

L
I
N
G
U
I
S
T
I
C
S

**A MODERN
COURSE IN
ENGLISH
SYNTAX**

**HERMAN WEKKER
AND LILIANE HAEGEMAN**

A MODERN COURSE IN ENGLISH SYNTAX

HERMAN WEKKER AND
LILIANE HAEGEMAN



London • New York

First published in 1985 by Croom Helm Ltd

Routledge is an imprint of the Taylor & Francis Group

This edition published in the Taylor & Francis e-Library, 2009.

To purchase your own copy of this or any of
Taylor & Francis or Routledge's collection of thousands of eBooks
please go to www.eBookstore.tandf.co.uk.

© 1985 Herman Wekker and Liliane Haegeman

All rights reserved. No part of this book may be
reprinted or reproduced or utilized in any form
or by any electronic, mechanical or other means,
now known or hereafter invented, including
photocopying and recording, or in any
information storage or retrieval system, without
permission in writing from the publishers.

British Library Cataloguing in Publication Data
Wekker, Herman

A modern course in English syntax.

1. English language—Syntax

I. Title II. Haegeman, Liliane

425 PE1361

Library of Congress Cataloguing in Publication Data
Wekker, Herman.

A modern course in English syntax.

1. English language—Syntax. I. Haegeman,
Liliane M.V. II. Title

PE1361.W37 1985 428.2 85—13994

ISBN 0-203-13343-9 Master e-book ISBN

ISBN 0-203-17560-3 (Adobe ebook Reader Format)

ISBN 0-415-03684-4 (Print Edition)

CONTENTS

Preface	xi
1: Aims and Methods	
1.1 What is Syntax?	1
1.2 Representing Sentence Structure	2
1.2.1 Bracketing	2
1.2.2 Labelled Bracketing	4
1.2.3 Tree Diagrams and Phrase Structure Rules	5
1.2.4 Sentence and Discourse	9
1.2.5 Exercises	11
1.3 Constituency Tests	11
1.3.1 Variations on Basic Sentence Patterns	11
1.3.2 Clefting	12
1.3.3 Pseudo-clefting	13
1.3.4 Passivisation	13
1.3.5 Pronominalisation	14
1.3.6 Fronting	14
1.3.7 Exercises	14
2: Constituents	
2.1 Introduction	17
2.1.1 Exercise	17
2.1.2 Word Order	18
2.1.3 Constituents	19
2.2 Sentences	21
2.2.1 Simple Declarative and Interrogative Sentences	21
2.2.2 Compound and Complex Sentences	22
2.2.3 Exercise	25
2.3 Clauses	25
2.3.1 Exercises	30

2.4	Phrases	31
2.4.1	Introduction	31
2.4.2	Noun Phrases	32
2.4.2.1	Elements before the Head	32
2.4.2.2	Pronouns	34
2.4.2.3	One-word Phrases	36
2.4.2.4	Elements after the Head	37
2.4.2.5	Summing up the NP	38
2.4.2.6	Exercise	39
2.4.3	Verb Phrases	39
2.4.3.1	Introduction	39
2.4.3.2	Verb Complements	40
2.4.3.3	Elements before the Head	43
2.4.3.4	The First Auxiliary in the VP	45
2.4.3.5	Summing up the VP	47
2.4.3.6	Exercises	48
2.4.4	Prepositional Phrases	48
2.4.4.1	Summing up the PP	49
2.4.4.2	Exercise	49
2.4.5	Adjective Phrases	50
2.4.5.1	Summing up the AdjP	51
2.4.5.2	Exercises	51
2.4.6	Adverb Phrases	52
2.4.6.1	Summing up the AdvP	53
2.4.6.2	Exercises	53
2.5	Words	53
2.5.1	Introduction	53
2.5.1.1	Word Classes: the Distribution of Words	54
2.5.2	Nouns	55
2.5.3	Verbs	56
2.5.4	Adjectives	56
2.5.5	Adverbs	57
2.5.6	Prepositions	58
2.5.7	Conjunctions	58
2.5.8	Exercises	59

3: Functions

3.1 Introduction: the Verb and its Complements	60
3.2 Grammatical Functions	63
3.2.1 Exercises	65
3.3 Grammatical Functions in S	65
3.3.1 Subject	65
3.3.2 Predicate	67
3.3.3 Sentence Adjuncts	67
3.4 Grammatical Functions in VP	68
3.4.1 Predicative Complements	68
3.4.1.1 Exercise	69
3.4.2 Direct Object	69
3.4.2.1 Exercise	70
3.4.3 Indirect Object	70
3.4.3.1 Exercises	71
3.4.4 Adverbial Complements with Intransitive Verbs	72
3.4.4.1 Exercise	74
3.4.5 Adverbial Complements with Transitive Verbs	74
3.4.5.1 Exercise	76
3.4.6 Predicative Complement +Adverbial Complement: a Complex Function	76
3.4.6.1 Exercise	77
3.4.7 Verb Patterns and Functions: a Summary	77
3.4.7.1 Exercise	79
3.4.8 Adjuncts and Complements	80
3.4.8.1 Exercise	80
3.4.9 Sentence Adjuncts and VP-Adjuncts	81
3.4.9.1 Exercise	83
3.4.10 Predicative Adjuncts	84
3.4.10.1 Exercise	88
3.4.11 Functions of Clauses	88
3.4.11.1 Finite Clauses	88

3.4.11.2 Non-finite and Verbless Clauses	92
3.4.11.3 Empty Subjects in Non-finite and Verbless Clauses	94
3.5 Grammatical Functions in Phrases: a Brief Survey	96
3.5.1 Exercises	98
4: Processes	
4.1 Introduction	99
4.2 Leftward Movement: <i>Wh</i> -movement and Related Transformations	101
4.2.1 Questions: Different Types	101
4.2.2 Subject-Auxiliary Inversion	101
4.2.3 <i>Do</i> -insertion	103
4.2.3.1 Exercise	104
4.2.4 <i>Wh</i> -movement	105
4.2.4.1 Subcategorisation Frames and <i>Wh</i> -movement	106
4.2.4.2 Exercise	110
4.2.4.3 Prepositional Phrases in <i>Wh</i> -questions	110
4.2.4.4 <i>Wh</i> -movement of Subjects	111
4.2.4.5 <i>Wh</i> -movement and \bar{S}	112
4.2.4.6 Exercises	112
4.2.5 Indirect Questions	112
4.2.5.1 Indirect Yes-No Questions	112
4.2.5.2 Indirect <i>Wh</i> -questions	114
4.2.5.3 Long <i>Wh</i> -movement	115
4.2.6 Relative Clauses	116
4.2.6.1 Exercise	119
4.2.6.2 The Meaning of Relative Clauses	119
4.2.6.3 Deletion in Relative Clauses	120
4.2.6.4 <i>That</i> in Relative Clauses	121
4.2.6.5 Adverbial Relative and Free Relative Clauses	122
4.2.6.6 Non-restrictive Relative Clauses	123
4.2.6.7 Exercises	124
4.2.7 Appositive Clauses	125
4.2.7.1 Exercises	126

4.2.8	Reduced Relative Clauses	127
4.2.8.1	Exercises	130
4.2.9	Fronting	131
4.2.9.1	Exercise	131
4.3	Passivisation	132
4.3.1	The Process	132
4.3.2	The <i>By</i> -phrase	134
4.3.3	Exercise	135
4.4	Movement to Focus Position	136
4.4.1	<i>It</i> -extraposition	136
4.4.1.1	<i>It</i> -extraposition in Passive Sentences	136
4.4.1.2	<i>It</i> -extraposition in Active Sentences	139
4.4.1.3	<i>It</i> -extraposition with Non-finite Clauses	140
4.4.1.4	Exercise	141
4.4.1.5	<i>It</i> -extraposition in Verbless Clauses	142
4.4.1.6	<i>It</i> -extraposition with Object Clauses	143
4.4.2	Subject-Verb Inversion	145
4.4.3	<i>There</i> -insertion	148
4.4.4	Extraposition from NP	150
4.4.5	Heavy Object-shift	151
4.4.6	Summing Up	152
4.5	Clefting and Pseudo-clefting	153
4.5.1	Clefting	153
4.5.2	Pseudo-clefting	154
4.6	Substitution by Pro-forms	155
4.7	Ellipsis	156
4.7.1	Subject Deletion	157
4.7.2	Verb Gapping and Backward Gapping	157
4.7.3	Other Examples	158
5:	Non-finite Clauses: Raising and Control	
5.1	Introduction	160
5.2	Finite and Non-finite Clauses	161
5.2.1	Introduction	161
5.2.2	Finite Subordinate Clauses	161

5.2.3 Non-finite Subordinate Clauses with Lexical Subjects	163
5.2.4 Non-finite Clauses without Lexical Subjects	165
5.2.5 <i>For...to</i> Filter	165
5.3 Raising	166
5.3.1 Raising with Passive Verbs	166
5.3.2 Raising with Intransitive Verbs and with Adjectives	169
5.3.3 Summing Up	170
5.3.4 Exercises	171
5.4 Control	172
5.4.1 Non-lexical Subjects	172
5.4.2 Control vs Raising	173
5.4.3 Subject and Object Control	174
5.4.4 Indefinite Control	175
5.4.5 Control Adjectives	176
5.4.6 Exercises	177
6: Reanalysis: a Problem of Bracketing	
6.1 Introduction	179
6.2 Prepositions and Particles	179
6.2.1 Exercises	185
6.3 Prepositional Verbs	186
6.3.1 Exercise	190
6.4 Phrasal-prepositional Verbs and Prepositional Idioms	190
6.4.1 Exercise	192
7: Levels of Structure	
7.1 Introduction	193
7.2 Noun Phrases	194
7.2.1 Flat Structures and Layered Structures	194
7.2.2 <i>One</i> -substitution	196
7.2.3 Layering	197
7.2.4 One-word Phrases	198
7.3 Verb Phrases	199
7.3.1 Layering in VPs	199
7.3.2 Formal Evidence	199

7.4 Prepositional Phrases	200
7.5 Adjective Phrases	200
7.5.1 Complements of Adjs	200
7.5.2 Discontinuous Strings	201
7.6 Adverb Phrases	202
7.7 Summing Up: Layered Structure	202
7.8 Exercise	203
Index	204

PREFACE

This book is an attempt to incorporate some of the less controversial insights from modern theoretical linguistics into the teaching of English syntax. It seeks to strike a balance between the traditional and structural approaches and transformational generative syntax, which is currently the most consistent and influential linguistic framework.

We take the view that it is important for students of English—whether or not they will go on to specialise in English syntax—to be able to interpret the structure of sentences and their component parts. The practice of syntactic analysis, or ‘parsing’ as it is traditionally called, will improve students’ understanding of the structure of English, it will facilitate their comprehension and production of the language and it will enable them to compare and contrast English structures and to perform syntactic operations on them. The ability to analyse structures in their context is a very useful and important skill for undergraduates to acquire and master. It is a skill which involves familiarity with, amongst other things, the principal categories and functions in English, the main syntactic operations, and the methods of syntactic argumentation.

Although the importance of syntactic analysis is generally recognised, none of the presently available grammars of English provides students with a systematic introduction to the methods of English syntactic analysis. Most of these grammars contain excellent inventories of the facts of English, but fail to show students explicitly how to formulate syntactic arguments, and how to apply relevant objective criteria and syntactic tests in the analysis of sentences. This book is meant to fill this gap. It offers an account of what we consider to be the basic rules and principles of English syntax, and it tries to make students aware of syntactic structures and their relations. It also tries to give some insight into the general methodology of syntactic description.

This coursebook is mainly intended for intermediate students of English at universities or colleges who have done one or more introductory grammar courses, but are not yet ready to enter a specialist course. Since in our experience students need a great deal of systematic guidance and practice in the analysis of sentences, we have provided numerous examples and exercises of different types.

A Modern Course in English Syntax also prepares students for further reading in the literature on English syntax, i.e. both the more ‘traditional’ accounts and the current transformational analyses. It offers some new ideas and insights, especially from the Extended Standard Theory (EST), which we have here tried to deal with in an informal manner; theory-internal discussions and irrelevant technicalities are generally avoided. We have found that the material of this book can be used profitably in English syntax courses in conjunction with reference grammars such as those by Randolph Quirk and his associates. In an attempt to make this book as self-contained as possible, we have refrained from including references to English grammars or the linguistic literature.

It must be emphasised that it is not the technical terminology or the grammatical labels that are of central importance in this book, but the approach developed for talking about sentences in English. While the book does not go into all the descriptive or theoretical detail of a fully developed syntax, and will therefore leave many issues undiscussed, it is intended to offer students the techniques for going into more complex matters. We expect that the book will sharpen students’ observations, and make them aware of what are valid arguments and methods for meaningful analyses.

Parts of the material of this volume have been used in English syntax courses at University College, London (mainly students of English literature and language), at the Universities of Nijmegen and Ghent (Dutch and Belgian students of English), and at a teacher training college in Nijmegen.

In developing this course, we have assumed that the student has some knowledge of English grammar. Exercises are included and integrated into the structure of the book. They are meant to give students insight into the main patterns and problems of English syntax, and to appeal to their own linguistic intuitions and creativity.

Our thanks are due to colleagues in several English departments in England, Belgium and The Netherlands for useful comments on an earlier draft of this work. In particular we would mention Ton Broeders, Alasdair MacDonald, Sidney Greenbaum, Annemoon van Hest, Frits Stuurman and Jan Verdonck, who have given us detailed criticisms and sound advice on how to ‘solve’ and present certain problems. The book has greatly benefited from their suggestions. Needless to say, we accept full responsibility for the views expressed here and for any shortcomings that remain. We are also grateful to Diane Crook, Mies Faber and Mariette Wauters for their patience in typing successive versions of the manuscript.

THE BACKGROUND

The approach presented here adopts as its main theoretical background the multi-level analysis of the EST-framework, which distinguishes levels of D-structure and S-structure, and levels of Phonetic form and Logical form. D-structure (or, roughly, ‘underlying structure’) is determined by (a) phrase structure (PS) rules and (b) the Lexicon. PS rules generate the underlying structure of the sentence. The Lexicon contains the store of lexical items of the language, each with a specification of its syntactic category and an insertion frame. Verbs, for example, are specified for the type of complementation they require. Inserting lexical material into the structures derived by the PS rules is restricted by a matching condition: nouns can only be inserted into slots marked N; verbs can only be inserted into V slots, etc. The matching condition has recently been replaced in EST by mechanisms such as Case assignment and the Theta criterion, but these notions are not incorporated into this book. The constituent structure generated by the PS rules and filled in by lexical items gives us the D-structure.

The underlying structures may be affected by operations such as ‘passivisation’, ‘extraposition’, etc. Following recent developments in the transformational framework, this book distinguishes between elementary operations on syntactic patterns and the functional effect that is achieved by the operations. The process of *wh*-question formation, for example, is not seen as a single transformation but rather as the compounding of

- (a) Subject—Auxiliary Inversion, and
- (b) *Wh*-fronting

Hence, the transition from *John saw who* to *Who did John see?* is not a single transformation, but a process involving two transformations. Subject—Auxiliary

Inversion and *wh*-fronting may both operate independently in other environments, producing other functional effects. For example:

Not one moment *did he* regret his decision.

The girl *whom* he had invited did not come.

We have also followed the recent trend in EST towards reducing the types of transformations to the single operation of 'movement': movement may be either to the left (for example, *wh*-movement, movement of the Object of passives, and raising) or to the right (as in extraposition). Furthermore, we assume that any moved category leaves behind an empty position. The movement of *who* in *Who did Peter invite---*?, for example, has left behind an empty Object position marked---. In current EST this approach is covered by Trace theory.

The various movement transformations referred to above yield S-structures. S-structure is assumed to be the level at which Case is assigned and at which coindexing takes place. Since Case assignment and coindexing are not discussed in this book, D-structure and S-structure will look identical for many sentences.

This is a book about English syntax in the narrow sense of that term, and the levels of Phonetic form and Logical form are referred to only occasionally. The level of Phonetic form, in EST, is that at which a number of additional morphological/phonological changes are performed. For example, the affixing of the ending of the finite verb is assumed to take place at the phonetic level of the grammar rather than at the syntactic level, i.e. after the level of S-structure. The operations performed on S-structures in the Phonetic component lead to surface structures. Although S-structures and surface structures, on the one hand, and D-structures and underlying structures, on the other, are not identical in EST, we shall make no rigid distinction between the two pairs, and mainly deal with surface structures and underlying structures. Logical form, too, is only introduced discursively, with respect to the 'binding' relations in sentences. The notion of binding plays an important part in the treatment of reflexive pronouns, which are usually bound to an antecedent in the same clause. This notion is also used with respect to the linking of relative and interrogative pronouns with the 'gap' at the vacated site. However, the treatment of such interdependencies at the level of Logical form is kept very informal and the standard EST treatment of such links in terms of coindexing is not introduced. By reducing abstract formalisations to a minimum, we focus on English sentences and their analysis rather than on the formal technicalities of the analysis.

In keeping with the basic EST approach of this coursebook, we shall not normally refer to semantics or pragmatics, unless information of this type might clarify the analysis. In this book we have devoted more space to grammatical functions than is customary in the EST tradition, adopting the configurational approach to grammatical functions, which sees notions such as 'Subject', 'Object', etc., not as primitives of syntax, but rather as fully configurationally determined, and thus derivative. This occasionally leads to analyses which deviate from the more traditional descriptive tradition, but we feel that

these modern analyses are more illuminating than traditional ones. With respect to the motivation of the analysis of constituent structure and function, we consistently rely on standard arguments of movement, substitution and analogy. Analogy has been used to motivate the presence of empty categories, and it may at least partly be seen as an informal translation of the Projection principle in EST, which states that distributional information encoded in the Lexicon (e.g. subcategorisation frames) must be preserved at each level of structure.

Herman Wekker
Liliane Haegeman
Nijmegen
Geneva

1

AIMS AND METHODS

1.1 WHAT IS SYNTAX?

Syntax, or **syntactic analysis**, may be defined as:

- (a) determining the relevant component parts of a sentence
- (b) describing these parts grammatically.

The component parts of a sentence are called **constituents**. In other words, syntax involves the two closely related tasks of:

- (a) breaking down the sentence into its **constituents**
- (b) assigning some grammatical label to each constituent, stating what type of constituent (or **grammatical category**) it is, and what **grammatical function** it has.

This definition of syntax implies that we start from what is regarded as the largest unit of syntactic description—the sentence—and proceed until we arrive at the smallest meaningful unit. This is called a ‘top to bottom’ analysis. The units smaller than the sentence will be referred to as **clauses**, **phrases**, **words** and **morphemes** respectively. However, instead of saying that a sentence can be broken down into smaller and smaller constituents, we might also look at the sentence the other way round—that is, ‘from bottom to top’—and say that constituents at different levels can combine to form increasingly larger units: we proceed then from the morpheme to the sentence as a whole. Constituents are like building blocks which pattern in certain ways to form larger and larger units, the largest unit being the sentence. Each constituent (except the smallest) can be broken down into its component parts. The purpose of doing syntax is to discover the ways in which constituents combine to form the **structure** of sentences.

In this book we adopt the (traditional) hierarchy of sentence constituents, as shown in the following diagram:

SENTENCE↔CLAUSE↔PHRASE↔WORD↔MORPHEME

This diagram represents the hierarchical scale of constituents. The four double-pointed arrows in the diagram indicate that it may be read ‘from left to right’, or ‘from right to left.’

The arrows pointing to the right indicate that a sentence may consist of one or more than one clause, that a clause may consist of one or more than one phrase, that a phrase may consist of one or more than one word, and that a word may consist of one or more than one morpheme. Morphemes are the minimal, indivisible units in syntax.

Conversely, as indicated by the arrows pointing to the left, we might also say that one or more than one morpheme may constitute a word, one or more than one word may form a phrase, one or more than one phrase may form a clause, and one or more than one clause may form a sentence.

In what follows we shall also see that a clause may contain one or more constituent clauses, and that a phrase may contain one or more constituent phrases or clauses.

To illustrate the hierarchical structure of sentences, let us consider sentence (1):

- (1) The snake killed the rat and swallowed it.

This sentence consists of two **coordinate** clauses, joined together by the **coordinator** *and*. The first clause is: *The snake killed the rat*, and the second is: *swallowed it*. The second clause has a reduced form. Its complete form would be *it swallowed it*.

The first clause in (1) consists of two phrases, and the second, as it stands, consists of only one phrase. The two phrases in the first clause are *The snake* and *killed the rat*, and the phrase in the second clause is *swallowed it*. As we shall see later, *the snake* is a **noun phrase** and *killed the rat* and *swallowed it* are both **verb phrases**.

Each phrase is made up of words. *The snake* consists of two words: *the* and *snake*; *killed the rat* consists of three words, of which the last two (*the* + *rat*) again constitute a noun phrase; and *swallowed it* consists of two words, of which the second in itself constitutes a noun phrase. The constituents *the rat* and *it* are examples of (noun) phrases within (verb) phrases.

Sentence (1) contains eight words, including the coordinator *and*. Each word consists of one or two morphemes: *the*, *snake*, *rat*, *and* and *it* are one-morpheme words, whereas *killed* and *swallowed* are both two-morpheme words. *The*, *snake*, *rat*, etc. are full words and morphemes at the same time: the word and morpheme boundaries coincide. The two morphemes of *killed* and *swallowed* are *kill* and—*ed*, and *swallow* and—*ed* (see 1.2.2).

1.2 REPRESENTING SENTENCE STRUCTURE

1.2.1 BRACKETING

The syntactic structure of sentence (1) above may be represented provisionally by marking off each constituent from sentence level to word level by square brackets: []. To simplify matters, we shall ignore the morpheme boundaries here. This convention of **bracketing** yields the following analysis, which looks rather daunting at first sight:

- (2) [[[[The] [snake]] [[killed] [[the] [rat]]]]]
 [and]
 [[[swallowed] [[it]]]]]

Analysis (2) is the result of first bracketing the sentence, then the two clauses, then the phrases, and finally the words, as follows:

Sentence:

- (3a) [The snake killed the rat and swallowed it]

Clauses:

- (3b) [[The snake killed the rat]
 and
 [swallowed it]]

Phrases:

- (3c) [[[The snake] [killed [the rat]]]
 and
 [[swallowed [it]]]]

Words:

- (3d) [[[[The] [snake]] [[killed] [[the] [rat]]]]
 [and]
 [[[swallowed] [[it]]]]]

Analysis (3d) is of course identical with (2) above.

The bracketing has here been done on a purely intuitive basis. In the following chapters we shall deal with the formal arguments which justify those choices. Check through the above analysis carefully again, and try to bracket the following sentences from sentence level to word level in the same way:

- (4) The terrorists assassinated the ambassador
(5) Her husband is an aristocrat.
(6) He gave his mother a present.

In (2) above we can see that word and morpheme boundaries may coincide: *the*, *snake*, *rat*, etc. are all one word and one morpheme, as opposed to *killed* and *swallowed*, which are words consisting of two morphemes each. Words and phrases may also coincide, as in:

- (7) John laughed.

In this sentence *John* is both a phrase (a noun phrase) and a word (a noun); *laughed* is also both a phrase (a verb phrase) and a word (a verb). The units sentence and clause also coincide in (7). Bracketing from sentence level to word level yields (8):

- (8) [[[[John]] [[laughed]]]]

Word, phrase, clause and sentence may also coincide, as in:

(9) Run!

The bracketing of sentence (9) is as follows:

(10) [[[Run]]]]

It is structurally one sentence, one clause, one phrase and one word (also one morpheme).

The above examples show that a sentence is not necessarily longer than a clause, a clause not necessarily longer than a phrase, and a phrase not necessarily longer than a word (in general, we shall not go beyond the level of the word in our analyses). We shall see that sentences may vary in length and complexity from one clause to indefinitely many clauses, clauses from one phrase to indefinitely many phrases, and phrases from one word to indefinitely many words.

In Chapter 2 we shall look at each of these grammatical units in more detail.

1.2.2 LABELLED BRACKETING

The system of bracketing which we have used so far is not very satisfactory. It is difficult to see, for example, which brackets go together to mark off a constituent. The notation introduced above could be improved by adding an appropriate grammatical label to each pair of square brackets. The label indicates what type of constituent (or grammatical category) is contained within the brackets. This convention is called **labelled bracketing**. To illustrate the new system, let us consider again the structure of the sentence *The snake killed the rat and swallowed it* ((1) above).

We shall use square brackets with the label S (short for ‘sentence’) to mark off the boundaries of the whole sentence, as follows:

(11) [_S **The snake killed the rat and swallowed it**]

The **category label** is inserted in the bottom corner of the left-hand bracket. This sentence, as we have seen, consists of two clauses: the clause *The snake killed the rat* and the clause *(it) swallowed it*. The two clauses are joined together by the coordinator *and*. The clauses can be bracketed and labelled as follows:

(12) [_{S₁} [_{S₂} **The snake killed the rat**] and [_{S₃} **(it)swallowed it**]]

The two clauses are labelled S₂ and S₃ here, and the whole sentence is labelled S₁. S₂ and S₃ are clauses inside S₁ (the numbers 1, 2, 3 etc. are added here and elsewhere just for convenience: they enable us to refer unambiguously to the different Ss). We use the label S for sentences as well as clauses, since clauses can be defined as Ss inside an S or inside a phrase. The clauses in (12) can be further analysed into phrases as follows:

- (13)
$$\left[S_1 \left[S_2 \left[{}_{NP} \text{The snake} \right] \left[{}_{VP} \text{killed} \left[{}_{NP} \text{the rat} \right] \right] \right] \right]$$

 and
$$\left[S_3 \left[{}_{NP} \text{(it)} \right] \left[{}_{VP} \text{swallowed} \left[{}_{NP} \text{it} \right] \right] \right]$$

The label NP stands for noun phrase, and the label VP for verb phrase. *The snake* and *the rat* are NPs, because their most important element is a noun (N); *it*, which replaces an NP, is also an NP (cf. 1.2.4 and 2.4.2.2). The VPs in (13) consist of a verb (V) followed by an NP. The verbs are *killed* and *swallowed*. The NPs *the rat* in *killed the rat* and *it* in *swallowed it* are parts of the VPs, and act as **Complements** of the verbs *killed* and *swallowed* respectively.

The words that make up the above phrases can also be bracketed and labelled as follows:

- (14)
$$\left[S_1 \left[S_2 \left[{}_{NP} \left[{}_{Det} \text{The} \right] \left[{}_N \text{snake} \right] \right] \left[{}_{VP} \left[{}_V \text{killed} \right] \left[{}_{NP} \left[{}_{Det} \text{the} \right] \left[{}_N \text{rat} \right] \right] \right] \right] \right]$$

 [and]

$$\left[S_3 \left[{}_{NP} \text{(it)} \right] \left[{}_{VP} \left[{}_V \text{swallowed} \right] \left[{}_{NP} \left[{}_N \text{it} \right] \right] \right] \right]$$

The label Det stands for **determiner**, the label N for noun, and the label V for verb. Noun and verb are major **word classes**, and Det is a collective term for various items preceding the noun, e.g.: *the*, *a*, *that*, *this*, *some*, *any*. Now compare (14) with (2) in section 1.2.1 above. The only difference between the two is that pairs of brackets are labelled here.

Our syntactic analysis of the sentence does not usually go below the level of the word or even the phrase, but occasionally it will be useful to mark off the morpheme structure of a given word, for example the structure of the past tense forms of verbs, or of the plurals and genitives of nouns. This, too, can be done by means of labelled bracketing, as, for example, in the case of *killed*:

- (16)
$$\left[{}_V \left[{}_B \text{kill} \right] \left[{}_{Suff} \text{ed} \right] \right]$$

Here the label B is used for the **base** (of the verb), and the label Suff for **suffix** (see section 1.1).

Throughout this book we shall frequently use this kind of labelled bracketing to represent sentence structure.

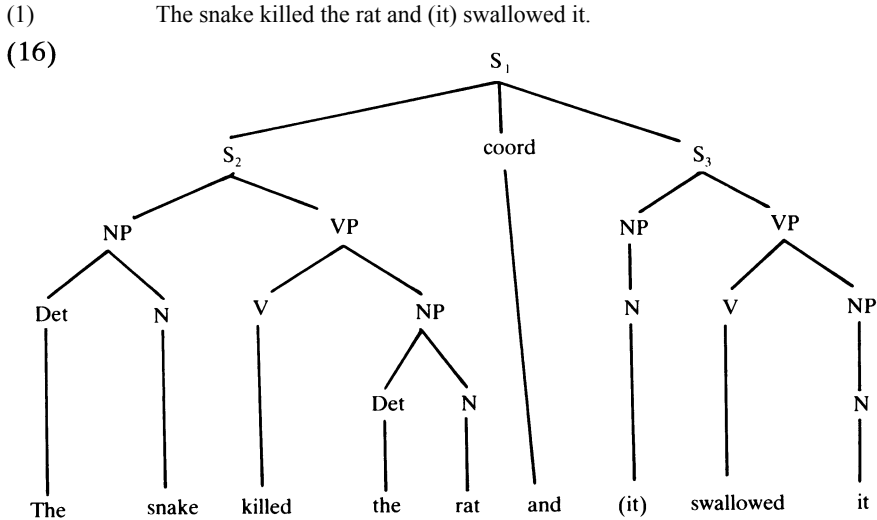
1.2.3 TREE DIAGRAMS AND PHRASE STRUCTURE RULES

Labelled bracketing is one of the most common ways of representing the constituent structure of sentences. However, there are many other methods of marking

diagrammatically what elements in a sentence go together and what elements do not. One other very common representation is the **tree diagram**.

The tree diagram is a notational device which is entirely equivalent to labelled bracketing: although it looks different, it provides the same information about the syntactic structure of a sentence.

Consider, for example, the following (simplified) tree diagram of the sentence:



The tree diagram provides the analysis of sentence (1) down to word level (determiner, noun, verb, etc.); in principle, it would also be possible to stop at phrase level (NP, VP, etc.), or to go beyond word level and indicate the morphological structure of each of the words. You can make your grammatical analysis as detailed as you like, or as is necessary for a specific purpose. The syntactic information contained in diagram (16) is essentially the same as that provided by the labelled bracketing in (14) in section 1.2.2. It is largely a matter of taste or practical convenience whether one chooses the notation of (14) in section 1.2.2 or that of (16) above.

To read tree diagram (16) we require some additional terminology. For example, we say that S_1 (the whole sentence) is **expanded** as S_2 and S_3 (two clauses), which are coordinated by *and*. S_1 is said to contain as its **immediate constituents** S_2 , S_3 and the coordinator *and*. S_2 is expanded as NP-VP, and so is S_3 . The VP in each case is expanded as V-NP. The constituents mentioned so far are considered to be the main constituents of the sentence. Further down the tree, NP may again be expanded as either Det-N (*the—snake*) or as N (*it*).

There is a convention which is generally used to sum up the system of expanding one unit into other units. It is a set of instructions called **phrase structure rules** (PS rules).

Let us provisionally formulate the following four PS rules:

(17a) $S \longrightarrow \bar{S} - \text{coord} - S$

- (17b) $S \longrightarrow NP-VP$
 (17c) $VP \longrightarrow V-NP$
 (17d) $NP \longrightarrow \left\{ \begin{array}{c} \text{Det-N} \\ N \end{array} \right\}$

The arrows mean: ‘expand’ or ‘rewrite’ X as Y (where X represents any element on the left of the arrow, and Y any element on the right). The curly brackets in rule (17d) indicate that NP may be expanded or rewritten as *either* Det-N *or* N.

The category labels in tree diagram (16), such as S_1 , S_2 , S_3 , NP, VP, V, etc., are all attached to the **nodes** of that tree, and the lines connecting these nodes are called **branches**. The node labels in the tree diagram correspond to the labels in the labelled bracketing in 1.2.2. The node labelled S_1 in (16) **dominates** the nodes labelled S_2 , coord, and S_3 . The node labelled S_2 dominates the nodes labelled NP and VP, but also all the other nodes further down that half of the tree. The same applies to S_3 . NP, VP, etc. are said to be dominated by S_2 or S_3 . We also see that S_2 and S_3 both contain two NPs. One NP is **immediately dominated** by S_2/S_3 , the other by the VP-node. The NP immediately dominated by the VP-node is also dominated by S_2/S_3 , but it is not immediately dominated by S_2/S_3 . It is important to distinguish between **dominance** and **immediate dominance**. In the latter case there must be no further nodes intervening between the nodes considered.

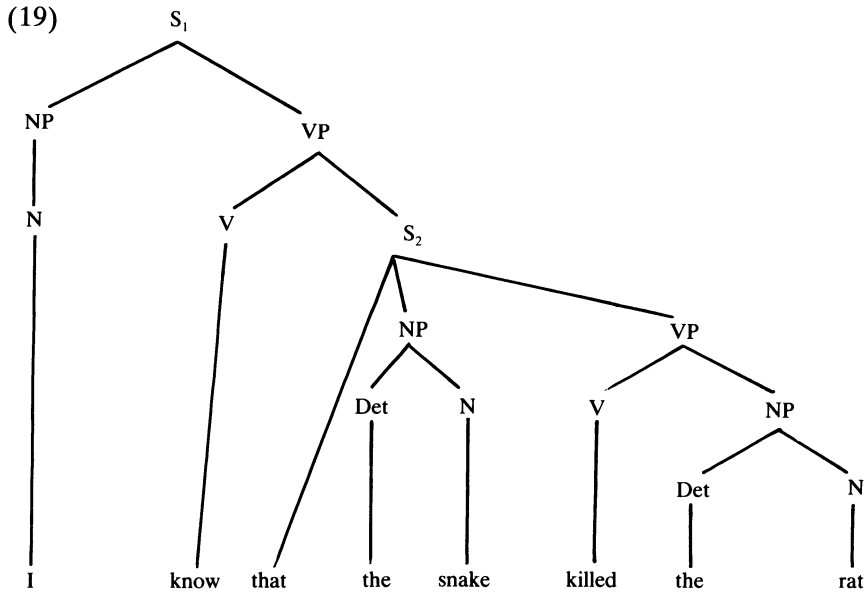
The **lexical items** (i.e. the words) *the*, *snake*, *killed*, etc., are attached to the so-called **terminal** nodes of the tree diagram, i.e. the bottom nodes. The other nodes in the tree diagram are **non-terminal**.

Tree diagram (16) shows, among other things, that *the* (Det) and *snake* (N) combine to form one constituent (an NP), that *killed* (V) and *the rat* (NP) constitute a VP, that the NP *the snake* and the VP *killed the rat* are a sentence/clause (S_2) inside another sentence (S_1), etc. In the same way the right half of the tree diagram also specifies the internal structure of clause S_3 .

Subordination of a clause, as in:

- (18) I know $\left[{}_S \text{that the snake killed the rat} \right]$

may be represented as follows:



Tree diagram (19) is provisional. We shall look in greater detail at how we should treat the sequence *that-the snake killed the rat* later (see section 2.3).

The provisional set of PS rules in (17) cannot fully describe structures like (19). For example, rule (17a) is optional: it need not apply (after all, not all sentences contain coordinate clauses). We start the rewriting operations for (19) at rule (17b). Also, instead of rewriting VP as V-NP, as stipulated by rule (17c), we must rewrite the VP in (19) as V-S. Consequently, our PS rule (17c) must allow for this possibility: we must adapt it in the following way (the curly brackets again indicate a choice: ‘either... or’):

(17c)
$$VP \longrightarrow V - \left\{ \begin{array}{c} NP \\ S \end{array} \right\}$$

The S, which now also occurs on the right-hand side of the arrow, may in its turn be expanded as NP-VP, as indicated in rule (17b), and this VP may now become either V-NP or V-S, and so on. It is possible to have an S embedded inside another S. This kind of embedding (or subordination) may, in principle, be repeated indefinitely many times: there may be indefinitely many Ss embedded inside other Ss. For example:

(20) I know that you think that she hopes that you will say to her that you love her

Draw a tree diagram representing the syntactic structure of (20).

In diagram (19) S_2 is a subordinate clause, which functions inside S_1 : S_2 is dominated by S_1 , and immediately dominated by VP.

Ss (clauses) may also appear inside NPs as in: