was a great man and I will miss him terribly.

This, the third edition of this book, is quite different from the first two. A reasonably complete list of changes can be found in the instructor's handbook. These include some important changes to definitions that instructors who have used previous editions will want to look at. The major changes to this volume are:

- A companion workbook, with answers, for students to practice assignments.
- New exercises in almost every chapter.
- New chapters on Auxiliaries, Ellipsis and Non-configurational Languages.
- The chapters on LFG and HPSG are now to be found for free on the book's companion website: www.wiley.com/go/carnie.

I hope that instructors and students will find these revisions helpful. I have attempted where possible to take into account all the many comments and suggestions I received from people using the first and second editions, although of course, in order to maintain consistency, I was unable to implement them all.

Acknowledgments:

I'd like to thank the many people who taught me syntax through the years: Barb Brunson, Noam Chomsky, Elizabeth Cowper, Ken Hale, Alec Marantz, Diane Massam, Jim McCloskey, Shigeru Miyagawa, and David Pesetsky. A number of people have read through this book or the previous editions and have given me helpful comments; others have helped on smaller issues but have had no less of an

impact on the work and still others have contributed problem sets or editorial advice. This long list includes: David Adger, William Alexander, Dean Allemang, Diana Archangeli, Ash Asudeh, Brett Baker, Uldis Balodis, Mark Baltin, Luis Barragan, Andy Barss, Dane Bell, Emily Bender, Abbas Benmamoun, Jeff Berry, Tom Bever, Claire Bown, Michael Bauer, Laura Blumenthal, Joan Bresnan, Aaron Broadwell, Dirk Bury, Roy Chan, Ronald Charles, Danny Chen, Deborah Chen-Pichler, Jaehoon Choi, Barbara Citko, Peter Cole, Chris Collins, Jennifer Columbus, Andrew Comrie, Lorenzo Demery, Sheila Dooley, Yehuda Falk, Muriel Fisher, Sandiway Fong, Leslie Ford, Amy Fountain, Stefan Frisch, Alexandra Galani, Jila Ghomeshi, David Gil, Carrie Gillion, Erin Good-Ament, Andrea Haber, Paul Hagstrom, Ken Hale, John Halle, Mike Hammond, Jack Hardy, Heidi Harley, Josh Harrison, Rachel Hayes-Harb, David Heap, Bernhard Heigl, One-Soon Her, Caroline Heycock, Stephan Hurtubise, John Ivens, Eloise Jelinek, Alana Johns, Mark Johnson, Hyun Kyoung Jung, Arsalan Kahnemuyipour, Dalina Kalluli, Simin Karimi, Andreas Kathol, Chris Kennedy, Greg Key, Amy LaCross, Péter Lazar, Carlos Gelormini Lezama, Jeff Lidz, Anne Lobeck, Leila Lomashvili, Sarah Longstaff, Alicia Lopez, Ahmad Reza Lotfi, Ricardo Mairal, Joan Maling, Jack Martin, Diane Massam, Martha McGinnis, Nathan McWhorter, Dave Medeiros, Mirjana Miskovic-Lukovic, Alan Munn, MaryLou Myers, Chris Nicholas, Janet Nicol, Jon Nissenbaum, Peter Norquest, Diane Ohala, Kazutoshi Ohno, Heidi Orcutt-Gachiri, Hiroyuki Oshita, Panayiotis Pappas, Jaime Parchment, Hyeson Park, Barbara Partee, Matt Pearson, David Pesetsky, Colin Phillips, Carl Pollard, Bill Poser, Kristen Pruett, Jeff Punske, Mike Putnam, Janet Randall, Marlita Reddy-Hjelmfelt, Sylvia Reed, Norvin Richards, Frank Richter, Betsy Ritter, Ed Rubin, Jeff Runner, Ivan Sag, Nathan Sanders, Yosuke Sato and his students, Theresa Satterfield, Leslie Saxon, Kevin Schluter, Carson Schütze, Jim Scobbie, Deborah Shapiro, Leah Shocket, Dan Siddiqi, Echo Ki Sihui, Peter Slomanson, Kyle Smith, Norvel Smith, Nick Sobin, Peggy Speas, Megan Stone, Tania Strahan, Joshua Strauss, Maggie Tallerman, Takashi Tanaka, Chris Tancredi, Deniz Tat, Brian ten Eyck, Lisa deMena Travis, Alex Trueman, Adam Ussishkin, Robert Van Valin, Enwei Wang, Shan Wang, Natasha Warner, Andy Wedel, Jennifer Wees, Mary Ann Willie, Marian Wiseley, Dainon Woudstra, Susi Wurmbrand, Kimberley Young, Kim Youngroung, and several anonymous Blackwell and Wiley reviewers. I'm absolutely convinced I've left someone off this large list. If it's you many apologies – I really did appreciate the help you gave me. The students in my *Introduction to Syntax* classes in Michigan in 1997, and in Arizona in 1998–2012, have used all or parts of this textbook. Glynis Baguley, Ada Brunstein, Sarah Coleman, Danielle Descoteaux, Lisa Eaton, Simon Eckley, Charlotte Frost, Graham Frankland, Tami Kaplan, Becky Kennison, Julia Kirk, Leah Morin, Allison Medoff, Anna Oxbury, Rhonda Pearce, Iain Potter, Beth Remmes, and Steve Smith of Wiley-Blackwell and their subcontractors all deserve many thanks for help getting this and the previous two editions to press. My family (my mother Jean, my late father Bob, Morag, Fiona, Pangur) were all incredible in their support and love. Go raibh maith agaibh agus tapadh leibh!

The artwork in chapters 3 and 6 was created by Dane Bell and is used with permission.

Part 1

Preliminaries

chapter 1

Generative Grammar

Learning Objectives

After reading chapter 1 you should walk away having mastered the following ideas and skills:

1. Explain why Language is a psychological property of humans.
2. Distinguish between prescriptive and descriptive rules.
3. Explain the scientific method as it applies to syntax.
4. Explain the differences between the kinds of data gathering, including corpora and linguistic judgments.
5. Explain the difference between competence and performance.
6. Provide at least three arguments for Universal Grammar.
7. Explain the logical problem of language acquisition.
8. Distinguish between learning and acquisition.
9. Distinguish among observational, descriptive and explanatory adequacy.

0. PRELIMINARIES

Although we use it every day, and although we all have strong opinions about its proper form and appropriate use, we rarely stop to think about the wonder of language. So-called language “experts” like William Safire tell

us about the misuse of *hopefully* or lecture us about the origins of the word *boondoggle*, but surprisingly, they never get at the true wonder of language: how it actually works as a complex machine. Think about it for a minute. You are reading this and understanding it, but you have no conscious knowledge of how you are doing it. The study of this mystery is the science of linguistics. This book is about one aspect of how language works: how sentences are structured, or the study of *syntax*.

Language is a psychological or cognitive property of humans. That is, there is some set of neurons in my head firing madly away that allows me to sit here and produce this set of letters, and there is some other set of neurons in your head firing away that allows you to translate these squiggles into coherent ideas and thoughts. There are several subsystems at work here. If you were listening to me speak, I would be producing sound waves with my vocal cords and articulating particular speech sounds with my tongue, lips, and vocal cords. On the other end of things you'd be hearing those sound waves and translating them into speech sounds using your auditory apparatus. The study of the acoustics and articulation of speech is called *phonetics*. Once you've translated the waves of sound into mental representations of speech sounds, you analyze them into syllables and pattern them appropriately. For example, speakers of English know that the made-up word *bluve* is a possible word of English, but the word *bnuck* is not. This is part of the science called *phonology*. Then you take these groups of sounds and organize them into meaningful units (called morphemes) and words. For example, the word *dancer* is made up of two meaningful bits: *dance* and the suffix *-er*. The study of this level of Language is called *morphology*. Next you organize the words into phrases and sentences. *Syntax* is the cover term for studies at this level of Language. Finally, you take the sentences and phrases you hear and translate them into thoughts and ideas. This last step is what we refer to as the *semantic* level of Language.

Syntax studies the level of Language that lies between words and the meaning of utterances: sentences. It is the level that mediates between sounds that someone produces (organized into words) and what they intend to say.

Perhaps one of the truly amazing aspects of the study of Language is not the origins of the word *demerit*, or how to properly punctuate a quote inside parentheses, or how kids have, like, destroyed the English language, eh? Instead it's the question of how we subconsciously get from sounds and words to meaning. This is the study of syntax.

Language vs. language

When I utter the term *language*, most people immediately think of some particular language such as English, French, or KiSwahili. But this is not the way linguists use the term; when linguists talk about *Language* (also known as i-language), they are generally talking about the *ability* of humans to speak any (particular) language. Noam Chomsky also calls this the *Human Language Capacity*. *Language* (written with a capital L) is the part of the mind or brain that allows you to speak, whereas *language* (with a lower-case l) (also known as e-language) is an instantiation of this ability (like French or English). In this book we'll be using *language* as our primary data, but we'll be trying to come up with a model of *Language*.

1. SYNTAX AS A COGNITIVE SCIENCE

Cognitive science is a cover term for a group of disciplines that all have the same goal: describing and explaining human beings' ability to think (or more particularly, to think about abstract notions like subatomic particles, the possibility of life on other planets or even how many angels can fit on the head of a pin, etc.). One thing that distinguishes us from other animals, even relatively smart ones like chimps and elephants, is our ability to use productive, combinatory *Language*. *Language* plays an important role in how we think about abstract notions, or, at the very least, *Language* appears to be structured in such a way that it allows us to express abstract notions.¹ The discipline of linguistics is thus one of the important subdisciplines of cognitive science.² Sentences are how we get at expressing abstract thought processes, so the study of syntax is an important foundation stone for understanding how we communicate and interact with each other as humans.

¹ Whether language constrains what abstract things we can think about (this idea is called the Sapir–Whorf hypothesis) is a matter of great debate and one that lies outside the domain of syntax per se.

² Along with psychology, neuroscience, communication, philosophy, and computer science.

2. MODELING SYNTAX

The dominant theory of syntax is due to Noam Chomsky and his colleagues, starting in the mid 1950s and continuing to this day. This theory, which has had many different names through its development (Transformational Grammar (TG), Transformational Generative Grammar, Standard Theory, Extended Standard Theory, Government and Binding Theory (GB), Principles and Parameters approach (P&P) and Minimalism (MP)), is often given the blanket name *Generative Grammar*. A number of alternate theories of syntax have also branched off of this research program. These include Lexical-Functional Grammar (LFG) and Head-Driven Phrase Structure Grammar (HPSG). These are also considered part of generative grammar; but we won't cover them extensively in this book. But I have included two additional chapters on these theories in the web resources for this book at www.wiley.com/go/carnie. The particular version of generative grammar that we will mostly look at here is roughly the *Principles and Parameters* approach, although we will occasionally stray from this into the more recent version called *Minimalism*.

The underlying thesis of generative grammar is that sentences are generated by a subconscious set of procedures (like computer programs). These procedures are part of our minds (or of our cognitive abilities if you prefer). The goal of syntactic theory is to model these procedures. In other words, we are trying to figure out what we subconsciously know about the syntax of our language.

In generative grammar, the means for modeling these procedures is through a set of formal grammatical *rules*. Note that these rules are nothing like the rules of grammar you might have learned in school. These rules don't tell you how to properly punctuate a sentence or not to split an infinitive. Instead, they tell you the order in which to put your words. In English, for example, we put the subject of a sentence before its verb. This is the kind of information encoded in generative rules. These rules are thought to generate the sentences of a language, hence the name *generative* grammar. You can think of these rules as being like the command lines in a computer program. They tell you step by step how to put together words into a sentence. We'll look at precise examples of these rules in the next few chapters. But first, let's look at some of the underlying assumptions of generative grammar.

Noam Chomsky

Avram Noam Chomsky was born on 7 December 1928, in Philadelphia. His father was a Hebrew grammarian and his mother a teacher. Chomsky got his Ph.D. from the University of Pennsylvania, where he studied linguistics under Zellig Harris. He took a position in machine translation and language teaching at the Massachusetts Institute of Technology. Eventually his ideas about the structure of language transformed the field of linguistics. Reviled by some and admired by others, Chomsky's ideas have laid the groundwork for the discipline of linguistics, and have been very influential in computer science and philosophy. Outside of linguistics, Chomsky is also one of the leading intellectuals in the anarchist socialist movement. His writings about the media and political injustice are also widely read. Chomsky is among the most quoted authors in the world (among the top ten and the only living person on the list).

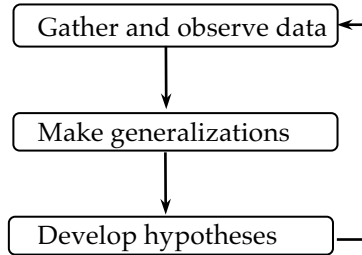
3. SYNTAX AS SCIENCE – THE SCIENTIFIC METHOD

For many people, the study of language properly belongs in the humanities. That is, the study of language is all about the beauty of its usage in fine (and not so fine) literature. However, there is no particular reason, other than our biases, that the study of language should be confined to a humanistic approach. It is also possible to approach the study of language from a scientific perspective; this is the domain of linguistics. People who study literature often accuse linguists of abstracting away from the richness of good prose and obscuring the beauty of language. Nothing could be further from the truth. Most linguists, including the present author, enjoy nothing more than reading a finely crafted piece of fiction, and many linguists often study, as a sideline, the more humanistic aspects of language. This doesn't mean, however, that one can't appreciate and study the formal properties (or rules) of language and do it from a scientific perspective. The two approaches to language study are both valid; they complement each other; and neither takes away from the other.

Science is perhaps one of the most poorly defined words of the English language. We regularly talk of scientists as people who study bacteria, particle physics, and the formation of chemical compounds, but ask your average Joe or Jill on the street what science means, and you'll be hard pressed to get a decent definition. But among scientists themselves, *science* refers to a particular methodology for study: the scientific method. The scientific method dates back to the ancient Greeks, such as Aristotle,

Euclid, and Archimedes. The method involves observing some data, making some generalizations about patterns in the data, developing hypotheses that account for these generalizations, and testing the hypotheses against more data. Finally, the hypotheses are revised to account for any new data and then tested again. A flow chart showing the method is given in (1):

1)



In syntax, we apply this methodology to sentence structure. Syntacticians start³ by observing data about the language they are studying, then they make generalizations about patterns in the data (e.g., in simple English declarative sentences, the subject precedes the verb). They then generate a hypothesis and test the hypothesis against more syntactic data, and if necessary go back and re-evaluate their hypotheses.

Hypotheses are only useful to the extent that they make *predictions*. A hypothesis that makes no predictions (or worse yet, predicts everything) is useless from a scientific perspective. In particular, the hypothesis must be *falsifiable*. That is, we must in principle be able to look for some data, which, if true, show that the hypothesis is wrong. This means that we are often looking for the cases where our hypotheses predict that a sentence will be grammatical (and it is not), or the cases where they predict that the sentence will be ungrammatical (contra to fact).

In syntax, hypotheses are called *rules*, and the group of hypotheses that describe a language's syntax is called a *grammar*.

The term *grammar* can strike terror into the hearts of people. But you should note that there are two ways to go about writing grammatical rules. One is to tell people how they *should* speak (this is of course the domain of English teachers and copy-editors); we call these kinds of rules *prescriptive rules* (as they prescribe how people should speak according

³ This is a bit of an oversimplification. We really have a “chicken and the egg” problem here. You can’t know what data to study unless you have a hypothesis about what is important, and you can’t have a hypothesis unless you have some basic understanding of the data. Fortunately, as working syntacticians this philosophical conundrum is often irrelevant, as we can just jump feet-first into both the hypothesis-forming and the data-analysis at the same time.

to some standard). Some examples of prescriptive rules include “never end a sentence with a preposition”, “use *whom* not *who*” and “don’t split infinitives”. These rules tell us how we are supposed to use our language. The other approach is to write rules that describe how people *actually* speak, whether or not they are speaking “correctly”. These are called **descriptive rules**. Consider for a moment the approach we’re taking in this book. Which of the two types (descriptive or prescriptive) is more scientific? Which kind of rule is more likely to give us insight into how the mind uses Language? We focus on descriptive rules. This doesn’t mean that prescriptive rules aren’t important (in fact, in the problem sets section of this chapter you are asked to critically examine the question of descriptive vs. prescriptive rules), but for our purposes descriptive rules are more important. For an interesting discussion of the prescriptive/descriptive debate, see Pinker’s (1995) book: *The Language Instinct*.

You now have enough information to answer General Problem Sets GPS1 & 2, as well as Challenge Problem Set CPS1 at the end of this chapter. For practice try Workbook Exercise WBE1 in chapter 1 of The Syntax Workbook, an optional companion book to this text.

Do Rules Really Exist?

Generative grammar claims to be a theory of cognitive psychology, so it’s reasonable to ask whether formal rules *really* exist in the brain/minds of speakers. After all, a brain is a mass of neurons firing away, so how can formal mathematical rules exist up there? Remember, however, that we are attempting to *model* Language; we aren’t trying to describe Language exactly. This question confuses two disciplines: psychology and neurology. Psychology is concerned with the mind, which represents the output and the abstract organization of the brain. Neurology is concerned with the actual firing of the neurons and the physiology of the brain. Generative grammar doesn’t try to be a theory of neurology. Instead it is a model of the psychology of Language. Obviously, the rules per se don’t exist in our brains, but they do model the external behavior of the mind. For more discussion of this issue, look at the readings in the further reading section of this chapter.

3.1 An Example of the Scientific Method as Applied to Syntax

Let’s turn now to a real-world application of the scientific method to some language data. The following data concern the form of a specific kind of

noun, called an *anaphor* (plural: *anaphors*; the phenomenon is called *anaphora*). These include the nouns that end with *-self* (e.g., *himself*, *herself*, *itself*). In chapter 5, we look at the distribution of anaphors in detail; here we'll only consider one superficial aspect of them. In the following sentences, as is standard in the syntactic literature, a sentence that isn't well-formed is marked with an *asterisk* (*) before it. For these sentences assume that *Bill* is male and *Sally* is female.

- 2) a) Bill kissed himself.
- b) *Bill kissed herself.
- c) Sally kissed herself.
- d) *Sally kissed himself.
- e) *Kiss himself.

To the unskilled eye, the ill-formed sentences in (2b and d) just look silly. It is obvious that Bill can't kiss herself, because Bill is male. However, no matter how matter-of-factly obvious this is, it is part of a bigger generalization about the distribution of anaphors. In particular, the generalization we can draw about the sentences in (2) is that an anaphor must agree in *gender* with the noun it refers to (its *antecedent*). So in (2a and b) we see that the anaphor must agree in gender with *Bill*, its antecedent. The anaphor must take the masculine form *himself*. The situation in (2c and d) is the same; the anaphor must take the form *herself* so that it agrees in gender with the feminine *Sally*. Note further that a sentence like (2e) shows us that anaphors must have an antecedent. An anaphor without an antecedent is unacceptable. A plausible hypothesis (or rule) given the data in (2), then, is stated in (3):

- 3) An anaphor must (i) have an antecedent and (ii) agree in gender (masculine, feminine, or neuter) with that antecedent.

The next step in the scientific method is to test this hypothesis against more data. Consider the additional data in (4):

- 4) a) The robot kissed itself.
- b) She knocked herself on the head with a zucchini.
- c) *She knocked himself on the head with a zucchini.
- d) The snake flattened itself against the rock.
- e) ?The snake flattened himself/herself against the rock.
- f) The Joneses think themselves the best family on the block.
- g) *The Joneses think himself the most wealthy guy on the block.
- h) Gary and Kevin ran themselves into exhaustion.
- i) *Gary and Kevin ran himself into exhaustion.

Sentences (4a, b, and c) are all consistent with our hypothesis that anaphors must agree in gender with their antecedents, which at least confirms that the hypothesis is on the right track. What about the data in (4d and e)? It appears as if any gender is compatible with the antecedent *the snake*. This appears, on the surface, to be a contradiction to our hypothesis. Think about these examples a little more closely, however. Whether sentence (4e) is well-formed or not depends upon your assumptions about the gender of the snake. If you assume (or know) the snake to be male, then *The snake flattened himself against the rock* is perfectly well-formed. But under the same assumption, the sentence *The snake flattened herself against the rock* seems very odd indeed, although it is fine if you assume the snake is female. So it appears as if this example also meets the generalization in (3); the vagueness about its well-formedness has to do with the fact that we are rarely sure what gender a snake is and not with the actual structure of the sentence.

Now, look at the sentences in (4f–i); note that the ill-formedness of (g) and (i) is not predicted by our generalization. In fact, our generalization predicts that sentence (4i) should be perfectly grammatical, since *himself* agrees in gender (masculine) with its antecedents *Gary* and *Kevin*. Yet there is clearly something wrong with this sentence. The hypothesis needs revision. It appears as if the anaphor must agree in gender and **number** with the antecedent. Number refers to the quantity of individuals involved in the sentence; English primarily distinguishes singular number from plural number. (5) reflects our revised hypothesis.

5) An anaphor must agree in gender and number with its antecedent.

If there is more than one person or object mentioned in the antecedent, then the anaphor must be plural (i.e., *themselves*).

Testing this against more data, we can see that this partially makes the right predictions (6a), but it doesn't properly predict the acceptability of sentences (6b–e):

- 6) a) People from Tucson think very highly of themselves.
- b) *I gave yourself the bucket of ice cream.
- c) I gave myself the bucket of ice cream.
- d) *She hit myself with a hammer.
- e) She hit herself with a hammer.

Even more revision is in order. The phenomenon seen in (6b–e) revolves around a grammatical distinction called **person**. Person refers to the perspective of the speaker with respect to the other participants in the speech act. **First person** refers to the speaker. **Second person** refers to the addressee. **Third person** refers to people being discussed that aren't participating in the

conversation. Here are the English pronouns associated with each person: (*Nominative* refers to the *case* form the pronouns take when in subject position like *I* in “*I* love peanut butter”; *accusative* refers to the form they take when in object positions like *me* in “John loves *me*”. We will look at case in much more detail in chapter 9, so don't worry if you don't understand it right now.)

7)	Nominative		Accusative		Anaphoric	
	Singular	Plural	Singular	Plural	Singular	Plural
1	I	we	me	us	myself	ourselves
2	you	you	you	you	yourself	yourselves
3 masc	he	they	him	them	himself	themselves
3 fem	she		her		herself	
3 neut	it		it		itself	

As you can see from this chart, the form of the anaphor seems also to agree in person with its antecedent. So once again we revise our hypothesis (rule):

- 8) An anaphor must agree in person, gender and number with its antecedent.

With this hypothesis, we have a straightforward statement of the distribution of this noun type, derived using the scientific method. In the problem sets below, and in chapter 6, you'll have an opportunity to revise the rule in (8) with even more data.

You now have enough information to try WBE2, and CPS2 & 3.

3.2 Sources of Data

If we are going to apply the scientific method to syntax, it is important to consider the sources of our data. One obvious source is in collections of either spoken or written texts. Such data are called *corpora* (singular: *corpus*). There are many corpora available, including some searchable through the internet. For languages without a literary tradition or ones spoken by a small group of people, it is often necessary for the linguist to go and gather data and compile a corpus in the field. In the early part of the last century, this was the primary occupation of linguists, and it is proudly carried on today by many researchers.

The linguist Heidi Harley reports in her blog⁴ on an example of using search engines to do linguistic analysis. Harley notes that to her ear, the

⁴ <http://heideas.blogspot.com/2005/10/scalar-adjectives-with-arguments.html>.