

Textbooks in Language Sciences

Editors: Stefan Müller, Martin Haspelmath

Editorial Board: Claude Hagège, Marianne Mithun, Anatol Stefanowitsch, Foong Ha Yap

In this series:

1. Müller, Stefan. Grammatical theory: From transformational grammar to constraint-based approaches.
2. Schäfer, Roland. Einführung in die grammatische Beschreibung des Deutschen.
3. Freitas, Maria João & Ana Lúcia Santos (eds.). Aquisição de língua materna e não materna: Questões gerais e dados do português.
4. Roussarie, Laurent. Sémantique formelle: Introduction à la grammaire de Montague.
5. Kroeger, Paul. Analyzing meaning: An introduction to semantics and pragmatics.
6. Ferreira, Marcelo. Curso de semântica formal.
7. Stefanowitsch, Anatol. Corpus linguistics: A guide to the methodology.


Stefan Müller. 2020. *Grammatical theory: From transformational grammar to constraint-based approaches. Fourth revised and extended edition.* (Textbooks in Language Sciences 1). Berlin: Language Science Press.

This title can be downloaded at:

<http://langsci-press.org/catalog/book/255>

© 2020, Stefan Müller

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):

<http://creativecommons.org/licenses/by/4.0/> 

ISBN: 978-3-96110-202-0 (Digital)

978-3-96110-220-4 (Hardcover)

978-3-96110-221-1 (Softcover)

ISSN: 2364-6209

DOI: 10.5281/zenodo.3364215

Source code available from www.github.com/langsci/25

Collaborative reading: paperhive.org/documents/remote?type=langsci&id=25

Cover and concept of design: Ulrike Harbort

Translator: Andrew Murphy, Stefan Müller

Typesetting: Stefan Müller

Proofreading: Viola Auermann, Armin Buch, Andreea Calude, Rong Chen, Matthew Czuba, Leonel de Alencar, Christian Döhler, Joseph T. Farquharson, Andreas Hölzl, Gianina Iordăchioaia, Paul Kay, Anne Kilgus, Sandra Kübler, Timm Lichte, Antonio Machicao y Priemer, Michelle Natolo, Stephanie Natolo, Sebastian Nordhoff, Elizabeth Pankratz, Parviz Parsafar, Conor Pyle, Daniela Schröder, Eva Schultze-Berndt, Alec Shaw, Benedikt Singpiel, Anelia Stefanova, Neal Whitman, Viola Wiegand

Open reviewing: Armin Buch, Leonel de Alencar, Andreas Hölzl, Dick Hudson, Gianina Iordăchioaia, Paul Kay, Timm Lichte, Antonio Machicao y Priemer, Andrew McIntyre, Arne Nymos, Sebastian Nordhoff, Neal Whitman

Fonts: Libertinus, Arimo, DejaVu Sans Mono

Typesetting software: Xe_{La}TeX

Language Science Press

Xhain

Grünberger Str. 16

10243 Berlin, Germany

langsci-press.org

Storage and cataloguing done by FU Berlin

For Max

Contents

Preface

This book is an extended and revised version of my German book *Grammatiktheorie* (MuellerGTBuch2). It introduces various grammatical theories that play a role in current theorizing or have made contributions in the past which are still relevant today. I explain some foundational assumptions and then apply the respective theories to what can be called the “core grammar” of German. I have decided to stick to the object language that I used in the German version of this book since many of the phenomena that will be dealt with cannot be explained with English as the object language. Furthermore, many theories have been developed by researchers with English as their native language and it is illuminative to see these theories applied to another language. I show how the theories under consideration deal with arguments and adjuncts, active/passive alternations, local reorderings (so-called scrambling), verb position, and fronting of phrases over larger distances (the verb second property of the Germanic languages without English).

The second part deals with foundational questions that are important for developing theories. This includes a discussion of the question of whether we have innate domain specific knowledge of language (UG), the discussion of psycholinguistic evidence concerning the processing of language by humans, a discussion of the status of empty elements and of the question whether we construct and perceive utterances holistically or rather compositionally, that is, whether we use phrasal or lexical constructions. The second part is not intended as a standalone book although the printed version of the book is distributed this way for technical reasons (see below). Rather it contains topics that are discussed again and again when frameworks are compared. So instead of attaching these discussions to the individual chapters they are organized in a separate part of the book.

Unfortunately, linguistics is a scientific field with a considerable amount of terminological chaos. I therefore wrote an introductory chapter that introduces terminology in the way it is used later on in the book. The second chapter introduces phrase structure grammars, which plays a role for many of the theories that are covered in this book. I use these two chapters (excluding the Section 2.3 on interleaving phrase structure grammars and semantics) in introductory courses of our

BA curriculum for German studies. Advanced readers may skip these introductory chapters. The following chapters are structured in a way that should make it possible to understand the introduction of the theories without any prior knowledge. The sections regarding new developments and classification are more ambitious: they refer to chapters still to come and also point to other publications that are relevant in the current theoretical discussion but cannot be repeated or summarized in this book. These parts of the book address advanced students and researchers. I use this book for teaching the syntactic aspects of the theories in a seminar for advanced students in our BA. The slides are available on my web page. The second part of the book, the general discussion, is more ambitious and contains the discussion of advanced topics and current research literature.

This book only deals with relatively recent developments. For a historical overview, see for instance **Robins97a-u**; **JL2006a-u**. I am aware of the fact that chapters on Integrational Linguistics (**Lieb83a-u**; **Eisenberg2004a**; **Nolda2007a-u**), Optimality Theory (**PS93a-u**; **Grimshaw97a-u**; **G. GMueller2000a-u**), Role and Reference Grammar (**vanValin93a-ed**) and Relational Grammar (**Perlmutter83a-ed**; **Perlmutter84b-ed**) are missing. I will leave these theories for later editions.

The original German book was planned to have 400 pages, but it finally was much bigger: the first German edition has 525 pages and the second German edition has 564 pages. I added a chapter on Dependency Grammar and one on Minimalism to the English version and now the book has ?? pages. I tried to represent the chosen theories appropriately and to cite all important work. Although the list of references is over 85 pages long, I was probably not successful. I apologize for this and any other shortcomings.

Acknowledgments

I would like to thank David Adger, Jason Baldridge, Felix Bildhauer, Emily M. Bender, Stefan Evert, Gisbert Fanselow, Sandiway Fong, Hans”=Martin Gärtner, Kim Gerdes, Adele Goldberg, Bob Levine, Paul Kay, Jakob Maché, Guido Mensching, Laura Michaelis, Geoffrey Pullum, Uli Sauerland, Roland Schäfer, Jan Strunk, Remi van Trijp, Shravan Vasishth, Tom Wasow, and Stephen Wechsler for discussion and Monika Budde, Philippa Cook, Laura Kallmeyer, Tibor Kiss, Gisela Klann”=Delius, Jonas Kuhn, Timm Lichte, Anke Lüdeling, Jens Michaelis, Bjarne Ørsnes, Andreas Pankau, Christian Pietsch, Frank Richter, Ivan Sag, and Eva Wittenberg for comments on earlier versions of the German edition of this book and Thomas Groß, Dick Hudson, Sylvain Kahane, Paul Kay, Haitao Liu (☒ ☒), Andrew McIntyre, Sebastian Nordhoff, Tim Osborne, Andreas Pankau, and

Christoph Schwarze for comments on earlier versions of this book. Thanks to Leonardo Boiko and Sven Verdoolaege for pointing out typos. Special thanks go to Martin Haspelmath for very detailed comments on an earlier version of the English book.

This book was the first Language Science Press book that had an open review phase (see below). I thank Dick Hudson, Paul Kay, Antonio Machicao y Priemer, Andrew McIntyre, Sebastian Nordhoff, and one anonymous open reviewer for their comments. These comments are documented at the download page of this book. In addition the book went through a stage of community proofreading (see also below). Some of the proofreaders did much more than proofreading, their comments are highly appreciated and I decided to publish these comments as additional open reviews. Armin Buch, Leonel de Alencar, Andreas Hölzl, Gianina Iordăchioaia, Timm Lichte, Antonio Machicao y Priemer, and Neal Whitman deserve special mention here.

I thank Wolfgang Sternefeld and Frank Richter, who wrote a detailed review of the German version of this book (SR2012a). They pointed out some mistakes and omissions that were corrected in the second edition of the German book and which are of course not present in the English version.

Thanks to all the students who commented on the book and whose questions lead to improvements. Lisa Deringer, Aleksandra Gabryszak, Simon Lohmiller, Theresa Kallenbach, Steffen Neuschulz, Reka Meszaros-Segner, Lena Terhart and Elodie Winckel deserve special mention.

Since this book is built upon all my experience in the area of grammatical theory, I want to thank all those with whom I ever discussed linguistics during and after talks at conferences, workshops, summer schools or via email. Werner Abraham, John Bateman, Dorothee Beermann, Rens Bod, Miriam Butt, Manfred Bierwisch, Ann Copestake, Holger Diessel, Kerstin Fischer, Dan Flickinger, Peter Gallmann, Petter Haugereid, Lars Hellan, Tibor Kiss, Wolfgang Klein, Hans=Ulrich Krieger, Andrew McIntyre, Detmar Meurers, Gereon Müller, Martin Neef, Manfred Sailer, Anatol Stefanowitsch, Peter Svenonius, Michael Tomasello, Hans Uszkoreit, Gert Webelhuth, Daniel Wiechmann and Arne Zeschel deserve special mention.

I thank Sebastian Nordhoff for a comment regarding the completion of the subject index entry for *recursion*.

Andrew Murphy translated part of Chapter 1 and the Chapters 2–3, 5–10, and 12–23. Many thanks for this!

I also want to thank the 27 community proofreaders (Viola Auermann, Armin Buch, Andreea Calude, Rong Chen, Matthew Czuba, Leonel de Alencar, Christian Döhler, Joseph T. Farquharson, Andreas Hölzl, Gianina Iordăchioaia, Paul Kay,

Preface

Anne Kilgus, Sandra Kübler, Timm Lichte, Antonio Machicao y Priemer, Michelle Natolo, Stephanie Natolo, Sebastian Nordhoff, Elizabeth Pankratz, Parviz Parsa-far, Conor Pyle, Daniela Schröder, Eva Schultze-Berndt, Alec Shaw, Benedikt Singpiel, Anelia Stefanova, Neal Whitman, Viola Wiegand) that each worked on one or more chapters and really improved this book. I got more comments from every one of them than I ever got for a book done with a commercial publisher. Some comments were on content rather than on typos and layout issues. No proofreader employed by a commercial publisher would have spotted these mistakes and inconsistencies since commercial publishers do not have staff that knows all the grammatical theories that are covered in this book.

During the past years, a number of workshops on theory comparison have taken place. I was invited to three of them. I thank Helge Dyvik and Torbjørn Nordgård for inviting me to the fall school for Norwegian PhD students *Languages and Theories in Contrast*, which took place 2005 in Bergen. Guido Mensching and Elisabeth Stark invited me to the workshop *Comparing Languages and Comparing Theories: Generative Grammar and Construction Grammar*, which took place in 2007 at the Freie Universität Berlin and Andreas Pankau invited me to the workshop *Comparing Frameworks* in 2009 in Utrecht. I really enjoyed the discussion with all participants of these events and this book benefited enormously from the interchange.

I thank Peter Gallmann for the discussion of his lecture notes on GB during my time in Jena. The Sections 3.1.3–3.4 have a structure that is similar to the one of his script and take over a lot. Thanks to David Reitter for the \LaTeX macros for Combinatorial Categorical Grammar, to Mary Dalrymple and Jonas Kuhn for the LFG macros and example structures, and to Laura Kallmeyer for the \LaTeX sources of most of the TAG analyses. Most of the trees have been adapted to the forest package because of compatibility issues with \XeLaTeX , but the original trees and texts were a great source of inspiration and without them the figures in the respective chapters would not be half as pretty as they are now.

I thank Sašo Živanović for implementing the \LaTeX package forest. It really simplifies typesetting of trees, dependency graphs, and type hierarchies. I also thank him for individual help via email and on stackexchange. In general, those active on stackexchange could not be thanked enough: most of my questions regarding specific details of the typesetting of this book or the implementation of the \LaTeX classes that are used by Language Science Press now have been answered within several minutes. Thank you! Since this book is a true open access book under the CC-BY license, it can also be an open source book. The interested reader finds a copy of the source code at <https://github.com/langsci/25>. By

making the book open source I pass on the knowledge provided by the \LaTeX gurus and hope that others benefit from this and learn to typeset their linguistics papers in nicer and/or more efficient ways.

Viola Auermann and Antje Bahlke, Sarah Dietzfelbinger, Lea Helmers, and Chiara Jancke cannot be thanked enough for their work at the copy machines. Viola also helped a lot with proof reading prefinal stages of the translation. I also want to thank my (former) lab members Felix Bildhauer, Philippa Cook, Janna Lipenkova, Jakob Maché, Bjarne Ørsnes and Roland Schäfer, which were mentioned above already for other reasons, for their help with teaching. During the years from 2007 until the publication of the first German edition of this book two of the three tenured positions in German Linguistics were unfilled and I would have not been able to maintain the teaching requirements without their help and would have never finished the *Grammatiktheorie* book.

I thank Tibor Kiss for advice in questions of style. His diplomatic way always was a shining example for me and I hope that this is also reflected in this book.

On the way this book is published

I started to work on my dissertation in 1994 and defended it in 1997. During the whole time the manuscript was available on my web page. After the defense, I had to look for a publisher. I was quite happy to be accepted to the series *Linguistische Arbeiten* by Niemeyer, but at the same time I was shocked about the price, which was 186.00 DM for a paperback book that was written and typeset by me without any help by the publisher (twenty times the price of a paperback novel).¹ This basically meant that my book was depublished: until 1998 it was available from my web page and after this it was available in libraries only. My Habilitationsschrift was published by CSLI Publications for a much more reasonable price. When I started writing textbooks, I was looking for alternative distribution channels and started to negotiate with no-name print on demand publishers. Brigitte Narr, who runs the Stauffenburg publishing house, convinced me to publish my HPSG textbook with her. The copyrights for the German version of the book remained with me so that I could publish it on my web page. The collaboration was successful so that I also published my second textbook about grammatical theory with Stauffenburg. I think that this book has a broader relevance and should be accessible for non-German-speaking readers as well. I therefore decided to have it translated into English. Since Stauffenburg is focused on books in German, I

¹As a side remark: in the meantime Niemeyer was bought by de Gruyter and closed down. The price of the book is now 139.95 €/ \$ 196.00. The price in Euro corresponds to 273.72 DM.

had to look for another publisher. Fortunately the situation in the publishing sector changed quite dramatically in comparison to 1997: we now have high profile publishers with strict peer review that are entirely open access. I am very glad about the fact that Brigitte Narr sold the rights of my book back to me and that I can now publish the English version with Language Science Press under a CC-BY license.

Language Science Press: scholar-owned high quality linguistic books

In 2012 a group of people found the situation in the publishing business so unbearable that they agreed that it would be worthwhile to start a bigger initiative for publishing linguistics books in platinum open access, that is, free for both readers and authors. I set up a web page and collected supporters, very prominent linguists from all over the world and all subdisciplines and Martin Haspelmath and I then founded Language Science Press. At about the same time the DFG had announced a program for open access monographs and we applied (MH2013a) and got funded (two out of 18 applications got funding). The money is used for a coordinator (Dr. Sebastian Nordhoff) and an economist (Debora Siller), two programmers (Carola Fanselow and Dr. Mathias Schenner), who work on the publishing platform Open Monograph Press (OMP) and on conversion software that produces various formats (ePub, XML, HTML) from our \LaTeX code. Svantje Lilienthal works on the documentation of OMP, produces screencasts and does user support for authors, readers and series editors.

OMP is extended by open review facilities and community-building gamification tools (MuellerOA; MH2013a). All Language Science Press books are reviewed by at least two external reviewers. Reviewers and authors may agree to publish these reviews and thereby make the whole process more transparent (see also Pullum84a for the suggestion of open reviewing of journal articles). In addition there is an optional second review phase: the open review. This review is completely open to everybody. The whole community may comment on the document that is published by Language Science Press. After this second review phase, which usually lasts for two months, authors may revise their publication and an improved version will be published. This book was the first book to go through this open review phase. The annotated open review version of this book is still available via the web page of this book.

Currently, Language Science Press has 17 series on various subfields of linguistics with high profile series editors from all continents. We have 18 published and

17 forthcoming books and 146 expressions of interest. Series editors and authors are responsible for delivering manuscripts that are typeset in L^AT_EX, but they are supported by a web-based typesetting infrastructure that was set up by Language Science Press and by volunteer typesetters from the community. Proofreading is also community-based. Until now 53 people helped improving our books. Their work is documented in the Hall of Fame: <http://langsci-press.org/hallOfFame>.

If you think that textbooks like this one should be freely available to whoever wants to read them and that publishing scientific results should not be left to profit-oriented publishers, then you can join the Language Science Press community and support us in various ways: you can register with Language Science Press and have your name listed on our supporter page with almost 600 other enthusiasts, you may devote your time and help with proofreading and/or typesetting, or you may donate money for specific books or for Language Science Press in general. We are also looking for institutional supporters like foundations, societies, linguistics departments or university libraries. Detailed information on how to support us is provided at the following webpage: <http://langsci-press.org/supportUs>. In case of questions, please contact me or the Language Science Press coordinator at contact@langsci-press.org.

Berlin, March 11, 2016

Stefan Müller

Foreword of the second edition

The first edition of this book was published almost exactly two years ago. The book has app. 15,000 downloads and is used for teaching and in research all over the world. This is what every author and every teacher dreams of: distribution of knowledge and accessibility for everybody. The foreword of the first edition ends with a description of Language Science Press in 2016. This is the situation now:² We have 324 expressions of interest and 58 published books. Books are published in 20 book series with 263 members of editorial boards from 44 different countries from six continents. We have a total of 175,000 downloads. 138 linguists from all over the world have participated in proofreading. There are currently 296 proofreaders registered with Language Science Press. Language Science Press is a community-based publisher, but there is one person who manages everything: Sebastian Nordhoff. His position has to be paid. We were successful in acquir-

²See <http://userblogs.fu-berlin.de/langsci-press/2018/01/18/achievements-2017/> for the details and graphics.

ing financial support by almost 100 academic institutions including Harvard, the MIT, and Berkeley.³ If you want to support us by just signing the list of supporters, by publishing with us, by helping as proofreader or by convincing your librarian/institution to support Language Science Press financially, please refer to <http://langsci-press.org/supportUs>.

After these more general remarks concerning Language Science Press I describe the changes I made for the second edition and I thank those who pointed out mistakes and provided feedback.

I want to thank Wang Lulu for pointing out some typos that she found while translating the book to Chinese. Thanks for both the typos and the translation.

Fritz Hamm noticed that the definition of Intervention (see p. 148) was incomplete and pointed out some inconsistencies in translations of predicates in Section 2.3. I turned some straight lines in Chapter 3 into triangles and added a discussion of different ways to represent movement (see Figure 3.8 on p. 110). I now explain what SpecIP stands for and I added footnote 8 on SpecIP as label in trees. I extended the discussion of Pirahã in Section ?? and added lexical items that show that Pirahã-like modification without recursion can be captured in a straightforward way in Categorical Grammar.

I reorganized the HPSG chapter to be in line with more recent approaches assuming the valence features *SPR* and *COMPS* (Sag97a; MuellerGermanic) rather than a single valence feature. I removed the section on the *LOCAL* feature in Sign-based Construction Grammar (Section 10.6.2.2 in the first edition) since it was build on the wrong assumption that the filler would be identical to the representation in the valence specification. In Sag2012a only the information in *SYN* and *SEM* is shared.

I added the example (??) on page ?? that shows a difference in choice of preposition in a prepositional object in Dutch vs. German. Since the publication of the first English edition of the Grammatical Theory textbook I worked extensively on the phrasal approach to benefactive constructions in LFG (AGT2014a). Section ?? was revised and adapted to what will be published as MuellerLFGphrasal. There is now a brief chapter on complex predicates in TAG and Categorical Grammar/HPSG (Chapter ??), that shows that valence"-based approaches allow for an underspecification of structure. Valence is potential structure, while theories like TAG operate with actual structure.

Apart from this I fixed several minor typos, added and updated some references and URLs. Thanks to Philippa Cook, Timm Lichte, and Antonio Machicao y

³A full list of supporting institutions is available here: <http://langsci-press.org/knowledgeunlatched>.

Priemer for pointing out typos. Thanks to Leonel Figueiredo de Alencar, Francis Bond, John Carroll, Alexander Koller, Emily M. Bender, and Glenn C. Slayden for pointers to literature. Sašo Živanović helped adapting version 2.0 of the forest package so that it could be used with this large book. I am very grateful for this nice tree typesetting package and all the work that went into it.

The source code of the book and the version history is available on GitHub. Issues can be reported there: <https://github.com/langsci/25>. The book is also available on paperhive, a platform for collective reading and annotation: <https://paperhive.org/documents/remote?type=langsci&id=25>. It would be great if you would leave comments there.

Berlin, 21st March 2018

Stefan Müller

Foreword of the third edition

Since more and more researchers and students are using the book now, I get feedback that helps improve it. For the third edition I added references, expanded the discussion of the passive in GB (Section 3.4) a bit and fixed typos.⁴

Chapter 4 contained figures from different chapters of **Adger2003a**. Adger introduces the DP rather late in the book and I had a mix of NPs and DPs in figures. I fixed this in the new edition. I am so used to talking about NPs that there were references to NP in the general discussion that should have been references to DP. I fixed this as well. I added a figure explaining the architecture in the Phase model of Minimalism and since the figures mention the concept of *numeration*, I added a footnote on numerations. I also added a figure depicting the architecture assumed in Minimalist theories with Phases (right figure in Figure 4.1).

I thank Frank Van Eynde for pointing out eight typos in his review of the first edition. They have been fixed. He also pointed out that the placement of ARG-ST in the feature geometry of signs in HPSG did not correspond to **GSag2000a-u**, where ARG-ST is on the top level rather than under CAT. Note that earlier versions of this book had ARG-ST under CAT and there had never been proper arguments for why it should not be there, which is why many practitioners of HPSG have kept it in that position (**MuellerLFGphrasal**). One reason to keep ARG-ST on the top level is that ARG-ST is appropriate for lexemes only. If ARG-ST is on the sign level, this can be represented in the type hierarchy: lexemes and word have an

⁴A detailed list of issues and fixes can be found in the GitHub repository of this book at <https://github.com/langsci/25/>.

ARG-ST feature, phrases do not. If ARG-ST is on the CAT level, one would have to distinguish between CAT values that belong to lexemes and words on the one hand and phrasal CAT values on the other hand, which would require two additional subtypes of the type *cat*. The most recent version of the computer implementation done in Stanford by Dan Flickinger has ARG-ST under LOCAL (2019-01-24). So, I was tempted to leave everything as it was in the second edition of the book. However, there is a real argument for not having ARG-ST under CAT. CAT is assumed to be shared in coordinations and CAT contains valence features for subjects and complements. The values of these valence features are determined by a mapping from ARG-ST. In some analyses, extracted elements are not mapped to the valence features and the same is sometimes assumed for omitted elements. To take an example consider (1):

- (1) He saw and helped the hikers.

saw and *helped* are coordinated and the members in the valence lists have to be compatible. Now if one coordinates a ditransitive verb with one omitted argument with a strictly transitive verb, this would work under the assumption that the omitted argument is not part of the valence representation. But if ARG-ST is part of CAT, coordination would be made impossible since a three-place argument structure list would be incompatible with a two-place list. Hence I decided to change this in the third edition and represent ARG-ST outside of CAT from now on.

I changed the section about Sign-Based Construction Grammar (SBCG) again. An argument about nonlocal dependencies and locality was not correct, since **Sag2012a** does not share all information between filler and extraction side. The argument is now revised and presented as Section ?? . Reviewing **MuellerCxG**, Bob Borsley pointed out to me that the XARG feature is a way to circumvent locality restrictions that is actually used in SBCG. I added a footnote to the section on locality in SBCG.

A brief discussion of **Welke2019a-u**'s (**Welke2019a-u**) analysis of the German clause structure was added to the chapter about Construction Grammar (see Section ??).

The analysis of a verb-second sentence in LFG is now part of the LFG chapter (Figure ?? on page ??) and not just an exercise in the appendix. A new exercise was designed instead of the old one and the old one was integrated into the main text.

I added a brief discussion of **Osborne2019a**'s (**Osborne2019a**) claim that Dependency Grammars are simpler than phrase structure grammars (p. ??).

Geoffrey Pullum pointed out at the HPSG conference in 2019 that the label *constraint* “=based may not be the best for the theories that are usually referred to with it. Changing the term in this work would require to change the title of the book. The label *model theoretic* may be more appropriate but some implementational work in HPSG and LFG not considering models may find the term inappropriate. I hence decided to stick to the established term.

I followed the advice by Lisbeth Augustinus and added a preface to Part II of the book that gives the reader some orientation as to what to expect.

I thank Mikhail Knyazev for pointing out to me that the treatment of V to I to C movement in the German literature differs from the lowering that is assumed for English and that some further references are needed in the chapter on Government & Binding.

Working on the Chinese translation of this book, Wang Lulu pointed out some typos and a wrong example sentence in Chinese. Thanks for these comments!

I thank Bob Borsley, Gisbert Fanselow, Hubert Haider and Pavel Logacev for discussion and Ina Baier for a mistake in a CG proof and Jonas Benn for pointing out some typos to me. Thanks to Tabea Reiner for a comment on gradedness. Thanks also to Antonio Machicao y Priemer for yet another set of comments on the second edition and to Elizabeth Pankratz for proofreading parts of what I changed.

Berlin, 15th August 2019

Stefan Müller

Foreword of the fourth edition

I fixed several typos, added and updated URLs and DOIs in the book and in the list of references. I added a footnote to Chapter 3 concerning the assignment of semantic roles across phrase boundaries (footnote 20 on p. 121). I thank Andreas Pankau for discussion on this point.

I added a paragraph discussing John Torr’s implementational work (pages 190–192). I thank Shalom Lappin and Richard Sproat for discussion of implementation issues.

Language Science Press will publish a handbook on Head-Driven Phrase Structure Grammar hopefully later this year (**HPSGHandbook**). It contains several chapters comparing other syntactic theories to HPSG. I added the respective references to the further readings sections of the chapters for Lexical Functional Grammar, Categorical Grammar, Construction Grammar, and Minimalism.

Preface

This edition is the first edition that uses precompiled trees. Setting this up was not straight-forward. I am really grateful to Sašo Živanović for helping me and adapting the forest package so that everything runs smoothly and efficiently. This saves me a lot of time and reduces the energy consumption of my computer dramatically.

Berlin, May 6, 2020

Stefan Müller

Part I

Background and specific theories

1 Introduction and basic terms

The aim of this chapter is to explain why we actually study syntax (Section 1.1) and why it is important to formalize our findings (Section 1.2). Some basic terminology will be introduced in Sections 1.3–1.8: Section 1.3 deals with criteria for dividing up utterances into smaller units. Section 1.4 shows how words can be grouped into classes; that is I will introduce criteria for assigning words to categories such as verb or adjective. Section 1.5 introduces the notion of heads, in Section 1.6 the distinction between arguments and adjuncts is explained, Section 1.7 defines grammatical functions and Section 1.8 introduces the notion of topological fields, which can be used to characterize certain areas of the clause in languages such as German.

Unfortunately, linguistics is a scientific field with a considerable amount of terminological chaos. This is partly due to the fact that terminology originally defined for certain languages (e.g., Latin, English) was later simply adopted for the description of other languages as well. However, this is not always appropriate since languages differ from one another considerably and are constantly changing. Due to the problems caused by this, the terminology started to be used differently or new terms were invented. When new terms are introduced in this book, I will always mention related terminology or differing uses of each term so that readers can relate this to other literature.

1.1 Why do syntax?

Every linguistic expression we utter has a meaning. We are therefore dealing with what has been referred to as form-meaning pairs (Saussure16a). A word such as *tree* in its specific orthographical form or in its corresponding phonetic form is assigned the meaning *tree'*. Larger linguistic units can be built up out of smaller ones: words can be joined together to form phrases and these in turn can form sentences.

The question which now arises is the following: do we need a formal system which can assign a structure to these sentences? Would it not be sufficient to formulate a pairing of form and meaning for complete sentences just as we did

1 Introduction and basic terms

for the word *tree* above?

That would, in principle, be possible if a language were just a finite list of word sequences. If we were to assume that there is a maximum length for sentences and a maximum length for words and thus that there can only be a finite number of words, then the number of possible sentences would indeed be finite. However, even if we were to restrict the possible length of a sentence, the number of possible sentences would still be enormous. The question we would then really need to answer is: what is the maximum length of a sentence? For instance, it is possible to extend all the sentences in (1):

- (1) a. This sentence goes on and on and on and on ...
- b. [A sentence is a sentence] is a sentence.
- c. that Max thinks that Julius knows that Otto claims that Karl suspects that Richard confirms that Friederike is laughing

In (1b), something is being said about the group of words *a sentence is a sentence*, namely that it is a sentence. One can, of course, claim the same for the whole sentence in (1b) and extend the sentence once again with *is a sentence*. The sentence in (1c) has been formed by combining *that Friederike is laughing* with *that, Richard* and *confirms*. The result of this combination is a new sentence *that Richard confirms that Friederike is laughing*. In the same way, this has then been extended with *that, Karl* and *suspects*. Thus, one obtains a very complex sentence which embeds a less complex sentence. This partial sentence in turn contains a further partial sentence and so on. (1c) is similar to those sets of Russian nesting dolls, also called *matryoshka*: each doll contains a smaller doll which can be painted differently from the one that contains it. In just the same way, the sentence in (1c) contains parts which are similar to it but which are shorter and involve different nouns and verbs. This can be made clearer by using brackets in the following way:

- (2) that Max thinks [that Julius knows [that Otto claims [that Karl suspects [that Richard confirms [that Friederike is laughing]]]]]]

We can build incredibly long and complex sentences in the ways that were demonstrated in (1).¹

¹It is sometimes claimed that we are capable of constructing infinitely long sentences (NKN2001a; KS2008a-u; Dan Everett in OW2012a at 25:19; Chesi2015a-u; Martorell2018a; Wikipedia entry of Biolinguistics, 2018-09-27) or that Chomsky made such claims (Leiss2003a). This is, however, not correct since every sentence has to come to an end at some point. Even in

It would be arbitrary to establish some cut-off point up to which such combinations can be considered to belong to our language (Harris57a; Chomsky57a). It is also implausible to claim that such complex sentences are stored in our brains as a single complex unit. While evidence from psycholinguistic experiments shows that highly frequent or idiomatic combinations are stored as complex units, this could not be the case for sentences such as those in (1). Furthermore, we are capable of producing utterances that we have never heard before and which have also never been uttered or written down previously. Therefore, these utterances must have some kind of structure, there must be patterns which occur again and again. As humans, we are able to build such complex structures out of simpler ones and, vice-versa, to break down complex utterances into their component parts. Evidence for humans' ability to make use of rules for combining words into larger units has now also been provided by research in neuroscience (Pulvermueller2010a).

It becomes particularly evident that we combine linguistic material in a rule-governed way when these rules are violated. Children acquire linguistic rules by generalizing from the input available to them. In doing so, they produce some utterances which they could not have ever heard previously:

- (3) Ich festhalte die. (Friederike, 2;6)
 I PART.hold them
 Intended: 'I hold them tight.'

Friederike, who was learning German, was at the stage of acquiring the rule for the position of the finite verb (namely, second position). What she did here, however, was to place the whole verb, including a separable particle *fest* 'tight', in the second position although the particle should be realized at the end of the clause (*Ich halte die fest.*).

If we do not wish to assume that language is merely a list of pairings of form and meaning, then there must be some process whereby the meaning of complex utterances can be obtained from the meanings of the smaller components of those utterances. Syntax reveals something about the way in which the words involved can be combined, something about the structure of an utterance. For instance, knowledge about subject-verb agreement helps with the interpretation of the following sentences in German:

the theory of formal languages developed in the Chomskyan tradition, there are no infinitely long sentences. Rather, certain formal grammars can describe a set containing infinitely many finite sentences (Chomsky57a). See also PS2010a and Section ?? on the issue of recursion in grammar and for claims about the infinite nature of language.

1 Introduction and basic terms

- (4) a. Die Frau schläft.
the woman sleep.3SG
'The woman sleeps.'
- b. Die Mädchen schlafen.
the girls sleep.3PL
'The girls sleep.'
- c. Die Frau kennt die Mädchen.
the woman know.3SG the girls
'The woman knows the girls.'
- d. Die Frau kennen die Mädchen.
the woman know.3PL the girls
'The girls know the woman.'

The sentences in (4a,b) show that a singular or a plural subject requires a verb with the corresponding inflection. In (4a,b), the verb only requires one argument so the function of *die Frau* 'the woman' and *die Mädchen* 'the girls' is clear. In (4c,d) the verb requires two arguments and *die Frau* 'the woman' and *die Mädchen* 'the girls' could appear in either argument position in German. The sentences could mean that the woman knows somebody or that somebody knows the woman. However, due to the inflection on the verb and knowledge of the syntactic rules of German, the hearer knows that there is only one available reading for (4c) and (4d), respectively.

It is the role of syntax to discover, describe and explain such rules, patterns and structures.

1.2 Why do it formally?

The two following quotations give a motivation for the necessity of describing language formally:

Precisely constructed models for linguistic structure can play an important role, both negative and positive, in the process of discovery itself. By pushing a precise but inadequate formulation to an unacceptable conclusion, we can often expose the exact source of this inadequacy and, consequently, gain a deeper understanding of the linguistic data. More positively, a formalized theory may automatically provide solutions for many problems other than those for which it was explicitly designed. Obscure and intuition-bound notions can neither lead to absurd conclusions nor provide new and

correct ones, and hence they fail to be useful in two important respects. I think that some of those linguists who have questioned the value of precise and technical development of linguistic theory have failed to recognize the productive potential in the method of rigorously stating a proposed theory and applying it strictly to linguistic material with no attempt to avoid unacceptable conclusions by ad hoc adjustments or loose formulation. (Chomsky57a)

As is frequently pointed out but cannot be overemphasized, an important goal of formalization in linguistics is to enable subsequent researchers to see the defects of an analysis as clearly as its merits; only then can progress be made efficiently. (Dowty79a)

If we formalize linguistic descriptions, it is easier to recognize what exactly a particular analysis means. We can establish what predictions it makes and we can rule out alternative analyses. A further advantage of precisely formulated theories is that they can be written down in such a way that computer programs can process them. When a theoretical analysis is implemented as a computationally processable grammar fragment, any inconsistency will become immediately evident. Such implemented grammars can then be used to process large collections of text, so-called corpora, and they can thus establish which sentences a particular grammar cannot yet analyze or which sentences are assigned the wrong structure. For more on using computer implementation in linguistics see Bierwisch63a, Mueller99a and Bender2008c as well as Section 3.6.2.

1.3 Constituents

If we consider the sentence in (5), we have the intuition that certain words form a unit.

- (5) Alle Studenten lesen während dieser Zeit Bücher.
all students read during this time books
‘All the students are reading books at this time.’

For example, the words *alle* ‘all’ and *Studenten* ‘students’ form a unit which says something about who is reading. *während* ‘during’, *dieser* ‘this’ and *Zeit* ‘time’ also form a unit which refers to a period of time during which the reading takes place, and *Bücher* ‘books’ says something about what is being read. The first unit is itself made up of two parts, namely *alle* ‘all’ and *Studenten* ‘students’. The unit

während dieser Zeit ‘during this time’ can also be divided into two subcomponents: *während* ‘during’ and *dieser Zeit* ‘this time’. *dieser Zeit* ‘this time’ is also composed of two parts, just like *alle Studenten* ‘all students’ is.

Recall that in connection with (1c) above we talked about the sets of Russian nesting dolls (*matryoshkas*). Here, too, when we break down (5) we have smaller units which are components of bigger units. However, in contrast to the Russian dolls, we do not just have one smaller unit contained in a bigger one but rather, we can have several units which are grouped together in a bigger one. The best way to envisage this is to imagine a system of boxes: one big box contains the whole sentence. Inside this box, there are four other boxes, which each contain *alle Studenten* ‘all students’, *lesen* ‘reads’, *während dieser Zeit* ‘during this time’ and *Bücher* ‘books’, respectively. Figure 1.1 illustrates this.

[frbox,baseline] [frbox,baseline] [frbox,baseline] alle; [frbox,baseline]
Studenten;; [frbox,baseline] lesen; [frbox,baseline] [frbox,baseline] während;
[frbox,baseline] [frbox,baseline] dieser; [frbox,baseline] Zeit;; [frbox,baseline]
Bücher;;

Figure 1.1: Words and phrases in boxes

In the following section, I will introduce various tests which can be used to show how certain words seem to “belong together” more than others. When I speak of a *word sequence*, I generally mean an arbitrary linear sequence of words which do not necessarily need to have any syntactic or semantic relationship, e.g., *Studenten lesen während* ‘students read during’ in (5). A sequence of words which form a structural entity, on the other hand, is referred to as a *phrase*. Phrases can consist of words as in *this time* or of combinations of words with other phrases as in *during this time*. The parts of a phrase and the phrase itself are called *constituents*. So all elements that are in a box in Figure 1.1 are constituents of the sentence.

Following these preliminary remarks, I will now introduce some tests which will help us to identify whether a particular string of words is a constituent or not.

1.3.1 Constituency tests

There are a number of ways to test the constituent status of a sequence of words. In the following subsections, I will present some of these. In Section 1.3.2, we will see that there are cases when simply applying a test “blindly” leads to unwanted results.

1.3.1.1 Substitution

If it is possible to replace a sequence of words in a sentence with a different sequence of words and the acceptability of the sentence remains unaffected, then this constitutes evidence for the fact that each sequence of words forms a constituent.

In (6), *den Mann* ‘the man’ can be replaced by the string *eine Frau* ‘a woman’. This is an indication that both of these word sequences are constituents.

- (6) a. Er kennt [den Mann].
 he knows the man
 ‘He knows the man.’
 b. Er kennt [eine Frau].
 he knows a woman
 ‘He knows a woman.’

Similarly, in (7a), the string *das Buch zu lesen* ‘the book to read’ can be replaced by *dem Kind das Buch zu geben* ‘the woman the book to give’.

- (7) a. Er versucht, [das Buch zu lesen].
 he tries the book to read
 ‘He is trying to read the book.’
 b. Er versucht, [dem Kind das Buch zu geben].
 he tries the child the book to give
 ‘He is trying to give the child the book.’

This test is referred to as the *substitution test*.

1.3.1.2 Pronominalization

Everything that can be replaced by a pronoun forms a constituent. In (8), one can for example refer to *der Mann* ‘the man’ with the pronoun *er* ‘he’:

- (8) a. [Der Mann] schläft.
 the man sleeps
 ‘The man is sleeping.’
 b. Er schläft.
 he sleeps
 ‘He is sleeping.’

1 Introduction and basic terms

It is also possible to use a pronoun to refer to constituents such as *das Buch zu lesen* ‘the book to read’ in (7a), as is shown in (9):

- (9) a. Peter versucht, [das Buch zu lesen].
Peter tries the book to read
‘Peter is trying to read the book.’
b. Klaus versucht das auch.
Klaus tries that also
‘Klaus is trying to do that as well.’

The pronominalization test is another form of the substitution test.

1.3.1.3 Question formation

A sequence of words that can be elicited by a question forms a constituent:

- (10) a. [Der Mann] arbeitet.
the man works
‘The man is working.’
b. Wer arbeitet?
who works
‘Who is working?’

Question formation is a specific case of pronominalization. One uses a particular type of pronoun (an interrogative pronoun) to refer to the word sequence.

Constituents such as *das Buch zu lesen* in (7a) can also be elicited by questions, as (11) shows:

- (11) Was versucht er?
what tries he
‘What does he try?’

1.3.1.4 Permutation test

If a sequence of words can be moved without adversely affecting the acceptability of the sentence in which it occurs, then this is an indication that this word sequence forms a constituent.

In (12), *keiner* ‘nobody’ and *dieses Kind* ‘this child’ exhibit different orderings, which suggests that *dieses* ‘this’ and *Kind* ‘child’ belong together.

- (12) a. dass keiner [dieses Kind] kennt
 that nobody this child knows
 b. dass [dieses Kind] keiner kennt
 that this child nobody knows
 ‘that nobody knows this child’

On the other hand, it is not plausible to assume that *keiner dieses* ‘nobody this’ forms a constituent in (12a). If we try to form other possible orderings by trying to move *keiner dieses* ‘nobody this’ as a whole, we see that this leads to unacceptable results:²

- (13) a. *dass Kind keiner dieses kennt
 that child nobody this knows
 b. *dass Kind kennt keiner dieses
 that child knows nobody this

Furthermore, constituents such as *das Buch zu lesen* ‘to read the book’ in (7a) can be moved:

- (14) a. Er hat noch nicht [das Buch zu lesen] versucht.
 he has PART not the book to read tried
 ‘He has not yet tried to read the book.’
 b. Er hat [das Buch zu lesen] noch nicht versucht.
 he has the book to read PART not tried
 c. Er hat noch nicht versucht, [das Buch zu lesen].
 he has PART not tried the book to read

1.3.1.5 Fronting

Fronting is a further variant of the movement test. In German declarative sentences, only a single constituent may normally precede the finite verb:

- (15) a. [Alle Studenten] lesen während der vorlesungsfreien Zeit Bücher.
 all students read.3PL during the lecture.free time books
 ‘All students read books during the semester break.’

²I use the following notational conventions for all examples: ‘**’ indicates that a sentence is ungrammatical, ‘#’ denotes that the sentence has a reading which differs from the intended one and finally ‘\$’ should be understood as a sentence which is deviant for semantic or information-structural reasons, for example, because the subject must be animate, but is in fact inanimate in the example in question, or because there is a conflict between constituent order and the marking of given information through the use of pronouns.

1 Introduction and basic terms

- b. [Bücher] lesen alle Studenten während der vorlesungsfreien Zeit.
books read all students during the lecture.free time
- c. * [Alle Studenten] [Bücher] lesen während der vorlesungsfreien Zeit.
all students books read during the lecture.free time
- d. * [Bücher] [alle Studenten] lesen während der vorlesungsfreien Zeit.
books all students read during the lecture.free time

The possibility for a sequence of words to be fronted (that is to occur in front of the finite verb) is a strong indicator of constituent status.

1.3.1.6 Coordination

If two sequences of words can be conjoined then this suggests that each sequence forms a constituent.

In (16), *der Mann* ‘the man’ and *die Frau* ‘the woman’ are conjoined and the entire coordination is the subject of the verb *arbeiten* ‘to work’. This is a good indication of the fact that *der Mann* and *die Frau* each form a constituent.

- (16) [Der Mann] und [die Frau] arbeiten.
the man and the woman work.3PL
‘The man and the woman work.’

The example in (17) shows that phrases with *to* = infinitives can be conjoined:

- (17) Er hat versucht, [das Buch zu lesen] und [es dann unauffällig verschwinden
he had tried the book to read and it then secretly disappear
zu lassen].
to let
‘He tried to read the book and then make it quietly disappear.’

1.3.2 Some comments on the status of constituent tests

It would be ideal if the tests presented here delivered clear-cut results in every case, as the empirical basis on which syntactic theories are built would thereby become much clearer. Unfortunately, this is not the case. There are in fact a number of problems with constituent tests, which I will discuss in what follows.

1.3.2.1 Expletives

There is a particular class of pronouns – so-called *expletives* – which do not denote people, things, or events and are therefore non-referential. An example of this is *es* ‘it’ in (18).

- (18) a. Es regnet.
it rains
‘It is raining.’
b. Regnet es?
rains it
‘Is it raining?’
c. dass es jetzt regnet
that it now rains
‘that it is raining now’

As the examples in (18) show, *es* can either precede the verb, or follow it. It can also be separated from the verb by an adverb, which suggests that *es* should be viewed as an independent unit.

Nevertheless, we observe certain problems with the aforementioned tests. Firstly, *es* ‘it’ is restricted with regard to its movement possibilities, as (19a) and (20b) show.

- (19) a. *dass jetzt es regnet
that now it rains
Intended: ‘that it is raining now’
b. dass jetzt keiner klatscht
that now nobody claps
‘that nobody is clapping now’
- (20) a. Er sah es regnen.
he saw it.ACC rain
‘He saw that it was raining.’
b. *Es sah er regnen.
it.ACC saw he rain
Intended: ‘he saw that it was raining.’
c. Er sah einen Mann klatschen.
he saw a.ACC man clap
‘He saw a man clapping.’

1 Introduction and basic terms

- d. Einen Mann sah er klatschen.
a.ACC man saw he clap
'A man, he saw clapping.'

Unlike the accusative object *einen Mann* 'a man' in (20c,d), the expletive in (20b) cannot be fronted.

Secondly, substitution and question tests also fail:

- (21) a. *Der Mann / er regnet.
the man he rains
b. *Wer / was regnet?
who what rains

Similarly, the coordination test cannot be applied either:

- (22) *Es und der Mann regnet / regnen.
it and the man rains rain

The failure of these tests can be easily explained: weakly stressed pronouns such as *es* are preferably placed before other arguments, directly after the conjunction (*dass* in (18c)) and directly after the finite verb in (20a) (see Abraham95a-u). If an element is placed in front of the expletive, as in (19a), then the sentence is rendered ungrammatical. The reason for the ungrammaticality of (20b) is the general ban on accusative *es* appearing in clause-initial position. Although such cases exist, they are only possible if *es* 'it' is referential (Lenerz94a; GS97a).

The fact that we could not apply the substitution and question tests is also no longer mysterious as *es* is not referential in these cases. We can only replace *es* 'it' with another expletive such as *das* 'that'. If we replace the expletive with a referential expression, we derive a different semantic interpretation. It does not make sense to ask about something semantically empty or to refer to it with a pronoun.

It follows from this that not all of the tests must deliver a positive result for a sequence of words to count as a constituent. That is, the tests are therefore not a necessary requirement for constituent status.

1.3.2.2 Movement

The movement test is problematic for languages with relatively free constituent order, since it is not always possible to tell what exactly has been moved. For example, the string *gestern dem Mann* 'yesterday the man' occupies different positions in the following examples:

- (23) a. weil keiner gestern dem Mann geholfen hat
 because nobody yesterday the man helped has
 ‘because nobody helped the man yesterday’
 b. weil gestern dem Mann keiner geholfen hat
 because yesterday the man nobody helped has
 ‘because nobody helped the man yesterday’

One could therefore assume that *gestern* ‘yesterday’ and *dem Mann* ‘the man’, which of course do not form a constituent, have been moved together. An alternative explanation for the ordering variants in (23) is that adverbs can occur in various positions in the clause and that only *dem Mann* ‘the man’ has been moved in front of *keiner* ‘nobody’ in (23b). In any case, it is clear that *gestern* and *dem Mann* have no semantic relation and that it is impossible to refer to both of them with a pronoun. Although it may seem at first glance as if this material had been moved as a unit, we have seen that it is in fact not tenable to assume that *gestern dem Mann* ‘yesterday the man’ forms a constituent.

1.3.2.3 Fronting

As mentioned in the discussion of (15), the position in front of the finite verb is normally occupied by a single constituent. The possibility for a given word sequence to be placed in front of the finite verb is sometimes even used as a clear indicator of constituent status, and even used in the definition of *Satzglied*³. An example of this is taken from **Bussmann83a**, but is no longer present in **Bussmann90a**.⁴

Satzglied test A procedure based on \rightarrow topicalization used to analyze complex constituents. Since topicalization only allows a single constituent to be moved to the beginning of the sentence, complex sequences of constituents, for example adverb phrases, can be shown to actually consist of one or more constituents. In the example *Ein Taxi quält sich im Schrittempo durch den*

³ *Satzglied* is a special term used in grammars of German, referring to a constituent on the clause level (**Duden2005-Authors**).

⁴ The original formulation is: **Satzgliedtest** [Auch: Konstituententest]. Auf der \rightarrow Topikalisierung beruhendes Verfahren zur Analyse komplexer Konstituenten. Da bei Topikalisierung jeweils nur eine Konstituente bzw. ein \rightarrow Satzglied an den Anfang gerückt werden kann, lassen sich komplexe Abfolgen von Konstituenten (z. B. Adverbialphrasen) als ein oder mehrere Satzglieder ausweisen; in *Ein Taxi quält sich im Schrittempo durch den Verkehr* sind *im Schrittempo* und *durch den Verkehr* zwei Satzglieder, da sie beide unabhängig voneinander in Anfangsposition gerückt werden können.

1 Introduction and basic terms

Verkehr ‘A taxi was struggling at walking speed through the traffic’, *im Schrittempo* ‘at walking speed’ and *durch den Verkehr* ‘through the traffic’ are each constituents as both can be fronted independently of each other. (Bussmann83a)

The preceding quote has the following implications:

- Some part of a piece of linguistic material can be fronted independently → This material does not form a constituent.
- Linguistic material can be fronted together → This material forms a constituent.

It will be shown that both of these prove to be problematic.

The first implication is cast into doubt by the data in (24):

- (24) a. Keine Einigung erreichten Schröder und Chirac über den Abbau
no agreement reached Schröder and Chirac about the reduction
der Agrarsubventionen.⁵
of.the agricultural.subsidies
‘Schröder and Chirac could not reach an agreement on the reduction
of agricultural subsidies.’
- b. [Über den Abbau der Agrarsubventionen] erreichten Schröder
about the reduction of.the agricultural.subsidies reached Schröder
und Chirac keine Einigung.
and Chirac no agreement

Although parts of the noun phrase *keine Einigung über den Abbau der Agrarsubventionen* ‘no agreement on the reduction of agricultural subsidies’ can be fronted individually, we still want to analyze the entire string as a noun phrase when it is not fronted as in (25):

- (25) Schröder und Chirac erreichten [keine Einigung über den Abbau
Schröder and Chirac reached no agreement about the reduction
der Agrarsubventionen].
of.the agricultural.subsidies

The prepositional phrase *über den Abbau der Agrarsubventionen* ‘on the reduction of agricultural subsidies’ is semantically dependent on *Einigung* ‘agreement’ cf. (26):

⁵tagesschau, 15.10.2002, 20:00.

- (26) Sie einigen sich über die Agrarsubventionen.
they agree REFL about the agricultural.subsidies
'They agree on the agricultural subsidies.'

This word sequence can also be fronted together:

- (27) [Keine Einigung über den Abbau der Agrarsubventionen] erreichten
no agreement about the reduction of.the agricultural.subsidies reached
Schröder und Chirac.
Schröder and Chirac

In the theoretical literature, it is assumed that *keine Einigung über den Abbau der Agrarsubventionen* forms a constituent which can be “split up” under certain circumstances.

In such cases, the individual constituents can be moved independently of each other (deKuthy2002a) as we have seen in (25).

The second implication is problematic because of examples such as (28):

- (28) a. [Trocken] [durch die Stadt] kommt man am Wochenende auch
dry through the city comes one at.the weekend also
mit der BVG.⁶
with the BVG
'With the BVG, you can be sure to get around town dry at the weekend.'
- b. [Wenig] [mit Sprachgeschichte] hat der dritte Beitrag in dieser
little with language.history has the third contribution in this
Rubrik zu tun, [...] ⁷
section to do
'The third contribution in this section has little to do with language history.'

In (28), there are multiple constituents preceding the finite verb, which bear no obvious syntactic or semantic relation to each other. Exactly what is meant by a "syntactic or semantic relation" will be fully explained in the following chapters. At this point, I will just point out that in (28a) the adjective *trocken* 'dry' has *man* 'one' as its subject and furthermore says something about the action of 'travelling through the city'. That is, it refers to the action denoted by the verb. As (29b) shows, *durch die Stadt* 'through the city' cannot be combined with the adjective *trocken* 'dry'.

- (29) a. Man ist / bleibt trocken.
one is stays dry
'One is/stays dry.'
- b. * Man ist / bleibt trocken durch die Stadt.
one is stays dry through the city

Therefore, the adjective *trocken* 'dry' does not have a syntactic or semantic relationship with the prepositional phrase *durch die Stadt* 'through the city'. Both phrases have in common that they refer to the verb and are dependent on it.

One may simply wish to treat the examples in (28) as exceptions. This approach would, however, not be justified, as I have shown in an extensive empirical study (Mueller2003b).

⁶taz berlin, 10.07.1998, p. 22.

⁷Zeitschrift für Dialektologie und Linguistik, LXIX, 3/2002, p. 339.

If one were to classify *trocken durch die Stadt* as a constituent due to it passing the fronting test, then one would have to assume that *trocken durch die Stadt* in (30) is also a constituent. In doing so, we would devalue the term *constituent* as the whole point of constituent tests is to find out which word strings have some semantic or syntactic relationship.⁸

- (30) a. Man kommt am Wochenende auch mit der BVG trocken durch
 one comes at.the weekend also with the BVG dry through
 die Stadt.
 the city
 ‘With the BVG, you can be sure to get around town dry at the weekend.’
 b. Der dritte Beitrag in dieser Rubrik hat wenig mit Sprachgeschichte
 the third contribution in this section has little with language.history
 zu tun.
 to do
 ‘The third contribution in this section has little to do with language
 history.’

The possibility for a given sequence of words to be fronted is therefore not a sufficient diagnostic for constituent status.

We have also seen that it makes sense to treat expletives as constituents despite the fact that the accusative expletive cannot be fronted (cf. (20a)):

- (31) a. Er bringt es bis zum Professor.
 he brings EXPL until to.the professor
 ‘He makes it to professor.’
 b. # Es bringt er bis zum Professor.
 it brings he until to.the professor

There are other elements that can also not be fronted. Inherent reflexives are a good example of this:

- (32) a. Karl hat sich nicht erholt.
 Karl has REFL not recovered
 ‘Karl hasn’t recovered.’

⁸These data can be explained by assuming a silent verbal head preceding the finite verb and thereby ensuring that there is in fact just one constituent in initial position in front of the finite verb (Mueller2005d; MuellerGS). Nevertheless, this kind of data are problematic for constituent tests since these tests have been specifically designed to tease apart whether strings such as *trocken* and *durch die Stadt* or *wenig* and *mit Sprachgeschichte* in (30) form a constituent.

1 Introduction and basic terms

- b. *Sich hat Karl nicht erholt.
REFL has Karl not recovered

It follows from this that fronting is not a necessary criterion for constituent status. Therefore, the possibility for a given word string to be fronted is neither a necessary nor sufficient condition for constituent status.

1.3.2.4 Coordination

Coordinated structures such as those in (33) also prove to be problematic:

- (33) Deshalb kaufte der Mann einen Esel und die Frau ein Pferd.
therefore bought the man a donkey and the woman a horse
'Therefore, the man bought a donkey and the woman a horse.'

At first glance, *der Mann einen Esel* 'the man a donkey' and *die Frau ein Pferd* 'the woman a horse' in (33) seem to be coordinated. Does this mean that *der Mann einen Esel* and *die Frau ein Pferd* each form a constituent?

As other constituent tests show, this assumption is not plausible. This sequence of words cannot be moved together as a unit:⁹

- (34) *Der Mann einen Esel kaufte deshalb.
the man a donkey bought therefore

Replacing the supposed constituent is also not possible without ellipsis:

- (35) a. #Deshalb kaufte er.
therefore bought he
b. *Deshalb kaufte ihn.
therefore bought him

The pronouns do not stand in for the two logical arguments of *kaufen* 'to buy', which are realized by *der Mann* 'the man' and *einen Esel* 'a donkey' in (33), but rather for one in each. There are analyses that have been proposed for examples such as (33) in which two verbs *kauft* 'buys' occur, where only one is overt,

⁹The area in front of the finite verb is also referred to as the *Vorfeld* 'prefield' (see Section 1.8). Apparent multiple fronting is possible under certain circumstances in German. See the previous section, especially the discussion of the examples in (28) on page 18. The example in (34) is created in such a way that the subject is present in the prefield, which is not normally possible with verbs such as *kaufen* 'to buy' for reasons which have to do with the information-structural properties of these kinds of fronting constructions. Compare also dKM2003a on subjects in fronted verb phrases and BC2010a on frontings of subjects in apparent multiple frontings.

however (Crysmann2003c). The example in (33) would therefore correspond to:

- (36) Deshalb kaufte der Mann einen Esel und kaufte die Frau ein
 therefore bought the man a donkey and bought the woman a
 Pferd.
 horse

This means that although it seems as though *der Mann einen Esel* ‘the man a donkey’ and *die Frau ein Pferd* ‘the woman a horse’ are coordinated, it is actually *kauft der Mann einen Esel* ‘buys the man a donkey’ and (*kauft*) *die Frau ein Pferd* ‘buys the woman a horse’ which are conjoined.

We should take the following from the previous discussion: even when a given word sequence passes certain constituent tests, this does not mean that one can automatically infer from this that we are dealing with a constituent. That is, the tests we have seen are not sufficient conditions for constituent status.

Summing up, it has been shown that these tests are neither sufficient nor necessary for attributing constituent status to a given sequence of words. However, as long as one keeps the problematic cases in mind, the previous discussion should be enough to get an initial idea about what should be treated as a constituent.

1.4 Parts of speech

The words in (37) differ not only in their meaning but also in other respects.

- (37) Der große Biber schwimmt jetzt.
 the big beaver swims now
 ‘The big beaver swims now.’

Each of the words is subject to certain restrictions when forming sentences. It is common practice to group words into classes with other words which share certain salient properties. For example, *der* ‘the’ is an article, *Biber* ‘beaver’ is a noun, *schwimmt* ‘swims’ is a verb and *jetzt* ‘now’ is an adverb. As can be seen in (38), it is possible to replace all the words in (37) with words from the same word class.

- (38) Die kleine Raupe frisst immer.
 the small caterpillar eats always
 ‘The small caterpillar is always eating.’

1 Introduction and basic terms

This is not always the case, however. For example, it is not possible to use a verb such as *verschlingt* ‘devours’ or the second-person form *schwimmst* in (38). This means that the categorization of words into parts of speech is rather coarse and that we will have to say a lot more about the properties of a given word. In this section, I will discuss various word classes/”parts of speech and in the following sections I will go into further detail about the various properties which characterize a given word class.

The most important parts of speech are *verbs*, *nouns*, *adjectives*, *prepositions* and *adverbs*. In earlier decades, it was common among researchers working on German (see also Section ?? on Tesnière’s category system) to speak of *action words*, *describing words*, and *naming words*. These descriptions prove problematic, however, as illustrated by the following examples:

- (39) a. die *Idee*
the idea
b. die *Stunde*
the hour
c. das laute *Sprechen*
the loud speaking
‘(the act of) speaking loudly’
d. Die *Erörterung* der Lage dauerte mehrere Stunden.
the discussion of the situation lasted several hours
‘The discussion of the situation lasted several hours.’

(39a) does not describe a concrete entity, (39b) describes a time interval and (39c) and (39d) describe actions. It is clear that *Idee* ‘idea’, *Stunde* ‘hour’, *Sprechen* ‘speaking’ and *Erörterung* ‘discussion’ differ greatly in terms of their meaning. Nevertheless, these words still behave like *Raupe* ‘caterpillar’ and *Biber* ‘beaver’ in many respects and are therefore classed as nouns.

The term *action word* is not used in scientific linguistic work as verbs do not always need to denote actions:

- (40) a. Ihm gefällt das Buch.
him pleases the book
‘He likes the book.’
b. Das Eis schmilzt.
the ice melts
‘The ice is melting.’

- c. Es regnet.
it rains
'It is raining.'

One would also have to class the noun *Erörterung* 'discussion' as an action word.

Adjectives do not always describe properties of objects. In the following examples, the opposite is in fact true: the characteristic of being a murderer is expressed as being possible or probable, but not as being true properties of the modified noun.

- (41) a. der mutmaßliche Mörder
the suspected murderer
b. Soldaten sind potenzielle Mörder.
soldiers are potential murderers

The adjectives themselves in (41) do not actually provide any information about the characteristics of the entities described. One may also wish to classify *lachende* 'laughing' in (42) as an adjective.

- (42) der lachende Mann
the laughing man

If, however, we are using properties and actions as our criteria for classification, *lachend* 'laughing' should technically be an action word.

Rather than semantic criteria, it is usually formal criteria which are used to determine word classes. The various forms a word can take are also taken into account. So *lacht* 'laughs', for example, has the forms given in (43).

- (43) a. Ich lache.
I laugh
b. Du lachst.
you.SG laugh
c. Er lacht.
he laughs
d. Wir lachen.
we laugh
e. Ihr lacht.
you.PL laugh
f. Sie lachen.
they laugh

1 Introduction and basic terms

In German, there are also forms for the preterite, imperative, present subjunctive, past subjunctive and non-finite forms (participles and infinitives with or without *zu* 'to'). All of these forms constitute the inflectional paradigm of a verb. Tense (present, preterite, future), mood (indicative, subjunctive, imperative), person (1st, 2nd, 3rd) and number (singular, plural) all play a role in the inflectional paradigm. Certain forms can coincide in a paradigm, as (43c) and (43e) and (43d) and (43f) show.

Parallel to verbs, nouns also have an inflectional paradigm:

- (44)
- a. der Mann
the.NOM man
 - b. des Mannes
the.GEN man.GEN
 - c. dem Mann
the.DAT man
 - d. den Mann
the.ACC man
 - e. die Männer
the.NOM men
 - f. der Männer
the.GEN men
 - g. den Männern
the.DAT men.DAT
 - h. die Männer
the.ACC men

We can differentiate between nouns on the basis of gender (feminine, masculine, neuter). The choice of gender is often purely formal in nature and is only partially influenced by biological sex or the fact that we are describing a particular object:

- (45)
- a. die Tüte
the.F bag(F)
'the bag'
 - b. der Krampf
the.M cramp(M)
'cramp'

- c. das Kind
the.N child(N)
'the child'

As well as gender, case (nominative, genitive, dative, accusative) and number are also important for nominal paradigms.

Like nouns, adjectives inflect for gender, case and number. They differ from nouns, however, in that gender marking is variable. Adjectives can be used with all three genders:

- (46) a. eine schöne Blume
a.F beautiful.F flower
b. ein schöner Strauß
a beautiful.M bunch
c. ein schönes Bouquet
a beautiful.N bouquet

In addition to gender, case and number, we can identify several inflectional classes. Traditionally, we distinguish between strong, mixed and weak inflection of adjectives. The inflectional class that we have to choose is dependent on the form or presence of the article:

- (47) a. ein alter Wein
an old wine
b. der alte Wein
the old wine
c. alter Wein
old wine

Furthermore, adjectives have comparative and superlative wordforms:

- (48) a. klug
clever
b. klüg-er
clever-er
c. am klüg-sten
at.the clever-est

This is not always the case. Especially for adjectives which make reference to some end point, a degree of comparison does not make sense. If a particular

1 Introduction and basic terms

solution is optimal, for example, then no better one exists. Therefore, it does not make sense to speak of a “more optimal” solution. In a similar vein, it is not possible to be “deader” than dead.

There are some special cases such as color adjectives ending in ” *a* in German *lila* ‘purple’ and *rosa* ‘pink’. These inflect optionally (49a), and the uninflected form is also possible:

- (49) a. eine lilan-e Blume
 a purple-F flower
 b. eine lila Blume
 a purple flower

In both cases, *lila* is classed an adjective. We can motivate this classification by appealing to the fact that both words occur at the same positions as other adjectives that clearly behave like adjectives with regard to inflection.

The parts of speech discussed thus far can all be differentiated in terms of their inflectional properties. For words which do not inflect, we have to use additional criteria. For example, we can classify words by the syntactic context in which they occur (as we did for the non-inflecting adjectives above). We can identify prepositions, adverbs, conjunctions, interjections and sometimes also particles. Prepositions are words which occur with a noun phrase whose case they determine:

- (50) a. in diesen Raum
 in this.ACC room
 b. in diesem Raum
 in this.DAT room

wegen ‘because’ is often classed as a preposition although it can also occur after the noun and in these cases would technically be a postposition:

- (51) des Geldes wegen
 the money.GEN because
 ‘because of the money’

It is also possible to speak of *adpositions* if one wishes to remain neutral about the exact position of the word.

Unlike prepositions, adverbs do not require a noun phrase.

- (52) a. Er schläft in diesem Raum.
 he sleeps in this room

- b. Er schläft dort.
he sleeps there

Sometimes adverbs are simply treated as a special variant of prepositions (see page 105). The explanation for this is that a prepositional phrase such as *in diesem Raum* ‘in this room’ shows the same syntactic distribution as the corresponding adverbs. *in* differs from *dort* ‘there’ in that it needs an additional noun phrase. These differences are parallel to what we have seen with other parts of speech. For instance, the verb *schlafen* ‘sleep’ requires only a noun phrase, whereas *erkennen* ‘recognize’ requires two.

- (53) a. Er schläft.
he sleeps
b. Peter erkennt ihn.
Peter recognizes him

Conjunctions can be subdivided into subordinating and coordinating conjunctions. Coordinating conjunctions include *und* ‘and’ and *oder* ‘or’. In coordinate structures, two units with the same syntactic properties are combined. They occur adjacent to one another. *dass* ‘that’ and *weil* ‘because’ are subordinating conjunctions because the clauses that they introduce can be part of a larger clause and depend on another element of this larger clause.

- (54) a. Klaus glaubt, dass er lügt.
Klaus believes that he lies
‘Klaus believes that he is lying.’
b. Klaus glaubt ihm nicht, weil er lügt.
Klaus believes him not because he lies
‘Klaus doesn’t believe him because he is lying.’

Interjections are “clause”-like expressions such as *Ja!* ‘Yes!’, *Bitte!* ‘Please!’, *Hallo!* ‘Hello!’, *Hurra!* ‘Hooray!’, *Bravo!* ‘Bravo!’, *Pst!* ‘Psst!’, *Plumps!* ‘Clonk!’.

If adverbs and prepositions are not assigned to the same class, then adverbs are normally used as a kind of “left over” category in the sense that all non-“inflecting” words which are neither prepositions, conjunctions nor interjections are classed as adverbs. Sometimes this category for “left overs” is subdivided: only words which can appear in front of the finite verb when used as a constituent are referred to as adverbs. Those words which cannot be fronted are dubbed *particles*. Particles themselves can be subdivided into various classes based on their function, e.g., degree particles and illocutionary particles. Since these functionally

1 Introduction and basic terms

defined classes also contain adjectives, I will not make this distinction and simply speak of *adverbs*.

We have already sorted a considerable number of inflectional words into word classes. When one is faced with the task of classifying a particular word, one can use the decision diagram in Figure 1.2, which is taken from the Duden grammar of German (**Duden2005-Authors**).¹⁰

Figure 1.2: Decision tree for determining parts of speech following **Duden2005-Authors**

If a word inflects for tense, then it is a verb. If it displays different case forms, then one has to check if it has a fixed gender. If this is indeed the case, then we know that we are dealing with a noun. Words with variable gender have to be checked to see if they have comparative forms. A positive result will be a clear indication of an adjective. All other words are placed into a residual category, which the Duden refers to as pronouns/article words. Like in the class of non-inflectional elements, the elements in this remnant category are subdivided according to their syntactic behavior. The Duden grammar makes a distinction between pronouns and article words. According to this classification, pronouns are words which can replace a noun phrase such as *der Mann* ‘the man’, whereas article words normally combine with a noun. In Latin grammars, the notion of ‘pronoun’ includes both pronouns in the above sense and articles, since the forms with and without the noun are identical. Over the past centuries, the forms have undergone split development to the point where it is now common in contemporary Romance languages to distinguish between words which replace a noun phrase and those which must occur with a noun. Elements which belong to the latter class are also referred to as *determiners*.

If we follow the decision tree in Figure 1.2, the personal pronouns *ich* ‘I’, *du* ‘you’, *er* ‘he’, *sie* ‘her’, *es* ‘it’, *wir* ‘we’, *ihr* ‘you’, and *sie* ‘they’, for example, would be grouped together with the possessive pronouns *mein* ‘mine’, *dein* ‘your’, *sein* ‘his’/”‘its’, *ihr* ‘her’/”‘their’, *unser* ‘our’, and *euer* ‘your’. The corresponding reflexive pronouns, *mich* ‘myself’, *dich* ‘yourself’, *sich* ‘himself’/”‘herself’/”‘itself’, ‘themselves’, *uns* ‘ourselves’, *euch* ‘yourself’, and the reciprocal pronoun *einander* ‘each other’ have to be viewed as a special case in German as there are no differing gender forms of *sich* ‘himself’/”‘herself’/”‘itself’ and *einander* ‘each other’.

¹⁰The Duden is the official document for the German orthography. The Duden grammar does not have an official status but is very influential and is used for educational purposes as well. I will refer to it several times in this introductory chapter.

Case is not expressed morphologically by reciprocal pronouns. By replacing genitive, dative and accusative pronouns with *einander*, it is possible to see that there must be variants of *einander* ‘each other’ in these cases, but these variants all share the same form:

- (55) a. Sie gedenken seiner / einander.
they commemorate him.GEN each.other
b. Sie helfen ihm / einander.
they help him.DAT each.other
c. Sie lieben ihn / einander.
they love him.ACC each.other

So-called pronominal adverbs such as *darauf* ‘on there’, *darin* ‘in there’, *worauf* ‘on where’, *worin* ‘in where’ also prove problematic. These forms consist of a preposition (e.g., *auf* ‘on’) and the elements *da* ‘there’ and *wo* ‘where’. As the name suggests, *pronominal adverbs* contain something pronominal and this can only be *da* ‘there’ and *wo* ‘where’. However, *da* ‘there’ and *wo* ‘where’ do not inflect and would therefore, following the decision tree, not be classed as pronouns.

The same is true of relative pronouns such as *wo* ‘where’ in (56):

- (56) a. Ich komme eben aus der Stadt, wo ich Zeuge eines Unglücks
I come PART from the city where I witness of.an accident
gewesen bin.¹¹
been am
‘I come from the city where I was witness to an accident.’
b. Studien haben gezeigt, daß mehr Unfälle in Städten passieren, wo
studies have shown that more accidents in cities happen where
die Zebrastreifen abgebaut werden, weil die Autofahrer unaufmerksam
the zebra.crossings removed become because the drivers unattentive
werden.¹²
become
‘Studies have shown that there are more accidents in cities where they
do away with zebra crossings, because drivers become unattentive.’
c. Zufällig war ich in dem Augenblick zugegen, wo der Steppenwolf
coincidentally was I in the moment present where the Steppenwolf

¹¹Duden84-Authors.

¹²taz berlin, 03.11.1997, p. 23.

zum erstenmal unser Haus betrat und bei meiner Tante sich
 to.the first.time our house entered and by my aunt REFL
 einmietete.¹³
 took lodgings

‘Coincidentally, I was present at the exact moment in which Steppen-
 wolf entered our house for the first time and took lodgings with my
 aunt.’

If they are uninflected, then they cannot belong to the class of pronouns according to the decision tree above. Eisenberg2004a notes that *wo* ‘where’ is a kind of *uninflected relative pronoun* (he uses quotation marks) and remarks that this term runs contrary to the exclusive use of the term pronoun for nominal, that is, inflected, elements. He therefore uses the term *relative adverb* for them (see also Duden2005-Authors).

There are also usages of the relatives *dessen* ‘whose’ and *wessen* ‘whose’ in combination with a noun:

- (57) a. der Mann, dessen Schwester ich kenne
 the man whose sister I know
 b. Ich möchte wissen, wessen Schwester du kennst.
 I would.like know whose sister you know
 ‘I would like to know whose sister you know.’

According to the classification in the Duden, these should be covered by the terms *Relativartikelwort* ‘relative article word’ and *Interrogativartikelwort* ‘interrogative article word’. They are mostly counted as part of the relative pronouns and question pronouns (see for instance Eisenberg2004a). Using Eisenberg’s terminology, this is unproblematic as he does not make a distinction between articles, pronouns and nouns, but rather assigns them all to the class of nouns. But authors who do make a distinction between articles and pronouns sometimes also speak of interrogative pronouns when discussing words which can function as articles or indeed replace an entire noun phrase.

One should be prepared for the fact that the term *pronoun* is often simply used for words which refer to other entities and, this is important, not in the way that nouns such as *book* and *John* do, but rather dependent on context. The personal pronoun *er* ‘he’ can, for example, refer to either a table or a man. This usage of the term *pronoun* runs contrary to the decision tree in Figure 1.2 and includes uninflected elements such as *da* ‘there’ and *wo* ‘where’.

¹³Herman Hesse, *Der Steppenwolf*. Berlin und Weimar: Aufbau-Verlag. 1986, p. 6.

Expletive pronouns such as *es* ‘it’ and *das* ‘that’, as well as the *sich* ‘him’/””‘her’/””‘itself’ belonging to inherently reflexive verbs, do not make reference to actual objects. They are considered pronouns because of the similarity in form. Even if we were to assume a narrow definition of pronouns, we would still get the wrong results as expletive forms do not vary with regard to case, gender and number. If one does everything by the book, expletives would belong to the class of uninflected elements. If we assume that *es* ‘it’ as well as the personal pronouns have a nominative and accusative variant with the same form, then they would be placed in with the nominals. We would then have to admit that the assumption that *es* has gender would not make sense. That is we would have to count *es* as a noun by assuming neuter gender, analogous to personal pronouns.

We have not yet discussed how we would deal with the italicized words in (58):

- (58) a. das *geliebte* Spielzeug
the beloved toy
- b. das *schlafende* Kind
the sleeping child
- c. die Frage des *Sprechens* und *Schreibens* über Gefühle
the question of the talking and writing about feelings
‘the question of talking and writing about feelings’
- d. Auf dem Europa-Parteitag fordern die *Grünen* einen ökosozialen
on the Europe-party.conference demand the Greens a eco-social
Politikwechsel.
political.change
‘At the European party conference, the Greens demanded eco-social
political change.’
- e. Max lacht *laut*.
Max laughs loudly
- f. Max würde *wahrscheinlich* lachen.
Max would probably laugh

geliebte ‘beloved’ and *schlafende* ‘sleeping’ are participle forms of *lieben* ‘to love’ and *schlafen* ‘to sleep’. These forms are traditionally treated as part of the verbal paradigm. In this sense, *geliebte* and *schlafende* are verbs. This is referred to as lexical word class. The term *lexeme* is relevant in this case. All forms in a given inflectional paradigm belong to the relevant lexeme. In the classic sense, this term also includes the regularly derived forms. That is participle forms and

nominalized infinitives also belong to a verbal lexeme. Not all linguists share this view, however. Particularly problematic is the fact that we are mixing verbal with nominal and adjectival paradigms. For example, *Sprechens* ‘speaking.GEN’ is in the genitive case and adjectival participles also inflect for case, number and gender. Furthermore, it is unclear as to why *schlafende* ‘sleeping’ should be classed as a verbal lexeme and a noun such as *Störung* ‘disturbance’ is its own lexeme and does not belong to the lexeme *stören* ‘to disturb’. I subscribe to the more modern view of grammar and assume that processes in which a word class is changed result in a new lexeme being created. Consequently, *schlafende* ‘sleeping’ does not belong to the lexeme *schlafen* ‘to sleep’, but is a form of the lexeme *schlafend*. This lexeme belongs to the word class ‘adjective’ and inflects accordingly.

As we have seen, it is still controversial as to where to draw the line between inflection and derivation (creation of a new lexeme). SWB2003a view the formation of the present participle (*standing*) and the past participle (*eaten*) in English as derivation as these forms inflect for gender and number in French.

Adjectives such as *Grünen* ‘the Greens’ in (58d) are nominalized adjectives and are written with a capital like other nouns in German when there is no other noun that can be inferred from the immediate context:

- (59) A: Willst du den roten Ball haben?
 want you the red ball have
 ‘Do you want the red ball?’
 B: Nein, gib mir bitte den grünen.
 no give me please the green
 ‘No, give me the green one, please.’

In the answer to (59), the noun *Ball* has been omitted. This kind of omission is not present in (58d). One could also assume here that a word class change has taken place. If a word changes its class without combination with a visible affix, we refer to this as *conversion*. Conversion has been treated as a sub-case of derivation by some linguists. The problem is, however, that *Grüne* ‘greens’ inflects just like an adjective and the gender varies depending on the object it is referring to:

- (60) a. Ein Grüner hat vorgeschlagen, ...
 a green.M has suggested
 ‘A (male) member of the Green Party suggested ...’
 b. Eine Grüne hat vorgeschlagen, ...
 a green.F has suggested
 ‘A (female) member of the Green Party suggested ...’

We also have the situation where a word has two properties. We can make life easier for ourselves by talking about *nominalized adjectives*. The lexical category of *Grüne* is adjective and its syntactic category is noun.

The word in (58e) can inflect like an adjective and should therefore be classed as an adjective following our tests. Sometimes, these kinds of adjectives are also classed as adverbs. The reason for this is that the uninflected forms of these adjectives behave like adverbs:

- (61) Max lacht immer / oft / laut.
 Max laughs always often loud
 ‘Max (always/often) laughs (loudly).’

To capture this dual nature of words some researchers distinguish between lexical and syntactic category of words. The lexical category of *laut* ‘loud(ly)’ is that of an adjective and the syntactic category to which it belongs is ‘adverb’. The classification of adjectives such as *laut* ‘loud(ly)’ in (61) as adverbs is not assumed by all authors. Instead, some speak of adverbial usage of an adjective, that is, one assumes that the syntactic category is still adjective but it can be used in a different way so that it behaves like an adverb (see Eisenberg2004a, for example). This is parallel to prepositions, which can occur in a variety of syntactic contexts:

- (62) a. Peter schläft im Büro.
 Peter sleeps in.the office
 ‘Peter sleeps in the office.’
 b. der Tisch im Büro
 the table in.the office
 ‘the table in the office’

We have prepositional phrases in both examples in (62); however, in (62a) *im Büro* ‘in the office’ acts like an adverb in that it modifies the verb *schläft* ‘sleeps’ and in (62b) *im Büro* modifies the noun *Tisch* ‘table’. In the same way, *laut* ‘loud’ can modify a noun (63) or a verb (61).

- (63) die laute Musik
 the loud music

1.5 Heads

The head of a constituent/phrase is the element which determines the most important properties of the constituent/phrase. At the same time, the head also determines the composition of the phrase. That is, the head requires certain other

1 Introduction and basic terms

elements to be present in the phrase. The heads in the following examples have been marked in *italics*:

- (64) a. *Träumt* dieser Mann?
dreams this.NOM man
'Does this man dream?'
b. *Erwartet* er diesen Mann?
expects he.NOM this.ACC man
'Is he expecting this man?'
c. *Hilft* er diesem Mann?
helps he.NOM this.DAT man
'Is he helping this man?'
d. *in* diesem Haus
in this.DAT house
e. ein *Mann*
a.NOM man

Verbs determine the case of their arguments (subjects and objects). In (64d), the preposition determines which case the noun phrase *diesem Haus* 'this house' bears (dative) and also determines the semantic contribution of the phrase (it describes a location). (64e) is controversial: there are linguists who believe that the determiner is the head (VH77a-u; Hellan86a; Abney87a; Netter94; Netter98a) while others assume that the noun is the head of the phrase (vanLangendonck94a; ps2; Demske2001a; MuellerLehrbuch1; Hudson2004a; Bruening2009a).

The combination of a head with another constituent is called a *projection of the head*. A projection which contains all the necessary parts to create a well-formed phrase of that type is a *maximal projection*. A sentence is the maximal projection of a finite verb.

Figure 1.3 on the facing page shows the structure of (65) in box representation.

- (65) Der Mann liest einen Aufsatz.
the man reads an essay
'The man is reading an essay.'

Unlike Figure 1.1, the boxes have been labelled here.

The annotation includes the category of the most important element in the box. VP stands for *verb phrase* and NP for *noun phrase*. VP and NP are maximal projections of their respective heads.

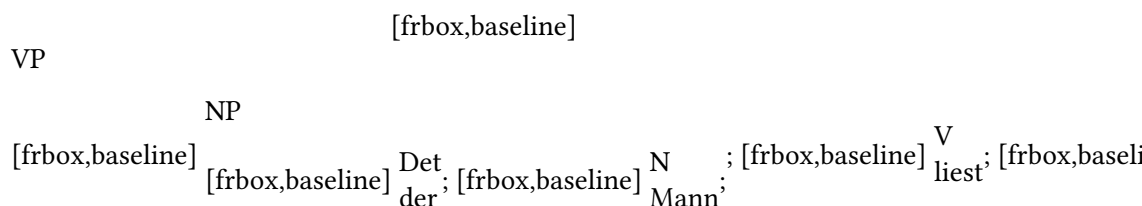


Figure 1.3: Words and phrases in annotated boxes

Anyone who has ever faced the hopeless task of trying to find particular photos of their sister's wedding in a jumbled, unsorted cupboard can vouch for the fact that it is most definitely a good idea to mark the boxes based on their content and also mark the albums based on the kinds of photos they contain.

An interesting point is that the exact content of the box with linguistic material does not play a role when the box is put into a larger box. It is possible, for example, to replace the noun phrase *der Mann* 'the man' with *er* 'he', or indeed the more complex *der Mann aus Stuttgart, der das Seminar zur Entwicklung der Zebrafincken besucht* 'the man from Stuttgart who takes part in the seminar on the development of zebra finches'. However, it is not possible to use *die Männer* 'the men' or *des Mannes* 'of the man' in this position:

- (66) a. *Die Männer liest einen Aufsatz.
 the men reads an essay
- b. *Des Mannes liest einen Aufsatz.
 of.the man.GEN reads an essay

The reason for this is that *die Männer* 'the men' is in plural and the verb *liest* 'reads' is in singular. The noun phrase bearing genitive case *des Mannes* can also not occur, only nouns in the nominative case. It is therefore important to mark all boxes with the information that is important for placing these boxes into larger boxes. Figure 1.4 shows our example with more detailed annotation.

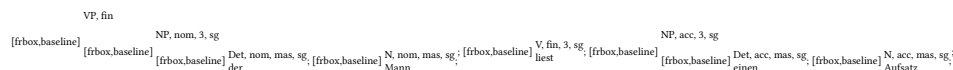


Figure 1.4: Words and word strings in annotated boxes

The features of a head which are relevant for determining in which contexts a phrase can occur are called *head features*. The features are said to be *projected* by the head.

1.6 Arguments and adjuncts

The constituents of a given clause have different relations to their head. It is typical to distinguish between arguments and adjuncts. The syntactic arguments of a head correspond for the most part to their logical arguments. We can represent the meaning of (67a) as (67b) using predicate logic.

- (67) a. Peter helps Maria.
b. $help'(peter', maria')$

The logical representation of (67b) resembles what is expressed in (67a); however, it abstracts away from constituent order and inflection. *Peter* and *Maria* are syntactic arguments of the verb *help* and their respective meanings (*Peter'* and *Maria'*) are arguments of the logical relation expressed by *help'*. One could also say that *help* assigns semantic roles to its arguments. Semantic roles include agent (the person carrying out an action), patient (the affected person or thing), beneficiary (the person who receives something) and experiencer (the person experiencing a psychological state). The subject of *help* is an agent and the direct object is a beneficiary. Arguments which fulfil a semantic role are also called *actants*. This term is also used for inanimate objects.

This kind of relation between a head and its arguments is covered by the terms *selection* and *valence*. Valence is a term borrowed from chemistry. Atoms can combine with other atoms to form molecules with varying levels of stability. The way in which the electron shells are occupied plays an important role for this stability. If an atom combines with others atoms so that its electron shell is fully occupied, then this will lead to a stable connection. Valence tells us something about the number of hydrogen atoms which an atom of a certain element can be combined with. In forming H_2O , oxygen has a valence of 2. We can divide elements into valence classes. Following Mendeleev, elements with a particular valence are listed in the same column in the periodic table.

The concept of valence was applied to linguistics by **Tesnière**^{59a-u}: a head needs certain arguments in order to form a stable compound. Words with the same valence – that is which require the same number and type of arguments – are divided into valence classes. Figure 1.5 shows examples from chemistry as well as linguistics.

Figure 1.5: Combination of hydrogen and oxygen and the combination of a verb with its arguments

We used (67) to explain logical valence. Logical valence can, however, some-

times differ from syntactic valence. This is the case with verbs like *rain*, which require an expletive pronoun as an argument. Inherently reflexive verbs such as *sich erholen* ‘to recover’ in German are another example.

- (68) a. Es regnet.
 it rains
 ‘It is raining.’
 b. Klaus erholt sich.
 Klaus recovers REFL
 ‘Klaus is recovering.’

The expletive *es* ‘it’ with weather verbs and the *sich* of so-called inherent reflexives such as *erholen* ‘to recover’ have to be present in the sentence. Germanic languages have expletive elements that are used to fill the position preceding the finite verb. These positional expletives are not realized in embedded clauses in German, since embedded clauses have a structure that differs from canonical unembedded declarative clauses, which have the finite verb in second position. (69a) shows that *es* cannot be omitted in *dass*”=clauses.

- (69) a. *Ich glaube, dass regnet.
 I think that rains
 Intended: ‘I think that it is raining.’
 b. *Ich glaube, dass Klaus erholt.
 I believe that Klaus recovers
 Intended: ‘I believe that Klaus is recovering.’

Neither the expletive nor the reflexive pronoun contributes anything semantically to the sentence. They must, however, be present to derive a complete, well-formed sentence. They therefore form part of the valence of the verb.

Constituents which do not contribute to the central meaning of their head, but rather provide additional information are called *adjuncts*. An example is the adverb *deeply* in (70):

- (70) John loves Mary deeply.

This says something about the intensity of the relation described by the verb. Further examples of adjuncts are attributive adjectives (71a) and relative clauses (71b):

- (71) a. a *beautiful* woman
 b. the man *who Mary loves*

1 Introduction and basic terms

Adjuncts have the following syntactic/semantic properties:

- (72) a. Adjuncts do not fulfil a semantic role.
- b. Adjuncts are optional.
- c. Adjuncts can be iterated.

The phrase in (71a) can be extended by adding another adjunct:

- (73) a beautiful clever woman

If one puts processing problems aside for a moment, this kind of extension by adding adjectives could proceed infinitely (see the discussion of (38) on page 76). Arguments, on the other hand, cannot be realized more than once:

- (74) * The man the boy sleeps.

If the entity carrying out the sleeping action has already been mentioned, then it is not possible to have another noun phrase which refers to a sleeping individual. If one wants to express the fact that more than one individual is sleeping, this must be done by means of coordination as in (75):

- (75) The man and the boy are sleeping.

One should note that the criteria for identifying adjuncts proposed in (72) is not sufficient, since there are also syntactic arguments that do not fill semantic roles (e.g., *es* 'it' in (68a) and *sich* (REFL) in (68b)) or are optional as *pizza* in (76).

- (76) Tony is eating (pizza).

Heads normally determine the syntactic properties of their arguments in a relatively fixed way. A verb is responsible for the case which its arguments bear.

- (77) a. Er gedenkt des Opfers.
 he remembers the.GEN victim.GEN
 'He remembers the victim.'
- b. *Er gedenkt dem Opfer.
 he remembers the.DAT victim
- c. Er hilft dem Opfer.
 he helps the.DAT victim
 'He helps the victim.'
- d. *Er hilft des Opfers.
 he helps the.GEN victim.GEN

The verb *governs* the case of its arguments.

The preposition and the case of the noun phrase in the prepositional phrase are both determined by the verb:¹⁴

- (78) a. Er denkt an seine Modelleisenbahn.
 he thinks on his.ACC model.railway
 ‘He is thinking of his model railway.’
 b. # Er denkt an seiner Modelleisenbahn.
 He thinks on his.DAT model.railway
 c. Er hängt an seiner Modelleisenbahn.
 He hangs on his.DAT model.railway
 ‘He clings to his model railway.’
 d. * Er hängt an seine Modelleisenbahn.
 he hangs on his.ACC model.railway

The case of noun phrases in modifying prepositional phrases, on the other hand, depends on their meaning. In German, directional prepositional phrases normally require a noun phrase bearing accusative case (79a), whereas local PPs (denoting a fixed location) appear in the dative case (79b):

- (79) a. Er geht in die Schule / auf den Weihnachtsmarkt / unter
 he goes in the.ACC school on the.ACC Christmas.market under
 die Brücke.
 the.ACC bridge
 ‘He is going to school/to the Christmas market/under the bridge.’
 b. Er schläft in der Schule / auf dem Weihnachtsmarkt / unter
 he sleeps in the.DAT school on the.DAT Christmas.market under
 der Brücke.
 the.DAT bridge
 ‘He is sleeping at school/at the Christmas market/under the bridge.’

An interesting case is the verb *sich befinden* ‘to be located’, which expresses the location of something. This cannot occur without some information about the location pertaining to the verb:

- (80) * Wir befinden uns.
 we are.located REFL

The exact form of this information is not fixed – neither the syntactic category nor the preposition inside of prepositional phrases is restricted:

¹⁴For similar examples, see Eisenberg94a.

- (81) Wir befinden uns hier / unter der Brücke / neben dem Eingang / im
we are REFL here under the bridge next.to the entrance in
Bett.
bed
'We are here/under the bridge/next to the entrance/in bed.'

Local modifiers such as *hier* 'here' or *unter der Brücke* 'under the bridge' are analyzed with regard to other verbs (e.g., *schlafen* 'sleep') as adjuncts. For verbs such as *sich befinden* 'to be (located)', we will most likely have to assume that information about location forms an obligatory syntactic argument of the verb.

The verb selects a phrase with information about location, but does not place any syntactic restrictions on its type. This specification of location behaves semantically like the other adjuncts we have seen previously. If I just consider the semantic aspects of the combination of a head and adjunct, then I also refer to the adjunct as a *modifier*.¹⁵ Arguments specifying location with verbs such as *sich befinden* 'to be located' are also subsumed under the term *modifier*. Modifiers are normally adjuncts, and therefore optional, whereas in the case of *sich befinden* they seem to be (obligatory) arguments.

In conclusion, we can say that constituents that are required to occur with a certain head are arguments of that head. Furthermore, constituents which fulfil a semantic role with regard to the head are also arguments. These kinds of arguments can, however, sometimes be optional.

Arguments are normally divided into subjects and complements.¹⁶ Not all heads require a subject (see **MuellerLehrbuch1**). The number of arguments of a head can therefore also correspond to the number of complements of a head.

1.7 Grammatical functions

In some theories, grammatical functions such as subject and object form part of the formal description of language (see Chapter ?? on Lexical Functional Grammar, for example). This is not the case for the majority of the theories discussed here, but these terms are used for the informal description of certain phenomena. For this reason, I will briefly discuss them in what follows.

¹⁵See Section 1.7.2 for more on the grammatical function of adverbials. The term adverbial is normally used in conjunction with verbs. *modifier* is a more general term, which normally includes attributive adjectives.

¹⁶In some schools the term complement is understood to include the subject, that is, the term complement is equivalent to the term argument (see for instance **Gross2003a**). Some researchers treat some subjects, e.g., those of finite verbs, as complements (**Pollard90a-Eng; Eisenberg94b**).

1.7.1 Subjects

Although I assume that the reader has a clear intuition about what a subject is, it is by no means a trivial matter to arrive at a definition of the word *subject* which can be used cross-linguistically. For German, Reis82 suggested the following syntactic properties as definitional for subjects:

- agreement of the finite verb with it
- nominative case in non-copular clauses
- omitted in infinitival clauses (control)
- optional in imperatives

I have already discussed agreement in conjunction with the examples in (4). Reis82 argues that the second bullet point is a suitable criterion for German. She formulates a restriction to non-copular clause because there can be more than one nominative argument in sentences with predicate nominals such as (82):

- (82) a. Er ist ein Lügner.
 he.NOM ist a liar.NOM
 ‘He is a liar.’
 b. Er wurde ein Lügner genannt.
 he.NOM was a liar.NOM called
 ‘He was called a liar.’

Following this criterion, arguments in the dative case such as *den Männern* ‘the men’ cannot be classed as subjects in German:

- (83) a. Er hilft den Männern.
 he helps the.DAT men.DAT
 ‘He is helping the men.’
 b. Den Männern wurde geholfen.
 the.DAT men.DAT were.3SG helped
 ‘The men were helped.’

Following the other criteria, datives should also not be classed as subjects – as Reis82 has shown. In (83b), *wurde*, which is the 3rd person singular form, does not agree with *den Männern*. The third of the aforementioned criteria deals with infinitive constructions such as those in (84):

1 Introduction and basic terms

- (84) a. Klaus behauptet, den Männern zu helfen.
Klaus claims the.DAT men.DAT to help
'Klaus claims to be helping the men.'
- b. Klaus behauptet, dass er den Männern hilft.
Klaus claims that he the.DAT men.DAT helps
'Klaus claims that he is helping the men.'
- c. * Die Männer behaupten, geholfen zu werden.
the men claim helped to become
Intended: 'The men are claiming to be helped.'
- d. * Die Männer behaupten, elegant getanzt zu werden.
the men claim elegantly danced to become
Intended: 'The men claim that there is elegant dancing.'

In the first sentence, an argument of the verb *helfen* 'to help' has been omitted. If one wishes to express it, then one would have to use the subordinate clause beginning with *dass* 'that' as in (84b). Examples (84c,d) show that infinitives which do not require a nominative argument cannot be embedded under verbs such as *behaupten* 'to claim'. If the dative noun phrase *den Männern* 'the men' were the subject in (83b), we would expect the control construction (84c) to be well-formed. This is, however, not the case. Instead of (84c), it is necessary to use (85):

- (85) Die Männer behaupten, dass ihnen geholfen wird.
the men.NOM claim that them.DAT helped becomes
'The men claim that they are being helped.'

In the same way, imperatives are not possible with verbs that do not require a nominative. (86) shows some examples from **Reis82**.

- (86) a. Fürchte dich nicht!
be.scared REFL not
'Don't be scared!'
- b. * Graue nicht!
dread not
'Don't dread it!'
- c. Werd einmal unterstützt und ...
be once supported and
'Let someone support you for once and ...'

- d. * Werd einmal geholfen und ...
 be once helped and
 ‘Let someone help you and ...’

The verb *sich fürchten* ‘to be scared’ in (86a) obligatorily requires a nominative argument as its subject (87a). The similar verb *grauen* ‘to dread’ in (86b) takes a dative argument (87b).

- (87) a. Ich fürchte mich vor Spinnen.
 I.NOM be.scared REFL before spiders
 ‘I am scared of spiders.’
 b. Mir graut vor Spinnen.
 me.DAT scares before spiders
 ‘I am dreading spiders.’

Interestingly, dative arguments in Icelandic behave differently. ZMT85a discuss various characteristics of subjects in Icelandic and show that it makes sense to describe dative arguments as subjects in passive sentences even if the finite verb does not agree with them (Section 3.1) or they do not bear nominative case. An example of this is infinitive constructions with an omitted dative argument (p. 457):

- (88) a. Ég vonast til að verða hjálpað.
 I hope for to be helped
 ‘I hope that I will be helped.’
 b. Að vera hjálpað í prófinu er óleyfilegt.
 to be helped on the.exam is not.allowed
 ‘It is not allowed for one to be helped during the exam.’

In a number of grammars, clausal arguments such as those in (89) are classed as subjects as they can be replaced by a noun phrase in the nominative (90) (see e.g., Eisenberg2004a).

- (89) a. Dass er schon um sieben kommen wollte, stimmt nicht.
 that he already at seven come wanted is.true not
 ‘It’s not true that he wanted to come as soon as seven.’
 b. Dass er Maria geheiratet hat, gefällt mir.
 that he Maria married has pleases me
 ‘I’m glad that he married Maria.’

- (90) a. Das stimmt nicht.
that is.true not
'That isn't true.'
- b. Das gefällt mir.
that pleases me
'I like that.'

It should be noted that there are different opinions on the question of whether clausal arguments should be treated as subjects or not. As recent publications show, there is still some discussion in Lexical Function Grammar (see Chapter ??) (DL2000a-u; Berman2003b-u; Berman2007a-u; AMM2005a-u; Forst2006a-u).

If we can be clear about what we want to view as a subject, then the definition of object is no longer difficult: objects are all other arguments whose form is directly determined by a given head. As well as clausal objects, German has genitive, dative, accusative and prepositional objects:

- (91) a. Sie gedenken des Mannes.
they remember the.GEN man.GEN
'They remember the man.'
- b. Sie helfen dem Mann.
they help the.DAT man.DAT
'They are helping the man.'
- c. Sie kennen den Mann.
they know the.ACC man.ACC
'They know the man.'
- d. Sie denken an den Mann.
they think on the man
'They are thinking of the man.'

As well as defining objects by their case, it is commonplace to talk of *direct objects* and *indirect objects*. The direct object gets its name from the fact that – unlike the indirect object – the referent of a direct object is directly affected by the action denoted by the verb. With ditransitives such as the German *geben* 'to give', the accusative object is the direct object and the dative is the indirect object.

- (92) dass er dem Mann den Aufsatz gibt
that he.NOM the.DAT man.DAT the.ACC essay.ACC gives
'that he gives the man the essay'

For trivalent verbs (verbs taking three arguments), we see that the verb can take either an object in the genitive case (93a) or, for verbs with a direct object in the accusative, a second accusative object (93b):

- (93) a. dass er den Mann des Mordes bezichtigte
 that he the.ACC man.ACC the.GEN murder.GEN accused
 ‘that he accused the man of murder’
 b. dass er den Mann den Vers lehrte
 that he the.ACC man.ACC the.ACC verse.ACC taught
 ‘that he taught the man the verse’

These kinds of objects are sometimes also referred to as indirect objects.

Normally, only those objects which are promoted to subject in passives with *werden* ‘to be’ are classed as direct objects. This is important for theories such as LFG (see Chapter ??) since passivization is defined with reference to grammatical function. With two-place verbal predicates, the dative is not normally classed as a direct object (Cook2006a-u).

- (94) dass er dem Mann hilft
 that he the.DAT man.DAT helps
 ‘that he helps the man’

In many theories, grammatical function does not form a primitive component of the theory, but rather corresponds to positions in a tree structure. The direct object in German is therefore the object which is first combined with the verb in a configuration assumed to be the underlying structure of German sentences. The indirect object is the second object to be combined with the verb. On this view, the dative object of *helfen* ‘to help’ would have to be viewed as a direct object.

In the following, I will simply refer to the case of objects and avoid using the terms direct object and indirect object.

In the same way as with subjects, we consider whether there are object clauses which are equivalent to a certain case and can fill the respective grammatical function of a direct or indirect object. If we assume that *dass du sprichst* ‘that you are speaking’ in (95a) is a subject, then the subordinate clause must be a direct object in (95b):

- (95) a. Dass du sprichst, wird erwähnt.
 that you speak is mentioned
 ‘The fact that you’re speaking is being mentioned.’

- b. Er erwähnt, dass du sprichst.
he mentions that you speak
'He mentions that you are speaking.'

In this case, we cannot really view the subordinate clause as the accusative object since it does not bear case. However, we can replace the sentence with an accusative-marked noun phrase:

- (96) Er erwähnt diesen Sachverhalt.
he mentions this.ACC matter
'He mentions this matter.'

If we want to avoid this discussion, we can simply call these arguments clausal objects.

1.7.2 The adverbial

Adverbials differ semantically from subjects and objects. They tell us something about the conditions under which an action or process takes place, or the way in which a certain state persists. In the majority of cases, adverbials are adjuncts, but there are – as we have already seen – a number of heads which also require adverbials. Examples of these are verbs such as *to be located* or *to make one's way*. For *to be located*, it is necessary to specify a location and for *to proceed to* a direction is needed. These kinds of adverbials are therefore regarded as arguments of the verb.

The term *adverbial* comes from the fact that adverbials are often adverbs. This is not the only possibility, however. Adjectives, participles, prepositional phrases, noun phrases and even sentences can be adverbials:

- (97) a. Er arbeitet sorgfältig.
he works carefully
b. Er arbeitet vergleichend.
he works comparatively
'He does comparative work.'
c. Er arbeitet in der Universität.
he works in the university
'He works at the university.'
d. Er arbeitet den ganzen Tag.
he works the whole day.ACC
'He works all day.'

- e. Er arbeitet, weil es ihm Spaß macht.
 he works because it him.DAT fun makes
 'He works because he enjoys it.'

Although the noun phrase in (97d) bears accusative case, it is not an accusative object. *den ganzen Tag* 'the whole day' is a so-called temporal accusative. The occurrence of accusative in this case has to do with the syntactic and semantic function of the noun phrase, it is not determined by the verb. These kinds of accusatives can occur with a variety of verbs, even with verbs that do not normally require an accusative object:

- (98) a. Er schläft den ganzen Tag.
 he sleeps the whole day
 'He sleeps the whole day.'
- b. Er liest den ganzen Tag diesen schwierigen Aufsatz.
 he reads the.ACC whole.ACC day this.ACC difficult.ACC essay
 'He spends the whole day reading this difficult essay.'
- c. Er gibt den Armen den ganzen Tag Suppe.
 he gives the.DAT poor.DAT the.ACC whole.ACC day soup
 'He spends the whole day giving soup to the poor.'

The case of adverbials does not change under passivization:

- (99) a. weil den ganzen Tag gearbeitet wurde
 because the.ACC whole.ACC day worked was
 'because someone worked all day'
- b. * weil der ganze Tag gearbeitet wurde
 because the.NOM whole.NOM day worked was

1.7.3 Predicatives

Adjectives like those in (100a,b) as well as noun phrases such as *ein Lügner* 'a liar' in (100c) are counted as predicatives.

- (100) a. Klaus ist *klug*.
 Klaus is clever
- b. Er isst den Fisch *roh*.
 he eats the fish raw

1 Introduction and basic terms

- c. Er ist *ein Lügner*.
he is a liar

In the copula construction in (100a,c), the adjective *klug* ‘clever’ and the noun phrase *ein Lügner* ‘a liar’ is an argument of the copula *sein* ‘to be’ and the depictive adjective in (100b) is an adjunct to *isst* ‘eats’.

For predicative noun phrases, case is not determined by the head but rather by some other element.¹⁷ For example, the accusative in (101a) becomes nominative under passivization (101b):

- (101) a. Sie nannte ihn einen Lügner.
she called him.ACC a.ACC liar
‘She called him a liar.’
b. Er wurde ein Lügner genannt.
he.NOM was a.NOM liar called
‘He was called a liar.’

Only *ihn* ‘him’ can be described as an object in (101a). In (101b), *ihn* becomes the subject and therefore bears nominative case. *einen Lügner* ‘a liar’ refers to

¹⁷There is some dialectal variation with regard to copula constructions: in Standard German, the case of the noun phrase with *sein* ‘to be’ is always nominative and does not change when embedded under *lassen* ‘to let’. According to **Duden95-Authors**, in Switzerland the accusative form is common which one finds in examples such as (ii.a).

- (i) a. Ich bin dein Tanzpartner.
I am your.NOM dancing.partner
b. Der wüste Kerl ist ihr Komplize.
the wild guy is her.NOM accomplice
c. Laß den wüsten Kerl [...] meinetwegen ihr Komplize sein.
let the.ACC wild.ACC guy for.all.I.care her.NOM accomplice be
‘Let’s assume that the wild guy is her accomplice, for all I care.’ (**Duden66-Authors**)
d. Baby, laß mich dein Tanzpartner sein.
baby let me.ACC your.NOM dancing.partner be
‘Baby, let me be your dancing partner!’ (Funny van Dannen, Benno-Ohnesorg-Theater, Berlin, Volksbühne, 11.10.1995)
- (ii) a. Er läßt den lieben Gott ‘n frommen Mann sein.
he lets the.ACC dear.ACC god a pious.ACC man be
‘He is completely lighthearted/unconcerned.’
b. *Er läßt den lieben Gott ‘n frommer Mann sein.
he lets the.ACC dear.ACC god a pious.NOM man be

ihn ‘him’ in (101a) and to *er* ‘he’ in (101b) and agrees in case with the noun over which it predicates. This is also referred to as *agreement case*.

For other predicative constructions see Duden2005-Authors and Mueller2002b and Mueller2008a.

1.7.4 Valence classes

It is possible to divide verbs into subclasses depending on how many arguments they require and on the properties these arguments are required to have. The classic division describes all verbs which have an object which becomes the subject under passivization as *transitive*. Examples of this are verbs such as *love* or *beat*. Intransitive verbs, on the other hand, are verbs which have either no object, or one that does not become the subject in passive sentences. Examples of this type of verb are *schlafen* ‘to sleep’, *helfen* ‘to help’, *gedenken* ‘to remember’. A subclass of transitive verbs are ditransitive verbs such as *geben* ‘to give’ and *zeigen* ‘to show’.

Unfortunately, this terminology is not always used consistently. Sometimes, two-place verbs with dative and genitive objects are also classed as transitive verbs. In this naming tradition, the terms intransitive, transitive and ditransitive are synonymous with one-place, two-place and three-place verbs.

The fact that this terminological confusion can lead to misunderstandings between even established linguistics is shown by Culicover and Jackendoff’s (CJ2005a) criticism of Chomsky. Chomsky states that the combination of the English auxiliary *be* + verb with passive morphology can only be used for transitive verbs. Culicover and Jackendoff claim that this cannot be true because there are transitive verbs such as *weigh* and *cost*, which cannot undergo passivization:

- (102) a. This book weighs ten pounds / costs ten dollars.
 b. * Ten pounds are weighed / ten dollar are cost by this book.

Culicover and Jackendoff use *transitive* in the sense of a verb requiring two arguments. If we only view those verbs whose object becomes the subject of a passive clause as transitive, then *weigh* and *cost* no longer count as transitive verbs and Culicover and Jackendoff’s criticism no longer holds.¹⁸ That noun phrases

¹⁸Their criticism also turns out to be unjust even if one views transitives as being two-place predicates. If one claims that a verb must take at least two arguments to be able to undergo passivization, one is not necessarily claiming that all verbs taking two or more arguments have to allow passivization. The property of taking multiple arguments is a condition which must be fulfilled, but it is by no means the only one.

such as those in (102) are no ordinary objects can also be seen by the fact they cannot be replaced by pronouns. It is therefore not possible to ascertain which case they bear since case distinctions are only realized on pronouns in English. If we translate the English examples into German, we find accusative objects:

- (103) a. Das Buch kostete einen Dollar.
the book costs one.ACC dollar
'The book costs one dollar.'
b. Das Buch wiegt einen Zentner.
the book weighs one.ACC centner
'The book weighs one centner.'

In the following, I will use *transitive* in the former sense, that is for verbs with an object that becomes the subject when passivized (e.g., with *werden* in German). When I talk about the class of verbs that includes *helfen* 'to help', which takes a nominative and dative argument, and *schlagen* 'to hit', which takes a nominative and accusative argument, I will use the term *two-place* or *bivalent verb*.

1.8 A topological model of the German clause

In this section, I introduce the concept of so-called *topological fields* (*topologische Felder*). These will be used frequently in later chapters to discuss different parts of the German clause. One can find further, more detailed introductions to topology in Reis80a, Hoehle86 and Askedal86. Woellstein2010a-u is a textbook about the topological field model.

1.8.1 The position of the verb

It is common practice to divide German sentences into three types pertaining to the position of the finite verb:

- verb-final clauses
- verb-first (initial) clauses
- verb-second (V2) clauses

The following examples illustrate these possibilities:

- (104) a. (Peter hat erzählt,) dass er das Eis gegessen hat.
Peter has told that he the ice.cream eaten has
'Peter said that he has eaten the ice cream.'

- b. *Hat* Peter das Eis gegessen?
has Peter the ice.cream eaten
'Has Peter eaten the ice cream?'
- c. Peter *hat* das Eis gegessen.
Peter has the ice.cream eaten
'Peter has eaten the ice cream.'

1.8.2 The sentence bracket, prefield, middle field and postfield

We observe that the finite verb *hat* 'has' is only adjacent to its complement *gegessen* 'eaten' in (104a). In (104b) and (104c), the verb and its complement are separated, that is, discontinuous. We can then divide the German clause into various sub-parts on the basis of these distinctions. In (104b) and (104c), the verb and the auxiliary form a "bracket" around the clause. For this reason, we call this the *sentence bracket* (*Satzklammer*). The finite verbs in (104b) and (104c) form the left bracket and the non-finite verbs form the right bracket. Clauses with verb-final order are usually introduced by conjunctions such as *weil* 'because', *dass* 'that' and *ob* 'whether'. These conjunctions occupy the same position as the finite verb in verb-initial or verb-second clauses. We therefore also assume that these conjunctions form the left bracket in these cases. Using the notion of the sentence bracket, it is possible to divide the structure of the German clause into the prefield (*Vorfeld*), middle field (*Mittelfeld*) and postfield (*Nachfeld*). The prefield describes everything preceding the left sentence bracket, the middle field is the section between the left and right bracket and the postfield describes the position after the right bracket. The Tables 1.1 and 1.2 give some examples of this. The right bracket can contain multiple verbs and is often referred to as a *verbal complex* or *verb cluster*. The assignment of question words and relative pronouns to the prefield will be discussed in the following section.

1.8.3 Assigning elements to fields

As the examples in the Tables 1.1 and 1.2 show, it is not required that all fields are always occupied. Even the left bracket can be empty if one opts to leave out the copula *sein* 'to be' such as in the examples in (105):

- (105) a. [...] egal, was noch passiert, der Norddeutsche Rundfunk
regardless what still happens the north.German broadcasting.company

1 Introduction and basic terms

steht schon jetzt als Gewinner fest.¹⁹

stands already now as winner PART

‘Regardless of what still may happen, the North German broadcasting company is already the winner.’

- b. Interessant, zu erwähnen, daß ihre Seele völlig in Ordnung
interesting to mention that her soul completely in order
war.²⁰

was

‘It is interesting to note that her soul was entirely fine.’

- c. Ein Treppenwitz der Musikgeschichte, daß die Kollegen von
an afterwit of.the history.of.music that the colleagues of

¹⁹Spiegel, 12/1999, p. 258.

²⁰Michail Bulgakow, *Der Meister und Margarita*. München: Deutscher Taschenbuch Verlag. 1997, p. 422.

Table 1.1: Examples of how topological fields can be occupied in declarative main clauses

Prefield Left b.		Middle field	Right b.	Postfield
Karl	schläft.			
Karl	sleeps			
Karl	hat		geschlafen.	
Karl	has		slept	
Karl	erkennt	Maria.		
Karl	recognizes	Maria		
Karl	färbt	den Mantel	um	den Maria kennt.
Karl	dies	the coat	PARTICLE	that Maria knows
Karl	hat	Maria	erkannt.	
Karl	has	Maria	recognized	
Karl	hat	Maria als sie aus dem Bus stieg sofort	erkannt.	
Karl	has	Maria when she got.off the bus immediately	recognized	
Karl	hat	Maria sofort	erkannt	als sie aus dem Bus stieg.
Karl	has	Maria immediately	recognized	when she got.off the bus
Karl	hat	Maria zu erkennen	behauptet.	
Karl	has	Maria to recognize	claimed	
Karl	hat		behauptet	Maria zu erkennen.
Karl	has		claimed	Maria to know

Table 1.2: Examples of how topological fields can be occupied in yes/no questions, imperatives, exclamatives and various verb final sentences including adverbial clauses, interrogative and relative clauses

Prefield	Left b.	Middle field	Right b.	Postfield
	Schläft sleeps	Karl? Karl		
	Schlaf! sleep			
	Iss eat	jetzt deinen Kuchen now your cake	auf! up	
	<i>Hat</i> has	er doch den ganzen Kuchen alleine he after.all the whole cake alone	gegessen! eaten	
	<i>weil</i> because	<i>er den ganzen Kuchen alleine</i> he the whole cake alone	<i>gegessen hat</i> eaten has	<i>ohne es zu bereuen</i> without it to regret
	<i>weil</i> because	<i>er den ganzen Kuchen alleine</i> he the whole cake	<i>essen können will</i> eat can wants.to	<i>ohne gestört zu werden</i> without distrurbed to be
<i>wer</i> who		<i>den ganzen Kuchen alleine</i> the whole cake alone	<i>gegessen hat</i> eaten has	
<i>der</i> who.REL		<i>den ganzen Kuchen alleine</i> the whole cake alone	<i>gegessen hat</i> eaten has	
<i>mit wem</i> with whom		<i>du</i> you	<i>geredet hast</i> spoken have	
<i>mit dem</i> with whom.REL		<i>du</i> you	<i>geredet hast</i> spoken have	

Rammstein vor fünf Jahren noch im Vorprogramm von Sandow
Rammstein before five years still in.the.pre.programme of Sandow
spielten.²¹
played

‘One of the little ironies of music history is that five years ago their
colleagues of Rammstein were still an opening act for Sandow.’

The examples in (105) correspond to those with the copula in (106):

- (106) a. Egal ist, was noch passiert, ...
regardless is what still happens
‘It is not important what still may happen ...’

²¹Flüstern & Schweigen, taz, 12.07.1999, p. 14.

- b. Interessant ist zu erwähnen, dass ihre Seele völlig in Ordnung
interesting is to mention that her soul completely in order
war.
was
'It is interesting to note that her soul was completely fine.'
- c. Ein Treppenwitz der Musikgeschichte ist, dass die Kollegen von
an afterwit of.the music.history is that the colleagues of
Rammstein vor fünf Jahren noch im Vorprogramm von Sandow
Rammstein before five years still in pre.programme of Sandow
spielten.
played
'It is one of the little ironies of music history that five years ago their
colleagues of Rammstein were still an opening act for Sandow.'

When fields are empty, it is sometimes not clear which fields are occupied by certain constituents. For the examples in (105), one would have to insert the copula to be able to ascertain that a single constituent is in the prefield and, furthermore, which fields are occupied by the other constituents.

In the following example taken from **Paul1919a**, inserting the copula obtains a different result:

- (107) a. Niemand da?
nobody there
b. Ist niemand da?
is nobody there
'Is nobody there?'

Here we are dealing with a question and *niemand* 'nobody' in (107a) should therefore not be analyzed as in the prefield but rather the middle field.

In (108), there are elements in the prefield, the left bracket and the middle field. The right bracket is empty.²²

- (108) Er gibt der Frau das Buch, die er kennt.
he.M gives the woman(F) the book.(N) that.F he knows
'He gives the book to the woman that he knows.'

How should we analyze relative clauses such as *die er kennt* 'that he knows'? Do they form part of the middle field or the postfield? This can be tested using a test

²²The sentence requires emphasis on *der* 'the'. *der Frau, die er kennt* 'the woman' is contrasted with another woman or other women.

developed by **Bech55a** (*Rangprobe*): first, we modify the example in (108) so that it is in the perfect. Since non-finite verb forms occupy the right bracket, we can clearly see the border between the middle field and postfield. The examples in (109) show that the relative clause cannot occur in the middle field unless it is part of a complex constituent with the head noun *Frau* ‘woman’.

- (109) a. Er hat [der Frau] das Buch gegeben, [die er kennt].
 he has the woman the book given that he knows
 ‘He has given the book to the woman that he knows.’
 b. *Er hat [der Frau] das Buch, [die er kennt,] gegeben.
 he has the woman the book that he knows given
 c. Er hat [der Frau, die er kennt,] das Buch gegeben.
 he has the woman that he knows the book given

This test does not help if the relative clause is realized together with its head noun at the end of the sentence as in (110):

- (110) Er gibt das Buch der Frau, die er kennt.
 he gives the book the woman that he knows
 ‘He gives the book to the woman that he knows.’

If we put the example in (110) in the perfect, then we observe that the lexical verb can occur before or after the relative clause:

- (111) a. Er hat das Buch [der Frau] gegeben, [die er kennt].
 he has the book the woman given that he knows
 ‘He has given the book to the woman he knows.’
 b. Er hat das Buch [der Frau, die er kennt,] gegeben.
 he has the book the woman that he knows given

In (111a), the relative clause has been extraposed. In (111b) it forms part of the noun phrase *der Frau, die er kennt* ‘the woman that he knows’ and therefore occurs inside the NP in the middle field. It is therefore not possible to rely on this test for (110). We assume that the relative clause in (110) also belongs to the NP since this is the most simple structure. If the relative clause were in the postfield, we would have to assume that it has undergone extraposition from its position inside the NP. That is, we would have to assume the NP”=structure anyway and then extraposition in addition.

We have a similar problem with interrogative and relative pronouns. Depending on the author, these are assumed to be in the left bracket (**Kathol2001a**; **Duerscheid2003a-u**; **Eisenberg2004a**; **Pafel2011a-u**) or the prefield (**Duden2005-Authors**;

Woellstein2010a-u) or even in the middle field (**AH2004a-u**). In Standard German interrogative or relative clauses, both fields are never simultaneously occupied. For this reason, it is not immediately clear to which field an element belongs. Nevertheless, we can draw parallels to main clauses: the pronouns in interrogative and relative clauses can be contained inside complex phrases:

- (112) a. der Mann, [mit dem] du gesprochen hast
 the man with whom you spoken have
 ‘the man you spoke to’
 b. Ich möchte wissen, [mit wem] du gesprochen hast.
 I want.to know with whom you spoken have
 ‘I want to know who you spoke to.’

Normally, only individual words (conjunctions or verbs) can occupy the left bracket,²³ whereas words and phrases can appear in the prefield. It therefore makes sense to assume that interrogative and relative pronouns (and phrases containing them) also occur in this position.

Furthermore, it can be observed that the dependency between the elements in the *Vorfeld* of declarative clauses and the remaining sentence is of the same kind as the dependency between the phrase that contains the relative pronoun and the remaining sentence. For instance, *über dieses Thema* ‘about this topic’ in (113a) depends on *Vortrag* ‘talk’, which is deeply embedded in the sentence: *einen Vortrag* ‘a talk’ is an argument of *zu halten* ‘to hold’, which in turn is an argument of *gebeten* ‘asked’.

- (113) a. Über dieses Thema habe ich ihn gebeten, einen Vortrag zu halten.
 about this topic have I him asked a talk to hold
 ‘I asked him to give a talk about this topic.’
 b. das Thema, über das ich ihn gebeten habe, einen Vortrag zu
 the topic about which I him asked have a talk to
 halten
 hold
 ‘the topic about which I asked him to give a talk’

²³Coordination is an exception to this:

- (i) Er [kennt und liebt] diese Schallplatte.
 he knows and loves this record
 ‘He knows and loves this record.’

The situation is similar in (113b): the relative phrase *über das* ‘about which’ is a dependent of *Vortrag* ‘talk’ which is realized far away from it. Thus, if the relative phrase is assigned to the *Vorfeld*, it is possible to say that such nonlocal frontings always target the *Vorfeld*.

Finally, the Duden grammar (Duden2005-Authors) provides the following examples from non-standard German (mainly southern dialects):

- (114) a. Kommt drauf an, mit wem dass sie zu tun haben.
comes there.upon PART with whom that you to do have
‘It depends on whom you are dealing with.’
- (115) a. Lotti, die wo eine tolle Sekretärin ist, hat ein paar merkwürdige
Lotti who where a great secretary is has a few strange
Herren empfangen.
gentlemen welcomed
‘Lotti, who is a great secretary, welcomed a few strange gentlemen.’
b. Du bist der beste Sänger, den wo ich kenn.
you are the best singer who where I know
‘You are the best singer whom I know.’

These examples of interrogative and relative clauses show that the left sentence bracket is filled with a conjunction (*dass* ‘that’ or *wo* ‘where’ in the respective dialects). So if one wants to have a model that treats Standard German and the dialectal forms uniformly, it is reasonable to assume that the relative phrases and interrogative phrases are located in the *Vorfeld*.

1.8.4 Recursion

As already noted by Reis80a, when occupied by a complex constituent, the pre-field can be subdivided into further fields including a postfield, for example. The constituents *für lange lange Zeit* ‘for a long, long time’ in (116b) and *daß du kommst* ‘that you are coming’ in (116d) are inside the prefield but occur to the right of the right bracket *verschüttet* ‘buried’ / *gewußt* ‘knew’, that is they are in the postfield of the prefield.

- (116) a. Die Möglichkeit, etwas zu verändern, ist damit verschüttet
the possibility something to change is there.with buried
für lange lange Zeit.
for long long time
‘The possibility to change something will now be gone for a long, long

time.'

- b. [Verschüttet für lange lange Zeit] ist damit die Möglichkeit,
buried for long long time ist there.with the possibility
etwas zu ver"-ändern.
something to change
- c. Wir haben schon seit langem gewußt, daß du kommst.
we have PART since long known that you come
'We have known for a while that you are coming.'
- d. [Gewußt, daß du kommst,] haben wir schon seit langem.
known that you come have we PART since long

Like constituents in the prefield, elements in the middle field and postfield can also have an internal structure and be divided into subfields accordingly. For example, *daß* 'that' is the left bracket of the subordinate clause *daß du kommst* in (116c), whereas *du* 'you' occupies the middle field and *kommst* 'come' the right bracket.



Comprehension questions

1. How does the head of a phrase differ from non"-heads?
2. What is the head in the examples in (117)?

- (117)
- a. he
 - b. Go!
 - c. quick

3. How do arguments differ from adjuncts?
4. Identify the heads, arguments and adjuncts in the following sentence (118) and in the subparts of the sentence:

- (118) Er hilft den kleinen Kindern in der Schule.
he helps the small children in the school
'He helps small children at school.'

5. How can we define the terms prefield (*Vorfeld*), middle field (*Mittelfeld*), postfield (*Nachfeld*) and the left and right sentence brackets (*Satzklammer*)?



Exercises

1. Identify the sentence brackets, prefield, middle field and postfield in the following sentences. Do the same for the embedded clauses!

- (119) a. Karl isst.
Karl eats
'Karl is eating.'
- b. Der Mann liebt eine Frau, den Peter kennt.
the man loves a woman who Peter knows
'The man who Peter knows loves a woman.'
- c. Der Mann liebt eine Frau, die Peter kennt.
the man loves a woman that Peter knows
'The man loves a woman who Peter knows.'
- d. Die Studenten haben behauptet, nur wegen der Hitze
the students have claimed only because of the heat
einzuschlafen.
to fall asleep
'The students claimed that they were only falling asleep because of the heat.'
- e. Dass Peter nicht kommt, ärgert Klaus.
that Peter not comes annoys Klaus
'(The fact) that Peter isn't coming annoys Klaus.'
- f. Einen Mann küssen, der ihr nicht gefällt, würde sie nie.
a man kiss that her not pleases would she never
'She would never kiss a man she doesn't like.'



Further reading

Reis80a gives reasons for why field theory is important for the description of the position of constituents in German.

Hoehle86 discusses fields to the left of the prefield, which are needed for left”=dislocation structures such as with *der Mittwoch* in (120), *aber* in (121a) and *denn* in (121b):

- (120) Der Mittwoch, der passt mir gut.
the Wednesday that fits me good
‘Wednesday, that suits me fine.’
- (121) a. Aber würde denn jemand den Hund füttern morgen Abend?
but would PART anybody the dog feed tomorrow evening
‘But would anyone feed the dog tomorrow evening?’
b. Denn dass es regnet, damit rechnet keiner.
because that it rains there.with reckons nobody
‘Because no-one expects that it will rain.’

Höhle also discusses the historical development of field theory.

2 Phrase structure grammar

This chapter deals with phase structure grammars (PSGs), which play an important role in several of the theories we will encounter in later chapters.

2.1 Symbols and rewrite rules

Words can be assigned to a particular part of speech on the basis of their inflectional properties and syntactic distribution. Thus, *weil* ‘because’ in (1) is a conjunction, whereas *das* ‘the’ and *dem* ‘the’ are articles and therefore classed as determiners. Furthermore, *Buch* ‘book’ and *Mann* ‘man’ are nouns and *gibt* ‘gives’ is a verb.

- (1) *weil er das Buch dem Mann gibt*
because he the book the man gives
‘because he gives the man the book’

Using the constituency tests we introduced in Section 1.3, we can show that individual words as well as the strings *das Buch* ‘the book’ and *dem Mann* ‘the man’, form constituents. These get then assigned certain symbols. Since nouns form an important part of the phrases *das Buch* and *dem Mann*, these are referred to as *noun phrases* or NPs, for short. The pronoun *er* ‘he’ can occur in the same positions as full NPs and can therefore also be assigned to the category NP.

Phrase structure grammars come with rules specifying which symbols are assigned to certain kinds of words and how these are combined to create more complex units. A simple phrase structure grammar which can be used to analyze (1) is given in (2):^{1,2}

¹I ignore the conjunction *weil* ‘because’ for now. Since the exact analysis of German verb”=first and verb”=second clauses requires a number of additional assumptions, we will restrict ourselves to verb”=final clauses in this chapter.

²The rule $NP \rightarrow er$ may seem odd. We could assume the rule $PersPron \rightarrow er$ instead but then would have to posit a further rule which would specify that personal pronouns can replace full NPs: $NP \rightarrow PersPron$. The rule in (2) combines the two aforementioned rules and states that *er* ‘he’ can occur in positions where noun phrases can.

2 Phrase structure grammar

- | | | | |
|-----|----------------------------------|-------------------------------------|-----------------------------|
| (2) | $NP \rightarrow \text{Det } N$ | $NP \rightarrow \text{er}$ | $N \rightarrow \text{Buch}$ |
| | $S \rightarrow NP \ NP \ NP \ V$ | $\text{Det} \rightarrow \text{das}$ | $N \rightarrow \text{Mann}$ |
| | | $\text{Det} \rightarrow \text{dem}$ | $V \rightarrow \text{gibt}$ |

We can therefore interpret a rule such as $NP \rightarrow \text{Det } N$ as meaning that a noun phrase, that is, something which is assigned the symbol NP, can consist of a determiner (Det) and a noun (N).

We can analyze the sentence in (1) using the grammar in (2) in the following way: first, we take the first word in the sentence and check if there is a rule in which this word occurs on the right-hand side of the rule. If this is the case, then we replace the word with the symbol on the left-hand side of the rule. This happens in lines 2–4, 6–7 and 9 of the derivation in (3). For instance, in line 2 *er* is replaced by NP. If there are two or more symbols which occur together on the right-hand side of a rule, then all these words are replaced with the symbol on the left. This happens in lines 5, 8 and 10. For instance, in line 5 and 8, Det and N are rewritten as NP.

(3)	words and symbols						rules that are applied
1	er	das	Buch	dem	Mann	gibt	
2	NP	das	Buch	dem	Mann	gibt	$NP \rightarrow \text{er}$
3	NP	Det	Buch	dem	Mann	gibt	$\text{Det} \rightarrow \text{das}$
4	NP	Det	N	dem	Mann	gibt	$N \rightarrow \text{Buch}$
5	NP		NP	dem	Mann	gibt	$NP \rightarrow \text{Det } N$
6	NP		NP	Det	Mann	gibt	$\text{Det} \rightarrow \text{dem}$
7	NP		NP	Det	N	gibt	$N \rightarrow \text{Mann}$
8	NP		NP		NP	gibt	$NP \rightarrow \text{Det } N$
9	NP		NP		NP	V	$V \rightarrow \text{gibt}$
10						S	$S \rightarrow NP \ NP \ NP \ V$

In (3), we began with a string of words and it was shown that we can derive the structure of a sentence by applying the rules of a given phrase structure grammar. We could have applied the same steps in reverse order: starting with the sentence symbol S, we would have applied the steps 9–1 and arrived at the string of words. Selecting different rules from the grammar for rewriting symbols, we could use the grammar in (2) to get from S to the string *er dem Mann das Buch gibt* ‘he the man the book gives’. We can say that this grammar licenses (or generates) a set of sentences.

The derivation in (3) can also be represented as a tree. This is shown by Figure 2.1 on the next page. The symbols in the tree are called *nodes*. We say that S immediately dominates the NP nodes and the V node. The other nodes in the tree

Figure 2.1: Analysis of *er das Buch dem Mann gibt* ‘he the book the man gives’

are also dominated, but not immediately dominated, by S. If we want to talk about the relationship between nodes, it is common to use kinship terms. In Figure 2.1, S is the *mother node* of the three NP nodes and the V node. The NP node and V are *sisters* since they have the same mother node. If a node has two daughters, then we have a binary branching structure. If there is exactly one daughter, then we have a unary branching structure. Two constituents are said to be *adjacent* if they are directly next to each other.

Phrase structure rules are often omitted in linguistic publications. Instead, authors opt for tree diagrams or the compact equivalent bracket notation such as (4).

- (4) $[_S [_{NP} \text{er}] [_{NP} [_{Det} \text{das}] [_N \text{Buch}]] [_{NP} [_{Det} \text{dem}] [_N \text{Mann}]] [_V \text{gibt}]]$
 he the book the man gives

Nevertheless, it is the grammatical rules which are actually important since these represent grammatical knowledge which is independent of specific structures. In this way, we can use the grammar in (2) to parse or generate the sentence in (5), which differs from (1) in the order of objects:

- (5) [weil] er dem Mann das Buch gibt
 because he.NOM the.DAT man the.ACC book gives
 ‘because he gives the man the book’

The rules for replacing determiners and nouns are simply applied in a different order than in (1). Rather than replacing the first Det with *das* ‘the’ and the first noun with *Buch* ‘book’, the first Det is replaced with *dem* ‘the’ and the first noun with *Mann*.

At this juncture, I should point out that the grammar in (2) is not the only possible grammar for the example sentence in (1). There is an infinite number of possible grammars which could be used to analyze these kinds of sentences (see exercise ??). Another possible grammar is given in (6):

- (6) $NP \rightarrow Det\ N$ $NP \rightarrow er$ $N \rightarrow Buch$
 $V \rightarrow NP\ V$ $Det \rightarrow das$ $N \rightarrow Mann$
 $Det \rightarrow dem$ $V \rightarrow gibt$

This grammar licenses binary branching structures as shown in Figure 2.2 on the following page.

Figure 2.2: Analysis of *er das Buch dem Mann gibt* with a binary branching structure

Both the grammar in (6) and (2) are too imprecise. If we adopt additional lexical entries for *ich* ‘I’ and *den* ‘the’ (accusative) in our grammar, then we would incorrectly license the ungrammatical sentences in (7b–d):³

- (7) a. er das Buch dem Mann gibt
 he.NOM the.ACC book the.DAT man gives
 ‘He gives the book to the man.’
 b. *ich das Buch dem Mann gibt
 I.NOM the.ACC book the.DAT man gives
 c. *er das Buch den Mann gibt
 he.NOM the.ACC book the.ACC man gives
 d. *er den Buch dem Mann gibt
 he.NOM the.M book(N) the man gives

In (7b), subject=verb agreement has been violated, in other words: *ich* ‘I’ and *gibt* ‘gives’ do not fit together. (7c) is ungrammatical because the case requirements of the verb have not been satisfied: *gibt* ‘gives’ requires a dative object. Finally, (7d) is ungrammatical because there is a lack of agreement between the determiner and the noun. It is not possible to combine *den* ‘the’, which is masculine and bears accusative case, and *Buch* ‘book’ because *Buch* is neuter gender. For this reason, the gender properties of these two elements are not the same and the elements can therefore not be combined.

In the following, we will consider how we would have to change our grammar to stop it from licensing the sentences in (7b–d). If we want to capture subject=verb agreement, then we have to cover the following six cases in German,

³With the grammar in (6), we also have the additional problem that we cannot determine when an utterance is complete since the symbol V is used for all combinations of V and NP. Therefore, we can also analyze the sentence in (i) with this grammar:

- (i) a. *der Mann erwartet
 the man expects
 b. *des Mannes er das Buch dem Mann gibt
 the.GEN man.GEN he.NOM the.ACC book the.DAT man gives

The number of arguments required by a verb must be somehow represented in the grammar. In the following chapters, we will see exactly how the selection of arguments by a verb (valence) can be captured in various grammatical theories.

as the verb has to agree with the subject in both person (1, 2, 3) and number (sg, pl):

- (8) a. Ich schlafe. (1, sg)
I sleep
b. Du schläfst. (2, sg)
you sleep
c. Er schläft. (3, sg)
he sleeps
d. Wir schlafen. (1, pl)
we sleep
e. Ihr schlaft. (2, pl)
you sleep
f. Sie schlafen. (3, pl)
they sleep

It is possible to capture these relations with grammatical rules by increasing the number of symbols we use. Instead of the rule $S \rightarrow NP\ NP\ NP\ V$, we can use the following:

- (9) $S \rightarrow NP_1_sg\ NP\ NP\ V_1_sg$
 $S \rightarrow NP_2_sg\ NP\ NP\ V_2_sg$
 $S \rightarrow NP_3_sg\ NP\ NP\ V_3_sg$
 $S \rightarrow NP_1_pl\ NP\ NP\ V_1_pl$
 $S \rightarrow NP_2_pl\ NP\ NP\ V_2_pl$
 $S \rightarrow NP_3_pl\ NP\ NP\ V_3_pl$

This would mean that we need six different symbols for noun phrases and verbs respectively, as well as six rules rather than one.

In order to account for case assignment by the verb, we can incorporate case information into the symbols in an analogous way. We would then get rules such as the following:

- (10) $S \rightarrow NP_1_sg_nom\ NP_dat\ NP_acc\ V_1_sg_nom_dat_acc$
 $S \rightarrow NP_2_sg_nom\ NP_dat\ NP_acc\ V_2_sg_nom_dat_acc$
 $S \rightarrow NP_3_sg_nom\ NP_dat\ NP_acc\ V_3_sg_nom_dat_acc$
 $S \rightarrow NP_1_pl_nom\ NP_dat\ NP_acc\ V_1_pl_nom_dat_acc$
 $S \rightarrow NP_2_pl_nom\ NP_dat\ NP_acc\ V_2_pl_nom_dat_acc$
 $S \rightarrow NP_3_pl_nom\ NP_dat\ NP_acc\ V_3_pl_nom_dat_acc$

Since it is necessary to differentiate between noun phrases in four cases, we have a total of six symbols for NPs in the nominative and three symbols for NPs with other cases. Since verbs have to match the NPs, that is, we have to differentiate between verbs which select three arguments and those selecting only one or two (11), we have to increase the number of symbols we assume for verbs.

- (11) a. Er schläft.
he sleeps
'He is sleeping.'
- b. * Er schläft das Buch.
he sleeps the book
- c. Er kennt das Buch.
he knows the book
'He knows the book.'
- d. * Er kennt.
he knows

In the rules above, the information about the number of arguments required by a verb is included in the marking 'nom_dat_acc'.

In order to capture the determiner=noun agreement in (12), we have to incorporate information about gender (fem, mas, neu), number (sg, pl), case (nom, gen, dat, acc) and the inflectional classes (strong, weak)⁴.

- (12) a. der Mann, die Frau, das Buch (gender)
the.M man(M) the.F woman(F) the.N book(N)
- b. das Buch, die Bücher (number)
the book.SG the books.PL
- c. des Buches, dem Buch (case)
the.GEN book.GEN the.DAT book
- d. ein Beamter, der Beamte (inflectional class)
a civil.servant the civil.servant

Instead of the rule $NP \rightarrow Det N$, we will have to use rules such as those in (13):⁵

⁴These are inflectional classes for adjectives which are also relevant for some nouns such as *Beamter* 'civil servant', *Verwandter* 'relative', *Gesandter* 'envoy'. For more on adjective classes see page 25.

⁵To keep things simple, these rules do not incorporate information regarding the inflection class.

- (13) NP_3_sg_nom → Det_fem_sg_nom N_fem_sg_nom
 NP_3_sg_nom → Det_mas_sg_nom N_mas_sg_nom
 NP_3_sg_nom → Det_neu_sg_nom N_neu_sg_nom
 NP_3_pl_nom → Det_fem_pl_nom N_fem_pl_nom
 NP_3_pl_nom → Det_mas_pl_nom N_mas_pl_nom
 NP_3_pl_nom → Det_neu_pl_nom N_neu_pl_nom
- NP_3_sg_nom → Det_fem_sg_nom N_fem_sg_nom
 NP_3_sg_nom → Det_mas_sg_nom N_mas_sg_nom
 NP_3_sg_nom → Det_neu_sg_nom N_neu_sg_nom
 NP_3_pl_nom → Det_fem_pl_nom N_fem_pl_nom
 NP_3_pl_nom → Det_mas_pl_nom N_mas_pl_nom
 NP_3_pl_nom → Det_neu_pl_nom N_neu_pl_nom

(13) shows the rules for nominative noun phrases. We would need analogous rules for genitive, dative, and accusative. We would then require 24 symbols for determiners ($3 * 2 * 4$), 24 symbols for nouns and 24 rules rather than one. If inflection class is taken into account, the number of symbols and the number of rules doubles.

2.2 Expanding PSG with features

Phrase structure grammars which only use atomic symbols are problematic as they cannot capture certain generalizations. We as linguists can recognize that NP_3_sg_nom stands for a noun phrase because it contains the letters NP. However, in formal terms this symbol is just like any other symbol in the grammar and we cannot capture the commonalities of all the symbols used for NPs. Furthermore, unstructured symbols do not capture the fact that the rules in (13) all have something in common. In formal terms, the only thing that the rules have in common is that there is one symbol on the left-hand side of the rule and two on the right.

We can solve this problem by introducing features which are assigned to category symbols and therefore allow for the values of such features to be included in our rules. For example, we can assume the features person, number and case for the category symbol NP. For determiners and nouns, we would adopt an additional feature for gender and one for inflectional class. (14) shows two rules augmented by the respective values in brackets:⁶

⁶Chapter 6 introduces attribute value structures. In these structure we always have pairs of a

2 *Phrase structure grammar*

- (14) $\text{NP}(3, \text{sg}, \text{nom}) \rightarrow \text{Det}(\text{fem}, \text{sg}, \text{nom}) \text{N}(\text{fem}, \text{sg}, \text{nom})$
 $\text{NP}(3, \text{sg}, \text{nom}) \rightarrow \text{Det}(\text{mas}, \text{sg}, \text{nom}) \text{N}(\text{mas}, \text{sg}, \text{nom})$

If we were to use variables rather than the values in (14), we would get rule schemata as the one in (15):

- (15) $\text{NP}(3, \text{Num}, \text{Case}) \rightarrow \text{Det}(\text{Gen}, \text{Num}, \text{Case}) \text{N}(\text{Gen}, \text{Num}, \text{Case})$

The values of the variables here are not important. What is important is that they match. For this to work, it is important that the values are ordered; that is, in the category of a determiner, the gender is always first, number second and so on. The value of the person feature (the first position in the $\text{NP}(3, \text{Num}, \text{Case})$) is fixed at ‘3’ by the rule. These kind of restrictions on the values can, of course, be determined in the lexicon:

- (16) $\text{NP}(3, \text{sg}, \text{nom}) \rightarrow \text{es}$
 $\text{Det}(\text{mas}, \text{sg}, \text{nom}) \rightarrow \text{des}$

The rules in (10) can be collapsed into a single schema as in (17):

- (17) $S \rightarrow \text{NP}(\text{Per1}, \text{Num1}, \text{nom})$
 $\text{NP}(\text{Per2}, \text{Num2}, \text{dat})$
 $\text{NP}(\text{Per3}, \text{Num3}, \text{acc})$
 $\text{V}(\text{Per1}, \text{Num1}, \text{ditransitive})$

The identification of Per1 and Num1 on the verb and on the subject ensures that there is subject=verb agreement. For the other NPs, the values of these features are irrelevant. The case of these NPs is explicitly determined.

2.3 **Semantics**

In the introductory chapter and the previous sections, we have been dealing with syntactic aspects of language and the focus will remain very much on syntax for the remainder of this book. It is, however, important to remember that we use language to communicate, that is, to transfer information about certain situations, topics or opinions. If we want to accurately explain our capacity for language, then we also have to explain the meanings that our utterances have. To this end, it is necessary to understand their syntactic structure, but this alone

feature name and a feature value. In such a setting, the order of values is not important, since every value is uniquely identified by the corresponding feature name. Since we do not have a feature name in schemata like (13), the order of the values is important.

is not enough. Furthermore, theories of language acquisition that only concern themselves with the acquisition of syntactic constructions are also inadequate. The syntax=semantics interface is therefore important and every grammatical theory has to say something about how syntax and semantics interact. In the following, I will show how we can combine phrase structure rules with semantic information. To represent meanings, I will use first=order predicate logic and λ =calculus. Unfortunately, it is not possible to provide a detailed discussion of the basics of logic so that even readers without prior knowledge can follow all the details, but the simple examples discussed here should be enough to provide some initial insights into how syntax and semantics interact and furthermore, how we can develop a linguistic theory to account for this.

To show how the meaning of a sentence is derived from the meaning of its parts, we will consider (18a). We assign the meaning in (18b) to the sentence in (18a).

- (18) a. Max schläft.
 Max sleeps
 ‘Max is sleeping.’
 b. *schlafen'*(*max'*)

Here, we are assuming *schlafen'* to be the meaning of *schläft* ‘sleeps’. We use prime symbols to indicate that we are dealing with word meanings and not actual words. At first glance, it may not seem that we have really gained anything by using *schlafen'* to represent the meaning of (18a), since it is just another form of the verb *schläft* ‘sleeps’. It is, however, important to concentrate on a single verb form as inflection is irrelevant when it comes to meaning. We can see this by comparing the examples in (19a) and (19b):

- (19) a. Jeder Junge schläft.
 every boy sleeps
 ‘Every boy sleeps.’
 b. Alle Jungen schlafen.
 all boys sleep
 ‘All boys sleep.’

To enhance readability I use English translations of the predicates in semantic representations from now on.⁷ So the meaning of (18a) is represented as (20) rather than (18b):

⁷Note that I do not claim that English is suited as representation language for semantic relations and concepts that can be expressed in other languages.

(20) $sleep'(max')$

When looking at the meaning in (20), we can consider which part of the meaning comes from each word. It seems relatively intuitive that max' comes from *Max*, but the trickier question is what exactly *schläft* ‘sleeps’ contributes in terms of meaning. If we think about what characterizes a ‘sleeping’ event, we know that there is typically an individual who is sleeping. This information is part of the meaning of the verb *schlafen* ‘to sleep’. The verb meaning does not contain information about the sleeping individual, however, as this verb can be used with various subjects:

- (21) a. Paul schläft.
Paul sleeps
‘Paul is sleeping.’
b. Mio schläft.
Mio sleeps
‘Mio is sleeping.’
c. Xaver schläft.
Xaver sleeps
‘Xaver is sleeping.’

We can therefore abstract away from any specific use of $sleep'$ and instead of, for example, max' in (20), we use a variable (e.g., x). This x can then be replaced by $paul'$, mio' or $xaver'$ in a given sentence. To allow us to access these variables in a given meaning, we can write them with a λ in front. Accordingly, *schläft* ‘sleeps’ will have the following meaning:

(22) $\lambda x sleep'(x)$

The step from (20) to (22) is referred to as *lambda abstraction*. The combination of the expression (22) with the meaning of its arguments happens in the following way: we remove the λ and the corresponding variable and then replace all instances of the variable with the meaning of the argument. If we combine (22) and max' as in (23), we arrive at the meaning in (20), namely $sleep'(max')$.

(23) $\lambda x sleep'(x) max'$

The process is called β ”=reduction or λ ”=conversion. To show this further, let us consider an example with a transitive verb. The sentence in (24a) has the meaning given in (24b):

- (24) a. Max mag Lotte.
 Max likes Lotte
 ‘Max likes Lotte.’
 b. $like'(max', lotte')$

The λ ”=abstraction of *mag* ‘likes’ is shown in (25):

- (25) $\lambda y \lambda x like'(x, y)$

Note that it is always the first λ that has to be used first. The variable y corresponds to the object of *mögen* ‘to like’. For languages like English it is assumed that the object forms a verb phrase (VP) together with the verb and this VP is combined with the subject. German differs from English in allowing more freedom in constituent order. The problems that result for form meaning mappings are solved in different ways by different theories. The respective solutions will be addressed in the following chapters.

If we combine the representation in (25) with that of the object *Lotte*, we arrive at (26a), and following β ”=reduction, (26b):

- (26) a. $\lambda y \lambda x like'(x, y)lotte'$
 b. $\lambda x like'(x, lotte')$

This meaning can in turn be combined with the subject and we then get (27a) and (27b) after β ”=reduction:

- (27) a. $\lambda x like'(x, lotte')max'$
 b. $like'(max', lotte')$

After introducing lambda calculus, integrating the composition of meaning into our phrase structure rules is simple. A rule for the combination of a verb with its subject has to be expanded to include positions for the semantic contribution of the verb, the semantic contribution of the subject and then the meaning of the combination of these two (the entire sentence). The complete meaning is the combination of the individual meanings in the correct order. We can therefore take the simple rule in (28a) and turn it into (28b):

- (28) a. $S \rightarrow NP(nom) V$
 b. $S(V' NP') \rightarrow NP(nom, NP') V(V')$

V' stands for the meaning of V and NP' for the meaning of the $NP(nom)$. $V' NP'$ stands for the combination of V' and NP' . When analyzing (18a), the meaning of V' is $\lambda x sleep'(x)$ and the meaning of NP' is max' . The combination of $V' NP'$ corresponds to (29a) or after β ”=reduction to (18b) – repeated here as (29b):

- (29) a. $\lambda x \text{ sleep}'(x) \text{max}'$
 b. $\text{sleep}'(\text{max}')$

For the example with a transitive verb in (24a), the rule in (30) can be proposed:

- (30) $S(V' \text{ NP2}' \text{ NP1}') \rightarrow \text{NP}(\text{nom}, \text{NP1}') \text{ V}(V') \text{ NP}(\text{acc}, \text{NP2}')$

The meaning of the verb (V') is first combined with the meaning of the object ($\text{NP2}'$) and then with the meaning of the subject ($\text{NP1}'$).

At this point, we can see that there are several distinct semantic rules for the phrase structure rules above. The hypothesis that we should analyze language in this way is called the *rule="to"=rule hypothesis* (Bach76a). A more general process for deriving the meaning of linguistic expression will be presented in Section 5.1.4.

2.4 Phrase structure rules for some aspects of German syntax

Whereas determining the direct constituents of a sentence is relative easy, since we can very much rely on the movement test due to the somewhat flexible order of constituents in German, it is more difficult to identify the parts of the noun phrase. This is the problem we will focus on in this section. To help motivate assumptions about \bar{X} syntax to be discussed in Section 2.5, we will also discuss prepositional phrases.

2.4.1 Noun phrases

Up to now, we have assumed a relatively simple structure for noun phrases: our rules state that a noun phrase consists of a determiner and a noun. Noun phrases can have a distinctly more complex structure than (31a). This is shown by the following examples in (31):

- (31) a. eine Frau
 a woman
 b. eine Frau, die wir kennen
 a woman who we know
 c. eine Frau aus Stuttgart
 a woman from Stuttgart

2.4 Phrase structure rules for some aspects of German syntax

- d. eine kluge Frau
a smart woman
- e. eine Frau aus Stuttgart, die wir kennen
a woman from Stuttgart who we know
- f. eine kluge Frau aus Stuttgart
a smart woman from Stuttgart
- g. eine kluge Frau, die wir kennen
a smart woman who we know
- h. eine kluge Frau aus Stuttgart, die wir kennen
a smart woman from Stuttgart who we know

As well as determiners and nouns, noun phrases can also contain adjectives, prepositional phrases and relative clauses. The additional elements in (31) are adjuncts. They restrict the set of objects which the noun phrase refers to. Whereas (31a) refers to a being which has the property of being a woman, the referent of (31b) must also have the property of being known to us.

Our previous rules for noun phrases simply combined a noun and a determiner and can therefore only be used to analyze (31a). The question we are facing now is how we can modify this rule or which additional rules we would have to assume in order to analyze the other noun phrases in (31). In addition to rule (32a), one could propose a rule such as the one in (32b).^{8,9}

- (32) a. $NP \rightarrow Det\ N$
b. $NP \rightarrow Det\ A\ N$

However, this rule would still not allow us to analyze noun phrases such as (33):

- (33) alle weiteren schlagkräftigen Argumente
all further strong arguments
'all other strong arguments'

In order to be able to analyze (33), we require a rule such as (34):

- (34) $NP \rightarrow Det\ A\ A\ N$

It is always possible to increase the number of adjectives in a noun phrase and setting an upper limit for adjectives would be entirely arbitrary. Even if we opt for the following abbreviation, there are still problems:

⁸See Eisenberg2004a for the assumption of flat structures in noun phrases.

⁹There are, of course, other features such as gender and number, which should be part of all the rules discussed in this section. I have omitted these in the following for ease of exposition.

(35) $NP \rightarrow \text{Det } A^* N$

The asterisk in (35) stands for any number of iterations. Therefore, (35) encompasses rules with no adjectives as well as those with one, two or more.

The problem is that according to the rule in (35) adjectives and nouns do not form a constituent and we can therefore not explain why coordination is still possible in (36):

(36) alle [[großen Seeelefanten] und [grauen Eichhörnchen]]
 all big elephant.seals and grey squirrels
 ‘all the big elephant seals and grey squirrels’

If we assume that coordination involves the combination of two or more word strings with the same syntactic properties, then we would have to assume that the adjective and noun form a unit.

The following rules capture the noun phrases with adjectives discussed thus far:

(37) a. $NP \rightarrow \text{Det } \bar{N}$
 b. $\bar{N} \rightarrow A \bar{N}$
 c. $\bar{N} \rightarrow N$

These rules state the following: a noun phrase consists of a determiner and a nominal element (\bar{N}). This nominal element can consist of an adjective and a nominal element (37b), or just a noun (37c). Since \bar{N} is also on the right-hand side of the rule in (37b), we can apply this rule multiple times and therefore account for noun phrases with multiple adjectives such as (33). Figure 2.3 shows the structure of a noun phrase without an adjective and that of a noun phrase with one or two adjectives. The adjective *grau* ‘grey’ restricts the set of referents

Figure 2.3: Noun phrases with differing numbers of adjectives

for the noun phrase. If we assume an additional adjective such as *groß* ‘big’, then it only refers to those squirrels who are grey as well as big. These kinds of noun phrases can be used in contexts such as the following:

(38) A: Alle grauen Eichhörnchen sind groß.
 all grey squirrels are big
 ‘All grey squirrels are big.’ B: Nein, ich habe ein kleines graues Eichhörnchen
 no I have a small grey squirrel

gesehen.

seen

‘No, I saw a small grey squirrel.’

We observe that this discourse can be continued with *Aber alle kleinen grauen Eichhörnchen sind krank* ‘but all small grey squirrels are ill’ and a corresponding answer. The possibility to have even more adjectives in noun phrases such as *ein kleines graues Eichhörnchen* ‘a small grey squirrel’ is accounted for in our rule system in (37). In the rule (37b), \bar{N} occurs on the left as well as the right”=hand side of the rule. This kind of rule is referred to as *recursive*.

We have now developed a nifty little grammar that can be used to analyze noun phrases containing adjectival modifiers. As a result, the combination of an adjective and noun is given constituent status. One may wonder at this point if it would not make sense to also assume that determiners and adjectives form a constituent, as we also have the following kind of noun phrases:

- (39) diese schlaun und diese neugierigen Frauen
these smart and these curious women

Here, we are dealing with a different structure, however. Two full NPs have been conjoined and part of the first conjunct has been deleted.

- (40) diese schlaun ~~Frauen~~ und diese neugierigen Frauen
these smart women and these curious women

One can find similar phenomena at the sentence and even word level:

- (41) a. dass Peter dem Mann das Buch ~~gibt~~ und Maria der Frau die
that Peter the man the book gives and Maria the woman the
Schallplatte gibt
record gives
‘that Peters gives the book to the man and Maria the record to the
woman’
b. be- und ent”=laden
PRT and PRT”=load
‘load and unload’

Thus far, we have discussed how we can ideally integrate adjectives into our rules for the structure of noun phrases. Other adjuncts such as prepositional phrases or relative clauses can be combined with \bar{N} in an analogous way to adjectives:

- (42) a. $\bar{N} \rightarrow \bar{N} PP$

- b. $\bar{N} \rightarrow \bar{N}$ relative clause

With these rules and those in (37), it is possible – assuming the corresponding rules for PPs and relative clauses – to analyze all the examples in (31).

(37c) states that it is possible for \bar{N} to consist of a single noun. A further important rule has not yet been discussed: we need another rule to combine nouns such as *Vater* ‘father’, *Sohn* ‘son’ or *Bild* ‘picture’, so-called *relational nouns*, with their arguments. Examples of these can be found in (43a–b). (43c) is an example of a nominalization of a verb with its argument:

- (43) a. der Vater von Peter
the father of Peter
‘Peter’s father’
b. das Bild vom Gleimtunnel
the picture of.the Gleimtunnel
‘the picture of the Gleimtunnel’
c. das Kommen des Installateurs
the coming of.the plumber
‘the plumber’s visit’

The rule that we need to analyze (43a,b) is given in (44):

- (44) $\bar{N} \rightarrow N PP$

Figure 2.4 shows two structures with PP”=arguments. The tree on the right also contains an additional PP”=adjunct, which is licensed by the rule in (42a).

Figure 2.4: Combination of a noun with PP complement *vom Gleimtunnel* to the right with an adjunct PP

In addition to the previously discussed NP structures, there are other structures where the determiner or noun is missing. Nouns can be omitted via ellipsis. (45) gives an example of noun phrases, where a noun that does not require a complement has been omitted. The examples in (46) show NPs in which only one determiner and complement of the noun has been realized, but not the noun itself. The underscore marks the position where the noun would normally occur.

- (45) a. eine kluge _
a smart
‘a smart one’

2.4 Phrase structure rules for some aspects of German syntax

- b. eine kluge große _
 a smart tall
 ‘a smart tall one’
- c. eine kluge _ aus Hamburg
 a smart from Hamburg
 ‘a smart one from Hamburg’
- d. eine kluge _, die alle kennen
 a smart who everyone knows
 ‘a smart one who everyone knows’
- (46) a. (Nein, nicht der Vater von Klaus), der _ von Peter war gemeint.
 no not the father of Klaus the of Peter was meant
 ‘No, it wasn’t the father of Klaus, but rather the one of Peter that was meant.’
- b. (Nein, nicht das Bild von der Stadtautobahn), das _ vom Gleimtunnel
 no not the picture of the motorway the of the Gleimtunnel
 war beeindruckend.
 was impressive
 ‘No, it wasn’t the picture of the motorway, but rather the one of the Gleimtunnel that was impressive.’
- c. (Nein, nicht das Kommen des Tischlers), das _ des Installateurs
 no not the coming of the carpenter the of the plumber
 ist wichtig.
 is important
 ‘No, it isn’t the visit of the carpenter, but rather the visit of the plumber that is important.’

In English, the pronoun *one* must often be used in the corresponding position,¹⁰ but in German the noun is simply omitted. In phrase structure grammars, this can be described by a so-called *epsilon production*. These rules replace a symbol with nothing (47a). The rule in (47b) is an equivalent variant which is responsible for the term *epsilon production*:

- (47) a. $N \rightarrow$
 b. $N \rightarrow \epsilon$

The corresponding trees are shown in Figure 2.5 on the following page. Going

¹⁰See FLGR2012a for English examples without the pronoun *one*.

Figure 2.5: Noun phrases without an overt head

back to boxes, the rules in (47) correspond to empty boxes with the same labels as the boxes of ordinary nouns. As we have considered previously, the actual content of the boxes is unimportant when considering the question of where we can incorporate them. In this way, the noun phrases in (31) can occur in the same sentences. The empty noun box also behaves like one with a genuine noun. If we do not open the empty box, we will not be able to ascertain the difference to a filled box.

It is not only possible to omit the noun from noun phrases, but the determiner can also remain unrealized in certain contexts. (48) shows noun phrases in plural:

- (48)
- a. Frauen
women
 - b. Frauen, die wir kennen
women who we know
 - c. kluge Frauen
smart women
 - d. kluge Frauen, die wir kennen
smart women who we know

The determiner can also be omitted in singular if the noun denotes a mass noun:

- (49)
- a. Getreide
grain
 - b. Getreide, das gerade gemahlen wurde
grain that just ground was
'grain that has just been ground'
 - c. frisches Getreide
fresh grain
 - d. frisches Getreide, das gerade gemahlen wurde
fresh grain that just ground was
'fresh grain that has just been ground'

Finally, both the determiner and the noun can be omitted:

2.4 Phrase structure rules for some aspects of German syntax

- (50) a. Ich helfe klugen.
I help smart
'I help smart ones.'
- b. Dort drüben steht frisches, das gerade gemahlen wurde.
there over stands fresh that just ground was
'Over there is some fresh (grain) that has just been ground.'

Figure 2.6 shows the corresponding trees.

Figure 2.6: Noun phrases without overt determiner

It is necessary to add two further comments to the rules we have developed up to this point: up to now, I have always spoken of adjectives. However, it is possible to have very complex adjective phrases in pre”=nominal position. These can be adjectives with complements (51a,b) or adjectival participles (51c,d):

- (51) a. der seiner Frau treue Mann
the his.DAT wife faithful man
'the man faithful to his wife'
- b. der auf seinen Sohn stolze Mann
the on his.ACC son proud man
'the man proud of his son'
- c. der seine Frau liebende Mann
the his.ACC woman loving man
'the man who loves his wife'
- d. der von seiner Frau geliebte Mann
the by his.DAT wife loved man
'the man loved by his wife'

Taking this into account, the rule (37b) has to be modified in the following way:

- (52) $\bar{N} \rightarrow AP \bar{N}$

An adjective phrase (AP) can consist of an NP and an adjective, a PP and an adjective or just an adjective:

- (53) a. $AP \rightarrow NP A$
b. $AP \rightarrow PP A$
c. $AP \rightarrow A$

2 Phrase structure grammar

There are two imperfections resulting from the rules we have developed thus far. These are the rules for adjectives or nouns without complements in (53c) as well as (37c) – repeated here as (54):

$$(54) \quad \bar{N} \rightarrow N$$

If we apply these rules, then we will generate unary branching subtrees, that is trees with a mother that only has one daughter. See Figure 2.6 for an example of this. If we maintain the parallel to the boxes, this would mean that there is a box which contains another box which is the one with the relevant content.

In principle, nothing stops us from placing this information directly into the larger box. Instead of the rules in (55), we will simply use the rules in (56):

$$(55) \quad \begin{array}{l} \text{a. } A \rightarrow \text{kluge} \\ \text{b. } N \rightarrow \text{Mann} \end{array}$$

$$(56) \quad \begin{array}{l} \text{a. } AP \rightarrow \text{kluge} \\ \text{b. } \bar{N} \rightarrow \text{Mann} \end{array}$$

(56a) states that *kluge* ‘smart’ has the same properties as a full adjective phrase, in particular that it cannot be combined with a complement. This is parallel to the categorization of the pronoun *er* ‘he’ as an NP in the grammars (2) and (6).

Assigning \bar{N} to nouns which do not require a complement has the advantage that we do not have to explain why the analysis in (57b) is possible as well as (57a) despite there not being any difference in meaning.

$$(57) \quad \begin{array}{l} \text{a. } [_{NP} \text{ einige } [_{\bar{N}} \text{ kluge } [_{\bar{N}} [_{N} \text{ Frauen }] \text{ und } [_{\bar{N}} [_{N} \text{ Männer }]]]]] \\ \quad \text{some} \quad \text{smart} \quad \text{women} \quad \text{and} \quad \text{men} \\ \text{b. } [_{NP} \text{ einige } [_{\bar{N}} \text{ kluge } [_{\bar{N}} [_{N} [_{N} \text{ Frauen }] \text{ und } [_{N} \text{ Männer }]]]]] \\ \quad \text{some} \quad \text{smart} \quad \text{women} \quad \text{and} \quad \text{men} \end{array}$$

In (57a), two nouns have projected to \bar{N} and have then been joined by coordination. The result of coordination of two constituents of the same category is always a new constituent with that category. In the case of (57a), this is also \bar{N} . This constituent is then combined with the adjective and the determiner. In (57b), the nouns themselves have been coordinated. The result of this is always another constituent which has the same category as its parts. In this case, this would be N. This N becomes \bar{N} and is then combined with the adjective. If nouns which do not require complements were categorized as \bar{N} rather than N, we would not have the problem of spurious ambiguities. The structure in (58) shows the only possible analysis.

- (58) $[_{NP} \text{ einige } [_{N} \text{ kluge } [_{N} \text{ Frauen }] \text{ und } [_{N} \text{ Männer }]]]]$
 some smart women and men

2.4.2 Prepositional phrases

Compared to the syntax of noun phrases, the syntax of prepositional phrases (PPs) is relatively straightforward. PPs normally consist of a preposition and a noun phrase whose case is determined by that preposition. We can capture this with the following rule:

- (59) $PP \rightarrow P \text{ NP}$

This rule must, of course, also contain information about the case of the NP. I have omitted this for ease of exposition as I did with the NP"-rules and AP"-rules above.

The Duden grammar (Duden2005-Authors) offers examples such as those in (60), which show that certain prepositional phrases serve to further define the semantic contribution of the preposition by indicating some measurement, for example:

- (60) a. $[[\text{Einen Schritt}] \text{ vor } \text{dem Abgrund}] \text{ blieb } \text{er stehen.}$
 one step before the abyss remained he stand
 ‘He stopped one step in front of the abyss.’
 b. $[[\text{Kurz}] \text{ nach dem Start}] \text{ fiel die Klimaanlage } \text{aus.}$
 shortly after the take.off fell the air.conditioning out
 ‘Shortly after take off, the air conditioning stopped working.’
 c. $[[\text{Schräg}] \text{ hinter der Scheune}] \text{ ist ein Weiher.}$
 diagonally behind the barn is a pond
 ‘There is a pond diagonally across from the barn.’
 d. $[[\text{Mitten}] \text{ im } \text{Urwald}] \text{ stießen die Forscher } \text{auf einen alten}$
 middle in.the jungle stumbled the researchers on an old
 Tempel.
 temple
 ‘In the middle of the jungle, the researches came across an old temple.’

To analyze the sentences in (60a,b), one could propose the following rules in (61):

- (61) a. $PP \rightarrow NP \text{ PP}$
 b. $PP \rightarrow AP \text{ PP}$

These rules combine a PP with an indication of measurement. The resulting constituent is another PP. It is possible to use these rules to analyze prepositional phrases in (60a,b), but it unfortunately also allows us to analyze those in (62):

- (62) a. * [_{PP} einen Schritt [_{PP} kurz [_{PP} vor dem Abgrund]]]
 one step shortly before the abyss
 b. * [_{PP} kurz [_{PP} einen Schritt [_{PP} vor dem Abgrund]]]
 shortly one step before the abyss

Both rules in (61) were used to analyze the examples in (62). Since the symbol PP occurs on both the left and right-hand side of the rules, we can apply the rules in any order and as many times as we like.

We can avoid this undesired side-effect by reformulating the previously assumed rules:

- (63) a. $PP \rightarrow NP \bar{P}$
 b. $PP \rightarrow AP \bar{P}$
 c. $PP \rightarrow \bar{P}$
 d. $\bar{P} \rightarrow P NP$

Rule (59) becomes (63d). The rule in (63c) states that a PP can consist of \bar{P} . Figure 2.7 shows the analysis of (64) using (63c) and (63d) as well as the analysis of an example with an adjective in the first position following the rules in (63b) and (63d):

- (64) vor dem Abgrund
 before the abyss
 ‘in front of the abyss’

Figure 2.7: Prepositional phrases with and without measurement

At this point, the attentive reader is probably wondering why there is no empty measurement phrase in the left figure of Figure 2.7, which one might expect in analogy to the empty determiner in Figure 2.6. The reason for the empty determiner in Figure 2.6 is that the entire noun phrase without the determiner has a meaning similar to those with a determiner. The meaning normally contributed by the visible determiner has to somehow be incorporated in the structure of the noun phrase. If we did not place this meaning in the empty determiner, this would lead to more complicated assumptions about semantic combination: we only really require the mechanisms presented in Section 2.3 and these are very

general in nature. The meaning is contributed by the words themselves and not by any rules. If we were to assume a unary branching rule such as that in the left tree in Figure 2.7 instead of the empty determiner, then this unary branching rule would have to provide the semantics of the determiner. This kind of analysis has also been proposed by some researchers. See Chapter ?? for more on empty elements.

Unlike determiner-less NPs, prepositional phrases without an indication of degree or measurement do not lack any meaning component for composition. It is therefore not necessary to assume an empty indication of measurement, which somehow contributes to the meaning of the entire PP. Hence, the rule in (63c) states that a prepositional phrase consists of \bar{P} , that is, a combination of P and NP.

2.5 \bar{X} theory

If we look again at the rules that we have formulated in the previous section, we see that heads are always combined with their complements to form a new constituent (65a,b), which can then be combined with further constituents (65c,d):

- (65) a. $\bar{N} \rightarrow N \text{ PP}$
 b. $\bar{P} \rightarrow P \text{ NP}$
 c. $\text{NP} \rightarrow \text{Det } \bar{N}$
 d. $\text{PP} \rightarrow \text{NP } \bar{P}$

Grammarians working on English noticed that parallel structures can be used for phrases which have adjectives or verbs as their head. I discuss adjective phrases at this point and postpone the discussion of verb phrases to Chapter 3. As in German, certain adjectives in English can take complements with the important restriction that adjective phrases with complements cannot realize these pre-nominally in English. (66) gives some examples of adjective phrases:

- (66) a. He is proud.
 b. He is very proud.
 c. He is proud of his son.
 d. He is very proud of his son.

Unlike prepositional phrases, complements of adjectives are normally optional. *proud* can be used with or without a PP. The degree expression *very* is also optional.

The rules which we need for this analysis are given in (67), with the corresponding structures in Figure 2.8.

- (67) a. $AP \rightarrow \bar{A}$
 b. $AP \rightarrow AdvP \bar{A}$
 c. $\bar{A} \rightarrow A PP$
 d. $\bar{A} \rightarrow A$

Figure 2.8: English adjective phrases

As was shown in Section 2.2, it is possible to generalize over very specific phrase structure rules and thereby arrive at more general rules. In this way, properties such as person, number and gender are no longer encoded in the category symbols, but rather only simple symbols such as NP, Det and N are used. It is only necessary to specify something about the values of a feature if it is relevant in the context of a given rule. We can take this abstraction a step further: instead of using explicit category symbols such as N, V, P and A for lexical categories and NP, VP, PP and AP for phrasal categories, one can simply use a variable for the word class in question and speak of X and XP.

This form of abstraction can be found in so-called \bar{X} theory (or X'' =bar theory, the term *bar* refers to the line above the symbol), which was developed by Chomsky70a and refined by Jackendoff77a. This form of abstract rules plays an important role in many different theories. For example: Government & Binding (Chapter 3), Generalized Phrase Structure Grammar (Chapter 5) and Lexical Functional Grammar (Chapter ??). In HPSG (Chapter ??), \bar{X} theory also plays a role, but not all restrictions of the \bar{X} schema have been adopted.

(68) shows a possible instantiation of \bar{X} rules, where the category X has been used in place of N, as well as examples of word strings which can be derived by these rules:

(68) \bar{X} rule	with specific categories	example strings
$\bar{X} \rightarrow \text{specifier } \bar{X}$	$\bar{N} \rightarrow \text{DET } \bar{N}$	the [picture of Paris]
$\bar{X} \rightarrow \bar{X} \text{ adjunct}$	$\bar{N} \rightarrow \bar{N} \text{ REL_CLAUSE}$	[picture of Paris] [that everybody knows]
$\bar{X} \rightarrow \text{adjunct } \bar{X}$	$\bar{N} \rightarrow \bar{A} \bar{N}$	beautiful [picture of Paris]
$\bar{X} \rightarrow X \text{ complement*}$	$\bar{N} \rightarrow N \bar{P}$	picture [of Paris]

Any word class can replace X (e.g., V , A or P). The X without the bar stands for a lexical item in the above rules. If one wants to make the bar level explicit, then it is possible to write X^0 . Just as with the rule in (15), where we did not specify the case value of the determiner or the noun but rather simply required that the values on the right”=hand side of the rule match, the rules in (68) require that the word class of an element on the right”=hand side of the rule (X or \bar{X}) matches that of the element on the left”=hand side of the rule (\bar{X} or $\bar{\bar{X}}$).

A lexical element can be combined with all its complements. The “*” in the last rule stands for an unlimited amount of repetitions of the symbol it follows. A special case is zero”fold occurrence of complements. There is no PP complement of *Bild* ‘picture’ present in *das Bild* ‘the picture’ and thus N becomes \bar{N} . The result of the combination of a lexical element with its complements is a new projection level of X : the projection level 1, which is marked by a bar. \bar{X} can then be combined with adjuncts. These can occur to the left or right of \bar{X} . The result of this combination is still \bar{X} , that is the projection level is not changed by combining it with an adjunct. Maximal projections are marked by two bars. One can also write XP for a projection of X with two bars. An XP consists of a specifier and \bar{X} . Depending on one’s theoretical assumptions, subjects of sentences (Haider95b-u; Haider97a; Berman2003a) and determiners in NPs (Chomsky70a) are specifiers. Furthermore, degree modifiers (Chomsky70a) in adjective phrases and measurement indicators in prepositional phrases are also counted as specifiers.

Non”=head positions can only host maximal projections and therefore complements, adjuncts and specifiers always have two bars. Figure 2.9 gives an overview of the minimal and maximal structure of phrases.

Figure 2.9: Minimal and maximal structure of phrases

Some categories do not have a specifier or have the option of having one. Adjuncts are optional and therefore not all structures have to contain an \bar{X} with an adjunct daughter. In addition to the branching shown in the right”=hand figure, adjuncts to XP and head”=adjuncts are sometimes possible. There is only a single rule in (68) for cases in which a head precedes the complements, however an order in which the complement precedes the head is of course also possible. This is shown in Figure 2.9.

Figure 2.10 on the next page shows the analysis of the NP structures *das Bild* ‘the picture’ and *das schöne Bild von Paris* ‘the beautiful picture of Paris’. The NP structures in Figure 2.10 and the tree for *proud* in Figure 2.8 show examples of minimally populated structures. The left tree in Figure 2.10 is also an example of

a structure without an adjunct. The right”=hand structure in Figure 2.10 is an example for the maximally populated structure: specifier, adjunct, and complement are present.

Figure 2.10: \bar{X} analysis of *das Bild* ‘the picture’ and *das schöne Bild von Paris* ‘the beautiful picture of Paris’

The analysis given in Figure 2.10 assumes that all non”=heads in a rule are phrases. One therefore has to assume that there is a determiner phrase even if the determiner is not combined with other elements. The unary branching of determiners is not elegant but it is consistent.¹¹ The unary branchings for the NP *Paris* in Figure 2.10 may also seem somewhat odd, but they actually become more plausible when one considers more complex noun phrases:

- (69) a. *das Paris der dreißiger Jahre*
 the Paris of.the thirty years
 ‘30’s Paris’
 b. *die Maria aus Hamburg*
 the Maria from Hamburg
 ‘Maria from Hamburg’

Unary projections are somewhat inelegant but this should not concern us too much here, as we have already seen in the discussion of the lexical entries in (56) that unary branching nodes can be avoided for the most part and that it is indeed desirable to avoid such structures. Otherwise, one gets spurious ambiguities. In the following chapters, we will discuss approaches such as Categorical Grammar and HPSG, which do not assume unary rules for determiners, adjectives and nouns.

Furthermore, other \bar{X} theoretical assumptions will not be shared by several theories discussed in this book. In particular, the assumption that non”=heads always have to be maximal projections will be disregarded. Pullum85a and KP90a have shown that the respective theories are not necessarily less restrictive than theories which adopt a strict version of the \bar{X} theory. See also the discussion in Section ??.

¹¹For an alternative version of \bar{X} theory which does not assume elaborate structure for determiners see Muysken82a.



Comprehension questions

1. Why are phrase structure grammars that use only atomic categories inadequate for the description of natural languages?
2. Assuming the grammar in (6), state which steps (replacing symbols) one has to take to get to the symbol V in the sentence (70).

(70) er das Buch dem Mann gibt
 he the book the man gives
 'He gives the book to the man.'

Your answer should resemble the analysis in (3).

3. Give a representation of the meaning of (71) using predicate logic:

(71) a. Ulrike kennt Hans.
 Ulrike knows Hans
 b. Joshi freut sich.
 Joshi is.happy REFL
 'Joshi is happy.'



Exercises

1. On page 65, I claimed that there is an infinite number of grammars we could use to analyze (1). Why is this claim correct?
2. Try to come up with some ways in which we can tell which of these possible grammars is or are the best?
3. A fragment for noun phrase syntax was presented in Section 2.4.1. Why is the interaction of the rules in (72) problematic?

- (72)
- a. $NP \rightarrow Det \bar{N}$
 - b. $\bar{N} \rightarrow N$
 - c. $Det \rightarrow \epsilon$
 - d. $N \rightarrow \epsilon$

4. Why is it not a good idea to mark *books* as NP in the lexicon?
5. Can you think of some reasons why it is not desirable to assume the following rule for nouns such as *books*:

- (73) $NP \rightarrow \text{Modifier}^* \text{books} \text{Modifier}^*$

The rule in (73) combines an unlimited number of modifiers with the noun *books* followed by an unlimited number of modifiers. We can use this rule to derive phrases such as those in (74):

- (74)
- a. *books*
 - b. *interesting books*
 - c. *interesting books from Stuttgart*

Make reference to coordination data in your answer. Assume that symmetric coordination requires that both coordinated phrases or words have the same syntactic category.

6. **FLGR2012a** suggested treating nounless structures like those in (75) as involving a phrasal construction combining the determiner *the* with an adjective.

- (75) a. Examine the plight of the very poor.
 b. Their outfits range from the flamboyant to the functional.
 c. The unimaginable happened.

(76) shows a phrase structure rule that corresponds to their construction:

(76) $NP \rightarrow \text{the Adj}$

Adj stands for something that can be a single word like *poor* or complex like *very poor*.

Revisit the German data in (45) and (46) and explain why such an analysis and even a more general one as in (77) would not extend to German.

(77) $NP \rightarrow \text{Det Adj}$

7. Why can \bar{X} theory not account for German adjective phrases without additional assumptions? (This task is for (native) speakers of German only.)
8. Come up with a phrase structure grammar that can be used to analyze the sentence in (78), but also rules out the sentences in (79).

- (78) a. Der Mann hilft dem Kind.
 the.NOM man helps the.DAT child
 'The man helps the child.'

- b. Er gibt ihr das Buch.
he.NOM gives her.DAT the book
'He gives her the book.'
- c. Er wartet auf ein Wunder.
he.NOM waits on a miracle
'He is waiting for a miracle.'

- (79) a. *Der Mann hilft er.
the.NOM man helps he.NOM
- b. *Er gibt ihr den Buch.
he.NOM gives her.DAT the.M book.N

9. Consider which additional rules would have to be added to the grammar you developed in the previous exercise in order to be able to analyze the following sentences:

- (80) a. Der Mann hilft dem Kind jetzt.
the.NOM man helps the.DAT child now
'The man helps the child now.'
- b. Der Mann hilft dem Kind neben dem Bushäuschen.
the.NOM man helps the.DAT child next.to the bus.shelter
'The man helps the child next to the bus shelter.'
- c. Er gibt ihr das Buch jetzt.
he.NOM gives her.DAT the.ACC book now
'He gives her the book now.'
- d. Er gibt ihr das Buch neben dem Bushäuschen.
he.NOM gives her.DAT the.ACC book next.to the bus.shelter
'He gives her the book next to the bus shelter.'
- e. Er wartet jetzt auf ein Wunder.
he.NOM waits now on a miracle
'He is waiting for a miracle now.'

f. Er wartet neben dem Bushäuschen auf ein Wunder.
 he.NOM waits next.to the.DAT bus.shelter on a miracle
 ‘He is waiting for a miracle next to the bus shelter.’

10. Install a Prolog system (e.g., SWI^a=Prolog^a) and try out your grammar. SWI-Prolog also comes with an online version where you can input your grammar directly without any installation.^b Details regarding the notation can be found in the English Wikipedia entry for Definite Clause Grammar (DCG).^c

^a<http://www.swi-prolog.org>, 2018-02-20.

^b<https://swish.swi-prolog.org/>, 2018-02-20.

^chttps://en.wikipedia.org/wiki/Definite_clause_grammar, 2018-02-20.



Further reading

The expansion of phrase structure grammars to include features was proposed as early as 1963 by **Harman63a**.

The phrase structure grammar for noun phrases discussed in this chapter covers a large part of the syntax of noun phrases but cannot explain certain NP structures. Furthermore, it has the problem, which exercise ?? is designed to show. A discussion of these phenomena and a solution in the framework of HPSG can be found in **Netter98a** and **Kiss2005a**.

The discussion of the integration of semantic information into phrase structure grammars was very short. A detailed discussion of predicate logic and its integration into phrase structure grammars – as well as a discussion of quantifier scope – can be found in **BB2005a**.

3 Transformational Grammar – Government & Binding

Transformational Grammar and its subsequent incarnations (such as Government and Binding Theory and Minimalism) were developed by Noam Chomsky at MIT in Boston (Chomsky57a; Chomsky65a; Chomsky75a; Chomsky81a; Chomsky86b; Chomsky95a-u). Manfred Bierwisch63a was the first to implement Chomsky's ideas for German. In the 60s, the decisive impulse came from the *Arbeitsstelle Strukturelle Grammatik* 'Workgroup for Structural Grammar', which was part of the Academy of Science of the GDR. See Bierwisch92 and Vater2010a for a historic overview. As well as Bierwisch's work, the following books focusing on German or the Chomskyan research program in general should also be mentioned: Fanselow87a, FF87a, SS88a, Grewendorf88a, Haider93a, Sternefeld2006a-u.

The different implementations of Chomskyan theories are often grouped under the heading *Generative Grammar*. This term comes from the fact that phrase structure grammars and the augmented frameworks that were suggested by Chomsky can generate sets of well-formed expressions (see p. 64). It is such a set of sentences that constitutes a language (in the formal sense) and one can test if a sentence forms part of a language by checking if a particular sentence is in the set of sentences generated by a given grammar. In this sense, simple phrase structure grammars and, with corresponding formal assumptions, GPSG, LFG, HPSG and Construction Grammar (CxG) are generative theories. In recent years, a different view of the formal basis of theories such as LFG, HPSG and CxG has emerged such that the aforementioned theories are now *model theoretic* theories rather than generative"=enumerative ones¹ (See Chapter ?? for discussion). In 1965, Chomsky defined the term *Generative Grammar* in the following way (see also Chomsky95a-u):

A grammar of a language purports to be a description of the ideal speaker-hearer's intrinsic competence. If the grammar is, furthermore, perfectly explicit – in other words, if it does not rely on the intelligence of the under-

¹Model theoretic approaches are always constraint"=based and the terms *model theoretic* and *constraint"=based* are sometimes used synonymously.

standing reader but rather provides an explicit analysis of his contribution – we may call it (somewhat redundantly) a *generative grammar*. (Chomsky65a)

In this sense, all grammatical theories discussed in this book would be viewed as generative grammars. To differentiate further, sometimes the term *Mainstream Generative Grammar* (MGG) is used (CJ2005a) for Chomskyan models. In this chapter, I will discuss a well-developed and very influential version of Chomskyan grammar, GB theory. More recent developments following Chomsky's Minimalist Program are dealt with in Chapter 4.

3.1 General remarks on the representational format

This section provides an overview of general assumptions. I introduce the concept of transformations in Section 3.1.1. Section 3.1.2 provides background information about assumptions regarding language acquisition, which shaped the theory considerably, Section 3.1.3 introduces the so-called T model, the basic architecture of GB theory. Section 3.1.4 introduces the \bar{X} theory in the specific form used in GB and Section 3.1.5 shows how this version of the \bar{X} theory can be applied to English. The discussion of the analysis of English sentences is an important prerequisite for the understanding of the analysis of German, since many analyses in the GB framework are modeled in parallel to the analyses of English. Section 3.1.6 introduces the analysis of German clauses in a parallel way to what has been done for English in Section 3.1.5.

3.1.1 Transformations

In the previous chapter, I introduced simple phrase structure grammars. Chomsky57a criticized this kind of rewrite grammars since – in his opinion – it is not clear how one can capture the relationship between active and passive sentences or the various ordering possibilities of constituents in a sentence. While it is of course possible to formulate different rules for active and passive sentences in a phrase structure grammar (e.g., one pair of rules for intransitive (1), one for transitive (2) and one for ditransitive verbs (3)), it would not adequately capture the fact that the same phenomenon occurs in the example pairs in (1)–(3):

- (1) a. weil dort noch jemand arbeitet
because there still somebody works
'because somebody is still working there'

3.1 General remarks on the representational format

- b. weil dort noch gearbeitet wurde
because there still worked was
'because work was still being done there'
- (2) a. weil er den Weltmeister schlägt
because he the world.champion beats
'because he beats the world champion'
- b. weil der Weltmeister geschlagen wurde
because the world.champion beaten was
'because the world champion was beaten'
- (3) a. weil der Mann der Frau den Schlüssel stiehlt
because the man the woman the key steals
'because the man is stealing the key from the woman'
- b. weil der Frau der Schlüssel gestohlen wurde
because the woman the key stolen was
'because the key was stolen from the woman'

Chomsky57a suggests a transformation that creates a connection between active and passive sentences. The transformation that he suggests for English corresponds to (4), which is taken from **Klenk2003a**:

- (4) NP V NP \rightarrow 3 [_{AUX} be] 2en [_{PP} [_P by] 1]
1 2 3

This transformational rule maps a tree with the symbols on the left"=hand side of the rule onto a tree with the symbols on the right"=hand side of the rule. Accordingly, 1, 2 and 3 on the right of the rule correspond to symbols, which are under the numbers on the left"=hand side. *en* stands for the morpheme which forms the participle (*seen*, *been*, ..., but also *loved*). Both trees for (5a,b) are shown in Figure 3.1.

- (5) a. John loves Mary.
- b. Mary is loved by John.

The symbols on the left of transformational rules do not necessarily have to be in a local tree, that is, they can be daughters of different mothers as in Figure 3.1.

Rewrite grammars were divided into four complexity classes based on the properties they have. The simplest grammars are assigned to the class 3, whereas



Figure 3.1: Application of passive transformation

the most complex are of Type-0. The so-called context-free grammars we have dealt with thus far are of Type-2. Transformational grammars which allow symbols to be replaced by arbitrary other symbols are of Type-0 (PR73a-u). Research on the complexity of natural languages shows that the highest complexity level (Type-0) is too complex for natural language. It follows from this – assuming that one wants to have a restrictive formal apparatus for the description of grammatical knowledge (Chomsky65a) – that the form and potential power of transformations has to be restricted.² Another criticism of early versions of transformational grammar was that, due to a lack of restrictions, the way in which transformations interact was not clear. Furthermore, there were problems associated with transformations which delete material (see PR73a-u; Klenk2003a). For this reason, new theoretical approaches such as Government & Binding (Chomsky81a) were developed. In this model, the form that grammatical rules can take is restricted (see Section 3.1.4). Elements moved by transformations are still represented in their original position, which makes them recoverable at the original position and hence the necessary information is available for semantic interpretation. There are also more general principles, which serve to restrict transformations.

After some initial remarks on the model assumed for language acquisition in GB theory, we will take a closer look at phrase structure rules, transformations and constraints.

3.1.2 The hypothesis regarding language acquisition: Principles & Parameters

Chomsky65a assumes that linguistic knowledge must be innate since the language system is, in his opinion, so complex that it would be impossible to learn a language from the given input using more general cognitive principles alone

²For more on the power of formal languages, see Chapter ??.

(see also Section ??). If it is not possible to learn language solely through interaction with our environment, then at least part of our language ability must be innate. The question of exactly what is innate and if humans actually have an innate capacity for language remains controversial and the various positions on the question have changed over the course of the last decades. Some notable works on this topic are **Pinker94a**, **Tomasello95a**, **Wunderlich2004a**, **HCF2002a**, **Chomsky2007a**, and **PS2001a** and other papers in the same volume. For more on this discussion, see Chapter ??.

Chomsky81a also assumes that there are general, innate principles which linguistic structure cannot violate. These principles are parametrized, that is, there are options. Parameter settings can differ between languages. An example for a parametrized principle is shown in (6):

- (6) Principle: A head occurs before or after its complement(s) depending on the value of the parameter POSITION.

The Principles & Parameters model (P&P model) assumes that a significant part of language acquisition consists of extracting enough information from the linguistic input in order to be able to set parameters. **Chomsky2000a-u** compares the setting of parameters to flipping a switch. For a detailed discussion of the various assumptions about language acquisition in the P&P-model, see Chapter ??. Speakers of English have to learn that heads occur before their complements in their language, whereas a speaker of Japanese has to learn that heads follow their complements. (7) gives the respective examples:

- (7) a. be showing pictures of himself
b. zibun -no syasin-o mise-te iru
REFL from picture showing be

As one can see, the Japanese verb, noun and prepositional phrases are a mirror image of the corresponding phrases in English. (8) provides a summary and shows the parametric value for the position parameter:

- | | | | |
|-----|----------|--------------------------------|-------------------------|
| (8) | Language | Observation | Parameter: head initial |
| | English | Heads occur before complements | + |
| | Japanese | Heads occur after complements | - |

Investigating languages based on their differences with regard to certain assumed parameters has proven to be a very fruitful line of research in the last few decades and has resulted in an abundance of comparative cross-linguistic studies.

After these introductory comments on language acquisition, the following sections will discuss the basic assumptions of GB theory.

3.1.3 The T model

Chomsky criticized simple PSGs for not being able to adequately capture certain correlations. An example of this is the relationship between active and passive sentences. In phrase structure grammars, one would have to formulate active and passive rules for intransitive, transitive and ditransitive verbs (see the discussion of (1)–(3) above). The fact that the passive can otherwise be consistently described as the suppression of the most prominent argument is not captured by phrase structure rules. Chomsky therefore assumes that there is an underlying structure, the so-called *Deep Structure*, and that other structures are derived from this. The general architecture of the so-called T model is discussed in the following subsections.

3.1.3.1 D-structure and S-structure

During the derivation of new structures, parts of the Deep Structure can be deleted or moved. In this way, one can explain the relationship between active and passive sentences. As the result of this kind of manipulation of structures, also called transformations, one derives a new structure, the *Surface Structure*, from the original Deep Structure. Since the Surface Structure does not actually mirror the actual use of words in a sentence in some versions of the theory, the term *S-structure* is sometimes used instead as to avoid misunderstandings.

- (9) *Surface Structure* = S-structure
Deep Structure = D-structure

Figure 3.2 gives an overview of the GB architecture: phrase structure rules and the lexicon license the D-structure from which the S-structure is derived by means of transformations. S-structure feeds into Phonetic Form (PF) and Log-

Figure 3.2: The T model

ical Form (LF). The model is referred to as the *T-model* (or Y-model) because D-structure, S-structure, PF and LF form an upside-down T (or Y). We will look at each of these individual components in more detail.

Using phrase structure rules, one can describe the relationships between individual elements (for instance words and phrases, sometimes also parts of words). The format for these rules is \bar{X} syntax (see Section 2.5). The lexicon, together with the structure licensed by \bar{X} syntax, forms the basis for D-structure. D-structure is then a syntactic representation of the selectional grid (= valence classes) of individual word forms in the lexicon.

3.1 General remarks on the representational format

The lexicon contains a lexical entry for every word which comprises information about morphophonological structure, syntactic features and selectional properties. This will be explained in more detail in Section 3.1.3.4. Depending on one's exact theoretical assumptions, morphology is viewed as part of the lexicon. Inflectional morphology is, however, mostly consigned to the realm of syntax. The lexicon is an interface for semantic interpretation of individual word forms.

The surface position in which constituents are realized is not necessarily the position they have in D"-structure. For example, a sentence with a ditransitive verb has the following ordering variants:

- (10) a. [dass] der Mann der Frau das Buch gibt
that the.NOM man the.DAT woman the.ACC book gives
'that the man gives the woman the book'
- b. Gibt der Mann der Frau das Buch?
gives the.NOM man the.DAT woman the.ACC book
'Does the man give the woman the book?'
- c. Der Mann gibt der Frau das Buch.
the.NOM man gives the.DAT woman the.ACC book
'The man gives the woman the book.'

The following transformational rules for the movements above are assumed: (10b) is derived from (10a) by fronting the verb, and (10c) is derived from (10b) by fronting the nominative noun phrase. In GB theory, there is only one very general transformation: Move α = "Move anything anywhere!". The nature of what exactly can be moved where and for which reason is determined by principles. Examples of such principles are the Theta"-Criterion and the Case Filter, which will be dealt with below.

The relations between a predicate and its arguments that are determined by the lexical entries have to be accessible for semantic interpretation at all representational levels. For this reason, the base position of a moved element is marked with a trace. This means, for instance, that the position in which the fronted *gibt* 'gives' originated is indicated in (11b). The respective marking is referred to as a *trace* or a *gap*. Such empty elements may be frightening when one encounters them first, but I already motivated the assumption of empty elements in nominal structures in Section 2.4.1 (page 79).

- (11) a. [dass] der Mann der Frau das Buch gibt
that the man the woman the book gives
'that the man gives the woman the book'

- b. Gibt_i der Mann der Frau das Buch _{-i}?
 gives the man the woman the book
 ‘Does the man give the woman the book?’
- c. [Der Mann]_j gibt_i _{-j} der Frau das Buch _{-i}.
 the man gives the woman the book
 ‘The man gives the woman the book.’

(11c) is derived from (11a) by means of two movements, which is why there are two traces in (11c). The traces are marked with indices so it is possible to distinguish the moved constituents. The corresponding indices are then present on the moved constituents. Sometimes, *t* (for *trace*) is used to represent traces.

The S”=structure derived from the D”=structure is a surface”=like structure but should not be equated with the structure of actual utterances.

3.1.3.2 Phonetic Form

Phonological operations are represented at the level of Phonetic Form (PF). PF is responsible for creating the form which is actually pronounced. For example, so”=called *wanna*”=contraction takes place at PF (Chomsky81a).

- (12) a. The students want to visit Paris.
 b. The students wanna visit Paris.

The contraction in (12) is licensed by the optional rule in (13):

- (13) want + to → wanna

3.1.3.3 Logical Form

Logical Form is the syntactic level which mediates between S”=structure and the semantic interpretation of a sentence. Some of the phenomena which are dealt with by LF are anaphoric reference of pronouns, quantification and control.

Syntactic factors play a role in resolving anaphoric dependencies. An important component of GB theory is Binding Theory, which seeks to explain what a pronoun can or must refer to and when a reflexive pronoun can or must be used.

(14) gives some examples of both personal and reflexive pronouns:

- (14) a. Peter kauft einen Tisch. Er gefällt ihm.
 Peter buys a table(_M) he likes him
 ‘Peter is buying a table. He likes it/him.’

3.1 General remarks on the representational format

- b. Peter kauft eine Tasche. Er gefällt ihm.
Peter buys a bag(*F*) he likes him
'Peter is buying a bag. He likes it/him.'
- c. Peter kauft eine Tasche. Er gefällt sich.
Peter buys a bag(*F*) he likes himself
'Peter is buying a bag. He likes himself.'

In the first example, *er* 'he' can refer to either Peter, the table or something/someone else that was previously mentioned in the context. *ihm* 'him' can refer to Peter or someone in the context. Reference to the table is restricted by world knowledge. In the second example, *er* 'he' cannot refer to *Tasche* 'bag' since *Tasche* is feminine and *er* is masculine. *er* 'he' can refer to Peter only if *ihm* 'him' does not refer to Peter. *ihm* would otherwise have to refer to a person in the wider context. This is different in (14c). In (14c), *er* 'he' and *sich* 'himself' must refer to the same object. This is due to the fact that the reference of reflexives such as *sich* is restricted to a particular local domain. Binding Theory attempts to capture these restrictions.

LF is also important for quantifier scope. Sentences such as (15a) have two readings. These are given in (15b) and (15c).

- (15) a. Every man loves a woman.
- b. $\forall x \exists y (man(x) \rightarrow (woman(y) \wedge love(x, y)))$
- c. $\exists y \forall x (man(x) \rightarrow (woman(y) \wedge love(x, y)))$

The symbol \forall stands for a *universal quantifier* and \exists stands for an *existential quantifier*. The first formula corresponds to the reading that for every man, there is a woman who he loves and in fact, these can be different women. Under the second reading, there is exactly one woman such that all men love her. The question of when such an ambiguity arises and which reading is possible when depends on the syntactic properties of the given utterance. LF is the level which is important for the meaning of determiners such as *a* and *every*.

Control Theory is also specified with reference to LF. Control Theory deals with the question of how the semantic role of the infinitive subject in sentences such as (16) is filled.

- (16) a. Der Professor schlägt dem Studenten vor, die Klausur noch mal
the professor suggests the student PART the test once again
zu schreiben.
to write
'The professor advises the student to take the test again.'

- b. Der Professor schlägt dem Studenten vor, die Klausur nicht zu
the professor suggests the student PART the test not to
bewerten.
grade
'The professor suggests to the student not to grade the test.'
- c. Der Professor schlägt dem Studenten vor, gemeinsam ins Kino
the professor suggests the student PART together into cinema
zu gehen.
to go
'The professor suggests to the student to go to the cinema together.'

3.1.3.4 The lexicon

The meaning of words tells us that they have to be combined with certain roles like “acting person” or “affected thing” when creating more complex phrases. For example, the fact that the verb *beat* needs two arguments belongs to its semantic contribution. The semantic representation of the contribution of the verb *beat* in (17a) is given in (17b):

- (17) a. Judit beats the grandmaster.
b. *beat'*(x,y)

Dividing heads into valence classes is also referred to as *subcategorization*: *beat* is subcategorized for a subject and an object. This term comes from the fact that a head is already categorized with regard to its part of speech (verb, noun, adjective, ...) and then further sub"-classes (e.g., intransitive or transitive verb) are formed with regard to valence information. Sometimes the phrase *X subcategorizes for Y* is used, which means *X selects Y*. *beat* is referred to as the predicate since *beat'* is the logical predicate. The subject and object are the arguments of the predicate. There are several terms used to describe the set of selectional requirements such as *argument structure*, *valence frames*, *subcategorization frame*, *thematic grid* and *theta"-grid* or *θ-grid*.

Adjuncts modify semantic predicates and when the semantic aspect is emphasized they are also called *modifiers*. Adjuncts are not present in the argument structure of predicates.

Following GB assumptions, arguments occur in specific positions in the clause – in so"-called argument positions (e.g., the sister of an X^0 element, see Section 2.5). The Theta"-Criterion states that elements in argument positions have

to be assigned a semantic role – a so-called theta-role – and each role can be assigned only once (Chomsky81a):

Principle 1 (Theta-Criterion)

- *Each theta-role is assigned to exactly one argument position.*
- *Every phrase in an argument position receives exactly one theta-role.*

The arguments of a head are ordered, that is, one can differentiate between higher- and lower-ranked arguments. The highest-ranked argument of verbs and adjectives has a special status. Since GB assumes that it is often (and always in some languages) realized in a position outside of the verb or adjective phrase, it is often referred to as the *external argument*. The remaining arguments occur in positions inside of the verb or adjective phrase. These kind of arguments are dubbed *internal arguments* or *complements*. For simple sentences, this often means that the subject is the external argument.

When discussing types of arguments, one can identify three classes of theta-roles:

- Class 1: agent (acting individual), the cause of an action or feeling (stimulus), holder of a certain property
- Class 2: experiencer (perceiving individual), the person profiting from something (beneficiary) (or the opposite: the person affected by some kind of damage), possessor (owner or soon-to-be owner of something, or the opposite: someone who has lost or is lacking something)
- Class 3: patient (affected person or thing), theme

If a verb has several theta-roles of this kind to assign, Class 1 normally has the highest rank, whereas Class 3 has the lowest. Unfortunately, the assignment of semantic roles to actual arguments of verbs has received a rather inconsistent treatment in the literature. This problem has been discussed by Dowty91a, who suggests using proto-roles. An argument is assigned the proto-agent role if it has sufficiently many of the properties that were identified by Dowty as prototypical properties of agents (e.g., animacy, volitionality).

The mental lexicon contains *lexical entries* with the specific properties of syntactic words needed to use that word grammatically. Some of these properties are the following:

- form
- meaning (semantics)

3 Transformational Grammar – Government & Binding

- grammatical features: syntactic word class + morphosyntactic features
- theta”=grid

(18) shows an example of a lexical entry:

(18)

form	<i>helf</i> - ‘help’
semantics	<i>helfen</i> ’
grammatical features	verb
theta”=grid	
theta”=roles	<u>agent</u> beneficiary
grammatical particularities	dative

Assigning semantic roles to specific syntactic requirements (beneficiary = dative) is also called *linking*.

Arguments are ordered according to their ranking: the highest argument is furthest left. In the case of *helfen*, the highest argument is the external argument, which is why the agent is underlined. With so”=called unaccusative verbs,³ the highest argument is not treated as the external argument. It would therefore not be underlined in the corresponding lexical entry.

3.1.4 \bar{X} theory

In GB, it is assumed that all syntactic structures licensed by the core grammar⁴ correspond to the \bar{X} schema (see Section 2.5).⁵ In the following sections, I will comment on the syntactic categories assumed and the basic assumptions with regard to the interpretation of grammatical rules.

3.1.4.1 Syntactic categories

The categories which can be used for the variable X in the \bar{X} schema are divided into lexical and functional categories. This correlates roughly with the difference between open and closed word classes. The following are lexical categories:

³See **Perlmutter78** for a discussion of unaccusative verbs. The term *ergative verb* is also common, albeit a misnomer. See **Burzio81-u**; **Burzio86a-u-gekauft** for the earliest work on unaccusatives in the Chomskyan framework and **Grewendorf89a** for German. Also, see **Pullum88a** on the usage of these terms and for a historical evaluation.

⁴**Chomsky81a** distinguishes between a regular area of language that is determined by a grammar that can be acquired using genetically determined language”=specific knowledge and a periphery, to which irregular parts of language such as idioms (e.g., *to pull the wool over sb.’s eyes*) belong. See Section ??.

⁵**Chomsky70a** allows for grammatical rules that deviate from the \bar{X} schema. It is, however, common practice to assume that languages exclusively use \bar{X} structures.

3.1 General remarks on the representational format

- V = verb
- N = noun
- A = adjective
- P = preposition/postposition
- Adv = adverb

Lexical categories can be represented using binary features and a cross”=classification:⁶

Table 3.1: Representation of four lexical categories using two binary features

	-V	+V
-N	P = [-N, -V]	V = [-N, +V]
+N	N = [+N, -V]	A = [+N, +V]

Adverbs are viewed as intransitive prepositions and are therefore captured by the decomposition in the table above.

Using this cross”=classification, it is possible to formulate generalizations. One can, for example, simply refer to adjectives and verbs: all lexical categories which are [+V] are either adjectives or verbs. Furthermore, one can say of [+N] categories (nouns and adjectives) that they can bear case.

Apart from this, some authors have tried to associate the head position with the feature values in Table 3.1 (see e.g., Grewendorf88a; Haftka96a; G. GMueller2011a). With prepositions and nouns, the head precedes the complement in German:

- (19) a. *für* Marie
 for Marie
 b. *Bild* von Maria
 picture of Maria

With adjectives and verbs, the head is final:

⁶See Chomsky70a for a cross”=classification of N, A and V, and Jackendoff77a for a cross”=classification that additionally includes P but has a different feature assignment.

- (20) a. dem König *treu*
 the king loyal
 ‘Loyal to the king’
 b. der [dem Kind *helfende*] Mann
 the the child helping man
 ‘the man helping the child’
 c. dem Mann *helfen*
 the man help
 ‘help the man’

This data seems to suggest that the head is final with [+V] categories and initial with [–V] categories. Unfortunately, this generalization runs into the problem that there are also postpositions in German. These are, like prepositions, not verbal, but do occur after the NP they require:

- (21) a. des Geldes *wegen*
 the money because
 ‘because of the money’
 b. die Nacht *über*
 the night during
 ‘during the night’

Therefore, one must either invent a new category, or abandon the attempt to use binary category features to describe ordering restrictions. If one were to place postpositions in a new category, it would be necessary to assume another binary feature.⁷ Since this feature can have either a negative or a positive value, one

⁷Martin Haspelmath has pointed out that one could assume a rule that moves a post-head argument into a pre-head position (see Riemsdijk78a for the discussion of a transformational solution). This would be parallel to the realization of prepositional arguments of adjectives in German:

- (i) a. auf seinen Sohn stolz
 on his son proud
 ‘proud of his son’
 b. stolz auf seinen Sohn
 proud of his son

But note that the situation is different with postpositions here, while all adjectives that take prepositional objects allow for both orders, this is not the case for prepositions. Most prepositions do not allow their object to occur before them. It is an idiosyncratic feature of some postpositions that they want to have their argument to the left.

would then have four additional categories. There are then eight possible feature combinations, some of which would not correspond to any plausible category.

For functional categories, GB does not propose a cross”=classification. Usually, the following categories are assumed:

- C Complementizer (subordinating conjunctions such as *dass* ‘that’)
- I Finiteness (as well as Tense and Mood);
also Infl in earlier work (inflection),
T in more recent work (Tense)
- D Determiner (article, demonstrative)

3.1.4.2 Assumptions and rules

In GB, it is assumed that all rules must follow the \bar{X} format discussed in Section 2.5. In other theories, rules which correspond to the \bar{X} format are used along other rules which do not. If the strict version of \bar{X} theory is assumed, this comes with the assumption of *endocentricity*: every phrase has a head and every head is part of a phrase (put more technically: every head projects to a phrase).

Furthermore, as with phrase structure grammars, it is assumed that the branches of tree structures cannot cross (*Non-Tangling Condition*). This assumption is made by the majority of theories discussed in this book. There are, however, some variants of TAG, HPSG, Construction Grammar, and Dependency Grammar which allow crossing branches and therefore discontinuous constituents (BJR91a; Reape94a; BC2005a; Heringer96a-u; Eroms2000a).

In \bar{X} theory, one normally assumes that there are at most two projection levels (X' and X''). However, there are some versions of Mainstream Generative Grammar and other theories which allow three or more levels (Jackendoff77a; Uszkoreit87a). In this chapter, I follow the standard assumption that there are two projection levels, that is, phrases have at least three levels:

- X^0 = head
- X' = intermediate projection (\bar{X} , read: X bar)
- XP = highest projection ($= X'' = \bar{\bar{X}}$), also called *maximal projection*

3.1.5 CP and IP in English

Most work in Mainstream Generative Grammar is heavily influenced by previous publications dealing with English. If one wants to understand GB analyses

of German and other languages, it is important to first understand the analyses of English and, for this reason, this will be the focus of this section. The CP/IP system is also assumed in LFG grammars of English and thus the following section also provides a foundation for understanding some of the fundamentals of LFG presented in Chapter ??.

In earlier work, the rules in (22a) and (22b) were proposed for English sentences (Chomsky81a).

- (22) a. $S \rightarrow NP VP$
 b. $S \rightarrow NP Infl VP$

Infl stands for *Inflection* as inflectional affixes are inserted at this position in the structure. The symbol AUX was also used instead of Infl in earlier work, since auxiliary verbs are treated in the same way as inflectional affixes. Figure 3.3 shows a sample analysis of a sentence with an auxiliary, which uses the rule in (22b).

Figure 3.3: Sentence with an auxiliary verb following Chomsky81a

Figure 3.4: Sentence with auxiliary verb in the CP/IP system

Together with its complements, the verb forms a structural unit: the VP. The constituent status of the VP is supported by several constituent tests and further differences between subjects and objects regarding their positional restrictions.

The rules in (22) do not follow the \bar{X} template since there is no symbol on the right-hand side of the rule with the same category as one on the left-hand side, that is, there is no head. In order to integrate rules like (22) into the general theory, Chomsky86b developed a rule system with two layers above the verb phrase (VP), namely the CP/IP system. CP stands for *Complementizer Phrase*. The head of a CP can be a complementizer. Before we look at CPs in more detail, I will discuss an example of an IP in this new system. Figure 3.4 shows an IP with an auxiliary in the I^0 position. As we can see, this corresponds to the structure of the \bar{X} template: I^0 is a head, which takes the VP as its complement and thereby forms I' . The subject is the specifier of the IP. Another way to phrase this is to say that the subject is in the specifier position of the IP. This position is usually referred to as SpecIP.⁸

⁸Sometimes SpecIP and similar labels are used in trees (for instance by Haegeman94a-u, Meinunger2000a and Lohnstein2014a). I avoid this in this book since SpecIP, SpecAdvP are not categories like NP or AP or AdvP but positions that items of a certain category can take. See Chapter 2 on the phrase structure rules that license trees.

3.1 General remarks on the representational format

The sentences in (23) are analyzed as complementizer phrases (CPs), the complementizer is the head:

- (23) a. that Ann will read the newspaper
b. that Ann reads the newspaper

In sentences such as (23), the CPs do not have a specifier. Figure 3.5 shows the analysis of (23a).

Figure 3.5: Complementizer phrase

Figure 3.6: Polar question

Yes/no-questions in English such as those in (24) are formed by moving the auxiliary verb in front of the subject.

- (24) Will Ann read the newspaper?

Let us assume that the structure of questions corresponds to the structure of sentences with complementizers. This means that questions are also CPs. Unlike the sentences in (23), however, there is no subordinating conjunction. In the D"-structure of questions, the C⁰ position is empty and the auxiliary verb is later moved to this position. Figure ?? shows an analysis of (24). The original position of the auxiliary is marked by the trace *-_k*, which is coindexed with the moved auxiliary.

wh-questions are formed by the additional movement of a constituent in front of the auxiliary; that is into the specifier position of the CP. Figure ?? on page ?? shows the analysis of (25):

- (25) What will Ann read?

Figure 3.7: *wh*"=question

As before, the movement of the object of *read* is indicated by a trace. This is important when constructing the meaning of the sentence. The verb assigns some semantic role to the element in its object position. Therefore, one has to be able to "reconstruct" the fact that *what* actually originates in this position. This is ensured by coindexation of the trace with *what*.

Several ways to depict traces are used in the literature. Some authors assume a trace instead of the object NP as in Figure 3.8a (Grewendorf88a; Haegeman94a-u). Others have the object NP in the tree and indicate the movement by a trace that is dominated by the NP as in Figure 3.8b (SS88a; Grewendorf88a; Haegeman94a-u; Sternefeld2006a-u). The first proposal directly reflects the assumption that a

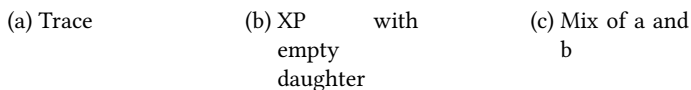


Figure 3.8: Alternative ways of depicting movement: the moved constituent can be represented by a trace or by an XP dominating a trace

complete phrase is moved and leaves a trace that represents the thing that is moved. If one thinks about the properties of the trace it is clear that it has the same category as the element that was at this position before movement. Hence the second way to represent the moved category is appropriate as well. Figure 3.8b basically says that the object that is moved is an NP but that there is nothing to pronounce. Given what was just said the most appropriate way to represent movement would be the one in Figure 3.8c. This picture is a mix of the two other pictures. The index is associated with the category and not with the empty phonology. In my opinion this best depicts the fact that trace and filler are related. However, I never saw this way of depicting movement in the GB literature and hence I will stick to the more common notation in Figure 3.8b. This way to depict movement is also more similar to the representation that is used by all authors for the movement of words (so-called head-movement). For example the trace $_{-k}$, which stands for a moved I^0 in Figure ?? is never depicted as daughter of I' but always as a daughter of I^0 .

Until now, I have not yet discussed sentences without auxiliaries such as (23b). In order to analyze this kind of sentences, it is usually assumed that the inflectional affix is present in the I^0 position. An example analysis is given in Figure 3.9. Since the inflectional affix precedes the verb, some kind of move-

Figure 3.9: Sentence without auxiliary

ment operation still needs to take place. There are two suggestions in the literature: one is to assume lowering, that is, the affix moves down to the verb (Pollock89a-uChomsky91a-uHaegeman94a-uSKS2013a-u). The alternative is to assume that the verb moves up to the affix (FF87a). Since theories with lowering of inflectional affixes are complicated for languages in which the verb ultimately ends up in C (basically in all Germanic languages except English), I follow FF87a's (FF87a) suggestion for English and Grewendorf93's (Grewendorf93) suggestion for German and assume that the verb moves from V to I in English and from V to I to C in German.⁹

⁹SKS2013a-u argue for an affix lowering approach by pointing out that approaches assuming that the verb stem moves to I (their T) predict that adverbs appear to the right of the verb

Following this excursus on the analysis of English sentences, we can now turn to German.

3.1.6 The structure of the German clause

The CP/IP model has been adopted by many scholars for the analysis of German.¹⁰ The categories C, I and V, together with their specifier positions, can be linked to the topological fields as shown in Figure 3.10.

Figure 3.10: CP, IP and VP and the topological model of German

Note that SpecCP and SpecIP are not category symbols. They do not occur in grammars with rewrite rules. Instead, they simply describe positions in the tree.

As shown in Figure 3.10, it is assumed that the highest argument of the verb (the subject in simple sentences) has a special status. It is taken for granted that the subject always occurs outside of the VP, which is why it is referred to as the external argument. The VP itself does not have a specifier. In more recent work, however, the subject is generated in the specifier of the VP (FS86a-u; KS91a-u). In some languages, it is assumed that it moves to a position outside of the VP. In other languages such as German, this is the case at least under certain conditions (e.g., definiteness, see Diesing92a). I am presenting the classical GB analysis here, where the subject is outside the VP. All arguments other than the subject are complements of the V, that are realized within the VP, that is, they are internal arguments. If the verb requires just one complement, then this is the sister of the

rather than to the left:

- (i) a. John will carefully study Russian.
- b. John carefully studies Russian.
- c. * John studies carefully Russian.

If the affix "s" is in the position of the auxiliary and the verb moves to the affix, one would expect (i.c) to be grammatical rather than (i.b).

A third approach is to assume empty I (or more recently T) heads for present and past tense and have these heads select a fully inflected verb. See Carnie2013a-u for such an approach to English.

For German it was also suggested not to distinguish between I and V at all and treat auxiliaries like normal verbs (see footnote 10 below). In such approaches verbs are inflected as V, no I node is assumed (Haider93a; Haider97a).

¹⁰For GB analyses without IP, see BK89a, Hoehle91, Haider93a; Haider97a and Sternefeld2006a-u. Haider assumes that the function of I is integrated into the verb. In LFG, an IP is assumed for English (Bresnan2001a; Dalrymple2001a-u), but not for German (Berman2003a). In HPSG, no IP is assumed.

head V^0 and the daughter of V' according to the \bar{X} schema. The accusative object is the prototypical complement.

Following the \bar{X} template, adjuncts branch off above the complements of V' . The analysis of a VP with an adjunct is shown in Figure 3.11.

- (26) weil der Mann morgen den Jungen trifft
 because the man tomorrow the boy meets
 ‘because the man is meeting the boy tomorrow’

Figure 3.11: Analysis of adjuncts in GB theory

3.2 Verb position

In German, the position of the heads of VP and IP (V^0 and I^0) are to the right of their complements and V^0 and I^0 form part of the right sentence bracket. The subject and all other constituents (complements and adjuncts) all occur to the left of V^0 and I^0 and form the middle field. It is assumed that German – at least in terms of D”=structure – is an SOV language (= a language with the base order Subject–Object–Verb). The analysis of German as an SOV language is almost as old as Transformational Grammar itself. It was originally proposed by Bierwisch63a.¹¹ Unlike German, Germanic languages like Danish, English and Romance languages like French are SVO languages, whereas Welsh and Arabic are VSO languages. Around 40 % of all languages belong to the SOV languages, around 35 % are SVO (wals-81).

The assumption of verb”=final order as the base order is motivated by the following observations:¹²

1. Verb particles form a close unit with the verb.

- (27) a. weil er morgen an-fängt
 because he tomorrow PART-starts
 ‘because he is starting tomorrow’

¹¹Bierwisch attributes the assumption of an underlying verb”=final order to Fourquet57a. A German translation of the French manuscript cited by Bierwisch can be found in Fourquet70a. For other proposals, see Bach62a, Reis74a, Koster75a and Thiersch78a. Analyses which assume that German has an underlying SOV pattern were also suggested in GPSG (Jacobs86a), LFG (Berman96a-u) and HPSG (KW91a; Oliva92a; Netter92; Kiss93; Frank94; Kiss95a; Feldhaus97; Meurers2000b; Mueller2005c; MuellerGS).

¹²For points 1 and 2, see Bierwisch63a. For point 4 see Netter92.

- b. Er fängt morgen an.
 he starts tomorrow PART
 ‘He is starting tomorrow.’

This unit can only be seen in verb”=final structures, which speaks for the fact that this structure reflects the base order.

Verbs which are derived from a noun by back-formation (e.g., *uraufführen* ‘to perform something for the first time’), can often not be divided into their component parts and V2 clauses are therefore ruled out (This was first mentioned by **Hoehle91b** in unpublished work. The first published source is **Haider93a**):

- (28) a. weil sie das Stück heute ur-auf-führen
 because they the play today PREF-PART-lead
 ‘because they are performing the play for the first time today’
 b. * Sie ur-auf-führen heute das Stück.
 they PREF-PART-lead today the play
 c. * Sie führen heute das Stück ur-auf.
 they lead today the play PREF-PART

The examples show that there is only one possible position for this kind of verb. This order is the one that is assumed to be the base order.

2. Verbs in non”=finite clauses and in finite subordinate clauses with a conjunction are always in final position (I am ignoring the possibility of extra-posing constituents):

- (29) a. Der Clown versucht, Kurt-Martin die Ware zu geben.
 the clown tries Kurt-Martin the goods to give
 ‘The clown is trying to give Kurt-Martin the goods.’
 b. dass der Clown Kurt-Martin die Ware gibt
 that the clown Kurt-Martin the goods gives
 ‘that the clown gives Kurt-Martin the goods’

3. If one compares the position of the verb in German with Danish (Danish is an SVO language like English), then one can clearly see that the verbs in German form a cluster at the end of the sentence, whereas they occur before any objects in Danish (**Oersnes2009b**):

- (30) a. dass er ihn gesehen₃ haben₂ muss₁
 that he him seen have must
 b. at han må₁ have₂ set₃ ham
 that he must have seen him
 ‘that he must have seen him’

4. The scope relations of the adverbs in (31) depend on their order: the left”=most adverb has scope over the two following elements.¹³ This was explained by

¹³ At this point, it should be mentioned that there seem to be exceptions from the rule that modifiers to the left take scope over those to their right. **Kasper94a** discusses examples such as (i), which go back to **BV72**.

- (i) a. Peter liest gut wegen der Nachhilfestunden.
 Peter reads well because.of the tutoring
 b. Peter liest wegen der Nachhilfestunden gut.
 Peter reads because.of the tutoring well
 ‘Peter can read well thanks to the tutoring.’

As **Koster75a** and **Reis80a** have shown, these are not particularly convincing counter”=examples as the right sentence bracket is not filled in these examples and therefore the examples are not necessarily instances of normal reordering inside of the middle field, but could instead involve extraposition of the PP. As noted by Koster and Reis, these examples become ungrammatical if one fills the right bracket and does not extrapose the causal adjunct:

- (ii) a. * Hans hat gut wegen der Nachhilfestunden gelesen.
 Hans has well because.of the tutoring read
 b. Hans hat gut gelesen wegen der Nachhilfestunden.
 Hans has well read because.of the tutoring
 ‘Hans has been reading well because of the tutoring.’

However, the following example from **Crysmann2004a** shows that, even with the right bracket occupied, one can still have an order where an adjunct to the right has scope over one to the left:

- (iii) Da muß es schon erhebliche Probleme mit der Ausrüstung gegeben haben, da
 there must EXPL already serious problems with the equipment given have since
 wegen schlechten Wetters ein Reinhold Messmer niemals aufgabe.
 because.of bad weather a Reinhold Messmer never would.give.up
 ‘There really must have been some serious problems with the equipment because someone
 like Reinhold Messmer would never give up just because of some bad weather.’

Nevertheless, this does not change anything regarding the fact that the corresponding cases in (31) and (32) have the same scope relations regardless of the position of the verb. The general means of semantic composition may well have to be implemented in the same way as in Crysmann’s analysis.

assuming the following structure:

- (31) a. dass er [absichtlich [nicht lacht]]
 that he intentionally not laughs
 ‘that he is intentionally not laughing’
 b. dass er [nicht [absichtlich lacht]]
 that he not intentionally laughs
 ‘that he is not laughing intentionally’

It is interesting to note that scope relations are not affected by verb position. If one assumes that sentences with verb”=second order have the underlying structure in (31), then this fact requires no further explanation. (32) shows the derived S”=structure for (31):

- (32) a. Er lacht_i [absichtlich [nicht _{-i}]].
 he laughs intentionally not
 ‘He is intentionally not laughing.’
 b. Er lacht_i [nicht [absichtlich _{-i}]].
 he laughs not intentionally
 ‘He is not laughing intentionally.’

After motivating and briefly sketching the analysis of verb”=final order, I will now look at the CP/IP analysis of German in more detail. C^0 corresponds to the left sentence bracket and can be filled in two different ways: in subordinate clauses introduced by a conjunction, the subordinating conjunction (the complementizer) occupies C^0 as in English. The verb remains in the right sentence bracket, as illustrated by (33).

- (33) dass jeder diesen Mann kennt
 that everybody this man knows
 ‘that everybody knows this man’

Figure 3.12 gives an analysis of (33). In verb”=first and verb”=second clauses, the

Figure 3.12: Sentence with a complementizer in C^0

finite verb is moved to C^0 via the I^0 position: $V^0 \rightarrow I^0 \rightarrow C^0$ (Grewendorf93). Figure 3.13 on the next page shows the analysis of (34):

- (34) Kennt jeder diesen Mann?
knows everybody this man
'Does everybody know this man?'

Figure 3.13: Verb position in GB

The C^0 position is empty in the D"-structure of (34). Since it is not occupied by a complementizer, the verb can move there.

3.3 Long"-distance dependencies

The SpecCP position corresponds to the prefield and can be filled by any XP in declarative clauses in German. In this way, one can derive the sentences in (36) from (35) by moving a constituent in front of the verb:

- (35) Gibt der Mann dem Kind jetzt den Mantel?
gives the.NOM man the.DAT child now the.ACC coat
'Is the man going to give the child the coat now?'
- (36) a. Der Mann gibt dem Kind jetzt den Mantel.
the.NOM man gives the.DAT child now the.ACC coat
'The man is giving the child the coat now.'
- b. Dem Kind gibt der Mann jetzt den Mantel.
the.DAT child gives the.NOM man now the.ACC coat
- c. Den Mantel gibt der Mann dem Kind jetzt.
the.ACC coat gives the.NOM man the.DAT child now
- d. Jetzt gibt der Mann dem Kind den Mantel.
now gives the.NOM man the.DAT child the.ACC coat

Since any constituent can be placed in front of the finite verb, German is treated typologically as one of the verb"-second languages (V2). Thus, it is a verb"-second language with SOV base order. English, on the other hand, is an SVO language without the V2 property, whereas Danish is a V2 language with SVO as its base order (see Oersnes2009b for Danish).

Figure 3.14 shows the structure derived from Figure 3.13. The crucial factor

Figure 3.14: Fronting in GB theory

for deciding which phrase to move is the *information structure* of the sentence. That is, material connected to previously mentioned or otherwise”=known information is placed further left (preferably in the prefield) and new information tends to occur to the right. Fronting to the prefield in declarative clauses is often referred to as *topicalization*. But this is rather a misnomer, since the focus (informally: the constituent being asked for) can also occur in the prefield. Furthermore, expletive pronouns can occur there and these are non”=referential and as such cannot be linked to preceding or known information, hence expletives can never be topics.

Transformation”=based analyses also work for so”=called *long*”=*distance dependencies*, that is, dependencies crossing several phrase boundaries:

- (37) a. [Um zwei Millionen Mark]_i soll er versucht haben,
 around two million Deutsche.Marks should he tried have
 [eine Versicherung _{-i} zu betrügen].¹⁴
 an insurance.company to deceive
 ‘He apparently tried to cheat an insurance company out of two million Deutsche Marks.’
- b. ”Wer_i glaubt er, daß er _{-i} ist?” erregte sich ein Politiker vom
 who believes he that he is retort REFL a politician from.the
 Nil.¹⁵
 Nile
 ‘“Who does he think he is?”, a politician from the Nile exclaimed.’
- c. Wen_i glaubst du, daß ich _{-i} gesehen habe?¹⁶
 who believe you that I seen have
 ‘Who do you think I saw?’
- d. [Gegen ihn]_i falle es den Republikanern hingegen schwerer,
 against him fall it the Republicans however more.difficult
 [[Angriffe _{-i}] zu lancieren].¹⁷
 attacks to launch
 ‘It is, however, more difficult for the Republicans to launch attacks against him.’

The elements in the prefield in the examples in (37) all originate from more deeply embedded phrases. In GB, it is assumed that long”=distance dependencies across sentence boundaries are derived in steps (Grewendorf88a), that is, in the analysis of (37c), the interrogative pronoun is moved to the specifier position of the *dass*”=clause and is moved from there to the specifier of the matrix clause. The reason for this is that there are certain restrictions on movement which must be checked locally.

¹⁴taz, 04.05.2001, p. 20.

¹⁵Spiegel, 8/1999, p. 18.

¹⁶Scherpenisse86a.

¹⁷taz, 08.02.2008, p. 9.

3.4 Passive

Before I turn to the analysis of the passive in Section 3.4.2, the first subsection will elaborate on the differences between structural and lexical case.

3.4.1 Structural and lexical case

The case of many case-marked arguments is dependent on the syntactic environment in which the head of the argument is realized. These arguments are referred to as arguments with *structural case*. Case"-marked arguments, which do not bear structural case, are said to have *lexical case*.¹⁸

The following are examples of structural case:¹⁹

- (38) a. Der Installateur kommt.
the.NOM plumber comes
'The plumber is coming.'
- b. Der Mann lässt den Installateur kommen.
the man lets the.ACC plumber come
'The man is getting the plumber to come.'
- c. das Kommen des Installateurs
the coming of.the plumber
'the plumber's visit'

In the first example, the subject is in the nominative case, whereas *Installateur* 'plumber' is in accusative in the second example and even in the genitive in the third following nominalization. The accusative case of objects is normally structural case. This case becomes nominative under passivization:

¹⁸Furthermore, there is a so"-called *agreeing case* (see page 48) and *semantic case*. Agreeing case is found in predicatives. This case also changes depending on the structure involved, but the change is due to the antecedent element changing its case. Semantic case depends on the function of certain phrases (e.g., temporal accusative adverbials). Furthermore, as with lexical case of objects, semantic case does not change depending on the syntactic environment. For the analysis of the passive, which will be discussed in this section, only structural and lexical case will be relevant.

¹⁹Compare HM94a.

(38b) is a so"-called *AcI* construction. *AcI* stands for *Accusativus cum infinitivo*, which means "accusative with infinitive". The logical subject of the embedded verb (*kommen* 'to come' in this case) becomes the accusative object of the matrix verb *lassen* 'to let'. Examples for *AcI*"=verbs are perception verbs such as *hören* 'to hear' and *sehen* 'to see' as well as *lassen* 'to let'.

- (39) a. Karl schlägt den Weltmeister.
Karl beats the.ACC world.champion
'Karl beats the world champion.'
b. Der Weltmeister wird geschlagen.
the.NOM world.champion is beaten
'The world champion is being beaten.'

Unlike the accusative, the genitive governed by a verb is a lexical case. The case of a genitive object does not change when the verb is passivized.

- (40) a. Wir gedenken der Opfer.
we remember the.GEN victims
b. Der Opfer wird gedacht.
the.GEN victims are remembered
'The victims are being remembered.'

(40b) is an example of the so-called *impersonal passive*. Unlike example (39b), where the accusative object became the subject, there is no subject in (40b). See Section 1.7.1.

Similarly, there is no change in case with dative objects:

- (41) a. Der Mann hat ihm geholfen.
the man has him.DAT helped
'The man has helped him.'
b. Ihm wird geholfen.
him.DAT is helped
'He is being helped.'

It still remains controversial as to whether all datives should be treated as lexical or whether some or all of the datives in verbal environments should be treated as instances of structural case. For reasons of space, I will not recount this discussion but instead refer the interested reader to Chapter 14 of **MuellerLehrbuch1**. In what follows, I assume – like **Haider86** – that the dative is in fact a lexical case.

3.4.2 Case assignment and the Case Filter

In GB, it is assumed that the subject receives case from (finite) I and that the case of the remaining arguments comes from V (**Chomsky81a**; **Haider84b**; **FF87a**).

Principle 2 (Case Principle)

- *V assigns objective case (accusative) to its complement if it bears structural case.*

- When finite, INFL assigns case to the subject.

The Case Filter rules out structures where case has not been assigned to an NP.

Figure 3.15 shows the Case Principle in action with the example in (42a).²⁰

- (42) a. [dass] der Mann der Frau den Jungen zeigt
 that the man the.DAT woman the.ACC boy shows
 ‘that the man shows the boy to the woman’
- b. [dass] der Junge der Frau gezeigt wird
 that the boy.NOM the.DAT woman shown is
 ‘that the boy is shown to the woman’
- [dotted](0,1ex)–(1,1ex); just case
 (0,1ex)–(1,1ex); just theta”=role
 [dashed](0,1ex)–(1,1ex); case and theta”=role

Figure 3.15: Case and theta-role assignment in active clauses

The passive morphology blocks the subject and absorbs the structural accusative. The object that would get accusative in the active receives only a semantic role in its base position in the passive, but it does not get the absorbed case. Therefore, it has to move to a position where case can be assigned to it (Chomsky81a). Figure 3.16 on the following page shows how this works for example (42b). This movement”=based analysis works well for English since the underlying object always has to move:

- (43) a. The mother gave [the girl] [a cookie].
 b. [The girl] was given [a cookie] (by the mother).
 c. * It was given [the girl] [a cookie].

²⁰The figure does not correspond to \bar{X} theory in its classic form, since *der Frau* ‘the woman’ is a complement which is combined with V' . In classical \bar{X} theory, all complements have to be combined with V^0 . This leads to a problem in ditransitive structures since the structures have to be binary (see Larson88a for a treatment of double object constructions). Furthermore, in the following figures the verb has been left in V^0 for reasons of clarity. In order to create a well”=formed S' =structure, the verb would have to move to its affix in I^0 . Note also that the assignment of the subject theta-role by the verb crosses a phrase boundary. This problem can be solved by assuming that the subject is generated within the VP, gets a theta role there and then moves to SpecIP. An alternative suggestion was to assume that the VP assigns a semantic role to SpecIP.

[dotted](0,1ex)–(1,1ex);	just case
(0,1ex)–(1,1ex);	just theta”=role
[dashed](0,1ex)–(1,1ex);	case and theta”=role

Figure 3.16: Case and theta-role assignment in passive clauses

(43c) shows that filling the subject position with an expletive is not possible, so the object really has to move. However, **Lenerz**⁷⁷ showed that such a movement is not obligatory in German:

- (44) a. weil das Mädchen dem Jungen den Ball schenkte
because the.NOM girl the.DAT boy the.ACC ball gave
‘because the girl gave the ball to the boy’
- b. weil dem Jungen der Ball geschenkt wurde
because the.DAT boy the.NOM ball given was
‘because the ball was given to the boy’
- c. weil der Ball dem Jungen geschenkt wurde
because the.NOM ball the.DAT boy given was

In comparison to (44c), (44b) is the unmarked order. *der Ball* ‘the ball’ in (44b) occurs in the same position as *den Ball* in (44a), that is, no movement is necessary. Only the case differs. (44c) is, however, somewhat marked in comparison to (44b). So, if one assumed (44c) to be the normal order for passives and (44b) is derived from this by movement of *dem Jungen* ‘the boy’, (44b) should be more marked than (44c), contrary to the facts. To solve this problem, an analysis involving abstract movement has been proposed for cases such as (44b): the elements stay in their positions, but are connected to the subject position and receive their case information from there. (**Grewendorf**^{88a}**Grewendorf**⁹³) assumes that there is an empty expletive pronoun in the subject position of sentences such as (44b) as well as in the subject position of sentences with an impersonal passive such as (45):²¹

- (45) weil heute nicht gearbeitet wird
because today not worked is
‘because there will be no work done today’

²¹See **Koster**^{86a} for a parallel analysis for Dutch as well as **Lohnstein**^{2014a} for a movement”=based account of the passive that also involves an empty expletive for the analysis of the impersonal passive.

A silent expletive pronoun is something that one cannot see or hear and that does not carry any meaning. For discussion of this kind of empty element, see Section ?? and Chapter ??.

In the following chapters, I describe alternative treatments of the passive that do without mechanisms such as empty elements that are connected to argument positions and that seek to describe the passive in a more general, cross-linguistically consistent manner as the suppression of the most prominent argument.

A further question which needs to be answered is why the accusative object does not receive case from the verb. This is captured by a constraint, which goes back to **Burzio86a-u-gekauft** and is therefore referred to as *Burzio's Generalization*.²²

(46) Burzio's Generalization (modified):

If V does not have an external argument, then it does not assign (structural) accusative case.

Koster86a has pointed out that the passive in English cannot be derived by Case Theory since if one allowed empty expletive subjects for English as well as German and Dutch, then it would be possible to have analyses such as the following in (47) where np is an empty expletive:

(47) np was read the book.

Koster rather assumes that subjects in English are either bound by other elements (that is, non-expletive) or lexically filled, that is, filled by visible material.

²²Burzio's original formulation was equivalent to the following: a verb assigns accusative if and only if it assigns a semantic role to its subject. This claim is problematic from both sides. In (i), the verb does not assign a semantic role to the subject; however there is nevertheless accusative case:

- (i) Mich friert.
me.ACC freezes
'I am freezing.'

One therefore has to differentiate between structural and lexical accusative and modify Burzio's Generalization accordingly. The existence of verbs like *begegnen* 'to bump into' is problematic for the other side of the implication. *begegnen* has a subject but still does not assign accusative but rather dative:

- (ii) Peter begegnete einem Mann.
Peter met a.DAT man
'Peter met a man.'

See **Haider99a** and **Webelhuth95a** as well as the references cited there for further problems with Burzio's Generalization.

Therefore, the structure in (47) would be ruled out and it would be ensured that *the book* would have to be placed in front of the finite verb so that the subject position is filled.

3.5 Local reordering

Arguments in the middle field can, in principle, occur in an almost arbitrary order. (48) exemplifies this:

- (48) a. [weil] der Mann der Frau das Buch gibt
 because the man the woman the book gives
 ‘because the man gives the book to the woman’
 b. [weil] der Mann das Buch der Frau gibt
 because the man the book the woman gives
 c. [weil] das Buch der Mann der Frau gibt
 because the book the man the woman gives
 d. [weil] das Buch der Frau der Mann gibt
 because the book the woman the man gives
 e. [weil] der Frau der Mann das Buch gibt
 because the woman the man the book gives
 f. [weil] der Frau das Buch der Mann gibt
 because the woman the book the man gives

In (48b–f), the constituents receive different stress and the number of contexts in which each sentence can be uttered is more restricted than in (48a) (Hoehle82a). The order in (48a) is therefore referred to as the *neutral order* or *unmarked order*.

Two proposals have been made for analyzing these orders: the first suggestion assumes that the five orderings in (48b–f) are derived from a single underlying order by means of Move- α (Frey93a). As an example, the analysis of (48c) is given in Figure ?? on page ?. The object *das Buch* ‘the book’ is moved to the left and

Figure 3.17: Analysis of local reordering as adjunction to IP

adjoined to the topmost IP.

An argument that has often been used to support this analysis is the fact that scope ambiguities exist in sentences with reorderings which are not present in sentences in the base order. The explanation of such ambiguities comes from the assumption that the scope of quantifiers can be derived from their position in the

surface structure as well as their position in the deep structure. If the position in both the surface and deep structure are the same, that is, when there has not been any movement, then there is only one reading possible. If movement has taken place, however, then there are two possible readings (Frey93a):

- (49) a. Es ist nicht der Fall, daß er mindestens einem Verleger fast jedes
 it is not the case that he at.least one publisher almost every
 Gedicht anbot.
 poem offered
 'It is not the case that he offered at least one publisher almost every poem.'
- b. Es ist nicht der Fall, daß er fast jedes Gedicht_i mindestens einem
 it is not the case that he almost every poem at.least one
 Verleger _{-i} anbot.
 publisher offered
 'It is not the case that he offered almost every poem to at least one publisher.'

It turns out that approaches assuming traces run into problems as they predict certain readings for sentences with multiple traces which do not exist (see Kiss2001a and Fanselow2001a). For instance in an example such as (50), it should be possible to interpret *mindestens einem Verleger* 'at least one publisher' at the position of _{-i}, which would lead to a reading where *fast jedes Gedicht* 'almost every poem' has scope over *mindestens einem Verleger* 'at least one publisher'. However, this reading does not exist.

- (50) Ich glaube, dass mindestens einem Verleger_i fast jedes Gedicht_j nur
 I believe that at.least one publisher almost every poem only
 dieser Dichter _{-i -j} angeboten hat.
 this poet offered has
 'I think that only this poet offered almost every poem to at least one publisher.'

SE2002a discuss analogous examples from Japanese, which they credit to Kazuko Yatsushiro. They develop an analysis where the first step is to move the accusative object in front of the subject. Then, the dative object is placed in front of that and then, in a third movement, the accusative is moved once more. The last movement can take place to construct either the S"=structure²³ or as a move-

²³The authors are working in the Minimalist framework. This means there is no longer S"=structure strictly speaking. I have simply translated the analysis into the terms used here.

ment to construct the Phonological Form. In the latter case, this movement will not have any semantic effects. While this analysis can predict the correct available readings, it does require a number of additional movement operations with intermediate steps.

The alternative to a movement analysis is so-called *base generation*: the starting structure generated by phrase structure rules is referred to as the *base*. One variant of base generation assumes that the verb is combined with one argument at a time and each θ -role is assigned in the respective head-argument configuration. The order in which arguments are combined with the verb is not specified, which means that all of the orders in (48) can be generated directly without any transformations.²⁴ Fanselow 2001a suggested such an analysis within the framework of GB.²⁵ Note that such a base-generation analysis is incompatible with an IP approach that assumes that the subject is realized in the specifier of IP. An IP approach with base-generation of different argument orders would allow the complements to appear in any order within the VP but the subject would be first since it is part of a different phrase. So the orders in (51a,b) could be analyzed, but the ones in (51c–f) could not:

- (51) a. dass der Mann der Frau ein Buch gibt
that the.NOM man the.DAT woman a.ACC book gives
b. dass der Mann ein Buch der Frau gibt
that the.NOM man a.ACC book the.DAT woman gives
c. dass der Frau der Mann ein Buch gibt
that the.DAT woman the.NOM man a.ACC book gives
d. dass der Frau ein Buch der Mann gibt
that the.DAT woman a.ACC book the.NOM man gives
e. dass ein Buch der Frau der Mann gibt
that a.ACC book the.DAT woman the.NOM man gives
f. dass ein Buch der Mann der Frau gibt
that a.ACC book the.NOM man the.DAT woman gives

²⁴ Compare this to the grammar in (6) on page 65. This grammar combines a V and an NP to form a new V. Since nothing is said about the case of the argument in the phrase structure rule, the NPs can be combined with the verb in any order.

²⁵ The base generation analysis is the natural analysis in the HPSG framework. It has already been developed by Gunji in 1986 for Japanese and will be discussed in more detail in Section ?? . SE2002a claim that they show that syntax has to be derivational, that is, a sequence of syntactic trees has to be derived. I am of the opinion that this cannot generally be shown to be the case. There is, for example, an analysis by Kiss 2001a which shows that scope phenomena can be explained well by constraint-based approaches.

For the discussion of different approaches to describing constituent position, see Fanselow93a.

3.6 Summary and classification

Works in GB and some contributions to the Minimalist Program (see Chapter 4) have led to a number of new discoveries in both language”=specific and cross”=linguistic research. In the following, I will focus on some aspects of German syntax.

The analysis of verb movement developed in Transformational Grammar by Bierwisch63a, Reis74a, Koster75a, Thiersch78a and denBesten83a has become the standard analysis in almost all grammar models (possibly with the exception of Construction Grammar and Dependency Grammar).

The work by Lenerz77 on constituent order has influenced analyses in other frameworks (the linearization rules in GPSG and HPSG go back to Lenerz’ descriptions). Haider’s work on constituent order, case and passive (Haider84b; Haider85; Haider85b; Haider86; Haider90a; Haider93a) has had a significant influence on LFG and HPSG analyses of German.

The entire configurationality discussion, that is, whether it is better to assume that the subject of finite verbs in German is inside or outside the VP, was important (for instance Haider82; Grewendorf83a; Kratzer84a; Kratzer96a; Webelhuth85a; Sternefeld85b; Scherpenisse86a; Fanselow87a; Grewendorf88a; Duerscheid89a; Webelhuth90; Oppenrieder91a; Wilder91a; Haider93a; Grewendorf93; Frey93a; Lenerz94a; Meinunger2000a) and German unaccusative verbs received their first detailed discussion in GB circles (Grewendorf89a; Fanselow92). The works by Fanselow and Frey on constituent order, in particular with regard to information structure, have advanced German syntax quite considerably (Fanselow88; Fanselow90; Fanselow93a; Fanselow2000a; Fanselow2001a; Fanselow2003d; Fanselow2003a; Fanselow2004a; Frey2000a-u; Frey2001a; Frey2004a; Frey2005a). Infinitive constructions, complex predicates and partial fronting have also received detailed and successful treatments in the GB/MP frameworks (Bierwisch63a; Evers75a; Haider82; Haider86c; Haider90b; Haider91; Haider93a; Grewendorf83a; Grewendorf87a; Grewendorf88a; denBesten85b; Sternefeld85b; Fanselow87a; Fanselow2002a; SS88a; BK89a; G. GMueller96a; GMueller98a; VS98a). In the area of secondary predication, the work by Winkler97a is particularly noteworthy.

This list of works from subdisciplines of grammar is somewhat arbitrary (it corresponds more or less to my own research interests) and is very much focused on German. There are, of course, a wealth of other articles on other languages and phenomena, which should be recognized without having to be individually

listed here.

In the remainder of this section, I will critically discuss two points: the model of language acquisition of the Principles & Parameters framework and the degree of formalization inside Chomskyan linguistics (in particular the last few decades and the consequences this has). Some of these points will be mentioned again in Part ??.

3.6.1 Explaining language acquisition

One of the aims of Chomskyan research on grammar is to explain language acquisition. In GB, one assumed a very simple set of rules, which was the same for all languages (\bar{X} theory), as well as general principles that hold for all languages, but which could be parametrized for individual languages or language classes. It was assumed that a parameter was relevant for multiple phenomena. The Principles & Parameters model was particularly fruitful and led to a number of interesting studies in which commonalities and differences between languages were uncovered. From the point of view of language acquisition, the idea of a parameter which is set according to the input has often been criticized as it cannot be reconciled with observable facts: after setting a parameter, a learner should have immediately mastered certain aspects of that language. **Chomsky86a** uses the metaphor of switches which can be flipped one way or the other. As it is assumed that various areas of grammar are affected by parameters, setting one parameter should have a significant effect on the rest of the grammar of a given learner. However, the linguistic behavior of children does not change in an abrupt fashion as would be expected (**Bloom93a**; **Haider93a**; **Abney96a**; **AW98a**; **Tomasello2000a**; **Tomasello2003a**). Furthermore, it has not been possible to prove that there is a correlation between a certain parameter and various grammatical phenomena. For more on this, see Chapter ??.

The Principles & Parameters model nevertheless remains interesting for cross-linguistic research. Every theory has to explain why the verb precedes its objects in English and follows them in Japanese. One can name this difference a parameter and then classify languages accordingly, but whether this is actually relevant for language acquisition is being increasingly called in question.

3.6.2 Formalization

In his 1963 work on Transformational Grammar, Bierwisch writes the following:²⁶

It is very possible that the rules that we formulated generate sentences which are outside of the set of grammatical sentences in an unpredictable way, that is, they violate grammaticality due to properties that we did not deliberately exclude in our examination. This is meant by the statement that a grammar is a hypothesis about the structure of a language. A systematic check of the implications of a grammar that is appropriate for natural languages is surely a task that cannot be done by hand any more. This task could be solved by implementing the grammar as a calculating task on a computer so that it becomes possible to verify to which degree the result deviates from the language to be described. (Bierwisch63a)

Bierwisch's claim is even more valid in light of the empirical progress made in the last decades. For example, **Ross67a** identified restrictions for movement and long"=distance dependencies and **Perlmutter78** discovered unaccusative verbs in the 70s. For German, see **Grewendorf89a** and **Fanselow92**. Apart from analyses of these phenomena, restrictions on possible constituent positions have been developed (**Lenerz77**), as well as analyses of case assignment (**YMJ87**; **Meurers99b**; **Prze99**) and theories of verbal complexes and the fronting of parts of phrases (**Evers75a**; **Grewendorf88a**; **HN94a**; **Kiss95a**; **G. GMueller98a**; **Meurers99c**; **Mueller99a**; **Mueller2002b**; **deKuthy2002a**). All these phenomena interact!

Consider another quote:

A goal of earlier linguistic work, and one that is still a central goal of the linguistic work that goes on in computational linguistics, is to develop grammars that assign a reasonable syntactic structure to every sentence of English, or as nearly every sentence as possible. This is not a goal that is cur-

²⁶Es ist also sehr wohl möglich, daß mit den formulierten Regeln Sätze erzeugt werden können, die auch in einer nicht vorausgesehenen Weise aus der Menge der grammatisch richtigen Sätze herausfallen, die also durch Eigenschaften gegen die Grammatikalität verstoßen, die wir nicht wissentlich aus der Untersuchung ausgeschlossen haben. Das ist der Sinn der Feststellung, daß eine Grammatik eine Hypothese über die Struktur einer Sprache ist. Eine systematische Überprüfung der Implikationen einer für natürliche Sprachen angemessenen Grammatik ist sicherlich eine mit Hand nicht mehr zu bewältigende Auf"gabe. Sie könnte vorgenommen werden, indem die Grammatik als Rechenprogramm in einem Elektronenrechner realisiert wird, so daß überprüft werden kann, in welchem Maße das Resultat von der zu beschreibenden Sprache abweicht.

rently much in fashion in theoretical linguistics. Especially in Government-Binding theory (GB), the development of large fragments has long since been abandoned in favor of the pursuit of deep principles of grammar. The scope of the problem of identifying the correct parse cannot be appreciated by examining behavior on small fragments, however deeply analyzed. Large fragments are not just small fragments several times over – there is a qualitative change when one begins studying large fragments. As the range of constructions that the grammar accommodates increases, the number of undesired parses for sentences increases dramatically. (Abney96a)

So, as Bierwisch and Abney point out, developing a sound theory of a large fragment of a human language is a really demanding task. But what we aim for as theoretical linguists is much more: the aim is to formulate restrictions which ideally hold for all languages or at least for certain language classes. It follows from this, that one has to have an overview of the interaction of various phenomena in not just one but several languages. This task is so complex that individual researchers cannot manage it. This is the point at which computer implementations become helpful as they immediately flag inconsistencies in a theory. After removing these inconsistencies, computer implementations can be used to systematically analyze test data or corpora and thereby check the empirical adequacy of the theory (Müller, Mueller99a; MuellerCoreGram; MuellerKernigkeit; OF98; Bender2008c, see Section 1.2).

More than 60 years after the first important published work by Chomsky, it is apparent that there has not been one large”=scale implemented grammatical fragment on the basis of Transformational Grammar analyses. Chomsky has certainly contributed to the formalization of linguistics and developed important formal foundations which are still relevant in the theory of formal languages in computer science and in theoretical computational linguistics (Chomsky59a-u). However, in 1981, he had already turned his back on rigid formalization:

I think that we are, in fact, beginning to approach a grasp of certain basic principles of grammar at what may be the appropriate level of abstraction. At the same time, it is necessary to investigate them and determine their empirical adequacy by developing quite specific mechanisms. We should, then, try to distinguish as clearly as we can between discussion that bears on leading ideas and discussion that bears on the choice of specific realizations of them. (Chomsky81a)

This is made explicit in a letter to *Natural Language and Linguistic Theory*:

Even in mathematics, the concept of formalization in our sense was not developed until a century ago, when it became important for advancing research and understanding. I know of no reason to suppose that linguistics is so much more advanced than 19th century mathematics or contemporary molecular biology that pursuit of Pullum's injunction would be helpful, but if that can be shown, fine. For the present, there is lively interchange and exciting progress without any sign, to my knowledge, of problems related to the level of formality of ongoing work. (Chomsky90a)

This departure from rigid formalization has led to there being a large number of publications inside Mainstream Generative Grammar with sometimes incompatible assumptions to the point where it is no longer clear how one can combine the insights of the various publications. An example of this is the fact that the central notion of government has several different definitions (see AS83a for an overview²⁷).

This situation has been criticized repeatedly since the 80s and sometimes very harshly by proponents of GPSG (GKPS85a; Pullum85a; Pullum89b; Pullum91b; KP90a).

The lack of precision and working out of the details²⁸ and the frequent modification of basic assumptions²⁹ has led to insights gained by Mainstream Generative Grammar rarely being translated into computer implementations. There are some implementations that are based on Transformational Grammar/GB/MP models or borrow ideas from Mainstream Generative Grammar (Petrack65a-u; ZFHW65a; Kay67a; Friedman69a; FBDPM71a-u; Plath73a; Morin73a-u; Marcus80a-u; AC86a; Kuhns86a; Corraera87a; Stabler87a; Stabler92a-u; Stabler2001a; KT91a; Fong91a-u; CL92a; Lohnstein93a-u; Lin93a; FC94a; Nordgard94a; Veenstra98a; FG2012a),³⁰ but these implementations often do not use transformations or differ greatly from the theoretical assumptions of the publications. For example, Marcus80a-u and Stabler87a use special purpose rules for auxiliary inversion.³¹

²⁷ A further definition can be found in AL84a-u. This is, however, equivalent to an earlier version as shown by PP86a.

²⁸ See e.g., Kuhns86a, CL92a, KT91a, Kolb97a and Freidin97a-u, Veenstra98a, LLJ2000b and Stabler2010a for the latter.

²⁹ See e.g., Kolb97a, Fanselow2009a and the quote from Stabler on page 189.

³⁰ See FC94a for a combination of a GB approach with statistical methods.

³¹ NF86a-u; NF87a-u has shown that Marcus' parser can only parse context"-free languages. Since natural languages are of a greater complexity (see Chapter ??) and grammars of corresponding complexity are allowed by current versions of Transformational Grammar, Marcus' parser can be neither an adequate implementation of the Chomskyan theory in question nor a piece of software for analyzing natural language in general.

These rules reverse the order of *John* and *has* for the analysis of sentences such as (52a) so that we get the order in (52b), which is then parsed with the rules for non”=inverted structures.

- (52) a. Has John scheduled the meeting for Wednesday?
b. John has scheduled the meeting for Wednesday?

These rules for auxiliary inversion are very specific and explicitly reference the category of the auxiliary. This does not correspond to the analyses proposed in GB in any way. As we have seen in Section 3.1.5, there are no special transformational rules for auxiliary inversion. Auxiliary inversion is carried out by the more general transformation Move- α and the associated restrictive principles. It is not unproblematic that the explicit formulation of the rule refers to the category *auxiliary* as is clear when one views Stabler’s GB”=inspired phrase structure grammar:

- (53) a. $s \rightarrow \text{switch}(\text{aux_verb}, \text{np}), \text{vp}$.
b. $s([\text{First}|\text{L0}], \text{L}, \text{X0}, \text{X}) :- \text{aux_verb}(\text{First}),$
 $\text{np}(\text{L0}, \text{L1}, \text{X0}, \text{X1}),$
 $\text{vp}([\text{First}|\text{L1}], \text{L}, \text{X1}, \text{X}).$

The rule in (53a) is translated into the Prolog predicate in (53b). The expression $[\text{First}|\text{L0}]$ after the s corresponds to the string, which is to be processed. The ‘|’=operator divides the list into a beginning and a rest. *First* is the first word to be processed and L0 contains all other words. In the analysis of (52a), *First* is *has* and L0 is *John scheduled the meeting for Wednesday*. In the Prolog clause, it is then checked whether *First* is an auxiliary ($\text{aux_verb}(\text{First})$) and if this is the case, then it will be tried to prove that the list L0 begins with a noun phrase. Since *John* is an NP, this is successful. L1 is the sublist of L0 which remains after the analysis of L0 , that is *scheduled the meeting for Wednesday*. This list is then combined with the auxiliary (*First*) and now it will be checked whether the resulting list *has scheduled the meeting for Wednesday* begins with a VP. This is the case and the remaining list L is empty. As a result, the sentence has been successfully processed.

The problem with this analysis is that exactly one word is checked in the lexicon. Sentences such as (54) can not be analyzed:³²

³²For a discussion that shows that the coordination of lexical elements has to be an option in linguistic theories, see Abeille2006a.

- (54) Could or should we pool our capital with that of other co-ops to address the needs of a regional “neighborhood”?³³

In this kind of sentence, two modal verbs have been coordinated. They then form an X^0 and – following GB analyses – can be moved together. If one wanted to treat these cases as Stabler does for the simplest case, then we would need to divide the list of words to be processed into two unlimited sub”=lists and check whether the first list contains an auxiliary or several coordinated auxiliaries. We would require a recursive predicate *aux_verbs* which somehow checks whether the sequence *could or should* is a well”=formed sequence of auxiliaries. This should not be done by a special predicate but rather by syntactic rules responsible for the coordination of auxiliaries. The alternative to a rule such as (53a) would be the one in (55), which is the one that is used in theories like GPSG (GKPS85a), LFG (Falk84a-u), some HPSG analyses (GSag2000a-u), and Construction Grammar (Fillmore99a):

- (55) $s \rightarrow v(\text{aux+}), np, vp.$

This rule would have no problems with coordination data like (54) as coordination of multiple auxiliaries would produce an object with the category *v(aux+)* (for more on coordination see Section ??). If inversion makes it necessary to stipulate a special rule like (53a), then it is not clear why one could not simply use the transformation”=less rule in (55).

In the MITRE system (ZFHW65a), there was a special grammar for the surface structure, from which the deep structure was derived via reverse application of transformations, that is, instead of using one grammar to create deep structures which are then transformed into other structures, one required two grammars. The deep structures that were determined by the parser were used as input to a transformational component since this was the only way to ensure that the surface structures can actually be derived from the base structure (Kay2011a).

The REQUEST system by Plath73a also used a surface grammar and inverse transformations to arrive at the deep structure, which was used for semantic interpretation.

There are other implementations discussed in this chapter that differ from transformation”=based analyses. For example, KT91a arrive at the conclusion that a declarative, constraint”=based approach to GB is more appropriate than a derivational one. Johnson89a suggests a *Parsing as Deduction* approach which reformulates sub”=theories of GB (\bar{X} theory, Theta”=Theory, Case Theory, ...) as

³³<http://www.cooperativegrocer.coop/articles/index.php?id=595>. 2010-03-28.

logical expressions.³⁴ These can be used independently of each other in a logical proof. In Johnson's analysis, GB theory is understood as a constraint"-based system. More general restrictions are extracted from the restrictions on S"- and D"-structure which can then be used directly for parsing. This means that transformations are not directly carried out by the parser. As noted by Johnson, the language fragment he models is very small. It contains no description of *wh*"-movement, for example (p. 114).

Lin93a implemented the parser *PrinciParse*. It is written in C++ and based on GB and Barriers – the theoretical stage after GB (see **Chomsky86b**). The system contains constraints like the Case Filter, the Theta-Criterion, Subjacency, the Empty Category Principle and so on. The Theta-Criterion is implemented with binary features +/-theta, there is no implementation of Logical Form (p. 119). The system organizes the grammar in a network that makes use of the object-oriented organization of C++ programs, that is, default-inheritance is used to represent constraints in super and subclasses (**Lin93a**). This concept of inheritance is alien to GB theory: it does not play any role in the main publications. The grammar networks license structures corresponding to \bar{X} theory, but they code the possible relations directly in the network. The network contains categories like IP, Ibar, I, CP, Cbar, C, VP, Vbar, V, PP, PSpec, Pbar, P and so on. This corresponds to simple phrase structure grammars that fully specify the categories in the rules (see Section 2.2) rather than working with abstract schemata like the ones assumed in \bar{X} theory (see Section 2.5). Furthermore Lin does not assume transformations but uses a GPSG-like feature passing approach to nonlocal dependencies (p. 116, see Section 5.4 on the GPSG approach).

Probably the most detailed implementation in the tradition of GB and Barriers is Stabler's Prolog implementation (**Stabler92a-u**). Stabler's achievement is certainly impressive, but his book confirms what has been claimed thus far: Stabler has to simply stipulate many things which are not explicitly mentioned in *Barriers* (e.g., using feature"-value pairs when formalizing \bar{X} theory, a practice that was borrowed from GPSG) and some assumptions cannot be properly formalized and are simply ignored (see **Briscoe97a** for details).

GB analyses which fulfill certain requirements can be reformulated so that they no longer make use of transformations. These transformation"-less approaches are also called *representational*, whereas the transformation"-based approaches are referred to as *derivational*. For representational analyses, there are only surface structures augmented by traces but none of these structures is connected to an underlying structure by means of transformations (see e.g., Koster **Koster78b-u**;

³⁴See **CL92a** and **FC94a** for another constraint"-based Parsing-as-Deduction approach.

Koster87a-u; KT91a; Haider93a; Frey93a; Lohnstein93a-u; FC94a; Veenstra98a). These analyses can be implemented in the same way as corresponding HPSG analyses (see Chapter ??) as computer-processable fragments and this has in fact been carried out for example for the analysis of verb position in German.³⁵ However, such implemented analyses differ from GB analyses with regard to their basic architecture and in small, but important details such as how one deals with the interaction of long”=distance dependencies and coordination (Gazdar81). For a critical discussion and classification of movement analyses in Transformational Grammar, see Borsley2012a.

Following this somewhat critical overview, I want to add a comment in order to avoid being misunderstood: I do not demand that all linguistic work shall be completely formalized. There is simply no space for this in a, say, thirty page essay. Furthermore, I do not believe that all linguists should carry out formal work and implement their analyses as computational models. However, there has to be *somebody* who works out the formal details and these basic theoretical assumptions should be accepted and adopted for a sufficient amount of time by the research community in question.



Comprehension questions

1. Give some examples of functional and lexical categories.
2. How can one represent lexical categories with binary features and what advantages does this have?

³⁵This shows that ten Hacken’s contrasting of HPSG with GB and LFG (TenHacken2007a) and the classification of these frameworks as belonging to different research paradigms is completely mistaken. In his classification, ten Hacken refers mainly to the model”=theoretic approach that HPSG assumes. However, LFG also has a model”=theoretic formalization (Kaplan89b). Furthermore, there is also a model”=theoretic variant of GB (Rogers98a-u). For further discussion, see Chapter ??.



Exercises

1. Draw syntactic trees for the following examples:

- (56) a. dass die Frau den Mann liebt
that the.NOM woman the.ACC man loves
'that the woman loves the man'
- b. dass der Mann geliebt wird
that the.NOM man loved is
'that the man is loved'
- c. Der Mann wird geliebt.
the.NOM man is loved
'The man is loved.'
- d. dass der Mann der Frau hilft
that the.NOM man the.DAT woman helps
'that the man helps the woman'
- e. Der Mann hilft der Frau.
the man.NOM helps the.DAT woman
'The man is helping the woman.'

For the passive sentences, use the analysis where the subject noun phrase is moved from the object position, that is, the analysis without an empty expletive as the subject.

3.6 Summary and classification



Further reading

For Sections 3.1–3.5, I used material from Peter Gallmann from 2003 (**Gallmann2003a**). This has been modified, however, at various points. I am solely responsible for any mistakes or inadequacies. For current materials by Peter Gallmann, see <http://www.syntax-theorie.de>.

In the book *Syntaktische Analyseperspektiven*, **Lohnstein2014a** presents a variant of GB which more or less corresponds to what is discussed in this chapter (CP/IP, movement”=based analysis of the passive). The chapters in said book have been written by proponents of various theories and all analyze the same newspaper article. This book is extremely interesting for all those who wish to compare the various theories out there.

Haegeman94a-u is a comprehensive introduction to GB. Those who do read German may consider the textbooks by **FF87a**, **SS88a** and **Grewendorf88a** since they are also addressing the phenomena that are covered in this book.

In many of his publications, Chomsky discusses alternative, transformation”=less approaches as “notational variants”. This is not appropriate, as analyses without transformations can make different predictions to transformation”=based approaches (e.g., with respect to coordination and extraction. See Section 5.5 for a discussion of GPSG in this respect). In **Gazdar81b**, one can find a comparison of GB and GPSG as well as a discussion of the classification of GPSG as a notational variant of Transformational Grammar with contributions from Noam Chomsky, Gerald Gazdar and Henry Thompson.

Borsley99a-u and **KS2008a-u** have parallel textbooks for GB and HPSG in English. For the comparison of Transformational Grammar and LFG, see **BK82a**. **Kuhn2007a** offers a comparison of modern derivational analyses with constraint”=based LFG and HPSG approaches. **Borsley2012a** contrasts analyses of long”=distance dependencies in HPSG with movement”=based analyses as in GB/Minimalism. Borsley discusses four types of data which are problematic for movement”=based approaches: extraction without fillers, extraction with multiple gaps (see also the discussion of (57) on p. 182 and

of (55) on p. 214 of this book), extractions where fillers and gaps do not match and extraction without gaps.

4 Transformational Grammar – Minimalism

Like the Government & Binding framework that was introduced in the previous chapter, the Minimalist framework was initiated by Noam Chomsky at the MIT in Boston. **Chomsky93b-u**; **Chomsky95a-u** argued that the problem of language evolution should be taken seriously and that the question of how linguistic knowledge could become part of our genetic endowment should be answered. To that end he suggested refocusing the theoretical developments towards models that have to make minimal assumptions regarding the machinery that is needed for linguistic analyses and hence towards models that assume less language specific innate knowledge.

Like GB, Minimalism is wide-spread: theoreticians all over the world are working in this framework, so the following list of researchers and institutions is necessarily incomplete. *Linguistic Inquiry* and *Syntax* are journals that almost exclusively publish Minimalist work and the reader is referred to these journals to get an idea about who is active in this framework. The most prominent researchers in Germany are Artemis Alexiadou, Humboldt University Berlin; Günther **Grewendorf2002a**, Frankfurt am Main; Joseph Bayer, Konstanz; and Gereon Müller, Leipzig.

While innovations like \bar{X} theory and the analysis of clause structure in GB are highly influential and can be found in most of the other theories that are discussed in this book, this is less so for the technical work done in the Minimalist framework. It is nevertheless useful to familiarize with the technicalities since Minimalism is a framework in which a lot of work is done and understanding the basic machinery makes it possible to read empirically interesting work in that framework.

While the GB literature of the 1980s and 1990s shared a lot of assumptions, there was an explosion of various approaches in the Minimalist framework that is difficult to keep track of. The presentation that follows is based on David Adger's textbook (**Adger2003a**).

4.1 General remarks on the representational format

The theories that are developed in the framework of the Minimalist Program build on the work done in the GB framework. So a lot of things that were explained in the previous chapter can be taken over to this chapter. However, there have been some changes in fundamental assumptions. The general parametrized principles were dropped from the theory and instead the relevant distinctions live in features. Languages differ in the values that certain features may have and in addition to this, features may be strong or weak and feature strength is also a property that may vary from language to language. Strong features make syntactic objects move to higher positions. The reader is familiar with this feature”=driven movement already since it was a component of the movement”=based analysis of the passive in Section 3.4. In the GB analysis of passive, the object had to move to the specifier position of IP in order to receive case. Such movements that are due to missing feature values are a key component in Minimalist proposals.

4.1.1 Basic architecture

Chomsky assumes that there are just two operations (rules) for combining linguistic objects: External Merge and Internal Merge. External Merge simply combines two elements like *the* and *book* and results in a complex phrase. Internal Merge is used to account for movement of constituents. It applies to one linguistic object and takes some part of this linguistic object and adjoins it to the left of the respective object. The application of External Merge and Internal Merge can apply in any order. For instance, two objects can be combined with External Merge and then one of the combined items is moved to the left by applying Internal Merge. The resulting object can be externally merged with another object and so on. As an example consider the Determiner Phrase (DP) in (1):¹

- (1) the man who we know

To derive this DP the verb *know* is externally merged with its object *who*. After several intermediate merges that will be discussed below, *know who* will be merged with *we* and finally the *who* is moved to the left by Internal Merge, resulting in *who we know*. This relative clause can be externally merged with *man* and so on.

¹Most researchers working in Minimalism follow Abney87a in assuming that the determiner rather than the noun is the head of nominal structures. Hence sequence like (1) are determiner phrases rather than noun phrases.

So, Minimalist theories differ from GB in not assuming a Deep Structure that is generated by some \bar{X} grammar and a Surface Structure that is derived from the Deep Structure by Move- α . Instead, it is assumed that there is a phase in which External and Internal Merge (combination and movement) apply in any order to derive a certain structure that is then said to be spelled out. It is said that the structure is sent to the interfaces: the articulatory-perceptual system (AP) on the one hand and the conceptual-intentional system (CI) on the other side. AP corresponds to the level of Phonological Form (PF) and CI to the level of Logical Form (LF) in GB. The new architecture is depicted in Figure 4.1 (left figure). Syntax

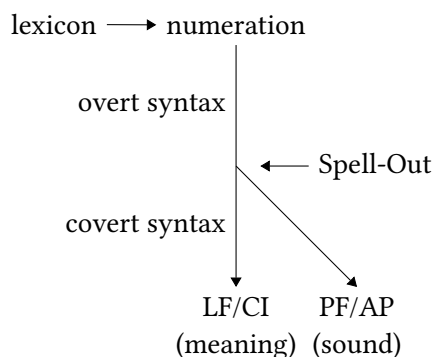


Figure 4.1: Architecture assumed in Minimalist theories before the Phase model (left) and in the Phase model (right) according to **Richards2015a**

is assumed to operate on so-called numerations, selections of lexical items that are relevant for a derivation.² *Overt syntax* stands for syntactic operations that usually have a visible effect. After overt syntax the syntactic object is sent off to the interfaces and some transformations may take place after this Spell-Out point. Since such transformations do not affect pronunciation, this part of syntax is called *covert syntax*. Like in GB's LF, the covert syntax can be used to derive certain scope readings.

This architecture was later modified to allow Spell-Out at several points in the

²It is unclear to me how numerations are determined. Since empty elements play a crucial role in the analysis of sentences in Minimalism and since it is not known which empty elements are needed in an actual analysis until the analysis is carried out, there will be infinitely many numerations that potentially could be used in the analysis of a given string. Which numeration is chosen and how numerations could be integrated in psycholinguistically plausible models of human sentence comprehension is unclear to me. Numerations will be ignored in what follows and it will be assumed that lexical items come from the lexicon directly. See **HNG2005a** for more on numerations.

derivation (right figure). It is now assumed that there are *phases* in a derivation and that a completed phase is spelled out once it is used in a combination with a head (Chomsky2008a). For instance, a subordinated sentence like *that Peter comes* in (2) is one phase and is sent to the interfaces before the whole sentence is completed.³

- (2) He believes that Peter comes.

There are different proposals as to what categories form complete phases. Since the concept of phases is not important for the following introduction, I will ignore this concept in the following. See Section ?? on the psycholinguistic plausibility of phases in particular and the Minimalist architecture in general.

4.1.2 Valence, feature checking, and agreement

The basic mechanism in Minimalist theories is feature checking. For instance, the noun *letters* may have a P feature, which means that it has to combine with a PP in order to form a complete phrase.

- (3) letters to him

It is assumed that there are interpretable and uninterpretable features. An example of an interpretable feature is the number feature of nouns. The singular/plural distinction is semantically relevant. The category features for part of speech information are purely syntactic and hence cannot be interpreted semantically. Minimalism assumes that all uninterpretable features have to be used up during the derivation of a complex linguistic object. This process of eating up the features is called *checking*. As an example, let us consider the noun *letters* again. The analysis of (3) is depicted in Figure 4.2. The fact that the P feature

Figure 4.2: Valence representation via uninterpretable features

of *letters* is uninterpretable is represented by the little *u* in front of the P. The uninterpretable P feature of *letters* can be checked against the P feature of *to him*. *him* is assumed to be of category D since it can appear in the same places as full nominal phrases, which are assumed to be DPs here. All checked features are said to delete automatically. The deletion is marked by striking the features out

³ Andreas Pankau (p. c. 2015) pointed out to me that there is a fundamental problem with such a conception of phases, since if it is the case that only elements that are in a relation to a head are sent off to the interface then the topmost phrase in a derivation would never be sent to the interfaces, since it does not depend on any head.

in the figures. Strings like (4) are ruled out as complete derivations since the D feature of P is not checked. This situation is shown in Figure 4.3.

- (4) * letters to

Figure 4.3: Illegitimate syntactic object due to an uninterpretable feature

If this structure would be used in a larger structure that is spelled out, the derivation would *crash* since the conceptual system could not make sense of the D feature that is still present at the P node.

Selectional features are atomic, that is, the preposition cannot select an DP[*acc*] as in GB and the other theories in this book unless DP[*acc*] is assumed to be atomic. Therefore, an additional mechanism is assumed that can check other features in addition to selectional features. This mechanism is called *Agree*.

- (5) a. * letters to he
b. letters to him

The analysis of (5b) is shown in Figure 4.4. There is an interesting difference be-

Figure 4.4: Feature checking via Agree

tween the checking of selectional features and the checking of features via Agree. The features that are checked via Agree do not have to be at the top node of the object that is combined with a head. This will play a role later in the analysis of the passive and local reordering.

4.1.3 Phrase structure and \bar{X} theory

The projections of \bar{X} structures were given in Figure 2.9 on page 87. According to early versions of the \bar{X} theory, there could be arbitrarily many complements that were combined with X^0 to form an \bar{X} . Arbitrarily many adjuncts could attach to \bar{X} and then at most one specifier could be combined with the \bar{X} yielding an XP. Minimalist theories assume binary branching and hence there is at most one complement, which is the first-merged item. Furthermore, it is not assumed that there is a unique specifier position. Chomsky rather assumes that all items that are not complements are specifiers. That is, he distinguishes between first-merged (complements) and later-merged items (specifiers). Figure 4.5 shows an example with two specifiers. It is also possible to have just a complement and no specifier or to have one or three specifiers. What structures are ultimately licensed depends

on the features of the items that are involved in the Merge operations. Whether a phrasal projection counts as an \bar{X} or an XP depends on whether the phrase is used as a complement or specifier of another head or whether it is used as head in further Merge operations. If a phrase is used as specifier or complement its status is fixed to be a phrase (XP), otherwise the projectional status of resulting phrases is left underspecified. Lexical head daughters in Merge operations have the category X and complex head daughters in Merge operations have the category \bar{X} . This solves the problem that standard \bar{X} theoretic approaches had with pronouns and proper names: a lot of unary branching structure had to be assumed (See left picture in Figure 2.9). This is not necessary any longer in current Minimalist theories.⁴

4.1.4 Little ν

In Section 3.4, I used \bar{X} structures in which a ditransitive verb was combined with its accusative object to form a \bar{V} , which was then combined with the dative object to form a further \bar{V} . Such binary branching structures and also flat structures in which both objects are combined with the verb to form a \bar{V} are rejected by many practitioners of GB and Minimalism since the branching does not correspond to branchings that would be desired for phenomena like the binding of reflexives and negative polarity items. A binding in which *Benjamin* binds *himself* in (6a) is impossible:

- (6) a. *Emily showed himself Benjamin in the mirror.
- b. Peter showed himself Benjamin in the mirror.

What is required for the analysis of Binding and NPI phenomena in theories that analyze these phenomena in terms of tree configurations is that the reflexive pronoun is “higher” in the tree than the proper name *Benjamin*. More precisely, the reflexive pronoun *himself* has to c-command *Benjamin*. c-command is defined as follows (Adger2003a):⁵

⁴For problems with this approach see Brosziewski2003a-u.

⁵c-command also plays a prominent role in GB. In fact, one part of Government & Binding is the Binding Theory, which was not discussed in the previous chapter since binding phenomena do not play a role in this book.

Figure 4.5: Complements and specifiers in Minimalist theories

- (7) A node A c-commands B if, and only if A's sister either:
- a. is B, or
 - b. contains B

In the trees to the left and in the middle of Figure 4.6 the c-command relations are not as desired: in the left-most tree both DPs c-command each other and in the middle one *Benjamin* c-commands *himself* rather than the other way round. Hence it is assumed that the structures at the left and in the middle are inap-

Figure 4.6: Three possible analyses of ditransitives

propriate and that there is some additional structure involving the category v , which is called *little v* (Adger2003a). The sister of *himself* is \bar{V} and \bar{V} contains *Benjamin*, hence *himself* c-commands *Benjamin*. Since the sister of *Benjamin* is V and V neither is nor contains *himself*, *Benjamin* does not c-command *himself*.

The analysis of ditransitives involving an additional verbal head goes back to Larson88a. HK93a-u assume that this verbal head contributes a causative semantics. The structure in Figure 4.7 is derived by assuming that the verb *show* starts out in the V position and then moves to the v position. *show* is assumed to mean *see* and in the position of little v it picks up the causative meaning, which results in a *cause-see'* meaning (Adger2003a).

Figure 4.7: Analysis of ditransitives involving movement to little v

While the verb shell analysis with an empty verbal head was originally invented by Larson88a for the analysis of ditransitive verbs, it is now also used for the analysis of strictly transitive and even intransitive verbs.

Adger2003a argues that semantic roles are assigned uniformly in certain tree configurations:

- (8)
- a. DP daughter of $vP \rightarrow$ interpreted as agent
 - b. DP daughter of VP \rightarrow interpreted as theme
 - c. PP daughter of $\bar{v} \rightarrow$ interpreted as goal

Adger assumes that such uniformly assigned semantic roles help in the process of language acquisition and from this, it follows that little v should also play a role in the analysis of examples with strictly transitive and intransitive verbs. The Figures 4.8 and 4.9 show the analysis of sentences containing the verbs *burn* and *laugh*, respectively.⁶

⁶If all intransitive verbs of this type are supposed to have agents as subjects, a very broad

Figure 4.8: Analysis of strictly transitives involving little *v*

Figure 4.9: Analysis of intransitives involving little *v*

Adger2003a assumes that intransitive and transitive verbs move from V to little *v* as well. This will be reflected in the following figures.

4.1.5 CP, TP, *v*P, VP

Section 3.1.5 dealt with the CP/IP system in GB. In the course of the development of Minimalism, the Inflectional Phrase was split into several functional projections (Chomsky89a-u) of which only the Tense Phrase is assumed in current Minimalist analyses. So, the TP of Minimalism corresponds to IP in the GB analysis. Apart from this change, the core ideas of the CP/IP analysis have been transferred to the Minimalist analysis of English. This subsection will first discuss special features that are assumed to trigger movement (Subsection 4.1.5.1) and then case assignment (Subsection 4.1.5.2).

4.1.5.1 Features as triggers for movement: The EPP feature on T

In GB approaches, the modals and auxiliaries were analyzed as members of the category I and the subjects as specifiers of IP. In the previous section, I showed how subjects are analyzed as specifiers of *v*P. Now, if one assumes that a modal verb combines with such a *v*P, the subject follows the modal, which does not correspond to the order that is observable in English. This problem is solved by assuming a strong uninterpretable D feature at T. Since the feature is strong, a suitable D has to move to the specifier of T and check the D locally. Figure 4.10 shows the TP that plays a role in the analysis of (9):

- (9) Anna will read the book.

Figure 4.10: Analysis of *Anna will read the book.* involving a modal and movement of the subject from *v* to T

The Determiner Phrase (DP) *the book* is the object of *read* and checks the D feature of *read*. little *v* selects for the subject *Anna*. Since T has a strong D feature

conception of agent has to be assumed that also subsumes the subject of verbs like *sleep*. Usually sleeping is not an activity that is performed intentionally.

(marked by an asterisk '*'), *Anna* must not remain inside of the *vP* but moves on to the specifier position of TP.

Full sentences are CPs. For the analysis of (9), an empty C head is assumed that is combined with the TP. The empty C contributes a clause type feature Decl. The full analysis of (9) is shown in Figure 4.11.

Figure 4.11: Analysis of *Anna will read the book.* as CP with an empty C with the clause-type feature Decl

The analysis of the question in (10) involves an unvalued clause-type feature on T for the sentence type *question*.

(10) What will Anna read?

The empty complementizer C has a Q feature that can value the clause-type feature on T. Since clause-type features on T that have the value Q are stipulated to be strong, the T element has to move to C to check the feature locally. In addition, the *wh* element is moved. This movement is enforced by a strong *wh* feature on C. The analysis of (10) is given in Figure 4.12.

Figure 4.12: Analysis of *What will Anna read?* with an empty C with a strong *wh* feature

4.1.5.2 Case assignment

In the GB analysis that was presented in Chapter 3, nominative was assigned by (finite) I and the other cases by the verb (see Section 3.4.2). The assignment of nominative is taken over to Minimalist analyses, so it is assumed that nominative is assigned by (finite) T. However, in the Minimalist theory under consideration, there is not a single verb projection, but there are two verbal projections: *vP* and VP. Now, one could assume that V assigns accusative to its complement or that *v* assigns accusative to the complement of the verb it dominates. Adger2003a assumes the latter approach, since it is compatible with the analysis of so-called unaccusative verbs and the passive. Figure 4.13 shows the TP for (11):

(11) Anna reads the book.

Figure 4.13: Case assignment by T and *v* in the TP for of *Anna reads the book*.

The two DPs *Anna* and *the book* start out with unvalued uninterpretable case features: [*ucase:*]. The features get valued by T and *v*. It is assumed that only one feature is checked by Merge, so this would be the D feature on T, leaving the case feature for the other available checking mechanism: Agree. Agree can be used to check features in sister nodes, but also features further away in the tree. The places that are possible candidates for Agree relations have to stand in a certain relation to each other. The first node has to c-command the node it Agrees with. c-command roughly means: one node up and then arbitrarily many nodes down. So *v* c-commands VP, V, the DP *the book*, and all the nodes within this DP. Since Agree can value features of c-commanded nodes, the accusative on *v* can value the case feature of the DP *the book*.

The non-locality that is built into Agree raises a problem: why is it that (12) is ungrammatical?

(12) * Him likes she.

The accusative of *v* could be checked with its subject and the nominative of T with the object of *likes*. Both DPs stand in the necessary c-command relations to T and *v*. This problem is solved by requiring that all Agree relations have to involve the closest possible element. Adger2003a formulates this constraint as follows:

(13) Locality of matching: Agree holds between a feature F on X and a matching feature F on Y if and only if there is no intervening Z[F].

Intervention is defined as in (14):

(14) Intervention: In a structure [X ... Z ... Y], Z intervenes between X and Y iff X c-commands Z and Z c-commands Y.

So, since T may Agree with *Anna* it must not Agree with *the book*. Hence nominative assignment to *she* in (12) is impossible and (12) is correctly ruled out.

4.1.6 Adjuncts

Adger2003a assumes that adjuncts attach to XP and form a new XP. He calls this operation *Adjoin*. Since this operation does not consume any features it is different from External Merge and hence a new operation would be introduced into the theory, contradicting Chomsky's claim that human languages use only Merge as a structure building operation. There are proposals to treat adjuncts as elements in special adverbial phrases with empty heads (see Section 4.6.1) that are also assumed to be part of a hierarchy of functional projections. Personally,

I prefer Adger’s solution that corresponds to what is done in many other frameworks: there is a special rule or operation for the combination of adjuncts and heads (see for instance Section ?? on the HPSG schema for head adjunct combinations).

4.2 Verb position

The analysis of verb first sentences in German is straightforward, given the machinery that was introduced in the previous section. The basic idea is the same as in GB: the finite verb moves from V to *v* to T and then to C. The movement to T is forced by a strong tense feature on T and the movement of the T complex to C is enforced by a clause-type feature on T that is valued as a strong Interrogative feature (Int) or a strong Declarative feature (Decl) by C. The analysis of the interrogative clause in (15) is shown in Figure 4.14 on the next page.

- (15) Kennt jeder diesen Mann?
 knows everybody this man
 ‘Does everybody know this man?’

Figure 4.14: Analysis of *Kennt jeder diesen Mann?* ‘Does everybody know this man?’ following the analysis of Adger2003a

4.3 Long”=distance dependencies

Having explained the placement of the verb in initial position, the analysis of V2 sentences does not come with a surprise: Adger2003a assumes a feature that triggers the movement of a constituent to a specifier position of C. Adger calls this feature *top*, but this is a misnomer since the initial position in German declarative sentences is not restricted to topics (maybe he would assume an alternative C head with a *foc* feature). Figure 4.15 shows the analysis of (16):

- (16) Diesen Mann kennt jeder.
 this man knows everybody
 ‘Everybody knows this man.’

As in the verb-initial clause in Figure 4.14 a feature on C triggers verb movement. This time it is a Decl feature since we are dealing with a declarative clause. The *top* feature triggers movement of *diesen Mann* ‘this man’ to the specifier position of C.

Figure 4.15: Analysis of *Diesen Mann kennt jeder*. ‘This man, everybody knows.’ following the analysis of Adger2003a

4.4 Passive

Adger2003a suggests an analysis for the passive in English, which I adapted here to German. Like in the GB analysis that was discussed in Section 3.4 it is assumed that the verb does not assign accusative to the object of *schlagen* ‘to beat’. In Minimalist terms, this means that little *v* does not have an *acc* feature that has to be checked. This special version of little *v* is assumed to play a role in the analysis of sentences of so-called unaccusative verbs (Perlmutter78). Unaccusative verbs are a subclass of intransitive verbs that have many interesting properties. For instance, they can be used as adjectival participles although this is usually not possible with intransitive verbs:

- (17) a. * der getanzte Mann
 the danced man
 b. der gestorbene Mann
 the died man
 ‘the dead man’

The explanation of this difference is that adjectival participles predicate over what is the object in active sentences:

- (18) a. dass der Mann das Buch gelesen hat
 that the man the book read has
 ‘that the man read the book’
 b. das gelesene Buch
 the read book

Now the assumption is that the argument of *gestorben* ‘died’ behaves like an object, while the argument of *getanzt* ‘danced’ behaves like a subject. If adjectival passives predicate over the object it is explained why (17b) is possible, while (17a) is not.

Adger2003a assumes the structure in Figure 4.16 for *v*Ps with unaccusative verbs. It is assumed that this unaccusative variant of little *v* plays a role in the

Figure 4.16: Structure of *v*P with unaccusative verbs like *fall*, *collapse*, *wilt* according to Adger2003a

analysis of the passive. Unaccusative verbs are similar to passivized verbs in that they do have a subject that somehow also has object properties. The special version of little *v* is selected by the Passive head *werden* ‘be’, which forms a Passive Phrase (abbreviated as PassP). See Figure 4.17 for the analysis of the example in (19):

- (19) *dass er geschlagen wurde*
 that he beaten was
 ‘that he was beaten’

Figure 4.17: Minimalist analysis of the passive without movement but with non-local case assignment via Agree

The Pass head requires the Infl feature of little *v* to have the value Pass, which results in participle morphology at spellout. Hence the form that is used is *geschlagen* ‘beaten’. The auxiliary moves to T to check the strong Infl feature at T and since the Infl feature is past, the past form of *werden* ‘be’, namely *wurde* ‘was’, is used at spellout. T has a nom feature that has to be checked. Interestingly, the Minimalist approach does not require the object of *schlagen* to move to the specifier position of T in order to assign case, since case assignment is done via Agree. Hence in principle, the pronominal argument of *schlagen* could stay in its object position and nevertheless get nominative from T. This would solve the problem of the GB analysis that was pointed out by Lenerz77. See page 122 for Lenerz’ examples and discussion of the problem. However, Adger2003a assumes that German has a strong EPP feature on T. If this assumption is upheld, all problems of the GB account will carry over to the Minimalist analysis: all objects have to move to T even when there is no reordering taking place. Furthermore, impersonal passives of the kind in (20) would be problematic, since there is no noun phrase that could be moved to T in order to check the EPP feature:

- (20) *weil getanzt wurde*
 because danced was
 ‘because there was dancing there’

4.5 Local reordering

Adger2003a does not treat local reordering. But there are several other suggestions in the literature. Since all reorderings in Minimalist theories are feature-

driven, there must be an item that has a feature that triggers reorderings like those in (21b):

- (21) a. [weil] jeder diesen Mann kennt
 because everyone this man knows
 ‘because everyone knows this man’
 b. [weil] diesen Mann jeder kennt
 because this man everyone knows

There have been various suggestions involving functional projections like Topic Phrase (Laenzlinger2004a) or AgrS and AgrO (Meinunger2000a) that offer places to move to. G. GMueller2014a-u offers a leaner solution, though. In his approach, the object simply moves to a second specifier position of little *v*. The analysis is depicted in Figure 4.18.⁷

Figure 4.18: Analysis of *dass diesen Mann jeder kennt* ‘that everybody knows this man’ as movement of the object to a specifier position of *v*

An option that was suggested by Laenzlinger2004a is to assume several Object Phrases for objects that may appear in any order. The objects move to the specifier positions of these projections and since the order of the Object Phrases is not restricted, both orders in (22) can be analyzed:

- (22) a. dass Hans diesen Brief meinem Onkel gibt
 that Hans this letter my uncle gives
 ‘that Hans gives this letter to my uncle’
 b. dass Hans meinem Onkel diesen Brief gibt
 that Hans my uncle this letter gives
 ‘that Hans gives to my uncle this letter’

4.6 New developments and theoretical variants

At the start of the 90s, Chomsky suggested a major rethink of the basic theoretical assumptions of GB and only keeping those parts of the theory which are absolutely necessary. In the *Minimalist Program*, Chomsky gives the central motivations for the far-reaching revisions of GB theory (Chomsky93b-u;

⁷G. Müller assumes optional features on *v* and *V* that trigger local reorderings (p. 48). These are not given in the figure.

Chomsky95a-u). Until the beginning of the 90s, it was assumed that Case Theory, the Theta=Criterion, \bar{X} theory, Subjacency, Binding Theory, Control Theory etc. all belonged to the innate faculty for language (**Richards2015a**). This, of course, begs the question of how this very specific linguistic knowledge made its way into our genome. The Minimalist Program follows up on this point and attempts to explain properties of language through more general cognitive principles and to reduce the amount of innate language=specific knowledge postulated. The distinction between Deep Structure and Surface Structure, for example, was abandoned. Move still exists as an operation, but can be used directly to build sub=structures rather than after a complete D=structure has been created. Languages differ with regard to whether this movement is visible or not.

Although Chomsky's Minimalist Program should be viewed as a successor to GB, advocates of Minimalism often emphasize the fact that Minimalism is not a theory as such, but rather a research program (Chomsky **Chomsky2007a**; **Chomsky2013a**). The actual analyses suggested by **Chomsky95a-u** when introducing the research program have been reviewed by theoreticians and have sometimes come in for serious criticism (**Kolb97a**; **JL97a-u-platte**; **JL99a-u**; **LLJ2000b**; **LLJ2000a**; **LLJ2001a**; **Seuren2004a**; **PJ2005a**), however, one should say that some criticisms overshoot the mark.

There are various strains of Minimalism. In the following sections, I will discuss some of the central ideas and explain which aspects are regarded problematic.

4.6.1 **Move, Merge, feature=driven movement and functional projections**

Johnson, Lappin and Kolb have criticized the computational aspects of Chomsky's system. Chomsky suggested incorporating principles of economy into the theory. In certain cases, the grammatical system can create an arbitrary number of structures, but only the most economical, that is, the one which requires the least effort to produce, will be accepted as grammatical (transderivational economy). This assumption does not necessarily have to be taken too seriously and, in reality, does not play a role in many works in the Minimalist framework (although see **Richards2015a** for recent approaches with derivations which are compared in terms of economy). Nevertheless, there are other aspects of Chomsky's theory which can be found in many recent works. For example, Chomsky has proposed reducing the number of basic, structure building operations which license structures to two: Move and Merge (that is, Internal and External Merge). Move corresponds to the operation Move- α , which was already discussed in Chapter 3,

and Merge is the combination of (two) linguistic objects.

It is generally assumed that exactly two objects can be combined (Chomsky95a-u). For Move, it is assumed that there must be a reason for a given movement operation. The reason for movement is assumed to be that an element can check some feature in the position it is moved to. This idea was already presented in the analysis of the passive in Section 3.4: the accusative object does not bear case in passive sentences and therefore has to be moved to a position where it can receive case. This kind of approach is also used in newer analyses for a range of other phenomena. For example, it is assumed that there are phrases whose heads have the categories focus and topic. The corresponding functional heads are always empty in languages like German and English. Nevertheless, the assumption of these heads is motivated by the fact that other languages possess markers which signal the topic or focus of a sentence morphologically. This argumentation is only possible if one also assumes that the inventory of categories is the same for all languages. Then, the existence of a category in one language would suggest the existence of the same category in all other languages. This assumption of a shared universal component (Universal Grammar, UG) with detailed language”=specific knowledge is, however, controversial and is shared by few linguists outside of the Chomskyan tradition. Even for those working in Chomskyan linguistics, there have been questions raised about whether it is permissible to argue in this way since if it is only the ability to create recursive structures that is responsible for the human-specific ability to use language (faculty of language in the narrow sense) – as HCF2002a assume –, then the individual syntactic categories are not part of UG and data from other languages cannot be used to motivate the assumption of invisible categories in another language.

4.6.1.1 Functional projections and modularization of linguistic knowledge

The assumption that movement must be licensed by feature checking has led to an inflation of the number of (silent) functional heads.⁸ Rizzi97a-u suggests the structure in Figure 4.19 (see also Grewendorf Grewendorf2002a; Grewendorf2009a).

The functional categories Force, Top, Foc and Fin correspond to clause type,

⁸The assumption of such heads is not necessary since features can be “bundled” and then they can be checked together. For an approach in this vein, which is in essence similar to what theories such as HPSG assume, see Sternefeld2006a-u.

In so”=called cartographic approaches, it is assumed that every morphosyntactic feature corresponds to an independent syntactic head (CR2010a). For an explicitly formalized proposal in which exactly one feature is consumed during a combination operation see Stabler2001a. Stabler’s *Minimalist Grammars* are discussed in more detail in Section 4.6.4.

Figure 4.19: Syntactic structure of sentences following Rizzi97a-u

topic, focus and finiteness. It is assumed that movement always targets a specifier position. Topics and focused elements are always moved to the specifier position of the corresponding phrase. Topics can precede or follow focused elements, which is why there are two topic projections: one above and one below FocP. Topic phrases are recursive, that is, an arbitrary number of TopPs can appear at the positions of TopP in the figure. Following Grewendorf2002a, topic and focus phrases are only realized if they are required for particular information structural reasons, such as movement.⁹ Chomsky95a-u follows Pollock89a-u in assuming that all languages have functional projections for subject and object agreement as well as negation (AgrS, AgrO, Neg).¹⁰ Sternefeld95a, Stechow96a and Meinunger2000a differentiate between two agreement positions for direct and indirect objects (AgrO, AgrIO). As well as AgrS, AgrO and Neg, BS97a-u assume the functional heads Share and Dist in order to explain scope phenomena in English as feature"-driven movements at LF. For a treatment of scope phenomena without empty elements or movement, see Section ?? . BG2005a assume the categories -PolP, +PolP and %PolP for their discussion of polarity.

Webelhuth95a gives an overview of the functional projections that had been proposed up to 1995 and offers references for AgrA, AgrN, AgrV, Aux, Clitic Voices, Gender, Honorific, μ , Number, Person, Predicate, Tense, Z.

In addition to AdvP, NegP, AgrP, FinP, TopP and ForceP, WHBH2007a-u postulate an OuterTopP. Poletto2000a-u suggests both a HearerP and a SpeakerP for the position of clitics in Italian. BB2011b-u assume a BenefactiveP

Cinque99a-u adopts the 32 functional heads in Table 4.1 in his work. He assumes that all sentences contain a structure with all these functional heads. The specifier positions of these heads can be occupied by adverbs or remain empty. Cinque claims that these functional heads and the corresponding structures form part of Universal Grammar, that is, knowledge of these structures is innate (page 107).¹¹

⁹There are differing opinions as to whether functional projections are optional or not. Some authors assume that the complete hierarchy of functional projections is always present but functional heads can remain empty (e.g., Cinque99a-u and CR2010a).

¹⁰See Chomsky95a-u, however.

¹¹Table 4.1 shows only the functional heads in the clausal domain. Cinque94a-u also accounts for the order of adjectives with a cascade of projections: Quality, Size, Shape, Color, Nationality. These categories and their ordering are also assumed to belong to UG (p. 100).

Cinque94a-u claims that a maximum of seven attributive adjectives are possible and explains this with the fact that there are a limited number of functional projections in the nominal domain. As was shown on page 76, with a fitting context it is possible to use several adjectives

Table 4.1: Functional heads following Cinque99a-u

1. Mood _{Speech Act}	2. Mood _{Evaluative}	3. Mood _{Evidential}	4. Mood _{Epistemic}
5. T(Past)	6. T(Future)	7. Mood _{Irrealis}	8. Mod _{Necessity}
9. Mod _{Possibility}	10. Mod _{Volitional}	11. Mod _{Obligation}	12. Mod _{Ability/permission}
13. Asp _{Habitual}	14. Asp _{Repetitive(I)}	15. Asp _{Frequentative(I)}	16. Asp _{Celerative(I)}
17. T(Anterior)	18. Asp _{Terminative}	19. Asp _{Continuative}	20. Asp _{Perfect(?)}
21. Asp _{Retrospective}	22. Asp _{Proximative}	23. Asp _{Durative}	24. Asp _{Generic/progressive}
25. Asp _{Prospective}	26. Asp _{SgCompletive(I)}	27. Asp _{PlCompletive}	28. Asp _{Voice}
29. Asp _{Celerative(II)}	30. Asp _{SgCompletive(II)}	31. Asp _{Repetitive(II)}	32. Asp _{Frequentative(II)}

Laenzlinger2004a follows Cinque in proposing this sequence of functional heads for German. He also follows Kayne94a-u, who assumes that all syntactic structures have the order specifier head complement cross”=linguistically, even if the surface order of the constituents seems to contradict this.

The constituent orders that are visible in the end are derived by leftward”=movement.¹² Figure 4.20 shows the analysis of a verb”=final clause where the functional adverbial heads have been omitted.¹³ Subjects and objects are generated as arguments inside of vP and VP, respectively. The subject is moved to the specifier of the subject phrase and the object is moved to the specifier of the object phrase. The

of the same kind, which is why some of Cinque’s functional projections would have to be subject to iteration.

¹²This also counts for extraposition, that is, the movement of constituents into the postfield in German. Whereas this would normally be analyzed as rightward”=movement, Kayne94a-u analyzes it as movement of everything else to the left. Kayne assumes that (i.b) is derived from (i.a) by moving part of the DP:

- (i) a. just walked into the room [_{DP} someone who we don’t know].
- b. Someone_i just walked into the room [_{DP} _i who we don’t know].

(i.a) must have to be some kind of derived intermediate representation, otherwise English would not be SV(O) underlyingly but rather V(O)S. (i.a) is therefore derived from (ii) by fronting the VP *just walked into the room*.

- (ii) Someone who we don’t know just walked into the room

Such analyses have the downside that they cannot be easily combined with performance models (see Chapter ??).

¹³These structures do not correspond to \bar{X} theory as it was presented in Section 2.5. In some cases, heads have been combined with complements to form an XP rather than an X’. For more on \bar{X} theory in the Minimalist Program, see Section 4.6.3.

Figure 4.20: Analysis of sentence structure with leftward remnant movement and functional heads following **Laenzlinger2004a**

verbal projection (VP_k) is moved in front of the auxiliary into the specifier position of the phrase containing the auxiliary. The only function of SubjP and ObjP is to provide a landing site for the respective movements. For a sentence in which the object precedes the subject, Laenzlinger assumes that the object moves to the specifier of a topic phrase. Figure 4.20 contains only a ModP and an AspP, although Laenzlinger assumes that all the heads proposed by Cinque are present in the structure of all German clauses. For ditransitive verbs, Laenzlinger assumes multiple object phrases (page 230). A similar analysis with movement of object and subject from verb"-initial VPs to Agr positions was suggested by **Zwart1994a-u** for Dutch.

For general criticism of Kayne's model, see **Haider2000a**. Haider shows that a Kayne"-like theory makes incorrect predictions for German (for instance regarding the position of selected adverbials and secondary predicates and regarding verbal complex formation) and therefore fails to live up to its billing as a theory which can explain all languages. **Haider97a** has shown that the assumption of an empty Neg head, as assumed by **Pollock89a-u**, **Haegeman95a-u** and others, leads to problems. See **Bobaljik99a** for problems with the argumentation for Cinque's cascade of adverb"-projections.

Furthermore, it has to be pointed out that SubjP and ObjP, TraP (Transitive Phrase) and IntraP (Intransitive Phrase) (**Karimi-Doostan2005a**) and TopP (topic phrase), DistP (quantifier phrase), AspP (aspect phrase) (**EKiss2003a-u**; **Karimi2005a**), PathP and PlaceP (**Svenonius2004a-u**) encode information about grammatical function, valence, information structure and semantics in the category symbols.¹⁴ In a sense, this is a misuse of category symbols, but such a misuse of information structural and semantic categories is necessary since syntax, semantics, and information structure are tightly connected and since it is assumed that the semantics interprets the syntax, that is, it is assumed that semantics comes after syntax (see Figure 3.2 and Figure 4.1). By using semantically and pragmatically relevant categories in syntax, there is no longer a clean distinction between the levels of morphology, syntax, semantics and pragmatics: everything has been 'syntactified'. **Rizzi2014a** himself talks about syntactification. He points

¹⁴For further examples and references, see Newmeyer (**Newmeyer2004b**; **Newmeyer2005a**). Newmeyer references also works which stipulate a projection for each semantic role, e.g., Agent, Reciprocal, Benefactive, Instrumental, Causative, Comitative, and Reversive Phrase.

out that there are fundamental problems with the T-model and its current variants in Minimalism and concludes that a syntactification in terms of Rizzi-style functional heads or a proliferation of T heads with respective features (see also **Borsley2006a**; **BM2020Minimalism**) is the only way to save this architecture. Felix Bildhauer (p. c. 2012) has pointed out to me that approaches which assume a cascade of functional projections where the individual aspects of meaning are represented by nodes are actually very close to phrasal approaches in Construction Grammar (see **Adger2013a** also for a similar view). One simply lists configurations and these are assigned a meaning (or features which are interpreted post-syntactically, see **CR2010a** for the interpretation of TopP, for example).

4.6.1.2 Feature checking in specifier positions

If one takes the theory of feature checking in Specifier”=Head relations to its logical conclusion, then one arrives at an analysis such as the one suggested by **Radford97a-u**. Radford assumes that prepositions are embedded in an Agreement Phrase in addition to the structure in (23), which is usually assumed, and that the preposition adjoins to the head of the Agreement Phrase and the argument of the preposition is moved to the specifier position of the Agreement Phrase.

(23) [pp P DP]

The problem here is that the object now precedes the preposition. In order to rectify this, Radford assumes a functional projection p (read *little p*) with an empty head to which the preposition then adjoins. This analysis is shown in Figure 4.21. This machinery is only necessary in order to retain the assumption that feature

Figure 4.21: PP analysis following Radford with case assignment in specifier position and little p

checking takes place in specifier”=head relations. If one were to allow the preposition to determine the case of its object locally, then all this theoretical apparatus would not be necessary and it would be possible to retain the well-established structure in (23).

Sternefeld2006a-u is critical of this analysis and compares it to Swiss cheese (being full of holes). The comparison to Swiss cheese is perhaps even too positive since, unlike Swiss cheese, the ratio of substance to holes in the analysis is extreme (2 words vs. 5 empty elements). We have already seen an analysis of noun phrases on page 81, where the structure of an NP, which only consisted

of an adjective *klugen* ‘clever’, contained more empty elements than overt ones. The difference to the PP analysis discussed here is that empty elements are only postulated in positions where overt determiners and nouns actually occur. The little *p* projection, on the other hand, is motivated entirely theory”=internally. There is no theory”=external motivation for any of the additional assumptions made for the analysis in Figure 4.21 (see Sternefeld2006a-u).

A variant of this analysis has been proposed by HNG2005a. The authors do without little *p*, which makes the structure less complex. They assume the structure in (24), which corresponds to the AgrOP”=subtree in Figure 4.21.

$$(24) \quad [_{AgrP} DP_k [_{Agr'} P_i + Agr [_{PP} t_i t_k]]]$$

The authors assume that the movement of the DP to SpecAgrP happens invisibly, that is, covert. This solves Radford’s problem and makes the assumption of *pP* redundant.

The authors motivate this analysis by pointing out agreement phenomena in Hungarian: Hungarian postpositions agree with the preceding noun phrase in person and number. That is, the authors argue that English prepositional and Hungarian postpositional phrases have the same structure derived by movement, albeit the movement is covert in English.

In this way, it is possible to reduce the number and complexity of basic operations and, in this sense, the analysis is minimal. These structures are, however, still incredibly complex. No other kind of theory discussed in this book needs the amount of inflated structure to analyze the combination of a preposition with a noun phrase. The structure in (24) cannot be motivated by reference to data from English and it is therefore impossible to acquire it from the linguistic input. A theory which assumes this kind of structures would have to postulate a Universal Grammar with the information that features can only be checked in (certain) specifier positions (see Chapters ?? and ?? for more on Universal Grammar and language acquisition). For general remarks on (covert) movement see Haider2016a.

4.6.1.3 Locality of selection and functional projections

Another problem arises from the use of functional heads to encode linear order. In the classic CP/IP”=system and all other theories discussed here, a category stands for a class of objects with the same distribution, that is, NP (or DP) stands for pronouns and complex noun phrases. Heads select phrases with a certain category. In the CP/IP”=system, I selects a VP and an DP/NP, whereas C selects an IP. In newer analyses, this kind of selectional mechanism does not work as

easily. Since movement has taken place in (25b), we are dealing with a TopP or FocP in *das Buch dem Mann zu geben* ‘the book the man to give’. Therefore, *um* cannot simply select a non”=finite IP, but rather has to disjunctively be able to select a TopP, FocP or IP. It has to be ensured that TopPs and FocPs are marked with regard to the form of the verb contained inside them, since *um* can only be combined with *zu*=infinitives.

- (25) a. *um dem Mann das Buch zu geben*
 for the man the book to give
 ‘to give the man the book’
 b. *um das Buch dem Mann zu geben*
 for the book the man to give
 ‘to give the book to the man’

The category system, selectional mechanisms and projection of features would therefore have to be made considerably more complicated when compared to a system which simply base generates the orders or a system in which a constituent is moved out of the IP, thereby creating a new IP.

Proposals that follow **Cinque99a-u** are problematic for similar reasons: Cinque assumes the category AdverbP for the combination of an adverb and a VP. There is an empty functional head, which takes the verbal projection as its complement and the adverb surfaces in the specifier of this projection. In these systems, adverb phrases have to pass on inflectional properties of the verb since verbs with particular inflectional properties (finiteness, infinitives with *zu*, infinitives without *zu*, participles) have to be selected by higher heads (see page 197 and Section ??). There is of course the alternative to use Agree for this, but then all selection would be nonlocal and after all selection is not agreement. For further, more serious problems with this analysis like modification of adverbs by adverbs in connection with partial fronting and restrictions on non”=phrasality of preverbal adverbials in English, see **Haider97a**.

A special case of the adverb problem is the negation problem: **Ernst92a** studied the syntax of negation more carefully and pointed out that negation can attach to several different verbal projections (26a,b), to adjectives (26c) and adverbs (26d).

- (26) a. Ken could not have heard the news.
 b. Ken could have not heard the news.
 c. a [not unapproachable] figure
 d. [Not always] has she seasoned the meat.

If all of these projections are simply NegPs without any further properties (about verb form, adjective part of speech, adverb part of speech), it would be impossible to account for their different syntactic distributions. Negation is clearly just a special case of the more general problem, since adverbs may attach to adjectives forming adjectival phrases in the traditional sense and not adverb phrases in Chinque's sense. For instance, the adverb *oft* 'often' in (27) modifies *lachender* 'laughing' forming the adjectival phrase *oft lachender*, which behaves like the unmodified adjectival participle *lachender*: it modifies *Mann* 'man' and it precedes it.

- (27) a. ein lachender Mann
 a laughing man
 'a laughing man'
 b. ein oft lachender Mann
 a often laughing man
 'a man that laughs often'

Of course one could imagine solutions to the last three problems that use the Agree relation to enforce selectional constraints nonlocally, but such accounts would violate locality of selection (see **Ernst92a** and the discussion in Section ?? of this book) and would be much more complicated than accounts that assume a direct selection of dependents.

Related to the locality issues that were discussed in the previous paragraph is the assumption of special functional projections for the placement of clitics: if one uses SpeakerP so that a clitic for first person singular can be moved to the correct specifier positions and a HearerP so that the clitic for second person can be moved to the correct position (**Poletto2000a-u**), then what one has are special projections which need to encode in addition all features that are relevant for clauses (alternatively one could of course assume nonlocal Agree to be responsible for distributional facts). In addition to these features, the category labels contain information that allows higher heads to select clauses containing clitics. In other approaches and earlier variants of transformational grammar, selection was assumed to be strictly local so that higher heads only have access to those properties of embedded categories that are directly relevant for selection (**Abraham2005a; Sag2007a**) and not information about whether an argument of a head within the clause is the speaker or the hearer or whether some arguments in the clause are realized as clitics. Locality will be discussed further in Section ??.

4.6.1.4 Feature-driven movement

Finally, there is a conceptual problem with feature"-driven movement, which has been pointed out by Gisbert Fanselow: **Frey2004a** assumes a KontrP (contrastive phrase) and **Frey2004b-u** a TopP (topic phrase) (see **Rizzi97a-u** for TopP and FocP (focus phrase) in Italian and **Haftka95a**, **Grewendorf** (**Grewendorf2002a**); **Grewendorf2009a**, **Abraham2003a**, **Laenzlinger2004a** and **Hinterhoelzl2004a** for analyses of German with TopP and/or FocP). Constituents have to move to the specifier of these functional heads depending on their information structural status. **Fanselow2003b** has shown that such movement"-based theories for the ordering of elements in the middle field are not compatible with current assumptions of the Minimalist Program. The reason for this is that sometimes movement takes place in order to create space for other elements (altruistic movement). If the information structure of a sentence requires that the closest object to a verb is neither focused nor part of the focus, then the object closest to the verb should not receive the main stress in the clause. This can be achieved by deaccentuation, that is, by moving the accent to another constituent or even, as shown in (28b), by moving the object to a different position from the one in which it receives structural stress.

- (28) a. dass die Polizei gestern Linguisten verhaftete
that the police yesterday linguists arrested
'that the police arrested linguists yesterday'
- b. dass die Polizei Linguisten gestern verhaftete
that the police linguists yesterday arrested
'that the police arrested linguists yesterday'

In Spanish, partial focus can be achieved not by special intonation, but rather only by altruistic movement in order to move the object out of the focus. See also **BC2010a** for a discussion of "altruistic" multiple frontings in German.

It is therefore not possible to assume that elements are moved to a particular position in the tree in order to check some feature motivated by information structural properties. Since feature checking is a prerequisite for movement in current minimalist theory, one would have to postulate a special feature, which only has the function of triggering altruistic movement. Fanselow (**Fanselow2003b**; **Fanselow2006a**) has also shown that the ordering constraints that one assumes for topic, focus and sentence adverbs can be adequately described by a theory which assumes firstly, that arguments are combined (in minimalist terminology: *merged*) with their head one after the other and secondly, that adjuncts can be

adjoined to any projection level. The position of sentence adverbs directly before the focused portion of the sentence receives a semantic explanation: since sentence adverbs behave like focus”=sensitive operators, they have to directly precede elements that they refer to. It follows from this that elements which do not belong to the focus of an utterance (topics) have to occur in front of the sentence adverb. It is therefore not necessary to assume a special topic position to explain local reorderings in the middle field. This analysis is also pursued in LFG and HPSG. The respective analyses are discussed in more detail in the corresponding chapters.

4.6.2 Labeling

In the Minimalist Program, Chomsky tries to keep combinatorial operations and mechanisms as simple as possible. He motivates this with the assumption that the existence of a UG with less language”=specific knowledge is more plausible from a evolutionary point of view than a UG which contains a high degree of language”=specific knowledge (Chomsky2008a).

For this reason, he removes the projection levels of \bar{X} theory, traces, indices and “similar descriptive technology” (Chomsky2008a). All that remains is Merge and Move, that is, Internal and External Merge. Internal and External Merge combine two syntactic objects α and β into a larger syntactic object which is represented as a set $\{\alpha, \beta\}$. α and β can be either lexical items or internally complex syntactic objects. Internal Merge moves a part of an object to its periphery.¹⁵ The result of internally merging an element is a set $\{\alpha, \beta\}$ where α was a part of β . External Merge also produces a set with two elements. However, two independent objects are merged. The objects that are created by Merge have a certain category (a set of features). For instance, if one combines the elements α and β , one gets $\{l, \{\alpha, \beta\}\}$, where l is the category of the resulting object. This category is also called a *label*. Since it is assumed that all constituents are headed, the category that is assigned to $\{\alpha, \beta\}$ has to be either the category of α or the category of β . Chomsky2008a discusses the following two rules for the determination of the label of a set.

- (29) a. In $\{H, \alpha\}$, H an LI, H is the label.
- b. If α is internally merged to β , forming $\{\alpha, \beta\}$ then the label of β is the label of $\{\alpha, \beta\}$.

¹⁵To be more specific, part of a syntactic object is copied and the copy is placed at the edge of the entire object. The original of this copy is no longer relevant for pronunciation (*Copy Theory of Movement*).

As Chomsky notes, these rules are not unproblematic since the label is not uniquely determined in all cases. An example is the combination of two lexical elements. If both H and α in (29a) are lexical items (LI), then both H and α can be the label of the resulting structure. Chomsky notices that this could result in deviant structures, but claims that this concern is unproblematic and ignores it. Chomsky offered a treatment of the combination of two lexical items in his **Chomsky2013a** paper. The solution to the problem is to assume that all combinations of lexical elements consist of a functional element and a root (**Marantz97a; Borer2005a-u**). Roots are not considered as labels per definition¹⁶ and hence the category of the functional element determines the category of the combination (**Chomsky2013a**). Such an analysis can only be rejected: the goal of the Minimalist Program is to simplify the theoretical proposals to such an extent that the models of language acquisition and language evolution become plausible, but in order to simplify basic concepts it is stipulated that a noun cannot simply be a noun but needs a functional element to tell the noun what category it has. Given that the whole point of Chomsky's Bare Phrase Structure (**Chomsky95b-u**) was the elimination of the unary branching structures in \bar{X} theory, it is unclear why they are reintroduced now through the backdoor, only more complex with an additional empty element.¹⁷ Theories like Categorical Grammar and HPSG can combine lexical items directly without assuming any auxiliary projections or empty elements. See also **Rauh2016a** for a comparison of the treatment of syntactic categories in earlier versions of Transformational Grammar, HPSG, Construction Grammar, Role and Reference Grammar and root-based Neo-Constructivist proposals like the one assumed by **Chomsky2013a**. Rauh concludes that the direct connection of syntactic and semantic information is needed and that the Neo-Constructivism of Marantz and Borer has to be rejected. For further criticism of Neo-Constructivist approaches see **Wechsler2008a** and **MWArgSt**.

The combination of a pronoun with a verbal projection poses a problem that is related to what has been said above. In the analysis of *He left*, the pronoun *he* is a lexical element and hence would be responsible for the label of *He left*, since *left* is an internally complex verbal projection in Minimalist theories. The

¹⁶ Another category that is excluded as label per definition is *Conj*, which stands for conjunction (**Chomsky2013a**). This is a stipulation that is needed to get coordination to work. See below.

¹⁷ The old \bar{X} rule in (i.a) corresponds to the binary combination in (i.b).

- (i) a. $N' \rightarrow N$
- b. $N \rightarrow N\text{-func root}$

In (i.a) a lexical noun is projected to an N' and in (i.b), a root is combined with a functional nominal head into a nominal category.

result would be a nominal label rather than a verbal one. To circumvent this problem, Chomsky2013a assumes that *he* has a complex internal structure: ‘perhaps D-pro’, that is, *he* is (perhaps) composed out of an invisible determiner and a pronoun.

The case in which two non-LIs are externally merged (for instance a nominal and a verbal phrase) is not discussed in Chomsky2008a. Chomsky2013a suggests that a phrase XP is irrelevant for the labeling of {XP, YP} if XP is moved (or rather copied in the Copy Theory of Movement) in a further step. Chomsky assumes that one of two phrases in an {XP, YP} combination has to move, since otherwise labeling would be impossible (p. 12).¹⁸ The following coordination example will illustrate this: Chomsky assumes that the expression *Z and W* is analyzed as follows: first, Z and W are merged. This expression is combined with Conj (30a) and in the next step Z is raised (30b).

- (30) a. [_{α} Conj [_{β} Z W]]
 b. [_{γ} Z [_{α} Conj [_{β} Z W]]

Since Z in β is only a copy, it does not count for labeling and β can get the label of W. It is stipulated for the combination of Z and α that Conj cannot be the label and hence the label of the complete structure is Z.¹⁹

A special case that is discussed by Chomsky is the Internal Merge of an LI α with a non LI β . According to rule (29a) the label would be α . According to

¹⁸His explanation is contradictory: on p. 11 Chomsky assumes that a label of a combination of two entities with the same category is this category. But in his treatment of coordination, he assumes that one of the conjuncts has to be raised, since otherwise the complete structure could not be labeled.

¹⁹As Bob Borsley (p.c. 2013) pointed out to me, this makes wrong predictions for coordinations of two singular noun phrases with *and*, since the result of the coordination is a plural DP and not a singular one like the first conjunct. Theories like HPSG can capture this by grouping features in bundles that can be shared in coordinated structures (syntactic features and nonlocal features, see ps2).

Furthermore the whole account cannot explain why (i.b) is ruled out.

- (i) a. both Kim and Lee
 b. * both Kim or Lee

The information about the conjunction has to be part of the representation for *or Lee* in order to be able to contrast it with *and Lee*.

A further problem is that the label of α should be the label of W since Conj does not count for label determination. This would lead to a situation in which we have to choose between Z and W to determine the label of γ . Following Chomsky’s logic, either Z or W would have to move on to make it possible to label γ . Chomsky2013a mentions this problem in footnote 40, but does not provide a solution.

(29b), the label would be β (see also Donati2006a-u). Chomsky discusses the combination of the pronoun *what* with *you wrote* as an example.

(31) what [C [you wrote *t*]]

If the label is determined according to (29b), one then has a syntactic object that would be called a CP in the GB framework; since this CP is, moreover, interrogative, it can function as the complement of *wonder* as in (32a). If the label is determined according to (29a), one gets an object that can function as the accusative object of *read* in (32b), that is, something that corresponds to a DP in GB terminology.

- (32) a. I wonder what you wrote.
b. I read what you wrote.

what you wrote in (32b) is a so-called free relative clause.

Chomsky's approach to free relative clauses is interesting but is unable to describe the phenomenon in full breadth. The problem is that the phrase that contains the relative pronoun may be complex (contrary to Donati's claims, see also Citko2008a for a rejection of Donati's claim).²⁰ (33) provides an English example from BG78a. German examples from Bausewein90 and Mueller99b are given in (34).

(33) I'll read [whichever book] you give me.

- (34) a. Ihr könnt beginnen, [mit *wem*] ihr wollt.²¹
you can start with whom you want
'You can start with whoever you like.'
b. [*Wessen Birne*] noch halbwegs in der Fassung steckt, pflegt solcherlei
whose bulb/head yet halfway in the socket is uses such
Erloschene zu meiden;²²
extinct to avoid
'Those who still have their wits half way about them tend to avoid such
vacant characters;'

²⁰Chomsky2013a admits that there are many open questions as far as the labeling in free relative clauses is concerned and hence admits that there remain many open questions with labeling as such.

²¹Bausewein90.

²²Thomas Gsell, taz, 12.02.1997, p. 20.

- c. [*Wessen Schuhe*] "danach" besprenkelt sind, hat keinen Baum gefunden
 whose shoes after.that speckled are has no tree found
 und war nicht zu einem Bogen in der Lage.²³
 and was not to a bow in the position
 'Those whose shoes are spattered afterwards couldn't find a tree and
 were incapable of peeing in an arc.'

Since *wessen Schuhe* 'whose shoes' is not a lexical item, rule (29b) has to be applied, provided no additional rules are assumed to deal with such cases. This means that the whole free relative clause *wessen Schuhe danach besprenkelt sind* is labeled as CP. For the free relatives in (33) and (34) the labeling as a CP is an unwanted result, since they function as subjects or objects of the matrix predicates and hence should be labelled DP. However, since *wessen Schuhe* is a complex phrase and not a lexical item, (29a) does not apply and hence there is no analysis of the free relative clause as a DP. Therefore, it seems one must return to something like the GB analysis proposed by GR81, at least for the German examples. Gross and van Riemsdijk assume that free relatives consist of an empty noun that is modified by the relative clause like a normal noun. In such an approach, the complexity of the relative phrase is irrelevant. It is only the empty head that is relevant for labeling the whole phrase.²⁴ However, once empty heads

²³taz, taz mag, 08./09.08.1998, p. XII.

²⁴Assuming an empty head is problematic since it may be used as an argument only in those cases in which it is modified by an adjunct, namely the relative clause (Mueller99b). See also Ott2011a for a later rediscovery of this problem. It can be solved in HPSG by assuming a unary projection that projects the appropriate category from a relative clause. I also use the unary projection to analyze so-called *non-matching* free relative clauses (Mueller99b). In constructions with nonmatching free relative clauses, the relative clause fills an argument slot that does not correspond to the properties of the relative phrase (Bausewein90). Bausewein discusses the following example, in which the relative phrase is a PP but the free relative fills the accusative slot of *kocht* 'cooks'.

- (i) Sie kocht, worauf sie Appetit hat.
 she cooks where.on she appetite has
 'She cooks what she feels like eating.'

See Mueller99b for corpus examples.

Minimalist theories do not employ unary projections. Ott2011a develops an analysis in which the category of the relative phrase is projected, but he does not have a solution for nonmatching free relative clauses (p.187). The same is true for Citko's analysis, in which an internally merged XP can provide the label.

Many other proposals for labeling or, rather, non-labeling exist. For instance, some Minimalists want to eliminate labeling altogether and argue for a label-free syntax. As was pointed out by OPG2011a, such analyses bring Minimalism closer to Dependency Grammar. It is un-

are countenanced in the analysis, the application of (29a) to (31) is undesirable since the application would result in two analyses for (32b): one with the empty nominal head and one in which (31) is labeled as DP directly. One might argue that in the case of several possible derivations, the most economical one wins, but the assumption of transderivational constraints leads to undesired consequences (Pullum2013a).

Chomsky2013a abandons the labeling condition in (29b) and replaces it with general labeling rules that hold for both internal and external Merge of two phrases. He distinguishes two cases. In the first case, labeling becomes possible since one of the two phrases of the set { XP, YP } is moved away. This case was already discussed above. Chomsky writes about the other case: *X and Y are identical in a relevant respect, providing the same label, which can be taken as the label of the SO* (p. 11). He sketches an analysis of interrogative clauses on p. 13 in which the interrogative phrase has a Q feature and the remaining sentence from which the Q phrase was extracted has a Q feature as well. Since the two constituents share this property, the label of the complete clause will be Q. This kind of labeling will “perhaps” also be used for labeling normal sentences consisting of a subject and a verb phrase agreeing in person and number. These features would be responsible for the label of the sentence. The exact details are not worked out, but almost certainly will be more complex than (29b).

A property that is inherent in both Chomsky2005a and Chomsky2013a is that the label is exclusively determined from one of the merged objects. As Bob Borsley pointed out to me, this is problematic for interrogative/relative phrases like (35).

(35) with whom

The phrase in (35) is both a prepositional phrase (because the first word is a preposition) and an interrogative/relative phrase (because the second word is an interrogative/relative word). So, what is needed for the correct labeling of PPs

clear how any of these models could deal with non-matching free relative clauses. GO2009a provide an analysis of free relatives in their version of Dependency Grammar, but deny the existence of nonmatching ones (p. 78). They suggest an analysis in which the relative phrase is the root/label of the free relative clause and hence they have the same problem as Minimalist proposals have with non-matching free relative clauses. As GO2009a and OPG2011a state: empty heads are usually not assumed in (their version of) Dependency Grammar. Neither are unary branching projections. This seems to make it impossible to state that free relative clauses with a relative phrase YP can function as XP, provided XP is a category that is higher in the obliqueness hierarchy of KC77a, a generalization that was discovered by Bausewein90 (see also Mueller99b and Vogel2001a). In order to be able to express the relevant facts, an element or a label has to exist that is different from the label of *worauf* in (i).

like the one in (35) is a well-defined way of percolating different properties from daughters to the mother node.²⁵

For further problems concerning labeling and massive overgeneration by recent formulations of Merge see **FSP2016a**.

Summarizing, one can say that labeling, which was introduced to simplify the theory and reduce the amount of language specific innate knowledge that has to be assumed, can only be made to function with a considerable amount of stipulations. For instance, the combination of lexical elements requires the assumption of empty functional heads, whose only purpose is determining the syntactic category of a certain lexical element. If this corresponded to linguistic reality, knowledge about labeling, the respective functional categories, and information about those categories that have to be ignored for the labeling would have to be part of innate language specific knowledge and nothing would be gained. One would be left with bizarre analyses with an enormous degree of complexity without having made progress in the Minimalist direction. Furthermore, there are empirical problems and a large number of unsolved cases.

The conclusion is that the label of a binary combination should not be determined in the ways suggested by **Chomsky2008a**; **Chomsky2013a**. An alternative option for computing the label is to use the functor of a functor argument structure as the label (**BE95a**). This is the approach taken by **Categorial Grammar (Ajdukiewicz35a-u; Steedman2000a-u)** and in **Stabler's Minimalist Grammars (Stabler2010b)**.²⁶ **Stabler's** formalization of Merge will be discussed in Sec-

²⁵HPSG solves this problem by distinguishing head features including part of speech information and nonlocal features containing information about extraction and interrogative/relative elements. Head features are projected from the head, the nonlocal features of a mother node are the union of the nonlocal features of the daughters minus those that are bound off by certain heads or in certain configurations.

Citko2008a suggests an analysis in which both daughters can contribute to the mother node. The result is a complex label like { P, { D, N } }. This is a highly complex data structure and **Citko** does not provide any information on how the relevant information that it contains is accessed. Is an object with the label { P, { D, N } } a P, a D or an N? One could say that P has priority since it is in the least embedded set, but D and N are in one set. What about conflicting features? How does a preposition that selects for a DP decide whether { D, N } is a D or an N? In any case it is clear that a formalization will involve recursive relations that dig out elements of subsets in order to access their features. This adds to the overall complexity of the proposal and is clearly dispreferred over the HPSG solution, which uses one part of speech value per linguistic object.

²⁶For the **Categorial Grammar** approach to work, it is necessary to assign the category x/x to an adjunct, where x stands for the category of the head to which the adjunct attaches. For instance, an adjective combines with a nominal object to form a nominal object. Therefore its category is n/n rather than *adj*.

tion 4.6.4.

4.6.3 Specifiers, complements, and the remains of \bar{X} theory

Chomsky2008a assumes that every head has exactly one complement but an arbitrary number of specifiers. In standard \bar{X} theory, the restriction that there can be at most one complement followed from the general \bar{X} schema and the assumption that structures are at most binary branching: in standard \bar{X} theory a lexical head was combined with all its complements to form an X' . If there are at most two daughters in a phrase, it follows that there can be only one complement (Sentences with ditransitive verbs have been analyzed with an empty head licensing an additional argument; see Larson88a for the suggestion of an empty verbal head and MWArgSt for a critical assessment of approaches involving little v). In standard \bar{X} theory there was just one specifier. This restriction has now been abandoned. Chomsky writes that the distinction between specifier and complement can now be derived from the order in which elements are merged with their head: elements that are *first-merged* are complements and all others – those which are *later-merged* – are specifiers.

Such an approach is problematic for sentences with monovalent verbs: according to Chomsky's proposal, subjects of monovalent verbs would not be specifiers but complements.²⁷ This problem will be discussed in more detail in Section 4.6.4.

Similarly, Stabler's approach does not extend to adjuncts unless he is willing to assign the category noun to attributive adjectives. One way out of this problem is to assume a special combination operation for adjuncts and their heads (see FG2002a). Such a combination operation is equivalent to the Head"=Adjunct Schema of HPSG.

²⁷Pauline Jacobson (p.c. 2013) pointed out that the problem with intransitive verbs could be solved by assuming that the last-merged element is the specifier and all non-last-merged elements are complements. This would solve the problems with intransitive verbs and with the coordination of verbs in (36) but it would not solve the problem of coordination in head-final languages as in (39). Furthermore, current Minimalist approaches make use of multiple specifiers and this would be incompatible with the Jacobsonian proposal unless one would be willing to state more complicated restrictions on the status of non-first-merged elements.

Apart from this, theories assuming that syntactic objects merged with word groups are specifiers do not allow for analyses in which two lexical verbs are directly coordinated, as in (36):²⁸

(36) He [knows and loves] this record.

For example, in an analysis suggested by Steedman91a, *and* (being the head) is first merged with *loves* and then the result is merged with *knows*. The result of this combination is a complex object that has the same syntactic properties as the combined parts: the result is a complex verb that needs a subject and an object. After the combination of the conjunction with the two verbs, the result has to be combined with *this record* and *he*. *this record* behaves in all relevant respects like a complement. Following Chomsky's definition, however, it should be a specifier, since it is combined with the third application of Merge. The consequences are unclear. Chomsky assumes that Merge does not specify constituent order. According to him, the linearization happens at the level of Phonological Form (PF). The restrictions that hold there are not described in his recent papers. However, if the categorization as complement or specifier plays a role for linearization as in Kayne's work (Kayne2011a) and in Stabler's proposal (see Section 4.6.4), *this record* would have to be serialized before *knows and loves*, contrary to the facts. This means that a Categorical Grammar-like analysis of coordination is not viable and the only remaining option would seem to assume that *knows* is combined with an object and then two VPs are coordinated. Kayne94a-u follows WC80a-u in suggesting such an analysis and assumes that the object in the first VP is

²⁸Chomsky2013a suggests the coordination analysis in (30): according to this analysis, the verbs would be merged directly and one of the verbs would be moved around the conjunction in a later step of the derivation. As was mentioned in the previous section, such analyses do not contribute to the goal of making minimal assumptions about innate language specific knowledge since it is absolutely unclear how such an analysis of coordination would be acquired by language learners. Hence, I will not consider this coordination analysis here.

Another innovation of Chomsky's 2013 paper is that he eliminates the concept of specifier. He writes in footnote 27 on page 43: *There is a large and instructive literature on problems with Specifiers, but if the reasoning here is correct, they do not exist and the problems are unformulable.* This is correct, but this also means that everything that was explained with reference to the notion of specifier in the Minimalist framework until now does not have an explanation any longer. If one follows Chomsky's suggestion, a large part of the linguistic research of the past years becomes worthless and has to be redone.

Chomsky did not commit himself to a particular view on linearization in his earlier work, but somehow one has to ensure that the entities that were called specifier are realized in a position in which constituents are realized that used to be called specifier. This means that the following remarks will be relevant even under current Chomskyan assumptions.

deleted. However, Borsley2005a shows that such an analysis makes wrong predictions, since (37a) would be derived from (37b) although these sentences differ in meaning.²⁹

- (37) a. Hobbs whistled and hummed the same tune.
b. Hobbs whistled the same tune and hummed the same tune.

Since semantic interpretation cannot see processes such as deletion that happen at the level of Phonological Form (Chomsky95a-u), the differences in meaning cannot be explained by an analysis that deletes material.

In a further variant of the VP coordination analysis, there is a trace that is related to *this record*. This would be a *Right-Node-Raising* analysis. Borsley2005a has shown that such analyses are problematic. Among the problematic examples that he discusses is the following pair (see also Bresnan74a-u).

- (38) a. He tried to persuade and convince him.
b. * He tried to persuade, but couldn't convince, him.

The second example is ungrammatical if *him* is not stressed. In contrast, (38a) is well-formed even with unstressed *him*. So, if (38a) were an instance of Right-Node-Raising, the contrast would be unexpected. Borsley therefore excludes a Right-Node-Raising analysis.

The third possibility to analyze sentences like (36) assumes discontinuous constituents and uses material twice: the two VPs *knows this record* and *loves this record* are coordinated with the first VP being discontinuous. (See Crysmann2000a and BS2004a for such proposals in the framework of HPSG.) However, discontinuous constituents are not usually assumed in the Minimalist framework (see for instance Kayne94a-u). Furthermore, Abeille2006a showed that there is evidence for structures in which lexical elements are coordinated directly. This means that one needs analyses like the CG analysis discussed above, which would result in the problems with the specifier/complement status just discussed.

Furthermore, Abeillé has pointed out that NP/DP coordinations in head-final languages like Korean and Japanese present difficulties for Merge-based analyses. (39) shows a Japanese example.

- (39) Robin-to Kim
Robin-and Kim
'Kim and Robin'

²⁹See also BV72, Jackendoff77a, Dowty79a, denBesten83a, Klein85 and Eisenberg94a for similar observations and criticism of similar proposals in earlier versions of Transformational Grammar.

In the first step *Robin* is merged with *to*. In a second step *Kim* is merged. Since *Kim* is a specifier, one would expect that *Kim* is serialized before the head as it is the case for other specifiers in head-final languages.

Chomsky tries to get rid of the unary branching structures of standard \bar{X} theory, which were needed to project lexical items like pronouns and determiners into full phrases, referring to work by Muysken82a. Muysken used the binary features MIN and MAX to classify syntactic objects as minimal (words or word-like complex objects) or maximal (syntactic objects that stand for complete phrases). Such a feature system can be used to describe pronouns and determiners as [+MIN, +MAX]. Verbs like *give*, however, are classified as [+MIN, -MAX]. They have to project in order to reach the [+MAX]-level. If specifiers and complements are required to be [+MAX], then determiners and pronouns fulfill this requirement without having to project from X^0 via X' to the XP'' -level.

In Chomsky's system, the MIN/MAX distinction is captured with respect to the completeness of heads (complete = phrase) and to the property of being a lexical item. However, there is a small but important difference between Muysken's and Chomsky's proposal: the predictions with regard to the coordination data that was discussed above. Within the category system of \bar{X} theory, it is possible to combine two X^0 s to get a new, complex X^0 . This new object has basically the same syntactic properties that simple X^0 s have (see Jackendoff77a and GKPS85a). In Muysken's system, the coordination rule (or the lexical item for the conjunction) can be formulated such that the coordination of two +MIN items is a +MIN item. In Chomsky's system an analogous rule cannot be defined, since the coordination of two lexical items is not a lexical item any longer.

Like Chomsky in his recent Minimalist work, Categorical Grammar (Ajdukiewicz35a-u) and HPSG (Pollard and Sag ps; ps2) do not (strictly) adhere to \bar{X} theory. Both theories assign the symbol NP to pronouns (for CG see SB2006a-u, see Steedman2000a-u for the incorporation of lexical type raising in order to accommodate quantification). The phrase *likes Mary* and the word *sleeps* have the same category in Categorical Grammar ($\bar{s}\backslash np$). In both theories it is not necessary to project a noun like *tree* from N^0 to \bar{N} in order to be able to combine it with a determiner or an adjunct. Determiners and monovalent verbs in controlled infinitives are not projected from an X^0 level to the XP level in many HPSG analyses, since the valence properties of the respective linguistic objects (an empty SUBCAT or COMPS list) are sufficient to determine their combinatoric potential and hence their distribution (Mueller96a; Mueller99a). If the property of being minimal is needed for the description of a phenomenon, the binary feature LEX is used in HPSG (Pollard and Sag ps; ps2). However, this feature is not needed for the distinction

between specifiers and complements. This distinction is governed by principles that map elements of an argument structure list (ARG-ST) onto valence lists that are the value of the SPECIFIER and the COMPLEMENTS feature (abbreviated as SPR and COMPS respectively).³⁰ Roughly speaking, the specifier in a verbal projection is the least oblique argument of the verb for configurational languages like English. Since the argument structure list is ordered according to the obliqueness hierarchy of KC77a, the first element of this list is the least oblique argument of a verb and this argument is mapped to the SPR list. The element in the SPR list is realized to the left of the verb in SVO languages like English. The elements in the COMPS list are realized to the right of their head. Approaches like the one by GSag2000a-u that assume that head-complement phrases combine a word with its arguments have the same problem with coordinations like (36) since the head of the VP is not a word.³¹ However, this restriction for the head can be replaced by one that refers to the LEX feature rather than to the property of being a word or lexical item.

Pollard & Sag as well as Sag & Ginzburg assume flat structures for English. Since one of the daughters is marked as lexical, it follows that the rule does not combine a head with a subset of its complements and then apply a second time to combine the result with further complements. Therefore, a structure like (40a) is excluded, since *gave John* is not a word and hence cannot be used as the head daughter in the rule.

- (40) a. [[gave John] a book]
 b. [gave John a book]

Instead of (40a), only analyses like (40b) are admitted; that is, the head is combined with all its arguments all in one go. The alternative is to assume binary branching structures (MuellerHPSGHandbook; MOeDanish). In such an approach, the head complement schema does not restrict the word/phrase status of the head daughter. The binary branching structures in HPSG correspond to External Merge in the MP.

³⁰Some authors assume a three-way distinction between subjects, specifiers, and complements.

³¹As mentioned above, a multidomination approach with discontinuous constituents is a possible solution for the analysis of (36) (see Crysmann2000a and BS2004a). However, the coordination of lexical items has to be possible in principle as Abeille2006a has argued. Note also that the HPSG approach to coordination cannot be taken over to the MP. The reason is that the HPSG proposals involve special grammar rules for coordination and MP comes with the claim that there is only Merge. Hence the additional introduction of combinatorial rules is not an option within the MP.

In the previous two sections, certain shortcomings of Chomsky's labeling definition and problems with the coordination of lexical items were discussed. In the following section, I discuss Stabler's definition of Merge in Minimalist Grammar, which is explicit about labeling and in one version does not have the problems discussed above. I will show that his formalization corresponds rather directly to HPSG representations.

4.6.4 Minimalism, Categorical Grammar, and HPSG

In this section, I will relate Minimalism, Categorical Grammar and HPSG to one another. This section is based on **MuellerUnifying**. Readers who are not yet familiar with Categorical Grammar and HPSG should skim this section or consult the Chapters 6, ?? and ?? and return here afterwards.

In Section 4.6.2, it was shown that Chomsky's papers leave many crucial details about labeling unspecified. Stabler's work is relatively close to recent Minimalist approaches, but is worked out much more precisely (see also **Stabler2010a** on formalization of post GB approaches). **Stabler2001a** shows how Kayne's theory of remnant movement can be formalized and implemented. Stabler refers to his particular way of formalizing Minimalist theories as *Minimalist Grammars* (MG). There are a number of interesting results with regard to the weak capacity of Minimalist Grammars and variants thereof (**Michaelis2001a-u**). It has been shown, for instance, that the number of possible languages one could create with MGs includes the set of those which can be created by Tree Adjoining Grammars (see Chapter ??). This means that it is possible to assign a greater number of word strings to structures with MGs, however, the structures derived by MGs are not necessarily always the same as the structures created by TAGs. For more on the generative capacity of grammars, see Chapter ??.

Although Stabler's work can be regarded as a formalization of Chomsky's Minimalist ideas, Stabler's approach differs from Chomsky's in certain matters of detail. Stabler assumes that the results of the two Merge operations are not sets but pairs. The head in a pair is marked by a pointer ('<' or '>'). Bracketed expressions like $\{\alpha, \{\alpha, \beta\}\}$ (discussed in Section 4.6.2) are replaced by trees like the one in (41).

(41)

1 is the head in (41), 2 is the complement and 3 the specifier. The pointer points to the part of the structure that contains the head. The daughters in a tree are ordered, that is, 3 is serialized before 1 and 1 before 2.

Stabler2010a defines External Merge as follows:

$$(42) \quad \text{em}(t_1[=f], t_2[f]) = \begin{cases} \text{if } t_1 \text{ has exactly 1 node} \\ \text{otherwise} \end{cases}$$

$=f$ is a selection feature and f the corresponding category. When $t_1[=f]$ and $t_2[f]$ are combined, the result is a tree in which the selection feature of t_1 and the respective category feature of t_2 are deleted. The upper tree in (42) represents the combination of a (lexical) head with its complement. t_1 is positioned before t_2 . The condition that t_1 has to have exactly one node corresponds to Chomsky's assumption that the first Merge is a Merge with a complement and that all further applications of Merge are Merges with specifiers (Chomsky2008a).

Stabler defines Internal Merge as follows:³²

$$(43) \quad \text{im}(t_1[+f]) =$$

t_1 is a tree with a subtree t_2 which has the feature f with the value ' $-$ '. This subtree is deleted ($t_2[-f]^> \mapsto \epsilon$) and a copy of the deleted subtree without the $-f$ feature ($t_2^>$) is positioned in specifier position. The element in specifier position has to be a maximal projection. This requirement is visualized by the raised '>'.³³

Stabler provides an example derivation for the sentence in (44).

$$(44) \quad \text{who Marie praises}$$

praises is a two-place verb with two $=D$ features. This encodes the selection of two determiner phrases. *who* and *Marie* are two Ds and they fill the object and subject position of the verb. The resulting verbal projection *Marie praises who* is embedded under an empty complementizer which is specified as $+wH$ and hence provides the position for the movement of *who*, which is placed in the specifier position of CP by the application of Internal Merge. The $-wH$ feature of *who* is deleted and the result of the application of Internal Merge is *who Marie praises*.

This analysis has a problem that was pointed out by Stabler himself in unpublished work cited by Veenstra98a: it makes incorrect predictions in the case of monovalent verbs. If a verb is combined with an DP, the definition of External Merge in (42) treats this DP as a complement³³ and serializes it to the right of the head. Instead of analyses of sentences like (45a) one gets analyses of strings like (45b).

$$(45) \quad \text{a. Max sleeps.}$$

³²In addition to what is shown in (43), Stabler's definition contains a variant of the *Shortest Move Constraint* (SMC), which is irrelevant for the discussion at hand and hence will be omitted.

³³Compare also Chomsky's definition of specifier and complement in Section 4.6.3.

- b. * Sleeps Max.

To solve this problem, Stabler assumes that monovalent verbs are combined with a nonovert object (see **Veenstra98a** who, quoting Stabler's unpublished work, also adopts this solution). With such an empty object, the resulting structure contains the empty object as a complement. The empty object is serialized to the right of the verb and *Max* is the specifier and hence serialized to the left of the verb as in (46).

(46) Max sleeps \bar{e} .

Of course, any analysis of this kind is both stipulative and entirely ad hoc, being motivated only by the wish to have uniform structures. Moreover, it exemplifies precisely one of the methodological deficiencies of Transformational Generative Grammar discussed at length by **CJ2005a**: the excessive appeal to uniformity.

An alternative is to assume an empty verbal head that takes *sleeps* as complement and *Max* as subject. Such an analysis is often assumed for ditransitive verbs in Minimalist theories which assume Larsonian verb shells (**Larson88a**). Larsonian analyses usually assume that there is an empty verbal head that is called little *v* and that contributes a causative meaning. As was discussed in Section 4.1.4, **Adger2003a** adopts a little *v*"=based analysis for intransitive verbs. Omitting the TP projection, his analysis is provided in Figure 4.22. Adger argues

Figure 4.22: Little *v*"=based analysis of *Max sleeps*

that the analysis of sentences with unergative verbs involves a little *v* that selects an agent, while the analysis of unaccusative verbs involves a little *v* that does not select an N head. For unaccusatives, he assumes that the verb selects a theme. He states that little *v* does not necessarily have a causative meaning but introduces the agent. But note that in the example at hand the subject of *sleep* is neither causing an event, nor is it necessarily deliberately doing something. So it is rather an undergoer than an agent. This means that the assumption of the empty *v* head is made for purely theory"=internal reasons without any semantic motivation in the case of intransitives. If the causative contribution of little *v* in ditransitive constructions is assumed, this would mean that one needs two little *vs*, one with and one without a causative meaning. In addition to the lack of theory"=external motivation for little *v*, there are also empirical problems for such analyses (for instance with coordination data). The reader is referred to **MWArgSt** for further details.

Apart from the two operations that were defined in (42) and (43), there are no other operations in MG.³⁴ Apart from the problems with monovalent verbs, this results in the problem that was discussed in Section 4.6.3: there is no analysis with a direct combination of verbs for (36) – repeated here as (47).

(47) He [knows and loves] this record.

The reason is that the combination of *knows*, *and* and *loves* consists of three nodes and the Merge of *knows and loves* with *this record* would make *this record* the specifier of the structure. Therefore *this record* would be serialized before *knows and loves*, contrary to the facts. Since the set of languages that can be generated with MGs contains the languages that can be generated with certain TAGs and with Combinatorial Categorical Grammar (Michaelis2001a-u), the existence of a Categorical Grammar analysis implies that the coordination examples can be derived in MGs somehow. But for linguists, the fact that it is possible to generate a certain string at all (the weak capacity of a grammar) is of less significance. It is the actual structures that are licensed by the grammar that are important (the strong capacity).

4.6.4.1 Directional Minimalist Grammars and Categorical Grammar

Apart from reintroducing X^0 categories, the coordination problem can be solved by changing the definition of Merge in a way that allows heads to specify the direction of combination with their arguments: Stabler2010b suggests marking the position of an argument relative to its head together with the selection feature and gives the following redefinition of External Merge.

$$(48) \quad \text{em}(t_1[\alpha], t_2[x]) = \begin{cases} \text{if } \alpha \text{ is } =x \\ \text{if } \alpha \text{ is } x= \end{cases}$$

The position of the equal sign specifies on which side of the head an argument has to be realized. This corresponds to forward and backward Application in Categorical Grammar (see Section ??). Stabler calls this form of grammar Directional MG (DMG). This variant of MG avoids the problem with monovalent verbs and the coordination data is unproblematic as well if one assumes that the conjunction is a head with a variable category that selects for elements of the same category to the left and to the right of itself. *know* and *love* would both select an

³⁴For extensions see FG2002a.

object to the right and a subject to the left and this requirement would be transferred to *knows and loves*.³⁵ See **Steedman91a** for the details of the CG analysis and **BvN98a** for an earlier HPSG proposal involving directionality features along the lines suggested by Stabler for his DMGs.

4.6.4.2 Minimalist Grammars and Head-Driven Phrase Structure Grammar

The notation for marking the head of a structure with ‘>’ and ‘<’ corresponds directly to the HPSG representation of heads. Since HPSG is a sign-based theory, information about all relevant linguistic levels is represented in descriptions (phonology, morphology, syntax, semantics, information structure). (49) gives an example: the lexical item for the word *grammar*.

$$(49) \left[\begin{array}{l} \text{word} \\ \text{PHON} \end{array} \quad \langle 'gramər \rangle \right. \\ \left. \begin{array}{l} \text{SYNSEM|LOC} \end{array} \quad \left[\begin{array}{l} \text{loc} \\ \text{CAT} \quad \left[\begin{array}{l} \text{cat} \\ \text{HEAD} \quad \text{noun} \\ \text{SPR} \quad \langle \text{DET} \rangle \end{array} \right] \\ \text{CONT} \quad \dots \left[\begin{array}{l} \text{grammar} \\ \text{INST} \quad X \end{array} \right] \end{array} \right] \right]$$

The part of speech of *grammar* is *noun*. In order to form a complete phrase, it requires a determiner. This is represented by giving the *SPR* feature the value $\langle \text{DET} \rangle$. Semantic information is listed under *CONT*. For details see Chapter ??.

Since we are dealing with syntactic aspects exclusively, only a subset of the used features is relevant: valence information and information about part of speech and certain morphosyntactic properties that are relevant for the external distribution of a phrase is represented in a feature description under the path *SYNSEM|""LOC|""CAT*. The features that are particularly interesting here are the so-called head features. Head features are shared between a lexical head and its maximal projection. The head features are located inside *CAT* and are grouped together under the path *HEAD*. Complex hierarchical structure is also modelled

³⁵ Note however, that this transfer makes it necessary to select complex categories, a fact that I overlooked in **MuellerUnifying**. The selection of simplex features vs. complex categories will be discussed in Section 4.6.5.

with feature value pairs. The constituents of a complex linguistic object are usually represented as parts of the representation of the complete object. For instance, there is a feature HEAD-DAUGHTER the value of which is a feature structure that models a linguistic object that contains the head of a phrase. The Head Feature Principle (50) refers to this daughter and ensures that the head features of the head daughter are identical with the head features of the mother node, that is, they are identical to the head features of the complete object.

$$(50) \quad \text{headed} = \text{phrase} \Rightarrow \left[\begin{array}{l} \text{SYNSEM|LOC|CAT|HEAD } \boxed{1} \\ \text{HEAD-DTR|SYNSEM|LOC|CAT|HEAD } \boxed{1} \end{array} \right]$$

Identity is represented by boxes with the same number.

GSag2000a-u represent all daughters of a linguistic object in a list that is given as the value of the DAUGHTERS attribute. The value of the feature HEAD-DAUGHTER is identified with one of the elements of the DAUGHTERS list:

$$(51) \quad \begin{array}{ll} \text{a.} & \left[\begin{array}{l} \text{HEAD-DTR } \boxed{1} \\ \text{DTRS} \quad \langle \boxed{1} \alpha, \beta \rangle \end{array} \right] \\ \text{b.} & \left[\begin{array}{l} \text{HEAD-DTR } \boxed{1} \\ \text{DTRS} \quad \langle \alpha, \boxed{1} \beta \rangle \end{array} \right] \end{array}$$

α and β are shorthands for descriptions of linguistic objects. The important point about the two descriptions in (51) is that the head daughter is identical to one of the two daughters, which is indicated by the $\boxed{1}$ in front of α and β , respectively. In the first feature description, the first daughter is the head and in the second description, the second daughter is the head. Because of the Head Feature Principle, the syntactic properties of the whole phrase are determined by the head daughter. That is, the syntactic properties of the head daughter correspond to the label in Chomsky's definition. This notation corresponds exactly to the one that is used by Stabler: (51a) is equivalent to (52a) and (51b) is equivalent to (52b).

$$(52) \quad \begin{array}{ll} \text{a.} & \text{b.} \end{array}$$

An alternative structuring of this basic information, discussed by ps2, uses the two features HEAD-DAUGHTER and NON-HEAD-DAUGHTERS rather than HEAD-DAUGHTER and DAUGHTERS. This gives rise to feature descriptions like (53a), which corresponds directly to Chomsky's set-based representations, discussed in Section 4.6.2 and repeated here as (53b).

- (53) a. $\left[\begin{array}{ll} \text{HEAD-DTR} & \alpha \\ \text{NON-HEAD-DTRS} & \langle \beta \rangle \end{array} \right]$
 b. $\{ \alpha, \{ \alpha, \beta \} \}$

The representation in (53a) does not contain information about linear precedence of α and β . Linear precedence of constituents is constrained by linear precedence rules, which are represented independently from constraints regarding (immediate) dominance.

The definition of Internal Merge in (43) corresponds to the Head=Filler Schema in HPSG (**ps2**). Stabler's derivational rule deletes the subtree $t_2[-f]^>$. HPSG is monotonic, that is, nothing is deleted in structures that are licensed by a grammar. Instead of deleting t_2 inside of a larger structure, structures containing an empty element (NB – not a tree) are licensed directly.³⁶ Both in Stabler's definition and in the HPSG schema, t_2 is realized as filler in the structure. In Stabler's definition of Internal Merge, the category of the head daughter is not mentioned, but **ps2** restrict the head daughter to be a finite verbal projection. **Chomsky2007a** assumes that all operations but External Merge operate on phase level. Chomsky assumes that CP and v^*P are phases. If this constraint is incorporated into the definition in (43), the restrictions on the label of t_1 would have to be extended accordingly. In HPSG, sentences like (54) have been treated as VPs, not as CPs and hence Pollard & Sag's requirement that the head daughter in the Head Filler Schema be verbal corresponds to Chomsky's restriction.

(54) Bagels, I like.

Hence, despite minor presentational differences, we may conclude that the formalization of Internal Merge and that of the Head-Filler Schema are very similar.

An important difference between HPSG and Stabler's definition is that 'movement' is not feature driven in HPSG. This is an important advantage since feature-driven movement cannot deal with instances of so-called altruistic movement (**Fanselow2003b**), that is, movement of a constituent that happens in order to make room for another constituent in a certain position (see Section 4.6.1.4).

A further difference between general \bar{X} theory and Stabler's formalization of Internal Merge on the one hand and HPSG on the other is that in the latter case there is no restriction regarding the completeness (or valence 'saturation') of the filler daughter. Whether the filler daughter has to be a maximal projection (English) or not (German), follows from restrictions that are enforced locally when

³⁶See **BMS2001a** for a traceless analysis of extraction in HPSG and **MuellerGS** and Chapter ?? of this book for a general discussion of empty elements.

the trace is combined with its head. This makes it possible to analyze sentences like (55) without remnant movement.³⁷

- (55) Gelesen_i hat_j das Buch keiner _{-i -j}.
 read has the book nobody

In contrast, Stabler is forced to assume an analysis like the one in (56b) (see also G. GMueller98a for a remnant movement analysis). In a first step, *das Buch* is moved out of the VP (56a) and in a second step, the emptied VP is fronted as in (56b).

- (56) a. Hat [das Buch]_j [keiner [_{VP -j} gelesen]].
 b. [_{VP -j} Gelesen]_i hat [das Buch]_j [keiner _{-i}].

Haider93a, dKM2001a and Fanselow2002a showed that this kind of remnant movement analysis is problematic for German. The only phenomenon that Fanselow identified as requiring a remnant movement analysis is the problem of multiple fronting (see Mueller2003b for an extensive discussion of relevant data). Mueller2005c; Mueller2005d; MuellerGS develops an alternative analysis of these multiple frontings which uses an empty verbal head in the *Vorfeld*, but does not assume that adjuncts or arguments like *das Buch* in (56b) are extracted from the *Vorfeld* constituent. Instead of the remnant movement analysis, the mechanism of argument composition from Categorical Grammar (Geach70a; HN94a) is used to ensure the proper realization of arguments in the sentence. Chomsky2007a already uses argument composition as part of his analysis of TPs and CPs. Hence both remnant movement and argument composition are assumed in recent Minimalist proposals. The HPSG alternative, however, would appear to need less theoretical apparatus and hence has to be preferred for reasons of parsimony.

Finally, it should be mentioned that all transformational accounts have problems with Across the Board extraction like (57a) and (57) in which one element corresponds to several gaps.

- (57) a. Bagels, I like and Ellison hates.³⁸
 b. The man who_i [Mary loves _{-i}] and [Sally hates _{-i}] computed my tax.

This problem was solved for GPSG by Gazdar81 and the solution carries over to HPSG. The Minimalist community tried to address these problems by introducing operations like sideward movement (Nunes2004a-u) where constituents can

³⁷See also MOe2013b for an analysis of object shift in Danish that can account for verb fronting without remnant movement. The analysis does not have any of the problems that remnant movement analyses have.

³⁸ps2.

be inserted into sister trees. So in the example in (57a), *Bagels* is copied from the object position of *hates* into the object position of *like* and then these two copies are related to the fronted element. Kobele criticized such solutions since they overgenerate massively and need complicated filters. What he suggests instead is the introduction of a GPSG-style SLASH mechanism into Minimalist theories (Kobele2008a).

Furthermore, movement paradoxes (Bresnan2001a) can be avoided by not sharing all information between filler and gap, a solution that is not available for transformational accounts, which usually assume identity of filler and gap or – as under the Copy Theory of Movement – assume that a derivation contains multiple copies of one object only one of which is spelled out. See also Borsley2012a for further puzzles for, and problems of, movement-based approaches.

A further difference between MG and HPSG is that the Head=Filler Schema is not the only schema for analyzing long=distance dependencies. As was noted in footnote 12 on page 156, there is dislocation to the right (extraposition) as well as fronting. Although these should certainly be analyzed as long=distance dependencies, they differ from other long=distance dependencies in various respects (see Section ??). For analyses of extraposition in the HPSG framework, see Keller95b, Bouma96, and Mueller99a.

Apart from the schema for long=distance dependencies, there are of course other schemata in HPSG which are not present in MG or Minimalism. These are schemata which describe constructions without heads or are necessary to capture the distributional properties of parts of constructions, which cannot be easily captured in lexical analyses (e.g., the distribution of *wh*- and relative pronouns). See Section ??.

Chomsky2010a has compared a Merge=based analysis of auxiliary inversion to a HPSG analysis and critiqued that the HPSG analysis uses ten schemata rather than one (Merge). GSag2000a-u distinguish three types of construction with moved auxiliaries: inverted sentences such as those with fronted adverbial and with *wh*=questions (58a,b), inverted exclamatives (58c) and polar interrogatives (58d):

- (58) a. Under no circumstances *did she think they would do that*.
 b. Whose book *are you reading*?
 c. Am I tired!
 d. Did Kim leave?

Fillmore99a captures various different usage contexts in his Construction Grammar analysis of auxiliary inversion and shows that there are semantic and prag-

matic differences between the various contexts. Every theory must be able to account for these. Furthermore, one does not necessarily require ten schemata. It is possible to determine this – as Categorical Grammar does – in the lexical entry for the auxiliary or on an empty head (see Chapter ?? for a more general discussion of lexical and phrasal analyses). Regardless of this, every theory has to somehow account for these ten differences. If one wishes to argue that this has nothing to do with syntax, then somehow this has to be modelled in the semantic component. This means that there is no reason to prefer one theory over another at this point.

4.6.5 Selection of atomic features vs. selection of complex categories

BE95a pointed out that Minimalist theories are very similar to Categorical Grammar and I have discussed the similarities between Minimalist theories and HPSG in **MuellerUnifying** and in the previous subsections. However, I overlooked one crucial difference between the usual assumptions about selection in Minimalist proposals on the one hand and Categorical Grammar, Dependency Grammar, LFG, HPSG, TAG, and Construction Grammar on the other hand: what is selected in the former type of theory is a single feature, while the latter theories select for feature bundles. This seems to be a small difference, but the consequences are rather severe. Stabler's definition of External Merge that was given on page 176 removes the selection feature (=f) and the corresponding feature of the selected element (f). In some publications and in the introduction in this book, the selection features are called uninterpretable features and are marked with a *u*. The uninterpretable features have to be checked and then they are removed from the linguistic object as in Stabler's definition. The fact that they have been checked is represented by striking them out. It is said that all uninterpretable features have to be checked before a syntactic object is sent to the interfaces (semantics and pronunciation). If uninterpretable features are not checked, the derivation crashes. **Adger2003a** explicitly discusses the consequences of these assumptions: a selecting head checks a feature of the selected object. It is not possible to check features of elements that are contained in the object that a head combines with. Only features at the topmost node, the so-called root node, can be checked with external merge. The only way features inside complex objects can be checked is by means of movement. This means that a head may not combine with a partially saturated linguistic object, that is, with a linguistic object that has an unchecked selection feature. I will discuss this design decision with reference to an example provided by **Adger2003a**. The noun *letters* selects for a P and Ps select for an N. The analysis of (59a) is depicted left in Figure 4.23 on the facing page.

- (59) a. letters to Peter
 b. * letters to

Figure 4.23: The analysis of *letters to Peter* according to Adger2003a

The string in (59b) is ruled out since the uninterpretable N feature of the preposition *to* is not checked. So this integrates the constraint that all dependent elements have to be maximal into the core mechanism. This makes it impossible to analyze examples like (60) in the most straightforward way, namely as involving a complex preposition and a noun that is lacking a determiner:

- (60) vom Bus
 from.the bus

In theories in which complex descriptions can be used to describe dependants, the dependent may be partly saturated. So for instance in HPSG, fused prepositions like *vom* ‘from.the’ can select an \bar{N} , which is a nominal projection lacking a specifier:

- (61) $N[SPR \langle DET \rangle]$

The description in (61) is an abbreviation for an internally structured set of feature-value pairs (see Section ??). The example here is given for the illustration of the differences only, since there may be ways of accounting for such cases in a single-feature-Merge system. For instance, one could assume a DP analysis and have the complex preposition select a complete NP (something of category N with no uninterpretable features). Alternatively, one can assume that there is indeed a full PP with all the structure that is usually assumed and the fusion of preposition and determiner happens during pronunciation. The first suggestion eliminates the option of assuming an NP analysis as it was suggested by Bruening2009a in the Minimalist framework.

Apart from this illustrative example with a fused preposition, there are other cases in which one may want to combine unsaturated linguistic objects. I already discussed coordination examples above. Another example is the verbal complex in languages like German, Dutch, and Japanese. Of course there are analyses of these languages that do not assume a verbal complex (G. GMueller98a; Wurmbrand2003b), but these are not without problems. Some of the problems were discussed in the previous section as well.

Summing up this brief subsection, it has to be said that the feature checking mechanism that is built into the conception of Merge is more restrictive than

the selection that is used in Categorical Grammar, Lexical Functional Grammar, HPSG, Construction Grammar, and TAG. In my opinion, it is too restrictive.

4.6.6 Summary

In sum, one can say that the computational mechanisms of the Minimalist Program (e.g., transderivational constraints and labeling), as well as the theory of feature”=driven movement are problematic and the assumption of empty functional categories is sometimes ad hoc. If one does not wish to assume that these categories are shared by all languages, then proposing two mechanisms (Merge and Move) does not represent a simplification of grammar since every single functional category which must be stipulated constitutes a complication of the entire system.

The labeling mechanism is not yet worked out in detail, does not account for the phenomena it was claimed to provide accounts for, and hence should be replaced by the head/functor”=based labeling that is used in Categorical Grammar and HPSG.

4.7 Summary and classification

This section is similar to Section 3.6. I first comment on language acquisition and then on formalization.

4.7.1 Explaining language acquisition

Chomsky2008a counts theories in the MP as Principle & Parameter analyses and identifies MP parameters as being in the lexicon. Also, see Hornstein2013a. UG is defined as possibly containing non”=language”=specific components, which are genetically determined (Chomsky2007a). UG consists of unbounded Merge and the condition that expressions derived by a grammar must fulfill the restrictions imposed by the phonological and conceptual”=intentional interfaces. In addition, a specific repertoire of features is assumed to be part of UG (Chomsky2007a). The exact nature of these features has not been explained in detail and, as a result, the power of UG is somewhat vague. However, there is a fortunate convergence between various linguistic camps as Chomsky does not assume that the swathes of functional projections which we encountered in Section 4.6.1 also form part of UG (however, authors like CR2010a do assume that a hierarchy of functional projections is part of UG). Since there are still parameters, the same arguments

used against GB approaches to language acquisition that were mentioned in Section 3.6.1 are still relevant for theories of language acquisition in the Minimalist Program. See Chapter ?? for an in”=depth discussion of approaches to language acquisition and the Principles & Parameters model as well as input”=based approaches.

Chomsky’s main goal in the Minimalist Program is to simplify the theoretical assumptions regarding formal properties of language and the computational mechanisms that are used so much as to make it plausible that they or relevant parts of them are part of our genetic endowment. But if we recapitulate what was assumed in this chapter, it is difficult to believe that Minimalist theories achieve this goal. To derive a simple sentence with an intransitive verb, one needs several empty heads and movements. Features can be strong or weak, Agree operates nonlocally in trees across several phrase boundaries. And in order to make correct predictions, it has to be made sure that Agree can only see the closest possible element (13)–(14). This is a huge machinery in comparison to a Categorical Grammar that just combines adjacent things. Categorical Grammars can be acquired from input (see Section ??), while it is really hard to imagine how the fact that there are features that trigger movement when they are strong, but do not trigger it when they are weak, should be acquired from data alone.

4.7.2 Formalization

Section 3.6.2 commented on the lack of formalization in transformational grammar up until the 1990s. The general attitude towards formalization did not change in the minimalist era and hence there are very few formalizations and implementations of Minimalist theories.

Stabler2001a shows how it is possible to formalize and implement Kayne’s theory of remnant movement. In Stabler’s implementation³⁹, there are no transderivational constraints, no numerations⁴⁰, he does not assume Agree (see Fong2014a)

³⁹His system is available at: <http://linguistics.ucla.edu/people/stabler/coding.html>. 2018-02-20.

⁴⁰There is a numeration lexicon in Veenstra98a. This lexicon consists of a set of numerations, which contain functional heads, which can be used in sentences of a certain kind. For example, Veenstra assumes numerations for sentences with bivalent verbs and subjects in initial position, for embedded sentences with monovalent verbs, for *wh*=questions with monovalent verbs, and for polar interrogatives with monovalent verbs. An element from this set of numerations corresponds to a particular configuration and a phrasal construction in the spirit of Construction Grammar. Veenstra’s analysis is not a formalization of the concept of the numeration that one finds in Minimalist works. Normally, it is assumed that a numeration contains all the lexical entries which are needed for the derivation of a sentence. As (i) shows, complex

etc. The following is also true of Stabler’s implementation of Minimalist Grammars and GB systems: there are no large grammars. Stabler’s grammars are small, meant as a proof of concept and purely syntactic. There is no morphology⁴¹, no treatment of multiple agreement (Stabler2010b) and above all no semantics. PF and LF processes are not modelled.⁴² The grammars and the computational system developed by Sandiway Fong are of similar size and faithfulness to the theory (FG2012a; Fong2014a): the grammar fragments are small, encode syntactic aspects such as labeling directly in the phrase structure (FG2012a) and therefore, fall behind \bar{X} theory. Furthermore, they do not contain any morphology. Spell-Out is not implemented, so in the end it is not possible to parse or generate any utterances.⁴³ Herring2016a’s (Herring2016a) dissertation is a promising begin-

sentences can consist of combinations of sentences with various different sentence types:

- (i) Der Mann, der behauptet hat, dass Maria gelacht hat, steht neben der Palme,
the man who claimed has that Maria laughed has stands next.to the palm.tree
die im letzten Jahr gepflanzt wurde.
which in last year planted was
‘The man who claimed Maria laughed is standing next to the palm tree that was planted last year.’

In (i), there are two relative clauses with verbs of differing valence, an embedded sentence with a monovalent verb and the matrix clause. Under a traditional understanding of numerations, Veenstra would have to assume an infinite numeration lexicon containing all possible combinations of sentence types.

⁴¹The test sentences have the form as in (i).

- (i) a. the king will -s eat
b. the king have -s eat -en
c. the king be -s eat -ing
d. the king -s will -s have been eat -ing the pie

⁴²See SE2002a for suggestions of PF and LF=“movement and the deletion of parts of copies (p. 285). The implementation of this would be far from trivial.

⁴³The claim by BPYC2011a in reference to Fong’s work is just plain wrong: *But since we have sometimes adverted to computational considerations, as with the ability to “check” features of a head/label, this raises a legitimate concern about whether our framework is computationally realizable. So it is worth noting that the copy conception of movement, along with the locally oriented “search and labeling” procedure described above, can be implemented computationally as an efficient parser; see Fong, 2011, for details.* If one has a piece of software which cannot parse a single sentence, then one cannot claim that it is efficient since one does not know whether the missing parts of the program could make it extremely inefficient. Furthermore, one cannot compare the software to other programs. As has already been discussed, labeling is not carried out by Fong as was described in Chomsky’s work, but instead he uses a phrase structure grammar of the kind described in Chapter 2.

ning. Herring developed a system that can be used for grammar development in the Minimalist Program. In the version described in his thesis the system could generate but was unable to parse sentences (p. 138, 143). PF phenomena were not modeled (p. 142–143) and the two example fragments are small and come without a semantics (p. 143).

The benchmark here has been set by implementations of grammars in constraint-based theories; for example, the HPSG grammars of German (**MK2000a**), English (**FCS2000a**) and Japanese (**Siegel2000a**) that were developed in the 90s as part of *Verbmobil* (**Wahlster2000a-ed-not-crossreferenced**) for the analysis of spoken language or the LFG or CCG systems with large coverage. These grammars can analyze up to 83 % of utterances in spoken language (for *Verbmobil* from the domains of appointment scheduling and trip planning) or written language. Linguistic knowledge is used to generate and analyze linguistic structures. In one direction, one arrives at a semantic representation of a string of words and in the other one can create a string of words from a given semantic representation. A morphological analysis is indispensable for analyzing naturally occurring data from languages with elaborated morphological marking systems. In the remainder of this book, the grammars and computational systems developed in other theories will be discussed at the beginning of the respective chapters.

The reason for the lack of larger fragments inside of GB/MP could have to do with the fact that the basic assumptions of Minimalist community change relatively quickly:

In Minimalism, the triggering head is often called a *probe*, the moving element is called a *goal*, and there are various proposals about the relations among the features that trigger syntactic effects. **Chomsky95a-u** begins with the assumption that features represent requirements which are checked and deleted when the requirement is met. The first assumption is modified almost immediately so that only a proper subset of the features, namely the ‘formal’, ‘uninterpretable’ features are deleted by checking operations in a successful derivation (Collins, 1997; **Chomsky95a-u**). Another idea is that certain features, in particular the features of certain functional categories, may be initially unvalued, becoming valued by entering into appropriate structural configurations with other elements (**Chomsky2008a**; Hiraiwa, 2005). And some recent work adopts the view that features are never deleted (**Chomsky2007a**). These issues remain unsolved. (**Stabler2010a**)

In order to fully develop a grammar fragment, one needs at least three years (compare the time span between the publication of *Barriers* (1986) and Stabler’s

implementation (1992)). Particularly large grammars require the knowledge of several researchers working in international cooperation over the space of years or even decades. This process is disrupted if fundamental assumptions are repeatedly changed at short intervals.

As far as large-scale coverage is concerned, the more recent work by John Torr is an exception to what was said above.⁴⁴ **TSSC2019a** state that their parser is the first one to take up the Sproat & Lappin Challenge to the Minimalist community (**SL2005a**). The work of the authors is impressive and they really implemented a wide-coverage statistically trained parser based on Transformational Grammar, but what they did is different from standard Minimalism since they assume “around 45” versions of Move and Merge (p. 2488) in comparison to the two versions usually assumed in Minimalism (Move and Merge or Internal and External Merge).⁴⁵ **TS2016a** explain some of the schemata that are assumed: there are versions of Merge that combine a head with a complement and versions that combine a head with a specifier (see **MuellerUnifying** and Section 4.6.4.2 above for a comparison of Minimalist Grammars with HPSG. The respective variants of Merge correspond to HPSG’s Specifier-Head Schema and the Head-Complement Schema, respectively). **TS2016a** assume four schemata for adjunction (HPSG has one such schema and use underspecification with respect to order, see **MuellerOrder**). They assume a special rule for rightward movement (p. 5) corresponding to **Keller95b**’s (**Keller95b**) and **Mueller99a**’s (**Mueller99a**) Head-Extra Schema for extraposition. In addition the authors assume two schemata for head movement. HPSG assumes a lexical rule or a unary branching schema applying to words or coordinations of words (**MuellerGSMuellerOrder**)). Across the Board extraction (**Ross67a**) is taken care of by four special schemata. See **AC2020a** for an overview of treatments of coordination in HPSG. The treatment of Across the Board Extraction is non-standard Minimalism. For the analysis of examples like (62) in which one filler corresponds to two gaps in two conjuncts, the authors build on **Kobele2008a** who uses a SLASH passing mechanism going back to **Sag83a-u** and **Gazdar81a**. While Koeble assumes the SLASH passing mechanism of GPSG, **TS2016a** suggest an analysis of (62) with two instances of *who* in object positions, which are later unified into one when the second conjunct is merged into the main structure.

⁴⁴It is not an exception as far as theory development is concerned. Torr’s system is based on **Chomsky95a-u**, so he did not follow new trends but stayed within a certain setting.

⁴⁵Torr explained in p.c. 2019 that these 45 rules can be folded into two Merge functions and two Move functions. But in the end this is just a clever way of hiding complexity. It is like **Chomsky2005a** revising the theory with Move and Merge into one with just one operation Merge but assuming two subcases of Internal and External Merge.

(62) Who_i did Jack say Mary likes _{-i} and Pete hates _i?

An interesting property of the analysis is that *who Pete hates* forms a discontinuous constituent: *who* is combined with *hates* despite its sentence-initial position. Information about this *wh* element is passed up the tree in an GPSG-style way. The difference is that there is no trace but the extracted element is identical in phonological material with the filler. Interestingly, there is an HPSG variant of nonlocal dependencies that is very similar to what TS2016a suggest and together with a modified Filler Head Schema the analyses are parallel: HN94b suggested that the linguistic objects that are involved in nonlocal dependencies are of type *sign* rather than *local*. This makes it possible to pass up information about a daughter including its phonological make up. If one assumes a version of HPSG permitting discontinuous constituents (Reape94a; Kathol2001a; Mueller95c; Mueller2004b and Section ?? of this book) and a Filler Head Schema that requires that the phonology of the filler is identical to the phonology in the SLASH list and that does not insert the fronted element into the constituent order domain of the head (since it is in there already), we get an analysis of the type described in TS2016a. Figure ?? shows the analysis that was suggested by TS2016a and Figure ?? the HPSG analog. Directional Minimalist Grammars use

Figure 4.24: Derivation tree of *who Jack likes* in Directional Minimalist Grammar according to TS2016a

the '=' sign to indicate the direction in which an argument is required. =d means that a DP is required to the left of a head and d= encode the requirement of a DP to the right. This is like the '/' notation of Categorical Grammar (see Chapter ??). *likes* has the category d= =d v, which means that it is a verb requiring a d to its right (the object) and a d to its left (the subject). *who* is of category d and has a -wh feature, something that has to be checked for a derivation to be complete. *Jack* is the subject of *likes* and fulfills the =d requirement of *likes*. Items like [pres] and [int] are empty elements. [pres] has a +case feature and can make *Jack* move to its specifier. The movement consumes the -case feature and puts *Jack* to the front of the string. This looks like a unary projection in the derivation tree. The empty interrogative head [int] selects for a t to its right. The result is a C projection that has a +wh feature. In the final step *who*, which is -wh moves to the left and the wh features are removed. The important thing is that the information about the phonology of *who* and its wh feature is percolated up in the tree until it is finally bound off in the last derivation step.

Figure ?? shows the HPSG analog. The information about the local properties

of the *wh* word including its phonology are passed up in the tree until they are bound off in a filler head configuration. The Filler-Head Schema binds off the nonlocal dependency and makes sure that the phonology of the filler is not realized twice (see **Reape94a**, **MuellerOrder** on linearization domains and **AC2020a** on multi-dominance approaches in HPSG). An alternative to a binary branching

Figure 4.25: Possible HPSG analysis of *who Jack likes* using discontinuous constituents

Filler-Head Schema would be a unary branching rule that binds off the element in SLASH and adds the stored phonology to the phonology of the daughter. This would then be completely parallel to the unary branching assumed in Torr’s Directional Minimalist Grammar.

Concluding the discussion of Torr’s work, it can be said that it is truly impressive but that it shows a convergence between Minimalism (or rather Minimalist Grammar) and HPSG. Tools from GPSG/HPSG were adopted and the outcome differs in crucial aspects from what is taught in Minimalist textbooks (just one or two instances of Merge vs. 45, transformations vs. GPSG-style percolation of features).



Further reading

This chapter heavily draws on **Adger2003a**. Other textbooks on Minimalism are **Radford97a-u**, **Grewendorf2002a**, and **HNG2005a**.

Kuhn2007a offers a comparison of modern derivational analyses with constraint"-based LFG and HPSG approaches. **Borsley2012a** contrasts analyses of long"-distance dependencies in HPSG with movement"-based analyses as in GB/"Minimalism. Borsley discusses four types of data which are problematic for movement"-based approaches: extraction without fillers, extraction with multiple gaps, extractions where fillers and gaps do not match and extraction without gaps. **BM2020Minimalism** is another comparison of Minimalism and HPSG. The authors discuss differences of ap-

proach and outlook of the two frameworks (sociological differences, formalization and exhaustivity), empirical quality of the work, differences in assumed syntactic structures, psycholinguistic issues and the assumptions made in the frameworks regarding language acquisition.

The discussion of labeling, abandonment of \bar{X} theory and a comparison between Stabler's Minimalist Grammars and HPSG from Sections 4.6.2–4.6.4 can be found in **MuellerUnifying**.

Intonational Phrasing, Discontinuity, and the Scope of Negation by **BG2005a** is recommended for the more advanced reader. The authors compare analyses of negated quantifiers with wide scope in the framework of Minimalism (following Kayne) as well as Categorical Grammar (following Steedman).

Sternefeld2006a-u is a good, detailed introduction to syntax (839 pages) which develops a Transformational Grammar analysis of German which (modulo transformations) almost matches what is assumed in HPSG (feature descriptions for arguments ordered in a valence list according to a hierarchy). Sternefeld's structures are minimal since he does not assume any functional projections if they cannot be motivated for the language under discussion. Sternefeld is critical regarding certain aspects which some other analyses take for granted. Sternefeld views his book explicitly as a textbook from which one can learn how to argue coherently when creating theories. For this reason, this book is not just recommended for students and PhD students.

BM2020Minimalism compares Minimalism with HPSG.

SR2012a discuss the situation in theoretical linguistics with particular focus on the theories described in this and the previous chapter. I can certainly understand the frustration of the authors with regard to the vagueness of analyses, argumentation style, empirical base of research, rhetorical clichés, immunization attempts and general respect for scientific standards: a current example of this is the article *Problems of Projection* by **Chomsky2013a**.^a I, however, do not share the general, pessimistic tone of this article. In my opinion, the patient's condition is critical, but he is not dead yet. As a reviewer of the Sternefeld and Richter paper pointed out, the situation in linguistics has changed so much that now having a dissertation from MIT does not necessarily guarantee you a position (footnote 16) later on. One could view a reorientation of certain scientists with

regard to certain empirical questions, adequate handling of data (Fanselow Fanselow2004b; Fanselow2009a) and improved communication between theoretical camps as a way out of this crisis.

Since the 90s, it is possible to identify an increased empirical focus (especially in Germany), which manifests itself, for example, in the work of linguistic Collaborative Research Centers (SFBs) or the yearly *Linguistic Evidence* conference. As noted by the reviewer cited above, in the future, it will not be enough to focus on Chomsky's problems in determining the syntactic categories in sentences such as *He left* (see Section 4.6.2). Linguistic dissertations will have to have an empirical section, which shows that the author actually understands something about language. Furthermore, dissertations, and of course other publications, should give an indication that the author has not just considered theories from a particular framework but is also aware of the broad range of relevant descriptive and theoretical literature.

^aVagueness: in this article, *perhaps* occurs 19 times, *may* 17 as well as various *if*'s. Consistency: the assumptions made are inconsistent. See footnote 18 on page 165 of this book. Argumentation style: the term specifier is abolished and it is claimed that the problems associated with this term can no longer be formulated. Therefore, they are now not of this world. See footnote 28 on page 171 of this book. Immunization: Chomsky writes the following regarding the Empty Category Principle: *apparent exceptions do not call for abandoning the generalization as far as it reaches, but for seeking deeper reasons to explain where and why it holds* p. 9. This claim is most certainly correct, but one wonders how much evidence one needs in a specific case in order to disregard a given analysis. In particular regarding the essay *Problems of Projection*, one has to wonder why this essay was even published only five years after *On phases*. The evidence against the original approach is overwhelming and several points are taken up by Chomsky2013a himself. If Chomsky were to apply his own standards (for a quote of his from 1957, see page 6) as well as general scientific methods (Occam's Razor), the consequence would surely be a return to head"-based analyses of labeling.

For detailed comments on this essay, see Sections 4.6.2 and 4.6.3.

As I have shown in Section 4.6.4 and in MuellerUnifying and will also

show in the following chapters and the discussion chapters in particular, there are most certainly similarities between the various analyses on the market and they do converge in certain respects. The way of getting out of the current crisis lies with the empirically”=grounded and theoretically broad education and training of following generations.

In short: both teachers and students should read the medical record by Sternefeld and Richter. I implore the students not to abandon their studies straight after reading it, but rather to postpone this decision at least until after they have read the remaining chapters of this book.

5 Generalized Phrase Structure Grammar

Generalized Phrase Structure Grammar (GPSG) was developed as an answer to Transformational Grammar at the end of the 1970s. The book by **GKPS85a** is the main publication in this framework. Hans Uszkoreit has developed a largish GPSG fragment for German (**Uszkoreit87a**). Analyses in GPSG were so precise that it was possible to use them as the basis for computational implementations. The following is a possibly incomplete list of languages with implemented GPSG fragments:

- German (**Weisweber87a-u**; **WP92b**; **Naumann87a-u**; **Naumann88-u-gekauft**; **Volk88**)
- English (**Evans85a-u**; **PT85a-u**; **Phillips92a-u**; **GCB93a-u**)
- French (**EdSB96a**)
- Persian (**BSM2011a**)

As was discussed in Section 3.1.1, **Chomsky57a** argued that simple phrase structure grammars are not well-suited to describe relations between linguistic structures and claimed that one needs transformations to explain them. These assumptions remained unchallenged for two decades (with the exception of publications by **Harman63a** and **Freidin75a**) until alternative theories such as LFG and GPSG emerged, which addressed Chomsky's criticisms and developed non"-transformational explanations of phenomena for which there were previously only transformational analyses or simply none at all. The analysis of local reordering of arguments, passives and long"-distance dependencies are some of the most important phenomena that have been discussed in this framework. Following some introductory remarks on the representational format of GPSG in Section 5.1, I will present the GPSG analyses of these phenomena in some more detail.

5.1 General remarks on the representational format

This section has five parts. The general assumptions regarding features and the representation of complex categories is explained in Section 5.1.1, the assumptions regarding the linearization of daughters in a phrase structure rule is explained in Section 5.1.2. Section 5.1.3 introduces metarules, Section 5.1.4 deals with semantics, and Section 5.1.5 with adjuncts.

5.1.1 Complex categories, the Head Feature Convention, and \bar{X} rules

In Section 2.2, we augmented our phrase structure grammars with features. GPSG goes one step further and describes categories as sets of feature=value pairs. The category in (1a) can be represented as in (1b):

- (1) a. NP(3,sg,nom)
 b. { CAT n, BAR 2, PER 3, NUM sg, CASE nom }

It is clear that (1b) corresponds to (1a). (1a) differs from (1b) with regard to the fact that the information about part of speech and the \bar{X} level (in the symbol NP) are prominent, whereas in (1b) these are treated just like the information about case, number or person.

Lexical entries have a feature SUBCAT. The value is a number which says something about the kind of grammatical rules in which the word can be used. (2) shows examples for grammatical rules and lists some verbs which can occur in these rules.¹

- | | | |
|-----|---|---|
| (2) | V2 \rightarrow H[5] | (<i>kommen</i> ‘come’, <i>schlafen</i> ‘sleep’) |
| | V2 \rightarrow H[6], N2[CASE acc] | (<i>kennen</i> ‘know’, <i>suchen</i> ‘search’) |
| | V2 \rightarrow H[7], N2[CASE dat] | (<i>helfen</i> ‘help’, <i>vertrauen</i> ‘trust’) |
| | V2 \rightarrow H[8], N2[CASE dat], N2[CASE acc] | (<i>geben</i> ‘give’, <i>zeigen</i> ‘show’) |
| | V2 \rightarrow H[9], V3[+dass] | (<i>wissen</i> ‘know’, <i>glauben</i> ‘believe’) |

These rules license VPs, that is, the combination of a verb with its complements, but not with its subject. The numbers following the category symbols (V or N) indicate the \bar{X} projection level. For Uszkoreit, the maximum number of projections of a verbal projection is three rather than two as is often assumed.

The H on the right side of the rule stands for *head*. The *Head Feature Convention* (HFC) ensures that certain features of the mother node are also present on the node marked with H (for details see GKPS85a and Uszkoreit87a):

¹The analyses discussed in the following are taken from Uszkoreit87a.

Principle 1 (Head Feature Convention)

The mother node and the head daughter must bear the same head features unless indicated otherwise.

In (2), examples for verbs which can be used in the rules are given in brackets. As with ordinary phrase structure grammars, one also requires corresponding lexical entries for verbs in GPSG. Two examples are provided in (3):

- (3) V[5, VFORM *inf*] → einzuschlafen
V[6, VFORM *inf*] → aufzuessen

The first rule states that *einzuschlafen* ‘to fall asleep’ has a SUBCAT value of 5 and the second indicates that *aufzuessen* ‘to finish eating’ has a SUBCAT value of 6. It follows, then, that *einzuschlafen* can only be used in the first rule (2) and *aufzuessen* can only be used in the second. Furthermore, (3) contains information about the form of the verb (*inf* stands for infinitives with *zu* ‘to’).

If we analyze the sentence in (4) with the second rule in (2) and the second rule in (3), then we arrive at the structure in Figure 5.1.

- (4) Karl hat versucht, [den Kuchen aufzuessen].
Karl has tried the cake to.eat.up
‘Karl tried to finish eating the cake.’

Figure 5.1: Projection of head features in GPSG

The rules in (2) say nothing about the order of the daughters which is why the verb (H[6]) can also be in final position. This aspect will be discussed in more detail in Section 5.1.2. With regard to the HFC, it is important to bear in mind that information about the infinitive verb form is also present on the mother node. Unlike simple phrase structure rules such as those discussed in Chapter 2, this follows automatically from the Head Feature Convention in GPSG. In (3), the value of VFORM is given and the HFC ensures that the corresponding information is represented on the mother node when the rules in (2) are applied. For the phrase in (4), we arrive at the category V2[VFORM *inf*] and this ensures that this phrase only occurs in the contexts it is supposed to:

- (5) a. [Den Kuchen aufzuessen] hat er nicht gewagt.
the cake to.eat.up has he not dared
‘He did not dare to finish eating the cake.’

- b. * [Den Kuchen aufzuessen] darf er nicht.
 the cake to.eat.up be.allowed.to he not
 Intended: 'He is not allowed to finish eating the cake.'
- c. * [Den Kuchen aufessen] hat er nicht gewagt.
 the cake eat.up has he not dared
 Intended: 'He did not dare to finish eating the cake.'
- d. [Den Kuchen aufessen] darf er nicht.
 the cake eaten.up be.allowed.to he not
 'He is not allowed to finish eating the cake.'

gewagt 'dared' selects for a verb or verb phrase with an infinitive with *zu* 'to' but not a bare infinitive, while *darf* 'be allowed to' takes a bare infinitive.

This works in an analogous way for noun phrases: there are rules for nouns which do not take an argument as well as for nouns with certain arguments. Examples of rules for nouns which either require no argument or two PPs are given in (6) (GKPS85a):

- (6) $N1 \rightarrow H[30]$ (*Haus* 'house', *Blume* 'flower')
 $N1 \rightarrow H[31], PP[mit], PP[über]$ (*Gespräch* 'talk', *Streit* 'argument')

The rule for the combination of \bar{N} and a determiner is as follows:

- (7) $N2 \rightarrow Det, H1$

$N2$ stands for NP, that is, for a projection of a noun phrase on bar level two, whereas $H1$ stands for a projection of the head daughter on the bar level one. The Head Feature Convention ensures that the head daughter is also a nominal projection, since all features on the head daughter apart from the \bar{X} level are identified with those of the whole NP. When analyzing (8), the second rule in (6) licenses the \bar{N} *Gesprächs mit Maria über Klaus*. The fact that *Gesprächs* 'conversation' is in the genitive is represented in the lexical item of *Gesprächs* and since *Gesprächs* is the head, it is also present at \bar{N} , following the Head Feature Convention.

- (8) des Gespräch-s mit Maria über Klaus
 the.GEN conversation-GEN with Maria about Klaus
 'the conversation with Maria about Klaus'

For the combination of \bar{N} with the determiner, we apply the rule in (7). The category of the head determines the word class of the element on the left-hand side of the rule, which is why the rule in (7) corresponds to the classical \bar{X} rules that

we encountered in (65c) on page 85. Since *Gesprächs mit Maria über Klaus* is the head daughter, the information about the genitive of \bar{N} is also present at the NP node.

5.1.2 Local reordering

The first phenomenon to be discussed is local reordering of arguments. As was already discussed in Section 3.5, arguments in the middle field can occur in an almost arbitrary order. (9) gives some examples:

- (9) a. [weil] der Mann der Frau das Buch gibt
 because the.NOM man the.DAT woman the.ACC book gives
 ‘because the man gives the book to the woman’
 b. [weil] der Mann das Buch der Frau gibt
 because the.NOM man the.ACC book the.DAT woman gives
 c. [weil] das Buch der Mann der Frau gibt
 because the.ACC book the.NOM man the.DAT woman gives
 d. [weil] das Buch der Frau der Mann gibt
 because the.ACC book the.DAT woman the.NOM man gives
 e. [weil] der Frau der Mann das Buch gibt
 because the.DAT woman the.NOM man the.ACC book gives
 f. [weil] der Frau das Buch der Mann gibt
 because the.DAT woman the.ACC book the.NOM man gives

In the phrase structure grammars in Chapter 2, we used features to ensure that verbs occur with the correct number of arguments. The following rule in (10) was used for the sentence in (9a):

- (10) $S \rightarrow NP[nom] NP[dat] NP[acc] V_nom_dat_acc$

If one wishes to analyze the other orders in (9), then one requires an additional five rules, that is, six in total:

- (11) $S \rightarrow NP[nom] NP[dat] NP[acc] V_nom_dat_acc$
 $S \rightarrow NP[nom] NP[acc] NP[dat] V_nom_dat_acc$
 $S \rightarrow NP[acc] NP[nom] NP[dat] V_nom_dat_acc$
 $S \rightarrow NP[acc] NP[dat] NP[nom] V_nom_dat_acc$
 $S \rightarrow NP[dat] NP[nom] NP[acc] V_nom_dat_acc$
 $S \rightarrow NP[dat] NP[acc] NP[nom] V_nom_dat_acc$

In addition, it is necessary to postulate another six rules for the orders with verb=initial order:

- (12) $S \rightarrow V_nom_dat_acc\ NP[nom]\ NP[dat]\ NP[acc]$
 $S \rightarrow V_nom_dat_acc\ NP[nom]\ NP[acc]\ NP[dat]$
 $S \rightarrow V_nom_dat_acc\ NP[acc]\ NP[nom]\ NP[dat]$
 $S \rightarrow V_nom_dat_acc\ NP[acc]\ NP[dat]\ NP[nom]$
 $S \rightarrow V_nom_dat_acc\ NP[dat]\ NP[nom]\ NP[acc]$
 $S \rightarrow V_nom_dat_acc\ NP[dat]\ NP[acc]\ NP[nom]$

Furthermore, one would also need parallel rules for transitive and intransitive verbs with all possible valences. Obviously, the commonalities of these rules and the generalizations regarding them are not captured. The point is that we have the same number of arguments, they can be realized in any order and the verb can be placed in initial or final position. As linguists, we find it desirable to capture this property of the German language and represent it beyond phrase structure rules. In Transformational Grammar, the relationship between the orders is captured by means of movement: the Deep Structure corresponds to verb=final order with a certain order of arguments and the surface order is derived by means of Move- α . Since GPSG is a non=transformational theory, this kind of explanation is not possible. Instead, GPSG imposes restrictions on *immediate dominance* (ID), which differ from those which refer to *linear precedence* (LP): rules such as (13) are to be understood as dominance rules, which do not have anything to say about the order of the daughters (Pullum82a).

- (13) $S \rightarrow V, NP[nom], NP[acc], NP[dat]$

The rule in (13) simply states that S dominates all other nodes. Due to the abandonment of ordering restrictions for the right=hand side of the rule, we only need one rule rather than twelve.

Nevertheless, without any kind of restrictions on the right=hand side of the rule, there would be far too much freedom. For example, the following order would be permissible:

- (14) * Der Frau der Mann gibt ein Buch.
 the woman.DAT the.NOM man gives the.ACC book

Such orders are ruled out by so=called *Linear Precedence Rules* or LP=rules. LP=constraints are restrictions on local trees, that is, trees with a depth of one. It is, for example, possible to state something about the order of V, NP[nom], NP[acc] and NP[dat] in Figure 5.2 on the facing page using linearization rules. The following linearization rules serve to exclude orders such as those in (14):

Figure 5.2: Example of a local tree

- (15) $V[+MC] < X$
 $X < V[-MC]$

MC stands for *main clause*. The LP"-rules ensure that in main clauses (+MC), the verb precedes all other constituents and follows them in subordinate clauses (-MC). There is a restriction that says that all verbs with the MC"-value '+' also have to be (+FIN). This will rule out infinitive forms in initial position.

These LP rules do not permit orders with an occupied prefield or postfield in a local tree. This is intended. We will see how fronting can be accounted for in Section 5.4.

5.1.3 Metarules

We have previously encountered linearization rules for sentences with subjects, however our rules have the form in (16), that is, they do not include subjects:

- (16) $V2 \rightarrow H[7], N2[CASE\ dat]$
 $V2 \rightarrow H[8], N2[CASE\ dat], N2[CASE\ acc]$

These rules can be used to analyze the verb phrases *dem Mann das Buch zu geben* 'to give the man the book' and *das Buch dem Mann zu geben* 'to give the book to the man' as they appear in (17), but we cannot analyze sentences like (9), since the subject does not occur on the right"-hand side of the rules in (16).

- (17) a. Er verspricht, [dem Mann das Buch zu geben].
he promises the.DAT man the.ACC book to give
'He promises to give the man the book.'
b. Er verspricht, [das Buch dem Mann zu geben].
he promises the.ACC book the.DAT man to give
'He promises to give the book to the man.'

A rule with the format of (18) does not make much sense for a GPSG analysis of German since it cannot derive all the orders in (9) as the subject can occur between the elements of the VP as in (9c).

- (18) $S \rightarrow N2\ V2$

With the rule in (18), it is possible to analyze (9a) as in Figure 5.3 on the next page and it would also be possible to analyze (9b) with a different ordering of

Figure 5.3: VP analysis for German (not appropriate in the GPSG framework)

the NPs inside the VP. The remaining examples in (9) cannot be captured by the rule in (18), however. This has to do with the fact that only elements in the same local tree, that is, elements which occur on the right”=hand side of a rule, can be reordered. While we can reorder the parts of the VP and thereby derive (9b), it is not possible to place the subject at a lower position between the objects. Instead, a metarule can be used to analyze sentences where the subject occurs between other arguments of the verb. This rule relates phrase structure rules to other phrase structure rules. A metarule can be understood as a kind of instruction that creates another rule for each rule with a certain form and these newly created rules will in turn license local trees.

For the example at hand, we can formulate a metarule which says the following: if there is a rule with the form “V2 consists of something” in the grammar, then there also has to be another rule “V3 consists of whatever V2 consists + an NP in the nominative”. In formal terms, this looks as follows:

- (19) $V2 \rightarrow W \mapsto$
 $V3 \rightarrow W, N2[\text{CASE nom}]$

W is a variable which stands for an arbitrary number of categories ($W = \text{whatever}$). The metarule creates the following rule in (20) from the rules in (16):

- (20) $V3 \rightarrow H[7], N2[\text{CASE dat}], N2[\text{CASE nom}]$
 $V3 \rightarrow H[8], N2[\text{CASE dat}], N2[\text{CASE acc}], N2[\text{CASE nom}]$

Now, the subject and other arguments both occur in the right”=hand side of the rule and can therefore be freely ordered as long as no LP rules are violated.

5.1.4 Semantics

The semantics adopted by GKPS85a goes back to Richard Montague74a-u. Unlike a semantic theory which stipulates the combinatorial possibilities for each rule (see Section 2.3), GPSG uses more general rules. This is possible due to the fact that the expressions to be combined each have a semantic type. It is customary to distinguish between entities (e) and truth values (t). Entities refer to an object in the world (or in a possible world), whereas entire sentences are either true or false, that is, they have a truth value. It is possible to create more complex types from the types e and t . Generally, the following holds: if a and b are types, then $\langle a, b \rangle$ is also a type. Examples of complex types are $\langle e, t \rangle$ and $\langle e, \langle e, t \rangle \rangle$.

We can define the following combinatorial rule for this kind of typed expressions:

- (21) If α is of type $\langle b, a \rangle$ and β of type b , then $\alpha(\beta)$ is of type a .

This type of combination is also called *functional application*. With the rule in (21), it is possible that the type $\langle e, \langle e, t \rangle \rangle$ corresponds to an expression which still has to be combined with two expressions of type e in order to result in an expression of type t . The first combination step with e will yield $\langle e, t \rangle$ and the second step of combination with a further e will give us t . This is similar to what we saw with λ -expressions on page 73: $\lambda y \lambda x \text{ like}'(x, y)$ has to combine with a y and an x . The result in this example was $\text{mögen}'(\text{max}', \text{lotte}')$, that is, an expression that is either true or false in the relevant world.

In GKPS85a, an additional type is assumed for worlds in which an expression is true or false. For reasons of simplicity, I will omit this here. The types that we need for sentences, NPs and N's, determiners and VPs are given in (22):

- (22) a. $\text{TYP}(S) = t$
 b. $\text{TYP}(\text{NP}) = \langle \langle e, t \rangle, t \rangle$
 c. $\text{TYP}(\text{N}') = \langle e, t \rangle$
 d. $\text{TYP}(\text{Det}) = \langle \text{TYP}(\text{N}'), \text{TYP}(\text{NP}) \rangle$
 e. $\text{TYP}(\text{VP}) = \langle e, t \rangle$

A sentence is of type t since it is either true or false. A VP needs an expression of type e to yield a sentence of type t . The type of the NP may seem strange at first glance, however, it is possible to understand it if one considers the meaning of NPs with quantifiers. For sentences such as (23a), a representation such as (23b) is normally assumed:

- (23) a. All children laugh.
 b. $\forall x \text{ child}'(x) \rightarrow \text{laugh}'(x)$

The symbol \forall stands for the universal quantifier. The formula can be read as follows. For every object, for which it is the case that it has the property of being a child, it is also the case that it is laughing. If we consider the contribution made by the NP, then we see that the universal quantifier, the restriction to children and the logical implication come from the NP:

- (24) $\forall x \text{ child}'(x) \rightarrow P(x)$

This means that an NP is something that must be combined with an expression which has exactly one open slot corresponding to the x in (24). This is formulated in (22b): an NP corresponds to a semantic expression which needs something of type $\langle e, t \rangle$ to form an expression which is either true or false (that is, of type t).

An N' stands for a nominal expression for the kind $\lambda x \text{ child}(x)$. This means if there is a specific individual which one can insert in place of the x , then we arrive at an expression that is either true or false. For a given situation, it is the case that either John has the property of being a child or he does not. An N' has the same type as a VP.

$TYP(N')$ and $TYP(NP)$ in (22d) stand for the types given in (22c) and (22b), that is, a determiner is semantically something which has to be combined with the meaning of N' to give the meaning of an NP.

GKPS85a point out a redundancy in the semantic specification of grammars which follow the rule="to"=rule hypothesis (see Section 2.3) since, instead of giving rule="by"=rule instructions with regard to combinations, it suffices in many cases simply to say that the functor is applied to the argument. If we use types such as those in (22), it is also clear which constituent is the functor and which is the argument. In this way, a noun cannot be applied to a determiner, but rather only the reverse is possible. The combination in (25a) yields a well-formed result, whereas (25b) is ruled out.

- (25) a. $Det'(N')$
- b. $N'(Det')$

The general combinatorial principle is then as follows:

- (26) Use functional application for the combination of the semantic contribution of the daughters to yield a well-formed expression corresponding to the type of the mother node.

The authors of the GPSG book assume that this principle can be applied to the vast majority of GPSG rules so that only a few special cases have to be dealt with by explicit rules.

5.1.5 Adjuncts

For nominal structures in English, GKPS85a assume the \bar{X} analysis and, as we have seen in Section 2.4.1, this analysis is applicable to nominal structures in German. Nevertheless, there is a problem regarding the treatment of adjuncts in the verbal domain if one assumes flat branching structures, since adjuncts can freely occur between arguments:

- (27) a. weil der Mann der Frau das Buch *gestern* gab
 because the man the woman the book yesterday gave
 'because the man gave the book to the woman yesterday'
- b. weil der Mann der Frau *gestern* das Buch gab
 because the man the woman yesterday the book gave

- c. weil der Mann *gestern* der Frau das Buch gab
 because the man yesterday the woman the book gave
- d. weil *gestern* der Mann der Frau das Buch gab
 because yesterday the man the woman the book gave

For (27), one requires the following rule:

- (28) $V3 \rightarrow H[8], N2[\text{CASE dat}], N2[\text{CASE acc}], N2[\text{CASE nom}], \text{AdvP}$

Of course, adjuncts can also occur between the arguments of verbs from other valence classes:

- (29) weil (oft) die Frau (oft) dem Mann (oft) hilft
 because often the woman often the man often helps
 ‘because the woman often helps the man’

Furthermore, adjuncts can occur between the arguments of a VP:

- (30) Der Mann hat versucht, der Frau heimlich das Buch zu geben.
 the man has tried the woman secretly the book to give
 ‘The man tried to secretly give the book to the woman.’

In order to analyze these sentences, we can use a metarule which adds an adjunct to the right”=hand side of a V2 (Uszkoreit87a).

- (31) $V2 \rightarrow W \mapsto$
 $V2 \rightarrow W, \text{AdvP}$

By means of the subject introducing metarule in (19), the V3”=rule in (28) is derived from a V2”=rule. Since there can be several adjuncts in one sentence, a metarule such as (31) must be allowed to apply multiple times. The recursive application of metarules is often ruled out in the literature due to reasons of generative capacity (see Chapter ??) (Thompson82a-u; Uszkoreit87a). If one uses the Kleene star, then it is possible to formulate the adjunct metarule in such a way that it does not have to apply recursively (Uszkoreit87a):

- (32) $V2 \rightarrow W \mapsto$
 $V2 \rightarrow W, \text{AdvP}^*$

If one adopts the rule in (32), then it is not immediately clear how the semantic contribution of the adjuncts can be determined.² For the rule in (31), one can

²In LFG, an adjunct is entered into a set in the functional structure (see Section ??). This also works with the use of the Kleene Star notation. From the f”=structure, it is possible to compute

combine the semantic contribution of the AdvP with the semantic contribution of the V2 in the input rule. This is of course also possible if the metarule is applied multiple times. If this metarule is applied to (33a), for example, the V2"-node in (33a) contains the semantic contribution of the first adverb.

- (33) a. $V2 \rightarrow V, NP, AdvP$
 b. $V2 \rightarrow V, NP, AdvP, AdvP$

The V2"-node in (33b) receives the semantic representation of the adverb applied to the V2"-node in (33a).

WP92b have shown that it is possible to use metarules such as (31) if one does not use metarules to compute a set of phrase structure rules, but rather directly applies the metarules during the analysis of a sentence. Since sentences are always of finite length and the metarule introduces an additional AdvP to the right"-hand side of the newly licensed rule, the metarule can only be applied a finite number of times.

5.2 Passive as a metarule

The German passive can be described in an entirely theory"-neutral way as follows:³

- The subject is suppressed.
- If there is an accusative object, this becomes the subject.

This is true for all verb classes which can form the passive. It does not make a difference whether the verbs takes one, two or three arguments:

- (34) a. weil er noch gearbeitet hat
 because he.NOM still worked has
 'because he has still worked'

the semantic denotation with corresponding scope by making reference to the c"-structure. In HPSG, Kasper94a has made a proposal which corresponds to the GPSG proposal with regard to flat branching structures and an arbitrary number of adjuncts. In HPSG, however, one can make use of so"-called relational constraints. These are similar to small programs which can create relations between values inside complex structures. Using such relational constraints, it is then possible to compute the meaning of an unrestricted number of adjuncts in a flat branching structure.

³This characterization does not hold for other languages. For instance, Icelandic allows for dative subjects. See ZMT85a.

- b. weil noch gearbeitet wurde
because still worked was
'because there was still working there'
- (35) a. weil er an Maria gedacht hat
because he.NOM on Maria thought has
'because he thought of Maria'
b. weil an Maria gedacht wurde
because on Maria thought was
'because Maria was thought of'
- (36) a. weil sie ihn geschlagen hat
because she.NOM him.ACC beaten has
'because she has beaten him'
b. weil er geschlagen wurde
because he.NOM beaten was
'because he was beaten'
- (37) a. weil er ihm den Aufsatz gegeben hat
because he.NOM him.DAT the.ACC essay given has
'because he has given him the essay'
b. weil ihm der Aufsatz gegeben wurde
because him.DAT the.NOM essay given was
'because he was given the essay'

In a simple phrase structure grammar, we would have to list two separate rules for each pair of sentences making reference to the valence class of the verb in question. The characteristics of the passive discussed above would therefore not be explicitly stated in the set of rules. In GPSG, it is possible to explain the relation between active and passive rules using a metarule: for each active rule, a corresponding passive rule with suppressed subject is licensed. The link between active and passive clauses can therefore be captured in this way.

An important difference to Transformational Grammar/GB is that we are not creating a relation between two trees, but rather between active and passive rules. The two rules license two unrelated structures, that is, the structure of (38b) is not derived from the structure of (38a).

- (38) a. weil sie ihn geschlagen hat
 because she.NOM him.ACC beaten has
 ‘because she has beaten him’
 b. weil er geschlagen wurde
 because he.NOM beaten was
 ‘because he was beaten’

The generalization with regard to active/passive is captured nevertheless.

In what follows, I will discuss the analysis of the passive given in **GKPS85a** in some more detail. The authors suggest the following metarule for English (p. 59):⁴

- (39) $VP \rightarrow W, NP \mapsto$
 $VP[PAS] \rightarrow W, (PP[by])$

This rule states that verbs which take an object can occur in a passive VP without this object. Furthermore, a *by*-PP can be added. If we apply this metarule to the rules in (40), then this will yield the rules listed in (41):

- (40) $VP \rightarrow H[2], NP$
 $VP \rightarrow H[3], NP, PP[to]$
 (41) $VP[PAS] \rightarrow H[2], (PP[by])$
 $VP[PAS] \rightarrow H[3], PP[to], (PP[by])$

It is possible to use the rules in (40) to analyze verb phrases in active sentences:

- (42) a. [_S The man [_{VP} devoured the carcass]].
 b. [_S The man [_{VP} handed the sword to Tracy]].

The combination of a VP with the subject is licensed by an additional rule ($S \rightarrow NP, VP$).

With the rules in (41), one can analyze the VPs in the corresponding passive sentences in (43):

- (43) a. [_S The carcass was [_{VP[PAS]} devoured (by the man)]].
 b. [_S The sword was [_{VP[PAS]} handed to Tracy (by the man)]].

⁴See **WP92b** for a parallel rule for German which refers to accusative case on the left “=hand side of the metarule.

At first glance, this analysis may seem odd as an object is replaced inside the VP by a PP which would be the subject in an active clause. Although this analysis makes correct predictions with regard to the syntactic well-formedness of structures, it seems unclear how one can account for the semantic relations. It is possible, however, to use a lexical rule that licenses the passive participle and manipulates the semantics of the output lexical item in such a way that the *by*-PP is correctly integrated semantically (GKPS85a).

We arrive at a problem, however, if we try to apply this analysis to German since the impersonal passive cannot be derived by simply suppressing an object. The V2-rules for verbs such as *arbeiten* ‘work’ and *denken* ‘think’ as used for the analysis of (34a) and (35a) have the following form:

- (44) $V2 \rightarrow H[5]$
 $V2 \rightarrow H[13], PP[an]$

There is no NP on the right-hand side of these rules which could be turned into a *von*-PP. If the passive is to be analyzed as suppressing an NP argument in a rule, then it should follow from the existence of the impersonal passive that the passive metarule has to be applied to rules which license finite clauses, since information about whether there is a subject or not is only present in rules for finite clauses.⁵ In this kind of system, the rules for finite sentences (V3) are the basic rules and the rules for V2 would be derived from these.

It would only make sense to have a metarule which applies to V3 for German since English does not have V3 rules which contain both the subject and its object on the right-hand side of the rule.⁶ For English, it is assumed that a sentence consists of a subject and a VP (see GKPS85a). This means that we arrive at two very different analyses for the passive in English and German, which do not capture the descriptive insight that the passive is the suppression of the subject and the subsequent promotion of the object in the same way. The central difference between German and English seems to be that English obligatorily requires a subject,⁷ which is why English does not have an impersonal passive.

⁵GPSG differs from GB in that infinitive verbal projections do not contain nodes for empty subjects. This is also true for all other theories discussed in this book with the exception of Tree-Adjoining Grammar.

⁶GKPS85a suggest a metarule similar to our subject introduction metarule on page 202. The rule that is licensed by their metarule is used to analyze the position of auxiliaries in English and only licenses sequences of the form AUX NP VP. In such structures, subjects and objects are not in the same local tree either.

⁷Under certain conditions, the subject can also be omitted in English. For more on imperatives and other subject-less examples, see page ??.

This is a property independent of passives, which affects the possibility of having a passive structure, however.

The problem with the GPSG analysis is the fact that valence is encoded in phrase structure rules and that subjects are not present in the rules for verb phrases. In the following chapters, we will encounter approaches from LFG, Categorical Grammar, HPSG, Construction Grammar, and Dependency Grammar which encode valence separately from phrase structure rules and therefore do not have a principled problem with impersonal passive.

See **Jacobson87b** for more problematic aspects of the passive analysis in GPSG and for the insight that a lexical representation of valence – as assumed in Categorical Grammar, GB, LFG and HPSG – allows for a lexical analysis of the phenomenon, which is however unformulable in GPSG for principled reasons having to do with the fundamental assumptions regarding valence representations.

5.3 Verb position

Uszkoreit87a analyzed verb”=initial and verb”=final order as linearization variants of a flat tree. The details of this analysis have already been discussed in Section 5.1.2.

An alternative suggestion in a version of GPSG comes from **Jacobs86a**: Jacobs’s analysis is a rendering of the verb movement analysis in GB. He assumes that there is an empty verb in final position and links this to the verb in initial position using technical means which we will see in more detail in the following section.

5.4 Long”=distance dependencies as the result of local dependencies

One of the main innovations of GPSG is its treatment of long”=distance dependencies as a sequence of local dependencies (**Gazdar81**). This approach will be explained taking constituent fronting to the prefield in German as an example. Until now, we have only seen the GPSG analysis for verb-initial and verb-final position: the sequences in (45) are simply linearization variants.

- (45) a. [dass] der Mann der Frau das Buch gibt
that the man the woman the book gives
‘that the man gives the book to the woman’

5.4 Long”=distance dependencies as the result of local dependencies

- b. Gibt der Mann der Frau das Buch?
gives the man the woman the book
‘Does the man give the book to the woman?’

What we want is to derive the verb”=second order in the examples in (46) from V1 order in (45b).

- (46) a. Der Mann gibt der Frau das Buch.
the man gives the woman the book
‘The man gives the woman the book.’
b. Der Frau gibt der Mann das Buch.
the woman gives the man the book
‘The man gives the woman the book.’

For this, the metarule in (47) has to be used. This metarule removes an arbitrary category X from the set of categories on the right”=hand side of the rule and represents it on the left”=hand side with a slash (‘/’):⁸

- (47) $V3 \rightarrow W, X \mapsto$
 $V3/X \rightarrow W$

This rule creates the rules in (49) from (48):

- (48) $V3 \rightarrow H[8], N2[CASE\ dat], N2[CASE\ acc], N2[CASE\ nom]$

(49) $V3/N2[CASE\ nom] \rightarrow H[8], N2[CASE\ dat], N2[CASE\ acc]$
 $V3/N2[CASE\ dat] \rightarrow H[8], N2[CASE\ acc], N2[CASE\ nom]$
 $V3/N2[CASE\ acc] \rightarrow H[8], N2[CASE\ dat], N2[CASE\ nom]$

The rule in (50) connects a sentence with verb”=initial order with a constituent which is missing in the sentence:

- (50) $V3[+FIN] \rightarrow X[+TOP], V3[+MC]/X$

In (50), X stands for an arbitrary category which is marked as missing in V3 by the ‘/’. X is referred to as a *filler*.

The interesting cases of values for X with regard to our examples are given in (51):

⁸ An alternative to Uszkoreit’s trace”=less analysis (Uszkoreit87a), which is explained here, consists of using a trace for the extracted element as in GB.

- (51) $V3[+FIN] \rightarrow N2[+TOP, CASE\ nom], V3[+MC]/N2[CASE\ nom]$
 $V3[+FIN] \rightarrow N2[+TOP, CASE\ dat], V3[+MC]/N2[CASE\ dat]$
 $V3[+FIN] \rightarrow N2[+TOP, CASE\ acc], V3[+MC]/N2[CASE\ acc]$

(51) does not show actual rules. Instead, (51) shows examples for insertions of specific categories into the X"-position, that is, different instantiations of the rule.

The following linearization rule ensures that a constituent marked by [+TOP] in (50) precedes the rest of the sentence:

- (52) $[+TOP] < X$

TOP stands for *topicalized*. As was mentioned on page 117, the prefield is not restricted to topics. Focused elements and expletives can also occur in the prefield, which is why the feature name is not ideal. However, it is possible to replace it with something else, for instance *prefield*. This would not affect the analysis. X in (52) stands for an arbitrary category. This is a new X and it is independent from the one in (50).

Figure 5.4 shows the interaction of the rules for the analysis of (53).⁹

- (53) Dem Mann gibt er das Buch.
the.DAT man gives he.NOM the.ACC book
‘He gives the man the book.’

Figure 5.4: Analysis of fronting in GPSG

The metarule in (47) licenses a rule which adds a dative object into slash. This rule now licenses the subtree for *gibt er das Buch* ‘gives he the book’. The linearization rule $V[+MC] < X$ orders the verb to the very left inside of the local tree for V3. In the next step, the constituent following the slash is bound off. Following the LP"-rule $[+TOP] < X$, the bound constituent must be ordered to the left of the V3 node.

The analysis given in Figure 5.4 may seem too complex since the noun phrases in (53) all depend on the same verb. It is possible to invent a system of linearization rules which would allow one to analyze (53) with an entirely flat structure. One would nevertheless still need an analysis for sentences such as those in (37) on page 118 – repeated here as (54) for convenience:

⁹The FIN feature has been omitted on some of the nodes since it is redundant: +MC"-verbs always require the FIN value ‘+’.

5.4 Long”=distance dependencies as the result of local dependencies

- (54) a. [Um zwei Millionen Mark]_i soll er versucht haben,
 around two million Deutsche.Marks should he tried have
 [eine Versicherung _{-i} zu betrügen].¹⁰
 an insurance.company to deceive
 ‘He apparently tried to cheat an insurance company out of two million Deutsche Marks.’
- b. ”Wer_i glaubt er, daß er _{-i} ist?” erregte sich ein Politiker vom
 who believes he that he is retort REFL a politician from.the
 Nil.¹¹
 Nile
 ‘“Who does he think he is?”, a politician from the Nile exclaimed.’
- c. Wen_i glaubst du, daß ich _{-i} gesehen habe?¹²
 who believe you that I seen have
 ‘Who do you think I saw?’
- d. [Gegen ihn]_i falle es den Republikanern hingegen schwerer,
 against him fall it the Republicans however more.difficult
 [[Angriffe _{-i}] zu lancieren].¹³
 attacks to launch
 ‘It is, however, more difficult for the Republicans to launch attacks against him.’

The sentences in (54) cannot be explained by local reordering as the elements in the prefield are not dependent on the highest verb, but instead originate in the lower clause. Since only elements from the same local tree can be reordered, the sentences in (54) cannot be analyzed without postulating some kind of additional mechanism for long”=distance dependencies.¹⁴

Before I conclude this chapter, I will discuss yet another example of fronting, namely one of the more complex examples in (54). The analysis of (54c) consists

¹⁰ taz, 04.05.2001, p. 20.

¹¹ Spiegel, 8/1999, p. 18.

¹² Scherpenisse86a.

¹³ taz, 08.02.2008, p. 9.

¹⁴ One could imagine analyses that assume the special mechanism for nonlocal dependencies only for sentences that really involve dependencies that are nonlocal. This was done in HPSG by Kathol195a and Wetta2011a and by GO2009a in Dependency Grammar. I discuss the Dependency Grammar analyses in detail in Section ?? and show that analyses that treat simple V2 sentences as ordering variants of non-V2 sentences have problems with the scope of fronted adjuncts, with coordination of simple sentences and sentences with nonlocal dependencies and with so-called multiple frontings.

of several steps: the introduction, percolation and finally binding off of information about the long”=distance dependency. This is shown in Figure 5.5. Simpli-

Figure 5.5: Analysis of long”=distance dependencies in GPSG

fying somewhat, I assume that *gesehen habe* ‘have seen’ behaves like a normal transitive verb.¹⁵ A phrase structure rule licensed by the metarule in (47) licenses the combination of *ich* ‘I’ and *gesehen habe* ‘has seen’ and represents the missing accusative object on the V3 node. The complementizer *dass* ‘that’ is combined with *ich gesehen habe* ‘I have seen’ and the information about the fact that an accusative NP is missing is percolated up the tree. This percolation is controlled by the so-called *Foot Feature Principle*, which states that all foot features of all the daughters are also present on the mother node. Since the SLASH feature is a foot feature, the categories following the ‘/’ percolate up the tree if they are not bound off in the local tree. In the final step, the V3/N2[acc] is combined with the missing N2[acc]. The result is a complete finite declarative clause of the highest projection level.

5.5 Summary and classification

Some twenty years after Chomsky's criticism of phrase structure grammars, the first large grammar fragment in the GPSG framework appeared and offered analyses of phenomena which could not be described by simple phrase structure rules. Although works in GPSG essentially build on Harman's 1963 idea of a transformationless grammar, they also go far beyond this. A special achievement of GPSG is, in particular, the treatment of long"=distance dependencies as worked out by **Gazdar81**. By using the SLASH"=mechanism, it was possible to explain the simultaneous extraction of elements from conjuncts (Across the Board Extraction, **Ross67a**). The following examples from **Gazdar81** show that gaps in conjuncts must be identical, that is, a filler of a certain category must correspond to a gap in every conjunct:

- (55) a. The kennel which Mary made and Fido sleeps in has been stolen.
(= S/NP & S/NP)
- b. The kennel in which Mary keeps drugs and Fido sleeps has been stolen.
(= S/PP & S/PP)

¹⁵See Nerbonne86a and Johnson86a, for analyses of verbal complexes in GPSG.

- c. * The kennel (in) which Mary made and Fido sleeps has been stolen.
(= S/NP & S/PP)

GPSG can plausibly handle this with mechanisms for the transmission of information about gaps. In symmetric coordination, the SLASH elements in each conjunct have to be identical. On the one hand, a transformational approach is not straightforwardly possible since one normally assumes in such analyses that there is a tree and something is moved to another position in the tree thereby leaving a trace. However, in coordinate structures, the filler would correspond to two or more traces and it cannot be explained how the filler could originate in more than one place.

While the analysis of Across the Board extraction is a true highlight of GPSG, there are some problematic aspects that I want to address in the following: the interaction between valence and morphology, the representation of valence and partial verb phrase fronting, and the expressive power of the GPSG formalism.

5.5.1 Valence and morphology

The encoding of valence in GPSG is problematic for several reasons. For example, morphological processes take into account the valence properties of words. Adjectival derivation with the suffix ”*bar* ‘-able’ is only productive with transitive verbs, that is, with verbs with an accusative object which can undergo passivization:

- | | | | |
|------|----|--------------------------------------|-----------------------------------|
| (56) | a. | lös- <i>bar</i>
solv-able | (nominative, accusative) |
| | b. | vergleich- <i>bar</i>
compar-able | (nominative, accusative, PP[mit]) |
| | c. | * schlaf- <i>bar</i>
sleep-able | (nominative) |
| | d. | * helf- <i>bar</i>
help-able | (nominative, dative) |

A rule for derivations with *-bar* ‘-able’ must therefore make reference to valence information. This is not possible in GPSG grammars since every lexical entry is only assigned a number which says something about the rules in which this entry can be used. For ”*bar*”-derivations, one would have to list in the derivational rule all the numbers which correspond to rules with accusative objects, which of course does not adequately describe the phenomenon. Furthermore, the valence of the resulting adjective also depends on the valence of the verb. For example,

a verb such as *vergleichen* ‘compare’ requires a *mit* (with)-PP and *vergleichbar* ‘comparable’ does too (Riehemann 193a; Riehemann 198a). In the following chapters, we will encounter models which assume that lexical entries contain information as to whether a verb selects for an accusative object or not. In such models, morphological rules which need to access the valence properties of linguistic objects can be adequately formulated.

The issue of interaction of valence and derivational morphology will be taken up in Section ?? again, where approaches in LFG and Construction Grammar are discussed that share assumptions about the encoding of valence with GPSG.

5.5.2 Valence and partial verb phrase fronting

Nerbonne 1986a and Johnson 1986a investigate fronting of partial VPs in the GPSG framework. (57) gives some examples: in (57a) the bare verb is fronted and its arguments are realized in the middle field, in (57b) one of the objects is fronted together with the verb and in (57c) both objects are fronted with the verb.

- (57) a. Erzählen wird er seiner Tochter ein Märchen können.
 tell will he his daughter a fairy.tale can
 b. Ein Märchen erzählen wird er seiner Tochter können.
 a fairy.tale tell will he his daughter can
 c. Seiner Tochter ein Märchen erzählen wird er können.
 his daughter a fairy.tale tell will he can
 ‘He will be able to tell his daughter a fairy tale.’

The problem with sentences such as those in (57) is that the valence requirements of the verb *erzählen* ‘to tell’ are realized in various positions in the sentence. For fronted constituents, one requires a rule which allows a ditransitive to be realized without its arguments or with one or two objects. Furthermore, it has to be ensured that the arguments that are missing in the prefield are realized in the remainder of the clause. It is not legitimate to omit obligatory arguments or realize arguments with other properties like a different case, as the examples in (58) show:

- (58) a. Verschlungen hat er es nicht.
 devoured has he.NOM it.ACC not
 ‘He did not devour it.’
 b. *Verschlungen hat er nicht.
 devoured has he.NOM not

- c. * Verschlungen hat er ihm nicht.
 devoured has he.NOM him.DAT not

The obvious generalization is that the fronted and unfronted arguments must add up to the total set belonging to the verb. This is scarcely possible with the rule-based valence representation in GPSG. In theories such as Categorical Grammar (see Chapter ??), it is possible to formulate elegant analyses of (58) (Geach70a). Nerbonne and Johnson both suggest analyses for sentences such as (58) which ultimately amount to changing the representation of valence information in the direction of Categorical Grammar.

Before I turn to the expressive power of the GPSG formalism, I want to note that the problems that we discussed in the previous subsections are both related to the representation of valence in GPSG. We already run into valence"-related problems when discussing the passive in Section 5.2: since subjects and objects are introduced in phrase structure rules and since there are some languages in which subject and object are not in the same local tree, there seems to be no way to describe the passive as the suppression of the subject in GPSG.

5.5.3 Generative capacity

In GPSG, the system of linearization, dominance and metarules is normally restricted by conditions we will not discuss here in such a way that one could create a phrase structure grammar of the kind we saw in Chapter 2 from the specification of a GPSG grammar. Such grammars are also called context"-free grammars. In the mid-80s, it was shown that context"-free grammars are not able to describe natural language in general, that is it could be shown that there are languages that need more powerful grammar formalisms than context"-free grammars (Shieber85a; Culy85a; see Pullum86a for a historical overview). The so-called *generative capacity* of grammar formalisms is discussed in Chapter ??.

Following the emergence of constraint"-based models such as HPSG (see Chapter ??) and unification-based variants of Categorical Grammar (see Chapter ?? and Uszkoreit86d), most authors previously working in GPSG turned to other frameworks. The GPSG analysis of long"-distance dependencies and the distinction between immediate dominance and linear precedence are still used in HPSG and variants of Construction Grammar to this day. See also Section ?? for a Tree Adjoining Grammar variant that separates dominance from precedence.



Comprehension questions

1. What does it mean for a grammar to be in an ID/LP format?
2. How are linear variants of constituents in the middle field handled by GPSG?
3. Think of some phenomena which have been described by transformations and consider how GPSG has analyzed these data using other means.



Exercises

1. Write a small GPSG grammar which can analyze the following sentences:

(59) a. [dass] der Mann ihn liest
that the.NOM man him.ACC reads
'that the man reads it'
b. [dass] ihn der Mann liest
that him.ACC the.NOM man reads
'that the man reads it'

- c. Der Mann liest ihn.
the.NOM man reads him.ACC
'The man reads it.'

Include all arguments in a single rule without using the metarule for introducing subjects.



Further reading

The main publication in GPSG is **GKPS85a**. This book has been critically discussed by **Jacobson87b**. Some problematic analyses are contrasted with alternatives from Categorical Grammar and reference is made to the heavily Categorical Grammar influenced work of **Pollard84a-u**, which counts as one of the predecessors of HPSG. Some of Jacobson's suggestions can be found in later works in HPSG.

Grammars of German can be found in **Uszkoreit87a** and **Busemann92a-u**. **Gazdar81** developed an analysis of long-distance dependencies, which is still used today in theories such as HPSG.

A history of the genesis of GPSG can be found in **Pullum89a**.

6 Feature descriptions

In the previous chapter, we talked about sets of feature”=value pairs, which can be used to describe linguistic objects. In this chapter, we will introduce feature descriptions which play a role in theories such as LFG, HPSG, Construction Grammar, versions of Categorical Grammar and TAG (and even some formalizations of Minimalist theories (Veenstra98a)). This chapter will therefore lay some of the groundwork for the chapters to follow.

Feature structures are complex entities which can model properties of a linguistic object. Linguists mostly work with feature descriptions which describe only parts of a given feature structure. The difference between models and descriptions will be explained in more detail in Section ??.

Alternative terms for feature structures are:

- feature”=value structure
- attribute”=value structure

Other terms for feature description are the following:

- *attribute-value matrix* (AVM)
- *feature matrix*

In what follows, I will restrict the discussion to the absolutely necessary details in order to keep the formal part of the book as short as possible. I refer the interested reader to Shieber86a, ps, Johnson88, Carpenter92a, King94a and Richter2004a-u. Shieber’s book is an accessible introduction to Unification Grammars. The works by King and Richter, which introduce important foundations for HPSG, would most probably not be accessible for those without a good grounding in mathematics. However, it is important to know that these works exist and that the corresponding linguistic theory is build on a solid foundation.

6.1 Feature descriptions

When describing linguistic signs, we have to say something about their properties. For a noun, we can say that it has case, gender, number and person fea-

6 Feature descriptions

tures. For a word such as *Mannes* ‘man’, we can say that these features have the values *genitive*, *masculine*, *singular* and 3. If we were to write these as a list of feature”=value pairs, we would arrive at the following feature description:

- (1) Feature-value pair for *Mannes*:

CASE	<i>genitive</i>
GENDER	<i>masculine</i>
NUMBER	<i>singular</i>
PERSON	3

It is possible to describe a variety of different things using feature descriptions. For example, we can describe a person as in (2):

(2)

FIRSTNAME	<i>max</i>
LASTNAME	<i>meier</i>
DATE-OF-BIRTH	<i>10.10.1985</i>

People are related to other people – a fact that can also be expressed in feature”=value pairs. For example, the fact that Max Meier has a father called Peter Meier can be captured by expanding (2) as follows:

(3)

FIRSTNAME	<i>max</i>										
LASTNAME	<i>meier</i>										
DATE-OF-BIRTH	<i>10.10.1985</i>										
FATHER	<table> <tr> <td>FIRSTNAME</td><td><i>peter</i></td></tr> <tr> <td>LASTNAME</td><td><i>meier</i></td></tr> <tr> <td>DATE-OF-BIRTH</td><td><i>10.05.1960</i></td></tr> <tr> <td>FATHER</td><td>...</td></tr> <tr> <td>MOTHER</td><td>...</td></tr> </table>	FIRSTNAME	<i>peter</i>	LASTNAME	<i>meier</i>	DATE-OF-BIRTH	<i>10.05.1960</i>	FATHER	...	MOTHER	...
FIRSTNAME	<i>peter</i>										
LASTNAME	<i>meier</i>										
DATE-OF-BIRTH	<i>10.05.1960</i>										
FATHER	...										
MOTHER	...										
MOTHER	...										

The value of the FATHER feature is another feature description containing the same features as (2).

In feature descriptions, a *path* is a sequence of features which immediately follow each other. The *value of a path* is the feature description at the end of the path. Therefore, the value of FATHER|DATE-OF-BIRTH is *10.05.1960*.

One can think of many different features that could be included in representations such as (3). One may wonder how to integrate information about offspring into (3).

An obvious solution would be to add features for DAUGHTER and SON:

$$(4) \left[\begin{array}{ll} \text{FIRSTNAME} & \text{max} \\ \text{LASTNAME} & \text{meier} \\ \text{DATE-OF-BIRTH} & 10.10.1985 \\ \text{FATHER} & \dots \\ \text{MOTHER} & \dots \\ \text{DAUGHTER} & \dots \end{array} \right]$$

This solution is not satisfactory as it is not immediately clear how one could describe a person with several daughters. Should one really introduce features such as DAUGHTER-1 or DAUGHTER-3?

$$(5) \left[\begin{array}{ll} \text{FIRSTNAME} & \text{max} \\ \text{LASTNAME} & \text{meier} \\ \text{DATE-OF-BIRTH} & 10.10.1985 \\ \text{FATHER} & \dots \\ \text{MOTHER} & \dots \\ \text{DAUGHTER-1} & \dots \\ \text{DAUGHTER-2} & \dots \\ \text{DAUGHTER-3} & \dots \end{array} \right]$$

How many features do we want to assume? Where is the limit? What would the value of DAUGHTER=32 be?

For this case, it makes much more sense to use a list. Lists are indicated with angle brackets. Any number of elements can occur between these brackets. A special case is when no element occurs between the brackets. A list with no elements is also called *empty list*. In the following example, Max Meier has a daughter called Clara, who herself has no daughter.

$$(6) \left[\begin{array}{ll} \text{FIRSTNAME} & \text{max} \\ \text{LASTNAME} & \text{meier} \\ \text{DATE-OF-BIRTH} & 10.10.1985 \\ \text{FATHER} & \dots \\ \text{MOTHER} & \dots \\ \text{DAUGHTER} & \left\langle \begin{array}{ll} \text{FIRSTNAME} & \text{clara} \\ \text{LASTNAME} & \text{meier} \\ \text{DATE-OF-BIRTH} & 10.10.2004 \\ \text{FATHER} & \dots \\ \text{MOTHER} & \dots \\ \text{DAUGHTER} & \langle \rangle \end{array} \right\rangle \end{array} \right]$$

(7)

FIRSTNAME	<i>max</i>														
LASTNAME	<i>meier</i>														
DATE-OF-BIRTH	<i>10.10.1985</i>														
GENDER	<i>male</i>														
FATHER	...														
MOTHER	...														
CHILDREN	<table> <tr> <td>FIRSTNAME</td><td><i>clara</i></td></tr> <tr> <td>LASTNAME</td><td><i>meier</i></td></tr> <tr> <td>DATE-OF-BIRTH</td><td><i>10.10.2004</i></td></tr> <tr> <td>GENDER</td><td><i>female</i></td></tr> <tr> <td>FATHER</td><td>...</td></tr> <tr> <td>MOTHER</td><td>...</td></tr> <tr> <td>CHILDREN</td><td>⟨⟩</td></tr> </table>	FIRSTNAME	<i>clara</i>	LASTNAME	<i>meier</i>	DATE-OF-BIRTH	<i>10.10.2004</i>	GENDER	<i>female</i>	FATHER	...	MOTHER	...	CHILDREN	⟨⟩
FIRSTNAME	<i>clara</i>														
LASTNAME	<i>meier</i>														
DATE-OF-BIRTH	<i>10.10.2004</i>														
GENDER	<i>female</i>														
FATHER	...														
MOTHER	...														
CHILDREN	⟨⟩														

Now, we are left with the question of sons. Should we add another list for sons? Do we want to differentiate between sons and daughters? It is certainly the case that the gender of the children is an important property, but these are properties of the objects themselves, since every person has a gender. The description in (7) therefore offers a more adequate representation.

At this point, one could ask why the parents are not included in a list as well. In fact, we find similar questions also in linguistic works: how is information best organized for the job at hand? One could argue for the representation of descriptions of the parents under separate features, by pointing out that with such a representation it is possible to make certain claims about a mother or father without having to necessarily search for the respective descriptions in a list.

If the order of the elements is irrelevant, then we could use sets rather than lists. Sets are written inside curly brackets.¹

6.2 Types

In the previous section, we introduced feature descriptions consisting of feature= value pairs and showed that it makes sense to allow for complex values for features. In this section, feature descriptions will be augmented to include

¹The definition of a set requires many technicalities. In this book, I would use sets only for collecting semantic information. This can be done equally well using lists, which is why I do not introduce sets here and instead use lists.

types. Feature descriptions which are assigned a type are also called *typed feature descriptions*. Types say something about which features can or must belong to a particular structure. The description previously discussed describes an object of the type *person*.

$$(8) \left[\begin{array}{ll} \textit{person} & \\ \text{FIRSTNAME} & \textit{max} \\ \text{LASTNAME} & \textit{meier} \\ \text{DATE-OF-BIRTH} & 10.10.1985 \\ \text{GENDER} & \textit{male} \\ \text{FATHER} & \dots \\ \text{MOTHER} & \dots \\ \text{CHILDREN} & \langle \dots, \dots \rangle \end{array} \right]$$

Types are written in *italics*.

The specification of a type determines which properties a modelled object has. It is then only possible for a theory to say something about these properties. Properties such as OPERATING VOLTAGE are not relevant for objects of the type *person*. If we know the type of a given object, then we also know that this object must have certain properties even if we do not yet know their exact values. In this way, (9) is still a description of Max Meier even though it does not contain any information about Max' date of birth:

$$(9) \left[\begin{array}{ll} \textit{person} & \\ \text{FIRSTNAME} & \textit{max} \\ \text{LASTNAME} & \textit{meier} \\ \text{GENDER} & \textit{male} \end{array} \right]$$

We know, however, that Max Meier must have been born on some day since this is a description of the type *person*. The question *What is Max' date of birth?* makes sense for a structure such as (9) in a way that the question *Which operating voltage does Max have?* does not. If we know that an object is of the type *person*, then we have the following basic structure: