

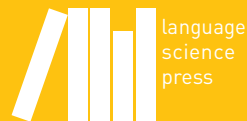
语法理论

从转换语法到基于约束的理论

Stefan Müller

DRAFT
of March 20, 2017, 16:47

Textbooks in Language Sciences ??



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.

语法理论

从转换语法到基于约束的理论

Stefan Müller

DRAFT
of March 20, 2017, 16:47

Stefan Müller. 2017. 语法理论: 从转换语法到基于约束的理论 (Textbooks in Language Sciences ??). Berlin: Language Science Press.

This title can be downloaded at:

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

© 2017, Stefan Müller

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

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

ISBN: Digital, complete work: ??;

Hardcover: vol1: ??; vol. 2 ??

Softcover: vol1: ??; vol. 2 ??

Softcover US: vol1: ??; vol. 2 ??

ISSN: 2364-6209

Cover and concept of design: Ulrike Harbort

Translators: 王璐璐、孙薇薇、黄思思

Typesetting: Stefan Müller

Fonts: Linux Libertine, Arimo, DejaVu Sans Mono

Typesetting software: Xe_{La}TeX

Language Science Press

Habelschwerdter Allee 45

14195 Berlin, Germany

langsci-press.org

Storage and cataloguing done by FU Berlin



Language Science Press has no responsibility for the persistence or accuracy of URLs for external or third-party Internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

献给 Max

目 录

前言	xv
一 研究背景与语法理论	1
1 导言与术语	3
1.1 为什么要研究句法?	3
1.2 为什么要形式化?	5
1.3 构成成分	6
1.3.1 组成成分测试	7
1.3.2 关于成分测试法的一些看法	10
1.4 词类	16
1.5 中心语	26
1.6 论元成分与说明语	28
1.7 语法功能	32
1.7.1 主语	32
1.7.2 状语	36
1.7.3 谓语	38
1.7.4 配价类型	39
1.8 德语小句的空间位置模型	40
1.8.1 动词的位置	40
1.8.2 句子的框架结构、前场、中场及后场	40
1.8.3 场内元素的排列	41
1.8.4 递归	45
2 短语结构语法	49
2.1 符号与重写规则	49
2.2 短语结构语法中特征的运用	54
2.3 语义	55
2.4 部分德语句法的短语结构规则	58
2.4.1 名词短语	58
2.4.2 介词短语	66

目 录

2.5	\bar{X} 理论	68
3	转换语法-管辖与约束理论	77
3.1	关于表示形式的一般说明	78
3.1.1	转换	78
3.1.2	有关语言习得的假说: 原则与参数理论	80
3.1.3	T 模型	81
3.1.4	\bar{X} 理论	86
3.1.5	英语中的 CP 和 IP	89
3.1.6	德语小句的结构	91
3.2	动词位置	92
3.3	长距离依存	96
3.4	被动	99
3.4.1	结构格与词汇格	99
3.4.2	格指派与格过滤	101
3.5	局部重新排序	104
3.6	总结	106
3.6.1	有关语言习得的解释	107
3.6.2	形式化	107
4	转换语法-最简方案	115
4.1	表示形式的一般说明	115
4.1.1	基本框架	116
4.1.2	配价、特征核对与一致关系	117
4.1.3	短语结构与 \bar{X} 理论	118
4.1.4	小 v	119
4.1.5	CP、TP、vP 和 VP	121
4.1.6	说明语	125
4.2	动词位置	125
4.3	长距离依存	126
4.4	被动	127
4.5	局部重新排序	129
4.6	新的发展与理论变体	130
4.6.1	移位、合并、特征驱动的移位与功能投射	131
4.6.2	标记	139
4.6.3	限定语、补足语和 \bar{X} 理论的剩余部分	144
4.6.4	最简方案、范畴语法和 HPSG 理论	148
4.6.5	原子特征的选择与复杂范畴的选择	155
4.6.6	小结	157
4.7	总结	157
4.7.1	关于语言习得的解释	157
4.7.2	形式化	158

5	广义短语结构语法	161
5.1	表示形式概说	161
5.1.1	复杂范畴、中心语特征原则, 以及 \bar{X} 规则	161
5.1.2	局部重新排序	164
5.1.3	元规则	166
5.1.4	语义	167
5.1.5	状语	169
5.2	作为元规则的被动	170
5.3	动词位置	173
5.4	作为局部依存结果的长距离依存	173
5.5	总结	176
5.5.1	定价与形态	177
5.5.2	定价与部分动词短语前置	178
5.5.3	普遍能力	179
6	特征描写	181
6.1	特征描写	181
6.2	类型	183
6.3	析取	186
6.4	结构共享	187
6.5	循环结构	189
6.6	合一	189
6.7	现象、模型和形式化理论	191
7	Lexical Functional Grammar	195
7.1	General remarks on the representational format	198
7.1.1	Functional structure	198
7.1.2	Completeness	201
7.1.3	Coherence	201
7.1.4	Restrictions on the c-structure/f-structure relation	202
7.1.5	Semantics	203
7.1.6	Adjuncts	205
7.2	Passive	206
7.3	Verb position	209
7.4	Local reordering	210
7.5	Long-distance dependencies and functional uncertainty	213
7.6	Summary and classification	216
8	Categorial Grammar	219
8.1	General remarks on the representational format	220
8.1.1	Representation of valence information	220
8.1.2	Semantics	222
8.1.3	Adjuncts	223
8.2	Passive	224

目 录

8.3	Verb position	225
8.4	Local reordering	227
8.5	Long-distance dependencies	228
8.5.1	Type Raising	228
8.5.2	Forward and backward composition	229
8.5.3	Analysis of long-distance dependencies	229
8.6	Summary and classification	231
9	中心语驱动的短语结构语法	237
9.1	有关表示形式的一般说明	240
9.1.1	配价信息的表示	240
9.1.2	组成成分结构的表示	242
9.1.3	线性顺序规则	246
9.1.4	中心语属性的投射	247
9.1.5	承继层级体系与概说	248
9.1.6	语义	251
9.1.7	说明语	253
9.2	被动	256
9.2.1	配价信息与格原则	259
9.3	动词位置	260
9.4	局部重新排序	265
9.5	长距离依存	267
9.6	新的进展与理论变体	272
9.6.1	限定语、补足语与论元结构	272
9.6.2	基于线性顺序的 HPSG 理论	275
9.7	总结	275
10	Construction Grammar	281
10.1	General remarks on the representational format	283
10.1.1	The head-complement construction	284
10.1.2	Representation of valence information	285
10.1.3	Semantics	285
10.1.4	Adjuncts	285
10.2	Passive	285
10.3	Verb position	290
10.4	Local reordering	292
10.5	Long-distance dependencies	293
10.6	New developments and theoretical variants	293
10.6.1	Berkeley Construction Grammar	293
10.6.2	Sign-Based Construction Grammar	297
10.6.3	Embodied Construction Grammar	309
10.6.4	Fluid Construction Grammar	313
10.6.5	Coordination	321
10.6.6	Discontinuous constituents and performance models	322

10.6.7	Discontinuity vs. Subject-Head and Head-Filler Schema	323
10.6.8	Restricting discontinuity	323
10.7	Summary and classification	333
11	依存语法	335
11.1	表示形式概述	337
11.1.1	价、结与卫星结构	337
11.1.2	说明语	339
11.1.3	线性	340
11.1.4	语义	341
11.2	被动	342
11.3	动词位置	344
11.4	局部重新排序	345
11.5	长距离依存	346
11.6	新的发展与理论变体	350
11.6.1	Tesnière 的词类划分	351
11.6.2	联系、联结与转用	352
11.6.3	辖域	356
11.7	总结	359
11.7.1	线性化	359
11.7.2	依存语法与短语结构语法	365
12	Tree Adjoining Grammar	377
12.1	General remarks on representational format	378
12.1.1	Representation of valence information	378
12.1.2	Substitution	378
12.1.3	Adjunction	379
12.1.4	Semantics	380
12.2	Local reordering	382
12.3	Verb position	385
12.4	Passive	387
12.5	Long-distance dependencies	388
12.6	New developments and theoretical variants	388
12.6.1	FTAG	389
12.6.2	V-TAG	392
12.6.3	The competence-performance distinction and the generative capacity of tree-local MC-LTAG	394
12.7	Summary and classification	400

二 讨论	403■
13 语言知识的天赋性	405
13.1 句法普遍性	407
13.1.1 中心语导向参数	407
13.1.2 \bar{X} 结构	409
13.1.3 主语和宾语的语法功能	411
13.1.4 约束原则	413
13.1.5 长距离依存的属性	414
13.1.6 表示时态、情态和体的语法语素	422
13.1.7 词类	423
13.1.8 递归与无限	424
13.1.9 小结	428
13.2 语言习得的速度	429
13.3 习得的关键期	429
13.4 非人类的灵长动物的习得的缺乏	430
13.5 克里奥尔语和手语	431
13.6 大脑部位的定位	432
13.7 语言跟一般认知的区别	433
13.7.1 威廉综合症	433
13.7.2 带有 FoxP2 基因突变的 KE 家族	433
13.8 刺激贫乏	434
13.8.1 Gold 定理	434
13.8.2 四个案例	437
13.8.3 无指导的数据驱动的剖析 (U-DOP)	448
13.8.4 负向证据	456
13.9 总结	457
14 Generative-enumerative vs. model-theoretic approaches	459
14.1 Graded acceptability	461
14.2 Utterance fragments	462
14.3 A problem for model-theoretic approaches?	464
15 The competence/performance distinction	469
15.1 The derivational theory of complexity	471
15.2 Incremental processing	475
16 Language acquisition	483
16.1 Principles & Parameters	483
16.2 Principles and the lexicon	491
16.3 Pattern-based approaches	491
16.4 Selection-based approaches	498
16.5 Summary	499

17	生成能力与语言形式描述	501
18	Binary branching, locality, and recursion	505
18.1	Binary branching	505
18.2	Locality	508
18.3	Recursion	517
19	Empty elements	519
19.1	Views on empty elements	519
19.2	Eliminating empty elements from grammars	521
19.3	Empty elements and semantic interpretation	527
19.4	Evidence for empty elements	533
19.5	Transformations, lexical rules, and empty elements	533
20	Extraction, scrambling, and passive: one or several descriptive devices?	539
21	Phrasal vs. lexical analyses	549
21.1	Some putative advantages of phrasal models	550
21.1.1	Usage-based theories	550
21.1.2	Coercion	552
21.1.3	Aspect as a clause level phenomenon	554
21.1.4	Simplicity and polysemy	555
21.2	Evidence for lexical approaches	557
21.2.1	Valence and coordination	557
21.2.2	Valence and derivational morphology	562
21.3	Radical underspecification: the end of argument structure?	568
21.3.1	Neo-Davidsonianism	568
21.3.2	Little <i>v</i> and idiom asymmetries	569
21.3.3	Deverbal nominals	572
21.3.4	Idiosyncratic syntactic selections	574
21.3.5	Expletives	577
21.3.6	An exoskeletal approach	577
21.3.7	Is there an alternative to lexical valence structure?	582
21.3.8	Summary	583
21.4	Relations between constructions	583
21.4.1	Inheritance hierarchies for constructions	583
21.4.2	Mappings between different levels of representations	588
21.4.3	Is there an alternative to lexical rules?	590
21.5	Further problems for phrasal approaches	591
21.5.1	Particle verbs and commitment to phrase structure configurations	591
21.6	Arguments from language acquisition	592
21.6.1	Recognizability of constructions	592
21.6.2	Coordination and discontinuousness	594

目 录

21.7	Arguments from psycho- and neurolinguistics	599
21.7.1	Lexical rules vs. phrasal constructions	599
21.7.2	Light verbs	601
21.7.3	Arguments from neurolinguistics	602
21.8	Arguments from statistical distribution	606
21.8.1	Unsupervised Data-Oriented Parsing	606
21.8.2	Collostructions	608
21.9	Conclusion	609
21.10	Why (phrasal) constructions?	610
21.10.1	Verbless directives	610
21.10.2	Serial verbs	617
21.10.3	Relative and interrogative clauses	618
21.10.4	The N-P-N construction	620
22	Universal Grammar and comparative linguistics without UG	623
22.1	Formal tools for capturing generalizations	623
22.2	How to develop linguistic theories that capture cross-linguistic general- izations	628
23	结论	637
24	练习题答案	641
24.1	导言与术语	641
24.2	短语结构语法	642
24.3	转换语法—管辖与约束理论	645
24.4	广义短语结构语法	646
24.5	特征描写	648
24.6	词汇功能语法	650
24.7	范畴语法	651
24.8	中心语驱动的短语结构语法	652
24.9	构式语法	653
24.10	依存语法	654
24.11	树邻接语法	655
	参考文献	659
	索引	747
	人名索引	747
	语言索引	747
	术语索引	747
	中文术语索引	747

前言

本书是在我的德语书 *Grammatiktheorie*(Müller 2013b)的基础上进行扩展与修订的版本。它介绍了在当代理论学界发挥了重要作用，或者在过去做出了重要贡献的如今仍是主流的几种不同的语法理论。我对一些基础的假设进行了解释，并将这些不同的理论应用到德语的所谓“核心语法”之中。我决定采用在本书的德语版本中使用的目标语言，因为很多要分析的现象不能用英语当作目标语言来解释。而且，研究者已经开发了许多英语的理论，因为他们的母语是英语，并且应用这些理论到其他语言之中是很有启发性的。我将展示这些理论是如何处理论元、说明语、主动/被动转换，局部重新排序（所谓的置换）、动词位置以及长距离的短语前置（除了英语的日耳曼语族的语言中动词位于第二位的属性）这些现象的。

第二部分探讨理论发展中重要的基础问题。这包括我们是否具有具体语言的内在知识的讨论、人类处理语言的心理语言学的证据多讨论、空语素的地位的讨论，以及我们是整体性地还是组合性低构造和获取话语的问题，即我们是使用短语结构还是词汇结构呢。

不过，语言学是一个相当多的术语混乱不清的科学领域。所以，我在导言部分专门介绍了本书后面章节中将会运用到术语。第二章介绍短语结构语法，该语法在本书介绍的许多理论中都发挥了重要的作用。我在德语专业本科生的导论课中讲解这两章的内容（除了第 2.3 节有关短语结构语法和语义之间的关系）。高级读者可以略过这些导论性质的章节。后续的章节安排也适用于没有前期知识的读者来理解这些理论的基本内容。有关最新的理论发展的内容是更有挑战性的：这些内容会指向后续将要介绍的章节，也会指向在现今的理论讨论中相关的其它文献，这些文献不会在本书中重复表示或者进行总结。这部分内容可供高水平的学生与学者欣赏。我将这本书作为我们本科课程中高水平学生的研讨课上讲解理论的句法方面的教材。这些课件可在我的网页上下载。本书的第二部分更有挑战性，它包括了难点问题和研究文献的讨论内容。

本书只介绍相对较近的理论发展。对于历史文献的回顾，请看Robins (1997); Jungen & Lohnstein (2006)。本书并不包括整合语言学 (Lieb 1983; Eisenberg 2004; Nolda 2007)、优选论 (Prince & Smolensky 1993、Grimshaw 1997、G. Müller 2000)、角色和参数语法 (Van Valin 1993)以及关系语法 (Perlmutter 1983, 1984)的内容。我将这些内容留到以后的版本中。

德语书的最初版本只计划有四百页，但是最后的成书内容多多了：德语教材的第一版有 529 页，第二版有 564 页。我在英语版中加入了依存语法和英语的最简语法这两章内容，现在本书有 746 页。我尽最大努力将所选的理论表示清楚，并列

前言

出所有重要的文献。尽管参考文献的列表超过了 85 页，我也有可能没有列出全部的文献。我对此和其他问题表示歉意。

本书的版本

本书的官方版本是 PDF 文档，它可在本书在语言科学出版社的网页上获取¹。本页还链接到打印版。由于本书的内容非常多，我们决定将其分成两卷。第一卷包括所有的理论的描述，第二卷是讨论部分。这两卷都包括了参考文献和索引的完整列表。第二卷从第 405 页开始。由此，打印版与 PDF 文档中的部分是相同的。

致谢

我要感谢跟我讨论本书早期德语版的 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, 和 Stephen Wechsler 以及对本书早期德语版本提出评论的 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, 和 Eva Wittenberg 我还感谢在本书的早期版本中提出评论的 Thomas Groß, Dick Hudson, Sylvain Kahane, Paul Kay, Haitao Liu (刘海涛), Andrew McIntyre, Sebastian Nordhoff, Tim Osborne, Andreas Pankau, and Christoph Schwarze 感谢 Leonardo Boiko 和 Sven Verdoolaege 挑出了错别字。特别感谢 Martin Haspelmath 对本书英文版的早期版本提出的详细评论。

我还要感谢在此书的中文译本中给予我帮助的学者们，他们是刘海涛、刘晓。

本书是语言科学出版社出版的具有公开评论阶段的第一本书（详见下面）。我感谢 Dick Hudson, Paul Kay, Antonio Machicao y Priemer, Andrew McIntyre, Sebastian Nordhoff 和一位匿名评论者对本书提出的评论。这些评论纪录在本书的下载页面中。除此之外，本书还历经了公开校对的阶段（也请看下面）。有些校对者不仅做了校对的动作，还提出了具有高度价值的评论。我决定将这些评论作为附加的公开评论发布出来。特别是 Armin Buch、Leonel de Alencar、Andreas Hölzl、Gianina Iordăchioaia、Timm Lichte、Antonio Machicao y Priemer 和 Neal Whitman 需要在这里明示。

我感谢 Wolfgang Sternefeld 和 Frank Richter，他们对本书的德语版写了详尽的评论。他们指出了一些错误和疏漏之处，我们在德语的第二版中进行了改正，这些都在英语版中有所显示。

感谢所有对本书进行评论和提出改进意见的学生们。特别是 Lisa Deringer, Aleksandra Gabryszak, Simon Lohmiller, Theresa Kallenbach, Steffen Neuschulz, Reka Meszaros-Segner, Lena Terhart 和 Elodie Winckel 需要在这里明示。

¹ <http://langsci-press.org/catalog/book/??>

由于本书是建立在我在语言理论领域中的实践经验之上，我想感谢那些在会议、工作坊、暑期学校期间以及通过邮件跟我讨论过语言学的学者们。特别值得列出的有 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 和 Arne Zeschel

我感谢 Sebastian Nordhoff 针对循环 (recursion) 这一主题的评论。

Andrew Murphy 翻译了英文版中的第一章到第三章，第五章到第十章，以及第十二章到第二十三章的内容。特别感谢！

我还要感谢 27 位校对着，他们的工作对本书的改进提供了极大的帮助。我从他们每个人那里获得的意见比从出版商那里获得的意见还要多。有些意见是针对内容的，而不是错别字和格式的。没有一位受雇于出版商的校对人员能够发现这些错误与不一致的地方，因为出版商没有雇员懂得本书所覆盖的所有语法理论。

过去的几年中，举办了一些理论比较的工作坊。我受邀参加了其中的三个工作坊。我感谢 Helge Dyvik 和 Torbjørn Nordgård 邀请我参加 2005 年在卑尔根举办的挪威博士生秋季学校“对比中的语言与理论” (*Languages and Theories in Contrast*)。Guido Mensching 和 Elisabeth Stark 邀请我参加了 2007 年在柏林自由大学举办的“比较语言与比较理论：生成语法与构式语法” (*Comparing Languages and Comparing Theories: Generative Grammar and Construction Grammar*) 工作坊。Andreas Pankau 邀请我参加 2009 年在乌得勒支举办的“比较框架” (*Comparing Frameworks*) 工作坊。我在与参加这些活动的学者们的讨论中受益良多，本书也受益于这些交流。

我感谢 Peter Gallmann 在我在耶拿期间与他讨论的课件内容。第 3.1.3–3.4 节与他的版本很像，并参考了其中很多内容。感谢 David Reitte 提供的组合性范畴语法的 \LaTeX 宏包，Mary Dalrymple 和 Jonas Kuhn 提供的 LFG 宏包和示例结构，以及 Laura Kallmeyer 提供的大部分 TAG 分析中的 \LaTeX 资源。由于与 \XeLaTeX 的兼容性问题，大部分树都调整为 forest 包的格式，但是原始的树和文本都给予了我很多灵感，没有他们，相应章节中的图绝不会像现在这样好看。

我感谢 Sašo Živanović 实现了 \LaTeX 的宏包 forest。这个宏包简化了树、依存图和类型层级的格式。我还要感谢他在邮件和 stackexchange 上的具体帮助。当然，那些在 stackexchange 上活跃的人所提供的帮助仅有感谢是不够的：大部分我的有关本书格式的细节问题或者现在由 Language Science Press 使用的 \LaTeX 类型的应用都在几分钟内得到的回答。谢谢你们！因为本书是一本 CC-BY 版权下的公开图书，它也是一本公开资源的书。感兴趣的读者可以在 <https://github.com/langsci/25> 拷贝这些资源。通过将本书的资源公开，我将 \LaTeX 大师们提供的资源传递下去，并希望其他人能够从中获益，并且学会按照更好看和更高效多方式来编写他们的语言学论文。

我还要感谢 Viola Auermann、Antje Bahlke、Sarah Dietzfelbinger、Lea Helmers 和 Chiara Jancke 所做的大量的复印工作。Viola 还在英译本定稿前的最后阶段帮我校对。我还要感谢我的（前任）实验室成员们 Felix Bildhauer, Philippa Cook, Janna Lipenkova, Jakob Maché, Bjarne Ørsnes 和 Roland Schäfer。他们在教学中以及其他方面都给予我很多帮助。从 2007 年直到本书第一版德语版的教材出版的这些年中，

德语语言学的三个终身教职中有两个职位都是空缺的，如果没有他们的帮助，我是无法完成教学任务并完成这本书的。

在本书出版的过程中

我从 1994 年开始写我的毕业论文，并在 1997 年成功通过答辩。这一阶段的手稿可以在我的网页上获取。在答辩之后，我必须找到出版商。我很高兴收到了 Niemeyer 的“语言学研究”系列丛中的邀请，但是同时我对价格感到震惊不已，当时是每本书需要 186 德国马克，这还是在我没有出版商的任何帮助的情况下自己写书和排版（这个价格是纸版小说的二十倍）。² 这基本上意味着我的书是没有出版的：直到 1998 年，它在我的网站上才可以看到，并随后在图书馆才能找到。我的教授转正著作由 CSLI 出版社出版，价格也相对来说合理多了。当我开始写教科书的时候，我就开始寻找不同的出版渠道，并开始跟无名印刷需求的出版社协商。Brigitte Narr 负责 Stauffenburg 出版集团，她说服我在他们的出版社出版 HPSG 的教材。这本书的德语版属于我，这样我就可以在我的主页上出版。这一合作是成功的，由此我还可以跟 Stauffenburg 出版我的第二本关于语法理论的教科书。我想这本书具有更为广泛的相关性，并且可以供非德语的读者欣赏。由此，我决定将之翻译为英语。因为 Stauffenburg 重点出版德语书记，我必须找到另一家出版社。幸运的是，出版界的情况与 1997 年相比发生了戏剧性的翻天覆地的变化：我们现在有高水平的出版社，不仅有严格的同行评审，还有着完全公开的途径。我很高兴 Brigitte Narr 将本书的版权卖回给我，我现在就可以在 CC-BY 版权下由语言科学出版社出版这本英文版教材。

语言科学出版社：归学者所有的高质量语言学书籍

在 2012 年，有一群人发现出版界的情况令人难以容忍，他们一致认为有必要在公开平台上发布语言学书籍。也就是说，针对所有读者和作者公开的平台。我建立了一个网页，并征集了支持者，来自全世界各地的著名语言学家、Martin Haspelmath 和我随后就成立了语言科学出版社。几乎同时，DFG 公布了一项公开专著的项目，我们申请 (Müller & Haspelmath 2013) 并获得了资助 (18 个申请中只有两家获得了资助)。这笔钱支付给一位主任 (Dr. Sebastian Nordhoff)、一位经济学家 (Debora Siller)、两位程序员 (Carola Fanselow 和 Dr. Mathias Schenner)，他们在公开专著出版社 (OMP) 出版平台工作，并应用转换软件来从我们的 \LaTeX 编码中生成不同的格式 (ePub、XML、HTML)。Svantje Lilienthal 负责 OMP 的文档，制作屏幕录像，并为作者、读者和编辑提供用户支持。

OMP 在公开评论方面和社区建设的游戏化工具方面进行了扩展。所有语言科学出版社出版的图中都至少由两位外部审稿人审稿。审稿人和作者可以同意出版这些审稿意见，并使得整个过更为透明（也可以看 Pullum (1984) 关于期刊文章的公开评论的建议）。另外，还有可选的第二轮评审过程：公开评审。这一阶段所有人都都是公开的。整个社团都可以评论语言科学出版社出版的书籍。在第二轮评论

² 与此同时，Niemeyer 被 de Gruyter 收购，并关门大吉了。这本书的价格现在是 139.95 欧元 / 196.00 美元。欧元的价格相当于 273.72 德国马克。

阶段后，这通常需要持续两个月的时间，作者会进行修订，进而改进的版本会出版出来。这本书是经历了这个公开评论阶段的第一本书。本书的标注了公开评论意见的版本可以通过[web page of this book](#)获得。

如今，语言科学出版社拥有 17 个语言学不同领域的系列书籍，这些高水平的编辑来自各个大陆。我们有 18 本已经出版的书籍，还有 17 本即将出版的书籍，还有 146 本的作者表示出了极大的兴趣。系列编辑和作者主要负责用 \LaTeX 编辑的手稿，但是他们受到基于网络的由语言科学出版社建立的格式模版以及社区里的志愿者的支持。校对也是基于社区的。截至目前有 53 位学者帮助我们出版的图书。他们的工作被记录在名人堂中：<http://langsci-press.org/about/hallOfFame>。

如果你认为这类教科书应该对那些想阅读这些书籍的人免费获得，而且出版社不应该变成利益驱动的出版社，那么你就应该加入语言科学出版社社区，并且在以下几个方面支持我们：你可以在语言科学出版社上注册，并将你的名字列在其他将近 600 名热心学者之中，你可以用你的时间帮助校对或者修改格式，或者你可以为某本书活着语言科学出版社捐钱。我们也在寻找基金会、社团、语言学系和大学图书馆等机构的支持。如何帮助我们的详细信息列在下面的网页中：<http://langsci-press.org/about/support>。如有问题，请联系我或者语言学科学出版社的主任contact@langsci-press.org。

柏林, March 20, 2017

Stefan Müller

第一部分 研究背景与语法理论

第一章 导言与术语

本章旨在探讨两个问题：一是为什么要研究句法（第1.1节），二是为什么说形式化的表述是非常重要的（第1.2节）。我们将在第1.3节到第1.8节的内容中介绍基本概念：第1.3节讲解将语流切分成一个个小单位的准则。第1.4节展示了词汇聚类的规律，其中我将重点介绍动词或形容词的分类标准。第1.5节介绍中心语的概念，而有关论元和说明语的区别将在第1.6节中展开。第1.7节界定了语法功能。第1.8节利用空间位置理论来分析德语这类语言中小句的空间位置。

需要说明的是，语言学领域里术语纷杂，难以统一。造成这一现象的部分原因是，这些术语最早来源于对某些语言（如拉丁语，英语）的分析，而后又被用于描述其他语言现象。然而，这样做并不合适，因为有些语言与其他语言区别很大，而且语言本身也在不断地发展变化。基于以上原因，有些术语的用法发生了变化或者是有新的术语被创造出来。

本书在介绍新术语时，会援引相关的术语或者区分出每一条术语的不同用法，这样可以方便读者将这些术语与其他文献的术语联系起来。

1.1 为什么要研究句法？

所有的语言表达式都有意义。据此，我们研究形式与意义的结合体 (de Saussure 1916b)。例如，*tree*（树）这个词的字形与其相应的语音形式被赋予了 *tree'*（树）的含义。相对大的语音单位可以由相对小的单位构成：词与词可以一起构成词组并且这些词组可以继而构成句子。

接下来的问题是：我们是否需要一个形式化的系统来表述这些句子的结构呢？就像上面，我们将 *tree*（树）这个词的形式与意义对儿形式化地表述出来对于完成这些句子来说是不是不够的呢？

理论上来说，这只适用于那些由有限词语序列构成的语言。如果我们假设句子的最大长度是有限的，词长是有限的，词语的数量也是有限的，那么句子的数量也应该是有限的。但是，即使我们能限制句子的长度，可以构成的句子数量也是无限的。由此，我们真正需要回答的问题是：句子的最大长度是多少呢？举例来说，我们可以把下面的句子（1）进行扩展：

- (1) a. This sentence goes on and on and on and on ...
 这 句子 长 PREP 和 PREP 和 PREP 和 PREP ...
- b. [A sentence is a sentence] is a sentence.
 一 句子 COP 一 句子 COP 一 句子

1 导言与术语

- c. that Max thinks that Julius knows that Otto claims that Karl suspects
CONJ Max 认为 CONJ Julius 知道 CONJ Otto 声称 CONJ Carl 怀疑
that Richard confirms that Friederike is laughing
CONJ Richard 承认 CONJ Friederike AUX 笑

例(1b)是对“a sentence is a sentence”这组词的说明,即它是一个句子。我们也可以按照(1b)的方式来扩展句子,将这一整句话也看作是一句话。(1c)是将“Friederike is laughing”和“Richard confirms”组合起来构成了一个新句子“Richard confirms, that Friederike is laughing”。按照同样的方式,我们还可以将“Karl suspects”扩展进来。由此,我们就可以将一个不太复杂的句子嵌套进来构成一个非常复杂的句子。这样就可以一句句地嵌套下去。(1c)这类句子类似于俄罗斯套娃(matryoshka):每个娃娃都由一个套一个的不同颜色的更小的娃娃构成。(1c)中的句子也是以同样的方式构成的,即他们都由更小的成分,而且是不同的名词和动词组成。我们可以用方括号很清楚地表示出来,如下所示:

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

我们可以像例(1)一样通过扩展的方式构建出又长又复杂的句子。¹

对于这些组合来说,我们很难轻易地说截止到哪里是我们的语言可以接受的。(Harris 1957: 208; Chomsky 1957: 23)。同样,如果认为这些复杂的句子被当作一个复杂的单位储存在大脑中,这样的观点也是难以令人信服的。虽然神经语言学的实验显示,高频词语或固定搭配往往以复杂单位的形式储存在大脑中,但是这与例(1)所示的情况是不同的。再者,我们能够造出我们从未听过、说过或者写下的话语。这些话语一定是有结构的,一定会有可以一遍一遍重复的模式。作为人类,我们有能力将简单的成分构成复杂的句子,也可以将复杂的话语分解出他们的构成成分。目前,也有神经科学相关的研究可以证明,人类具有利用规则将词构成更大单位的能力。(Pulvermüller 2010: 170)

这些规则被违反时,反而更能证明我们是按照规则来组织语言材料的。儿童是通过他们能够接触到的语言输入来习得语言的。在这样做的时候,他们可以说出他们以前从没听过的话语:

- (3) Ich festhalte die.
我 紧-握 他们(Friederike, 2;6)
想说:‘我紧紧地握住了。’

Friederike 说这句话的时候,正在习得德语句中定式动词(finite verb)位置的规则,即定式动词位于第二位。但是,她在这句话中将整个动词放在了第二位,其中就包括可分前缀 fest(紧)。正确的用法应该是将这个可分前缀放在句子的末尾。

¹ 也有学者认为我们有能力构建出无限长的句子(Nowak, Komarova & Niyogi 2001: 117; Kim & Sells 2008: 3; Dan Everett 在 O'Neill & Wood (2012) 中的 25 时 19 分),乔姆斯基也这样认为(Leiss 2003: 341)。但这是不正确的,因为每个句子必然要在某个节点结束。即使是在乔姆斯基传统下发展起来的形式语言也没有无限长的句子。不过,有些形式语法可以描述出一系列无限地包含着有限句子的情况。(Chomsky 1957: 13)也可以参看 Pullum & Scholz (2010) 和第 13.1.8 节中关于语法的递归性与语言无限论的观点。

如果我们并不希望语言仅仅是由一个个形式意义对儿构成的，那么必然需要将话语中小的意义成分构成复杂话语的意义的过程。句法就是用来揭示词汇组合的方式和话语的结构。比如说，主谓一致的知识有助于我们解释下面的德语句子：

- (4) a. Die Frau schläft.
DET 女人睡觉.3SG
‘这个女人睡着了。’
- b. Die Mädchen schlafen.
DET 女孩儿.3PL 睡觉.3PL
‘这些女孩儿睡着了。’
- c. Die Frau kennt die Mädchen.
DET 女人认识.3SG 这 女孩儿.3PL
‘这个女人认识这些女孩儿。’
- d. Die Frau kennen die Mädchen.
DET 女人.3PL 认识.3PL 这个女孩儿.3PL
‘这些女人认识这些女孩儿。’

例句（4a,b）显示，主语的单数或复数形式需要由相应的动词屈折形式来搭配。在（4c,d）中，动词带两个论元成分，而“die Frau”（这个女人）和“die Mädchen”（这些女孩儿）在德语中可以出现在任意一个论元位置上。这些句子可以理解为这个女人认识某人或者某人认识这个女人。不过，按照动词的屈折变化以及德语的句法规则，听话人知道对于（4c）和（4d）来说都只有一种解读。

所以说，句法就是用来发现、描写和解释这些规则、范式和结构的。

1.2 为什么要形式化？

为什么要对语言进行形式化的描述呢？我们先来看两条经典的论述：

在发现语言结构的过程中，对语言结构的精确建模，不管是正面的、还是负面的，都起到十分重要的作用。通过对一条不可接受的结论进行精确但是不充分的建模过程，可以暴露出这种不充分的确切来源，并最终对语言数据有更深层次的理解。乐观来看，一个形式化的理论可以为许多问题自动提供解决方案，而不是像其他那些被精细设计的理论那样。那些晦涩的，凭直觉判定的理念既不能得到荒谬的结论，也不能得到新的、正确的结论。所以说，他们在这两个方面都没有实际的用处。我认为那些质疑过语言学理论的发展在精确和技术层面上的价值的学者们都无法认识到这种方法的巨大潜力，即严密地论证观点以及严格地将之应用在语言事实上，而不是为了避免由特设的调整以及模糊的组成方式造成的不合格的结论。(Chomsky 1957: 5)

正如我们经常指出，但是不能过度强调的是，语言形式化的一个重要的目的就是可以使得研究者们看到一种观点的缺点的同时也看到它的优点。只有这样，我们才能使研究变得有效率。(Dowty 1979: 322)

1 导言与术语

如果我们将语言的描写形式化，就会易于我们认识某一特定分析所表示的确切含义。我们可以构建起该分析下可预测的内容，并排除其他的分析。另一个优势在于，精确的形式化理论可以用计算机程序能够理解的方式记录下来。这样，一个理论分析就可以作为计算过程中的语法部分来实现，如有不一致的地方就会更快地显现出来。这种语法可以用来分析大规模的数据，也叫做语料库，而且他们可以构建出语法尚无法分析的句子或者组配错误的结果。更多的在语言学中应用计算机实现方面的研究可以参看Bierwisch (1963: 163), Müller (1999a: Chapter 22) 和Bender (2008b) 这几篇文献，也可参看第3.6.2节的内容。

1.3 构成成分

拿例（5）来说，我们可以凭直觉判断出句中有些词构成了一个单位。

- (5) Alle Studenten lesen während dieser Zeit Bücher.
所有学生 读 在 这 时间书
‘在这个时候，所有的学生都在读书。’

例如，alle（所有）和 Studenten（学生）这两个词组成了一个单位，说的是有人在读书的意思。während（在……时候），dieser（这）和 Zeit（时间）这三个词也组成了一个单位，意思是读这个动作发生的这段时间，而 Bücher（书）是指读的对象。第一个单位本身由两部分组成，即 alle（所有）和 Studenten（学生）。“während dieser Zeit”（在这个时候）这个单位也可以分成两个小部分：während（在……时候）和 “dieser Zeit”（这个时候）。“dieser Zeit”（这个时候）也由两部分组成，跟 “alle Studenten”（所有的学生）是一样的。

在前面的例（1c）中，我们用俄罗斯套娃（matryoshkas）来比喻语言中的嵌套现象。这里，我们也可以将（5）分成更小的单位，从而组成更大的单位。但是，与俄罗斯套娃的隐喻不同的是，我们不能只将小的单位放在大的单位里，而是我们可以将几个单位组成更大的单位。最好的方法就是把它想象成一套盒子系统：一个大盒子装着整个句子。在这个盒子里，还有四个小盒子，每个盒子分别装着 “alle Studenten”（所有的学生），“lesen”（读），“während dieser Zeit”（在这个时候）和 “Bücher”（书）。如图 1.1 所示。

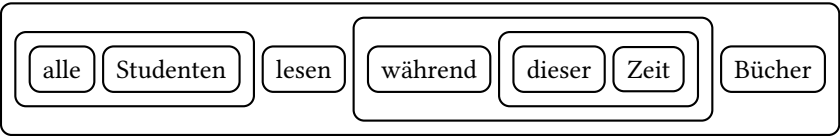


图 1.1: 盒子中的词与短语

在下一节，我会介绍几种不同的测验方法来判断出哪些词与其他词相比是“在一起的”。每当我说到一个词语序列（word sequence）的时候，通常是指一个任意的线性词语序列，它并不必须具有句法或语义上的联系，比如说例（5）中 “students read during”（学生们读在）。而由一组词组成的结构单位通常叫做短语（phrase）。

短语由词构成，如“this time”（这个时候），或者由词与短语组成，如“during this time”（在这个时候）。短语内的部分和短语本身叫做构成成分（constitutnes）。所以说，在图 1.1 中，盒子中的所有元素都是句子的构成成分。

基于上述的基本说明，我现在将介绍一些可以帮助我们判断出一个词串是否是构成成分的测试方法。

1.3.1 组成成分测试

有许多方法可以用来测试词汇序列的构成成分。在下面的章节中，我会介绍其中一些方法。在第 1.3.2 节，我们还将看到简单地盲目测试只会得到一些无用的结果。

1.3.1.1 替换

如果能将一个句子内的词语序列替换为另一套不同的词语序列，而且句子的可接受程度不变，那么就可以证明这些词语序列是一个构成成分。在例（6）中，“der Mann”（男人）可以被替换为“eine Frau”（一个女人）。这就表明了这些词语序列都是组成成分。

- (6) a. Er kennt [den Mann].
他认识 那 男人
‘他认识那个男人。’
b. Er kennt [eine Frau].
他认识 一 女人
‘他认识一个女人。’

与（7a）类似的是，字符串“das Buch zu lesen”可以被替换为“der Frau das Buch zu geben”。

- (7) a. Er versucht, [das Buch zu lesen].
他试图 这 书 AUX 读
‘他试着读这本书。’
b. Er versucht, [der Frau das Buch zu geben].
他试图 这 女人这 书 AUX 给
‘他试着把这本书给这位女士。’

这类测验叫做替换测试（substitution test）。

1.3.1.2 代词化

凡是能够由代词所替代的也构成一个成分。在例（8）中，我们可以用代词 er（他）替代“der Mann”（这个男人）：

1 导言与术语

- (8) a. [Der Mann] schläft.
这 男人 睡觉
‘这个男人在睡觉。’
b. Er schläft.
他 睡觉
‘他在睡觉。’

我们也可以用代词来指代诸如 “das Buch zu lesen” 这样的成分，如下面的例 (9) 所示：

- (9) a. Peter versucht, [das Buch zu lesen].
Peter 试图 这 书 AUX 读
‘Peter 试着读这本书。’
b. Klaus versucht das auch.
Klaus 试图 这 也
‘Klaus 也试着这么做。’

代词化测试是替换测试的另一种形式。

1.3.1.3 疑问结构

如果一个词语序列能够用问句来提问，那么它就构成一个成分：

- (10) a. [Der Mann] arbeitet.
这 男人 工作
‘这个男人在工作。’
b. Wer arbeitet?
谁 工作
‘谁在工作？’

疑问结构属于一种特殊的代词化。我们可以用一种特殊类型的代词（疑问代词）来指代词语序列。像 (7a) “das Buch zu lesen” 这类成分就可以用疑问词来提问，如例 (11) 所示：

- (11) Was versucht er?
什么 试图 他
‘他想试什么？’

1.3.1.4 变换测试

如有一组词语序列可以移动，而不会影响到其所在句子的合格性，那么这就意味着这个词语序列组成了一个成分。在例 (12) 中，keiner（没有人）和 diese Frau（这个女人）可以有不同的语序排列方式，这就意味着 diese（这）和 Frau（女人）是一个成分。

- (12) a. dass keiner [diese Frau] kennt
CONJ 没有 这 女人 认识
b. dass [diese Frau] keiner kennt
CONJ 这 女人 没有 认识
‘没有人认识这个女人’

从另一个角度来看, *keiner diese* (没有这) 在例 (12a) 中是无法构成一个成分的。如我们将 *keiner diese* (没有这) 整体移动的话, 我们会得到不合格的句子:²

- (13) a. *dass Frau keiner diese kennt
CONJ 女人没有 这 认识
b. *dass Frau kennt keiner diese
CONJ 女人认识 没有 这

再者, 诸如例 (7a) 中 “the das Buch zu lesen” 这一组成成分是可以移动的:

- (14) a. Er hat noch nicht [das Buch zu lesen] versucht.
他 AUX PART 不 DET 书 AUX 读 试图
‘他还没有试着读这本书。’
b. Er hat [das Buch zu lesen] noch nicht versucht.
他 AUX DET 书 AUX 读 PART 不 试图
c. Er hat noch nicht versucht, [das Buch zu lesen].
他 AUX PART 不 试图 DET 书 AUX 读

1.3.1.5 前置

前置是更深层次的移位测试。在德语的陈述句中, 只有一个成分能够前置到定式动词的前面:

- (15) a. [Alle Studenten] lesen während der vorlesungsfreien Zeit Bücher.
所有学生 读.3PL 在 DET 课程. 空闲 时间书
‘所有的学生都在学期放假的时候看书。’
b. [Bücher] lesen alle Studenten während der vorlesungsfreien Zeit.
书 读 所有学生 在 DET 课程. 空闲 时间
c. * [Alle Studenten] [Bücher] lesen während der vorlesungsfreien Zeit.
所有学生 书 读 在 DET 课程. 空闲 时间
d. * [Bücher] [alle Studenten] lesen während der vorlesungsfreien Zeit.
书 所有学生 读 在 DET 课程. 空闲 时间

一个词语序列能否前置, 即出现在定式动词的前面, 是确定其为一个成分的重要依据。

² 我在所有的例句中都使用如下的符号: ‘*’表示句子是不合乎语法的, ‘#’表示句子有着不同于常规用法的解读, 最后 ‘\$’ 是指那些因语义或信息结构等方面的原因可以被解读的句子, 比如说, 主语必须是生有的, 但是实际上我们提问的是非生有的主语, 或者由于代词的使用, 成分序列和已知信息的标记之间存在矛盾。

1 导言与术语

1.3.1.6 并列

如有两组词语序列可以连在一起，那么每一组词都是一个构成成分。

在例(16)中，*der Mann*（这个男人）和 *die Frau*（这个女人）连在一起使用，整个短语作动词 *arbeiten*（工作）的主语。这个事实可以证明 *der Mann*（这个男人）和 *die Frau*（这个女人）构成一个成分。

- (16) [Der Mann] und [die Frau] arbeiten.
这 男人 和 这 女人 工作.3PL
‘这个男人和这个女人都工作。’

例(17)说明了带 *zu* 不定式的短语可以并列：

- (17) Er hat versucht, [das Buch zu lesen] und [es dann unauffällig verschwinden
他AUX 试图 这 书 CONJ 读 和 它然后 悄悄地 消失
zu lassen].
CONJ 让
‘他试着读这本书，然后让它悄悄地消失。’

1.3.2 关于成分测试法的一些看法

如果上述测试法可以对每一种情况给出明确的结果，那就太理想了，正如基于经验主义的句法理论也会变得更加明确一样。然而，实际情况并不是这样。成分测试法实际上存在一些问题，我将在下面具体讨论。

1.3.2.1 虚位成分

在代词中有一类特殊的词，叫做虚位成分（*expletives*），它们并不指称人或者事物，即它们是无指的。例(18)中的 *es*（它）就是一个例子。

- (18) a. Es regnet.
它 下雨
‘下雨了。’
b. Regnet es?
下雨 它
‘下雨了吗？’
c. dass es jetzt regnet
CONJ 它 现在 下雨
‘现在在下雨’

如例(18)所示，*es*（它）可以用在动词前面，也可以用在动词后面。有副词的时候，它也可以与动词分开，这就意味着 *es*（它）可以被看作是一个独立的成分。无论如何，我们观察到上述测试存在一些问题。首先，在例(19a)和(20b)中，*es*（它）在移位方面有限制。

- (19) a. *dass jetzt es regnet
CONJ 现在它下雨
想说：“现在正在下雨”
b. dass jetzt keiner klatscht
CONJ 现在没有人鼓掌
‘现在没有人在鼓掌’
- (20) a. Er sah es regnen.
他看见它.ACC 下雨
‘他看见在下雨。’
b. *Es sah er regnen.
它.ACC 看见他下雨
想说：“他看见下雨了。’
c. Er sah einen Mann klatschen.
他看见一.ACC 男人 鼓掌
‘他看见一个男人在鼓掌。’
d. Einen Mann sah er klatschen.
一.ACC 人 看见他鼓掌
‘一个男人，他看见在鼓掌。’

与(20c,d)中的宾格宾语 *einen Mann* (一个男人)不同的是,(20b)中的形式代词不能前置。第二,替换和疑问测试法也不适用:

- (21) a. *Der Mann / er regnet.
DET 男人 他下雨
b. *Wer / was regnet?
谁 什么下雨

类似地,并列测试法也不适用:

- (22) *Es und der Mann regnet / regnen.
它和 DET 男人 下雨.3SG 下雨.3PL

这些测试方法不适用的原因是:弱重音的代词 *es* (它)倾向于位于其他成分的前面,连词的后面(18c中的 *dass*),以及定式动词的后面(20a)(参考Abraham 1995: 570)。如果一个成分位于虚位成分的前面,如(19a)所示,那么,整个句子就是不合乎语法的。例(20b)不合乎语法的原因在于宾格 *es* (它)不能位于小句句首的位置。尽管有这样的情况,只有当 *es* 是有指(referential)的时候才是成立的(Lenerz 1994: 162; Gärtner & Steinbach 1997: 4)。

事实是,我们不能在上例中应用替换和疑问测试的方法的原因是,这些例子中的 *es* 是无指的。我们可以将 *es* 替换为另一个虚位成分,比如说 *das*。如果将虚位成分替换为一个有指的表达式,我们可以得到一个不同的语义解释。而语义上的空位概念或者用代词来指代都是没有意义的。

1 导言与术语

这样看来并不是所有的测试法都会将词语序列区分出不同的成分，也就是说，这些测试法并不是检验组成成分的必要条件。

1.3.2.2 移位

对于具有相对自由语序的语言来说，移位测试是有问题的，因为我们不可能总是能够准确地判断出移位的成分。例如，字符串“gestern dem Mann”就有着不同的排列顺序。

- (23) a. weil keiner gestern dem Mann geholfen hat
因为没有人昨天 DET 男人 帮助 AUX
‘因为昨天没有人帮助了那个男人’
b. weil gestern dem Mann keiner geholfen hat
因为昨天 DET 男人 没有人帮助 AUX
‘因为昨天没有人帮助了那个男人’

我们可以推断出，gestern（昨天）和 dem Mann（那个男人）虽是一起移位的，但是它们并不能构成一个成分。对于（23）中语序变化的另一个解释是副词可以在小句的不同位置出现，而且只有（23b）中的 dem Mann（那个男人）移到了 keiner（没有）的前面。不管在什么情况下，gestern（昨天）和 dem Mann（那个男人）都没有语义关系，而且不可能用一个代词来指称他们。尽管看上去是，这个部分是以一个单位来移动的，可实际上我们知道 gestern dem Mann（昨天那个男人）并不能构成一个成分。

1.3.2.3 前置

正如我们在（15）中前置这一部分所讨论的，定式动词前的位置一般由一个成分充当。在定式动词前能否放置一组词被用来作为判断成分状态的明确标记，这种方法还用在了 Satzglied 这个术语的定义中。（Satzglied 在德语语法中是指小句层面的句子成分）（Eisenberg et al. 2005: 783）。Bußmann (1983)提出了一个例子，而这个例子在Bußmann (1990)中已经找不到了：³

句子成分测试是一个应用话题化的方式来对复杂成分进行分析的过程。由于话题化只允许一个单独的成分移到句子的开头，复杂的成分序列，比如说副词短语，实际上包括一个或多个组成成分。在“Ein Taxi quält sich im Schrittempo durch den Verkehr（出租车正在以步行的速度顽强地前进）”这个例子中，im Schrittempo（以步行的速度）和 durch den Verkehr（通过交通）都是句子成分，因为他们都可以各自独立地前置。（Bußmann 1983: 446）

上面这段话可以得到如下的推论：

- 如果某一个语言片段中的某一部分断片可以各自独立地前置，那么该语言片段不构成一个成分。

³ 句子成分测试（Satzgliedtest，也写作 Konstituententest）

- 如果某一个语言片段可以整体前置，那么该语言片段构成一个成分。

接下来我们要指出的是，这两种说法都是有问题的。第一个观点在（24）中就站不住脚：

- (24) a. Keine Einigung erreichten Schröder und Chirac über den Abbau der
不 一致 达成 Schröder 和 Chirac 关于DET 减少 PREP.DET
Agrarsubventionen.⁴
农业. 补贴
‘Schröder 和 Chirac 不能就农业补贴的减少达成一致。’
- b. [Über den Abbau der Agrarsubventionen] erreichten Schröder
关于 DET 减少 DET 农业. 补贴 达成 Schröder 和 Chirac
und Chirac keine Einigung.
不 一致

尽管“keine Einigung über den Abbau der Agrarsubventionen”这个名词短语的一部分可以独立地前置，在例（25）中我们仍可将没有前置的整个词串分析为一个名词短语。

- (25) Schröder und Chirac erreichten [keine Einigung über den Abbau der Agrarsubventionen].
Schröder 和 Chirac 达成 不 一致 关于DET 减少 DET 农业. 补贴

“über den Abbau der Agrarsubventionen”（关于缩减农业补贴）这个介词短语在语义上依存于 Einigung（一致），如例（26）所示：

- (26) Sie einigen sich über die Agrarsubventionen.
他们同意 REFL 关于DET 农业. 补贴
‘他们在农业补贴方面达成了一致。’

这个词语序列也可以整体前置：

- (27) [Keine Einigung über den Abbau der Agrarsubventionen] erreichten Schröder
不 一致 关于DET 减少 DET 农业. 补贴 达成 Schröder
und Chirac.
和 Chirac

在理论文献中，人们普遍认为“keine Einigung über den Abbau der Agrarsubventionen”构成了一个成分，并且它可以在一定的情况下进行“分裂”。

⁴ 《每日新闻》，2002 年 10 月 15 日，晚八点。

1 导言与术语

在这种情况下，如我们在例（25）中看到的，次级组成成分是可以各自独立移位的。(De Kuthy 2002)

第二条推论也是有问题的，如例（28）所示：

- (28) a. [Trocken] [durch die Stadt] kommt man am Wochenende auch
变干的 通过 DET 城市 来 一 PREP.DET 周末 也
mit der BVG.⁵
PREP DET BVG
‘搭乘 BVG，你就可以在周末确保风雨无阻地穿越这个城市。’
- b. [Wenig] [mit Sprachgeschichte] hat der dritte Beitrag in dieser
少 PREP 语言. 历史AUX DET 第三 文章PREP
Rubrik zu tun, [...] ⁶
DET 章节AUX 做
‘这一节中的第三篇文章与语言历史没有太多的关系。’

在例（28）中，定式动词前有多个组成成分，而且这些成分之间没有明显的句法或语义上的联系。在下面的章节中，我们会详细解释什么叫做“句法或语义关系”。在这一点上，我仅指出（28a）中的形容词 trocken（干燥）的主语是 man（人），并且是该主语进一步说明了有关“穿越城市旅行”的动作行为，这是因为它指称了动词所指的动作。如（29b）所示，durch die Stadt（穿过城市）不能与形容词 trocken（干燥）相组合。

- (29) a. Man ist / bleibt trocken.
人 是 保持 干燥
‘有人是干的。’
- b. *Man ist / bleibt trocken durch die Stadt.
人 是 保持 干燥 通过 DET 城市

所以说，形容词 trocken（干燥）与介词短语 durch die Stadt（穿越城市）之间不具有句法或语义上的联系。这些短语的共性在于他们都指向动词并且与它有依存关系。

有学者认为应该把例（28）看作是例外。不过，正如我在相关的经验主义的研究中所指出的，这种方法也是有问题的 (Müller 2003a)。

如果我们根据是否能够通过前置测试而将 durch die Stadt（通过这个城市）看作是一个成分，那么我们必须承认例（30）中的 durch die Stadt（通过这个城市）也是一个成分。这么做的结果是，我们会低估组成成分这个术语，因为组成成分测试的目的就是为了找到词串间的语义与语法联系。⁷

⁵ taz berlin, 1998 年 7 月 10 日, p. 22.

⁶ Zeitschrift für Dialektologie und Linguistik, LXIX, 3/2002, p. 339.

⁷ 这些数据可以这样来解释，即假设一个空动词的中心语位于定式动词的前面，继而保证了在定式动词前只有一个成分位于首位的要求。(Müller 2005c, 2015b)无论如何，这类数据对于组成成分测试来说都是有问题的，因为这些测试是专门用来区分例（30）中的 trocken（变干），durch die Stadt（通过这个城市），以及 mit Sprachgeschichte（与语言历史）这类字符串是否是组成成分的。

- (30) a. Man kommt am Wochenende auch mit der BVG trocken durch
人 来 PREP.DET 周末 也 PREP DET BVG 干燥 通过
die Stadt.
DET 城市
‘搭乘 BVG, 你就可以在周末确保风雨无阻地穿越这个城市。’
- b. Der dritte Beitrag in dieser Rubrik hat wenig mit Sprachgeschichte
DET 第三 文章 在 DET 章节 AUX 少 PREP 语言. 历史
zu tun.
AUX 做
‘这一节中的第三篇文章与语言历史没有太多的关系。’

所以说, 一个给定词语序列是否能够前置并不足以用来判断组成成分的性质。

再有, 虚位成分也被看作是组成成分, 尽管事实上宾格的虚位成分并不能前置 (cf. (20a)):

- (31) a. Er bringt es bis zum Professor.
他 拿 EXPL 直到 PREP.DET 教授
‘他把它拿给教授。’
- b. # Es bringt er bis zum Professor.
它 拿 他 直到 PREP.DET 教授

还有其他成分也不能前置。自反身代词就是一个很好的例子:

- (32) a. Karl hat sich nicht erholt.
Karl AUX REFL 不 恢复
‘Karl 还没有恢复。’
- b. *Sich hat Karl nicht erholt.
REFL AUX Karl 不 恢复

由此可见, 前置并不是成分测试的必要条件。这样的话, 一个给定词串能否前置对于判断组成成分的性质来说既不是充分也不是必要的条件了。

1.3.2.4 并列

例 (33) 中的并列结构也被证明是有问题的:

- (33) Deshalb kaufte der Mann einen Esel und die Frau ein Pferd.
所以 买 DET 男人 一 驴子和 DET 女人 一 马
‘所以说, 这个男人买了一头驴, 这个女人买了一匹马。’

乍看上去, *der Mann einen Esel* (这个男人一头驴) 和 *die Frau ein Pferd* (这个女人一匹马) 在例 (33) 中是并列的。这是不是说 *der Mann einen Esel* (这个男人一头驴) 和 *die Frau ein Pferd* (这个女人一匹马) 分别构成一个组成成分呢?

1 导言与术语

正如利用其他成分测试方法所证明的那样，这个观点并不是像它看上去那样的。这组词不能作为一个单位来整体移动：⁸

- (34) *Der Mann einen Esel kaufte deshalb.
DET 男人 一 驴子买 所以

我们也不能替换这个组成成分，除非有省略的情况。

- (35) a. #Deshalb kaufte er.
所以 买 他
b. *Deshalb kaufte ihn.
所以 买 他

代词不能填充到 kaufen（买）的两个逻辑论元的位置上，而是由（33）中的 der Mann（这个男人）和（33）中 einen Esel（一头驴子）来填充的，只不过是每个位置有一个成分。对于（33）这类例子的分析也有不同的看法，即这里有两个动词 kauft（买），其中只有一个是显性的 (Crysmann 2008)。这样，例（33）应该是：

- (36) Deshalb kaufte der Mann einen Esel und kaufte die Frau ein Pferd.
所以 买 DET 男人 一 驴子和 买 DET 女人 一 马

这就意味着，即使 der Mann einen Esel（这个男人一头驴）和 die Frau ein Pferd（这个女人一匹马）看上去是并列结构，实际上并列的成分是 kauft der Mann einen Esel（买这个男人一头驴）和 (kauft) die Frau ein Pferd（买）这个女人一匹马。

通过上面的讨论，我们得到的结论是：即使一个给定的词语序列通过了某种成分测试法，这并不意味着我们能够自动从这个测试中推导出它是一个组成成分，也就是说，上面的测试并不是判断组成成分性质的充分条件。

综上所述，这些测试方法对于判断一组词的组成成分来说既不是充分条件也不是必要条件。但是，只要我们对有争议的地方保持清醒的认识，我们就会大概知道如何来判断组成成分了。

1.4 词类

例（37）中的词不仅意义不同，其他方面也有所不同。

- (37) Der dicke Mann lacht jetzt.
DET 胖 男人 笑 现在
‘那个胖男人在笑。’

⁸ 定式动词前的位置也叫作前场 (Vorfeld)（参看 1.8 节）。德语中，显性的多项成分前置在某些情况下是可能的。请参考前面的内容，尤其是第 14 页的例（28）。例（34）也是一个例子，对于 kaufen（买）这类动词来说，主语在前场的位置上是比较少见的，因为这种前置的结构与信息结构有关。我们也可以比较 De Kuthy & Meurers 2003b 有关前置动词短语的主语的研究，以及 Bildhauer & Cook 2010: 72 有关显性多重前置的主语前置方面的研究。

句中的每一个词都受到某种限制。常见的练习就是将具有共同属性的词归为一类。比如说，*der*（那）是一个冠词，*Mann*（男人）是一个名词，*lacht*（笑）是一个动词，*jetzt*（现在）是一个副词。如例（38）所示，我们可以将例（37）中的词替换为相同词类的词。

- (38) *Die dünne Frau lächelt immer.*
 DET 瘦 女人笑 一直
 ‘那个瘦女人一直在笑。’

但是，并不是所有的词都能替换。比如说，我们不能用反身动词 *erholt*（恢复）或者例（38）中第二人称的动词 *lächelst*（笑）来替换。把词归入相应的词类并不是这样简单的。我们还要考虑给定词的属性。在这一节，我们会讨论不同的词类，并在下一节详细描述一个给定词类的若干属性。

词类中最为重要的有动词（verb）、名词（noun）、形容词（adjective）、介词（preposition）和副词（adverb）。很多年前，研究德语的学者通常会区分动作类词（action words），描述类词（describing words）和命名类词（naming words）（参看第 11.6.1 节有关 *Tesnière's* 的范畴系统）。但是这些说法都被证明是有问题的，如下所示：

- (39) a. *die Idee*
 DET 主意
 b. *die Stunde*
 DET 小时
 c. *das laute Sprechen*
 DET 大 说话
 ‘大声地说话（动作）’
 d. *Die Erörterung der Lage dauerte mehrere Stunden.*
 DET 讨论 DET 情形持续 几个 小时
 ‘有关这个情形的讨论已经持续了几个小时。’

(39a) 并不能描述一个实体，(39b) 描述的是一段长时间，(39c) 和 (39d) 描述动作。显然，*Idee*、*Stunde*、*Sprechen* 和 *Erörterung* 在意义上有很大的区别。无论如何，这些词仍在很多方面与 *Mann* 和 *Frau* 有相同之处，所以它们被归为名词。

“动作类词”这个术语不能在科学的语法体系中用来指称动词了，因为动词并不一定指称动作：

- (40) a. *Ihm gefällt das Buch.*
 他 喜欢 DET 书
 ‘他喜欢这本书。’
 b. *Das Eis schmilzt.*
 DET 冰 融化
 ‘冰融化了。’

1 导言与术语

- c. Es regnet.
它 下雨
'下雨了。'

我们也可以将 *Erörterung* (讨论) 归入动作类动词。形容词 并不总是描述事物的属性。在下面的例子中, 相反的情况是真实存在的, 即将一个杀人犯的特质表示为一种可能或者猜测, 而不是被修饰名词的真实属性。

- (41) a. der mutmaßliche Mörder
DET 被怀疑的 杀人犯
b. Soldaten sind potenzielle Mörder.
士兵 是 潜在的 杀人犯

例 (41) 中的形容词实际上并没有提供所描述实体的特征信息。我们也希望将例 (42) 中的 *lachende* (笑) 看作是一个形容词。

- (42) der lachende Mann
DET 笑 男人

不过, 如果我们将属性和行为作为分类的标准, 那么 *lachend* (笑) 在技术层面上来说应该属于动作类词。

与语义标准 不同的是, 决定词类的标准通常是形式上的标准。词的不同形式也要纳入考察的范围。所以, 举例来说 *lacht* (笑) 在例 (43) 中有如下几种形式。

- (43) a. Ich lache.
我 笑
b. Du lachst.
你.SG 笑
c. Er lacht.
他 笑
d. Wir lachen.
我们笑
e. Ihr lacht.
你们.PL 笑
f. Sie lachen.
他们笑

德语也有过去式、命令式、虚拟语气和不定式 (助词和带 *zu* 与不带 *zu* 的不定式) 的形态变化。这些形式就构成了动词的屈折范式。时态 (现在式、过去式、将来式), 情态 (陈述语气、虚拟语气、命令语气), 人称 (第 1, 2, 3 人称) 和数 (单数、复数) 都在屈折范式中有所体现。在有些范式中这些形式有所重合, 如例 (43c), 例 (43e), 例 (43d) 和例 (43f) 所示。

与动词相似的是, 名词也有屈折范式:

- (44) a. der Mann
DEF.NOM 男人

- b. des Mannes
DEF.GEN 男人.GEN
- c. dem Mann
DEF.DAT 男人
- d. den Mann
DEF.ACC 男人
- e. die Männer
DEF.NOM 男人
- f. der Männer
DEF.GEN 男人
- g. den Männern
DEF.DAT 男人.DAT
- h. die Männer
DEF.ACC 男人

我们根据性（阴性、阳性和中性）来区分名词。性一般是纯形式上的性质，只是部分地受到生物性别或者我们描述特定物体的事实的影响。

- (45)
- a. die Tüte
DEF.F 包 (F)
‘包’
 - b. der Krampf
DEF.M 夹子 (M)
‘夹子’
 - c. das Kind
DEF.N 孩子 (N)
‘孩子’

与性相似的是，格（主格、属格、与格、宾格）与数对名词性范式来说也是同样重要的。

与名词相似的是，形容词也有性、数和格的屈折变化。不过，这些变化与名词不同，因为性的标示是可变的。形容词则可以与这三种性一同使用。

- (46)
- a. eine kluge Frau
一.F 聪明.F 女人
 - b. ein kluger Mann
一 聪明.M 男人
 - c. ein kluges Kind
一 聪明.N 孩子

除了性数格之外，我们还可以区分其他几种屈折类型。传统上来说，我们区分形容词的强、中和弱变化。这些屈折类与定冠词的形式或者有无是密切相关的：

1 导言与术语

- (47) a. ein alter Wein
一 老 酒
b. der alte Wein
DET 老 酒
c. alter Wein
老 酒

另外，形容词也有比较级和最高级：

- (48) a. klug
聪明
b. klüg-er
聪明-PTCP
c. am klüg-sten
SUPRL 聪明-PTCP

并不是所有的形容词都有比较级。对于那些指称终结点的形容词来说只能用在肯定式中，如果有一个最优解，那么就没有更好的了。所以说，我们不能说一个“更好的最优”方案。相似的是，也不能比死“更死”的了。

特殊情况是德语中一些以 a 结尾的颜色形容词，如 lila（紫色）和 rosa（粉色）。这些词的屈折形式是可选的（49a），没有屈折变化的形式也是可行的：

- (49) a. eine lilan-e Blume
一 紫色-F 花
b. eine lila Blume
一 紫色花

上述例子中，lila（紫色）可以归入形容词。这是因为他们与其他形容词处于同样的位置上，并且在屈折变化上与形容词的变化是一致的。

迄今我们所讨论的词类与屈折变化属性的概念是不同的。对于那些没有屈折变化的词而言，我们需要用到额外的标准。比如说，我们可以通过他们出现的句法环境来判别词类（正如我们对上面有屈折变化的形容词所做的那样）。我们可以区分出介词、副词、连词、感叹词，有时也可以区分出助词。介词是指那些与名词共现，并决定这些名词的格属性的一类词：

- (50) a. in diesen Raum
在DEF.ACC 房间
b. in diesem Raum
在DEF.DAT 房间

wegen（因为）通常被看作是前置词，尽管它也会出现在名词后，后者在技术层面上被处理为后置词：

- (51) des Geldes wegen
DET 钱.GEN 因为
“因为钱”

如果希望对词的位置的说明保持中立的话，那么你也可以将之称为 *adpositions*（介词）。

与介词不同的是，副词不需要带名词短语。

- (52) a. Er schläft in diesem Raum.
他睡觉 PREP 这 房间
b. Er schläft dort.
他睡觉 那儿

有时，副词仅是简单地被看作是前置词的一种特殊形式（参看第 87 页）。对于这个观点的解释是，前置词短语如 *in diesem Raum*（在这间房间里）与相应的副词的表现是完全一致的。与 *dort*（那儿）不同的是，它需要带一个额外的名词短语。我们也可以在其他词类中找到类似的区分。例如，动词 *schlafen*（睡觉）需要带一个名词短语，而 *erkennen*（认识）需要带两个。

- (53) a. Er schläft.
他睡觉
b. Peter erkennt ihn.
Peter 认识 他

连词分为从属连词和并列连词。并列连词包括 *und*（和）与 *oder*（或者）。在并列结构中，两组具有同样句法属性的词被组合起来。他们在形式上是彼此有联系的。*dass*（这个）与 *weil*（因为）这两个连词是从属连词，因为他们引导的小句从属于一个更大的句子。

- (54) a. Klaus glaubt, dass er lügt.
Klaus 认为 CONJ 他说谎
‘Klaus 认为他在说谎。’
b. Klaus glaubt ihm nicht, weil er lügt.
Klaus 相信 他 不 因为他说谎
‘Klaus 不相信他，因为他在说谎。’

叹词 (*interjection*) 类似于小句表达式。如 “Ja! (是的!)” “Bitte! (请!)” “Hallo! (你好!)” “Hurra! (好耶!)” “Bravo! (太棒啦!)” “Pst! (嘘!)” “Plumps! (扑通!)”。

如果副词和介词不能归入某个特定的类别，那么副词就通常被看作是一种没有屈折变化的“剩余”类，其他的介词，连词或者感叹词都不能归入副词。有时这种“剩余”类是需要进一步划分的：只有那些出现在定式动词前的用作一个成分的词可以被看作是副词。这些不能前置的词被称为助词。这些助词可以根据他们不同的功能而归入不同的类别中，如程度助词和言外助词。由于这些基于功能的分类标准也包括了介词，但是我不作这种区分，只是将之归为副词。

我们已经将一些具有屈折变化的词归入不同的词类中了。如果需要划分词类，我们可以利用图 1.2 on the next page 来判断，该图摘自德语杜登语法 (Eisenberg et al. 2005: 133)。⁹

⁹ 《杜登语法》是德语正字法的官方文件。杜登语法虽然没有取得官方的地位，但是它也非常具有影响力，并且较多用于教学之中。在导言这部分内容中，我会经常引用这一重要文献。

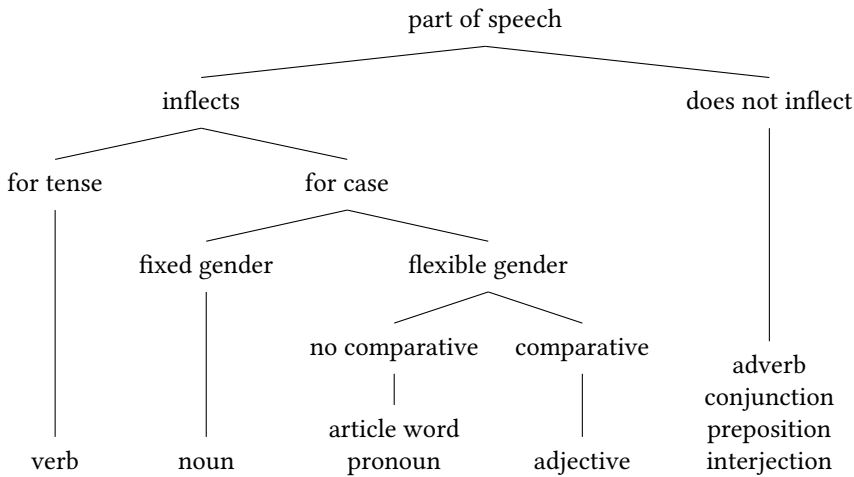


图 1.2: Eisenberg et al. (2005: 133)提出的词类决策树

如果一个词会随着时态发生变化，那么它就是动词。如果它有不同的变化式，那么就需要看它是否有固定的性。如果是这样的，那么就需要知道我们是不是在分析一个名词。具有不同性的词需要检查他们是否有比较级。如果是，那它可能是形容词。其他的词就归入剩余的类别中，杜登语法将之称为代词或冠词。对于没有屈折变化的元素来说，这个剩余类别中的元素可以根据他们的句法行为来进一步划分。杜登语法对代词和冠词进行了区分。基于这个分类标准，代词是那些可以替代诸如 *der Mann*（男人）的名词短语的词。而冠词通常与名词相组合。在拉丁语法中，代词包括上面所说的代词和冠词，因为它们带不带名词在形式上是一样的。在过去的几百年间，形式发生了分裂变化，在当代罗曼语族的语言中需要区分那些可以替代名词短语的词与必须与名词短语共现的词。后一种类别的词也叫做限定词（*determiner*）。

如果我们按照图 1.2 的决策树来分析代词（*pronoun*）的话，人称代词 *ich*（我），*du*（你），*er*（他），*sie*（她），*es*（它），*wir*（我们），*ihr*（你们）和 *sie*（他们）都可以跟属格代词 *mein*（我的），*dein*（你的），*sein*（他/它的），*ihr*（她的/他们的）和 *unser*（我们的）归为一类。相应的反身代词 *mich*（我自己），*dich*（你自己），*sich*（他/她/它/他们自己），*uns*（我们自己），*euch*（你们自己）和交互代词 *einander*（互相）在德语中是特殊的一类，因为他们没有不同的性的格式。交互代词没有格的形态变化。我们用交互代词替代属格、与格和宾格代词，无法看到 *einander*（交互）的格的变化形式，因为他们的格式是一样的：

- (55) a. *Sie gedenken seiner / einander.*
 他们想 他.GEN 互相
 b. *Sie helfen ihm / einander.*
 他们帮助 他.DAT 互相

- c. Sie lieben ihn / einander.
他们爱 他.ACC 互相

所谓的代副词 *darauf* (在那儿), *darin* (在这儿), *worauf* (在哪儿), *worin* (在哪儿) 也是有问题的。这些形式都包括一个介词, 如 *auf* (在……上), *da* (那儿), 以及 *wo* (哪儿)。正如其名称所示的, 代副词包括代词性成分, 而这个成分只能是 *da* (那儿) 和 *wo* (哪儿)。但是, *da* (那儿) 和 *wo* (哪儿) 并不具有屈折变化, 所以说按照决策树的划分, 应该将之归为代词。

例 (56) 中的相对代词也是类似的:

- (56) a. Ich komme eben aus der Stadt, wo ich Zeuge eines Unglücks
我 来 PART 自 DET 城市 CONJ 我 见证 PERP.ART 事故
gewesen bin.¹⁰
AUX COP
‘在我来自的那个城市里, 我亲眼目睹了一个事故。’
- b. Studien haben gezeigt, daß mehr Unfälle in Städten passieren, wo die
研究 AUX 表明 CONJ 更多 事故 在城市 发生 DET 斑马线
Zebrastreifen abgebaut werden, weil die Autofahrer unaufmerksam werden.¹¹
去除 CONJ 因为 DET 司机疏忽 AUX
‘研究表明, 城市中事故更多发生在没有斑马线的地方, 因为司机们更易疏忽。’
- c. Zufällig war ich in dem Augenblick zugegen, wo der Steppenwolf
恰巧地 AUX 我 在 DET 时候 出现 CONJ DET 荒原狼
zum erstenmal unser Haus betrat und bei meiner Tante sich einmietete.¹²
PREP.DET 第一次 我们的房子 进入 和 在我的 姑姑 REFL 租房
‘我恰好亲眼看见了荒原狼第一次闯进了我们的房子, 并在我姑姑那里租了一间房间。’

根据上面的决策树, 如果他们没有屈折变化, 则不能归入代词。Eisenberg (2004: 277) 指出 *wo* 是一种没有屈折变化的关系代词, 并且指明这种描述与名词性的用法是不同的, 因为名词是有屈折变化的元素。由此, 他用关系副词 (relative adverb) 来指称他们 (参见 Eisenberg et al. (2005: §856, §857))。

同样也有与名词相联系的关系词 *dessen* (他们) 和 *wessen* (他们) 的用法。

- (57) a. der Mann, dessen Schwester ich kenne
DET 男人 CONJ 姐妹 我 认识
- b. Ich möchte wissen, wessen Schwester du kennst.
我 想 知道 CONJ 姐妹 你 认识
‘我想知道你认识谁的姐妹。’

¹⁰ Drosdowski (1984: 672).

¹¹ taz berlin, 1997 年 11 月 3 日, 第 23 页。

¹² Herman Hesse, 《荒原狼》(Der Steppenwolf), Berlin und Weimar: Aufbau 出版社, 1986 年, 第 6 页。

1 导言与术语

根据杜登的分类标准，这些词被看作是“关系冠词”和“疑问冠词”。他们通常被看作是关系代词和疑问代词的一部分（参看Eisenberg (2004: 229)）。如果按照Eisenberg 提出的术语来看，这些词的归类问题是不具有争议的，因为他没有对冠词、代词和名词进行区分，而是将它们都归为名词类。但是对于那些提出需要区分出冠词和代词的学者来说，疑问代词也是一个经常探讨的问题，即他们可以作为冠词使用，也可以替代一个名词短语。

我们应须知，“代词”这个术语通常只是指那些能够指代其他实体的词，这点是非常重要的，这里所指的指代不是指像“书”和“约翰”等名词那样的指代，而是指依赖于语境的指代关系。例如，人称代词 *er*（他）既可以指桌子也可以指人。“代词”的这种用法与图 1.2 中的决策树是不同的，而且它还包括那些没有屈折变化的词，如 *da*（那儿）和 *wo*（哪儿）。

虚指代词，如 *es*（它），*das*（这个）及 *sich*（自己）这类反身动词，并不指代实际的物体。由于形式上的相似性，我们将之看作是代词。即使我们假定要对代词采用狭义的界定，我们也只会得到错误的结论，因为虚指词的形式并没有根据性、数和格而发生变化。如果我们按照教材中的分类标准，虚指成分可以归入无屈折变化的类型中。如果我们假定 *es*（它）与人称代词一定具有相同的主格和宾格形式，那么就可以将它们归入名词。这样我们就需要认为 *es* 是有性的，但是这样是讲不通的。这样，我们就将 *es* 看作是与人称代词相似的中性名词。

我们还没有讨论例（58）中斜体的词：

- (58) a. *das geliebte Spielzeug*
DET 喜欢 玩具
- b. *das schlafende Kind*
DET 睡着的 孩子
- c. *die Frage des Sprechens und Schreibens über Gefühle*
DET 问题 DET 谈论 和 书写 关于感情
'有关谈论和书写感情的问题'
- d. *Auf dem Europa-Parteitag fordern die Grünen einen ökosozialen Politikwechsel.*
在 DET 欧洲-党会 要求 DET 绿党 一 生态-社会 政治变革
'在欧洲党会中，绿党要求生态社会性的政治变革。'
- e. *Max lacht laut.*
Max 笑 大声
- f. *Max würde wahrscheinlich lachen.*
Max AUX 可能地 笑

geliebte（被爱）和 *schlafende*（睡着的）是 *lieben*（爱）和 *schlafen*（睡觉）的分词形式。这些形式传统上被看作是动词范式的一部分。从这个角度来看，*geliebte* 和 *schlafende* 是动词。这些形式可以归入词汇词的类别，即术语 *lexeme*（词位）。屈折范式的所有形式都属于相应的词位。传统意义上来看，这个术语还囊括了规则变化的屈折形式。也就是说，助词和名词化不定式也属于动词性词位。不过，不是所有的语言学家都持这一观点。问题在于我们将动词性范式与名词性和形容词性范式的概念混在一起了。比如说，*Sprechens*（谈论）是属格，并且形容词助词也有

性、数和格的屈折变化。进而，我们并不清楚为什么 *schlafende*（睡着的）应该归入动词词位，以及名词 *Störung*（混乱）单属于一个词位，而不是 *stören*（打扰）这个词位。我倾向于现代语法的解释，即随着词类发生变化，新的词位被创造出来。最终，*schlafende*（睡着的）不属于词位 *schlafen*（睡觉），而是词位 *schlafend* 在形式上的变化。这个词位属于形容词类并且有相应的屈折变化。

正如我们所看到，至今仍然难以区分屈折（inflection）与派生（derivation）（产生新的词位）。Sag, Wasow & Bender (2003: 263–264) 认为，英语中的现在分词与过去分词属于派生形式，而在法语中则随着性数而发生屈折变化。

例（58d）*Grünen*（绿色的）是名词化的形容词，在德语中，当文中没有其他名词时，它就与其他名词一样将首字母大写。

（59） A: Willst du den roten Ball haben?

要 你 DET 红色的球 有

“你想要这个红色的球吗？”

B: Nein, gib mir bitte den grünen.

不 给 我 请 DET 绿色的

“不，请给我那个绿色的，谢谢。”

在例（59）的答句中，名词 *ball*（球）被省略了。在例（58d）中则没有这种省略形式。我们也可以认为这里是词类发生了变化。如果一个词的词类发生了变化，但是并没有加上可见的词缀，那么，我们可以把这类现象叫做词类活用。有些语言学家认为转化是词汇派生的一个次级类型。但是，问题是 *Grüne*（绿色）就像形容词一样有屈折变化，并且会随着宾语的所指的性的不同而有所变化。

（60） a. Ein Grüner hat vorgeschlagen, ...

一 绿党成员.M AUX 提出建议

‘绿党的一位（男性）成员提出了建议 ...’

b. Eine Grüne hat vorgeschlagen, ...

一 绿党成员.F AUX 提出建议

‘绿党的一位（女性）成员提出了建议 ...’

我们也会遇到一个词兼有两种属性的情况。我们可以将之称为名词化的形容词。*Grüne*（绿色）的词汇范畴是形容词，而它的句法范畴（syntactic category）是名词。

在例（58e）中的词可以像形容词一样屈折变化，所以它在我们的测试中被归为形容词。有时，一些形容词也被叫做是副词。这是因为这些形容词的非屈折形式与副词很像：

（61） Max lacht immer / oft / laut.

Max 笑 总是 经常 大声

‘Max（总是 / 经常）（大声地）笑。’

为了描述这些词的双重特征，一些研究者认为有必要区分他们的词汇范畴与句法范畴。*loud* (*ly*) 在词汇范畴上属于形容词，而句法范畴上则属于副词。但是对于例（61）中 *loud* (*ly*) 的这种分类方法并不被所有的学者所认可。相反，有些人认

1 导言与术语

为这是形容词的副词性用法，也就是说，他们认为这些词的句法范畴仍是形容词，但是它们的用法可以不同，这样就像一个副词（参看Eisenberg 2004: Section 7.3）。这和介词的情况很像，介词也可以出现在不同的句法语境中：

- (62) a. Peter schläft im Büro.
Peter 睡觉 PERP.DET 办公室
‘Peter 在办公室里睡觉。’
b. der Tisch im Büro
DET 桌子 PERP.DET 办公室
‘办公室里的桌子’

在例（62）中，我们有两个介词短语的例子；但是，例（62a）中的 *im Büro*（在办公室里）像副词一样，因为它修饰的是动词 *schläft*（睡觉），而在例（62b）中的 *im Büro*（在办公室里）修饰是名词 *Tisch*（桌子）。同样，可以修饰名词（63）或动词（61）。

- (63) die laute Musik
DET 大声 音乐

1.5 中心语

一个构成成分或短语的中心语 决定了这个组成成分或短语的最重要的属性。同时，中心语也决定了这个短语的构成，即中心语要求在短语中有其他要素的共现。如下例所示（中心语用斜体表示）：

- (64) a. *Träumt* dieser Mann?
做梦 这.NOM 人
‘这个人做梦吗？’
b. *Erwartet* er diesen Mann?
等 他.NOM 这.ACC 人
‘他在等这个人吗？’
c. *Hilft* er diesem Mann?
帮助他.NOM 这.DAT 人
‘他在帮助这个人吗？’
d. *in* diesem Haus
在这.DAT 房间
e. ein Mann
一.NOM 人

动词决定它们的论元的格属性。在例（64d）中，介词决定名词短语（这个房子）的格属性（与格），也决定这个短语的语义（它描述了一个位置）。例（64e）是有争议的：有学者认为限定语是中心语（Vennemann & Harlow 1977; Hellan 1986; Abney 1987; Netter 1994, 1998），而其他学者认为名词是中心语（Van Langendonck

1994; Pollard & Sag 1994: 49; Demske 2001; Müller 2007b: Section 6.6.1; Hudson 2004; Bruening 2009)。

中心语与其他成分的组合叫做中心语的投射 (projection)。保证短语合法性所需的所有要素的投射叫做最大投射 (maximal projection)。一个句子就是定式动词的最大投射。

图 1.3 以方框的形式显示了例 (65) 的结构。

(65) Der Mann liest einen Aufsatz.

DET 人 读 一 论文

‘这个人在读一篇论文。’

与图 1.1不同的是，这些盒子都有标记。

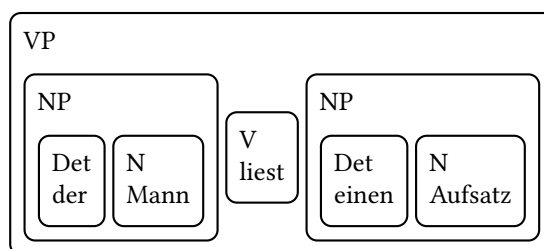


图 1.3: 有标记的盒子中的词与短语

这些标记包含了盒子中最重要的要素范畴。VP 表示动词短语，NP 表示名词短语。VP 和 NP 是他们各自中心语的最大投射。

试想你在自己姐妹的婚礼上被安排了找照片的任务，当那么多杂乱的、未分类的照片摆在你面前的时候，你一定会想说这些照片如果按照所含不同类型的照片标记的相册该有多好。对于上面的盒子来说，也是一样，如果能按照内容将它们分类将是一个好主意。

一个有趣的现象是如果装有语言素材的盒子放进更大的盒子中去时，这些盒子中的具体内容就不重要了。比如说，我们可以将名词短语 *der Mann* (男人) 用 *er* (他) 代替，或者用更为负责的形式 “*der Mann aus Stuttgart, der das Seminar zur Entwicklung der Zebrafinken besucht*” (从斯图加特来的那个参加斑马发展讨论班的男人) 代替。但是，我们不能用 *die Männer* (男人们) 来代替，也不能用 *des Mannes* (男人的) 来代替：

(66) a. **Die Männer liest einen Aufsatz.*

DET 男人 读 一 论文

b. **Des Mannes liest einen Aufsatz.*

DET 男人.GEN 读 一 论文

原因在于 *die Männer* (男人们) 是复数，而 *liest* (读) 是单数。带有属格的名词短语也不能出现，只有主格的名词才可以。所以说，我们有必要将那些对构成更大盒子的有用的信息标记出来。在下图中，我们加入了更多详细的标注信息。

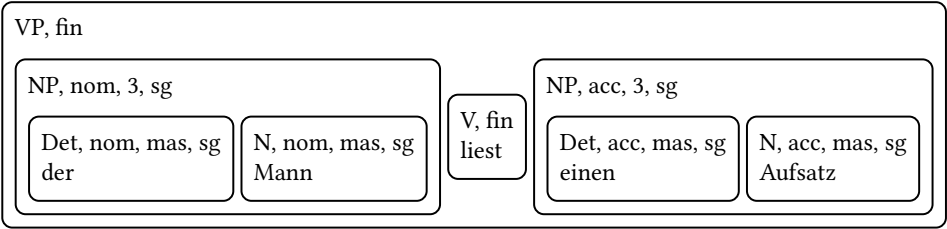


图 1.4: 在有标记盒子中的词与词串

那些决定中心语能带什么成分的特征叫做中心语特征（head features）。这些特征被认为是由中心语投射（projected）出来的。

1.6 论元成分与说明语

小句中的各成分与中心语具有不同的关系。最为典型的是区分论元成分和说明语。中心语的句法论元在很大程度上对应于其逻辑论元。我们可以用（67b）中的谓词逻辑来表示句子（67a）的意义。

- (67) a. Peter 帮助了 Maria.
- b. *help'(peter', maria')*

（67b）中的逻辑表达式与例（67a）所表示的十分相似；但是，它没有语序和屈折变化的信息。Peter 和 Maria 是动词（帮助）的句法论元，而它们各自的意义（*Peter'* 和 *Maria'*）是由（帮助）所表示的逻辑关系的论元。我们也可以说 *help*（帮助）指派给它的论元的语义角色。语义角色包括施事（发出动作的人），受事（受到影响的人或物），受益者（得到东西或经验的人），经事（经历某种心理状态的人）。*help*（帮助）的主语是事实，直接宾语是受益者。充当语义角色的论元也叫做行动元（actant）。这个术语用来指称无生的物体。

中心语和论元的这种关系叫做选择（selection）和价（valence）。价这个术语是从化学借来的。原子与原子组合成分子具有不同程度的稳定性。电子层的排列方式对这种稳定性起到了重要的作用。如果一个原子与其他原子相组合，这样它的电子层就被占满了，那么这就样就可以得到一个稳定的连接。价告诉我们要构成一个元素需要多少个氢原子。在构成水（H₂O）的时候，氧原子是二价。我们可以把元素按照价进行分类。按照 Mendeleev 的方法，带有一个特定价的元素被安排在元素周期表的同一个栏中。

价的概念被泰尼埃 *Tesnière* (1959) 用于语言学：一个中心语需要一定的论元以构成一个稳定的组合。具有相同价的词，即需要同样数目与种类的论元的词，被分成不同的类别。图 1.5 在下一页 分别展示了化学与语言学的例子。

我们用（67）来解释逻辑价。不过，逻辑价有时与句法价不同。对于动词 *rain*（下雨）来说，它需要一个虚指代词（expletive pronoun）作为论元。德语中的自反身动词也是一样的，如 *sich erholen*（复原）。



图 1.5: 氢原子和氧原子的组合与动词和其论元的组合

- (68) a. Es regnet.
它 下雨
‘下雨了。’
- b. Klaus erholt sich.
Klaus 恢复 REFL
‘Klaus 正在恢复健康。’

虚指的 es (它) 与表示天气的动词以及带 erholen (恢复) 这类带 sich 的内在反身动词都需要在句子里出现。日耳曼语言有虚指成分用来放在定式动词前的位置上。这些虚指成分并不能用在德语的嵌套句中, 因为嵌套句与常规的非嵌套的陈述句的结构不同, 即陈述句要求变位动词位于第二位。例 (69a) 说明了 es (它) 在 dass 引导的从句中不能省略。

- (69) a. *Ich glaube, dass regnet.
我 想 CONJ 下雨
想说: “我想是下雨了。”
- b. *Ich glaube, dass Klaus erholt.
我 想 CONJ Klaus 恢复
想说: “我相信 Klaus 在恢复之中。”

不管是虚指成分还是反身代词都对句子的语义没有贡献。但是, 为了构成一个完整的、合乎语法的句子, 它们必须出现。所以说, 它们也是动词的价的一部分。

那些对中心语对核心意义没有贡献的成分叫做说明语, 这些成分提供的是一些额外的信息。比如说例 (70) 中的副词 “深深”:

- (70) John 深深地爱着 Mary.

这里的副词说明了动词所描述的程度。此外, 还有属性形容词 (71a) 和关系从句 (71b) 的例子:

- (71) a. 一位漂亮的女人
b. the man who Mary loves
DET 男人 CONJ Mary 爱
‘Mary 爱着的那个男人’

说明语具有如下的句法和语义属性:

1 导言与术语

- (72) a. 说明语不构成语义角色。
b. 说明语是可选的。
c. 说明语可以重复。

例 (71a) 的短语可以通过增加说明语得到扩展:

- (73) 一位漂亮的聪明女人

如果我们先不考虑语言处理的问题, 这种通过增加说明语的方式可以无限扩展下去 (参看38) on page 60的讨论)。另一方面, 论元是不能多次实现的。

- (74) *The man the boy sleeps.
DET 男人DET 男孩睡觉

如果发出睡觉这个动作的实体已经被提及了, 那么就无法再用另一个名词短语来指称睡觉的个体。如果想要表达不只有一个个体在睡觉的话, 就必须采用例 (75) 中的并列式。

- (75) The man and the boy are sleeping.
DET 男人和 DET 男孩COP 睡觉
‘这个男人和这个男孩在睡觉。’

我们需要指出的是, 在72中提出的辨认说明语的标准并不充分, 因为还有句法论元不能充当语义角色。比如说68a中的 *es* (它), 68b中的 *sich* (自己), 以及例 (76) 中的可选成分, 如 “比萨饼”。

- (76) Tony 正在吃 (比萨饼)。

中心语通常以一种相对固定的方式决定其所带论元的句法属性。动词负责其所带动词的格属性。

- (77) a. Er gedenkt des Opfers.
他记得 DET.GEN 受害人.GEN
‘他记得受害人。’
b. *Er gedenkt dem Opfer.
他记得 DET.DAT 受害人
c. Er hilft dem Opfer.
他帮助DET.DAT 受害人
‘他帮助了受害人。’
d. *Er hilft des Opfers.
他帮助DET.GEN 受害人.GEN

动词管辖 (governs) 论元的格属性。介词短语中的介词和名词短语的格都由动词决定的:¹³

¹³ 相关例子参看 Eisenberg (1994b: 78).

- (78) a. Er denkt an seine Modelleisenbahn.
 他想 PERP 他的.ACC 火车模型
 ‘他在想他的火车模型。’
- b. # Er denkt an seiner Modelleisenbahn.
 他想 PREP 他的.DAT 火车模型
- c. Er hängt an seiner Modelleisenbahn.
 他附着 PREP 他的.DAT 火车模型
 ‘他紧贴着他的火车模型。’
- d. * Er hängt an seine Modelleisenbahn.
 他附着 PREP 他的.ACC 火车模型

另一方面，修饰介词短语的名词的格与他们的意义有关系。在德语中，例（79a）中表示趋向的介词短语通常要求其名词短语是第四格（宾格），而例（79b）中表示地点的介词短语则需要是第三格（与格）。

- (79) a. Er geht in die Schule / auf den Weihnachtsmarkt / unter
 他去 在 DET.ACC 学校 在 DET.ACC 圣诞市场 在..... 下
 die Brücke.
 DET.ACC 桥
 ‘他去学校 / 圣诞市场 / 桥下面。’
- b. Er schläft in der Schule / auf dem Weihnachtsmarkt / unter
 他睡觉 在 DET.DAT 学校 在 DET.DAT 圣诞市场 在..... 下
 der Brücke.
 DET.DAT 桥
 ‘他在学校 / 圣诞市场 / 桥下面睡觉。’

一个有趣的现象是，动词（位于）表示地点的信息。它不能单独使用，即无法在没有地点信息的情况下使用。

- (80) * Wir befinden uns.
 我们处于 REFL

这个信息的形式是不固定的，句法范畴或介词短语内的介词都不是限定的：

- (81) Wir befinden uns hier / unter der Brücke / neben dem Eingang / im
 我们处于 REFL 这儿 在..... 下 DET 桥 在..... 附近 DET 入口 在
 Bett.
 床
 ‘我们在这儿 / 桥下面 / 入口旁 / 床上。’

地点修饰词如 hier（这儿）或 unter der Brücke（桥下面）都可以看作是其他动词 schlafen（睡觉）的说明语。对于 sich befinden（位于）这类动词来说，我们更倾向于认为表示地点信息的格式说动词的句法论元的必有成分。动词选择表示地点信息的短语，但是并没有对其有任何句法上的限制。这种地点限制很像我们前面讲的说明语通过语义进行限制的方式。如果我只考虑中心语和说明语组合的语义层

1 导言与术语

面,那么我也会将附加语看作是修饰语(modifier)¹⁴。那些需要区分处所论元的动词,如 *sich befinden* (位于),也被看作是修饰语。修饰语通常是指说明语,所以说它也是可选的,然而在 *sich befinden* (位于)这个例子中,他们看起来是(必有)论元。

综上所述,我们可以说那些需要与中心语共现的句法成分是论元。并且,那些能够充当中心语的语义角色的句法成分也是论元。然而,这两类论元有时是可选的。

论元通常可以划分为主语和补足语。¹⁵不是所有的中心语都需要主语(参看 see Müller 2007b: Section 3.2)。由此,中心语所带论元的数量可以与中心语所带补足语的数量具有相关性。

1.7 语法功能

在有些理论中,诸如主语和宾语的语法功能构成了语言的形式化描述的一部分(如参考第7章有关词汇功能语法的内容)。但是本书中所讨论的主要理论并不这样看,这些术语被用来指称特定现象的非正式的描述。基于上述原因,我在下面的内容中进行简要的说明。

1.7.1 主语

尽管我认为读者对主语已经有了清晰的认识,但是给主语下一个跨语言的定义绝不是一件小事。对于德语来说,Reis (1982)提出下面的句法属性作为对主语对界定:

- 与变位动词构成主谓一致的关系
- 非名词性从句中的主格
- 在不定式中被省略(控制)
- 在祈使句中是可选的

我已经在例(4)中讨论过主谓一致的问题。Reis (1982)认为第二个要点适用于德语。她构建了非名词性从句的限制条件,因为名词作谓语的句子中可以有不止有一个名词性的论元成分,如例(82)所示:

- (82) a. Er ist ein Lügner.
他.NOM COP 一 骗子.NOM
'他是一个骗子。'
- b. Er wurde ein Lügner genannt.
他.NOM COP 一 骗子.NOM 叫做
'他被人叫做骗子。'

¹⁴ 参看第1.7.2节中更多有关状语的句法功能的内容。状语这个术语通常指与动词相关的成分。而修饰语是一个更为普遍的术语,通常还包括定语。

¹⁵ 有些学派认为补足语包括主语,即补足语的概念等同于论元(参看Groß 2003: 342)。有些学者将变位动词的主语看作是补足语(Pollard 1996b; Eisenberg 1994a: 376)

按照这一标准，德语中，诸如 *den Männern*（男人们）的与格论元不能作主语：

- (83) a. *Er hilft den Männern.*
 他 帮助DET.DAT 男人.DAT
 ‘他在帮助那个男人。’
 b. *Den Männern wurde geholfen.*
 DET.DAT 男人.DAT COP.3SG 帮助
 ‘那个男人被人帮助了。’

根据其他标准，与格也不应该被看作是主语（如Reis (1982)的观点）。在例（83b）中，*wurde*（想要）是一个第三人称单数的形式，不与 *den Männern*（男人们）搭配。前述指出的第三条标准有关不定式结构，如下例（mex1）所示：

- (84) a. *Klaus behauptet, den Männern zu helfen.*
 Klaus 声称 DET.DAT 男人.DAT AUX 帮助
 ‘Klaus 声称要帮助那个男人。’
 b. *Klaus behauptet, dass er den Männern hilft.*
 Klaus 声称 CONJ 他DET.DAT 男人.DAT 帮助
 ‘Klaus 声称他在帮助那个男人。’
 c. **Die Männer behaupten, geholfen zu werden.*
 DET 男人 声称 帮助 AUX COP
 想说：“那个男生声称得到了帮助。”
 d. **Die Männer behaupten, elegant getanzt zu werden.*
 DET 男人 声称 优雅地 跳舞 AUX COP
 想说：“那个男人声称在优雅地跳舞。”

在第一句中，动词 *helfen*（帮助）的论元被省略了。如果有人希望表达这个论元，那么久应该用例（84b）中 *dass* 引导的从句。例（84c,d）显示了不需要名词性论元的不定式不能嵌套在动词下，如 *behaupten*（声称）。如果与格名词短语（男人们）是（83b）的主语，我们应该回看到一个合乎语法的控制结构（84c）。但是，事实并非如此。与例（84c）不同的是，有必要用例（85）：

- (85) *Die Männer behaupten, dass ihnen geholfen wird.*
 DET 男人.NOM 声称 CONJ 他们.DAT 帮助 COP
 ‘这些男人们声称他们被帮助了。’

同理，祈使句也不能由不需要名词性成分的动词充当。例（86）列出了Reis (1982: 186)提出的一些例子。

- (86) a. *Fürchte dich nicht!*
 害怕 REFL 不
 ‘不要害怕！’

1 导言与术语

- b. *Graue nicht!
恐惧 不
想说：“不要感到恐惧！”
- c. Werd einmal unterstützt und ...
COP 一次 支持 和
‘让人支持你一次，并且...’
- d. *Werd einmal geholfen und ...
COP 一次 帮助 和
想说：“让人帮助你一次，并且...”

例（86a）中的动词（害怕）必须要带一个名词性的论元做主语（87a）。例（86b）中的与之类似的动词（怕）需要带一个与格论元（87b）。

- (87) a. Ich fürchte mich vor Spinnen.
我.NOM 害怕 REFL PERP 蜘蛛
‘我害怕蜘蛛。’
- b. Mir graut vor Spinnen.
我.DAT 害怕 PERP 蜘蛛
‘我怕蜘蛛。’

有趣的是，冰岛语中的与格论元表现不同。Zaenen et al. (1985) 讨论了冰岛语中主语的各种特征，并且可以在被动句中将与格论元看作是主语，即使变位动词与他们没有构成主谓一致的关系（第 3.1 节），或者他们并不必是主格。例如，下面就是带有被省略的与格论元的不定式结构（第 457 页）：

- (88) a. Ég vonast til að verða hjálpað.
我 希望 PERP AUX COP 帮助
‘我希望我能得到帮助。’
- b. Að vera hjálpað í prófinu er óleyfilegt.
AUX COP 帮助 PERP 考试 COP 不允许
‘在考试中是不允许被帮助的。’

在一些语法现象中，例（89）中的小句论元被看作是主语，因为它们能被主格的名词短语（90）所代替（参考 e.g., Eisenberg 2004: 63, 289）。

- (89) a. Dass er schon um sieben kommen wollte, stimmt nicht.
CONJ 他已经 在 七点 来 想 确定 不
‘他尽可能在七点赶来，这不是真的。’
- b. Dass er Maria geheiratet hat, gefällt mir.
CONJ 他Maria 娶 AUX 高兴 我
‘我很高兴他娶了 Maria。’

- (90) a. Das stimmt nicht.
DEM 确定 不
‘那不是真的。’
b. Das gefällt mir.
DEM 喜欢 我
‘我喜欢那个。’

需要指出的是，对于小句论元能否做主语这个问题上有不同的看法。最近发表的文献表明，在词汇功能语法中仍有相当多的讨论（参考第7章）（Dalrymple & Lødrup 2000; Berman 2003b, 2007; Alsina, Mohanan & Mohanan 2005; Forst 2006. ）。

如果我们知道如何界定主语，那么宾语的界定就不再困难了：宾语就是由给定中心语决定其形式的所有其他论元。以小句宾语为例，德语有属格、与格、宾格和介词宾语：

- (91) a. Sie gedenken des Mannes.
他们记得 DEF.GEN 人.GEN
‘他们记得这个人。’
b. Sie helfen dem Mann.
他们帮助 DEF.DAT 人.DAT
‘他们在帮助这个人。’
c. Sie kennen den Mann.
他们认识 DEF.ACC 人.ACC
‘他们认识这个人。’
d. Sie denken an den Mann.
他们想 PERP DET 人
‘他们正想起这个人。’

在对宾语按照格进行分类的同时，更为普遍的做法是区分直接宾语（direct objects）和间接宾语（indirect objects）。顾名思义，直接宾语与间接宾语不同，直接宾语的所指直接受到动词指定的动作的影响。带双宾语的动词，如德语的 *geben*（给），宾格宾语就是直接宾语，而与格宾语就是间接宾语。

- (92) dass er dem Mann den Aufsatz gibt
CONJ 他.NOM DEF.DAT 人.DAT DEF.ACC 论文.ACC 给
‘他给那个人这篇论文’

对于三价动词（带有三个论元的动词）来说，我们可以看到动词要么可以带一个属格宾语（93a），要么对于带宾格的直接宾语来说，再带一个宾格宾语（93b）：

- (93) a. dass er den Mann des Mordes bezichtigte
CONJ 他 DEF.ACC 人.ACC DEF.GEN 谋杀.GEN 控告
‘他控告了杀人的人’

1 导言与术语

- b. dass er den Mann den Vers lehrte
CONJ he DEF.ACC 人.ACC DEF.ACC 诗.ACC 教
‘他教那个人读了诗’

这类宾语有时也叫做间接宾语。通常，只有那些由 *werden* 引导的被动句中能够上升到主语位置上的成分才被看作是直接宾语。这对有些理论（如词汇功能语法，看第 7 章）来说非常重要，因为被动态被定义为语法功能。对于二元的动词性谓语来说，与格通常不被看作是直接宾语（Cook 2006）。

- (94) dass er dem Mann hilft
CONJ 他 DEF.DAT 人.DAT 帮助
‘他帮助那个人’

在很多理论中，语法功能并不是构成理论多原始成分，而是与树结构中的位置密切相关的。所以说，德语中的直接宾语在句法配置中首先与动词组合这个特点被认为是德语句子的底层结构。间接宾语是与动词组合的第二个宾语。按照这一观点，*helfen*（帮助）的与格宾语需要被看作是直接宾语。

下面，我就只用宾语的格属性来指称，而避免使用直接宾语和间接宾语的术语。我们对主语也采用相同的策略，有特定格属性的宾语小句能够分别对应于直接宾语或者间接宾语的语法功能。如果在例（95b）中，我们认为“*dass du sprichst*”（你在说话）这个小句是主语，那么从句就必然是直接宾语：

- (95) a. Dass du sprichst, wird erwähnt.
CONJ 你 说话 COP 提到
‘你在说话的事实被提及了。’
b. Er erwähnt, dass du sprichst.
他 提到 CONJ 你 说话
‘他提到你在说话。’

这种情况下，我们不能真的将从句看作是宾格宾语，因为它没有格属性。但是，我们用带有宾格标记的名词短语来替换这个句子：

- (96) Er erwähnt diesen Sachverhalt.
他 提到 这.ACC 事情
‘他提到这个事情。’

如果我们不想讨论这一问题，就可以简单地把这些论元称为小句宾语。

1.7.2 状语

状语与主语和宾语在语义上有很大的不同。它们告诉我们有关动作或过程发生地方的信息，或者是按照何种状态进行的方式。在大部分情况下，状语是说明语，但是，正如我们已经看到的，有些中心语也要有状语。这些动词的例子有 *to be located*（位于）或者 *to make one's way*（让路）。以 *to be located*（位于）为例，有必要明确一个地点，还需要一个前进的方向。由此，这类状语被看作是动词的论元。

状语这个术语的来源是因为大部分状语都是副词。但是，还有其他情况。介词、助词、介词短语、名词短语，甚至是句子都可以充当状语：

- (97)
- a. Er arbeitet sorgfältig.
他工作 认真
 - b. Er arbeitet vergleichend.
他工作 比较
‘他做比较的工作。’
 - c. Er arbeitet in der Universität.
他工作 在DET 大学
‘他在大学工作。’
 - d. Er arbeitet den ganzen Tag.
他工作 DET 整 天.ACC
‘他整天工作。’
 - e. Er arbeitet, weil es ihm Spaß macht.
他工作 因为它他.DAT 乐趣 作
‘他工作因为他喜欢工作。’

尽管例(97d)中的名词短语带有宾格，它并不是宾格宾格。*den ganzen Tag*（整天）号称是时间宾格。这种情况下宾格的出现与名词短语的句法和语义功能有关，它不是由动词决定的。这类宾格可以跟许多动词共现，甚至是那些通常不需要宾格宾语的动词：

- (98)
- a. Er schläft den ganzen Tag.
他睡觉 DET 整 天
‘他睡了一整天。’
 - b. Er liest den ganzen Tag diesen schwierigen Aufsatz.
他读 DEF.ACC 整.ACC 天 这.ACC 难的.ACC 论文
‘他花了一整天读这篇难懂的论文。’
 - c. Er gibt den Armen den ganzen Tag Suppe.
他给 DEF.DAT 穷人.DAT DEF.ACC 整天.ACC 天 汤
‘他花了一整天给穷人汤喝。’

被动态中状语的格不发生变化：

- (99)
- a. weil den ganzen Tag gearbeitet wurde
因为DEF.ACC 整.ACC 天 工作 AUX
‘因为有人整天工作’
 - b. *weil der ganze Tag gearbeitet wurde
因为DEF.NOM whole.NOM 天 工作 AUX

1 导言与术语

1.7.3 谓语

例 (100a,b) 中的形容词与例 (100c) 中的名词短语被看作是谓语。

- (100) a. Klaus ist *klug*.
Klaus COP 聪明
b. Er isst den Fisch *roh*.
他吃 DET 鱼 生的
c. Er ist *ein Lügner*.
他 COP 一 骗子

在例 (100a,c) 的系词结构中, 形容词 *klug* (聪明) 与名词短语 *ein Lügner* (一个骗子) 都是系词 *sein* (是) 的论元, 而例 (100b) 中的描述性形容词则作 *isst* (吃) 的状语。

对于名词性谓语来说, 格不是由中心语决定的, 而是由其他成分决定的。¹⁶
例如, 例 (101a) 中的宾格在例 (101b) 的被动句中变成了主格。

- (101) a. Sie nannte ihn *einen Lügner*.
她 叫 他.ACC 一.ACC 骗子
“她把他叫做骗子。”
b. Er *wurde ein Lügner genannt*.
他.NOM AUX 一.NOM 骗子 叫做
“他被人叫做骗子。”

¹⁶ 不同方言区的系词结构有所不同: 在标准德语中系词 *sein* (是) 所带的名词短语总是主格, 即使是嵌套在 *lassen* (让) 的下面也不发生变化。根据 Drosdowski (1995: § 1259), 瑞士地区则经常能发现例 (ii.a) 中用作宾格的情况。

- (i) a. Ich bin *dein Tanzpartner*.
I COP 你的.NOM 舞伴
b. Der *wüste Kerl* ist *ihr Komplize*.
DET 疯狂的家伙 COP 她的.NOM 从犯
c. Laß den *wüsten Kerl* [...] *meinetwegen ihr Komplize sein*.
让 DEF.ACC 疯狂的.ACC 家伙 我所关心的 她的.NOM 从犯 COP
‘我所关心的是, 让我们来假定那个疯狂的家伙是她的从犯。’ (Grebe & Gipper 1966: § 6925)
d. Baby, laß mich *dein Tanzpartner sein*.
宝贝 让 我.ACC 你的.NOM 舞伴 COP
‘宝贝, 让我当你的舞伴吧!’ (Funny van Dannen, Benno-Ohnesorg-Theater, Berlin, Volksbühne, 1995 年 10 月 11 日)
- (ii) a. Er *lässt den lieben Gott 'n frommen Mann sein*.
他让 DEF.ACC 亲爱的.ACC 上帝一虔诚的.ACC 人 COP
‘他全是不走心 (漠不关心) 啊。’
b. *Er *lässt den lieben Gott 'n frommer Mann sein*.
他让 DEF.ACC 亲爱的.ACC 上帝一虔诚的.NOM 人 COP

例(101a)中,只有 ihn (他)被描述成宾语。在例(101b)中,ihn (他)变成了主语,也就成为了主格。在例(101a)中,einen Lügner (一个骗子)指代 ihn (他),例(101b)中的 er (他)需要与作谓语的名词的格保持一致。这也叫做格的一致关系。

其他的谓词性结构可以参见Eisenberg et al. (2005: § 1206)、Müller (2002a: Chapter 4, Chapter 5), 以及Müller (2008a)。

1.7.4 配价类型

我们可以按照动词所带论元成分的数量与属性来对动词进行分类。一方面,那些可以带宾语且其宾语能变换为被动式的主语多动词叫做及物动词(transitive),如 love (爱)或 beat (击打)这类动词。另一方面,不及物动词不能带宾语,或者其在被动式中不能变成主语,如 schlafen (睡觉), helfen (帮忙)或 gedenken (纪念)。及物动词还包括双及物动词,如 geben (给)和 zeigen (展示)。

不过,这一术语的使用并不完全一致。有时,带与格和属格宾语的二位动词也被看作是及物动词。在这个命名系统中,不及物动词、及物动词以及双及物动词的术语与一位动词、二位动词和三位动词的术语在含义上是相同的。

这种术语混淆的情况导致对乔姆斯基的评论的误解,即使他们是由一些知名的语言学家(Culicover and Jackendoff's (2005: 59))提出来的。乔姆斯基指出,英语助动词“be”带动词的被动态只能用在及物动词上。Culicover 和 Jackendoff 则认为这是不对的,因为还有及物动词不能变换为被动式,如 weigh (称重)和 cost (花费)。

- (102) a. This book weighs ten pounds / costs ten dollars.
 这 书 重 十 磅 / 花费 十 美元
 ‘这本书重十磅/卖十美元。’
 b. *Ten pounds are weighed / ten dollar are cost by this book.
 十 磅 AUX 重 / 十 美元 AUX 花费PREP 这 书

Culicover 和 Jackendoff 这里说的及物动词是指带两个论元的动词。如果我们只把那些动词宾语可以变换为被动式主语动词看作是及物动词,那么 weigh (称重)和 cost (花费)都不能算是及物动词,这样 Culicover 和 Jackendoff 的观点就站不脚了。¹⁷ 诸如例(102)中的名词短语不是普通宾语,因为它们不能被代词替代。所以说它们的格属性无法确定,因为英语中只有代词有格的区分。如果我们将英语的例子翻译成德语,我们会发现它们是宾格宾语:

- (103) a. Das Buch kostete einen Dollar.
 DET 书 值 一.ACC 美元
 ‘这本书值一美元。’
 b. Das Buch wiegt einen Zentner.
 DET 书 重 一.ACC 公担
 ‘这本书重五十公斤。’

¹⁷ 即使我们将及物动词看作是二位谓词的话,他们的评论也是站不住的。如果我们认为动词至少带两个论元成分才能变换成被动式的话,那么我们也必须要认定所有带两个或两个以上论元的动词都可以变换成被动式。实际上,带多个论元这个属性只是一个充分条件,并不是唯一条件。

1 导言与术语

在下面，我会采用及物的前一个概念，即那些能在被动式中将宾语变换为主语的动词。对于 *helfen*（帮助）类动词，它们带一个主格和一个与格论元；而 *schlagen*（击打）类动词带有一个主格和一个宾格论元。我会用二位动词或二价动词（bivalent verb）动词来指称它们。

1.8 德语小句的空间位置模型

在这一节，我会介绍所谓的空间位置的概念。这一概念会在后续的章节中被经常用来讨论德语小句中的不同部分。此外，也可以在Reis (1980), Höhle (1986) and Askedal (1986)的文献中找到更多有关空间位置的知识。Wöllstein (2010)是一本有关空间位置模型的教科书。

1.8.1 动词的位置

最常见的是按照变位动词的位置将德语的句子分成三类：

- 动词位于末位的小句
- 动词位于首位的小句
- 动词位于第二位（V2）的小句

下面的例子说明了这些可能性：

- (104) a. (Peter *hat* erzählt,) dass er das Eis *gegessen hat*.
Peter AUX 告诉 CONJ 他DET 雪糕吃 AUX
‘Peter 说他把雪糕吃完了。’
- b. *Hat* Peter das Eis *gegessen*?
AUX Peter DET 雪糕吃
‘Peter 吃完雪糕了吗？’
- c. Peter *hat* das Eis *gegessen*.
Peter AUX DET 雪糕吃
‘Peter 吃完雪糕了。’

1.8.2 句子的框架结构、前场、中场及后场

我们观察到，在例（104a）中，变位动词 *hat* 只跟它的补足语 *gegessen*（吃完）挨着。在例（104b）和例（104c）中，动词跟它的补足语是分开的，也就是非连续的。这样，我们可以基于这些区别将德语小句分成不同的小类。在例（104b）和例（104c）中，动词和助词构成小句的一个基本框架。基于这个原因，我们称之为框型结构（sentence bracket）。例（104b）和例（104c）中的变位动词构成框架的左边界，非变位动词构成右边界。以动词为末位的小句通常由连词引入，如 *weil*（因为）、*dass*（这个），以及 *ob*（是否）。不管是在动词占首位，还是动词占末尾到小句中，这些连词占据了与变位动词一样的位置。由此，我们认为这些连词也构成了

这些句子的左边界。应用句子框架的概念，使得我们有可能将德语小句划分成前场、中场和后场这三个部分。前场指位于左边界之前的成分，中场是指位于左边界与右边界之间的成分，后场指右边界之后的成分。如图 1.1 on the following page 所示：右边界可以包括多个动词，通常被看作是动词性复合词（verbal complex）或动词词组（verb cluster）。下一节，我们将讨论疑问词和关系代词在前场的排列。

1.8.3 场内元素的排列

如图中的例子所示，并不是所有的位置都需要有成分来充当。如果想在例（105）中省略系词（copula）sein 的话，即使是左边界，也可以是空的：

- (105) a. [...] egal, was noch passiert, der Norddeutsche Rundfunk steht schon
 无论如何什么仍然发生 DET 北德 广播公司 存在 已经
 jetzt als Gewinner fest.¹⁸
 现在PREP 赢家 PART
 ‘不管将要发生什么，北德广播公司早就已经是赢家了。’
- b. Interessant, zu erwähnen, daß ihre Seele völlig in Ordnung war.¹⁹
 有趣 AUX 提到 CONJ 她的灵魂 完整地PREP 顺序 COP
 ‘值得注意的是，她的灵魂完好无损。’
- c. Ein Treppenwitz der Musikgeschichte, daß die Kollegen von Rammstein
 一 讽刺 DET 音乐历史 CONJ DET 成员 PREP Rammstein
 vor fünf Jahren noch im Vorprogramm von Sandow spielten.²⁰
 PREP 五 年 仍然 PREP.DET 开场 PREP Sandow 表演
 ‘音乐历史上具有讽刺意味的一件事是，五年前 Rammstein 的成员仍
 然为 Sandow 作开场演出。’

例（105）中的例子与例（106）中的系词具有相关性：

- (106) a. Egal ist, was noch passiert, ...
 不管COP 什么仍然 发生
 ‘今后将发生什么并不重要 ...’
- b. Interessant ist zu erwähnen, dass ihre Seele völlig in Ordnung
 有趣 COP AUX 提到 CONJ 她的灵魂 完整地PREP 顺序
 war.
 COP
 ‘值得注意的是，她的灵魂完好无损。’

¹⁸ 《明镜周刊》（*Spiegel*），1999 年 12 月，p. 258.

¹⁹ Michail Bulgakow, *Der Meister und Margarita*. München: Deutscher Taschenbuch Verlag. 1997, p. 422.

²⁰ Flüstern & Schweigen, taz, 12.07.1999, p. 14.

表 1.1: 空间位置分布例示

前场	左边界	中场	右边界	后场
Karl	schläft.			
Karl	hat		geschlafen.	
Karl	erkennt	Maria.		
Karl	färbt	den Mantel		den Maria kennt.
Karl	hat	Maria	um	
Karl	hat	Maria als sie aus dem Zug stieg sofort	erkannt.	
Karl	hat	Maria sofort	erkannt.	als sie aus dem Zug stieg.
Karl	hat	Maria zu erkennen	behauptet.	
Karl	hat		behauptet	Maria zu erkennen.
	Schläft	Karl?		
	Schlaf!			
	Iss	jetzt dein Eis	auf!	
	Hat	er doch das ganze Eis alleine	gegessen.	
	weil	er das ganze Eis alleine	gegessen hat	ohne mit der Wimper zu zucken.
wer	weil	er das ganze Eis alleine	essen können will	ohne gestört zu werden.
der		das ganze Eis alleine	gegessen hat	
mit wem		das ganze Eis alleine	gegessen hat	
mit dem		du	geredet hast	
		du	geredet hast	

- c. Ein Treppenwitz der Musikgeschichte ist, dass die Kollegen von
 一 讽刺 DET 音乐历史 COP CONJ DET 成员 PREP
 Rammstein vor fünf Jahren noch im Vorprogramm von Sandow
 Rammstein PREP 五 年 仍然 PREP.DET 开场 PREP Sandow
 spielten.
 表演

‘五年前，Rammstein 的成员仍为 Sandow 作开场演出是音乐史上的一件具有讽刺意义的事件。’

当某些位置为空时，就不太容易判断句中成分占据了哪些位置。如例（105）所示，我们需要将系词插入来确保哪个成分位于前场，以及其他成分所处的位置。

在下面Paul (1919: 13)引用的例子中，插入系词则得到了不同的结果：

- (107) a. Niemand da?
 没有人 那儿
 b. Ist niemand da?
 COP 没有人 那儿
 ‘没有人在那儿吗？’

这里，我们要分析的是一个问句，所以 niemand（没有人）不能分析为位于前场，而应是位于中场。

例（108）中，前场、左边界和中场都有成分充当，而右边界是空的。²¹

- (108) Er gibt der Frau das Buch, die er kennt.
 他.M 给 DET 女人 (F) DET 书.(N) CONJ.F 他认识
 ‘他把书交给了他认识的那个女人。’

我们应该怎么分析诸如“die er kennt”（他认识的吗？）这样的关系从句呢？它们构成了中场还是后场呢？我们可以应用Bech (1955: 72)开发的等级测试法（Rangprobe）来进行测试：首先，我们把例（108）改成完成时。由于非变位动词位于右边界，我们可以清楚地看出中场与后场的边界。例（109）中的例子显示了，关系从句不能位于中场，除非它是与中心语 Frau（女人）一起构成的一个复杂成分的一部分。

- (109) a. Er hat [der Frau] das Buch gegeben, [die er kennt].
 他 AUX DET 女人 DET 书 给 CONJ 他认识
 ‘他把书交给了他认识的那个女人。’
 b. *Er hat [der Frau] das Buch, [die er kennt,] gegeben.
 他 AUX DET 女人 DET 书 CONJ 他认识 给
 c. Er hat [der Frau, die er kennt,] das Buch gegeben.
 他 AUX DET 女人 CONJ 他认识 DET 书 给

如果修饰中心语名词的关系从句像例（110）这种位于句末的话，这个测试也没有用了。

²¹ 这个句子需要对 der（这个）进行强调。der Frau, die er kennt（他认识的这个女人）与另一个或其他女人进行区分。

1 导言与术语

- (110) Er gibt das Buch der Frau, die er kennt.
他给 DET 书 DET 女人 CONJ 他认识
‘他把书交给了他认识的那个女人。’

如果我们把例(110)的句子变成完成时,那么我们可以发现,动词可以出现在关系小句的前面或者后面:

- (111) a. Er hat das Buch [der Frau] gegeben, [die er kennt].
他 AUX DET 书 DET 女人 给 CONJ 他认识
‘他把书交给了他认识的那个女人。’
b. Er hat das Buch [der Frau, die er kennt,] gegeben.
他 AUX DET 书 DET 女人 CONJ 他认识 给

在例(111a)中,关系小句被提前了。在(111b)中,它构成了名词短语 *der Frau, die er kennt* (他认识的那个女人)的一部分,并位于中场的名词短语内部。这样,上面的测试方法也不适用于例(110)了。我们认为例(110)中的关系小句也术语名词短语,因为这样是最简单的结构。如果关系小句位于后场,我们则需要假定它从名词短语内部上升到这个位置上了。也就是说,我们需要假设其是一个名词短语结构,并且还涉及了移外变形。对于疑问词和冠词代词来说也有同样的问题。不同作者看法不同,有的认为它们位于左边界,有的认为是前场,也有的认为是中场。在标准德语中,疑问句和关系小句中的所有成分从来没有占全的时候。这就导致,无法直观地分清楚某个元素所在的位置。尽管如此,我们可以结合主句来判断:疑问句和关系小句中的代词可以包含在复杂短语中:

- (112) a. der Mann, [mit dem] du gesprochen hast
DET 人 PREP 谁 你 说话 AUX
‘你跟他说话的那个人’
b. Ich möchte wissen, [mit wem] du gesprochen hast.
我 想 知道 PREP 谁 你 说话 AUX
‘我想知道你在跟谁说话。’

通常,只有个别词语(连词或动词)能够占据左边界,²²而词和短语可以出现在前场中。由此可以推断,疑问词和关系代词(及其包含的短语)也能在这个位置上出现。

进而,我们观察到,在陈述句的前场位置的成分与句中其他成分的依存关系跟包括关系代词的短语与句中其他成分之间的依存关系是一样的。比如说,例(113a)中 *about this topic* (关于这个话题)依存于深深地嵌套在句中的 *Vortrag* (演讲): *einen Vortrag* (一个演讲)是 *zu halten* (发布)的一个论元成分,也是 *gebeten* (要求)的一个论元成分。

²² 并列结构是一个例外:

- (i) Er [kennt und liebt] diese Schallplatte.
他 知道 和 热爱 这 专辑
‘他了解并且热爱这张专辑。’

- (113) a. Über dieses Thema habe ich ihn gebeten, einen Vortrag zu halten.
关于这 题目 AUX 我 他 请求 一 演讲 AUX hold
‘我请他做个有关这个题目的演讲。’
- b. das Thema, über das ich ihn gebeten habe, einen Vortrag zu halten
DET 题目 关于我 他 请求AUX 一 演讲 AUX 作
‘这是那个我让他做一个演讲的题目’

例 (113b) 的情景是类似的：关系短语 *über das*（关于这个）是 *Vortrag*（演讲）的一个与之关系较远的依存成分。所以说，如果关系短语被放在前场，我们可以推断出，远距离的前置总是指向前场。

最后来看一下《杜登大辞典》(Eisenberg et al. 2005: §1347) 中应用的标准德语（大部分是南方方言）中的一些例子：

- (114) a. Kommt drauf an, mit wem dass sie zu tun haben.
来 PREP PART PREP 谁 CONJ 你 AUX 做 AUX
‘这要取决于是谁在处理。’
- (115) a. Lotti, die wo eine tolle Sekretärin ist, hat ein paar merkwürdige
Lotti 谁 哪儿 一 伟大秘书 COP 有 一 些 奇怪
Herren empfangen.
绅士 欢迎
‘Lotti 作为一名伟大的秘书，受到了一些奇怪的绅士的欢迎。’
- b. Du bist der beste Sänger, den wo ich kenn.
你 COP DET 最棒 歌手 谁 哪儿我 知道
‘你是我知道的最棒的歌手。’

这些有关疑问句和关系从句的例子说明句子的左边界在不同方言中分别由连词 *dass*（这个）和 *wo*（哪儿）充当。所以说，如果我们希望得到一个能够统一分析标准德语与方言的模型的话，就应该认为关系短语和疑问短语位于前场。和关系小句

1.8.4 递归

正如Reis (1980: 82)已经指出的，当前场由一个复杂成分充当的时候，可以将其进一步细化。例如，(116b) 中的 *für lange lange Zeit*（很长很长时间）和 (116d) 中的 *daß du kommst*（你要进来）都位于前场，但是出现在右边界 *gewußt*（知道）的右边，也就是说他们出现在前场内的后场。

- (116) a. Die Möglichkeit, etwas zu verändern, ist damit verschüttet für
DET 机会 东西 AUX 变化 COP 在那儿消失 PREP
lange lange Zeit.
长 长 时间
‘变革的机会将要消失很长很长时间了。’

1 导言与术语

- b. [Verschüttet für lange lange Zeit] ist damit die Möglichkeit, etwas
消失 PREP 长 长 时间 COP 在那儿DET 机会 东西
zu verändern.
AUX 变化
- c. Wir haben schon seit langem gewußt, daß du kommst.
我们AUX PART CONJ 长 知道 CONJ 你 来
‘我们早就知道你要来了。’
- d. [Gewußt, daß du kommst,] haben wir schon seit langem.
知道 CONJ 你 来 AUX 我们PART PREP 长

正如在前场的成分一样，中场和后场的成分也能有内部的结构，并且可以相应地划分到不同的次级结构中。如例（116c）中，daß（这个）是从句 daß du kommst（他知道）的左边界，而 du（你）占据了中场，kommst（来）是右边界。

思考题

1. 短语中的中心语与其他成分相比有何不同？
2. 请找出例（117）中的中心语：
(117) a. 他
b. 走！
c. 快
3. 论元与附加语有什么区别？
4. 请指出例（118）所示的整句及其小句中的中心语、论元和附加语：

- (118) Er hilft den kleinen Kindern in der Schule.
他 帮助DET 小 孩子 在DET 学校
‘他在学校帮助小孩子。’

5. 如何界定前场、中场、后场、左边界与右边界？

练习题

1. 请指出下列句子及其嵌套对小句的框架结构、前场、中场，以及后场。

- (119) a. Karl isst.
Karl 吃
‘Karl 正在吃。’

- b. Der Mann liebt eine Frau, den Peter kennt.
 DET 男人 爱 一 女人 谁 Peter 认识
 ‘Peter 认识的那个男人爱着这个女人。’
- c. Der Mann liebt eine Frau, die Peter kennt.
 DET 男人 爱 一 女人 CONJ Peter 认识
 ‘这个男人爱着那个彼得认识的女人。’
- d. Die Studenten haben behauptet, nur wegen der Hitze einzuschlafen.
 DET 学生们 AUX 声称 只 因为 DET 热
 睡着了
 ‘学生们声称他们是因为热才睡着了。’
- e. Dass Peter nicht kommt, ärgert Klaus.
 CONJ Peter 不 来 惹恼 Klaus
 ‘Peter 不能来的事实惹恼了 Klaus。’
- f. Einen Mann küssen, der ihr nicht gefällt, würde sie nie.
 一 男人 亲 CONJ 她 不 喜欢 AUX 她 从不
 ‘她绝不会亲一个她不喜欢的男人。’

延伸阅读

Reis (1980)解释了为什么空间位置理论对描述德语中成分之间关系如此重要。Höhle (1986)讨论了前场左边的位置由左移位结构充当，如例 (120) der Mittwoch (星期三)，例 (121a) 的 aber (但是)，以及例 (121b) 的 denn (那么)。

- (120) Der Mittwoch, der passt mir gut.
 DET 星期三 CONJ 适合 我 好
 ‘周三，我有空。’
- (121) a. Aber würde denn jemand den Hund füttern morgen Abend?
 但是 AUX PART 有人 DET 狗 喂 明天 晚上
 ‘但是明天晚上会有人喂狗吗？’
- b. Denn dass es regnet, damit rechnet keiner.
 因为 CONJ 它下雨 对此 想到 没有人
 ‘因为没人想到会下雨。’

Höhle 还讨论了空间位置理论的历史发展。

第二章 短语结构语法

本章讨论短语结构语法，该语法在我们接下来要讲解的理论中起到了至关重要的作用。

2.1 符号与重写规则

根据词的屈折形态和句法分布，我们可以判断出它们的词性。由此，例（1）中的 **weil**（因为）是一个连词，而 **das**（这）和 **dem**（这）是冠词，后两者也叫做限定词。进而，**Buch**（书）和 **Mann**（人）是名词，而 **gibt**（给）是动词。

- （1） **weil er das Buch dem Mann gibt**
因为他DET 书 DET 人 给
‘因为他给人这本书’

在第 1.3 节，我们介绍了组成成分测试的若干方法。采用这些方法，我们可以判断出 **das Buch**（这本书）和 **dem Mann**（这个人）这两组字符串分别构成了组成成分。那么这些成分需要特定的符号来指称他们。因为在这些短语中，名词起到了重要的作用，它们就叫做名词短语（**noun phrase**）或简称为 **NPs**。因为代词 **er**（他）也可以出现在完整的名词短语所在的位置中，所以代词也应归入 **NP** 这一类别中。

短语结构语法用规则来说明符号是如何指派到某类词中，并且这些词是如何构成更复杂的单位的。如（2）所示，这是用来分析例句（1）的一个简单的短语结构语法：^{1,2}

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

由此，我们将 $NP \rightarrow Det\ N$ 这一规则解读为一个名词短语，这个名词短语被赋予了 **NP** 这个符号，并且包括一个限定词（**D**）和一个名词（**N**）。

¹ 我们暂不分析 **weil**（因为）。因为对于德语的动词位于第一位和动词位于第二位的小句的确切分析来说，我们还需要一些额外的假设，所以在这一章只讨论动词位于句末的小句。

² $NP \rightarrow er$ 这条规则看起来有些奇怪。我们本可以用 $PersPron \rightarrow er$ 这条规则来说明，但是这就需要另一条规则来说明能够代替整个 **NP** 的人称代词 $NP \rightarrow PersPron$ 。（2）将前面两条规则整合成一条，并且说明 **er**（他）出现在名词短语可以出现的位置上。

2 短语结构语法

对于例（1）这个句子，我们可以应用（2）中的文法并按照如下的方式来分析：首先，我们选取句中的第一个词，然后看是否有规则来说明该词出现在规则的右边。如果有，那么我们就用规则左边的符号来替换它。这一过程可以参考（3）中的第 2-4，6-7 和 9 行的推导过程。比如说，第二行中 er 被 NP 代替。如果在规则的右边有两个以上的符号，那么这些符号都被左边的符号所代替。这一过程可以参考第 5、8 和 10 行。例如，在第 5 行和第 8 行中，Det 和 N 被重写为 NP。

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

如例（3）所示，我们从一串词开始，然后按照给定的短语结构文法，我们可以推导出句子的结构。我们也可以按照相反的方向来分析：从句子标号 S 开始，我们可以应用第 9-1 步，最后到词串。如要从文法中选择重写符号的不同规则，我们可以用（2）中的文法来从 S 分析到字符串“er dem Mann das Buch gibt”（他男人书给）。我们可以说，该文法允准（或生成）了一组句子。

例（3）中的推导过程可以表示为一棵树，如图 2.1 所示。

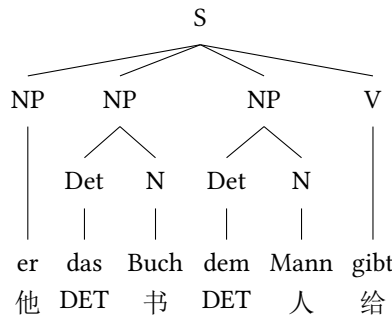


图 2.1: “er das Buch dem Mann gibt”（他这本书这个人给）的分析

树中的符号叫做节点（node）。S 直接统制 NP 节点和 V 节点。树中的其他节点也被统制，只不过不是受到 S 的直接统制。如果我们想讨论节点之间的关系，我们通常会用亲属词来表示。在图 2.1 中，S 是三个 NP 节点和 V 节点的母节点（mother node）。NP 节点和 V 节点是姊妹节点，因为他们有共同的母节点。如果一个节点有两个子节点，那么我们会得到一个二叉树结构。如果只有一个子节点，则会

得到一个单叉树结构。如果有两个成分是直接相连的，我们就说它们具有邻接性 (adjacent)。

语言学著作中通常会省略短语结构规则。相反，作者们倾向于使用树图或者紧缩的框式结构，如 (4) 所示。

- (4) [S [NP er] [NP [Det das] [N Buch]] [NP [Det dem] [N Mann]] [V gibt]]
 he the book the man gives

无论采取哪种形式，文法规则是最重要的，因为这些规则表示了语法知识，而这是与具体的结构没有太多关系的。这样，我们可以利用 (2) 中的文法来剖析或生成例 (5)，该句中的宾语顺序与前面例 (1) 的不同：

- (5) [weil] er dem Mann das Buch gibt
 因为 他.NOM DET.DAT 人 DET.ACC 书 给
 ‘因为他给这个人这本书’

与例 (1) 相比，这句话中用来替代限定词和名词的规则采用了不同的顺序。具体来说，这里并不是用 *das* (这) 来替代第一个 Det，也不是用 *Buch* (书) 来替代第一个名词，而是用 *dem* (这) 来替代第一个 Det，并用 *Mann* (人) 替代第一个名词。

在此，我需要指出的是，(2) 中的文法并不是针对 (1) 中的例句的唯一一套文法。实际上，对于这类句子 (参看练习1) 而言，还有无限多可能的文法。另一套文法如例 (6) 所示：

- (6) NP → Det N NP → er N → Buch
 V → NP V Det → das N → Mann
 Det → dem V → gibt

该文法允许了如图 2.2 on the next page 所示的二叉树结构。

实际上，(6) 和 (2) 中的文法都不够精确。如果我们在文法中加入额外的词条，如 *ich* (我) 和 *den* (DET 的宾格形式)，那么就会生成不合乎语法的句子，如例 (7b-d) 所示：³

- (7) a. er das Buch dem Mann gibt
 他.NOM DET.ACC 书 DET.DAT 人 给
 ‘他把书给这个人。’
 b. *ich das Buch dem Mann gibt
 我.NOM DET.ACC 书 DET.DAT 人 给

³ 应用 (6) 中的文法，我们还会遇到别的问题，即我们无法确定什么时候话语是完整的，因为符号 V 被用来表示 V 和 NP 的所有组合。由此，我们也可以用这个文法来分析 (i) 中的句子：

- (i) a. *der Mann erwartet
 DET 人 希望
 b. *des Mannes er das Buch dem Mann gibt
 DET.GEN 人.GEN 他.NOM DET.ACC 书 DET.DAT 人 给

动词所带的论元数量应该在文法中显示出来。在下面的章节中，我们将介绍动词对论元的选择 (价) 在不同的语法理论中是如何表述的。

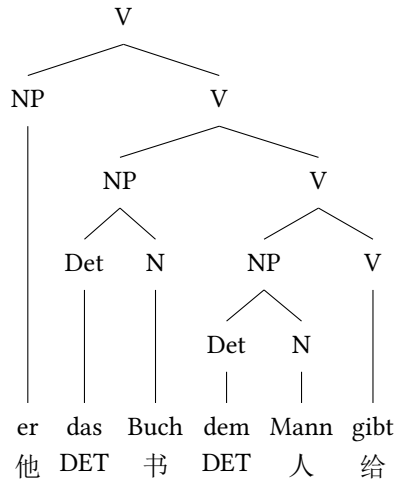


图 2.2: “er das Buch dem Mann gibt” (他这本书这个人给) 的二叉树结构分析

- c. *er das Buch den Mann gibt
 他.NOM DET.ACC 书 DET.ACC 人 给
- d. *er den Buch dem Mann gibt
 他.NOM DET.M 书 (N) DET 人 给

例 (7b) 违反了主谓一致原则，具体来说：ich (我) 和 gibt (给) 并不搭配。例 (7c) 是不合乎语法的，因为没有满足动词对格的约束条件。最后，例 (7d) 也是不合乎语法的，因为在限定词和名词之间缺少一致关系。我们不能将 den (这) 这个阳性的宾格成分与 Buch (书) 这个中性词搭配。基于上述原因，这两个成分的性质不同，所以说，二者不能搭配。

在下面的内容中，我们将来考虑如何让我们的语法可以杜绝生成例 (7b-d) 中的句子。如果我们想说明主谓一致关系，那么我们就需要说明德语中的六种格，并且动词必须与主语在人称 (1、2、3) 和数 (sg、pl) 上保持一致：

- (8) a. Ich schlafe. (1, sg)
 我 睡觉
- b. Du schläfst. (2, sg)
 你 睡觉
- c. Er schläft. (3, sg)
 他 睡觉
- d. Wir schlafen. (1, pl)
 我们睡觉
- e. Ihr schlaft. (2, pl)
 你们睡觉

- f. Sie schlafen. (3, pl)
他们睡觉

我们可以通过增加符号的数量来用语法规则描述这些关系。我们用下面的符号来替换 $S \rightarrow NP\ NP\ NP\ V$:

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

这样就意味着，我们需要六种不同的符号来分别表示名词短语和动词，还需要六条规则，而不是一条。

为了说明动词的格指派关系，我们可以按照类似的方式将格信息整合到符号中。这样，我们可以得到如下的规则：

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

考虑到名词需要区分四种不同的格，我们给主格 NP 设置了六种符号，其他格的 NP 需要三种符号。由于动词需要与 NP 相搭配，我们还需要区分 (11) 中带三个论元，或者只带一个或两个论元的动词，所以说，我们需要增加表示动词的符号数量。

- (11) a. Er schläft.
他睡觉
‘他正在睡觉。’
 b. *Er schläft das Buch.
他睡觉 DET 书
 c. Er kennt das Buch.
他认识 DET 书
‘他认识这本书。’
 d. *Er kennt.
他认识

在上面的规则中，有关动词所需论元数量的信息在 “nom_dat_acc” 这个标记中有所说明。

2 短语结构语法

为了说明例(12)中的限定词与名词的一致关系,我们需要整合性(阴性、阳性、中性)、数(单数、复数)、格(主格、属格、与格、宾格),以及屈折类型(强、弱)的信息⁴。

- (12) a. der Mann, die Frau, das Buch (性)
DET.M 人(M) DET.F 女人(F) DET.N 书(N)
b. das Buch, die Bücher (数)
DET 书.SG DET 书.PL
c. des Buches, dem Buch (格)
DET.GEN 书.GEN DET.DAT 书
d. ein Beamter, der Beamte (屈折类型)
一 公务员 DET 公务员

我们不用 $NP \rightarrow Det\ N$ 这条规则,我们将要用到下面这些规则,如例(13)所示:⁵

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

例(13)说明了主格名词短语的规则。我们还需要针对属格、与格和宾格的类似规则。这样,我们就需要24个符号来说明限定词($3 \times 2 \times 4$),24个符号说明名词,以及24条规则,而不是一条规则。如果要考虑屈折类型的话,符号的数量和规则的数量都要翻倍。

2.2 短语结构语法中特征的运用

短语结构语法如果只用原子式的符号是有问题的,因为他们无法描写一些特定的扩展形式。语言学家能够辨识出 $NP_3_sg_nom$ 指的是名词短语,因为这里面包括字母NP。但是,在形式化的术语中,这个符号与文法中的其他任意一个符号是一样的。我们并不能找到可以用来指代NP的所有符号的共同特征。进而,非结

⁴ 这些是形容词的屈折变化类型,它们也同样适用于某些名词,如 Beamter (公务员)、Verwandter (亲戚)和 Gesandter (公使)。有关形容词类型的更多信息请参考第19页。

⁵ 为了方便,这些规则没有包括屈折类型相关的信息。

构化的符号无法捕捉到这样的事实，即例（13）中的规则是有共同点的。在形式化的术语中，规则之间的唯一共性在于规则左边有一个符号，而右边有两个。

我们可以引进特征来解决这一问题。我们给范畴符号指派一些特征，并且允许我们的规则将这些特征的值包含进来。举例来说，我们可以假设范畴符号 NP 具有人称、数和格的特征。对于限定词和名词来说，我们可以采用一条附加的特征来表示性，另一条特征来表示屈折类型。例（14）给出了两条规则，并用括号分别表示出特征值：⁶

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

如果我们在例（14）中要用到变量而不是值，那么我们会得到例（15）中的规则格式：

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

这里变量的值并不重要，重要的是他们是搭配的。为了保证规则的有效性，这些值要按照顺序排列是非常重要的；也就是说，在限定词的范畴中，性永远排在第一位，数第二，其他往后排。人称特征的值在 NP (3, Num, Case) 的第一个位置，这是由规则决定的，并且是固定的，即第三人称。当然，值的这种限制也可以由词汇来决定：

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

(10) 中的规则可以如例（17）所示，归并到一条范式中：

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

Per1 和 Num1 在动词和主语中的实现保证了主谓一致。对于其他 NP 来说，这些特征的值是无关的。这些 NP 的格也是明确的。

2.3 语义

在引言和前面几个章节，我们主要分析了语言的句法方面，而且这本书后面的部分也主要探讨句法问题。但是，有必要提醒大家的是，我们是为了交流而使用语言的，也就是说，针对某些场景、话题或者观点来交换信息。如果我们想准确地解释我们的语言能力，那么我们还要解释我们所说的话的意义。为了达到这一目标，我们就有必要理解句法结构，但是单是这样是不够的。进而，有关语言习得的理论只是关注句法结构的习得也是不全面的。由此，句法语义接口（syntax-

⁶ 第 6 章介绍了属性值结构。这些结构总是包括一对属性名称和属性值。在这样的情况下，值的顺序就不重要了，因为每个值都是独一无二的，并由相应的特征名称来指定。因为我们在（13）这样的格式中没有属性名称，所以值的顺序是重要的。

2 短语结构语法

semantics interface) 问题的重要性尤为突出, 每个语法理论都要说明句法和语义是如何互动的。在下面, 我将讲解我们是如何将短语结构规则与语义信息相结合的。为了表示意义, 我会用一阶谓词逻辑和 λ 演算。不过遗憾的是, 我们无法介绍基础逻辑知识的详细信息, 以帮助没有相关经验的读者来理解下面的内容, 但是我们这里举的这些简单的例子应该可以为句法和语义的互动关系提供一些基本的认识, 进而得到一个能够对它进行解释的语言学理论。

为了显示句子的意义是如何从它的组成部分推导出来的, 我们来看一下例 (18a)。我们将例 (18b) 中的意义指派给例 (18a) 这个句子。

- (18) a. Max schläft.
Max 睡觉
'Max 正在睡觉。'
b. *schlafen'*(*max'*)

这里, 我们假定 *schlafen'* 表示了 *schläft* (睡觉) 的含义。我们用初始符号来表示我们处理的是词义而不是实在的词。乍看上去, 我们用 *schlafen'* 来表示 (19a) 的意义并没有太大的变化, 因为它只不过是动词 *schläft* (睡觉) 的另一种形式。但是, 我们有必要集中在一个单一的动词形式上, 因为屈折变化与意义是无关的。我们可以通过比较例 (19a) 和 (19b) 中的句子来对此进行分析:

- (19) a. Jeder Junge schläft.
每个 男孩 睡觉
'每个男孩都在睡觉。'
b. Alle Jungen schlafen.
所有男孩 睡觉
'所有的男孩都在睡觉。'

为了提高可读性, 我从现在开始在语义表示中使用谓词的英语释义。⁷ 所以说, (18a) 的意义表示为 (20), 而不是 (18b):

- (20) *sleep'*(*max'*)

当我们分析 (20) 的意义时, 我们可以看出每个词都表示了哪部分含义。凭直觉来看, *max'* 来自 Max。但是棘手的问题是 *schläft* 贡献了哪些含义。如果我们考虑一个“睡觉”事件的特征的话, 那么我们就知道典型的情况是有一个人在睡觉。这个信息属于动词 *schlafen* 的含义的一部分。但是, 动词的含义并不包括睡觉的个体, 因为这个动词可以跟不同的主语搭配:

- (21) a. Paul schläft.
Paul 睡觉
'Paul 正在睡觉。'

⁷ 需要注意的是, 我并不是说英语适合表示语言的语义关系与概念, 其实它们也可以用其他语言来表示。

- b. Mio schläft.
Mio 睡觉
'Mio 正在睡觉.'
- c. Xaver schläft.
Xaver 睡觉
'Xaver 正在睡觉.'

所以说，我们可以抽象出 *sleep'* 的任意一种具体的用法。相反，我们用变量（如 x ）来表示，如（19b）中的 *max'*。这个 x 可以在指定的句子中被替换为 *paul'*，*mio'* 或 *xaver'*。为了保证我们能在给定的含义中接触到这些变量，我们在它们前面写上 λ 。这样，*schläft* 的含义可以表示如下：

(22) $\lambda x \text{ sleep}'(x)$

从（20）到（22）的步骤被叫做 λ -抽象（lambda abstraction）。（22）这个表达式与它的论元的意义整合过程是按照下面的方式进行的：我们去除 λ 和相应的变量，然后将变量的所有实例替换为论元的意义。如果我们将（22）和（23）中的 *max'* 整合在一起，我们就会得到（18b）中的含义。

(23) $\lambda x \text{ sleep}'(x) \text{ max}'$

这一过程叫做 β -还原（ β -reduction）或 λ -变换（ λ -conversion）。为了深入说明这一概念，我们用及物动词的例子来分析。例（24a）中的句子具有（24b）所表示的含义：

- (24) a. Max mag Lotte.
Max 喜欢Lotte
'Max 喜欢 Lotte.'
- b. *like'*(*max'*, *lotte'*)

mag（喜欢）的 λ -抽象如例（25）所示：

(25) $\lambda y \lambda x \text{ like}'(x, y)$

需要注意的是，第一个 λ 总是需要用在第一位的。变量 y 对应于 *mögen* 的宾语。对于英语这种语言来说，宾语和动词一起构成动词短语（VP），并且这个 VP 是与主语相搭配的。德语与英语不同的地方在于，德语在语序方面允许更大的自由度。造成形式意义匹配的问题在不同的理论中有不同的解决方案。我们将在下面的章节中来说明具体的解决方案。

如果我们将（25）中的表达式与其宾语 *Lotte* 整合在一起，我们会得到（26a），然后通过 β -还原得到（26b）：

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

这一意义可以反过来与主语相整合，进而我们得到（27a），以及通过 β -还原后的（27b）：

2 短语结构语法

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

在介绍完 λ 演算后，我们将组合性语义整合进我们的短语结构中就比较简单了。动词跟主语组合的规则需要进一步扩展，以囊括进动词的语义贡献、主语的语义贡献，以及二者组合（整个句子）后的语义所占据的位置。完整的语义是按照正确顺序排列的个体语义的组合。由此，我们可以将（28a）中的简单规则转化为（28b）：（28b）：

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

V' 表示 V 的意义， NP' 表示 $\text{NP}(\text{nom})$ 的意义。 $V' \text{ NP}'$ 表示 V' 和 NP' 的组合性意义。当我们分析（18a）时， V' 的意义是 $\lambda x \text{ sleep}'(x)$ ，而 NP' 的意义是 max' 。 $V' \text{ NP}'$ 的组合对应于（29a），或者在 β -还原后对应于（18b）-这里重复表示为（29b）：

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

我们可以提出（30）中的规则来处理例（24a）中的及物动词：

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

动词（ V' ）的意义首先与宾语（ $\text{NP2}'$ ）的意义相组合，然后跟主语（ $\text{NP1}'$ ）的意义想组合。

在这一点，我们可以看到上述的短语结构规则有几条不同的语义规则。按照这一方式来分析语言的理论假说叫做规则到规则的假说（rule-to-rule hypothesis）（Bach 1976: 184）。针对语言表达式的意义的推导过程将在第 5.1.4 节详细展开。

2.4 部分德语句法的短语结构规则

因为德语中相对自由的语序允许我们通过移位测试来判断句中的直接成分，所以说对句中直接成分的判断是相对容易的。相较而言，名词短语内各部分的分析则比较困难的。我们在这一节就来集中解决这一问题。为了对第 2.5 节中有关 \bar{X} 句法理论的讲解做准备，我们还会讲到介词短语。

2.4.1 名词短语

到目前为止，我们提出了名词短语的一个相对简单的结构：我们的规则规定了一个名词短语包括一个限定词和一个名词。名词短语可以是像例（31a）那样较为复杂的结构，如例（31）所示：

- (31) a. eine Frau
ART 女人
b. eine Frau, die wir kennen
ART 女人 CONJ 我们认识

- c. eine Frau aus Stuttgart
ART 女人 PREP 斯图加特
- d. eine kluge Frau
ART 聪明 女人
- e. eine Frau aus Stuttgart, die wir kennen
ART 女人 PREP 斯图加特 CONJ 我们认识
- f. eine kluge Frau aus Stuttgart
ART 聪明 女人 PREP 斯图加特
- g. eine kluge Frau, die wir kennen
ART 聪明 女人 CONJ 我们认识
- h. eine kluge Frau aus Stuttgart, die wir kennen
ART 聪明 女人 PREP 斯图加特 CONJ 我们认识

与限定词和名词一样的是，名词短语也可以包括形容词、介词短语和关系从句。例(31)中附加的成分是说明语。它们限制了名词短语所指称的物体的集合。而(31a)是指具有女人的属性的某个实体，(31b)中的所指也必须具有为我们所知的属性。

前面的名词短语规则只是简单地将名词和限定词组合在一起，这样只能用来分析例(31a)。我们现在面临的问题是如何修改这条规则，以适应于分析例(31)中的其他名词短语。除了规则(32a)之外，我们可以提出跟(32b)一样的一条规则。^{8,9}

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

但是，这条规则并不能用来分析诸如(33)的名词短语：

- (33) alle weiteren schlagkräftigen Argumente
所有进一步 强有力 论证
'所有其他的强有力的论证'

为了分析(33)，我们需要例(34)中的规则：

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

在名词短语中，总是有可能增加形容词的数量。而给这些形容词制定一个上限是相当武断的。即使我们选择如下的简称，仍是有问题的：

- (35) $NP \rightarrow Det\ A^*\ N$

例(35)中的星号表示任意的迭代次数。所以说，(35)包括了没有形容词，有一个、两个，或更多个形容词的规则。

问题是，根据(35)中的规则，形容词和名词并不组成一个成分，这样我们就可以不解释为什么例(36)中的并列结构仍是可能的：

⁸ 参看Eisenberg (2004: 238)中有关名词短语中平铺结构的假说。

⁹ 当然，还有诸如性和数的其他特征也应该在这一节里讨论。但是，为了降低解释的难度，我们暂不考虑这些问题。

2 短语结构语法

- (36) alle [[geschickten Kinder] und [klugen Frauen]]
所有 有能力 孩子 和 聪明 女人
‘所有有能力的孩子和聪明的女人’

如果我们假设并列涉及到具有相同句法属性的两个或两个以上的字符串的组合的话，那么我们就需要认为形容词和名词构成了一个单位。

下面的规则说明了截至目前所讨论的带形容词的名词短语：

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

这些规则说明了：一个名词短语包括一个限定词和一个名词性成分 (\bar{N})。这个名词性成分包括一个形容词和一个名词性成分 (37b)，或者只包括一个名词 (37c)。因为 \bar{N} 也位于(37b)中规则的右边，所以我们可以多次应用这一规则，并且用它来说明诸如(33)的带有多个形容词的名词短语。图 2.3 给出了不带形容词的名词短语、带一个形容词的名词短语和带两个形容词的名词短语的结构。形容词 klug

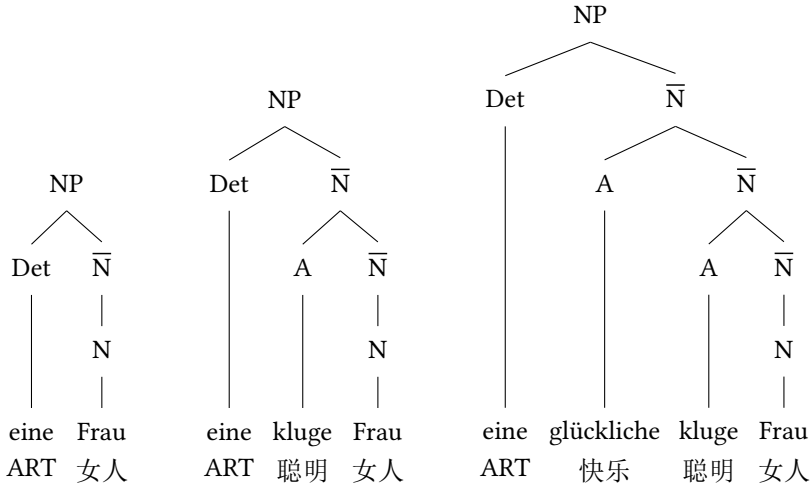


图 2.3: 带有不同数量的形容词的名词短语

(聪明) 限定了名词短语所指的集合。如果我们假设另一个形容词，如 glücklich (开心)，那么它只是指那些既开心又聪明的女人。这种名词短语可以用在如下的语境中：

- (38) A: Alle klugen Frauen sind unglücklich.
所有聪明 女人 COP 不开心
B: Nein, ich kenne eine glückliche kluge Frau.
不 我 认识 ART 开心 聪明 女人

我们观察到，这段话可以加上“Aber alle glücklichen klugen Frauen sind schön”（但是所有开心和聪明的女人都是漂亮的），以及一个相应的回答。在我们的规则系统中，(37)也允许了在名词短语中增加更多形容词的情况，如“eine glückliche kluge Frau”（一个开心的、聪明的女人）。在规则(37b)中， \bar{N} 既出现在规则的左右，也出现在规则的右边。这类规则被看作是递归的（recursive）。

我们现在构建了一套精巧的语法，可以用它来分析带有形容词性修饰语的名词短语。结果是，形容词和名词的组合构成一个组成成分。我们也可以想到这个问题，即限定词和形容词是否能够构成一个成分，如下类名词短语：

- (39) diese schlaun und diese neugierigen Frauen
这些 聪明 和 这些 好奇 女人

这里，我们分析一个不同的结构。虽然有两个完整的NP并列在一起，但是第一个并列成分的一部分被删除了。

- (40) diese schlaun ~~Frauen~~ und diese neugierigen Frauen
这些 聪明 女人 和 这些 好奇 女人

我们也可以在句子，甚至是词层面发现类似的现象：

- (41) a. dass Peter dem Mann das Buch ~~gibt~~ und Maria der Frau die Schallplatte
CONJ Peter DET 人 DET 书 给 和 Maria DET 女人 DET 唱片
gibt
给
‘Peters 给这个人这本书，并且 Maria 给这个女人这张唱片’
b. be- und ent-laden
PRT 和 PRT-装载
‘装载和卸载’

到目前为止，我们讨论了如何能够将形容词整合进我们的有关名词短语的规则当中。其他说明语，如介词短语或关系小句，也可以按照类似的方式将形容词加入 \bar{N} 之中。

- (42) a. $\bar{N} \rightarrow \bar{N} PP$
b. $\bar{N} \rightarrow \bar{N} \text{ relative clause}$

有了这些规则和(37)的那些规则，我们就可以设想PP和关系小句的规则，并用它们来分析(31)中的例子。

(37c)说明了 \bar{N} 有可能只带一个名词。但我们还没有讨论更为重要的规则：我们需要另一个规则来整合名词与它们所带的论元成分。如例(43a-b)所示，这些名词有Vater（爸爸）、Sohn（儿子）和Bild（照片）这类关系名词（relational nouns）。例(43c)说明了动词与其所带论元的名词化过程：

- (43) a. der Vater von Peter
DET 爸爸 PREP Peter
‘Peter 的爸爸’

2 短语结构语法

- b. das Bild vom Gleimtunnel
DET 照片PREP.DET 格莱姆隧道
‘格莱姆隧道的照片’
- c. das Kommen des Installateurs
DET 来 DET 水管工
‘水管工的检查’

下面是用来分析 (43a,b) 的规则，如 (44) 所示：(44):

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

图 2.4说明了带 PP 论元的两种结构。右边的树形图还包括一个额外的 PP 说明语，并由 (42a) 中的规则来允准。

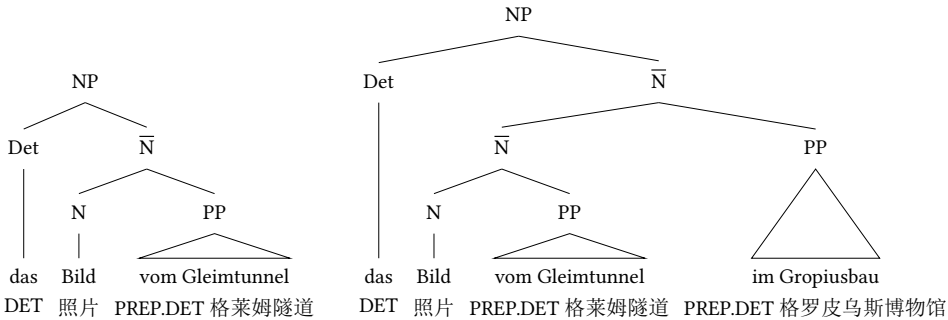


图 2.4: 名词与位于右侧的作为说明语的 PP 补足语 “vom Gleimtunnel” 的组合

除了上述讨论的 NP 结构，还有限定语或名词缺失的其他结构。名词可以通过省略而被删除。(45) 给出了名词短语的一个例子，这里不需要补足语的名词被省略了。例 (46) 说明了名词只有一个限定语，而且其补足语被实现了，而不是名词本身。下划线表明了名词正常应该出现的位置。

- (45)
- a. eine kluge _
ART 聪明
‘一个聪明的’
 - b. eine kluge große _
ART 聪明 高
‘一个聪明的高的’
 - c. eine kluge _ aus Hamburg
ART 聪明 PREP 汉堡
‘一位从汉堡来的聪明的’
 - d. eine kluge _, die alle kennen
ART 聪明 CONJ 所有人知道
‘一个大家都知道的聪明的’

- (46) a. (Nein, nicht der Vater von Klaus), der _ von Peter war gemeint.
不 不 DET 爸爸 PREP Klaus DET PREP Peter AUX 指的是
‘不, 指的的不是 Klaus 的爸爸, 而是 Peter 的爸爸。’
- b. (Nein, nicht das Bild von der Stadtautobahn), das _ vom Gleimtunnel
不 不 DET 图片PREP DET 公路 DET PREP.DET 格莱姆隧道
war beeindruckend.
AUX 印象深刻
‘不, 不是公路的照片让人印象深刻, 而是格莱姆隧道的照片。’
- c. (Nein, nicht das Kommen des Tischlers), das _ des Installateurs ist
不 不 DET 来 DET 木匠 DET DET 水管工 COP
wichtig.
重要
‘不, 木匠来不重要, 而是水管工来很重要。’

在英语中, 代词 one 必须用在相应的位置上,¹⁰ 但是德语中, 名词 就被简单地省略了。在短语结构语法中, 这就叫做 ϵ 生成式 (epsilon production)。例 ((47a) 中, 这些规则被一个空成分所替代。例 (47b) 中的规则是与术语 ϵ 生成式 (epsilon production) 相关的变量:

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

相应的树形图如图 2.5 所示。回到盒子的比喻, 例 (47) 中的规则像是贴有同样标

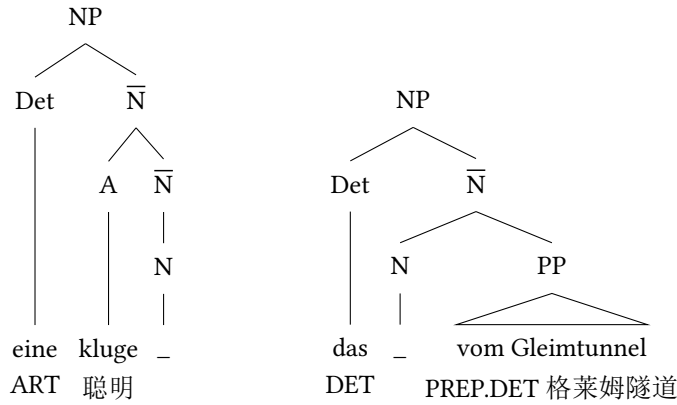


图 2.5: 没有明显中心语的名词短语

签的空盒子, 就好像是普通名词的盒子一样。正如我们在前面所强调的, 当我们要思考的是将这些盒子放在哪里的问题时, 装在盒子里的内容就不重要了。由此,

¹⁰ 参看 Fillmore et al. (2012: Section 4.12), 在英语中这些例子没有代词 one。

2 短语结构语法

例(31)中的名词短语可以在同样的句子中出现。空名词盒子也可以像一个真正的名词那样。如果我们不打开这个空盒子,我们就不会知道它与装满东西的盒子有什么不同。

我们不仅可以在名词短语中省略名词,在某些语境中,限定词也可以不出现。例(48)显示了复数的名词短语:

- (48)
- a. Frauen
女人
 - b. Frauen, die wir kennen
女人 CONJ 我们认识
 - c. kluge Frauen
聪明 女人
 - d. kluge Frauen, die wir kennen
聪明 女人 CONJ 我们认识

如果名词是集合名词的话,限定词在单数时可以被省略:

- (49)
- a. Getreide
庄稼
 - b. Getreide, das gerade gemahlen wurde
粮食 CONJ 刚 种 AUX
'刚种的庄稼'
 - c. frisches Getreide
新鲜 庄稼
 - d. frisches Getreide, das gerade gemahlen wurde
新鲜 粮食 CONJ 刚刚 种 AUX
'刚种的新鲜的庄稼'

最后,限定词和名词都可以被省略:

- (50)
- a. Ich helfe klugen.
我 帮助 聪明
'我帮助聪明人。'
 - b. Dort drüben steht frisches, das gerade gemahlen wurde.
那儿 那边 有 新鲜 CONJ 刚 种 AUX
'在那边有一些刚种的新鲜的(庄稼)。'

图 2.6 在下一页显示了相应的树形图。

针对我们目前提出的规则,有两点需要说明:截至目前,我一直在说形容词。但是,在名词前的位置上可以有非常复杂的形容词短语。这些可以是带补足语(51a,b)的形容词或者是形容词性助词(51c,d):

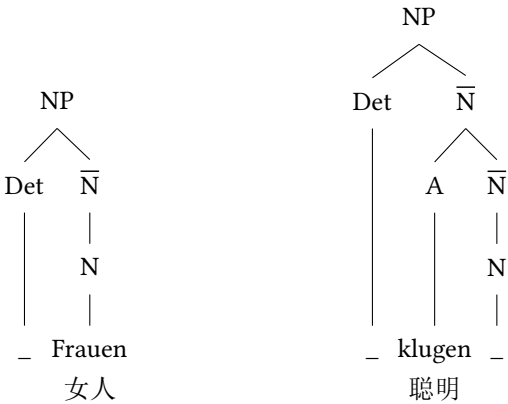


图 2.6: 没有明显限定词的名词短语

- (51) a. der seiner Frau treue Mann
DET 他的.DAT 妻子 忠诚 男人
‘忠于妻子的男人’
b. der auf seinen Sohn stolze Mann
DET PREP 他的.ACC 儿子 骄傲 男人
‘以儿子为骄傲的男人’
c. der seine Frau liebende Mann
DET 他的.ACC 女人 爱 男人
‘爱妻子的男人’
d. der von seiner Frau geliebte Mann
DET PREP 他的.DAT 妻子 爱 男人
‘被妻子爱的男人’

把这点考虑进来的话，我们需要将规则（37b）按照如下的方式来修改：

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

一个形容词短语（AP）可以包括一个 NP 和一个形容词，一个 PP 和一个形容词或者仅仅是一个形容词：

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

就目前我们提出的规则而言，还有两点不如意的结果。就是例（53c）和例（37c）中不带补足语的形容词或名词的规则，这里重复显示为例（54）：

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

2 短语结构语法

如果我们应用这些规则，那么我们会生成一元的子树，即只有一个母节点只有一个子节点的树。如图 2.6 所示。如果我们援引盒子的比喻，这就意味着一个盒子中只装有一个盒子。

原则上来说，没有人能阻止我们将这一信息放进更大的盒子中。我们不用 (55) 的规则，只是简单应用 (56) 中的规则：

- (55) a. $A \rightarrow \text{kluge}$
 b. $N \rightarrow \text{Mann}$

- (56) a. $AP \rightarrow \text{kluge}$
 b. $\bar{N} \rightarrow \text{Mann}$

例 (56a) 说明了 *kluge* (聪明) 与完整的形容词短语具有相同的属性。特别是，它不能带补足语。这与 (2) 和 (6) 的语法中，代词 *er* (他) 作为一个 NP 的范畴类型是一致的。

将 \bar{N} 指派给不需要补足语的名词是有优势的，我们不需要解释为什么 (57b) 和 (57a) 中的分析是可能的，尽管二者在意义上没有什么分别。

- (57) a. $[_{NP} \text{ einige } [_{\bar{N}} \text{ kluge } [_{\bar{N}} [_{N} \text{ Frauen }] \text{ und } [_{\bar{N}} [_{N} \text{ Männer }]]]]]]$
 一些 聪明 女人 和 男人
 b. $[_{NP} \text{ einige } [_{\bar{N}} \text{ kluge } [_{\bar{N}} [_{N} [_{N} \text{ Frauen }] \text{ und } [_{N} \text{ Männer }]]]]]]$
 一些 聪明 女人 和 男人

例 (57a) 中，两个名词被投射到 \bar{N} ，并且随后并列起来。属于同一范畴的两个成分并列的结果一定会是属于该范畴的一个新成分。在 (57a) 的例子中，这也是 \bar{N} 。这个成分随后与形容词和限定词相组合。在 (57b) 中，名词本身是一个并列结构。并列的结果一定是跟其组成成分具有一样的范畴。在这个例子中，这个范畴是 N。这个 N 变成 \bar{N} ，然后与形容词相组合。如果不带补足语的名词被看作是 \bar{N} ，而不是 N，我们就会遇到假歧义 (spurious ambiguities) 问题。例 (58) 中的结构是唯一可能的分析结果。

- (58) $[_{NP} \text{ einige } [_{\bar{N}} \text{ kluge } [_{\bar{N}} [_{\bar{N}} \text{ Frauen }] \text{ und } [_{\bar{N}} \text{ Männer }]]]]]]$
 一些 聪明 女人 和 男人

2.4.2 介词短语

与名词短语的句法相比，介词短语 (PPs) 的句法相对来说就比较简单了。PPs 通常包括一个介词和一个名词短语，这个名词短语的格属性由介词决定。我们可以用下面的规则来表示：

- (59) $PP \rightarrow P NP$

当然，这个规则也必须包括 NP 的格信息。为了方便，我在上面的 NP-规则和 AP-规则中省略了格信息。

下面 (60) 中的语法选自《杜登语法》(Eisenberg et al. 2005: §1300)，这里的介词短语进一步界定了介词所述方式的语义，如下所示：

- (60) a. [[Einen Schritt] vor dem Abgrund] blieb er stehen.
 ART 步 PREP DET 深渊 保留 他站
 ‘他在离深渊一步之遥的地方站住了。’
- b. [[Kurz] nach dem Start] fiel die Klimaanlage aus.
 短暂 PREP DET 起飞 坏 DET 空调 PTCP
 ‘起飞后不久，空调就停止工作了。’
- c. [[Schräg] hinter der Scheune] ist ein Weiher.
 斜 PREP DET 谷仓 COP ART 池塘
 ‘在谷仓的斜对面有一个池塘。’
- d. [[Mitten] im Urwald] stießen die Forscher auf einen alten Tempel.
 中间 PREP.DET 丛林 碰到 DET 研究人员 PREP ART 古老 寺庙
 ‘在丛林当中，研究者碰到了一栋古老的寺庙。’

为了分析例 (60a,b) 中的句子，我们可以提出下面这些规则，如 (61) 所示：

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

这些规则利用 PP 来表示动作的方式。最后得到的成分是另一个 PP。我们有必要在 (60a,b) 中利用这些规则来分析介词短语，但是我们无法应用这些规则来分析 (62) 中的句子：

- (62) a. * $[_{PP} \text{einen Schritt } [_{PP} \text{kurz } [_{PP} \text{vor dem Abgrund}]]]$
 ART 步 短 PREP DET 深渊
- b. * $[_{PP} \text{kurz } [_{PP} \text{einen Schritt } [_{PP} \text{vor dem Abgrund}]]]$
 短 ART 步 PREP DET 深渊

(61) 中的所有规则被用来分析 (62) 中的例子。由于在规则的左边和右边都有 PP，我们可以按照我们所希望的任意顺序和次数来应用这些规则。

为了避免这些不需要的副作用的产生，我们可以修改前面提出的规则：

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

规则 (59) 变成了 (63d))。(63c) 中的规则证明了 PP 可以包括 \bar{P} 。图 2.7 on the next page 显示了应用 (63c) 和 (63d) 对 (64) 的分析，以及在规则 (63b) 和 (63d) 的约束下形容词位于首位的例子：

- (64) vor dem Abgrund
 PREP DET 深渊
 ‘在深渊前’

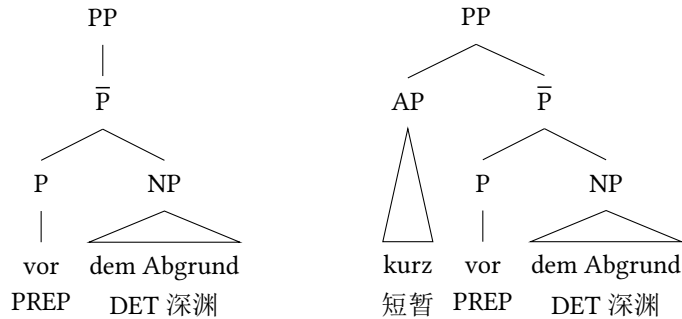


图 2.7: 带和不带方式的介词短语

有读者可能对图 2.7 有疑问，为什么左图中没有表示空方式的短语？这与图 2.6 中所示的空限定语是相似的。图 2.6 中的空限定词是指没有限定词的名词短语与有限定词的名词短语表示的意义是一样的。通常由可见的限定词表示的意义以某种方式整合进名词短语的结构之中了。如果我们在空限定语中没有表示这层意思的话，将会导致有关语义整合的更为复杂的问题：我们只是需要了解第 2.3 节表示的机制，而且这些信息是比较普遍的。意义是由这些词本身决定的，而不是由规则决定的。如果我们假设在图 2.7 中左边的那棵树那样有一条一元的规则，而不是空限定词的话，那么这个一元规则就应该提供限定词的语义。有些学者提出了这种分析。更多的有关空语类的内容参看第 19 章。

与没有限定词的 NP 不同的是，不带有关程度或方式的介词短语并不缺少任何意义的组成部分。所以说，我们没有必要假设有一个表示方式的空语类，并对整个 PP 有语义贡献。由此，(63c) 中的规则说明了介词短语包括 \bar{P} ，即由 P 和 NP 构成。

2.5 \bar{X} 理论

如果我们回顾前面章节中介绍的规则，我们会发现中心语永远和他们的补足语相组合，以构成一个新的成分 (65a,b)，然后再与其它成分 (65c,d) 相组合：

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

研究英语的语法学家发现平行结构可以用来表示那些以形容词或动词作为中心语的短语。在这里，我先讨论形容词短语，之后在第 3 章讨论动词短语。与德语一样的是，英语中可以带补足语的形容词是有严格限制的，即这些带补足语的形容词短语在英语中不能出现在名词性成分之前。如下例 (66) 所示：

- (66)
- a. He is proud.
他 COP 骄傲
‘他感到很骄傲。’
 - b. He is very proud.
他 COP 非常骄傲
‘他感到非常骄傲。’
 - c. He is proud of his son.
他 COP 骄傲 PREP 他的儿子
‘他为他的儿子感到骄傲。’
 - d. He is very proud of his son.
他 COP 非常骄傲 PREP 他的儿子
‘他为他的儿子感到非常骄傲。’

与介词短语不同的是，带补足语的形容词通常是可选的。proud（骄傲）可以带或不带 PP。表达程度的 very（很）也是可以选择的。

(67) 中有针对这一分析的规则，在图 2.8 中显示了相应的结构。

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

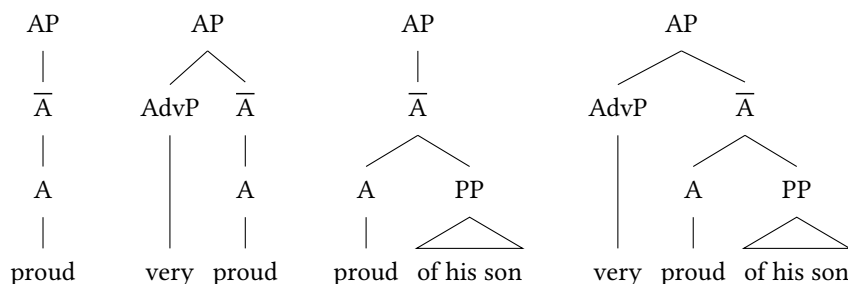


图 2.8: 英语的形容词短语

正如在第 2.2 节所显示的，我们可以对具体的短语结构规则进行扩展，并最终得到更具有普遍性的规则。按照这一方式，诸如人称、数和性这些属性不再按照范畴符号来进行编码，而是采用更为简单的符号，如 NP、Det 和 N。我们只需要明确那些在给定规则的语境中相关的特征的值。我们可以进一步抽象为：与其采用明确的范畴符号，如词范畴的 N、V、P 和 A，以及短语范畴的 NP、VP、PP 和 AP，不如简单地采用词类的变量，并用 X 和 XP 来表示。

2 短语结构语法

这种抽象的方式就叫做 \bar{X} theory (或者 X-bar 理论, *bar* 这个术语是指在符号上面的线)。该理论由Chomsky (1970) 提出, 并由Jackendoff (1977)完善。这种形式的抽象规则在很多不同的理论中发挥了重要的作用。比如说: 管辖与约束理论(第3章)、扩展的短语结构语法(第5章)和词汇功能语法(第7章)。在中心语驱动的短语结构语法(第9章)中, \bar{X} 理论也起到了重要的作用, 但是并没有采纳 \bar{X} 范式的所有限制条件。

例(68)说明了应用 \bar{X} 规则的可能的实例, X 范畴被放在了 N 的位置上, 就好像这些词串的例子都是由这些规则生成的:

(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]

任何一个词类都可以替代 X (如 V, A 或 P)。在上面的规则中, 没有横杠的 X 表示一个词汇项。如果想把横杠明确地显示出来, 就需要写成 X^0 。正如(15)中的规则所示, 我们并没有明确指定限定词或名词的格的值, 而是简单地要求在规则右边的值与之相符, (68)中的规则要求规则右边 (X 或 \bar{X}) 的元素的词类与规则左边 (\bar{X} 或 \bar{X}) 的元素的词类是一致的。

一个词汇项能跟它的补足语相组合。最后一条规则中的 “*”表示它后面的符号可以不受限制地重复。一个特殊的情况是补足语的缺失。在 *das Bild* 中, Bild 没有 PP 担当补足语, 由此变成了 \bar{N} 。词汇项及其补足语结合的结果是 X 的一个新的投射层: 投射层 1, 并由一个横杠来表示。之后, \bar{X} 可以跟说明语相组合。这一过程发生在 \bar{X} 的左边或右边。这一组合过程的结果仍然是一个 \bar{X} , 也就是说, 将一个说明语组合进来并没有改变投射层。最大投射由两个横杠表示。我们也可以将 XP 写成带有两个横杠的 X。一个 XP 包括一个指定语和 \bar{X} 。在有些理论中, 句子中的主语和 NP 中的限定词都是指定语。进而, 形容词短语中的程度修饰词和介词短语中的方式指示词也被算作是指定语。

非中心语的位置只能满足最大投射, 所以补足语、说明语和指定语总是有两个横杠。图 2.9 在下一页显示了短语结构的最小结构和最大结构。

有些范畴没有指定语或者只有一个。说明语也是可选的, 所以, 并不是所有结构都包括带有说明语子节点的 \bar{X} 。除了右边边图上的分支所示的, 带说明语的 XP 和中心语-说明语有时也是成立的。在(68)中, 只有一条规则是针对中心语在补足语之前的情况, 当然, 补足语在中心语之前的顺序也是可能的, 如图 2.9所示。

图 2.10 在下一页显示了对 NP 结构 “*das Bild*” (这张图片) 和 “*das schöne Bild von Paris*” (这张巴黎的美丽图片) 的分析。图 2.10 中的 NP 结构和图 2.8 中的树形图都是最小产出结构的例子。图 2.10 中左边的属性图也是没有说明语的结构的一个例子。图 2.10 中的右手结构是最大产出结构的一个例子: 指定语、说明语和补足语都是齐全的。

图 2.10 中的分析说明了规则中的所有非中心语都是短语。这样, 我们就需要认为这是一个限定短语, 即使限定词没有跟其他元素相组合。限定词的一元结构并

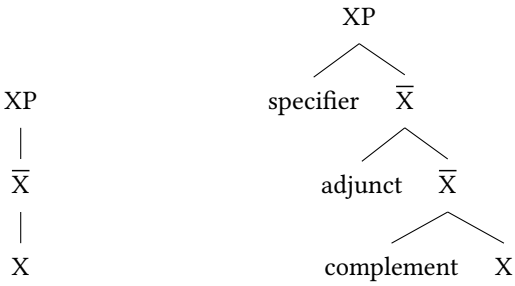


图 2.9: 短语的最小和最大结构

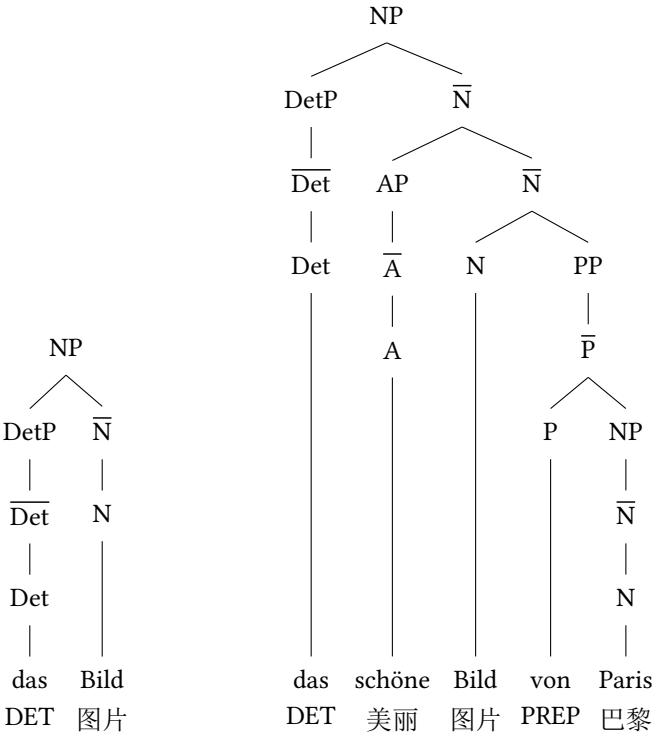


图 2.10: “das Bild”（这张图片）和 “das schöne Bild von Paris”（这张巴黎的美丽图片）的 \bar{X} 分析

2 短语结构语法

不好看，但是它是自足的。¹¹ 在图 2.10 中，Paris（巴黎）这个 NP 的一元分支看起来也有点奇怪，但是如果我们把它放进更为复杂的名词短语之中，他们就不那么奇怪了：

- (69) a. das Paris der dreißiger Jahre
DET 巴黎 DET 三十 年
‘30 年代的巴黎’
b. die Maria aus Hamburg
DET Maria PREP 汉堡
‘来自汉堡的 Maria’

虽然一元投射有些不够优雅，但是我们不用过于担心，因为在 (56) 有关词汇项的讨论中，大部分情况下一元分支节点都是可以避免的，或者说我们争取避免这样的结构。如果不这样做的话，我们就会得到假歧义结构。在下面的章节中，我们将讨论范畴语法和中心语驱动的短语结构语法，这些语法理论没有针对限定词、形容词和名词的一元规则。

再者，其他的 \bar{X} 理论与本书所讨论的理论也有不同的看法。特别是，我们将不予理会非中心语永远作为最大投射这一条假设。Pullum (1985) 和 Kornai & Pullum (1990) 认为，有些理论并不一定没有严格的 \bar{X} 理论的限制多。也可以参考第 13.1.2 节中的讨论。

思考题

1. 为什么只用原子式范畴的短语结构语法不足以描写自然语言呢？
2. 让我们来看一下 (6) 中的文法，为了分析到例 (70) 中的符号 V，我们应该采取哪些（替换符号的）步骤？

- (70) er das Buch dem Mann gibt
他 DET 书 DET 人 给
‘他给这个人这本书。’

你的回答应该与 (3) 中的分析保持一致。

3. 请应用谓词逻辑来表示例 (71) 的意义：

- (71) a. Ulrike kennt Hans.
Ulrike 认识 Hans
b. Joshi freut sich.
Joshi 高兴 REFL
‘Joshi 很高兴。’

¹¹ 有的 \bar{X} 理论并不认为限定词有复杂的结构，请看 Muysken (1982)。

练习题

1. 在第 51 页, 我认为句子 (1) 的分析可以有无限多种可能的文法。为什么这种说法是正确的?
2. 请说明哪个文法或哪些文法是最好的?
3. 在第 2.4.1 节, 我们给出了名词短语的一部分句法。为什么说例 (72) 中的规则是有问题的?

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

4. 为什么在词典中把 books 标记为 NP 是不合适的?
5. 为什么对于 books 这类名词来说, 下面的规则并不适用?

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

(73) 中的规则说明了有无限数量的修饰语与名词 books 相组合, 并且其后也有无限数量的修饰语。我们可以应用这个规则来生成例 (74) 中的短语:

- (74)
- a. books
书
‘书’
 - b. interesting books
有趣 书
‘有趣的书’
 - c. interesting books from Stuttgart
有趣 书 PREP 斯图加特
‘斯图加特的有趣的书’

请在你的回答中考虑到并列的语料。请假设对称的并列结构要求其并列的短语或词具有相同的句法属性。

6. Fillmore et al. (2012) 提出, 对于例 (75) 中的无名词结构, 应该将其看作是由限定词 the 和形容词相组合的短语结构。

- (75)
- a. Examine the plight of the very poor.
检查 DET 困境 PREP DET 非常穷
‘了解穷人的困境’

2 短语结构语法

- b. Their outfits range from the flamboyant to the functional.
他们的套装 范围 PREP DET 华丽 PREP DET 功能性
‘他们套装从华丽的到功能性的都有。’
- c. The unimaginable happened.
DET 不可想象的 发生
‘不可思议的事情发生了。’

例（76）说明了对应于其结构的短语结构规则：

（76） NP → the Adj

Adj 表示 poor 这样的单词或者 very poor 这样的复合体。

再来看（45）和（46）中的德语语料，并说明为什么这么分析，而且为什么例（77）中的更为普遍的分析不能扩展到德语的应用之中。

（77） NP → Det Adj

7. 为什么 \bar{X} 理论不能说明没有额外假设的德语形容词短语呢？（这一问题只适合于操德语（为母语）者。）
8. 请写出能够分析例（78）中的句子，并且能排除例（79）的句子的短语结构文法。

- （78） a. Der Mann hilft der Frau.
DET.NOM 男人 帮助DET.DAT 女人
‘这个男人帮助这个女人。’
- b. Er gibt ihr das Buch.
他.NOM 给 她.DAT DET 书
‘他给她这本书。’
- c. Er wartet auf ein Wunder.
他.NOM 等 PREP ART 奇迹
‘他在等奇迹发生。’

- （79） a. *Der Mann hilft er.
DET.NOM 人 帮助他.NOM
- b. *Er gibt ihr den Buch.
他.NOM 给 她.DAT DET.M 书.N

9. 为了分析下述句子，你应该对上述习题中完成的文法做哪些修改？

- （80） a. Der Mann hilft der Frau jetzt.
DET.NOM 男人 帮助DET.DAT 女人现在
‘这个男人在帮助这个女人。’

- b. Der Mann hilft der Frau neben dem Bushäuschen.
 DET.NOM 男人 帮助DET.DAT 女人PREP DET 公交车候车亭
 ‘这个男人帮助公交车候车亭旁边的那个女人。’
- c. Er gibt ihr das Buch jetzt.
 他.NOM 给 她.DAT DET.ACC 书 现在
 ‘他现在给她这本书。’
- d. Er gibt ihr das Buch neben dem Bushäuschen.
 他.NOM 给 她.DAT DET.ACC 书 PREP DET 公交车候车亭
 ‘他给公交车候车亭旁的那本书。’
- e. Er wartet jetzt auf ein Wunder.
 他.NOM 等 现在PREP ART 奇迹
 ‘他在等奇迹发生。’
- f. Er wartet neben dem Bushäuschen auf ein Wunder.
 他.NOM 等 PREP DET.DAT 公交车候车亭PREP ART 奇迹
 ‘他在等公交车候车亭旁的奇迹发生。’

10. 请安装 Prolog 系统（如 SWI-Prolog¹²）并测试你的语法。详细的使用说明可以在对应手册中定子句文法（Definite Clause Grammar, DCG）的关键词下面找到。

延伸阅读

在短语结构语法中加入特征的扩展最早是由Harman (1963)在 1963 年提出的。

本章所讨论的有关名词短语的短语结构语法覆盖了名词短语句法的大部分内容，但是仍不能解释某些特定的 NP 结构。而且，练习（3）也说明了目前文法的一些问题。Netter (1998)和Kiss (2005)在 HPSG 理论框架下讨论了这一问题，并提出了解决方案。

针对语义信息整合进短语结构语法中的讨论是比较少的。有关谓词逻辑的讨论以及它如何整合进短语结构语法中，还有量词辖域的讨论，可以参考 Blackburn & Bos (2005)。

¹² <http://www.swi-prolog.org>

第三章 转换语法-管辖与约束理论

转换语法 (Transformational Grammar) 和它的后续理论 (如管辖与约束理论、最简方案) 是由乔姆斯基在波士顿的麻省理工大学提出来的 (Chomsky 1957, 1965, 1975, 1981a, 1986a, 1995b)。Manfred Bierwisch (1963) 最早将乔姆斯基的想法应用到德语的分析中。在六十年代, 最有代表性的组织是德意志民主共和国 (GDR) 科学研究所的结构语法研究小组 (Arbeitsstelle Strukturelle Grammatik)。相关历史可以参考 Bierwisch 1992 和 Vater 2010。除了 Bierwisch 所做的工作, 下面这些专注于研究德语或者乔姆斯基的著作也是值得关注的: Fanselow (1987)、Fanselow & Felix (1987)、von Stechow & Sternefeld (1988)、Grewendorf (1988)、Haider (1993)、Sternefeld (2006)。

针对乔姆斯基理论的不同应用通常被归入到生成语法 (Generative Grammar) 的名头下面。这一术语来源于这样的事实, 乔姆斯基提出的短语结构语法和其扩充的理论框架可以生成合乎语法的表达式 (请看第50页)。正是这一组句子在形式上构成了一种语言, 我们可以通过检查一个特定的句子是否是由给定的文法生成的句子集合中的一部分来判断该句子是否属于该语言。从这个意义来看, 简单的短语结构语法与相应的形式化理论, GPSG、LFG、HPSG 和构式语法 (Construction Grammar, CxG) 都属于生成理论。近年来, 以形式化为基础的理论不断涌现, 如 LFG、HPSG 和 CxG, 前述的语法理论现在被称为模型理论 (model theoretic), 而不是生成组配的¹。(请看第 14 章的讨论)。在 1965 年, 乔姆斯基按照如下这种方式来界定生成语法 (Generative Grammar) 这一概念 (也可以参考 Chomsky 1995b: 162): ²

一种语言的语法应该是对理想的听者-说者的内在能力的描写。进而, 如果这种语法能够完美地表示出来-按照另一种说法, 如果它不依赖于读者的智力, 而只是针对他的贡献给出一个明确的分析的话-我们就可以将之称为生成语法 (generative grammar)。(Chomsky 1965: 4)

按照这一定义, 这本书所讨论的所有语法理论都应该被看作是生成语法。为了进一步区分, 有时会用主流的生成语法 (Mainstream Generative Grammar, MGG) 这一概念 (Culicover & Jackendoff 2005: 3) 来表示乔姆斯基模型。在这一章, 我将阐

¹ 模型理论的方法总是基于约束的, 所以说模型理论 (model theoretic) 和基于约束的理论 (constraint-based) 有时是同义的。

² 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 understanding reader but rather provides an explicit analysis of his contribution – we may call it (somewhat redundantly) a *generative grammar*.

3 转换语法-管辖与约束理论

述一个发展全面且影响深远的乔姆斯基式语法，即管辖和约束理论。最近的有关乔姆斯基的最简方案的内容将在第4章具体说明。

3.1 关于表示形式的一般说明

这一节说明本章的主要内容。在第3.1.1节，我将介绍转换的概念。第3.1.2节介绍了有关语言习得的假设的背景信息，这些信息在一定程度上构成了理论。第3.1.3节介绍了所谓的T模型，即管辖和约束理论的基本模型。第3.1.4节介绍了管辖和约束中用到的 \bar{X} 理论的特殊形式，第3.1.5节展示了 \bar{X} 理论的这个版本是如何应用在英语中的。有关英语句子的分析的讨论是理解对德语的分析的前提条件，因为在管辖和约束理论框架下的许多分析都是按照英语的分析来模拟的。第3.1.6节介绍了针对德语小句的分析，这些分析方法参考了第3.1.5节中对英语的分析。

3.1.1 转换

在前一章，我介绍了简单的短语结构语法。Chomsky (1957: Chapter 5)批判了这种重写文法，因为 – 按照他的说法 – 我们无法捕捉到主动句和被动句之间的关系，也无法说明句中成分排序的不同可能性。当然，我们可以在短语结构语法中给主动句和被动句不同的规则。例如，一对规则给不及物动词（1），一条规则给及物动词（2），再有一条规则给双及物动词（3）。但是这样仍然无法充分地说明同样的现象会发生在例（1）–（3）的句对儿中：

- (1) a. weil dort noch jemand arbeitet
因为那儿仍 有人 工作
‘因为有人还在那儿工作’
b. weil dort noch gearbeitet wurde
因为那儿仍 工作 AUX
‘因为那儿还有需要做的工作’
- (2) a. weil er den Weltmeister schlägt
因为他DET 世界冠军 击败
‘因为他击败了世界冠军’
b. weil der Weltmeister geschlagen wurde
因为DET 世界冠军 击败 AUX
‘因为世界冠军被击败了’
- (3) a. weil der Mann der Frau den Schlüssel stiehlt
因为DET 男人 DET 女人DET 钥匙 偷
‘因为这个男人正在从这个女人这儿偷钥匙’
b. weil der Frau der Schlüssel gestohlen wurde
因为DET 女人DET 钥匙 偷 AUX
‘因为这把钥匙从这个女人这里被偷走了’

Chomsky (1957: 43) 提出了一个转换过程，以在主动句和被动句之间创建联系。他认为，英语的被动式变换有着下面的形式，如 (4) 所示：

$$(4) \quad \begin{matrix} \text{NP} & \text{V} & \text{NP} & \rightarrow & 3 & [\text{AUX} \text{ be}] & 2\text{en} & [\text{PP} [\text{P} \text{ by}] 1] \\ 1 & 2 & 3 \end{matrix}$$

这个转换规则可以看成是带有规则左手边的符号的一棵树匹配到带有规则右手边的符号的另一棵树上。相应地，规则右边的 1、2 和 3 也对应于规则左手边的数字对应的符号。*en* 表示构成助词 (*seen*、*been*，还有 *loved*) 的语素。例 (5a,b) 的树形图如图 3.1 所示。

- (5) a. John loves Mary.
 John 爱 Mary
 ‘John 爱着 Mary。’
 b. Mary is loved by John.
 Mary AUX 爱 PREP John
 ‘Mary 正被 John 爱着。’

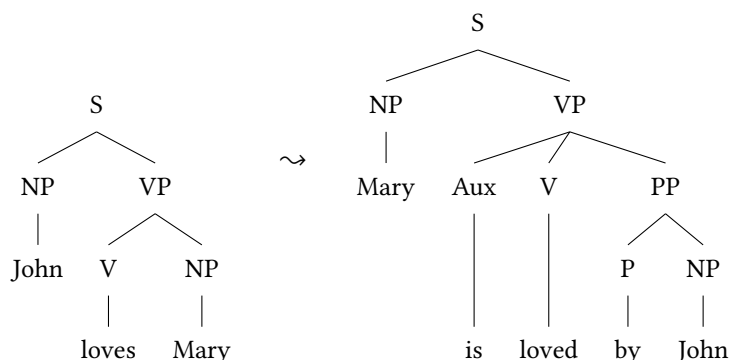


图 3.1: 被动式变换的应用

在转换规则左边的符号并不必是受限于局部的，也就是说它们可以是不同母节点的字节点，如图 3.1 所示。

重写文法按照他们的属性被分成四个复杂的类别。最简单的文法被分到第 3 类，最复杂的叫做 0 型。所谓的上下文无关文法 (context-free grammar) 是 context-free grammar 叫做 2 型文法。转换文法运用符号被任意其他符号所替换，叫做 0 型文法 (Peters & Ritchie 1973)。在对自然语言的复杂度进行研究的基础上，我们发现最复杂的 0 型文法对于自然语言来说太复杂了。按照这一观点，假设我们需要针对语言知识进行受限的形式化演算 (Chomsky 1965: 62)，我们需要限制转换的形式与潜在力量。³ 针对转换语法的早期版本的另一个争论是，由于缺乏限制，转换

³ 有关形式语言的力量的更多内容，请看第 17 章。

3 转换语法-管辖与约束理论

进行互动的方式是不够清楚的。而且，针对省略的材料而言，转换也是有问题的（请看Klenk 2003: Section 3.1.4）。基于上述原因，新的理论被提出来，如管辖与约束理论 (Chomsky 1981a)。在这个模型下，语法规则的形式受到严格的限制（请看第 3.1.4 节）。通过转换而移动定元素仍然在他们原始的位置上，这样就使得原始位置是可获得的，也就为语义解读提供了可能。也有更具有普遍性的原则来限制转换过程。

在了解了管辖与约束理论中有关语言习得模型的一些最初的看法之后，我们来仔细看一下短语结构语法、转换和限制。

3.1.2 有关语言习得的假说：原则与参数理论

Chomsky (1965: Section I.8)认为，语言知识一定是内置的，因为语言系统就是如此。所以说，我们不可能仅凭一般的认知规律依靠有限的输入而学会语言（也可以看第 13.8 节）。如果我们不能仅仅通过与环境的互动而学会语言的话，那么至少我们的部分语言能力是内置的。问题是到底有多少是内置的，还有人类是否真有一套语言的内在机制也是充满争议的，再有在过去的几十年中，对于这一问题的看法也发生了很多的变化。有关这一议题的代表性著作有 citewPinker94a、Tomasello (1995)、Wunderlich (2004)、Hauser, Chomsky & Fitch (2002)和Chomsky (2007)。更多的内容可以参看第 13 章。

Chomsky (1981a)还指出，还有一些普遍的、内在的原则是语法结构所不能违反的。这些原则是许多参数，即不只有一个选项。参数 设置在不同语言中是不同的。如下面的例（6）中的参数化原则：

（6） 原则：中心语是在补足语的前面或者后面，这取决于参数的值 POSITION。

原则与参数模型（P&P 模型）认为，语言习得中的一个重要部分在于从语言输入中提取足够的信息来设置参数。Chomsky (2000: 8) 用开关的开和关来比喻参数的设置。操英语者必须要学习的是，他们的语言中中心语出现在补足语的前面，而操日语者必须要学会的是，中心语位于补足语的后面，如下面的例（7）所示：

- （7） a. be showing pictures of himself
AUX 展示 图片 PREP 他自己
b. zibun -no syasin-o mise-te iru
REFL PREP 图片 展示 PTC

可见，日语动词、名词和形容词短语是对应的英语句子的镜像结构。（8）中的表格做了一个简单的总结，并且给出语序参数的参数值：

- （8） 语言 事实 参数：中心语位于前面
英语 中心语位于补足语之前 +
日语 中心语位于补足语之后 -

在过去的几十年中，根据特定的参数来研究语言之间的差异已经成为研究的一个重要方向，并且在跨语言的比较研究中取得了丰厚的成果。

在介绍完有关语言习得的内容之后，下面我们将讨论管辖与约束理论的基本观点。

3.1.3 T 模型

Chomsky指出,简单的短语结构语法不足以分析一些特定的关系。比如说,主动句和被动句之间的关系就是如此。在短语结构语法中,我们需要给不及物动词、及物动词和双及物动词创建主动和被动规则,请看上面(1)–(3)的讨论。事实上,短语结构语法无法描述被动式中主要的论元受到了抑制。所以,Chomsky提出还有一个潜在的结构,叫做深层结构(Deep Structure),其他结构都是由它生成的。在下面的章节中,我们将讨论T模型的构造。

3.1.3.1 深层结构与表层结构

在生成新结构的时候,深层结构的一部分可以被删除或者移动。按照这一方式,我们可以解释主动句和被动句之间的关系。对结构的这种操作结果也叫做转换。我们可以从原始的深层结构生成一个新的结构,即表层结构(Surface Structure)。在理论的有些版本中,表层结构并不是句中词语的实际使用情况的镜像,所以有时采用S-结构这一术语来避免误解。

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

图3.2给出了管辖与约束理论的整体框架:短语结构规则和词汇允准了D-结构转换为S-结构。S-结构包括语音形式(Phonetic Form, PF)和逻辑形式(Logical Form,

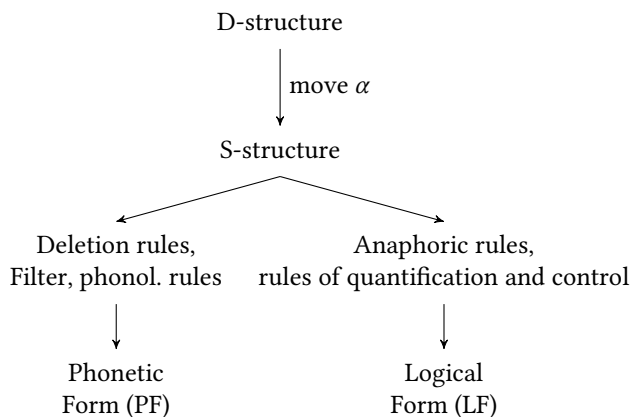


图 3.2: T 模型

LF)。这一模型被称为T模型(T-model),或Y模型(or Y-model),因为D-结构、S-结构、PF和LF从上倒下构成T(或Y)。我们会详细讲解这些具体的成分。

应用短语结构规则,我们可以描述具体元素之间的关系(比如说词和短语,有时也可以是词的一部分)。这些规则的形式是 \bar{X} 句法(请看第2.5节)。词汇与 \bar{X} 句法允准的结构一起构成了D-结构的基础。那么,D-结构就是由词库中的具体词形成的选择栅(= 配价类型)的句法表征。

3 转换语法-管辖与约束理论

词库包括每个人的词汇项，每个词汇项包含有关形态结构、句法特征和选择属性等信息。我们将在第 3.1.3.4 节详细解释这些内容。根据确切的理论假说，形态也被当作是词的一部分。但是，屈折形态更多是与句法有关的。词库是针对具体词的形式语义解释的一个接口。

成分在表层结构的位置并不一定是它们在深层结构中的位置。比如说，带有双及物动词的句子就有如下几种变换式：

- (10) a. [dass] der Mann der Frau das Buch gibt
CONJ DET.NOM 男人 DET.DAT 女人 DET.ACC 书 给
‘这个男人给这个女人这本书’
b. Gibt der Mann der Frau das Buch?
给 DET.NOM 男人 DET.DAT 女人 DET.ACC 书
‘这个男人给这个女人这本书了吗？’
c. Der Mann gibt der Frau das Buch.
DET.NOM 男人 给 DET.DAT 女人 DET.ACC 书
‘这个男人给这个女人这本书。’

针对上述移位过程，我们提出下列转换规则：(10b) 是通过将 (10a) 中的动词前置得出来的，而且 (10c) 是通过将主格对名词短语提前而得到的 (10b)。在管辖与约束理论中，只有一种非常普遍的转换：移动 α = “在任何地方移动任何东西”。而到底什么是可以移动的，以及移到哪里，和出于何种原因都是由原则所决定的。这类原则有题元准则 (Theta-Criterion) 和格过滤条件 (Case Filter)，详见下面。

谓词及其论元之间的关系是由词汇项决定的，这点在语义解读的所有表征层面上都是可及的。基于这一原因，位移元素的原本位置被标记为踪迹 (trace)。这就意味着，前置的 gibt (给) 的原始位置在例 (11b) 中有所标示。相应的标记叫做踪迹 (trace) 或 (gap)。这类空范畴乍看起来难以理解，但是我们早在第 2.4.1 节 (第 63 页) 就介绍了名词结构中的空语类。

- (11) a. [dass] der Mann der Frau das Buch gibt
CONJ DET 男人 DET 女人 DET 书 给
‘这个男人给这个女人这本书’
b. Gibt_i der Mann der Frau das Buch _{-i}?
给 DET 男人 DET 女人 DET 书
‘这个男人给这个女人这本书了吗？’
c. [Der Mann]_j gibt_i _{-j} der Frau das Buch _{-i}.
DET 男人 给 DET 女人 DET 书
‘这个男人给这个女人这本书。’

例 (11c) 是由 (11a) 经过两步移位构成的，这也就是为什么 (11c) 中有两个空语迹。这些踪迹用下标来表示，进而来区分所移动的成分。下标与所移动成分具有对应关系。有时， e (表示空) or t (表示痕迹) 也用来表示踪迹。

由 D-结构推导而来的 S-结构像表层的结构，但是它们与实际话语的结构并不完全相同。

3.1.3.2 语音形式

音位操作是在语音形式 (Phonetic Form, PF) 这一层来表示的。PF 负责实际发声的形式。比如说, wanna 这个简称就发生在 PF 层 (Chomsky 1981a: 20–21)。

- (12) a. The students want to visit Paris.
DET 学生 想要 INF 访问巴黎
‘这些学生想要访问巴黎。’
- b. The students wanna visit Paris.
DET 学生 想要 访问巴黎
‘这些学生想要访问巴黎。’

例 (12) 中的简称是由 (13) 中可选的规则所决定的:

- (13) want + to → wanna

3.1.3.3 逻辑形式

逻辑形式 (Logical Form, LF) 是介于 S-结构的句法层和句子的语义解释之间的。与 LF 有关的现象有代词的复指、量化和统制问题。

在解决复指依存的问题中, 语义因素起到了重要的作用。管辖与约束理论的一个重要部分是约束理论 (Binding Theory), 该理论用来解释代词能够或必须指称的对象, 以及什么时候能用或者必须用反身代词。(14) 中是人称代词和反身代词的例子:

- (14) a. Peter kauft einen Tisch. Er gefällt ihm.
Peter 买 一 桌子.M 他 喜欢 它
‘Peter 在买一个桌子。他喜欢它。’
- b. Peter kauft eine Tasche. Er gefällt ihm.
Peter 买 一 包.F 他 喜欢 它
‘Peter 在买一个包。他喜欢它。’
- c. Peter kauft eine Tasche. Er gefällt sich.
Peter 买 一 包.F he 喜欢 他自己
‘Peter 在买一个包。他喜欢他自己。’

在第一个例子中, er (他) 可以指代 Peter、桌子或者上文中提及的其他东西或其他人。ihm (它/他) 可以指代 Peter 或者上下文中的某人。根据世界知识, 这里限于指代桌子。在第二个例子中, er (他) 不能指代 Tasche (包), 因为 Tasche 是阴性的, 而 er 是阳性的。只有当 ihm (他) 不指向 Peter 的时候, er (他) 才可以指 Peter。否则, ihm 只能在更大范围的上下文内指称一个人。这与例 (14c) 是不同的。在 (14c) 中, er (他) 和 sich (他自己) 必须指向同一个目标。这是因为像 sich 这样的反身代词的所指受限于局部域之内。管约论主要负责描述这些限制条件。

LF 对于量词辖域来说也是重要的, 比如说 (15a) 有两种解释, 如 (15b) 和 (15c) 所示。

3 转换语法-管辖与约束理论

- (15) a. 每个男人都爱着一个女人。
b. $\forall x \exists y (\text{man}(x) \rightarrow (\text{love}(y) \wedge \text{love}(x, y)))$
c. $\exists y \forall x (\text{man}(x) \rightarrow (\text{love}(y) \wedge \text{love}(x, y)))$

符号 \forall 表示全称量词 (universal quantifier), 而 \exists 表示存在量词 (existential quantifier)。第一个公式是指对于每一个男人来说, 都有一个他爱的女人, 当然可以是不同的女人。第二种含义可以是有这样一个女人, 所有男人都爱她。什么时候会造成这种歧义以及哪种意思是可能的, 取决于给定话语的句法属性。LF 这一层次对于像 a 和 every 这样的限定词的意义来说是较为重要的。

控制理论 (Control Theory) 也跟 LF 密切相关。控制理论主要分析例 (16) 中的不定式主语的语义角色是如何实现的。

- (16) a. Der Professor schlägt dem Studenten vor, die Klausur noch mal zu
DET 教授 建议 DET 学生 PRT DET 考试 一次再 INF
schreiben.
写
‘教授建议学生再参加一次考试。’
b. Der Professor schlägt dem Studenten vor, die Klausur nicht zu bewerten.
DET 教授 建议 DET 学生 PRT DET 考试 不 INF 打分
‘教授建议学生这次考试不打分。’
c. Der Professor schlägt dem Studenten vor, gemeinsam ins Kino zu
DET 教授 建议 DET 学生 PRT 一起 进入电影院 INF
gehen.
去
‘教授建议学生一起去电影院。’

3.1.3.4 词库

为了创建更为复杂的短语, 词的意义告诉我们他们需要跟一定的角色, 如“发出动作的人”或“受到影响的事物”, 相组合。例如, 动词 beat 在语义上实际需要两个论元。(17a) 中动词 beat 的语义表征表示为 (17b) 中的形式:

- (17) a. Judit 打了大师。
b. 打'(x,y)

将中心语按照配价类型进行分类的过程也叫做次范畴化 (subcategorization): beat 可以次范畴化为一个主语和一个宾语。这一术语来自于这样的事实, 中心语已经按照它的词类 (动词、名词、形容词等) 进行分类, 并进一步根据配价信息分成小类 (如不及物动词或及物动词)。有时会用 “X subcategorizes for Y” (X 次范畴化 Y) 这个短语, 它表示 “X selects Y” (X 选择 Y)。beat 被看作是谓词, 因为 beat' 是逻辑谓词。主语和宾语都是谓词的论元。有很多术语用来描述这些选择性的必有元素的集合, 如论元结构 (argument structure)、配价框架 (valence frame)、次范畴框架 (subcategorization frame)、题元栅 (thematic grid), 以及西塔-栅或 θ -栅。

说明语修饰语义上的谓词，并且当语义部分被强调的时候，他们叫做修饰语(modifier)。在谓词的论元结构中没有说明语。

按照 GB 的理论假说，论元出现在小句的特定位置中，即所谓的论元位置，如一个 X^0 元素的子节点（请看第 2.5 节）。题元准则（Theta-准则）声称论元位置中的元素必须被赋予一个语义角色（semantic role），也叫做西塔角色（theta-role）。而且，每一个角色都只能被指派一次 (Chomsky 1981a: 36):

Principle 1 (题元准则)

- 每个题元角色只指派给一个论元位置。
- 一个论元位置内的任何短语只能充当一个题元角色。

中心语的论元成分是有主次之分的，也就是说，我们可以区分出高阶论元和低阶论元。动词和形容词的最高阶论元占有特殊的地位。因为 GB 理论认为，在有些语法中，这些高阶论元通常在动词或形容词短语之外的位置实现，通常叫做外部论元（external argument）。剩下的论元出现在动词和形容词短语内部的位置上。这类论元叫做内部论元（internal argument）或补足语（complements）。对于简单句来说，这通常意味着主语是外部论元。

当我们讨论论元类型的时候，我们可以指认出三种题元角色：

- 第一类：施事（发出动作的个体）、行动或感受的触发者（刺激），以及具有某种属性的个体
- 第二类：感事（感受的个体）、从某方面获益的人（受益者）（或者相反：受到某种伤害的人），以及领有者（所有者，或者即将拥有某物的人，或者相反的情况：失去或者缺少某物的人）
- 第三类：受事（受到影响的人或物）、客体

如果一个动词需要指派这些题元角色，那么第一类通常具有最高级别，而第三类是最低的。不过令人遗憾的是，现有文献中对动词论元所指派的语义角色的分析非常不一致。Dowty (1991) 针对这一问题作出了讨论，并提出应该应用原型角色。如果一个论元具有足够多的 Dowty 所定义的施事的原型特征（如有生的），那么该论元就被指派给原型施事。

心理词库应该包括这样的词条，即该词条具有这个词在合乎语法的使用中特定句法词的属性。比如说下列的一些属性：

- 形式
- 意义（语义）
- 语法属性：句法词类 + 形态特征
- theta-栅

(18) 给出了一个词条的例子：

3 转换语法-管辖与约束理论

(18)

形式	<i>hilft</i> ‘helps’
语义	<i>helfen</i> ’
语法特征	动词, 第三人称单数表示现在时
theta-栅	
theta-角色	<u>施事</u> 受益格
语法特性	与格

将语义角色指派给特殊的句法要求（受益者 = 与格）也叫做联接（linking）。

按照论元的主次顺序可以对论元进行排序：最高阶论元位于左边最远的位置上。以 *helfen* 为例，最高阶论元就是外部论元（external argument），所以施事下面用下划线表示。对于所谓的非宾格动词而言，⁴ 最高阶论元并不是外部论元。所以在相应的词条中没有用下划线表示。

3.1.4 \bar{X} 理论

在 GB 理论中，所有由核心语法⁵ 允准的句法结构都适用于 \bar{X} 范式（请看第 2.5 节）。⁶ 在下面的章节中，我将对其提出的句法范畴和有关语法规则的解释方面的基本论断进行评价。

3.1.4.1 句法范畴

在 \bar{X} 范式中，能用变量 X 表示的范畴可以分为词汇范畴 和功能范畴。这与开放和封闭的词汇类型是大致相关的。下面是词汇范畴：

- V = 动词
- N = 名词
- A = 形容词
- P = 介词/postposition
- Adv = 副词

词汇范畴可以用二元特征和交叉分类来表示：⁷

⁴ 有关非宾格动词的讨论请看 Perlmutter (1978)。作格动词（ergative verb）也是常见的术语，尽管用词不够准确。有关乔姆斯基框架下非宾格动词的早期研究请看 Burzio (1981, 1986)，有关德语的相关研究请看 Grewendorf (1989)。同样，有关这些术语的用法以及历史上的评介请看 Pullum (1988)。

⁵ Chomsky (1981a: 7–8) 将语言区分成常规区域和边缘区域，常规区域是指那些由语法决定的，基于基因自然习得的那部分固定的语言知识，而边缘区域是指语言中不规则的部分，如熟语 “to pull the wool over sb.’s eyes”（掩人耳目）。请看第 16.3 节。

⁶ Chomsky (1970: 210) 允许从 \bar{X} 范式中推导出语法规则。但是，更为普遍的看法是语言完全应用 \bar{X} 的结构。

⁷ 有关 N、A 和 V 的交叉分类，请看 Chomsky (1970: 199)，有关额外包含 P，但是具有不同的特征分布的内容请看 Jackendoff (1977: Section 3.2)。

表 3.1: 应用二元特征的四种词汇范畴表征

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

因为副词被看作是不及物介词，所以就在上面的表格中显示出来了。

应用这一交叉分类，我们就有可能构成普遍的规律。比如说，我们可以这样简单地指称形容词和动词：所有带有 [+V] 的词汇范畴要么是形容词，要么就是动词。进而，我们也可以说带有 [+N] 范畴（名词和形容词）的都可以有格属性。

除此之外，有些作者尝试将表 3.1 中的中心语位置与特征值联系起来（如 Grewendorf 1988: 52、Haftka 1996: 124和 G. Müller 2011a: 238）。在德语中，带有介词和名词的结构，中心语位于补足语之前：

- (19) a. *für* Marie
给 Marie
b. *Bild von* Maria
图片PREP Maria

带有形容词和动词的结构，中心语位于后面：

- (20) a. *dem König treu*
DET 国王 忠诚
‘对国王忠诚’
b. *der [dem Kind helfende] Mann*
DET DET 孩子 帮助 男人
‘帮助这个孩子的男人’
c. *dem Mann helfen*
DET 男人 帮助
‘帮助这个男人’

以上数据说明中心语带 [+V] 的位于后面，而带 [-V] 范畴的位于前面。不过，这条规律是有问题的，因为德语还有后置介词。这些成分跟前置词很像，不是动词性的，但是它们出现在所搭配的 NP 的后面：

- (21) a. *des Geldes wegen*
DET 钱 因为
‘因为钱’
b. *die Nacht über*
DET 晚上 PREP
‘在晚上’

3 转换语法-管辖与约束理论

所以说，我们要么必须发明一个新范畴，要么我们放弃用二元范畴特征来描述排序方面的限制。如果我们要把后置介词看作是一个新范畴，我们就必须假定还有一个二元特征。⁸ 因为这一属性可以是否定的或者是肯定的，这样就构成了另外四个范畴。一共有八个可能的特征组合，其中有一些不属于任何貌似合理的范畴。

对于功能范畴来说，GB 理论并没有提出交叉分类的方法。通常，我们会提出下列几种范畴：

- C 补足语（小句连接成分，如 dass）
- I 定式（也就是时态和情态）；
早期文献中也写作 Infl（曲折变化），
近期文献中写作的 T（时态）
- D 限定词（冠词、指示代词）

3.1.4.2 假设与规则

GB 理论认为，所有的规则都应该采用第 2.5 节中讨论的 \bar{X} 这种形式。在其他理论中，与 \bar{X} 形式相关的规则是这样应用的，但是其他的规则并不相同。如果按照严格的 \bar{X} 理论，就会得到短语的向心性特征（endocentricity）假设：每个短语都有一个中心语，而且每个中心语都是短语的一部分（用更为技术化的语言来说：每个中心语都投射到一个短语）。

进而，正如短语结构语法所指出的，树结构的分支之间不能交叉（Non-Tangling Condition）。这个观点在本书中的很多理论中都有所涉及。但是，在树邻接语法（TAG）和中心语驱动的短语结构语法（HPSG）、构式语法（Construction Grammar）和依存语法（Dependency Grammar）中，允许分支间交叉，由此构成非连续成分（Becker, Joshi & Rambow 1991; Reape 1994; Bergen & Chang 2005; Heringer 1996: 261; Eroms 2000: Section 9.6.2）。

在 \bar{X} 理论中，通常认为有两个投射层（ X' and X'' ）。但是，在主流的生成语法和其他理论中允许三层甚至是更多层（Jackendoff 1977; Uszkoreit 1987）。在这一章，我会按照标准说法的两个投射层，这样，短语就至少有三个层次：

- X^0 = 中心语
- X' = 中间投射（ \bar{X} ，读作：X bar）

⁸ Martin Haspelmath 指出，我们可以提出这样一条规则，将中心语后的论元成分放在中心语前的位置上（请看 Riemsdijk 1978: 89 有关转换方法的讨论）。这与德语形容词的前置论元的情况是一致的：

- (i) a. auf seinen Sohn stolz
PREP 他 儿子 骄傲
‘为他的儿子而骄傲’
- b. stolz auf seinen Sohn
骄傲 PREP 他的 儿子

但是需要注意的是，这里的情况与后置词并不相同，所有带前置宾语的形容词允许两种语序，而介词不是这样的。大部分的介词不允许它们的宾语出现在它们前面。对于后置词来说，它们的论元位于左边是一种特异性的特征。

- $XP = \text{最高投射} (= X'' = \overline{\overline{X}})$, 也叫做最大投射 (maximal projection)

3.1.5 英语中的 CP 和 IP

主流的生成语法的大部分著作都深受之前分析英语的文献的影响。如果想真正理解德语和其他语言的 GB 分析, 我们需要首先理解其对英语的分析, 并基于这个原因, 本节重在介绍这部分内容。在英语的词汇功能语法中, 也有 CP/IP 系统。由此, 下一节也介绍了第 7 章有关词汇功能语法的一些基础理论的理解。

在早期作品中, (22a) 和 (22b) 中的规则被用来分析英语句子 (Chomsky 1981a: 19)。

- (22) a. $S \rightarrow NP VP$
 b. $S \rightarrow NP Infl VP$

Infl 表示屈折变化 (Inflection), 就像插入结构中的屈折词缀一样。在早期文献中, 符号 AUX 也用来指代 Infl, 因为助动词跟屈折词汇一样。图 3.3 展示了应用 (22b) 中的规则对带有助词的句子的样本分析。

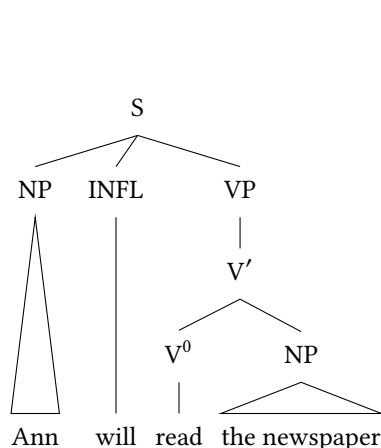


图 3.3: Chomsky (1981a: 19) 提出的带有助动词的句子

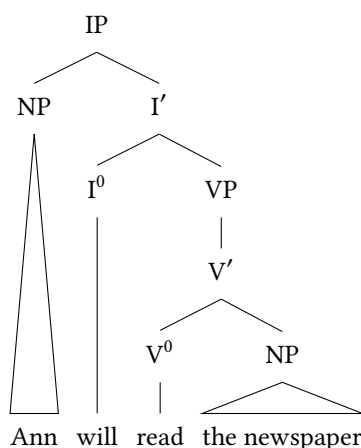


图 3.4: 在 CP/IP 系统中带有助动词的句子

动词跟它的补足语一起构成一个结构单元: VP。VP 的成分地位受到很多成分测试的支持, 也受到主语和宾语在位置上的限制的支持。

(23) 中的句子分析为补足语短语 (CPs), 补足语是中心语:

- (23) a. that Ann will read the newspaper
 CONJ Ann 将 读 DET 报纸

- b. that Ann reads the newspaper
CONJ Ann 读 DET 报纸

在像 (23) 的句子中，CPs 没有指定语。图 3.5 显示了 (23a) 的分析。

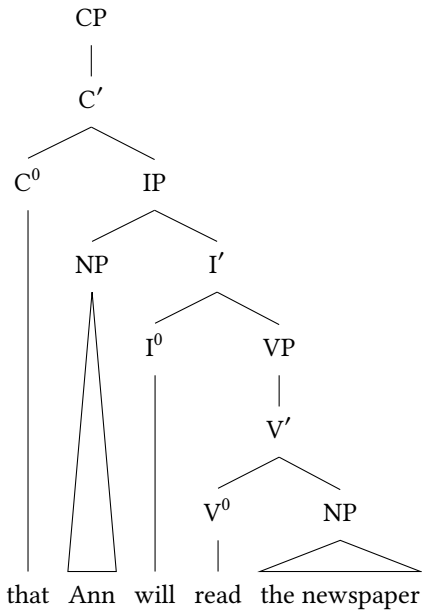


图 3.5: 补足语短语

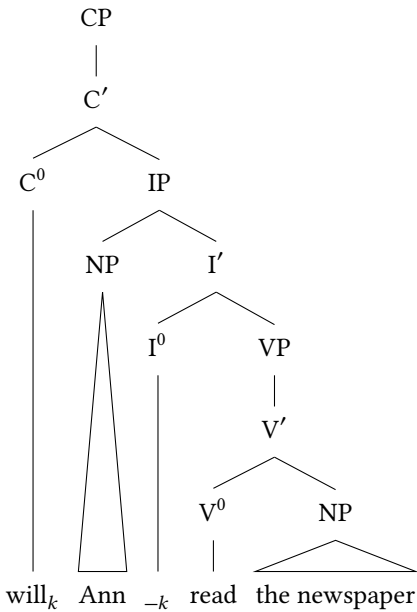


图 3.6: 极性问题

如 (24) 中的那些英语中的肯定与否定问题是通过将助动词前移到主语前而构成的。

- (24) Will Ann read the newspaper?
AUX Ann 读 DET 报纸

让我们来假设，问句的结构与带有补足语的句子结构是一致的。这就意味着问句也像 CPs。但是，与 (23) 中的句子不同，并没有从属连词。在问句的 D-结构中，C⁰ 位置是空的，而且助动词之后被移到这个位置上。图 3.6 表示了对 (24) 的分析。助动词的原始位置被标记为语迹 _{-k}，这与移动的助词的索引是相同的。

wh-问句是由助动词前的成分的额外移位构成的；也就是移到 CP 的说明语的位置上。图 3.7 在下一页显示了 (25) 的分析：

- (25) What will Ann read?
什么 AUX Ann 读

如前所述，read 的宾语的移位由语迹来说明。当我们构建句子的意思的时候，这是非常重要的。动词指派给它的宾语位置上的元素以语义角色。这样，我们就需

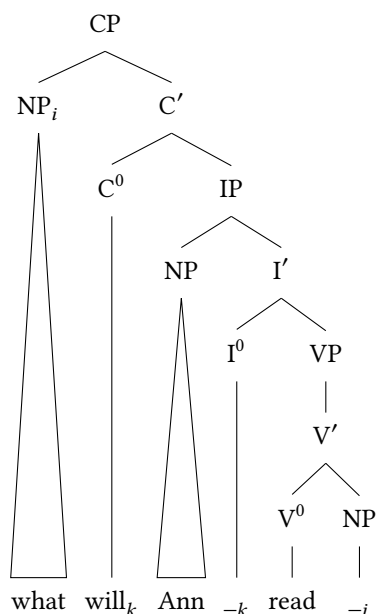
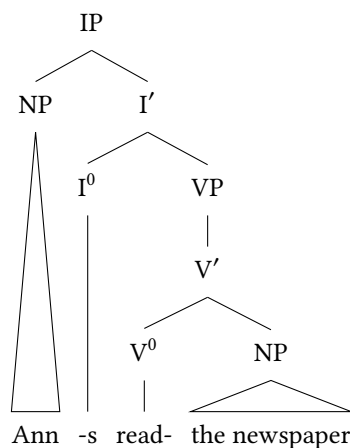
图 3.7: *wh*-question

图 3.8: 没有助动词的句子

要“重建”*what*（什么）在这个位置上的来源。这一点可以由 *what*（什么）的语迹的索引重合而得到保证。

直到现在，我们还没有讨论没有助动词的句子，如例（23b）。为了分析这类句子，我们必须假定屈折词缀在 I^0 位置上，如图 3.8 所示。因为屈折词缀出现在动词前，所以也需要一些移位操作。由于理论内部的原因，我们不需要移位到树的靠下的位置上，所示动词必须移到词缀处，而不是反过来。

按照这些针对英语句子的分析方法，我们来看德语。

3.1.6 德语小句的结构

很多学者都采用了 CP/IP 模型来分析德语。⁹ 范畴 C、I 和 V，与他们的特定位置一起，可以联接到空间位置的分布中，如图 3.9 on the next page 所示。

需要注意的是，SpecCP 和 SpecIP 并不是范畴符号。它们在语法中不与重写规则一同出现，它们仅是用来描述树上的某个位置的。

如图 3.9 所示，动词的最高论元（简单句的主语）具有特殊的地位。理所当然地，我们认为主语总是出现在 VP 的外面，并且指向外围论元。VP 本身没有指定

⁹ 对于没有 IP 的 GB 分析，请看 Bayer & Kornfilt (1989)、Höhle (1991a: 157)、Haider (1993, 1997a) 和 Sternefeld (2006: Section IV.3)。Haider 认为 I 的功能被整合进动词中。在 LFG 中，英语有 IP (Bresnan 2001: Section 6.2, Dalrymple 2001: Section 3.2.1)，而德语没有 (Berman 2003a: Section 3.2.3.2)。而在 HPSG 中，根本就没有 IP。

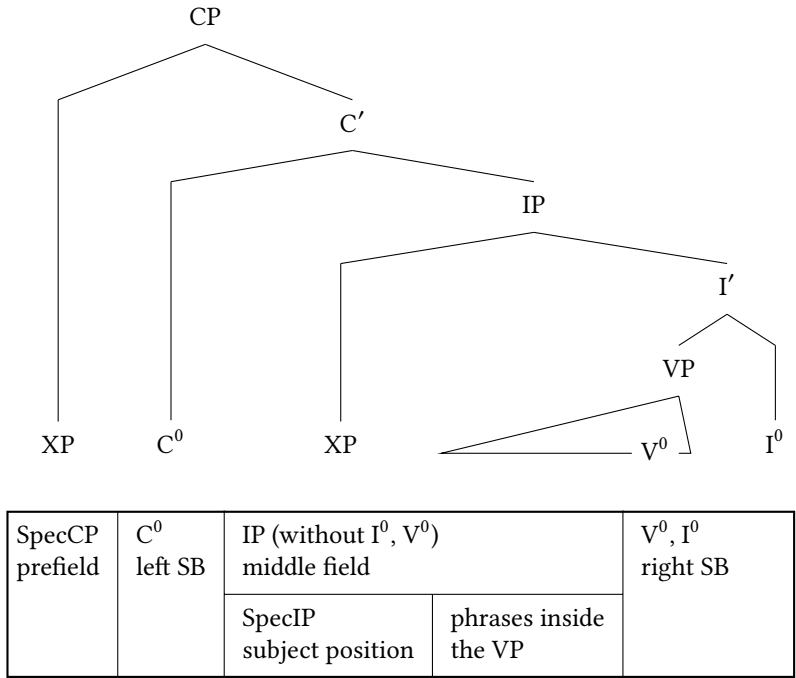


图 3.9: CP、IP 和 VP 以及德语的空间位置模型

语。但是，近期的一些文献认为主语由 VP 的指定语位置生成 (Fukui & Speas 1986; Koopman & Sportiche 1991)。在一些语言中，它被认为是移到了 VP 之外的位置上。在其他语言中，比如说德语，至少要在一定的条件下才是这样的（如有定性，请看 Diesing 1992）。这里，我将采用经典的 GB 分析，主语位于 VP 之外。所有不是主语的论元都是 V 的补足语，它们在 VP 内部实现，也就是说，它们是内部论元。如果动词只需要一个补足语，那么按照 \bar{X} 范式，这就是中心语 V^0 的姐妹节点和 V' 的子节点。非宾格宾语是原型补足语。

按照 \bar{X} 模版，说明语是 V' 的补足语之上的分支。带有一个说明语的 VP 的分析请见图 3.10 在下一页。

- (26) weil der Mann morgen den Jungen trifft
因为DET 男人 明天 DET 男孩 见面
‘因为这个男人明天要见这个男孩’

3.2 动词位置

在德语中，VP 和 IP 的中心语位置位于它们的补足语的右边， V^0 和 I^0 属于句子的右边界。主语和其他成分（补足语和说明语）都出现在 V^0 和 I^0 的左边，并且

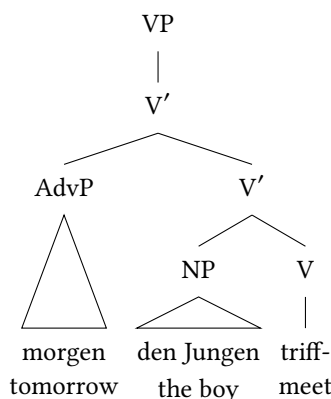


图 3.10: GB 理论中说明语的分析

构成中场。有人认为，德语-至少从 D-结构来看-是一个 SOV 类型的语言 (= 这种语言的基本语序是主语-宾语-动词)。有关德语作为一种 SOV 型语言的研究几乎跟转换语法一样古老。最早是由 Bierwisch (1963: 34) 提出的。¹⁰ 与德语不同的，其他日耳曼语言，如丹麦语、英语，以及罗曼语中的法语都是 SVO 型语言，而威尔士语、阿拉伯语是 VSO 型的语言。大概 40 % 的语言都属于 SOV 型的语言，35 % 是 SVO 型的 (Dryer 2013c)。

有关动词位于后面的观点是受到下面的数据的启发：¹¹

1. 助词与动词构成一个封闭的单位。

- (27) a. weil er morgen an-fängt
 因为他明天 PRT-开始
 ‘因为他明天开始’
 b. Er fängt morgen an.
 他开始 明天 PRT
 ‘他明天开始。’

这一单位只能用在动词末位的结构中，这也说明了这一结构反映了基本语序的事实。

通过后-构成 由名词推导出的动词 (如 *uraufführen*，表示第一次做某事) 通常不能分出他们的组成部分，这样就无法构成动词二位的小句 (Höhle (1991b)

¹⁰ Bierwisch 针对 Fourquet (1957) 的文献，提出了一个隐含的以动词为后的语序。我们可以在 Fourquet (1970: 117–135) 中找到由 Bierwisch 援引的法语手稿的德语译文。对于其他观点，请看 Bach (1962)、Reis (1974)、Koster (1975) 和 Thiersch (1978: Chapter 1)。有关德语具有隐含的 SOV 模式的观点也在 GPSG (Jacobs 1986: 110)、LFG (Berman 1996: Section 2.1.4) 和 HPSG (Kiss & Wesche 1991、Oliva 1992、Netter 1992、Kiss 1993、Frank 1994、Kiss 1995、Feldhaus 1997、Meurers 2000、Müller 2005b, 2015b) 中有所提及。

¹¹ 针对第 1 点和第 2 点，请看 Bierwisch (1963: 34–36)。针对 4，请看 Netter (1992: Section 2.3)。

3 转换语法-管辖与约束理论

在未发表的工作中首次提出这一观点。Haider (1993: 62)在出版物中首次提出。):

- (28) a. weil sie das Stück heute ur-auf-führen
因为他们DET 演出 今天 PREF-PART-引导
‘因为他们今天首次演出’
b. *Sie ur-auf-führen heute das Stück.
他们PREF-PART-引导今天 DET 演出
c. *Sie führen heute das Stück ur-auf.
他们引导 今天 DET 演出 PREF-PART

以上例子说明对于这类动词来说, 只有一个合适的位置。这个语序就被认为是基本语序。

2. 在非定式小句和定式从句中带连词的动词都处于动词末位 (我不考虑移位的可能性):

- (29) a. Der Clown versucht, Kurt-Martin die Ware zu geben.
DET 小丑 尝试 Kurt-Martin DET 物品 PCTP 给
‘小丑正在试着给 Kurt-Martin 这些物品。’
b. dass der Clown Kurt-Martin die Ware gibt
CONJ DET 小丑 Kurt-Martin DET 物品 给
‘小丑给 Kurt-Martin 这些物品’

3. 如果我们将德语中的动词位置与丹麦语 (丹麦语跟英语一样是 SVO 型的语言) 中的动词位置相比较的话, 我们可以清楚看到德语动词聚集在句子的末尾, 而丹麦语中它们出现在宾语的前面 (Ørnsnes 2009a: 146):

- (30) a. dass er ihn gesehen₃ haben₂ muss₁
CONJ 他他 看见 AUX 必须
b. at han må₁ have₂ set₃ ham
CONJ 他 必须AUX 看见他
‘他必须看见他’

4. 在 (31) 中, 副词的辖域决定于他们的语序: 位于最左边的副词辖域囊括后面的两个元素。¹² 这点可以在下面提出的结构中进行解释:

¹² 在这里, 我们需要指出的是, 左边的修饰语的辖域包括其右边的这条规则也是有例外的。Kasper (1994: 47)讨论了 (i) 中的这些例子, 也可以参考Bartsch & Vennemann (1972: 137)。

(i) a. Peter liest gut wegen der Nachhilfestunden.
Peter 阅读好 因为 DET 补习课

- (31) a. weil er [absichtlich [nicht lacht]]
 因为他 本想 不 笑
 ‘因为他本来想不笑’
 b. weil er [nicht [absichtlich lacht]]
 因为他 不 本想 笑
 ‘因为他本来不想笑’

值得一提的是，辖域关系并不受到动词位置的影响。如果我们假设动词二位的句子具有如(31)中的结构，那么这一事实就不需要再深入解释了。(32)显示了(31)推导出的S-结构：

- (32) a. Er lacht_i [absichtlich [nicht _{-i}]].
 他 笑 本想 不
 ‘他本来想不笑。’
 b. Er lacht_i [nicht [absichtlich _{-i}]].
 他 笑 不 本想
 ‘他本来没想笑。’

在简单介绍了动词末位语序的分析之后，我将详细说明德语的CP/IP分析。C⁰对应于句子的左边界，并且可以按照两个不同的方式来填充：英语中，在由连词引

-
- b. Peter liest wegen der Nachhilfestunden gut.
 Peter 阅读因为 DEF 补习课 好
 ‘Peter的阅读好，多亏上了补习班。’

正如Koster (1975: Section 6)和Reis (1980: 67)指出的，这些并不是特别可信的反例，因为这些例子中右边句子的边界没有被填满。所以，这些例子就不必是对中场内部重新排序的例子，而相反地是包括前置的PP。正如Koster和Reis所指出的，如果我们能够把右边界填上，并且不把致使的说明语前置的话，这些例子是不合乎语法的：

- (ii) a. *Hans hat gut wegen der Nachhilfestunden gelesen.
 Hans AUX 好 因为 DET 补习课 阅读
 b. Hans hat gut gelesen wegen der Nachhilfestunden.
 Hans AUX 好 阅读 因为 DET 辅导
 ‘Hans因为辅导而读得好。’

但是，下面的例子选自Crysmann (2004: 383)，即使右边界被填充了，我们还是可以得到这样的语序，位于右边的说明语的辖域囊括了左边的成分：

- (iii) Da muß es schon erhebliche Probleme mit der Ausrüstung gegeben haben, da wegen
 这儿 必须 它 已经 严重 问题 PREP DET 设备 有 AUX 因为 由于
 schlechten Wetters ein Reinhold Messmer niemals aufgab.
 坏 天气 ART Reinhold Messmer 从不 放弃
 ‘一定是设备出现了严重的问题，因为像Reinhold Messmer这样的人不会仅仅因为天气不好就放弃了。’

尽管如此，这并不能改变任何事。不管动词的位置在哪儿，(31)和(32)中的相关例子具有同样的含义。语义组合的方式可以在Crysmann的分析中按照同样的方式来实现。

导的从句中，从句连词（补足语）占据了 C^0 的位置。动词还是位于句子的右边界，如例（33）所示。

- (33) dass jeder diesen Mann kennt
 CONJ 每个人这 人 认识
 ‘每个人都认知这个人’

(33) 的分析如图 3.11 所示。在动词首位和动词二位的小句中，定式动词通过 I^0 位

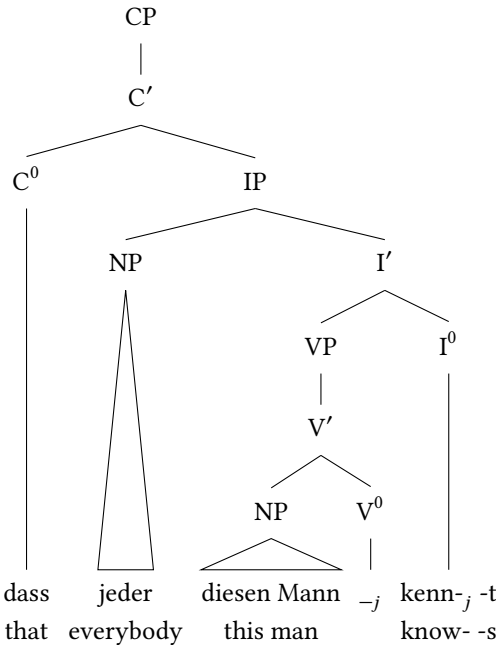


图 3.11: 在 C^0 中带有补足语的句子

置移动到了 C^0 : $V^0 \rightarrow I^0 \rightarrow C^0$ 。图显示了例（3.12 在下一页）的分析：

- (34) Kennt jeder diesen Mann?
 认识 每个人这 人
 ‘所有人都认知这个人吗？’

在例（34）的 D-结构中， C^0 位置是空的。因为这个位置不能由补足语充当，动词可以移到这里。

3.3 长距离依存

在德语的陈述句中，SpecCP 位置 对应于前场，并可以由任何 XP 成分来充当。这样，我们可以通过将成分移到动词前面这种方式，来从例（35）中推导出例（36）

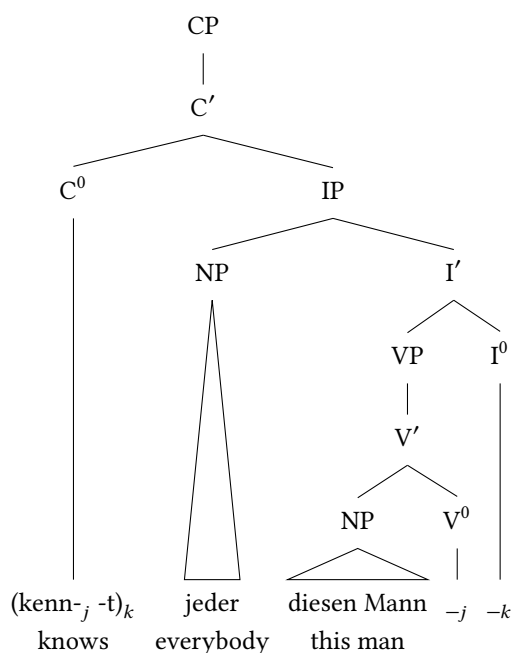


图 3.12: GB 中的动词位置

的句子:

- (35) Gibt der Mann dem Kind jetzt den Mantel?
 给 DET.NOM 人 DET.DAT 孩子 现在DET.ACC 大衣
 ‘这个人现在要把大衣给孩子了吗?’
- (36) a. Der Mann gibt dem Kind jetzt den Mantel.
 DET.NOM 人 给 DET.DAT 孩子 现在DET.ACC 大衣
 ‘这个人正在给孩子这件大衣.’
 b. Dem Kind gibt der Mann jetzt den Mantel.
 DET.DAT 孩子 给 DET.NOM 人 现在DET.ACC 大衣
 c. Den Mantel gibt der Mann dem Kind jetzt.
 DET.ACC 大衣 给 DET.NOM 人 DET.DAT 孩子 现在
 d. Jetzt gibt der Mann dem Kind den Mantel.
 现在 给 DET.NOM 人 DET.DAT 孩子 DET.ACC 大衣

由于任何成分都可以移到定式动词的前面，德语在类型学上被看作是一种动词位于第二位的语言（V2）。所以说，这是一种带有 SOV 基本语序的动词位于第二位的语言。另一方面，英语则是不具有 V2 属性的 SVO 语言，而丹麦语则是以 SOV 为基本语序的动词位于第二位的语言（请看Ørsnes 2009a有关丹麦语的研究）。

图 3.13 显示了由图 3.12推导出来的结构。决定短语移动的关键因素是句子的

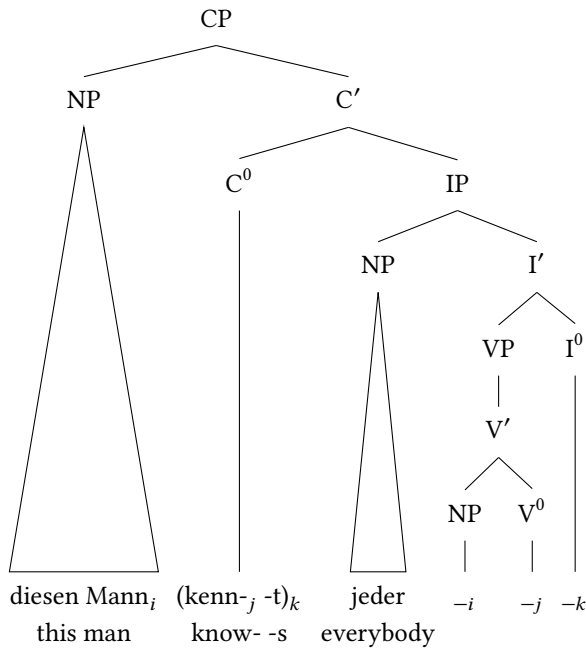


图 3.13: GB 理论中的前置

信息结构 (information structure)。具体来说，与前述和其他已知信息相联系的材料放在最左端（最好在前景），而新信息一般出现在右边。在陈述句中前置到前景的过程一般叫做话题化 (topicalization)。但是，这一术语不够准确，因为焦点 isfocus (非正式定义：所要表达的成分) 也可以出现在前景。进一步说，虚指代词 ispronoun!expletive 可以出现在那里，而这些成分是无指的，所以它们并不能联系到前面和已知的信息。据此，虚指成分永远不能作话题。

基于转换的分析也可以用在所谓的长距离依存中。所谓长距离依存，是指跨域几个短语边界的依存关系：

- (37) a. [Um zwei Millionen Mark]_i soll er versucht haben, [eine Versicherung
大约两 百万 德国马克应该他试图 AUX ART 保险公司
-_i zu betrügen].¹³
INF 欺骗
‘他显然试图从保险公司那里骗取两百万德国马克。’

¹³ taz, 2001 年 5 月 4 日, p. 20.

- b. „Wer_i glaubt er, daß er _i ist?“ erregte sich ein Politiker vom Nil.¹⁴
 谁 相信 他 CONJ 他 AUX 反驳 REFL ART 政客 PREP 尼罗河
 ‘“他认为他是谁呀？”，一个来自尼罗河来的政客这样反驳道。’
- c. Wen_i glaubst du, daß ich _i gesehen habe?¹⁵
 谁 认为 你 CONJ 我 看见 AUX
 ‘你认为我看见谁了？’
- d. [Gegen ihn]_i falle es den Republikanern hingegen schwerer,
 反对 他 陷阱它 DET 共和党人们 但是 更难
 [[Angriffe _i] zu lancieren].¹⁶
 攻击 INF 发起
 ‘但是，共和党人们很难对他发起攻击。’

在(37)的例子中，前场中的元素都是来自于嵌套更深的短语。在GB理论中，跨越句子界限的长距离依存是按照(Grewendorf 1988: 75-79)的步骤生成的，也就是说，在(37c)的分析中，疑问代词被移动到dass引导的从句的限定语位置，并且从那里移到主句的限定语位置上。这是因为移位受到的限制必须在局部进行核查。

3.4 被动

在介绍第3.4.2节的被动之前，第一小节将详细说明结构格与词汇格的区别。

3.4.1 结构格与词汇格

带有格标记的论元的格属性有赖于论元的中心语所实现的句法环境。这些论元被叫做是具有结构格的论元。而不具有结构格的带有格标记的论元叫做词汇格。¹⁷

下面是结构格的例子：¹⁸

- (38) a. Der Installateur kommt.
 DET.NOM 水管工 来
 ‘水管工要来。’

¹⁴ Spiegel, 1999 年 8 月, p. 18.

¹⁵ Scherpenisse (1986: 84).

¹⁶ taz, 2008 年 2 月 8 日, p. 9.

¹⁷ 另外，它们也叫做一致格 (agreeing case) (请看第 ?? 页) 和语义格 (semantic case)。一致格一般在谓词中。这个格属性也随着相应的结构发生变化，但是这个变化一般发生在前面的元素中。语义格与某些短语的功能具有相关性 (如时间宾格副词性结构)。进而，与宾语的词汇格类似的是，语义格并不受到句法环境的影响而发生变化。对于这一小节将讨论到的被动的分析而言，只有结构格和词汇格是相关的。

¹⁸ 我们来比较Heinz & Matiassek (1994: 200)。

例(38b)也叫做AcI结构。AcI是“Accusativus cum infinitivo”的缩写，它表示“无限的宾格”。所嵌套动词(kommen, 这里表示“来”)的逻辑主语变成了主动词(lassen, “让”)的宾格宾语。AcI-动词一般是感官动词, 如 hören (听)、sehen (看) 和 lassen (让)。

3 转换语法-管辖与约束理论

- b. Der Mann lässt den Installateur kommen.
DET 人 让 DET.ACC 水管工 来
‘这个人让水管工来。’
- c. das Kommen des Installateurs
DET 来 PREP.DET 水管工
‘水管工的拜访’

在第一个例子中，主语是主格，而在第二个例子中，Installateur（水管工）是宾格，在第三个例子中，它则是紧跟主格后面的属格。宾语的宾格通常是结构格。这个格在被动化后变成了主格：

- (39) a. Karl schlägt den Weltmeister.
Karl 打败 DET.ACC 世界冠军
‘Karl 打败了世界冠军。’
- b. Der Weltmeister wird geschlagen.
DET.NOM 世界冠军 AUX 打败
‘世界冠军被打败了。’

与宾格不同的是，由动词管辖的属格是一个词汇格。当动词被动化之后，具有属格的宾语的格是不变的。

- (40) a. Wir gedenken der Opfer.
我们记得 DET.GEN 受害者
- b. Der Opfer wird gedacht.
DET.GEN 受害者AUX 被记住
‘受害者被记住了。’

例（40b）是所谓 impersonal passive（人称被动）的一个例子。与例（39b）不同的是，宾格宾语变成了主语，而（40b）中没有主语。请看第 1.7.1 节。

同样，与格宾语也没有变化：

- (41) a. Der Mann hat ihm geholfen.
DET 人 AUX 他.DAT 帮助
‘这个人帮助了他。’
- b. Ihm wird geholfen.
他.DAT AUX 帮助
‘他被帮助了。’

对于与格来说，它应该被看作是词汇格，还是说在动词结构环境中的一些或者所有与格都应该被看作是结构格还是存在争论的。限于篇幅，我不在这里展开讨论，有兴趣的读者可以参考Müller (2007b)中第 14 章的内容。在后面的内容中，我将按照Haider (1986a: 20) 的假设，认为与格实际上是一个词汇格。

3.4.2 格指派与格过滤

在 GB 理论中, 主语从 (定式的) I 位置得到格, 而其他论元的格来自于 V (Chomsky 1981a: 50; Haider 1984: 26; Fanselow & Felix 1987: 71–73)。

Principle 2 (格指派原则)

- V 将宾格 (*accusative*) 指派给它的补足语, 如果这个宾格是结构格的话。
- 当动词为定式动词时, *INFL* 将格指派给主语。

格过滤 过滤掉那些没有将格指派到 NP 的结构。

图 3.14 on the following page 说明了例 (42a) 中的格指派原则。¹⁹

- (42) a. [dass] der Mann der Frau den Jungen zeigt
CONJ DET 男人 DET.DAT 女人 DET.ACC 男孩 展示
‘这个男人将这个男孩展示给这个女人’
- b. [dass] der Junge der Frau gezeigt wird
CONJ DET 男孩.NOM DET.DAT 女人 展示 AUX
‘这个男孩被展示给这个女人’

被动在形态上要求主语不出现, 并且吸收了结构化的宾格。在主动句中的宾语在被动式中会变成宾格, 而且其语义角色与其原有位置保持一致, 但是它不能得到吸收的格。所以, 它就必须移动到可以指派格的位置上 (Chomsky 1981a: 124)。图 3.15 on the next page 说明了对于例 (42b) 而言, 这一过程是如何进行的。

这种基于移位的分析适合于英语, 因为隐含的宾语总是要移动的:

- (43) a. The mother gave [the girl] [a cookie].
DET 妈妈 给 DET 女孩 ART 饼干
‘妈妈给这个女孩饼干。’
- b. [The girl] was given [a cookie] (by the mother).
DET 女孩 AUX 给 ART 饼干 PREP DET 妈妈
‘这个女孩从妈妈那里得到了饼干。’
- c. *It was given [the girl] [a cookie].
它 AUX 给 DET 女孩 ART 饼干

例 (43c) 说明了在主语位置上填上虚指代词是不可行的, 所以宾语必须要移位。但是 Lenerz (1977: Section 4.4.3) 指出, 这类移位在德语中并不是必需的:

- (44) a. weil das Mädchen dem Jungen den Ball schenkte
因为 DET.NOM 女孩 DET.DAT 男孩 DET.ACC 球 给
‘因为这个女孩将球给了这个男孩儿’

¹⁹ 上图并不符合 \bar{X} 理论的标准范式, 因为 *der Frau* (这个女人) 是一个与 V' 相组合的补足语。在标准的 \bar{X} 理论中, 所有的补足语都与 V^0 相组合。这对于双及物动词结构来说是有问题的, 因为我们要将这些结构处理为二分的 (参看 Larson (1988) 有关双宾语结构的分析)。进而, 在下面的图中, 出于明晰性的原因, 动词被置于 V^0 的位置上。为了创造出合格的 S-结构, 动词必须要移动到 I^0 位置的前缀处。

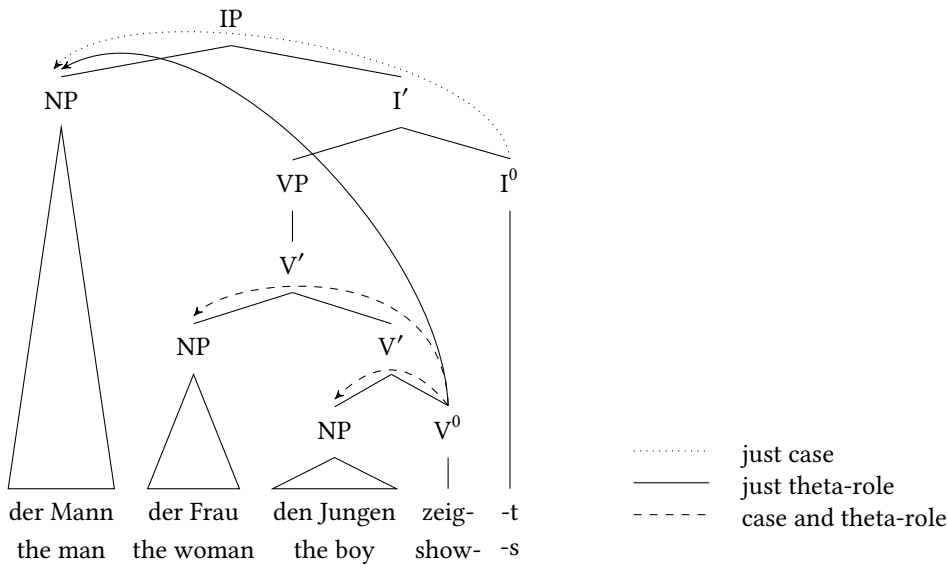


图 3.14: 主动句中的格与题元角色指派

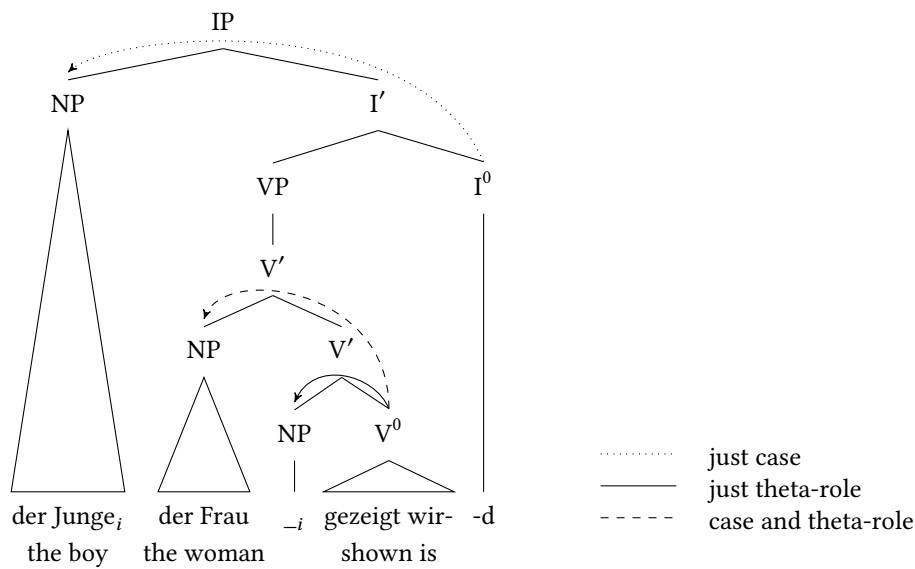


图 3.15: 被动句中的格与题元角色指派

- b. weil dem Jungen der Ball geschenkt wurde
 因为DET.DAT 男孩 DET.NOM 球 给 AUX
 ‘因为这个球被送给了这个男孩’
- c. weil der Ball dem Jungen geschenkt wurde
 因为DET.NOM 球 DET.DAT 男孩 给 AUX

与例(44c)相比,例(44b)是没有标记的顺序。例(44b)中的 *der Ball* (球)与例(44a)中的 *den Ball* 出现在相同的位置上。也就是说,移位并不是必需的。只有格是不同的。但是,例(44c)相对于(44b)而言,在某种程度上是有标记的。我们提出的诸如(44b)的分析中包括了抽象的移位过程:元素在他们的位置上,但是连接到主语的位置上,并在那里接收到他们的格信息。Grewendorf (1993: 1311)认为在诸如例(44b)的主语位置上有一个空的虚指代词,而且在诸如例(45)的句子的主语位置上包含了人称被动的过程:²⁰

- (45) weil heute nicht gearbeitet wird
 因为今天 不 工作 AUX
 ‘因为今天没有完成的工作’

空的虚指代词是指我们看不到也听不到的成分。对于这类空成分的讨论,可以参考第13.1.3节和第19章。

在下面的章节中,我将介绍被动的另一种分析方法,这种方法没有联系到论元位置的空范畴,并且试图按照一种更为普遍和一致的具有跨语言共性的方式(如对于最主要论元的压制)来分析被动。

进而,我们需要回答的问题是为什么宾格宾语没有从动词那里得到格指派。对于这一问题的解释可追溯到Burzio (1986: 178–185)提出的Burzio 概说(Burzio’s Generalization)。²¹

- (46) Burzio 概说(修订版):
 如果 V 没有外部论元,那么它并不指派(结构)宾格。

²⁰ 请看Koster (1986: 11–12)有关荷兰语的平行分析,还有Lohnstein (2014)基于移位的分析,即对于人称被动而言,也包括一个空的虚指代词。

²¹ Burzio 提出的早期公式对应于下面:一个动词指派宾格,当且仅当它指派给它的主语一个语义角色。这个概念在以下两个方面来看都是有问题的。(i),动词并没有指派给主语语义角色,而是不管怎样都是宾格:

- (i) Mich friert.
 我.ACC 冻
 ‘我冻极了。’

我们必须区分结构宾格和词汇宾格,并相应地修改Burzio 概说。诸如 *begegnen* (遇见)的存在动词从隐含的另一个方面来看也是有问题的。*begegnen* 有主语,但是仍然不能指派宾格,而是与格:

- (ii) Peter begegnete einem Mann.
 Peter 遇见 ART.DAT 人
 ‘Peter 遇见了个人。’

请看Haider (1999)和Webelhuth (1995: 89),以及这里针对Burzio 概说的问题的参考文献。

3 转换语法-管辖与约束理论

Koster (1986: 12)指出, 英语中的被动不能由格理论推导出来, 因为如果我们在英语中, 像德语和荷兰语那样允许空的虚指主语的话, 我们就可以分析下面的例子(47), 这里 np 是一个空的虚指成分:

- (47) np was read the book.
np AUX 读 DET 书

相反, Koster 提出, 英语中的主语要么受制于其他成分(即非虚指成分), 要么在词汇上被填充, 即由可见的材料填充。这样, 例(47)的结构就被排除了, 并且我们可以保证 the book 需要被位于定式动词的前面, 以使得主语位置被填充上。

3.5 局部重新排序

位于中场的论元在原则上几乎可以按照任意的顺序排列, 如例(48)所示:

- (48) a. [weil] der Mann der Frau das Buch gibt
因为 DET 男人 DET 女人 DET 书 给
'因为这个男人给这个女人这本书'
b. [weil] der Mann das Buch der Frau gibt
因为 DET 男人 DET 书 DET 女人 给
c. [weil] das Buch der Mann der Frau gibt
因为 DET 书 DET 男人 DET 女人 给
d. [weil] das Buch der Frau der Mann gibt
因为 DET 书 DET 女人 DET 男人 给
e. [weil] der Frau der Mann das Buch gibt
因为 DET 女人 DET 男人 DET 书 给
f. [weil] der Frau das Buch der Mann gibt
因为 DET 女人 DET 书 DET 男人 给

在例(48b-f)中, 组成成分具有不同的重音, 而且跟例(48a)相比, 每个能说的句子的语境的数量是受到限制的(Höhle 1982)。由此, 例(48a)中的语序被看作是中性语序(neutral order)或没有标记的语序(unmarked order)。

为了分析这些语序, 我们提出了两条建议: 第一条建议是假定(48b-f)中的五种语序是通过 Move- α (Frey 1993)将一个隐含的语序推导出来的。举例来说, 对(48c)的分析如图 3.16 在下一页所示。宾语 das Buch(书)被移到了左边, 并且连接到最高点的 IP。

一个通常用来支撑这一分析观点的论断是, 实际上在重新排序的句子中存在的辖域歧义在基本语序的句子中是不存在的。对于这种歧义的解释是基于这样的假设, 即限定词的辖域可以在表层结构的位置上推导出来, 也可以在深层结构的位置上推导出来。如果在表层结构和深层结构中的位置是一致的, 也就说没有任何的移位, 那么就只有一种解读的可能性。但是, 如果发生了移位, 那么就就有两种可能的解释(Frey 1993):

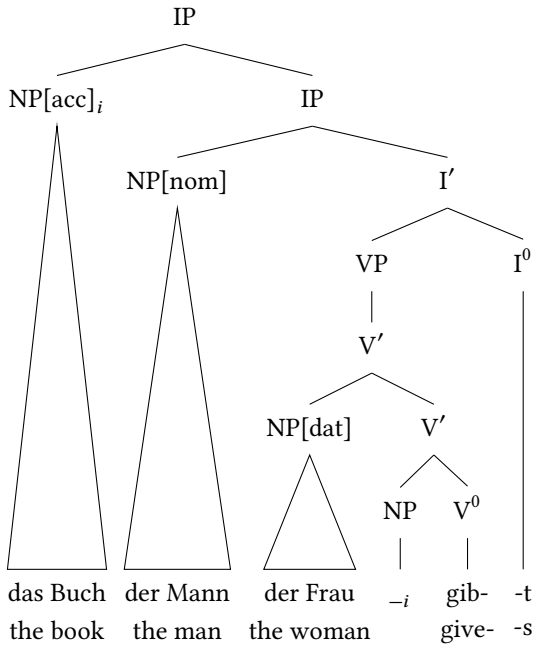


图 3.16: 作为 IP 附加语的局部重新排序分析

- (49) a. Es ist nicht der Fall, daß er mindestens einem Verleger fast jedes
它 COP 不 DET 事实 CONJ 他至少 ART 出版社 几乎每
Gedicht anbot.
诗 提供
‘事实并不是他给至少一家出版社提供了几乎每一首诗。’
- b. Es ist nicht der Fall, daß er fast jedes Gedicht_i mindestens einem
它 COP 不 DET 事实 CONJ 他几乎每 诗 至少 ART
Verleger _{-i} anbot.
出版社 提供
‘事实并不是他给至少一家出版社提供了几乎每一首诗。’

有些观点认为语迹会带来诸如对于句子有不同解读的问题，事实上，这些问题是不存在的（请看 Kiss 2001: 146 和 Fanselow 2001: Section 2.6）。以 (50) 的例子来说，可以将 “mindestens einem Verleger” (at least one publisher) 这句话在 _{-i} 位置上进行解释，这样就可以得到后面的解释，即 “fast jedes Gedicht” (almost every poem) 涵盖了 “mindestens einem Verleger” (at least one publisher)。但是，这样的解释是不存在的。

- (50) Ich glaube, dass mindestens einem Verleger_i fast jedes Gedicht_j nur dieser
我 相信 CONJ 至少 ART 出版社 几乎每 诗 只有这

3 转换语法-管辖与约束理论

Dichter $_{-i}$ $_{-j}$ angeboten hat.
诗人 提供 AUX

‘我想，只有这位诗人给至少一家出版社提供了几乎每一首诗。’

Sauerland & Elbourne (2002: 308) 讨论了日语中的相关例子，他们将之归功于 Kazuko Yatsushiro。他们提出了这样一个分析，其中第一步是将宾格宾语提到主语的前面。然后，与格宾语放在它的前面，然后，第三步移位是，宾语再次移动。最后一步既可以构成表层结构²² 也可以构成音位形式。对于后者来说，这种移位没有任何语义效应。但是，这种分析可以预测正确的有可能性的解读，它并不需要带有中间步骤的格外的移位操作。

移位分析的另一种说法叫做基础生成 (base generation)：由短语结构规则允准生成的最初结构叫做基础结构 (base)。基础生成的一个变体认为动词一次与一个论元相组合，而且每个 θ -角色被指派给相应的中心语-论元配置。其中，与动词相组合的论元的顺序不是固定的，这就意味着 (48) 中所有的顺序可以不通过任何转换而直接生成。²³ 这类观点由 Fanselow (2001) 提出来分析 GB 理论。²⁴ 有关组成成分的位置的不同观点，请看 Fanselow (1993)。

3.6 总结

GB 理论和最简方案 (参看第 4 章) 中的一些重要文献在语言的共性与个性方面都有一些新的发现。下面，我将重点分析德语句法。

在转换语法下有关动词移位的分析是由 Bierwisch (1963: 34) 提出的。Reis (1974)、Koster (1975)、Thiersch (1978: Chapter 1) 和 den Besten (1983) 几乎在所有语法模型 (除了构式语法 和 依存语法) 中都已经成为标准的分析了。

Lenerz (1977) 有关成分顺序的分析影响到其他框架下的分析 (GPSG 和 HPSG 中的线性规则可以追溯到 Lenerz 的分析)。Haider 有关成分序列、格和被动的研究 (1984; 1985b; 1985a; 1986a; 1990b; 1993) 对德语的 LFG 和 HPSG 分析产生了重要的影响。

句法配置方面的整个讨论都是非常重要的，比如说我们假定德语中定式动词的主语是在 VP 内部好一些，还是外部好一些 (例如 Haider 1982; Grewendorf 1983; Kratzer 1984, 1996; Webelhuth 1985; Sternefeld 1985b; Scherpenisse 1986; Fanselow 1987; Grewendorf 1988; Dürscheid 1989; Webelhuth 1990; Oppenrieder 1991; Wilder 1991; Haider 1993; Grewendorf 1993; Frey 1993; Lenerz 1994; Meinunger 2000)，以及德语的宾格动词第一次在 GB 中得到了详细的讨论 (Grewendorf 1989; Fanselow 1992a)。Fanselow 和 Frey 有关成分序列的工作，特别是有关信息结构的研究，在相当程度上提高了德语句法的研究水平 (Fanselow 1988, 1990, 1993, 2000a, 2001, 2003b,c, 2004a;

²² 这些作者是在最简方案下提出这些理论的。这就意味着没有严格意义上的表层结构。我只是简单地将他们的分析翻译为这里所使用的术语。

²³ 我们将这个与第 51 页 (6) 中的语法相比较。这个语法包括一个 V 和一个 NP 来构成一个新的 V。由于在短语结构规则中没有提及论元的格，NPs 可以按照任意顺序来与动词相组合。

²⁴ 基础生成分析在 HPSG 框架下是一种自然分析。它是由 Gunji 在 1986 年提出来分析日语的，我们会在第 9.4 节来详细讨论这一方法。Sauerland & Elbourne (2002: 313–314) 声称他们认为句法是派生的，即有一系列句法树需要被推导出来。我认为事实上这是不大可能的。比如说，Kiss (2001) 提出的分析就说明了辖域问题可以在基于约束的理论中得到很好的解释。

Frey 2000, 2001, 2004a, 2005)。不定式结构、复杂谓词结构和部分前置都在 GB/MP 框架下得到了详细讨论与成功的分析 (Bierwisch 1963; Evers 1975; Haider 1982, 1986b, 1990a, 1991, 1993; Grewendorf 1983, 1987, 1988; den Besten 1985; Sternefeld 1985b; Fanselow 1987, 2002; von Stechow & Sternefeld 1988; Bayer & Kornfilt 1989; G. Müller 1996a, 1998; Vogel & Steinbach 1998)。在第二谓词方面, 特别值得指出的是 Winkler (1997) 的工作。

以上有关语法的不同理论方面的著作并不全面 (与我个人的研究兴趣密切相关), 而且大部分是关于德语的。当然, 对于其他语言和现象来说, 肯定还有其他有价值的文献, 只不过我们在这里不一一列出了。

在这一节剩下的部分, 我将重点讨论两点: 原则与参数理论下语言习得的模型与乔姆斯基式语言学内形式化的程度 (特别是近几十年的研究成果) 其中有些内容还会在第二部分再次提及。

3.6.1 有关语言习得的解释

乔姆斯基式研究语法的一个重要目标就是解释语言习得。在 GB 理论中, 我们假定一套非常简单的规则, 它可以适用于所有的语言 (\bar{X} theory), 也就是适用于所有语言的普遍原则, 但是对于个别语言和语言类别来说还需要一些参数来调整。这样, 一个参数被认为与某些现象相联系。原则与参数模型特别丰富, 它在解释语言之间的共性与差异的时候得到了一些有趣的研究成果。在语言习得方面, 参数是由输入所决定的这一观点倍受争议, 因为它不能在观察到的现象中得到解释: 在设定参数后, 学习者应该立刻掌握语言的某些方面。citet[146]Chomsky86 应用开关的比喻来进行说明, 要么开启, 要么关闭。正如他们所想的, 语言中的很多方面都由参数锁定, 设定一个参数会对给定学习者的语法的剩下的部分产生重要的影响。但是, 儿童的语言行为并不像我们所期待的那样会根据参数而突然发生改变 (citealp[731]Bloom93a; Haider 1993: 6; Abney 1996: 3; Ackerman & Webelhuth 1998: Section 9.1; Tomasello 2000, 2003; Newmeyer 2005)。进而, 我们无法证明在某个参数不同的语法现象之间存在联系。更多的内容, 请参考第 16 章。

无论如何, 原则与参数模型对于跨语言的研究都保有极大的兴趣。每一个理论都必须解释为什么英语中动词在宾语前面, 而日语中则在宾语后面。我们可以将这个不同看作是一个参数, 并将语言进行相应的分类, 但是这是否真正地 with 语言认知是有关系的还是存在很大的疑问的。

3.6.2 形式化

Bierwisch 在他发表于 1963 年的有关转换语法的那篇文章中指出:²⁵

²⁵ Es 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 Aufgabe. 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.

3 转换语法-管辖与约束理论

很有可能的是，由我们构造的规则所生成的句子出乎意料地是不合乎语法的，也就是说，它们由于某些属性而违反了合格性，而这些属性是在我们的检验中要特别排除的。这就意味着，语法是对一种语言的结构假说。对于自然语言的语法的合格性的系统诊断显然不能再由人工来解决。这一任务可以由计算机按照计算任务来进行实现，这样就可以对所描述语言派生出的结果进行检验。(Bierwisch 1963: 163)

Bierwisch 所说的这些观点在过去几十年经验主义的研究中更为正确。比如说，Ross (1967)指明了移位和长距离依存所需的限制条件，Perlmutter (1978) 在 70 年代发现了非宾格动词。德语的情况请看Grewendorf (1989) 和 Fanselow (1992a)。除了对于这些现象的分析，(Lenerz 1977)提出了可能的短语成分位置的限制条件，还有格指派分析 (Yip, Maling & Jackendoff 1987; Meurers 1999c; Przepiórkowski 1999b)，动词性复杂结构的理论和短语中成分的前置问题 (Evers 1975; Grewendorf 1988; Hinrichs & Nakazawa 1994; Kiss 1995; G. Müller 1998; Meurers 1999b; Müller 1999a, 2002a; De Kuthy 2002)。所有这些现象都是彼此联系的！

再来看其他内容：

早期语言学的研究目标，也是计算语言学中语言学部分的核心目标，是开发出一种语法能够针对英语中的每一句话都指派给一个合理的句法结构。在理论语言学中，这一目标并不重要。特别是在管辖约束理论中，为了追求语法的深层原则早就抛弃了大规模的语法片段。问题在于，对于正确剖析的确认不能通过检查小片段的行而决定，但是它们是在深层分析的。大的片段不仅仅是小片段的倍数——当我们开始研究大片段的时候，就已经有质的变化了。当语法能够适应的结构增加，不需要的句子的剖析结构也显著地增加。(Abney 1996: 20)

所以，正如 Bierwisch 和 Abney 所指出的，开发一个人类语言的绝大部分的好用的理论是非常艰巨的任务。但是理论语言学家所要实现的则更多：他们的目标是构建出能够完美地解释所有语言的限制条件，或者至少对某些语言类型来说是这样的。这样的话，我们就需要对不仅仅是一种语言，而是许多语言的交互关系有一个整体的认识。这个任务太难了，以至于个体的研究者不能胜任。在这点上，计算机可以提供帮助，因为他们可以立即指出理论中的不一致。在去掉这些不一致之后，计算机实现程序可以用来系统地测试数据或者语料库，进而检查理论在经验上是否是自足的 (Müller, 1999a: Chapter 22; 2015a; 2014d; Oepen & Flickinger 1998; Bender 2008b, see Section 1.2)。

在乔姆斯基第一篇重要著作发表的 50 多年后，很显然目前并没有基于转换语法分析的大规模语法片段。毫无疑问，乔姆斯基对语言的形式化作出了重要的贡献，并且他开发出了重要的形式基础理论，这些方面与计算机科学和理论计算语言学中有关形式语言的理论仍是密切相关的 (Chomsky 1959)。但是，在 1981 年，他已经反对严格的形式化了：

我想我们实际上开始得到语法的一些基本原则，这些原则也许处于合适的抽象层面上。同时，我们有必要对它们进行检验，并且通过开发十分具体的机制来测试他们在事实方面的充分性。然后，我们应该尽量区分具有前瞻性的思想的讨论与选取了具体实现形式的讨论。(Chomsky 1981a: 2-3)

上面这段话在写给 *Natural Language and Linguistic Theory* (《自然语言与语言学理论》) 的一封信中进行了更为清晰的阐述:

即使在数学中, 我们所谓的形式化的概念直到一个世纪之前尚未提出来, 这一概念对提高研究和理解的水平都至关重要。毫无疑问, 语言学已经比 19 世纪的数学和支持 Pullum 的禁令有用的当代分子物理学更为先进了, 只要这是可以展示出来的, 自然是好的。目前, 据我所示, 我们有着生动的互动与令人激动的进展, 但是没有没有任何迹象表明目前所进行的工作中有关于形式化层面的问题。(Chomsky 1990: 146)

这种与严格意义上的形式化的背离导致在主流的生成语法下相当多的文献有时会得到不一致的观点, 这些观点导致我们无法清晰地将不同文献的观点整合起来。比如说核心概念管辖就有几种不同的界定(概述请看 Aoun & Sportiche 1983²⁶)。

这种情况从 80 年代开始就不断受到争议, 而且有时 GPSG 的支持者提出的意见更为尖锐(Gazdar, Klein, Pullum & Sag 1985: 6、Pullum 1985, 1989a、Pullum 1991: 48、Kornai & Pullum 1990)。

主流的生成语法内部这种缺乏精确的和细节的研究²⁷, 以及对基本假设不断的修订²⁸ 导致他们的研究成果极少能够应用到计算实现中。有一些基于转换语法、GB 理论或 MP 模型, 以及借鉴了主流的生成语法思想的一些实现系统(Petrick 1965; Zwicky, Friedman, Hall & Walker 1965; Kay 1967; Friedman 1969; Friedman, Bredt, Doran, Pollack & Martner 1971; Morin 1973; Marcus 1980; Abney & Cole 1986; Kuhns 1986; Correa 1987; Stabler 1987, 1992, 2001; Kolb & Thiersch 1991; Fong 1991; Crocker & Lewin 1992; Lohnstein 1993; Fordham & Crocker 1994; Nordgård 1994; Veenstra 1998; Fong & Ginsburg 2012),²⁹。但是, 这些实现系统通常不用转换或者很大程度上与出版提出的理论假设不一致。例如, Marcus (1980: 102–104) and Stabler (1987: 5) 应用特别的规则来描述助词置换的问题。³⁰ 为了分析例 (51a) 这个句子, 这些规则将 John 和 has 的顺序颠倒, 这样就得到了例 (51b) 中的顺序, 这句话就可以应用非转换结构的原则来剖析。

- (51) a. Has John scheduled the meeting for Wednesday?
AUX John 安排 DET 会议 PREP 星期三
'John 安排了星期三的会了吗?'
- b. John has scheduled the meeting for Wednesday?
John AUX 安排 DET 会议 PREP 星期三
'John 安排了星期三的会了吗?'

²⁶ 进一步的定义可在 Aoun & Lightfoot (1984) 中找到。不过, 这与 Postal & Pullum (1986: 104–106) 中的早期版本是一致的。

²⁷ 请看 e.g., Kuhns (1986: 550)、Crocker & Lewin (1992: 508)、Kolb & Thiersch (1991: 262)、Kolb (1997: 3)、Freidin (1997: 580), 以及针对后者的 Veenstra (1998: 25, 47)、Lappin et al. (2000a: 888) 和 Stabler (2011a: 397, 399, 400)。

²⁸ 参考 e.g., Kolb (1997: 4)、Fanselow (2009) 和第 159 页引用 Stabler 的话。

²⁹ 请看 Fordham & Crocker (1994) 将统计方法整合进 GB 理论中的研究。

³⁰ Nozohoor-Farshi (1986, 1987) 指出, Marcus 的剖析器只能剖析上下文无关语言。由于自然语言更为复杂(看第 17 章), 而且相应复杂度的语法可以在目前的转换语法的框架内被允准, Marcus 的剖析器既不是乔姆斯基理论下的充分的实现, 也不是分析自然语言整体的一个软件。

这些针对助动词置换的规则非常明确，而且清晰地指称助动词的范畴。然而，这些规则与 GB 理论下提出的分析没有丝毫关系。正如我们在第 3.1.5 节看到的，我们并没有针对助动词置换提出特别的转换规则。助动词置换是由更为普遍的转换 *Move- α* 和相关的限制原则而实现的。规则的构成指向助动词的类别并不是没有问题的，正如 Stabler 受到 GB 启发提出的短语结构语法：

- (52) a. $s \rightarrow \text{switch}(\text{aux_verb}, \text{np}), \text{vp}$.
 b. $s([\text{First}|\text{L0}], \text{L}, \text{X0}, \text{X}) \text{ :- } \text{aux_verb}(\text{First}),$
 $\text{np}(\text{L0}, \text{L1}, \text{X0}, \text{X1}),$
 $\text{vp}([\text{First}|\text{L1}], \text{L}, \text{X1}, \text{X})$.

我们将例 (52a) 转换为 (52b) 中的 Prolog 谓词表达式。s 后的表达式 $[\text{First}|\text{L0}]$ 对应的字符串需要被处理。‘|’-操作符将列表分为开头和其他剩余部分。“First” 是最先处理的词，L0 则包含其他词。在例 (51a) 的分析中，第一个词是 “has”，L0 是 “John scheduled the meeting for Wednesday”。在 Prolog 子句中，随后测试了 First 是否是助动词 ($\text{aux_verb}(\text{First})$)，而且如果是的话，那么就需要证明 L0 序列由一个名词短语开头。因为 “John” 是一个 NP，这样就是成功的。L1 是分析所剩下的 L0 的子集，也就是 “scheduled the meeting for Wednesday”。这组词随后跟助动词 (First) 相组合，现在就需要检查剩下的这组词 “has scheduled the meeting for Wednesday” 是否由 VP 开头。事实是这样的，而且剩下的列表 L 是空的。最后，这个句子被成功地处理了。

这一分析的问题在于每个词都在词典中进行了核查。诸如 (53) 这样的句子³¹就无法分析了：³²

- (53) Could or should we pool our capital with that of other co-ops to
 AUX 或 AUX 我们合作我们的资本 PREP CONJ PREP 其他的合作者 INF
 address the needs of a regional “neighborhood”?
 应对 DET 需求 PREP ART 区域的 邻居
 ‘我们是可以还是应该将我们的资金与其他合作者的资金一起用来解决区域性的 “邻居” 的需求?’

在这类句子中，两个情态动词并列在一起。二者构成了一个 X^0 ，且在 GB 理论下，二者可以一起移动。如果我们想将这些案例按照 Stabler 那样简单地处理，那么我们就需要将这些词处理为两个无限的子集，并且检查第一组词是否包含一个助动词或者几个并列的助动词。我们需要一个递归的谓词性助动词，它们可以在一定程度上检查 *could or should* 这个词语序列是否是合格的助动词序列。这并不能通过指定一个特殊的谓词来实现，而是需要通过表示助动词并列的句法规则来实现。与 (52a) 类似的另一条规则可以是 (54)，这条规则用在像 GPSG (Gazdar et al. 1985: 62)、LFG (Falk 1984: 491)、一些 HPSG 的分析 (Ginzburg & Sag 2000: 36) 和构式语法中 (Fillmore 1999)：

- (54) $s \rightarrow v(\text{aux+}), \text{np}, \text{vp}$.

³¹ <http://www.cooperativegrocer.coop/articles/index.php?id=595>. 2010 年 3 月 28 日。

³² 有关词素并列是语言学理论的一种观点的讨论，请看 Abeillé (2006)。

这条规则对于例(53)这类并列语料来说是没有问题的,因为多个助动词的并列可以构成 $v(\text{aux}+)$ 这个范畴的对象(更多有关并列的研究请看第21.6.2节)。如果置换需要一个像(52a)的特殊的规则来操作,那么就不清楚为什么我们不能简单地应用(54)的有更少转换的规则。

在MITRE系统(Zwicky et al. 1965)中,有针对表层结构的具体语法,这一表层结构是从深层结构通过将转换的应用反过来推导出来的,也就是说,不用一个文法来创造深层结构,然后转换为其他结构,我们需要两个文法。由剖析器决定的深层结构被用来当作转换成分的输入,因为这是真正确保表层结构是由基本结构推导出来的(Kay 2011: 10)。

本章还讨论了其他与基于转换的分析不同的应用实现。比如说,Kolb & Thiersch (1991: 265, Section 4)得到这样的结论,一个陈述的、基于约束的方法对于GB来说比一个推导的方法更为有效。Johnson (1989)提出了推导式剖析方法,该方法重建了GB的子理论(\bar{X} 理论、Theta-理论、格语法……)作为逻辑表达式。³³ 这些方法可以在逻辑的基础上各自独立地应用。在Johnson的分析中,GB理论被看作是基于约束的系统。更多普遍性的约束条件从S-结构和D-结构中限制条件抽取出来,这样就可以随后直接用于剖析了。这就意味着转换并不是直接由剖析器实现的。如Johnson所指出的,他所模拟的语言片段是非常小的。比如说,它没有针对 *wh*-移位的描述(第114页)。

在GB和语障(GB之后的理论,请看Chomsky 1986a)传统下的最详细的应用实现可能就是Stabler的Prolog实现了(1992)。Stabler的成就自然是十分显著的,但是他在书里也这样声称:Stabler必须简化语障(Barriers)理论中没有明确说明的很多事情(比如说在对 \bar{X} 理论形式化的时候,应用特征-值对儿,这是借鉴了GPSG理论),而且有些假设不能有效地形式化,并被简单地忽略了(更多细节请看Briscoe 1997)。

满足一定要求的GB分析可以被重建,这样他们就不再应用转换条件了。这些不用转换的方法也叫做表征(representational),而基于转换的方法叫做推导(derivational)。对于表征分析来说,只有表层结构有语迹,但是这些结构都没有通过转换与深层结构联系起来(请看e.g., Koster 1978; 1987: 235; Kolb & Thiersch 1991; Haider 1993: Section 1.4; Frey 1993: 14; Lohnstein 1993: 87–88, 177–178; Fordham & Crocker 1994: 38; Veenstra 1998: 58)。这些分析可以按照相应的HPSG分析来进行计算处理的实现(请看第9章),实际上我们有对德语动词位置的例子。³⁴ 但是,这类应用分析与GB分析在基本框架和一些小且重要的方面(比如说如何处理长距离依存和并列(Gazdar 1981b))是不同的。有关内容和转换语法内的移位分析的分类,请看Borsley (2012)。

在上面富有争议的概述内容之后,我要加一条评论以避免可能的误解:我不是要求所有的语言学的工作都应该是形式化的。比如说,对于一篇三十页的文章来说,这是不可能的。而且,我并不认为所有的语言学家都应该做形式化的工作,并且将他们的分析应用于计算模型之中。不过,总要有有人做形式化的细节工

³³ 请看Crocker & Lewin (1992: 511)和Fordham & Crocker (1994: 38)有关其他的基于约束的推导式的剖析方法。

³⁴ 这说明了ten Hacken将HPSG、GB和LFG相比较的不同的地方,而且这些不同研究范式下的框架的类型完全是错误的。在他的分类中,ten Hacken主要是指HPSG认定的理论模型的形式化方法。但是,LFG也是一个理论模型的形式化理论(Kaplan 1995)。而且,GB也有一个理论模型的变体(Rogers 1998)。更多的内容,请看第14章。

3 转换语法-管辖与约束理论

作，而且这些基础的理论假设也应该在我们共同的研究领域内在相当充分的时间
内被接受和认可。

思考题

1. 请举例说明功能范畴与词汇范畴。
2. 如何用二元特征来表示词汇范畴？这样做有什么好处？

练习题

1. 请画出下列句子的句法树：

- (55)
- a. dass die Frau den Mann liebt
CONJ DET.NOM 女人 DET.ACC 男人 爱
‘这个女人爱这个男人’
 - b. dass der Mann geliebt wird
CONJ DET.NOM 男人 爱 AUX
‘这个男人被人爱’
 - c. Der Mann wird geliebt.
DET.NOM 男人 AUX 爱
‘这个男人被人爱。’
 - d. dass der Mann der Frau hilft
CONJ DET.NOM 男人 DET.DAT 女人 帮助
‘这个男人帮助这个女人’
 - e. Der Mann hilft der Frau.
DET 男人.NOM 帮助 DET.DAT 女人
‘这个男人正在帮助这个女人。’

请这样分析被动句：位于主语的名词短语是从宾语的位置移动而来的，也就
是说，主语并不是一个空的虚指代词。

延伸阅读

在第 3.1 节到第 3.5 节，我参考了 Peter Gallmann 2003 年以来的文献材料。但是，
我在很多地方做了一些调整。对于其中的任何错误与不足，我个人负全部责任。有
关 Peter Gallmann 的材料，请看 <http://www.syntax-theorie.de>。

在《句法分析视点》(*Syntaktische Analyseperspektiven*) 这本书中，Lohnstein
(2014) 展示了 GB 理论的一个变体，这与本章所讨论的问题多少有些相关性（CP/IP
和被动的基于移位的分析）。这本书里的不同章节是由主张不同理论的学者针对同

一篇新闻报道的分析而写的。这对于希望对不同理论进行比较的读者来说是特别有意思的研究。

Haegeman (1994)是对 GB 理论的综合性导论。对于可以阅读德语的读者来说,可以看看 Fanselow & Felix (1987)、von Stechow & Sternefeld (1988)和 Grewendorf (1988)所著的教科书,因为他们也提到了本书所提及的一些现象。

在乔姆斯基的很多著作中,他将没有转换的方法叫做“显著变体”。这是不合适的,因为没有转换的分析可以对基于转换的方法做出不同的预测(比如说,与并列和提取相关的方面,可以参考第 5.5 节有关 GPSG 的讨论)。在 Gazdar (1981a)中,我们可以找到有关 GB 和 GPSG 相比较的内容,以及将 GPSG 的类型作为转换语法的一个显著变体,这方面的代表有 Noam Chomsky、Gerald Gazdar 和 Henry Thompson。

Borsley (1999)和 Kim & Sells (2008)出版了英语的 GB 和 HPSG 教材。对于转换语法和 LFG 的比较,请看 Bresnan & Kaplan (1982)。Kuhn (2007)将现代推导分析与基于约束的 LFG 和 HPSG 进行了比较分析。Borsley (2012)比较了 HPSG 理论下的长距离依存和 GB 与最简方案下的基于移位的分析。Borsley 讨论了四种对于基于移位的方法而言有问题的数据:没有填充的提取、多重空位的提取(请看本书第154页的脚注57和第177页的脚注55)、填充与空位不匹配的提取以及没有空位的提取。

第四章 转换语法-最简方案

跟上一章介绍的管辖与约束理论一样，最简方案是由乔姆斯基在波士顿的麻省理工学院提出来的。Chomsky (1993, 1995b)提出，我们应该认真地对待语言进化的问题，而且我们应该能够回答出语言知识是如何成为我们天赋的一部分的问题。为了回答这些问题，他认为我们应该重新关注那些语言学分析机制所需的最精简假设的理论发展，并最终得到那些更少地针对个别语言的内在知识的模型。

与 GB 理论一样，最简方案也广为流传：全世界的理论学家都在这一框架下工作，所以说下面列出的研究者与研究机构一定是不够全面的。*Linguistic Inquiry* (《语言探索》) 与 *Syntax* (《句法》) 这两本期刊几乎只刊登最简方案方面的工作，而且读者可以从这些期刊中获知在该领域中较为活跃的人物。德语方面最为著名的学者有 Artemis Alexiadou (柏林洪堡大学)、Günther Grewendorf (2002) (缅因的法兰克福)、Joseph Bayer (康斯坦斯)、和 Gereon Müller (莱比锡)。

尽管 \bar{X} 理论具有极大的创新性，而且 GB 理论下对小句的分析也具有非常重要的影响，我们可以在本书中讨论的大部分学说中找到这些内容，但是最简方案下技术层面的工作却非常少。无论如何，技术层面的工作是十分有用的，因为在最简方案这个框架下人们已经做了很多研究，而且理解该理论的基本机制可以有效地获知该理论下对语言事实的一些有趣的研究成果。

尽管 1980 年代和 1990 年代的 GB 文献有很多相同的假设，最简方案则引发了许多不同的方法，这些方法使得我们难以跟上理论的发展。下面的内容主要参考了 David Adger 编著的教材之中的内容 (Adger 2003)。

4.1 表示形式的一般说明

最简方案框架下的理论假说是以 GB 框架下的已经完成的工作为基础的。所以，在上一章解释的很多内容可以在这一章继续讨论。但是，在基础假设方面有一些改动。该理论摒弃了最为普遍的参数原则，而是将相关区别放在特征之中。语言间的差异表现在某些特征的值上面，而且，这些特征可以是强特征，也可以是弱特征，而且特征强度也是区分不同语言的一个属性。强特征使得句法对象移向更高的位置上。读者应该对这种特征驱动的移位方法较为熟悉了，因为它是第 3.4 节中介绍的基于移位的被动分析的一部分。在被动的 GB 分析中，宾语为了获得格，必须要移到 IP 的限定语位置上。这类由于特征值的缺失而导致的移位是最简方案中的核心部分。

4.1.1 基本框架

乔姆斯基提出，只有两条操作（规则）用来整合语言对象：外部合并与内部合并。外部合并是指将 *the* 和 *book* 这两类要素组合起来，从而得到一个复杂的短语。内部合并被用来指称组成成分的移位。它应用于一个语言对象，并且从这个语言对象中取出一部分，并将之链接到对应对象的左边。外部合并与内部合并的应用可以按照任意顺序来实现。比如说，两个对象可以按照外部合并组合起来，然后一个合并项通过应用内部合并而移动到左边。得到的对象可以再与其他对象进行外部的合并，并且继续合并。如例（1）所示：

- (1) *the man who we know*
 DET 男人 CONJ 我们认识
 ‘我们认识的那个男人’

为了得到这个 NP，动词 “know” 与它的宾语 “who” 通过外部合并到一起。在下面讨论的几个中间过程的合并之后，“know who” 将与 “we” 合并，最后 “who” 通过内部合并移到左边，得到了 “who we know”。这个关系小句还可以通过外部合并与 “man” 组合在一起，并且继续进行下去。

所以，最简方案与 GB 理论是不同的，它没有假设具有一个根据某个 \bar{X} 语法生成的深层结构以及一个通过 *Move- α* 从深层结构生成的表层结构。相反，它假定短语是由外部合并和内部合并（组合与移位）按照任意顺序生成一个结构，然后就被定型了。一般认为，该结构被发送到表层：一方面是发音-感知系统（AP），另一方面是概念-意念系统（CI）。AP 对应于 GB 理论中的语音形式（PF）层，CI 对应于逻辑形式（LF）层。图 4.1 描述了这种新的架构。显性句法表示通常具有可见效

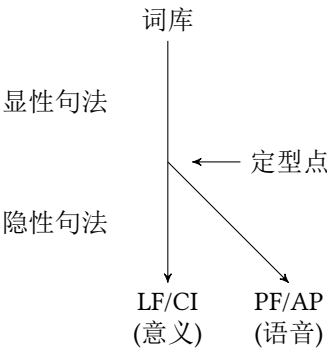


图 4.1: 在短语模型之前的最简方案的架构

应的句法操作。在显性句法之后，句法对象被发送到表层，这里在这个定型点之后有一些转换操作。因为这类转换并不会影响发音，这部分句法被叫做隐性句法。就像 GB 理论中的 LF，隐性句法被用来推导出由辖域决定的意义。

这一架构后来被修改成在推导过程中的很多点都可以作定型点。现在是这样假定的，在推导过程中包括很多阶段，一旦一个完整的短语被用在与中心语的组合

中 (Chomsky 2008), 它就被定型了。比如说, 像例 (2) 中的从句 “that Peter comes” 是一个短语, 它在整个句子完成之前被送到表层。¹

- (2) He believes that Peter comes.
 他认为 CONJ Peter 来
 ‘他认为 Peter 会来。’

对于截止到哪个范畴构成完整的短语这一问题有许多不同的看法。因为阶段这个概念对于下面的内容并不重要, 我会在下面忽略这一概念。请看第 15.1 节中有关阶段的心理语言学的可能性与普遍意义上的最简结构的讨论。

4.1.2 配价、特征核对与一致关系

最简方案的基本机制是特征核对。例如, 名词 *letters* (信) 有一个 P 特征, 这个特征是指它要与一个 PP 相结合以构成一个完整的短语。

- (3) letters to Peter
 信 PREP Peter
 ‘给 Peter 的信’

一般认为, 既有可预测的特征, 也有不可预测的特征。不可预测的特征的例子是名词的数的特征。单数和复数的区别在语义上是具有相关性的。词类的范畴特征纯粹是句法特征, 所以它不能在语义上被解读。最简方案认为所有不可预测的特征需要在复杂的语言对象的推导过程中被用到。这种对特征的穷尽过程被叫做核对 (checking)。比如说, 我们再来看名词 *letters* 的例子。例 (3) 的分析如图 4.2 所示。*letters* 的 P 特征是不可预测的, 这点由 P 前面的小 u 表示。*letters* 的 P 特征的

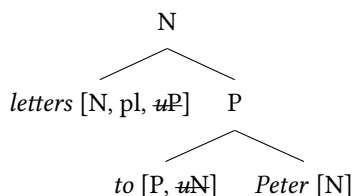


图 4.2: 不可预测特征的配价表征

不可预测性可以通过 “to Peter” 的 P 特征来进行核对。所有核对过的特征就被自动删除了。在图中, 删除通过将特征突显而得到表示。像例 (4) 的字符串被作为完整的推导被规则排除了, 因为 P 的 N 特征没有被核查。这一情况如图 4.3 on the next page 所示。

¹ Andreas Pankau (p.c. 2015) 跟我指出, 这种分段的概念有一个根本性的问题, 因为如果是这样的话, 那么只有与中心语有关系的要素被送到表层, 然后推导中的最顶层短语就再也不能被送到表层了, 因为它没有任何可供依存的中心语。

4 转换语法-最简方案

- (4) * letters to
信 PREP

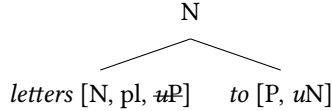


图 4.3: 基于不可预测特征的不合法的句法对象

如果这个结构能够用在拼读出的更大的结构中，那么这个推导就会崩溃了，因为概念系统不允许 N 特征还出现在 P 节点上。

选择性特征是原子式的，这就是说，在 GB 和本书中介绍的其他理论中，介词不能选择 NP[acc]，除非这个 NP[acc] 是原子式的。这样，一个附加的机制就在选择特征之外来核对其他特征。这个机制叫做一致（Agree）。

- (5) a. * letters to he
信 PREP 他
b. letters to him
信 PREP 他
‘给他的信’

例（5b）的分析如图 4.4 所示。在选择性特征核对与一致核对之间有着有趣的差别。

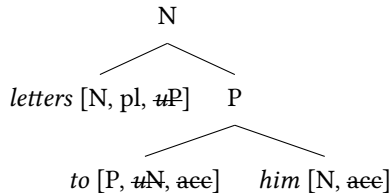


图 4.4: 一致的特征核对

一致所核查的特征不需要是与中心语相组合的对象的最高节点。这在后面被动和局部重新排序的分析中发挥着一定的作用。

4.1.3 短语结构与 \bar{X} 理论

在第 71 页的图 2.9 中给出了 \bar{X} 结构的投射。按照 \bar{X} 理论的早期版本，可以有任意多的补足语与 X^0 组合以构成 \bar{X} 。也可以有任意多的附加语附加到 \bar{X} 上，然后至多一个限定语可以组合到 \bar{X} 中，以构成一个 XP。最简方案采用二分叉结构，所以至多只有一个补足语，它是第一个合并的项目。进而，它并没有指定一个特殊的限定语位置。Chomsky 倾向于这样的认识，所有不是补足语的项目是限定语。也就

是说，他将首次合并（补足语）与后合并的项目（限定语）区分开来。图 4.5 显示了带有两个限定语的例子。也可以只有一个补足语且没有限定语，或者有一个或三

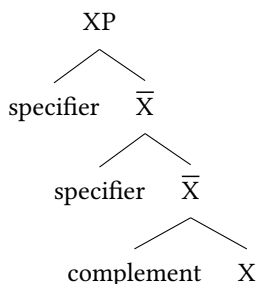


图 4.5: 最简方案下的补足语与限定语

个限定语。最终允准哪种结构在于参与到合并操作中的项目的特征。短语投射是 \bar{X} 还是 XP 取决于这个短语是否被用作另一个中心语的补足语或是限定语，或者它是否用作进一步合并操作的中心语。如果一个短语被用为限定语或补足语，它的状态就固定为一个短语 (XP)，否则得到的短语的投射状态被看作是未被限定的。在合并操作中的词汇中心语的子节点具有范畴 X ，而合并操作中的复杂中心语的子节点们具有范畴 \bar{X} 。这就解决了标准的 \bar{X} 理论方法具有代词和专有名词时的问题：必须假定出许多的一元结构（看 2.9 中的左图）。这在最简方案中就不是问题了。²

4.1.4 小 v

在第 3.4 节，我用 \bar{X} 结构来表示双及物动词，该动词带上宾格宾语构成一个 \bar{V} ，然后跟与格宾语构成一个 \bar{V} 。这种动词带宾语构成 \bar{V} 的二元结构以及平铺结构的分析被大部分采用 GB 理论和最简方案的学者所反对，因为这些分支并不是反身代词与否定极项的约束现象所需的分支结构。在例 (6a) 中，Benjamin 与 himself 具有约束关系是不可能的：

- (6) a. *Emily showed himself Benjamin in the mirror.
 Emily 展示 他自己 Benjamin PREP DET 镜子
 b. Peter showed himself Benjamin in the mirror.
 Peter 展示 他自己 Benjamin PREP DET 镜子
 ‘Peter 在镜子中给他自己看 Benjamin。’

约束关系所需要的分析，以及按照树的构造分析这些现象的这些理论中的 NPI 现象是，反身代词在树中比专有名词 Benjamin 的位置高。更准确地说，反身代词 himself 必须 c-统制 Benjamin。c-统制的定义如下所示 (Adger 2003: 117):³

² 有关这一问题请看 Brosziewski (2003: Chapter 2.1)。

³ c-统制在 GB 理论中也发挥着重要的作用。实际上，管辖与约束理论的一部分是约束理论，我们在前面的章节中没有讨论这一现象，因为本书中并不涉及管辖现象。

- (7) 节点 A c-统制 B, 当且仅当 A 的子节点满足以下条件之一:
- a. 是 B, 或者
 - b. 包括 B

在图 4.6 左边和中部的树中, 并没有所期待的 c-统制: 在最左边的树中, 所有的 NPs 互相 c-统制, 在中部的树中, “Benjamin” c-统制 “himself”, 而不是其他成分。所以说, 在左边和中间的结构是不合适的, 而且还有一些附加的结构包括范

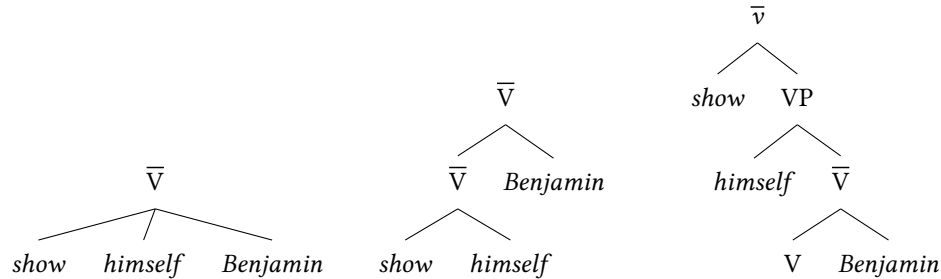


图 4.6: 双及物动词的三种可能分析

畴 v , 它被叫做小 v (Adger 2003: Section 4.4)。“himself”的子节点是 \bar{V} , 而且 \bar{V} 包括 “Benjamin”, 所以 “himself” c-统制 “Benjamin”。由于 “Benjamin” 的子节点是 V , 而且 V 既不是 “himself”, 也不包括 “himself”, “Benjamin” 并没有 c-统制 “himself”。

关于双及物动词, 早期的分析认为包括一个附加的动词性中心语 Larson (1988)。Hale & Keyser (1993a: 70) 认为, 这个动词性中心语具有致使的语义。图 4.7 中结构的生成被认为是动词 *show* 始于 V 位置, 然后移动到 v 的位置上。*show* 被认为具有 *see* 的意义, 而且在小 v 的位置上, 它具有致使含义, 这样就得到了致使-看见的含义 (Adger 2003: 133)。

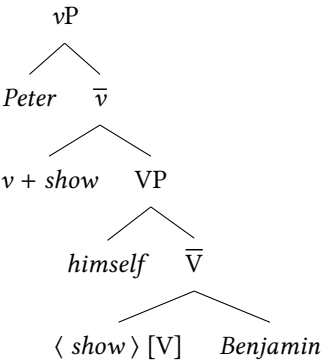


图 4.7: 移动 to 小 v 的双及物分析

尽管带有空动词中心语的动词壳分析最初是由Larson (1988)提出来分析双及物动词的, 现在它还被用来分析严格的及物动词, 甚至是不及物动词。

Adger (2003: Section 4.5)认为, 在具体的树的配置中, 语义角色的配置是不一致的。

- (8) a. vP 的 NP 子节点 → 被分析为施事
 b. VP 的 NP 子节点 → 被分析为主事
 c. \bar{v} 的 PP 子节点 → 被分析为目标

Adger 认为, 这种语义角色指派的不一致有助于语言认知的过程, 而且从此, 按照这一点, 小 v 在严格的及物和不及物动词的分析中也发挥了重要的作用。图 4.8和图 4.9显示了分别包括动词 *burn* 和 *laugh* 的句子的分析。⁴

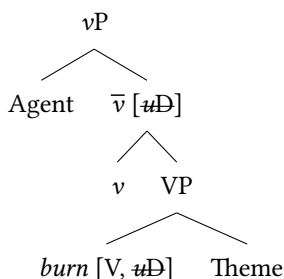


图 4.8: 包括小 v 的严格及物动词的分析

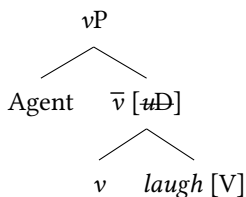


图 4.9: 包括小 v 的不及物动词分析

Adger (2003: 164)认为, 不及物和及物动词从 V 移动到小 v 的位置上。这点在下面的图中有所显示。

4.1.5 CP、TP、 vP 和 VP

第 3.1.5 节分析 GB 理论中的 CP/IP 系统。在最简方案的发展过程中, 屈折短语被分成几个功能性投射 (Chomsky 1989), 其中只有时态短语在目前的最简方案的

⁴ 如果这一类型的所有不及物动词都被认为是具有作主语的主事, 那么就需要对主事做出更为宽泛的界定, 以包括 *sleep* 这类动词的主语。通常, *sleeping* 不是一个有意为之的活动。

分析中有所涉及。所以，最简方案的 TP 对应于 GB 分析中的 IP。除了这一变化，CP/IP 分析的核心思想被转化为英语的最简方案分析。这一小节将先讨论触发移位的特殊特征（第 4.1.5.1 小节），然后是格指派（第 4.1.5.2 小节）。

4.1.5.1 特征作为移位的触发语：T 的 EPP 特征

在 GB 理论中，情态动词和助词被分析为范畴 I 的成员，主语是 IP 的限定语。在上一节，我说明了主语是如何被分析为 vP 的限定语的。现在，如果我们假设有一个情态动词，包括这样一个 vP，且主语在情态动词后面，这一语序与观察到的英语的语序要求是不一致的。要解决这一问题，可以假定在 T 上有一个强势的不可解释的 D 特征。因为特征很强，一个合适的 D 必须要移动到 T 的限定语的位置上，然后在域内核查 D。图 4.10 显示了在例（9）的分析中 TP 所发挥的作用：

- (9) Anna will read the book.
Anna 将 读 DET 书
‘Anna 将要读书。’

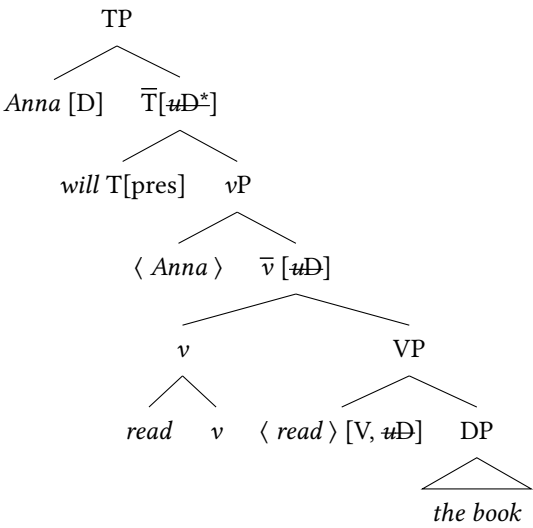


图 4.10: 包括情态词和从 v 到 T 的主语移位的“Anna will read the book”的句子分析

DP “the book” 是 read 的宾语，然后核查 read 的 D 特征。小 v 选择了主语 Anna。因为 T 有着强势的 D 特征（由星号“*”标记），Anna 一定不能在 vP 内部，但是会移动到 TP 的限定语位置上。

完整的句子是 CPs。针对例（9）的分析，空 C 的中心语被认为是与 TP 组合在一起。空 C 贡献了从句的类型特征 Decl。例（9）的完整分析如图 4.11 所示。

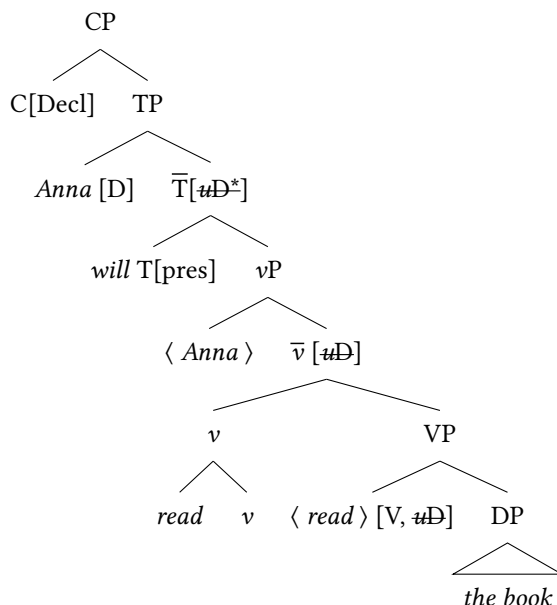


图 4.11: 带有空 C 和小句类型特征 Decl 的 CP 的 “Anna will read the book” 的句子分析

例 (10) 中间句的分析包括对于问句 (question) 的句子类型的未指派值的 T 的小句类型特征。

- (10) What will Anna read?
 什么 将 Anna 读
 ‘Anna 要读什么?’

空补足语 C 具有 Q 特征，它可以给 T 的小句类型特征赋值。由于 T 的小句类型特征具有强势的 Q 值，T 元素必须要移到 C 来局部核查。另外，wh 元素被移位了。这个移位是由 C 上的强 wh 特征决定的。例 (10) 的分析如图 4.12 on the following page 所示。

4.1.5.2 格指派

在第 3 章介绍的 GB 分析中，主格是由 (定式) I 所指派的，而其他格是由动词指派的 (请看第 3.4.2 节)。主格的指派由最简方案接管，所以一般认为主格由 (定式) T 所指派。但是，在我们考虑的最简方案中，没有一个单一的动词投射，但是有两个动词性投射：vP 和 VP。现在，我们可以认为 V 指派给它的补足语宾格，或者 v 将宾格指派给它所统制的动词的补足语。Adger (2003: Section 6.3.2, Section 6.4) 认同后一种方法，因为它与所谓的非宾格动词和被动的分析是一致的。图 4.13 on page 125 显示了例 (11) 中的 TP:

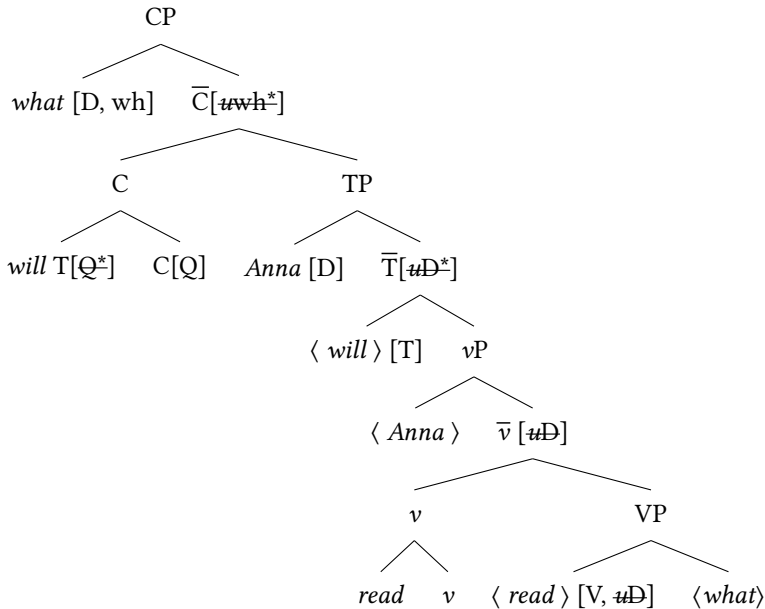


图 4.12: 带有空 C 和强 wh 特征的 “What will Anna read?” 的句子分析

- (11) Anna reads the book.
Anna 读 DET 书
‘Anna 读这本书。’

“Anna” 和 “the book” 这两个 NP 开始是没有赋值的不可解读的格特征: [u 格:]。这些特征被赋值为 T 和 v。一般认为, 只有一个特征通过合并得到核查, 所以这可以是 T 上的 D 特征, 并为其他可能的核对机制留下格特征: 一致。一致可以用来核对子节点的特征, 也可是树上较远距离的特征。第一个节点需要 c-统制跟它具有一致关系的节点。c-统制大概是指: 一个节点在上, 然后任意多节点在下。所以 v c-统制 VP、V 和 DP “the book”, 以及所有该 DP 内部的节点。由于一致可以给 c-统制的节点赋值, v 上的宾格可以给 DP “the book” 的格特征赋值。

一致内部的非局部性带来一个问题: 为什么例 (12) 是不合乎语法的?

- (12) *Him likes she.
他 喜欢她

v 的宾格可以通过它的主语得到核查, 而 T 的主格可以通过 likes 的宾语得到核查。所有的 DP 都与 T 和 v 具有必需的 c-关系。这一问题就通过要求所有的一致关系都包括最近的可能的要素而得到解决。Adger (2003: 218) 构建了如下的限制条件:

- (13) 匹配的局部性: 在 X 上的特征 F 与在 Y 上匹配的特征 F 具有一致关系, 当且仅当没有介于中间的 Z[F]。

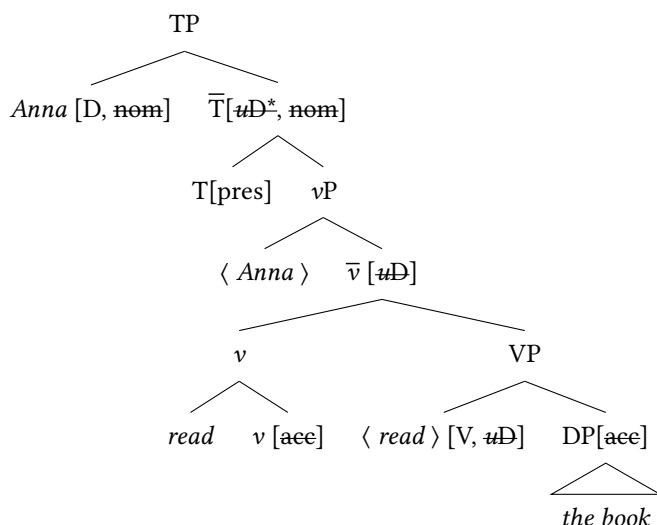


图 4.13: T 的格指派和 “Anna reads the book” 这句的 TP 中的 v

这种介于关系在例 (14) 中被界定为:

- (14) 介于关系: 在结构 $[X \dots Z \dots Y]$ 中, Z 介于 X 与 Y 中间, 当且仅当 X c-统制 Y。

所以说, 因为 T 可能与 Anna 具有一致关系, 它一定不能与 “the book” 具有一致关系。由此, (12) 中指派给 she 的主格是不可能的, 而且 (12) 也被准确地排除了。

4.1.6 说明语

Adger (2003: Section 4.2.3) 认为说明语附加在 XP 上, 并构成了一个新的 XP。他把这一操作叫做邻接 (Adjoin)。由于这一操作并不消耗任何特征, 它与外部合并是不同的, 所以说这是引入到理论中的一个新操作, 这与 Chomsky 所主张的人类语言只使用合并作为结构构建操作是相互矛盾的。也有人提议将说明语看作是带有空中心语的特殊的副词性短语 (请看第 4.6.1 节), 并将其看作是功能性投射层级中的一部分。我个人更倾向于 Adger 在许多其他的框架下提出的解决办法: 我们用一条特殊的规则和操作来解决说明语和中心语的组合问题 (请看第 9.1.7 节有关 HPSG 框架下针对中心语说明语的组合问题)。

4.2 动词位置

根据前一节介绍的机制, 德语动词位于首位的句子的分析是比较直接的。基本观点与 GB 理论中的是一致的: 定式动词从 V 移到 v, 再移到 T, 然后到 C。移

到 T 的移位是由 T 上的强势时态特征所控制的，由 T 复杂式到 C 的移位由 T 上的小句类型特征得到加强，该特征通过 C 被赋值为强势的 Decl。例（15）的分析由图 4.14 所示。

- (15) Kennt jeder diesen Mann?
认识 每人 这 男人
‘每个人都认识这个男人吗？’

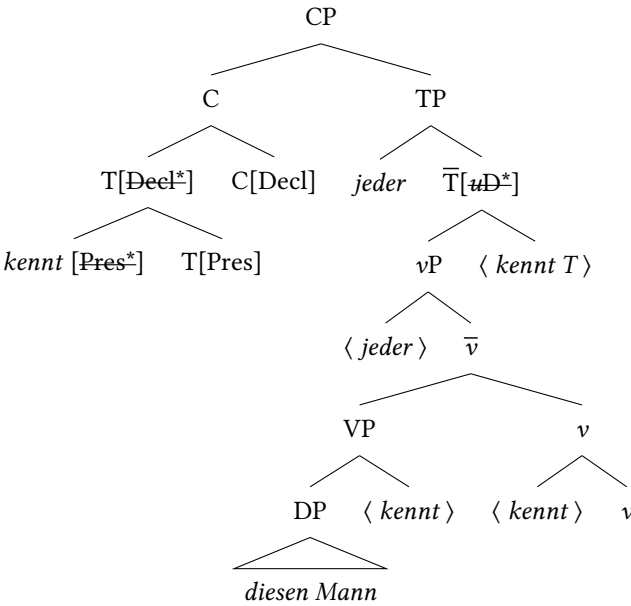


图 4.14: 在Adger (2003)的分析下有关“Kennt jeder diesen Mann?”（每个人都认识这个男人吗？）的分析

4.3 长距离依存

在解释完动词位于句首的句子之后，动词位于第二位的句子的分析就不稀奇了：Adger (2003: 331) 认为有一个特征触发了成分向 C 的限定语的位置上的移位。Adger 称这个特征为向上，但是这个术语是不恰当的，因为德语陈述句的首位并不限制为话题成分。图 4.15 在下一页显示了例（16）的分析：

- (16) Diesen Mann kennt jeder.
这 男人 认识 每人
‘每个人都认识这个男人。’

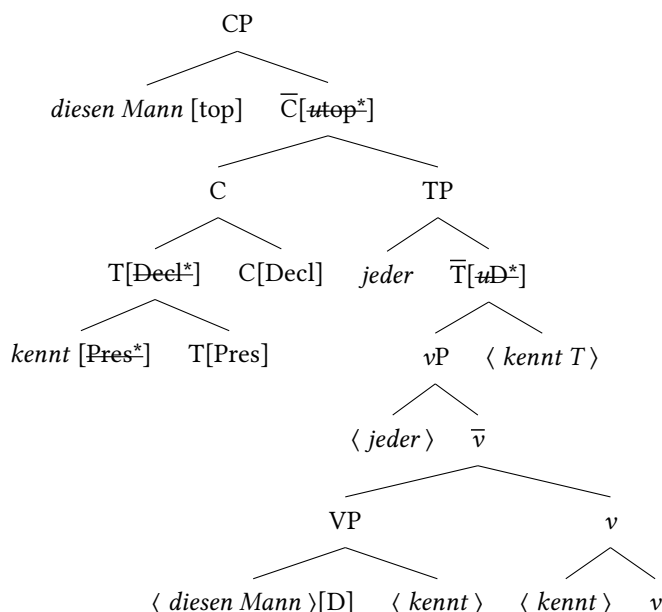


图 4.15: 在Adger (2003: 331)的分析下有关“Diesen Mann kennt jeder.”(这个男人, 每个人都认识。)的分析

4.4 被动

Adger (2003) 针对英语被动式提出了相关的分析, 这里我将其应用于德语。就像第 3.4 节讨论的 GB 分析中, 一般认为动词并没有将宾格指派给 *shalagen* (打) 的宾语。在最简方案的术语中, 这意味着小 *v* 不具有需要被核查的宾格特征。小 *v* 的这一特殊版本在所谓的非宾格动词的句子的分析中发挥了重要作用 (Perlmutter 1978)。非宾格动词是具有许多有趣特征的不及物动词的小类。比如说, 他们可以被用在形容词分词中, 尽管这在不及物动词中并不常见。

- (17) a. *der getanzte Mann
DET 跳舞 人
b. der gestorbene Mann
DET 死 人
'这个死人'

对于这一区别的解释是形容词分词说明了主动句中的宾语:

- (18) a. dass der Mann das Buch gelesen hat
CONJ DET 人 DET 书 读 AUX
'这个人读这本书'

b. das gelesene Buch
DET 读 书

现在的设想是 gestorben（死）的论元看上去像宾语，而 getanzt（跳舞）的论元像主语。如果形容词性被动式可以说明宾语，这就解释了为什么例（17b）是可能的，而例（17a）是不可能的。

Adger (2003: 140)提出了图 4.16中带非宾格动词的 vPs 的结构。一般认为，这个

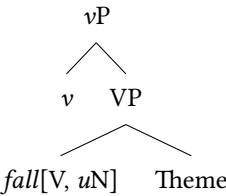


图 4.16: 在Adger (2003: 140)下，带 fall、collapse 和 wilt 这类非宾格动词的 vP 结构

小 v 的非宾格变量在被动的分析中起到了重要的作用。非宾格动词与被动动词相似，因为他们都有一个主语，这些主语在某种程度上也有宾语的特征。小 v 的特别版本由被动式中心语 werden 选择，这就构成了一个被动短语（缩写为 PassP）。请看图 4.17中针对（19）中的例子的分析：

(19) dass er geschlagen wurde
CONJ 他打 AUX
‘他被打了’

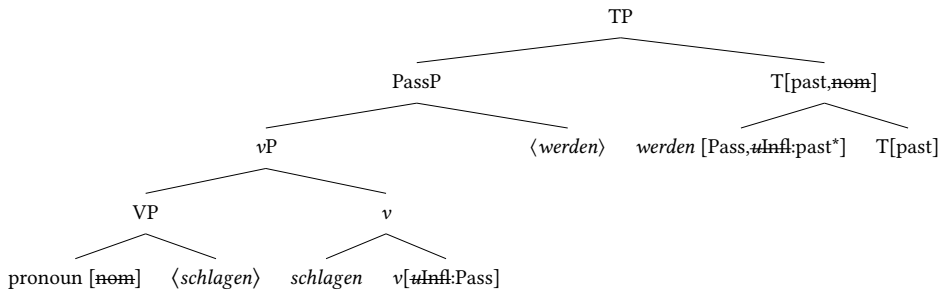


图 4.17: 基于一致的带有非域内格指派的没有移位的被动的最简方案分析

Pass 中心语要求小 v 的 Infl 特征具有 Pass 的值，这就导致了输出层面的分词形态变化。所以采用的形式是 geschlagen（打）。助词移动到 T，来核对 T 的 Infl 的强势特征，并且由于 Infl 特征是过去式，werden 的过去式形式是 wurde，该形式用于输出表层。T 有一个主格特征尚需被核查。有趣的是，最简方案并不要求 schlagen 的宾语移动到 T 的限定语位置上来指派格，因为格指派是通过一致来达成的。所

以说原则上, *schlagen* 的凸显论元可以在它的宾语位置上, 并且无论如何都会从 T 上得到主格。这就可以解决 Lenerz (1977: Section 4.4.3) 指出的 GB 分析中的问题。请看第 101 页有关 Lenerz 的例子和问题的讨论。但是, Adger (2003: 332) 认为德语在 T 上具有强 EPP 特征。如果这一假设得到支持, 那么 GB 理论下所有的问题都会延伸到最简方案的分析中: 所有的宾语都需要移到 T 上, 即使没有发生重新排序。进而, 例 (20) 这类人称被动式就会有问题, 因为没有名词短语能够为了核查 EPP 特征而移到 T 上:

- (20) *weil getanzt wurde*
 因为跳舞 AUX
 ‘因为那儿有人跳舞’

4.5 局部重新排序

Adger (2003) 并没有分析局部重新排序。但是文献中有一些其他方面的建议。因为最简方案中所有的重新排序都是特征驱动的, 所以就必须有一个特征可以触发例 (21b) 中的重新排序:

- (21) a. [*weil*] *jeder diesen Mann kennt*
 因为 每人 这 人 认识
 ‘因为每个人都认识这个人’
 b. [*weil*] *diesen Mann jeder kennt*
 因为 这 人 每人 认识

像话题短语 (Laenzlinger 2004: 222) 或提供可移动位置的 AgrS 和 AgrO (Meinunger 2000: Chapter 4) 这类功能性投射有着许多不同的看法。G. Müller (2014a: Section 3.5) 给出了一个简洁的解决方案。在他的方法中, 宾语简单地移到小 *v* 的第二个限定语位置上。这一分析在图 4.18 on the next page 中有所描述。⁵

Laenzlinger (2004: 229–230) 提出了一个观点来假定宾语的几个宾语短语可以按照任意次序排列。宾语移到这些投射的限定语位置上, 而且因为宾语短语的语序没有限制, 例 (22) 中的所有顺序都是可分析的:

- (22) a. *dass Hans diesen Brief meinem Onkel gibt*
 CONJ Hans 这 信 我的 舅舅 给
 ‘Hans 把这封信给我的舅舅’
 b. *dass Hans meinem Onkel diesen Brief gibt*
 CONJ Hans 我的 舅舅 这 信 给
 ‘Hans 给我的舅舅这封信’

⁵ G. Müller 提出了 *v* 的可选特征和触发局部重新排序的 V (第 48 页)。这些在图中没有显示。

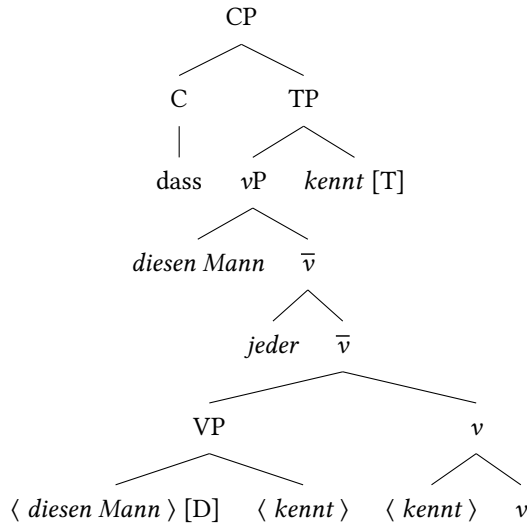


图 4.18: “dass diesen Mann jeder kennt” (每个人都认识这个人) 句中宾语移到 *v* 的限定语位置上的分析

4.6 新的发展与理论变体

在 90 年代初期, 乔姆斯基对 GB 理论的基本假设进行了重新思考, 并只保留了那些绝对必要的部分。在最简方案 (Minimalist Program) 中, 乔姆斯基说明了对 GB 理论进行修正的核心动因 (Chomsky 1993, 1995b)。直到 90 年代早期, 格理论、Theta-标准、 \bar{X} 理论、邻接理论、约束理论、控制理论等都属于语言的内在机制 (Richards 2015: 804)。当然, 这就涉及到了非常具体的语言知识是如何进入我们的基因组的问题。最简方案沿着这一思路, 并且试图解释更为普遍的认知原则下的语言属性, 以及减少具体的内在语言知识的数量。比如说, 深层结构和表层结构之间的区别就被取消了。移位仍是一种操作, 但是只直接用来构建子结构, 而不是在一个完整的 D-结构之后完成。语言之间的差别在于这种移位是可见的还是不可见的。

尽管乔姆斯基的最简方案应该被看作是 GB 理论的后续理论, 最简方案的支持者经常强调这样一个事实, 即最简方案并不是一种理论, 而是一个研究项目 (乔姆斯基 2007: 4、2013: 6)。在 Chomsky (1995b) 介绍这一研究项目时, 乔姆斯基提出的实际分析被理论家们热烈地评论, 并且有时会引发严重的质疑 (Kolb 1997; Johnson & Lappin 1997, 1999; Lappin, Levine & Johnson 2000a,b, 2001; Seuren 2004; Pinker & Jackendoff 2005)。不过, 我们应该承认有些评论偏离了问题的关键。

最简方案有很多分支。在下面的内容中, 我将讨论一些核心的观点, 并解释哪些部分被认为是有问题的。

4.6.1 移位、合并、特征驱动的移位与功能投射

Johnson、Lappin 和 Kolb 曾质疑过乔姆斯基系统的计算方面。乔姆斯基将经济原则引入了理论。在某些条件下，语法系统可以创造出一个任意数量的结构，但是只能是最经济的，也就是说，那些需要最少力气来产生的结构被认为是合乎语法的，也叫做转移派生式经济原则（transderivational economy）。这一假设并不需要被过分重视，实际上，它在最简方案框架下的很多研究中没有发挥重要的作用（尽管Richards (2015)在最新的有关生成的方法中用经济性进行了比较）。无论如何，乔姆斯基的理论的其他方面可以在很多近期的研究中有所发现。比如说，乔姆斯基提出将基本结构的构建允准规则的数量减少到两个：移位和合并（即内部和外部合并）。移位对应于 Move- α 操作，这点已经在第3章有所讨论，合并是将（两个）语言对象进行组合。

普遍认为的是，两个对象可以被组合起来 (Chomsky 1995b: 226)。对于移位来说，一个给定的移位操作一定会有一个动因。这个移位的原因被认为是可以核查它要移到的位置上的某些特征。这一观点早在第3.4节有关被动的分析中就有所说明：宾格宾语在被动句中不能带有格信息，进而必须移到能够接收格的位置上。这类方法也可以用在一系列其他现象中。例如，有的短语的中心语是焦点和话题范畴。德语和英语中相应的功能中心语永远是空的。尽管如此，提出这些中心语是受到别的语言中有表示话题和焦点的形态变化的启发。这一论断是合理的，只是建立在所有其他的语言中都有同样的范畴的假设上。不过，这一具有普遍性的共享部分（普遍语法，Universal Grammar, UG）跟具体的语言知识这种假设是冲突的，并且只被少数在乔姆斯基传统外的学者们所认可。即使在乔姆斯基式语言学下工作的学者而言，仍有很多像是否这样讨论问题是合适的问题被提出来，因为这只是创造循环结构的能力，这一能力负责人类应用语言的具体能力（狭义的语言的功用）——正如Hauser, Chomsky & Fitch (2002)所认为的——那么个别的句法范畴不属于普遍语法，其他语言的数据也不能用来解释另一种语言不可见的范畴。

4.6.1.1 功能投射与语言知识的模式化

移位必须由特征核查所允准这一思想导致了（静态）功能中心语数量的膨胀。

⁶ Rizzi (1997: 297) 提出了图 4.19 on the following page（也可以参考 Grewendorf 2002: 85, 240、2009）中的结构。

⁶ 这类中心语的假设并不是必要的，因为特征可以被集中，然后他们能被一起核查。在这一点，HPSG 理论在本质上是类似的，请看Sternefeld (2006: Section II.3.3.4, Section II.4.2)。

在所谓的制图方法中，每个形态句法特征都对应于一个独立的句法中心语 (Cinque & Rizzi 2010: 54, 61)。对于一个明显形式化的方法来说，在一个组合操作中有一个特征被利用了（请看Stabler (2001: 335)）。Stabler 的最简方案语法（Minimalist Grammars）在第 4.6.4 节中有更为详细的讨论。

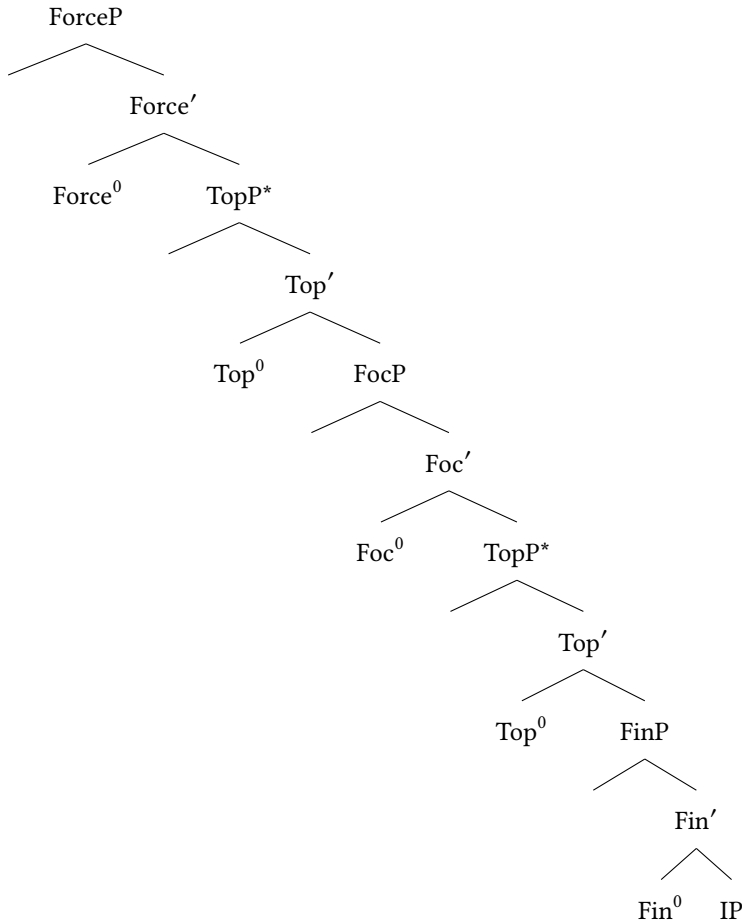


图 4.19: 在Rizzi (1997: 297)下句子的句法结构

功能范畴 Force、Top、Foc 和 Fin 对应于小句类型、话题、焦点和定式。一般认为移位总是锚定限定语位置。话题和焦点元素总是移动到相应短语的限定语位置上。话题可以在焦点元素的前面或后面，这就是为什么会有两个话题投射：一个在 FocP 上面、一个在 FocP 下面。话题短语是可循环的，也就是说，任意数目的 TopP 可以出现在图中的 TopP 位置上。根据 Grewendorf (2002: 70)，话题和焦点短语只因具体的信息结构的需要而得到实现，比如说移位。⁷ Chomsky (1995b: 147) 采纳了 Pollock (1989) 的观点，并认为所有语言都有主宾一致关系 isagreement/object 和否定 (AgrS、AgrO、Neg) 的功能投射。⁸ Sternefeld (1995: 78)、von Stechow (1996:

⁷ 对于功能投射是不是可选的具有不同的看法。有些作者认为功能投射的完整层级总是存在的，但是功能中心语可以是空的（如 Cinque 1999: 106 和 Cinque & Rizzi 2010: 55）。

⁸ 请看 Chomsky (1995b: Section 4.10.1)。

103)和Meinunger (2000: 100–101, 124)区分了直接宾语和间接宾语 (AgrO、AgrIO) 的两个一致关系的位置。同样对于 AgrS、AgrO和 Neg, Beghelli & Stowell (1997) 认为功能性中心语为了解释英语在 LF 层特征驱动的辖域现象来进行分享和分发的操作。对于没有空元素或移动的辖域现象的分析, 请看第 19.3 节。Błaszczak & Gärtner (2005: 13)提出了 -PolP、+PolP和%PolP范畴来讨论极性问题。

Webelhuth (1995: 76)对功能性投射进行了概述, 并在 1995 年提出了以下范畴的定义, 包括 AgrA AgrN、AgrV、Aux、Clitic 态、Gender、Honorific、 μ 、Number、Person、Predicate、Tense和 Z。

除了 AdvP、NegP、AgrP、FinP、TopP 和 ForceP, Wiklund, Hrafnbjargarson, Bentzen & Hróarsdóttir (2007)提出了外围 TopP。Poletto (2000: 31)针对意大利语中附着语的位置, 提出了 HearerP 和 SpeakerP。

Cinque (1999: 106)在他的研究中采用了 32 个功能性中心语, 如图 4.1所示。他认

表 4.1: Cinque (1999: 106)的功能性中心语

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)}

为所有的句子都包括一个具有所有这些功能中心语的结构。这些中心语的限定语位置可以由副词充当或者保持空位。Cinque 认为这些功能中心语和相应的结构构成了普遍语法中的一部分, 即这些结构的知识是内在的 (第 107 页)。⁹ Laenzlinger (2004)在 Cinque 的思想下, 提出德语的功能中心语语序。他还采用了Kayne (1994)的思想, 他认为所有的句法结构都有跨语言的具体语序的中心语补足语, 即使表层语序看起来是与之矛盾的。

组成成分序列最终是可见的是由左向移位推导而成的。¹⁰ 图 4.20 on page 135

⁹ 图 4.1显示了小句域内的功能中心语。Cinque (1994: 96, 99) 也说明了带有投射的形容词的顺序: 质量、大小、形状、颜色和国籍。这些范畴和他们的语序被认为与普遍语法是有关系的 (第 100 页)。

Cinque (1994: 96) 认为最有可能有七个属性形容词, 并解释了这样的事实, 在名词域中有着有限数量的功能投射。如第60页所示的, 在合适的语境中, 有可能有这七种形容词, 这就是为什么 Cinque 的功能性投射需要重复的原因。

¹⁰ 这也适用于外置, 即德语中将成分移到后场。通常这会被分析为右向移位, Kayne (1994: Chapter 9)将它分析为把所有其他成分向左移动。Kayne 认为 (i.b) 是通过移动一部分的 NP 来从 (i.a) 推导而来的:

(i) a. just walked into the room [NP someone who we don't know].
刚 走 进 DET 房间 NP 某人 谁 我们 不 认识
'刚走进房间 [NP 我们不认识的某人]'

显示了动词位于末位的小句的分析, 这里功能副词中心语被省略了。¹¹ 主语和宾语分别作为 *vP* 和 *VP* 内部的论元而生成。主语移到了主语短语的限定语位置, 而宾语移到了宾语短语的限定语位置。动词性投射 (*VP_k*) 移到了助词前, 并移到了包括助词的短语的限定语位置上。*SubjP* 和 *ObjP* 的唯一功能在于给各自移位提供落地点。对于宾语在宾语前的句子而言, Laenzlinger 认为宾语移到了话题短语的限定语位置上。图 4.20 只包含一个 *ModP* 和一个 *AspP*, 尽管 Laenzlinger 认为 Cinque 提出的所有中心语在所有德语小句的结构中都是存在的。Zwart (1994) 针对荷兰语提出了一个类似的分析, 宾语和主语从动词位于末位的 *VP* 移到了 *Agr* 位置上。

针对 Kayne 模型的争论, 请看 Haider (2000)。Haider 指出, Kayne 类的理论对德语作出了不正确的判断 (比如说对于选择的副词和第二谓词的位置, 以及动词的复杂构成), 由此就无法证明该理论可以解释所有的语言的论断。Haider (1997a: Section 4) 指出, 由 Pollock (1989)、Haegeman (1995) 和其他人提出的空否定中心语的论断是有问题的。请看 Bobaljik (1999) 关于 Cinque 的副词投射的论断的讨论。

进而, 还需要指出 *SubjP* 和 *ObjP*, *TraP* (及物短语) 和 *IntraP* (不及物短语) (Karimi-Doostan 2005: 1745) 和 *TopP* (话题短语), *DistP* (数量短语), *AspP* (情态短语) (Kiss 2003: 22、Karimi 2005: 35), *PathP* 和 *PlaceP* (Svenonius 2004: 246) 这些范畴符号都包含了语法功能、配价、信息结构和语义的信息。¹² 某种程度上, 这是对范畴符号的误用, 但是这种对信息结构和语义范畴的误用是必要的, 因为句法、语义和信息结构是紧密相连的, 而且, 语义解读句法信息, 即语义在句法之后进行 (看图 3.2 和图 4.1)。通过在句法中运用语义和语用相关的范畴, 在形态、句法、语义和语用层面上已经没有清晰的界限了: 每个元素都被“句法化”了。Felix Bildhauer (p.c. 2012) 向我指出, 意义的单个方面由节点表示的一系列功能投射的假设实际上与构式语法中的短语方法非常接近了 (Adger 2013: 470 也持这一观点)。我们只是简单地列出这些句法配置, 并且这些配置被赋予某种意义 (或在句法后被解读的特征, 如 Cinque & Rizzi (2010: 62) 有关 *TopP* 的分析)。

-
- b. Someone_i just walked into the room [*NP* _{-i} who we don't know].
 某人_i 刚 走 进 DET 房间 *NP* _{-i} 谁 我们 不 认识
 ‘某人刚走进房间 [*NP* _{-i} 我们不认识的]’

(i.a) 一定是某种推导出的中间表示, 否则英语就不会是 *SV(O)*, 而是 *V(O)S*。由此, (i.a) 是通过将 *VP* “just walked into the room” 前置由 (ii) 推导出来的。

- (ii) Someone who we don't know just walked into the room
 某人 谁 我们 不 认识 刚 走 进 DET 房间
 ‘我们不认识的某个人刚走进房间’

需要承认的是, 这类分析不能轻易地与运用模型相组合 (请看第 15 章)。

¹¹ 这些结构并不对应于第 2.5 节介绍的 \bar{X} 理论。在某些情况下, 中心语与补足语相组合以构成 *XP* 而不是 \bar{X} 。更多有关最简方案中 \bar{X} 理论的内容, 请看第 4.6.3 节。

¹² 更多例子和参考资料请看 Newmeyer (2004a: 194、2005: 82) Newmeyer 也研究了规定每个语义角色 (如施事、受益格、工具、致事、随伴格和逆行格短语) 投射的条件。

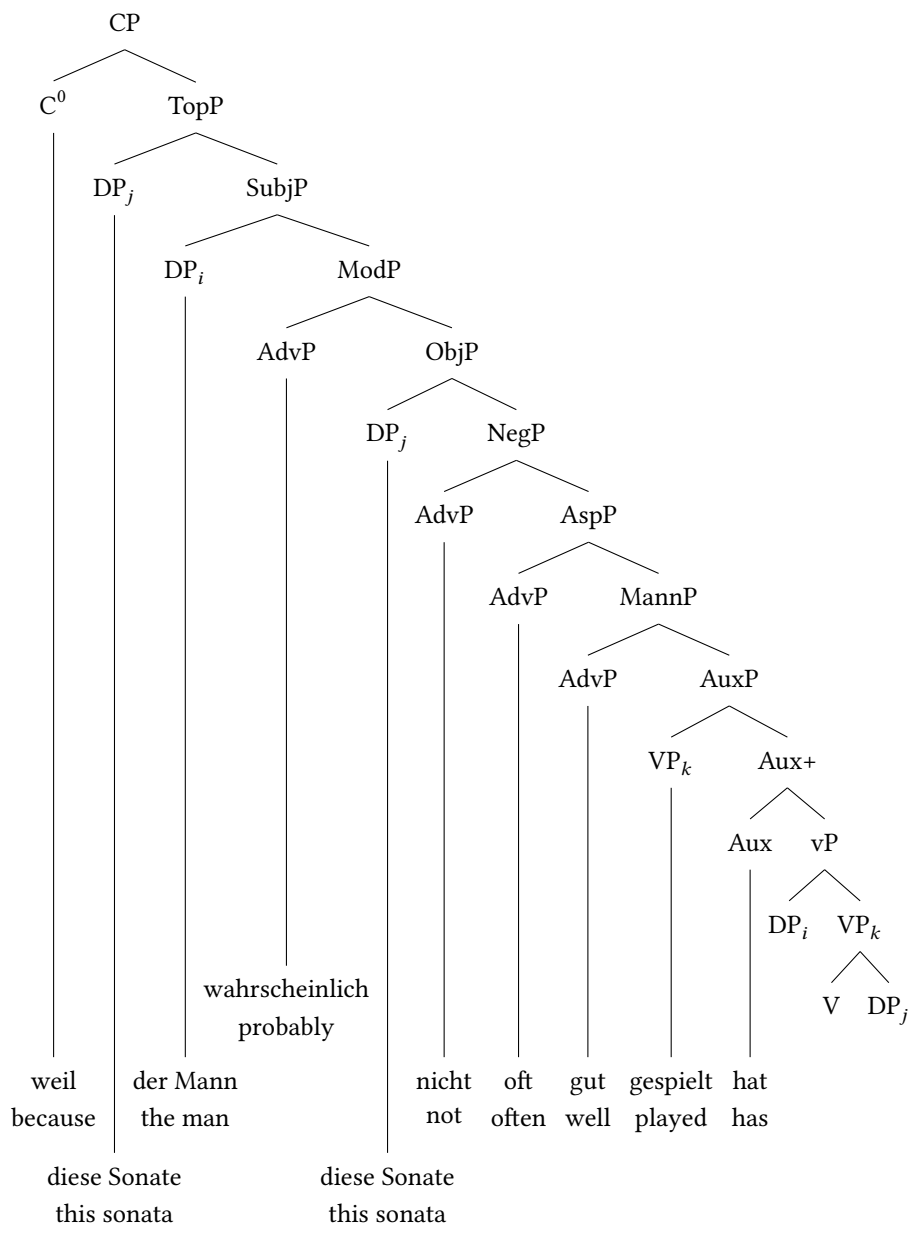


图 4.20: 在Laenzlinger (2004: 224)下的左向剩余移位和功能中心语的句子结构分析

4.6.1.2 限定语位置的特征核对

如果有人在从限定语-中心语关系到逻辑结论中采用了特征核对理论, 那么就可以得到诸如Radford (1997: 452)提出的分析结果。Radford 认为, 介词嵌套在附加于例 (23) 中的结构的一致关系短语中, 通常认为, 介词邻接于一致关系短语的中心语, 并且介词的论元移到了一致关系短语的限定语位置上。

(23) [PP P DP]

这里的问题是宾语在介词之前。为了校正这一点, Radford 提出了带有空中心语的功能投射 p (读作小 p), 这一中心语是介词连接的对象。这一分析如图 4.21所示。这一机制只在保证特征核查发生在限定语-中心语关系这一假设中是必需的。如果

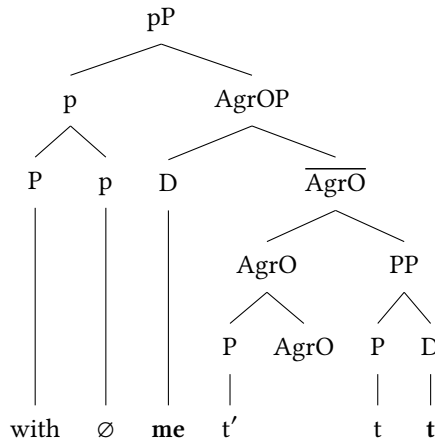


图 4.21: Radford 理论下有关限定语位置和小 p 的格指派分析

想允许介词决定它所带的宾语的格, 那么所有这些理论上的演算就没有必要了, 而且可以保证 (23) 中的结构是完好的。

Sternefeld (2006: 549–550)对这一分析提出质疑, 并将之比喻为瑞士奶酪 (充满了空洞)。这个瑞士奶酪的比喻也许太贴切了, 因为, 与瑞士奶酪不同的是, 这一分析中空洞的比率非常大 (两个词与五个空元素)。我们已经看过第 65页有关名词短语的分析了, 这里 NP 的结构, 只包括一个形容词 *klugen* (聪明), 它所包含的空元素比显性的词还多。与这里所讨论的 PP 的分析不同的是, 空成分只出现在显性的限定词和名词实际出现的位置上。另一方面, 小 p 投射是全部由理论内部推导的。对于图 4.21的分析没有任何一条附加的假设是由理论外推导而成的 (请看Sternefeld 2006: 549–550)。

这一分析的变体由Hornstein, Nunes & Grohmann (2005: 124)提出。这些作者没有利用小 p , 而是不那么复杂的结构。他们提出了例 (24) 中的结构, 对应于图 4.21中的 AgrOP-子树。

(24) [_{AgrP} DP_k [_{Agr'} P_i+Agr [_{PP} t_i t_k]]]

作者认为 DP 向 SpecAgrP 的移动是不可见的，即隐性的。这就解决了 Radford 的难题，并使得 pP 的假设多余了。

这些作者通过匈牙利语中的一致现象来推动这一分析：匈牙利语的名词短语后的后介词与其在人称和数上保持一致关系。也就是说，他们认为英语的前介词与匈牙利语的后介词短语具有同样由移位生成的结果，尽管英语中这一移位是隐性的。

按照这一思路，我们有可能减少基本操作的数量，并降低其复杂度，这一分析可以说是最简的。但是，这些结构仍然是非常复杂的。本书中所介绍的其他理论没有一种是需要用如此膨胀的结构来分析介词和名词短语的组合的。例 (24) 中的结果不能由英语数据来说明，由此无法从语言输入而获得。提出这类结构的理论需要提出一个特征只能在（特定的）限定语位置上来核查的普遍语法（看第 13 章和第 16 章有关普遍语法和语言习得的内容）。更多有关（隐性）移位的讨论请看 Haider (2014: Section 2.3)。

4.6.1.3 局部选择与功能投射

应用功能性中心语来表示语序会带来另一个问题。在经典的 CP/IP-系统和这里讨论的所有其他理论中，一个范畴代表带有同样分布的一类宾语，即 NP（或 DP）表示代词和复杂名词短语。中心语选择哪些带有一定范畴的短语。在 CP/IP-系统中，I 选择了一个 VP 和一个 NP，而 C 选择了一个 IP。在较新的分析中，这类选择机制并不能简单地奏效。因为例 (25b) 中已经有移位了，我们需要在 “das Buch dem Mann zu geben”（这个人要给的这本书）这句话中处理 TopP 或 FocP。由此，um 不能简单地选择非定式 IP，而是分别选择 TopP、FocP 或 IP。需要确定的是，TopPs 和 FocPs 标记了包含其中的动词的形式，因为 um 只能与 zu-不定式相搭配。

- (25) a. um dem Mann das Buch zu geben
 PREP DET 人 DET 书 INF 给
 ‘给那个人这本书’
 b. um das Buch dem Mann zu geben
 PREP DET 书 DET 人 INF 给
 ‘把这本书给那个人’

由此，范畴系统、选择机制和特征投射就需要被处理得更为复杂，尤其是将他们与简单基础生成的系统或从 IP 移出以构成一个新的 IP 的系统相比。

支持 Cinque (1999) 理论的观点是有问题的，原因是：Cinque 认为范畴 AdverbP 用来组合副词和 VP。这是一个空的功能中心语，它将动词性投射作为它的补足语，而且，副词表面上位于该投射的限定语位置。在这些系统中，副词短语需要传递动词的屈折属性，因为具有特定的屈折属性的动词（定式的、带 zu 的不定式、不带 zu 的不定式、助词）需要由更高位置的中心语来选择（请看第 163 页和第 9.1.4 节）。当然也有应用一致关系来表示的，但是这样的话，所有的选择关系就是非局部的，而且所有的选择关系都不是一致的。进而，这个分析的更为严重的问题有由副词修饰副词的结构，这与英语中的部分前置和位于动词前的副词性的非短语结构有关，请看 Haider (1997a: Section 5)。

4 转换语法-最简方案

副词问题的一个特殊情况是否定问题: Ernst (1992)更为细致地研究了否定的句法,并指出否定不能附加在几个不同的动词投射 (26a,b)、形容词 (26c) 和副词 (26d) 上。

- (26) a. Ken could not have heard the news.
Ken AUX 不 AUX 听 DET 新闻
'Ken 不可能听过这个新闻。'
- b. Ken could have not heard the news.
Ken AUX AUX 不 听 DET 新闻
Ken 可能没听过这个新闻。"
- c. a [not unapproachable] figure
ART 不 不可接触 人物
'一位不可接触的人物'
- d. [Not always] has she seasoned the meat.
AUX 总是 AUX 她 DET 肉
'她并不总是把肉调味。'

如果所有这些投射只是没有更多属性(有关动词形式、形容词词性、副词词性)的 NegP, 这就无法说明他们不同的句法分布。否定很明显只是更多普遍性问题的一个特殊的情况, 因为副词在传统意义上可以附加在形容词上构成形容词短语, 而不是 Chinque 意义上的副词短语。比如说, 例 (27) 中的副词 oft (经常) 修饰 lachender (笑) 构成了形容词短语 oft lachender, 这就像未被修饰的形容词助词 lachender: 它修饰 Mann (人), 并位于它前面。

- (27) a. ein lachender Mann
ART 笑 人
'一位大笑的人'
- b. ein oft lachender Mann
ART 经常大笑 人
'一位经常大笑的人'

当然我们可以想出针对上述三个问题的解决方案, 并且应用一致关系来加强非局部的选择限制, 但是这样就会破坏选择的局部限制(请看 Ernst 1992: 110 和本书第 18.2 节的讨论), 而且这比依存关系的直接选择来说要复杂得多。

与前述我们讨论的局部问题相关的是安放附着词的特殊功能投射的假设: 如果用 SpeakerP, 那么第一人称单数的附着词就可以移到正确的限定语位置上, 而对于 HearerP 来说, 第二人称附着词就可以移到正确的位置上 (Poletto 2000: 31), 然后我们需要的用来编码所有与小句有关的特征的特殊投射(另外, 我们当然可以假定有非局部的一致关系来表示分配事实)。在这些特征之外, 范畴标签包括允许更高的中心语选择包括附着词的小句的信息。在其他方法和转换语法早期变体中, 选择被认为是严格的局部限制, 这样更高的中心语只能接触到所嵌套范畴的属性, 这些范畴直接与选择相关 (citealp[223]Abraham2005a、Sag 2007), 而不是有关小句内中心语的论元是否是说话者还是听话者, 或者小句中的一些论元是否被实现为附着词。局部限制将在第 18.2 节讨论。

4.6.1.4 特征驱动的移位

最后, 对于特征驱动的移位来说有一个概念上的问题, 这一问题由 Gisbert Fanselow 指出: Frey (2004a: 27) 提出了 KontrP (对比短语), 并且 Frey (2004b) 提出了 TopP (话题短语) (请看 Rizzi (1997) 提出意大利语中的 TopP 和 FocP (焦点短语)、Haftka (1995)、Grewendorf (2002: 85, 240)、2009、Abraham (2003: 19)、Laenzlinger (2004: 224) 以及 Hinterhölzel (2004: 18) 有关德语中带 TopP 和 / 或 FocP 的分析)。构成成分根据信息结构的状态移到这些功能中心语的限定语位置上。Fanselow (2003a) 指出这种基于移位的理论对于中场的元素排序来说与最简方案的假设是不一致的。这一现象的原因是, 有时移动的发生是为了给其他元素创造出位置 (利他移位)。如果句子的信息结构要求离动词最近的宾语既不是焦点, 也不是焦点的一部分, 那么离动词最近的宾语不应在小句处得到最重音。这可以通过去重音化而得到, 即通过将重音移到另一个成分上, 或者进而, 如例 (28b) 所示, 通过将宾语从接受结构重音的位置上移到不同的位置上。

- (28) a. dass die Polizei gestern Linguisten verhaftete
 CONJ DET 经常 昨天 语言学家 逮捕
 ‘昨天警察逮捕了语言学家’
 b. dass die Polizei Linguisten gestern verhaftete
 CONJ DET 警察 语言学家 昨天 逮捕
 ‘昨天警察逮捕了语言学家’

在西班牙语中, 部分焦点可以不通过特殊的语调就能实现, 但是只能通过利他移位来将宾语移出焦点位置。请看德语中有关利他的多重前置的讨论。

为了核对由信息结构的属性驱动的一些特征, 就不可能认为成分移到了树中的特殊位置上。因为特征核对是现今最简方案中移位的前提条件, 我们就不得不假定有一个特殊的属性, 该属性只有触发利他移位的功能。Fanselow (2003a: Section 4、2006: 8) 还指出, 话题、焦点和句子副词的语序限制可以被一种理论来合理地描述, 其中, 论元首先跟他们的中心语一个接一个地组合 (最简方案中的术语是合并), 然后状语可以再附加在任意投射层上。在句中焦点位置前的句子副词有这样的语义解释: 由于句子副词像焦点敏感算子, 他们不得不直接位于他们所指的元素之前。这样, 话语 (话题) 中不属于焦点的元素需要在句子副词之前。这样就必须假定一个话题位置来解释中场的局部重新排序。这一分析在 LFG 和 HPSG 中也有所涉及。相应的分析在后面的章节中有更为详细的讨论。

4.6.2 标记

在最简方案中, 乔姆斯基尝试将组合操作及其基础简化到越简单越好。为了支撑这一假设, 他认为带有较少具体语言知识的普遍语法比带有较多具体语言支持的普遍语法而言在演化的角度来看更为可行 (Chomsky 2008: 135)。

基于此, 他取消了 X 理论中的投射层、语迹、指标和“相似的描述技术” (Chomsky 2008: 138)。剩下只有合并和移位, 即内部和外部合并。内部和外部合并将两个句法对象 α 和 β 合并进一个表示为集合 $\{\alpha, \beta\}$ 的更大的句法对象中。 α 和 β 要么是词汇项, 要么是内部复杂的句法对象。内部合并将宾语的一部分移到了它

的边缘。¹³内部合并元素的结果是集合 $\{\alpha, \beta\}$ ，其中， α 是 β 的一部分。外部合并也得到有两个元素的集合。但是，两个独立的对象被合并了。通过合并创造出的对象有一定的范畴（一套特征）。例如，如果我们将元素 α 和 β 相组合，那么我们会得到 $\{1, \{\alpha, \beta\}\}$ ，其中，1是结果对象的范畴。这一范畴也叫做标记（label）。因为我们认为所有成分都有中心语，指派到 $\{\alpha, \beta\}$ 的范畴要么是范畴 α ，要么是范畴 β 。Chomsky (2008: 145)讨论了下面两条决定集合标记的规则。

- (29) a. 在 $\{H, \alpha\}$ 中，H和LI，H是标记。
b. 如果 α 在内部合并到 β 上，构成了 $\{\alpha, \beta\}$ ，然后 β 的标记是 $\{\alpha, \beta\}$ 的标记。

正如Chomsky指出的，这些规则不是没有问题的，因为在所有的情况中，标记并不是唯一确定的。比如说有两个词汇元素的组合，如果例(29a)中的H和 α 都是词汇项(LI)，那么H和 α 可以是结果结构的标记。Chomsky意识到，这可以得到派生结构，但是他指出这一点是没有问题的，并且将之忽视了。Chomsky在他的2013这篇论文中给出了两个词汇项到组合的分析。解决这一问题的方法是界定所有的词汇项的组合都包括一个功能元素和一个根元素(Marantz 1997; Borer 2005)。根在定义上不是标记¹⁴而且功能元素范畴决定了组合的范畴(Chomsky 2013: 47)。这样的分析只能遭到反对：最简方案的目标是简化理论设想，使得语言习得和语言演化的模型变成可行的，但是为了简化基本概念，它就要设计名词不仅仅是一个名词，而是需要一个功能元素来说明名词所在的范畴。Chomsky的光杆短语结构(Chomsky 1995a)的所有设想是在 \bar{X} 理论中去除一元分支结构，我们不清楚他们为什么在现在通过借壳而被重新介绍出来，带有一个额外的空元素就更为复杂了。¹⁵诸如范畴语法和HPSG的理论可以将词汇项直接组合，无须假定任何助词投射或空元素。请看Rauh (2013)有关早期的转换语法、HPSG、构式语法、角色参数语法和Chomsky (2013)提出的基于根的新构建主义理论中句法范畴的处置的比较分析。Rauh总结道，句法和语义信息的直接连接是需要的，而且Marantz和Borer的新构建主义必须遭到反对。更多有关新构建主义的评论请看Wechsler (2008a)和Müller & Wechsler (2014a: Sections 6.1 and 7)。

代词与动词投射相组合带来的问题与上面讲的问题有关。在“He left”的分析中，代词he是一个词汇元素，所以需要为“He left”的标记负责，因为left是最简方案中的一个内在复杂动词投射。结果可以是一个名词标记，而不是动词的。为了解决这一问题，Chomsky (2013: 46)指出he有一个复杂的内部结构：“也许是D-pro”，即he（可能）是在一个可见的限定词和一个名词之外构成的。

¹³ 更为具体来说，句法对象的一部分被复制，然后该复制成分被放到整个对象的边缘位置。该复制的原型不再与发音有关（移位的复制理论）。

¹⁴ 另一个通过定义被排除为标记的范畴 Conj，它表示并列(Chomsky 2013: 45–46)。这是使得并列可以完成的规定。请看下面。

¹⁵ (i.a) 中旧的 \bar{X} 规则对应于(i.b)中的二元组合。

- (i) a. $N' \rightarrow N$
b. $N \rightarrow N\text{-func root}$

在(i.a)中，一个词汇名词投射到 N' 中，而且在(i.b)中，一个根元素与功能名词中心语相组合为一个名词范畴。

两个非词汇项外部合并的情况(如一个名词短语和动词短语)并没有在Chomsky (2008)中有所讨论。Chomsky (2013: 43–44)指出, 短语 XP 与 {XP, YP} 的标记无关, 如果 XP 进一步被移位的话(或者是在移位的复制理论中被复制)。Chomsky 认为在 {XP, YP} 组合中的两个短语中的一个要被移位, 因为不然的话就无法标记了(第 12 页)。¹⁶ 我们用下面并列的例子来进行说明。Chomsky 认为, “Z and W” 这个表达式可以这样进行分析: 首先, Z 和 W 合并。这一表达式再与 Conj 相组合 (30a), 然后在下一步 Z 被提升 (30b)。

- (30) a. [$_{\alpha}$ Conj [$_{\beta}$ Z W]]
 b. [$_{\gamma}$ Z [$_{\alpha}$ Conj [$_{\beta}$ Z W]]

由于 β 中的 Z 只是一个复制, 它并不能表示标记, 而且 β 可以获得 W 这个标记。并且要求 Z 和 α 相组合, Conj 不能是标记, 所以完整结构的标记是 Z。¹⁷

Chomsky 讨论的另一特殊情况是 $LI\alpha$ 与非 $LI\beta$ 的内部合并。按照规则 (29a), 标记应该是 α 。按照 (29b), 标记应为 β (也请看Donati (2006)) Chomsky 讨论了代词 “what” 与 “you wrote” 相组合的例子。

- (31) what [C [you wrote t]]
 什么 C 你写 t
 ‘你写了什么’

如果标记是由 (29b) 决定的, 那么在 GB 框架下就有一个被叫做 CP 的句法对象; 由于这个 CP 是疑问句, 它可以在 (32a) 中被用作 wonder 的补足语。如果标记是由 (29a) 来决定的, 我们就可以在 (32b) 中得到被用为 read 的宾格宾语的宾语, 即在 GB 术语中与 DP 相关的成分。

- (32) a. I wonder what you wrote.
 我想知道 什么 你 写
 ‘我想知道你写了什么’

¹⁶ 他的解释是自相矛盾的: 在第 11 页, Chomsky 认为两个具有相同范畴的实体的组合的标记就是这个范畴。但是在他对并列的操作中, 他认为其中一个并列项需要被移动, 否则完整的结构就不能得到标记了。

¹⁷ 正如 Bob Borsley (p.c. 2013) 跟我指出的, 这种方法对于两个用 and 连接的单一名词短语的并列来说会得出错误的结论, 因为并列的结果是一个复数 NP, 而不是像第一个并列成分的单一个体。HPSG 之类的理论可以通过将特征集中起来表示这一特点, 这些特征在并列结构中共享(句法特征和非域内特征, 请看Pollard & Sag (1994: 202))。

进而, 整个分析并不能解释为什么 (i.b) 被规则排除出去了。

- (i) a. both Kim and Lee
 都 Kim 和 Lee
 ‘Kim 和 Lee 两个人’
 b. *both Kim or Lee
 都 Kim 或 Lee

并列的信息必须是 “or Lee” 这个表达式的一部分, 这样可以用来与 “and Lee” 进行对比。

另一个问题是 α 这个标记应该是 W 的标记, 因为 Conj 并不能决定是哪个标记。这就导致, 我们需要在 Z 和 W 之间进行选择, 来决定 γ 的标记。按照 Chomsky 的逻辑, 要么是 Z, 要么是 W 必须要移位来使得标记 γ 成立。Chomsky (2013) 在脚注 40 中提到了这一问题, 但是并没有给出解决办法。

4 转换语法-最简方案

- b. I read what you wrote.
我读 什么 你 写
‘我读了你写的东西。’

例 (32b) 中的 “what you wrote” 被叫做自由关系从句。

Chomsky 关于自由关系从句的方法是有趣的，但是不能在全维度上来描述这一现象。问题是包括关系代词的短语可以是复杂的（与 Donati 的观点相比，请看 Citko (2008: 930–932)）。¹⁸ (33) 是从 Bresnan & Grimshaw (1978: 333) 中选出的英语例子。Bausewein (1990: 155) 和 Müller (1999b: 78) 中的德语例子在例 (34) 中。

- (33) I'll read [whichever book] you give me.
我-将读 任何一个 书 你 给 我
‘我会读你给我的任何一本书。’

- (34) a. Ihr könnt beginnen, [mit wem] ihr wollt.¹⁹
你 可以 开始 PREP 谁 你 想要
‘你想跟你想要的任何一个人开始。’
b. [Wessen Birne] noch halbwegs in der Fassung steckt, pflegt solcherlei
谁的 灯泡/头也 半路 PREP DET 插座 AUX 使用 这样
Erloschene zu meiden;²⁰
消逝 INF 避免
‘那些还有一半智慧的人想要避免这类空字符；’
c. [Wessen Schuhe] „danach“ besprenkelt sind, hat keinen Baum gefunden
谁的 鞋 此后 有斑点的 AUX AUX 不 树 找到
und war nicht zu einem Bogen in der Lage.²¹
和 AUX 不 INF ART 弓 PREP DET 位置
‘那些鞋上有斑点的人找不到树也不能按照弧形来小便。’

由于 “wessen Schuhe” (谁的鞋) 并不是一个词汇项，我们就需要应用规则 (29b)，而且对于这种情况没有额外的规则。这就意味着整个自由关系从句 “wessen Schuhe danach besprenkelt sind” 被标记为 CP。对于标记为 CP 的 (33) 和 (34) 中的自由关系从句并不是期望得到的结果，因为他们作为主谓词的主语或宾语，并由此被标记为 DP。但是，因为 “wessen Schuhe” 是一个复杂短语，而不是一个词汇项，(29a) 并不适用，所以就没有作为 DP 的自由关系小句的分析。所以说，看起来我们必须回到 Groos & van Riemsdijk (1981) 提出的 GB 分析，至少是针对德语的例子。Gross 和 van Riemsdijk 认为自由关系从句包括一个空名词，它像正常名词一样被关系小

¹⁸ Chomsky (2013: 47) 承认，目前关于自由关系小句中的标记还有很多问题，所以他承认这类标记相关的问题还有很多。

¹⁹ Bausewein (1990: 155).

²⁰ Thomas Gsella, taz, 1997 年 2 月 12 日，第 20 页。

²¹ taz, taz mag, 1998 年 8 月 8 日或 9 日，第 XII 页。

句来修饰。在这一方法中，关系短语的复杂度是无关的。只有空中心语是跟整个短语的标记是有关的。²²

但是，一旦空中心语在分析中被批准，(29a) 到 (31) 的应用就是不需要的了，因为应用可以导致针对 (32b) 的两个分析：一个是空名词性中心语，一个是被直接标记为 NP 的 (31)。有人可能会指出，在许多可能的推导式中，最为精简的会胜出，但是转移派生限制的假设会得到不想要的结果 (Pullum 2013: Section 5)。

Chomsky (2013) 放弃了 (29b) 中的标记条件，而是将它替换为普遍的标记规则，该规则约束两个短语的内部和外部合并。他区分了这两种情况。在第一种情况中，标记是可能的，因为集合 {XP, YP} 中的两个短语中的一个被移走了。这一情况早上面就有所讨论。Chomsky 还写了另一种情况：“X 和 Y 在相关方面是相同的，并具有相同的标记，可以被看作是 SO 的标记” (第 11 页)。他草拟了第 13 页的疑问小句的分析，这里疑问短语具有一个 Q 特征，提取出 Q 短语的剩余的句子也具有 Q 特征。这类标记也许也可以被用作正常句子的标记，它们包括主语和动词短语在人称和数上的一致关系。这些特征可以对句子的标记负责。具体的细节还没有制定出来，但是可以肯定的是比 (29b) 更为复杂。

在 Chomsky (2005) 和 Chomsky (2013) 中较为一致的属性是由合并对象之一专门确定的标记。正如 Bob Borsley 跟我指出的，这对 (35) 这类疑问/关系短语来说是有问题的。

- (35) with whom
PREP 谁
‘跟谁’

(35) 中的短语既是一个介词短语（因为第一词是介词），也是一个疑问/关系短语

²² 假设空中心语是有问题的，因为它只能在那些由状语修饰的情况下用作论元，叫做关系小句 (Müller 1999b: 97)。对于后期重新发现这一问题的也可以看 Ott (2011: 187)。这一问题可以在 HPSG 理论下得到解决，通过假定一个一元投射，投射到关系小句的合适的范畴。我也应用一元投射来分析所谓的非匹配的自由关系小句 (Müller 1999b)。在非匹配的自由关系小句的结构中，关系小句填充了论元槽的位置，该位置并不对应于关系短语的属性 (Bausewein 1990)。Bausewein 讨论了下面的例子，这里关系短语是一个 PP，但是自由关系代词填充了 kocht (做饭) 的宾格槽。

- (i) Sie kocht, worauf sie Appetit hat.
她 做饭 是什么她 胃口 AUX
‘她想吃的时候就做饭。’

请看 Müller (1999b: 60–62) 关于语料库的例子。

最简方案不能应用一元投射。Ott (2011) 提出了关系短语的范畴被投射的分析，但是他没有为非匹配自由关系小句提供解决方案 (第 187 页)。这跟 Citko 的分析是一样的，其中内部合并的 XP 可以提供标记。

还有许多其他有标记，或者更多没有标记的研究。比如说，一些最简方案研究者想要标记删除，进一步提出没有标记的句法。如 Osborne, Putnam & Groß (2011) 所指出的，这类分析使得最简方案更像依存语法了。不清楚的是这些模型中的哪些能够处理非匹配的自由关系小句。Groß & Osborne (2009: Section 5.3.3) 在他们的依存语法中给出了自由关系小句的分析，但是否认了非匹配部分的存在 (第 78 页)。他们提出，关系短语是自由关系小句的根或标记，所以他们具有与最简方案中非匹配自由小句相同的问题。正如 Groß & Osborne (2009: 73) 和 Osborne et al. (2011: 327) 所声称的：在 (他们所谓的) 依存语法中通常没有所谓的空中心语。也没有一元分支投射。这看起来使得带有关系短语 YP 的自由关系小句可以作为 XP，使得 XP 是一个在 Keenan & Comrie (1977) 的旁格层级中更高的范畴，Bausewein (1990) 发现的一般特征 (也可以看 Müller 1999b: 60–62 和 Vogel 2001: 4)。为了能够表示相关事实，需要有一个与 (i) 中的 worauf 标记不同的元素或者标记。

(因为第二个词是一个疑问/关系词)。所以,我们需要的是PPs的正确标记,像(35)中那种是覆盖了从子节点到母节点的不同属性的界定良好的形式。²³

更多最近有关合并的构成的标记和大量的过度分析请看Fabregas et al. (2016)。

总之,我们可以说,用来简化理论和减少具体语言内在知识所提出的标记,只能用在一定数量的规定中。例如,需要空功能词假设的词汇元素的组合,他们的唯一目的在于决定特定词汇元素的句法范畴。如果这与语言实体相关,有关标记的知识、各自的功能范畴,以及需要忽视标记的范畴可以被称为具体语言内在知识的一部分,而且什么都得不到。剩下的只是没有在最简方案的方向下取得进展的特别复杂的怪异分析。进而,还有很多以观察和实验为依据的问题以及大量未能解决的问题。

结论是,二元组合的标记不应该由Chomsky (2008, 2013)提出的方式来决定。另一种用来计算标记的选择是用功能论元结构中的功能函子作为标记(Berwick & Epstein 1995: 145)。这是范畴语法(Ajdukiewicz 1935; Steedman 2000)和Stabler的最简方案(2011b)中所采用的方法。²⁴ Stabler有关合并的形式化的内容将在第4.6.4节进行讨论。

4.6.3 限定语、补足语和 \bar{X} 理论的剩余部分

Chomsky (2008: 146)认为,每个中心语都只有一个补足语,但是有任意数量的限定语。在标准 \bar{X} 理论中,限制是在普遍的 \bar{X} 模式中最多只有一个补足语,而且假定这样的结构至多是二元分支的:在标准 \bar{X} 理论中,一个词汇中心语与所有它的补足语相组合以构成一个 X' 。如果在一个短语中最多有两个子节点,只能有一个补足语(双及物动词句子被分析为一个空中心语允准一个额外的论元;请看Larson (1988)有关空动词中心语的研究以及Müller & Wechsler (2014a: Sections 6.1 and 7)对包括小 v 方法的批判性评价)。在标准的 \bar{X} 理论中,只有一个限定语。这一限制现在被废黜了。Chomsky写道,限定语和补足语的区别现在可以从哪些元素与他们的中心语相合并的顺序推导出来:首先合并的元素是补足语,其他所有的(即后期合并的)是限定语。

这一方法对单价动词的句子来说是有问题的:根据Chomsky的观点,单价动词的主语不是限定语,而是补足语。²⁵有关这一问题的更多细节将在第4.6.4节详

²³ HPSG通过区分包括言语信息的中心语特征和包括提取和疑问/关系元素的信息的非局部特征来解决这一问题。中心语特征从中心语投射,母节点的非局部特征是子节点的非局部特征减去那些特定中心语或者在特定句法配置中没有约束的特征。

Citko (2008: 926)提出了一个分析,其中所有的子节点都可以为母节点贡献力量。结果是像 $\{P, \{D, N\}\}$ 的复杂标记。这是一个高度复杂的数据结构,而且Citko并没有给出任何信息,它所包含的相关信息是如何被接触到的。带有标记 $\{P, \{D, N\}\}$ 的宾语是P、D,还是N呢?我们可以说P有优先性,因为它在嵌套最少的集合中,但是D和N在一个集合中。冲突的特征怎么办呢?介词如何选择DP决定了 $\{D, N\}$ 是D,还是N?任何情况下,很清楚的是形式化会包括循环的关系,它们为了接触到他们的特征能够发掘出子集的元素。这增加了整体的复杂度,而且在HPSG中不受推崇,因为它对每个语言对象都使用了一种词性。

²⁴ 为了使范畴语法的方法奏效,有必要将范畴 x/x 指派给状语,这里 x 表示状语所附加的中心语的范畴。例如,形容词与名词对象组合成一个名词对象。这里它的范畴是 n/n ,而不是 adj 。

同样,Stabler的方法并不适用于状语,除非他想将名词范畴指派给定语形容词。解决这一问题的一个方法是假定对状语和他们的中心语有一个特殊的组合操作(请看Frey & Gärtner 2002: Section 3.2)。这一组合操作与HPSG理论中的中心语-状语范式是等效的。

²⁵ Pauline Jacobson (p.c. 2013)指出,不及物动词的问题可以这样来解决,即假定最后合并的元素是限

细讨论。

除此之外，认为与词组相合并的句法对象是限定语的分析不允许两个动词直接并列，如 (36) 所示：²⁶

- (36) He [knows and loves] this record.
他 知道 和 喜爱 这 专辑
'他知道并喜爱这张专辑。'

例如，在 Steedman (1991: 264) 提出的分析中，and（作为中心语）首先与 love 合并，然后得到的结果与 knows 合并。这一组合的结果是有所组合部分相同句法属性的复杂对象：结果是一个复杂动词，该动词需要一个主语和一个宾语。在与两个动词并列的组合之后，结果需要与“this record”和“he”组合。“this record”在所有相关方面都像补足语。但是，按照 Chomsky 的定义，它应该是一个限定语，因为它与合并的第三次应用所组合。得到的结果是不清楚的。Chomsky 认为合并不能说明成分序列。按照他的观点，线性化发生在语音形式（Phonological Form, PF）层。其中的限制在他近期的文章中并没有提及。但是，如果作为补足语或限定语的范畴在 Kayne 的工作中对线性化起到了作用，而且在 Stabler 的观点中（请看第 4.6.4 节），“this record”需要在“knows and loves”之前线性化，这是与事实相反的。这就意味着类似于范畴语法的有关并列的分析是不可行的，而且剩下的唯一选择看起来就是假定 knows 与宾语相组合，然后两个 VPs 并列在一起。Kayne (1994: 61, 67) 追随 Wexler & Culicover (1980: 303) 的观点，提出这一观点，并认为第一个 VP 中的宾语被删除了。但是，Borsley (2005: 471) 指出这一分析作出了错误的预测，因为 (37a) 可以从 (37b) 推导出来，尽管这些句子在意义上是不同的。²⁷

- (37) a. Hobbs whistled and hummed the same tune.
Hobbs 吹口哨 和 哼 DET 相同 调子
'Hobbs 吹着口哨，并哼着相同的曲调。'

定语，而所有非最后合并的元素是补足语。这就可以解决非及物动词的问题以及 (36) 中动词的并列问题。但是，它不能解决 (39) 中中心语后置语言的并列问题。进而，现在的最简方案允许多重限定语，这就与 Jacobson 式的思想不一致了，除非希望对非首先合并元素的状态施以更为复杂的限制。

²⁶ Chomsky (2013: 46) 在 (30) 中提出了并列分析：按照这一分析，动词可以直接合并，而且其中一个动词可以在生成的较后步骤中在连词前后移动。正如在上一节所指出的，这样的分析并不有利于达到关于具体语言的内在知识进行最小化的假设的目的，因为它完全无法清楚地解释这样的并列分析是如何由语言学习者习得的。所以说，我在这里不考虑并列分析。

Chomsky 2013 年的这篇文章的另一个创新之处在于，他取消了限定语的概念。他在第 43 页的脚注 27 写道：“关于限定语的问题有非常多有启发性的文献，但是如果这里的推理是正确的，它们就不存在了，而且这些问题不能用公式来表示。”这么说是对的，但是这就意味着在最简方案框架下关于限定语概念的解释到目前为止不再有意义了。如果我们按照 Chomsky 的说法，过去几十年一大部分的语言学研究都没有价值了，而且需要重新做。

Chomsky 在他早期的研究中没有严格遵守线性化的思路，但是某种程度上我们要确认那些叫做限定语的实体在被叫做限定语的成分所在的位置上实现的。这就意味着下面的评论即使在 Chomsky 式的假设中也是相关的。

²⁷ 也可以看 Bartsch & Vennemann (1972: 102)、Jackendoff (1977: 192–193)、Dowty (1979: 143)、den Besten (1983: 104–105)、Klein (1985: 8–9) 和 Eisenberg (1994b) 关于转换语法的早期版本中相似观察与相似观点的评论。

4 转换语法-最简方案

- b. Hobbs whistled the same tune and hummed the same tune.
Hobbs 吹口哨 DET 相同 调子和 哼 DET 相同 调子
'Hobbs 吹着相同的调子, 并且哼着相同的调子。'

因为语义解读不能在语音形式层面看到像删除这样的处理 (Chomsky 1995b: Chapter 3), 语义之间的区别不能通过删除材料的分析而得到解释。

在另一个 VP 并列结构的分析中, 有一个语迹与 "this record" 相关。这可以是一个右节点提升 (Right-Node-Raising) 分析。Borsley (2005) 指出这样的分析是有问题的。下面的例子选自他讨论的几个有问题的例子 (也可以看 Bresnan 1974: 615)。

- (38) a. He tried to persuade and convince him.
他 尝试 INF 劝说 和 说服 他
'他试着劝说并说服他。'
b. *He tried to persuade, but couldn't convince, him.
他 尝试 INF 劝说 但是 AUX-NEG 说服 他

如果 him 没有重读的话, 第二个例子是不合乎语法的。相较而言, (38a) 是合格的, 即使带有不重读的 him。所以, 如果 (38a) 是右节点提升的一个例子, 对比就是出乎意料的。由此, Borsley 排除了右节点提升分析。

针对 (36) 这类句子分析的第三个可能性提出了非连续的成分概念, 并应用了两次材料: 两个 VP *knows this record* 和 *loves this record* 与第一个 VP 的并列是非连续的。(请看 Crysmann (2001) 和 Beavers & Sag (2004) 在 HPSG 理论框架下提出的分析。) 但是, 非连续的成分通常不在最简方案框架下进行应用 (比如说 Kayne (1994: 67))。而且, Abeillé (2006) 指出, 有证据表明词汇元素直接并列的结构。这就意味着我们需要上面讨论的 CG 分析, 这就会导致刚刚讨论的限定语/补足语状态的问题。

进而, Abeillé 指出, 在中心语后置的语言中, 如韩语和日语中, NP 并列在基于合并的分析中是有困难的。(39) 是一个日语的例子。

- (39) Robin-to Kim
Robin-和 Kim
'Kim 和 Robin'

在第一步中, Robin 与 to 合并。在第二步, Kim 被合并。因为 Kim 是限定语, 我们可以认为 Kim 位于中心语之前, 因为在中心语后置的语言中, 限定语就是这样排列的。

Chomsky 试图去除标准 \bar{X} 理论的一元分支结构, 该结构用来将代词和限定词等词汇项投射到完整的短语中, 相关工作由 Muysken (1982) 完成。Muysken 应用二元特征 MIN 和 MAX 将句法对象区分为最小 (词或者类词的复杂对象) 或最大 (代表完整短语的句法对象)。这样的特征系统可以用来将代词和限定词描述为 [+MIN, +MAX]。但是, 像 give 一类的动词被归类为 [+MIN, -MAX]。它们必须为了达到 [+MAX]-层而投射。如果限定语和补足语需要是 [+MAX] 的, 那么限定词和代词满足了这一要求, 而不需要从 X^0 通过 X' 投射到 XP-层。

在 Chomsky 的系统里, MIN/MAX 的区别表现在中心语的完整性 (完整=短语), 以及作为词汇项的属性。但是, 在 Muysken 和 Chomsky 的观点中有一个小但是很重要的区别: 关于上面讨论的并列数据的认识问题。在 \bar{X} 理论的范畴系统中, 可

以将两个 X^0 s 相组合以得到一个新的、复杂的 X^0 。这一新对象基本上具有简单的 X^0 s 具有的相同的句法属性（请看 Jackendoff 1977: 51 和 Gazdar, Klein, Pullum & Sag 1985）。在 Muysken 的系统里，可以构成这样的并列规则（或连词的词汇项）来将两个 +MIN 并列为一个 +MIN 项。在 Chomsky 的系统里，无法定义一个相似的规则，因为两个词汇项的并列不再是一个词汇项了。

正如 Chomsky 最近提出的最简方案，范畴语法 (Ajdukiewicz 1935) 和 HPSG 理论 (Pollard and Sag 1987, 1994: 39–40) 并不（严格）遵循 \bar{X} 理论。所有的理论都将符号 NP 指派给代词（有关 CG 请看 Steedman & Baldridge (2006: p. 615)，有关为了量化来加入词汇类型的提升请看 Steedman (2000: Section 4.4)）。短语 “likes Mary” 和词 “sleeps” 在范畴语法 ($s \backslash np$) 中具有相同的范畴。在所有的理论中，没有必要为了将像 tree 的名词与限定词或状语相组合而从 N^0 投射到 \bar{N} 。在受限不定式中的限定词和单价动词在许多 HPSG 理论的分析中没有从 X^0 层投射到 XP 层，因为每个语言对象（一个空的 SUBCAT 或 COMPS list）的价的属性足够决定他们的组合潜力以及句法分布 (Müller 1996d, Müller 1999a)。如果最简的属性对于现象的描述是必要的，那么在 HPSG 理论中就应用了二元特征 LEX (Pollard and Sag 1987: 172, 1994: 22)。但是，这个特征对于区分限定语和补足语来说不是必要的。这一区分由将论元结构列表 (ARG-ST) 的元素匹配到限定语与补足语的特征（分别缩写为 SPR 和 COMPS）的配价列表的原则所决定。²⁸ 严格来说，动词投射中的限定语是英语类这类层次结构化语言中动词的最间接论元。由于论元结构列表是由 Keenan & Comrie (1977) 提出的间接层级所决定的，该列表中的第一个元素是动词的最间接论元，而且该论元匹配到 SPR list 上。SPR list 中的元素实现为英语等 SVO 型语言的动词左边的成分。COMPS list 中的元素实现为他们的中心语右边的成分。Ginzburg & Sag (2000: 34, 364) 提出的方法认为中心语-补足语短语将动词及其论元组合起来，这与 (36) 这类并列具有相同的问题，因为 VP 的中心语不是一个词。²⁹ 但是，中心语的这一限制可以由指称 LEX 的特征所替代，而不是词或词项的属性。

Pollard & Sag 和 Sag & Ginzburg 提出了英语的平铺结构。由于其中一个子节点被标记为词汇，它自然得到的规则就不会将中心语与它的补足语的子集相结合，然后第二次将结果与其他补足语组合起来。所以说，像 (40a) 的结构就被排除了，因为 “gave John” 不是一个词，所以不能用作规则中的中心语节点。

- (40) a. [[gave John] a book]
 给 John ART 书
 ‘给 John 一本书’
 b. [gave John a book]
 给 John ART 书
 ‘给 John 一本书’

与 (40a) 不同的是，只有像 (40b) 的分析是被允许的；也就是说，中心语与它所有的论元相组合。另一方面是假定二元分支结构 (Müller 2015c, Müller & Ørsnes 2015:

²⁸ 一些作者提出了主语、限定语和补足语的三分法。

²⁹ 正如上面提到的，对于非连续成分的多领域方法对于 (36) 的分析而言是可行的（请看 Crysmann 2001 和 Beavers & Sag 2004）。但是，正如 Abeillé (2006) 强调的，词汇项的合并需要在原则上是可行的。还需要注意的是，并列的 HPSG 方法不能代替 MP。原因是 HPSG 方法包括并列的特殊的语法规则，而 MP 则只有合并。所以说，组合型规则的额外介入并不是 MP 的一个选项。

4 转换语法-最简方案

Section 1.2.2)。在这一方法中，中心语补足语模式不能限制中心语节点的词或短语的状态。HPSG 理论中的二元分支结构对应于 MP 的外部合并。

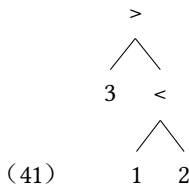
在前面两节中，我们讨论了 Chomsky 有关标记的界定问题以及词汇项的并列问题。在下面一节中，我将探讨 Stabler 对最简方案的合并的界定，他的界定是与标记相关的，而且其中一个版本没有涉及到上面讨论的问题。我将指出他的形式化方法与 HPSG 的表示非常一致。

4.6.4 最简方案、范畴语法和 HPSG 理论

在这一小节，我将最简方案、范畴语法和 HPSG 理论联系起来。对范畴语法和 HPSG 理论不太熟悉的读者可以略过这一节，或者在看完第 6 章、第 8 章和第 9 章之后再回来。

在第 4.6.2 节，我们指出 Chomsky 的论文在未确定的标记上留下了许多重要的细节尚未讨论。Stabler 的工作相对而言接近于近期的最简方案，但是在实现方面更为准确（也请看 Stabler (2011a: 397, 399, 400) 关于后 GB 方法的形式化）。Stabler (2001) 说明了 Kayne 的剩余移位理论是如何形式化和实现的。Stabler 将他对最简方案理论的形式化的方式叫做最简语法 (Minimalist Grammars, MG)。关于最简方案和其变体 (Michaelis 2001) 的能力不足有许多有趣的结论。比如说，可以用 MGs 创造的语言也可以用树邻接语法（请看第 12 章）来创造。这就意味着，可以将更多的词串指派到 MGs 的结构中，但是，由 MGs 推导出来的结构并不一定必须与 TAGs 创造的结构相同。更多有关语法的生成能力的内容，请看第 17 章。

尽管 Stabler 的工作可以看作是 Chomsky 的最简方案思想的形式化，Stabler 采用的方法与 Chomsky 的方法在某些细节方面是不同的。Stabler 认为两个合并操作的结果不是集合，而是偶对。偶对的中心语由指针（‘<’或‘>’）来标记。像（第 4.6.2 节讨论的）括号表达式 $\{\alpha, \{\alpha, \beta\}\}$ 替换为 (41) 中的树的表达形式。



在 (41) 中，1 是中心语，2 是补足语，3 是限定语。指针指向包括中心语的结构的那个部分。树中的子节点是按照顺序排列的，即 3 在 1 之前，1 在 2 之前。

Stabler (2011a: 402) 对外部合并的界定如下所示：

$$(42) \quad em(t_1[=f], t_2[f]) = \begin{cases} < \\ \begin{array}{c} \diagup \quad \diagdown \\ t_1 \quad t_2 \end{array} & \text{if } t_1 \text{ has exactly 1 node} \\ > \\ \begin{array}{c} \diagup \quad \diagdown \\ t_2 \quad t_1 \end{array} & \text{otherwise} \end{cases}$$

$=f$ 是一个选择特征, f 是对应的范畴。当 $t_1[=f]$ 和 $t_2[f]$ 组合时, 结果是一棵树, 其中 t_1 的选择特征和 t_2 的个别范畴特征被删除了。(42) 中的上层树表示一个 (词汇) 中心语与其补足语相组合。 t_1 位于 t_2 之前。只能有一个节点的条件与 Chomsky 的设想是一致的, 即第一个合并是与补足语的合并, 其他合并操作是与限定语的合并 (Chomsky 2008: 146)。

Stabler 将内部合并界定为:³⁰

$$(43) \quad im(t_1[+f]) = \begin{array}{c} > \\ \diagup \quad \diagdown \\ t_2^> \quad t_1\{t_2[-f]^> \mapsto \epsilon\} \end{array}$$

t_1 是一棵带有 t_2 子树的树, 它的特征 f 的值是 ‘-’。这一子树被删除了 ($t_2[-f]^> \mapsto \epsilon$), 然后没有特征 $-f$ ($t_2^>$) 的被删除子树的复制位于限定语的位置。在限定语位置上的元素必须是最大投射。这一要求由提升的 ‘>’ 来表示。

Stabler 为 (44) 中的句子提供了一个推导的例子。

- (44) who Marie praises
 谁 Marie 表扬
 ‘Marie 表扬的人’

praises 是一个带有两个 $=D$ 特征的二元动词。这就意味着有两个限定短语的选择。who 和 Marie 是两个 Ds, 而且它们填充到动词的宾语和主语位置上。得到的动词投射 “Marie praises who” 嵌套在空的补足语下面, 它被限定为 $+WH$, 这样就为 who 的移位提供了位置, 它被放在内部合并的操作下 CP 的限定语位置上。who 的 $-WH$ 特征被删除了, 应用内部合并得到的结果是 “Marie praises who”。

这一分析的问题由 Stabler 他自己在未发表的著作 Veenstra (1998: 124) 中提出来: 它对单价动词的处理是不正确的。如果动词与 NP 组合, 按照 (42) 中有关外部合并的定义, 就会将 NP 看作是补足语³¹, 并将其排到中心语的右边。与 (45a) 中的句子的分析相反, 我们会得到 (45b) 中的字符串的分析。

- (45) a. Max sleeps.
 Max 睡觉
 ‘Max 睡觉。’

³⁰ 除了 (43) 所示的情况, Stabler 的定义包括最短移位限制 (Shortest Move Constraint, SMC) 的变体, 这与目前所讨论的问题是无关的, 所以我们在这里就不说明了。

³¹ 再与第 4.6.3 节中 Chomsky 关于限定语和补足语的定义相比较。

4 转换语法-最简方案

- b. *Sleeps Max.
睡觉 Max

为了解决这一问题, Stabler 提出, 单价动词与非显性宾语相组合(请看 Veenstra (1998: 61, 124), 他引用了 Stabler 尚未发表的工作, 而且也采用了这一解决方法)。有了这一空宾语, 得到的结构和 Max 是限定语, 并由此被排列在 (46) 中动词的左边。

- (46) Max sleeps _.
Max 睡觉
'Max 睡觉。'

当然, 任何一种这类分析都是规定的, 而且完全是特设的。而且, 它恰巧可以佐证 Culicover & Jackendoff (2005: Section 2.1.2) 讨论的转换生成语法在方法论上的缺陷: 过度地追求统一。

另一种方法是, 假定有一个空的动词中心语, 它带有补足语 sleeps 和主语 Max。这种分析经常被用来说明最简方案中 Larson 式的动词壳 (Larson 1988) 理论下的双及物动词。Larson 式的分析一般认为有一个叫做小 v 的空的动词中心语, 该中心语具有致使的含义。正如我们在第 4.1.4 节讨论的, Adger (2003) 采用了基于小 v 的分析来讨论不及物动词。不考虑 TP 投射, 图 4.22 显示了他的分析。Adger 认为带

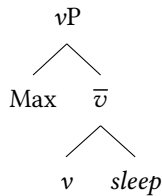


图 4.22: “Max sleeps” (Max 睡觉) 的基于小 v 的分析

有非作格动词的句子的分析包括选择了施事的小 v , 而带有非宾格动词的分析包括小 v , 但是它并不选择 N 中心语。对于非宾格动词而言, 他认为动词选择了主题。他证明了小 v 不必具有致使的含义, 但是要能引入施事。不过需要注意的是手头这个例子中, sleep 的主语既没有引发事件, 也不必故意做某事。所以, 它实际上是遭受者, 而不是施事。这意味着关于空 v 中心语的假设纯粹是基于理论的假设, 而不是针对不及物动词的语义所驱动的。如果假定了双及物构式中小 v 的致使贡献, 这就意味着我们需要两个小 v s, 一个具有致使含义, 另一个不具有致使含义。小 v 除了缺乏理论外的驱动力, 这类分析仍有一些实证方面的问题(比如说并列)。更多细节可以参考 citew[Sections 6.1 and 7]MWArgSt。

除了 (42) 和 (43) 中界定的两个操作, MG 中没有其他操作了。³² 除了 单价动词的问题, 它还会导致第 4.6.3 节讨论的问题: 对于例 (36) – 这里重复显示为 (47) – 的动词, 还没有对其直接组合的分析。

³² 扩展内容请看 Frey & Gärtner (2002: Section 3.2)

- (47) He [knows and loves] this record.
 他知道 和 喜爱 这 专辑
 ‘他知道并喜爱这张专辑。’

原因是“know、and 和 loves”的组合包括三个节点，而且“knows and loves”与“this record”的合并会让“this record”变成结构的限定语。由此“this record”就会排在“knows and loves”的前面，这是与事实相反的。由于可以由 MGs 生成的语言包括那些可以由 TAGs 和组合性范畴语法 (Michaelis 2001) 生成的语言，范畴语法分析的存在暗示了并列例子可以由 MGs 推导出来。但是对于语言学家而言，事实是无法生成不太重要的所有字符串（语法的能力较弱）。由语法所允准的实际结构大多是重要的（强势能力）。

4.6.4.1 导向性的最简语法与范畴语法

除了重新界定 X^0 范畴，并列问题还可以通过改变合并的定义来解决，改变后的定义允许中心语确定与之组合的论元的方向：Stabler (2011b: p. 635) 指出，决定论元的位置与选择特征一起重新定义了外部合并。

$$(48) \quad em(t_1[\alpha], t_2[x]) = \begin{cases} < \\ \begin{array}{c} \diagup \quad \diagdown \\ t_1 \quad t_2 \end{array} & \text{if } \alpha \text{ is } =x \\ > \\ \begin{array}{c} \diagup \quad \diagdown \\ t_2 \quad t_1 \end{array} & \text{if } \alpha \text{ is } x= \end{cases}$$

等号的位置决定了论元需要实现在中心语的哪一边。这对应于范畴语法的前向与后向应用（请看第 8.1.1 节）。Stabler 将这一形式叫做导向性的 MG (DMG)。这一 MG 的变体避免了单价动词的问题，而且如果我们假设连词是带有变量范畴的中心语，这些变量范畴选择与它左边或右边具有同样范畴的元素，这样并列数据也没有问题了。know 和 love 都会将宾语选在右边，主语选在左边，而且这一要求可以被传递到“knows and loves”。³³Stabler 针对 DMGs 提出的有关 CG 的更多细节请看 Steedman (1991: 264)。

4.6.4.2 最简语法与中心语驱动的短语结构语法

用“>”和“<”标记结构的中心语直接对应于中心语的 HPSG 表示。由于 HPSG 是一个基于符号的理论，所有相关的语言学层面都表示为描写（音位、形态、句法、语义、信息结构）。(49) 给出了一个例子：词 grammar 的词汇项。

³³ 不过，需要注意的是，这一传递使得选择复杂范畴变成了可能，这是我在 Müller (2013c) 忽略的事实。简单特征还是复杂特征的选择问题将在第 4.6.5 节讨论。

$$(49) \quad \left[\begin{array}{c} \text{word} \\ \text{PHON} \\ \\ \text{SYNSEM|LOC} \end{array} \left[\begin{array}{c} \langle \text{'gramər} \rangle \\ \text{loc} \\ \text{CAT} \left[\begin{array}{c} \text{cat} \\ \text{HEAD noun} \\ \text{SPR} \langle \text{DET} \rangle \end{array} \right] \\ \text{CONT} \dots \left[\begin{array}{c} \text{grammar} \\ \text{INST X} \end{array} \right] \end{array} \right] \right]$$

grammar 的词性是名词。为了构成一个完整的短语，它需要一个限定词。这由值为 $\langle \text{DET} \rangle$ 的 SPR 特征来表示。语义信息列于 CONT 下。更多细节请看第 9 章。

由于我们穷尽地处理句法方面，只有使用过的特征的子集是相关的：配价信息和词类信息以及与短语的外在分布有关的某些形态句法属性表现为 SYNSEM|LOC|CAT 路径下的特征描述。这里特别有趣的特征是所谓的中心语特征。在词汇中心语和它的最大投射之间共享中心语特征。复杂的层级结构也被模拟为特征值偶对。复杂语言对象的组成成分通常表现为复杂对象的表示的部分。比如说，特征 HEAD-DAUGHTER 的值是一个特征结构，该特征结构模拟了包括短语的中心语的语言学对象。中心语特征原则 (50) 指向这个子节点，并且保证中心语与完整对象的中心语特征是相同的。

$$(50) \quad \text{headed-phrase} \Rightarrow \left[\begin{array}{c} \text{SYNSEM|LOC|CAT|HEAD } \boxed{1} \\ \text{HEAD-DTR|SYNSEM|LOC|CAT|HEAD } \boxed{1} \end{array} \right]$$

指称是由具有相同数字的盒子表示的。

Ginzburg & Sag (2000: 30) 在给出了子节点属性的值的列表中表示了语言对象的所有子节点。特征 HEAD-DAUGHTER 的值与子节点列表中的一个元素是同指的：

$$(51) \quad \begin{array}{l} \text{a.} \quad \left[\begin{array}{c} \text{HEAD-DTR } \boxed{1} \\ \text{DTRS} \quad \langle \boxed{1} \alpha, \beta \rangle \end{array} \right] \\ \text{b.} \quad \left[\begin{array}{c} \text{HEAD-DTR } \boxed{1} \\ \text{DTRS} \quad \langle \alpha, \boxed{1} \beta \rangle \end{array} \right] \end{array}$$

α 和 β 是对语言学对象的描写的简称。(51) 中的两个描写的要点在于中心语子节点与两个子节点的一个是相同的，这分别由 α 和 β 前的 $\boxed{1}$ 决定。在第一个特征描述中，第一个子节点是中心语，而在第二个描述中，第二个子节点是中心语。由于中心语特征原则，整个短语的句法属性由中心语子节点决定。也就是说，中心语子节点的句法属性对应于 Chomsky 界定的标注。这一说法直接对应于 Stabler 使用的术语：(51a) 等同于 (52a)，而且 (51b) 等同于 (52b)。

$$(52) \quad \begin{array}{cc} \begin{array}{c} < \\ \diagup \quad \diagdown \\ \alpha \quad \beta \end{array} & \begin{array}{c} > \\ \diagup \quad \diagdown \\ \alpha \quad \beta \end{array} \\ \text{a.} & \text{b.} \end{array}$$

由Pollard & Sag (1994: Chapter 9)讨论的, 基本信息的另一种结构应用了 HEAD-DAUGHTER 和 NON-HEAD-DAUGHTERS 这两个特征, 而不是 HEAD-DAUGHTER 和 DAUGHTERS。这给出了像 (53a) 的特征描述, 它直接对应于 Chomsky 的基于集合的表示, 并在第 4.6.2 节进行了讨论, 而且这里重复为 (53b)。

- (53) a. $\left[\begin{array}{ll} \text{HEAD-DTR} & \alpha \\ \text{NON-HEAD-DTRS} & \langle \beta \rangle \end{array} \right]$
 b. $\{ \alpha, \{ \alpha, \beta \} \}$

例 (53a) 中的表示并不包括 α 和 β 的线性排列顺序。组成成分的线性排列顺序受到排列规则的限制, 这些规则是受到 (直接) 统制之外的独立的限制。

在 (43) 中有关内部合并的定义对应于 HPSG 中的中心语-填充语范式 (Pollard & Sag 1994: 164)。Stabler 的派生规则删除了 $t_2[-f]^>$ 子树。HPSG 是单调的, 即在语法允准的结构中是没有什么成分被删除的。与在更大的结构中删除 t_2 相反的是, 包括空元素 (NB, 不是树) 的结构是被直接允准的。³⁴在 Stabler 的定义和 HPSG 范式中, t_2 被实现为结构中的填充语。在 Stabler 关于内部合并的界定中, 没有提及中心语子节点的范畴, 但是 Pollard & Sag (1994: 164) 将中心语子节点限制为一个定式的动词投射。Chomsky (2007: 17) 认为, 所有的操作, 除了外部合并操作, 都是短语层面的。Chomsky 认为 CP 和 v^*P 是短语。如果这一限制被合并到 (43) 的定义中, 标记 t_1 的限制也会相应地得到扩展。在 HPSG 中, 像 (54) 的句子被看作是 VPs, 而不是 CPs。所以 Pollard & Sag 要求中心语-填充语范式中的中心语子节点应该是动词性的, 以对应于 Chomsky 提出的限制。

- (54) Bagels, I like.
 百吉饼我喜欢
 ‘百吉饼, 我喜欢。’

所以说, 除了较小的表现上的差异, 我们可以得到这样的结论, 内部合并与中心语-填充语模式的形式化是非常相似的。

在 HPSG 和 Stabler 的定义之间一个重要的区别是移位。在 HPSG 中不是特征驱动的。这是一个重要的优势, 因为特征驱动的移位不能处理所谓的排他移位 (Fanselow 2003a) 的例子, 也就是说, 组成成分移位的发生是为某个位置上的其他成分让地方 (请看第 4.6.1.4 节)。

在一般的 \bar{X} 理论和 Stabler 在内部合并上的形式化一方面与 HPSG 中的另一方面的差别是, 后者并没有有关填充语子节点的完整性 (或配价饱和度) 的限制。填充语子节点是否是最大投射或者不是, 当语迹是与它的中心语相组合的时候, 这一问题受到局部强加的限制。这就使得我们有可能分析 (55) 这类没有剩余移位的句子。³⁵

- (55) Gelesen_i hat_j das Buch keiner _{-i -j}.
 读 AUX DET 书 没有人

³⁴ 请看 Bouma, Malouf & Sag (2001a) 在 HPSG 模式中有关无语迹的研究, 以及本书中的第 Müller (2015b: Chapter 7) 章和第 19 章有关空元素的讨论。

³⁵ 也可以参考 Müller & Ørnsnes (2013b) 有关丹麦语中宾语转换的分析, 这种现象与没有剩余移位的动词前置是有关的。这一分析没有剩余分析所具有的任何问题。

相反, Stabler 被迫要假定 (56b) 中这样的分析 (请看 G. Müller (1998) 有关剩余移位的分析)。第一步, “das Buch” 被移出了 VP (56a), 而且在第二步, 空置的 VP 被提前, 如 (56b) 所示。

- (56) a. Hat [das Buch]_j [keiner [VP _{-j} gelesen]].
 b. [VP _{-j} Gelesen]_i hat [das Buch]_j [keiner _{-i}].

Haider (1993: 281), De Kuthy & Meurers (2001: Section 2) 和 Fanselow (2002) 指出, 这类剩余移位分析对于德语是有问题的。Fanselow 所指的需要剩余移位分析的唯一现象的问题是多重前置 (请看 Müller (2003a) 针对更多相关数据的讨论)。Müller (2005b,c, 2015b) 针对这些多重前置提出了另一种分析方法, 他提出了前场中的空动词中心语概念, 但是并没有假定例 (56b) 中 “das Buch” 这类附加语和论元是从前场成分中提取出来的。与剩余移位分析相反的是, 范畴语法 (Geach 1970; Hinrichs & Nakazawa 1994) 中的论元组成的机制被用来确保句中论元的合理实现。Chomsky (2007: 20) 已经将论元组成作为他的 TPs 和 CPs 的一部分。所以说, 这两个剩余移位和论元组合被看作是最近的最简方案的理论。但是, HPSG 看起来不需要太多的理论演算, 并由此在简约的推理方面更为可取。

最后, 需要指出的是, 所有的转换方法在跨域提取方面都是有问题的, 就像 (57a) 和 (57) 中, 一个元素对应于几个空位。

- (57) a. Bagels, I like and Ellison hates.³⁶
 b. The man who_i [Mary loves _{-i}] and [Sally hates _{-i}] computed my tax.
 DET 男人 CONJ_i Mary 爱 _{-i} 和 Sally 恨 _{-i} 计算 我的税
 ‘Mary 喜爱但是 Sally 憎恨的这个男人计算了我的税。’

Gazdar (1981b) 针对 GPSG 解决了这一问题, 而且该解决方案也被用在了 HPSG 中。最简方案的圈子试着通过引进侧边移位来解决这些问题 (Nunes 2004), 其中组成成分可以被插入子树中。所以 (57a) 的例子中, Bagels 从 hates 的宾语位置拷贝到 like 的宾语位置, 然后两个拷贝都与前置的元素相关。Kobele 批评这些方法, 因为他们过度生成, 而且需要复杂的过滤。相反, 他提出将 GPSG 式的 SLASH 机制引入到最简方案的理论中 (Kobele 2008)。

进而, 移位矛盾 (Bresnan 2001: Chapter 2) 可以通过不在填充语和空位之间分享所有的信息来避免, 这一方法在转换框架下是不可能的, 它通常会假定填充语和空位具有同一性 – 正如在移位的复制理论下 – 认为一个推导过程包括一个宾语的多重复制, 其中只有一个是被拼写出来的。也可以参考 Borsley (2012) 跟多有关基于移位的方法的问题与困惑。

MG 和 HPSG 更深一层的区分在于中心语-填充语不是分析长距离议论的唯一模式。正如在第 133 页的脚注 10 所指出的, 右侧 (外置) 和前置都有错位的情况。尽管这些问题一定会分析为长距离依存, 他们与其他长距离依存在很多方面都有不同之处 (请看第 13.1.5 节)。在 HPSG 框架下的有关外置的分析, 请看 Keller (1995)、Bouma (1996) 和 Müller (1999a)。

除了长距离依存, 在 HPSG 中当然还有其他不在 MG 和最简方案下的现象。这些模式包括描述那些没有中心语的结构或者需要捕捉结构内部的分布属性, 这些

³⁶ Pollard & Sag (1994: 205).

不能轻易地在词汇分析中有所表示（比如说 *wh-* 和关系代词的分布情况）。请看第 21.10 节。

Chomsky (2010) 比较过助词转换的基于合并的分析与 HPSG 式的分析，并且批判说 HPSG 应用了十个模式，而不是一个（合并）。Ginzburg & Sag (2000) 区分了三类带有移位的助词的结构：带有前置副词和 *wh-* 问句 (58a,b) 的倒置句、倒置的感叹句 (58c)，以及极性疑问句 (58d)：

- (58) a. Under no circumstances *did she think they would do that.*
 PREP 没有条件 AUX 她想他们 AUX 做那
 ‘在没有任何条件下，她想他们会那样做。’
- b. Whose book *are you reading?*
 谁的 书 AUX 你读
 ‘你在读谁的书？’
- c. Am I tired!
 AUX 我累
 ‘我累了！’
- d. Did Kim leave?
 AUX Kim 离开
 ‘Kim 离开了吗？’

Fillmore (1999) 在他的构式语法框架下分析助词倒装捕捉了几个不同的应用语境，并指出在不同的语境下有语义和语用的不同。每个理论都需要说明这些。进而，我们不需要十个模式。正如范畴语法那样，我们有可能在词汇项中来决定助词或空的中心语（请看第 21 章更多有关词汇和短语分析的讨论）。除此之外，每个理论都要在某种程度上解释这十种区别。如果有人认为这与句法无关，那么这就需要在语义组成上进行模拟。这就意味着在这一点上没有理由倾向于一种理论。

4.6.5 原子特征的选择与复杂范畴的选择

Berwick & Epstein (1995) 指出最简方案与范畴语法非常相似，而且我已经在 Müller (2013c) 和前面几节中讨论了最简方案和 HPSG 的相似之处。但是，我忽略了最简方案和范畴语法、依存语法、LFG、HPSG、TAG 和构式语法中关于选择的一个重要区别：在理论的前一版本中选择的是一个单一特征，而后面的理论选择的是特征束。这一区别看起来很小，但是后果很严重。在第 148 页给出的 Stabler 关于外部合并的定义移除了选择特征 (=f)，以及被选择的元素 (f) 的相应特征。在一些文献和本书的导言部分，选择特征作为不可预测的特征被标记为 *u*。不可预测的特征需要被核对，然后从 Stabler 界定的语言对象中移除。他们被核对的事实通过选择被表示出来。所有不可预测的特征都需要在句法对象被送到表层（语义和语音）之前核对。如果不可预测的特征没有被核对，推导就失败了。Adger (2003: Section 3.6) 明确地讨论了这些假说的后果：一个选择的中心语核对被选择宾语的特征。我们不可能去检查与中心语相组合的包含在宾语中的元素的特征。只有在最高点的特征，所谓的根节点，可以由外在合并核对。在复杂对象内部的特征被核对的唯一途径是通过移位。这就意味着中心语可以不与部分饱和的语言对

4 转换语法-最简方案

象相组合，即带有一个未核对选择特征的语言对象。我会讨论Adger (2003: 95)提供的例子所指的决策设计。名词 *letters* 选择了一个 P，而且 Ps 选了一个 N。例 (59a) 的分析如图 4.23 所示。

- (59) a. letters to Peter
信 PREP Peter
‘给 Peter 的信’
b. * letters to
信 PREP

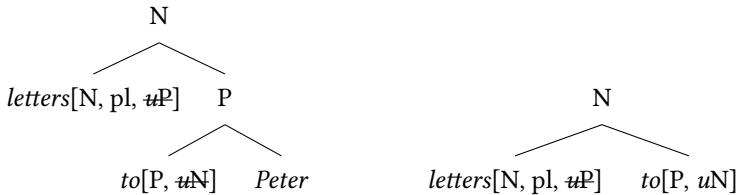


图 4.23: Adger (2003: 95)对 “letters to Peter” 的分析

例 (59b) 中的字符串被规则排除了，因为介词 *to* 的不可预测的 N 特征没有被核查。所以这就整合了这样的限制，所有的依存元素必须在核心机制中是最大的。这就使得 (60) 这类例子在最直接的方式上来分析是不可能的，即包括一个复杂的介词和缺少限定词的名词：

- (60) vom Bus
PREP.DET 公交车

在复杂描述可以用来描写依存关系的理论中，被依存的对象可以部分满足。所以比如说 HPSG 中，像 *vom* (*from.the*) 的复合介词可以选择一个 \bar{N} ，它是一个缺少限定语的名词性投射：

- (61) N[SPR < DET]]

例 (61) 中的描述是一个内部饱和了的特征值偶对的缩写（请看第 9.6.1 节）。这里的例子只针对于差异的解释，因为有很多方法用来解释单一特征合并系统的情况。例如，我们可以假设一个 DP 分析，并且让复杂介词选择一个完整的 NP（没有不可预测特征的范畴 N）。另外，我们可以假设确实有一个完整的 PP，带有通常所假定的所有结构，而且介词和限定词的组合在发音阶段形成。第一种方法省略了假定 NP 分析的选项，正如 Bruening (2009) 在最简方案框架下所提出的。

除了符合介词这个例子，还有其他想要将未饱和的语言对象相组合的例子。我已经讨论了上面的并列例子。另一个例子是像德语、荷兰语和日语这样的语言中的动词复杂形式。当然，还有一些看法并不认同动词复杂式 (G. Müller 1998; Wurmbrand 2003a)，但是他们也不是没有问题的。有些问题在前一节中也有所讨论。

简单总结这一小节的内容，需要说明的是构建进合并概念的特征核对机制比范畴语法、词汇功能语法、HPSG、构式语法和 TAG 中运用的选择来说更为严格。依我看，它有些太严格了。

4.6.6 小结

总之，我们可以说最简方案约束的计算机制与特征驱动的移位理论是有问题的（如转移派生限制与标记），而且空功能范畴的假设有时是特异的。如果我们不希望假定这些范畴在所有的语言中都存在的话，那么提出两种机制（合并与移位）并不能表示出语法的简化，因为每个必须约定的功能范畴都构成了整个系统的复杂度的一部分。

标记机制在细节上也没有完全实现，它并不能描述它所生成可以描述的现象，由此它应该被替换为范畴语法和 HPSG 中基于中心语/函项的标记。

4.7 总结

这一节与第 3.6 节类似。我首先评论语言习得，然后是形式化。

4.7.1 关于语言习得的解释

Chomsky (2008: 135) 将 MP 中的理论归纳为原则 & 参数分析，并将 MP 参数确定在词汇中。也可以参考 Hornstein (2013: 396)。UG 被界定为可能包括非具体语言的成分，他们是由基因决定的 (Chomsky 2007: 7)。UG 包括非绑定的合并，而且由语法生成的表达式的条件必须满足由语音和概念-意愿接口赋予的限制。另外，特征的具体集合被看作是 UG 的一部分 (Chomsky 2007: 6-7)。这些特征的具体本质没有被详细地解释，而且结果是，UG 的力量有时是模糊的。但是，在不同的语言学阵营里有达成一致的地方，就像 Chomsky 并不认为我们在第 4.6.1 节介绍的功能投射的部分也构成 UG 的一部分（但是，像 Cinque & Rizzi (2010) 这类作者认为功能投射的层级是 UG 的一部分）。由于参数的存在，第 3.6.1 节用来反对 GB 理论关于语言习得的论断与最简方案中语言习得的论断具有相关性。请看第 16 章有关语言习得和原则 & 参数模型，以及基于输入的方法的深入讨论。

Chomsky 在最简方案中的主要目标是简化语言的形式属性的理论假说，以及用来使得他们和相关部分是我们的基因天赋的一部分的计算机制。但是，如果我们总结一下本章的主要内容，我们很难相信最简方案可以达到这一目标。为了得到一个不及物动词的简单句，我们需要几个空的中心语和移位。特征可以是强的或者弱的，一致操作跨越了几个短语界限的句法树。而且为了得到正确的结果，必须要确保一致只能看到最近的元素 (13) - (14)。这与范畴语法相比是一个巨大的工程，范畴语法只组合临近的成分。范畴语法可以从输入获得（请看第 13.8.3 节），而它很难想象当输入很强时，特征驱动移位，而在输入很弱时就没有驱动这样的事实应该单独从数据获得。

4 转换语法-最简方案

4.7.2 形式化

第 3.6.2 节评论了转换语法直到上世纪九十年代都缺少形式化的研究内容。在最简方案中对形式化的态度没有变化，所以只有很少的最简方案理论的形式化与应用实现。

Stabler (2001) 说明了如何对 Kayne 的剩余移位理论进行形式化与实现。在 Stabler 的实现中³⁷，并没有转移派生限制，也没有计数³⁸，他不提倡一致等关系（请看 see Fong 2014: 132）。下述也是 Stabler 对于最简方案和 GB 理论的应用：没有大量的语法。Stabler 的语法是小型的，作为概念和句法的证据。没有形态变化³⁹，没有对多重一致关系的处理 (Stabler 2011b: Section 27.4.3)，尤其是没有语义。PF 和 LF 的过程也没有被模拟。⁴⁰ 由 Sandiway Fong 开发的语法和计算系统与理论的大小和相似度基本一致 (Fong & Ginsburg 2012; Fong 2014)：语法片段较小，句法编码，如直接在短语结构中标注 (Fong & Ginsburg 2012: Section 4)，而且远远位于 X 理论的后面。进而，他们不包括形态变化。没有实现拼写输出，所以最终无法剖析和生成任何话语。⁴¹ 这里的标准由基于约束的理论中语法的应用所决定；比如说，在上世纪九十年代开发的德语、英语和日语的 HPSG 语法作为 *Verbmobil* (Wahlster 2000a)

³⁷ 他的系统可以在以下网址找到：<http://www.linguistics.ucla.edu/people/stabler/coding.html>。2016 年 3 月 5 日。

³⁸ 在 Veenstra (1998: Chapter 9) 中有一个计数的词汇项。这个词汇项包括一个计数的集合，它包括功能中心语，可以用在某类句子中。比如说，Veenstra 认为计数可以用在带有二价动词的句子和主语位于首位的句子、带有单价动词的嵌套句、带有单价动词的 *wh*-疑问句，以及带有单价动词的极性疑问句。这一计数集合的一个元素对应于一个特别的句法配置以及在构式语法思想下的短语构式。Veenstra 的分析并不是在最简方案下找到的计数的概念的形式化。通常，它被认为是包含所有句子生成所需词汇项的计数。如 (i) 所示，复杂句可以包括带有不同句子类型的句子的组合：

- (i) Der Mann, der behauptet hat, dass Maria gelacht hat, steht neben der Palme, die im
DET 男人 CONJ 声称 AUX CONJ Maria 大笑 AUX 站 PREP DET 棕榈树 CONJ PREP
letzten Jahr gepflanzt wurde.
上次 年 种植 AUX
‘声称 Maria 大笑的男人站在去年种下的棕榈树旁。’

在 (i) 中，有两个带有不同价的动词的关系小句，一个带有单价动词的嵌套句和主句。在传统上对计数的理解中，Veenstra 需要假定有无限数量的词汇包含所有可能的句子类型的组合。

³⁹ 测试句具有 (i) 中的形式。

- (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

⁴⁰ 请看 Sauerland & Elbourne (2002) 有关 PF 和 LF-移位以及复制部分删除的建议（第 285 页）。对于这一点的应用并不是微不足道的。

⁴¹ Berwick, Pietroski, Yankama & Chomsky (2011: 1221) 关于 Fong 的研究的评论是错误的：“但是因为我们有时转而考虑计算机处理的问题，正如可以‘核对’中心语或标签的特征的能力一样，这就指出了一个合理的疑问，我们的框架是否是可计算的。所以我们值得指出移位的拷贝理论，与上述描述的‘搜索与标注’过程的域内导向，可以作为有效的剖析器来计算应用；更多细节内容请看 Fong (2011)。”如果一个软件不能剖析一个简单句，那么我们就不能说它是有效的，因为我们不知道程序所缺少的部分会不会使得它失效。进而，我们将这个软件与其他程序进行比较。正如上面已经讨论的，Fong 没有用 Chomsky 所谓的标签，但是相反他应用了第 2 章所描述的短语结构语法。

的一部分,用来对口语和大范围的 LFG 或 CCG 系统的分析。这些语法可以分析口语 (Verbmobil 涉及到日程安排和旅行计划领域) 或书面语中 83% 的话语。语言知识被用来生成和分析语言结构。在一个方向上,我们会得到词串的语义表示,而另一个方向上,可以创造出给定语义表征的词串。形态分析在分析有精细的形态标记的语言中自然发生的数据上是不可缺少的。在本书的其他部分,我们会在每章的开头讨论其他理论开发的语法和计算系统。

在 GB/MP 内部缺少大规模的语法片段的原因有可能跟最简方案团体的基本假设的变化相对较快有关系:

在最简方案中,驱动的中心语通常被叫做探针 (probe), 移动的元素叫做目标 (goal), 而且对于激活句法效应的特征之间的关系有许多不同的看法。Chomsky (1995b: p. 229) 首先提出, 当要求被满足, 特征表示核对与删除的特征。第一个设想几乎立即被修改, 这样只有特征的一部分子集, 即“形式化”、“不可预测”的特征在成功的生成过程中被核对操作所删除 (Collins, 1997、Chomsky 1995b: §4.5)。另一个想法是一些特征, 特别是某些功能范畴的特征, 也许受限是未被估值的, 通过与其他元素进入合适的句法配置而被估值 (Chomsky 2008、Hiraiwa, 2005)。而且近期的工作采取了这样的观点, 特征从未被删除 (Chomsky 2007: p. 11)。这些问题都没有得到解决。 (Stabler 2011a: 397)

为了开发全部的语法片段, 一般需要至少三年时间, 我们可以计算一下 Barriers (1986) 发表的文章和 Stabler 的应用 (1992) 发表的时间间隔。特别是大量的语法需要几个研究者在几年甚至是几十年间开展国际合作。如果基本假设在短期内重复变化, 这一过程就会被打断。

延伸阅读

这一章主要参考了 Adger (2003) 的内容。关于最简方案的其他教科书有 Radford (1997)、Grewendorf (2002) 和 Hornstein, Nunes & Grohmann (2005)。

Kuhn (2007) 对比了基于约束的 LFG 和 HPSG 的现代生成分析。Borsley (2012) 对比了 HPSG 中的长距离依存分析和 GB/最简方案中的基于移位的分析。Borsley 讨论了基于移位的方法的四种数据: 没有填充语的抽取、带有多个空位的抽取、填充语和空位不匹配的抽取, 以及没有空位的抽取。

有关标注、 \bar{X} 理论的弃用, 以及第 4.6.2–4.6.4 节 Stabler 的最简语法和 HPSG 的比较内容可以在 Müller (2013c) 找到。

Błaszczak & Gärtner (2005) 所著的《韵律短语、非连续及否定的辖域》(Intonational Phrasing, Discontinuity, and the Scope of Negation) 推荐给想深入了解的读者。作者将否定量词的分析与最简方案 (在 Kayne 的影响下) 和范畴语法 (在 Steedman 的影响下) 中的宽域进行了比较。

Sternefeld (2006) 对开发了德语转换语法分析的句法 (839 页) 进行了详细的介绍, 这些语法 (以转换为模) 几乎与 HPSG 理论下的语法完全一致 (按照层级体系在配价列表中的论元的特征描述)。Sternefeld 的结构是最小的, 因为他没有假定任何功能投射, 如果他们不能在所研究的语言中被激发的话。Sternefeld 对于其他研究中理所当然的某些方面特别严格。Sternefeld 把他的书看作是教科书, 通过这

本书，学生可以学到在提出理论时如何做到论述的一致性。基于此，这本书不仅仅推荐给学生和博士生。

Sternefeld & Richter (2012)讨论了理论语言学中尤其注重描述这一章和前面章节中的内容。我非常能够理解作者们在面对分析的模糊性、论述的方式、研究的实证基础、修辞的套路、免疫尝试以及科学标准的一般方面时感到的挫折：现成的例子就是Chomsky (2013)的文章《投射的问题》(*Problems of Projection*)。⁴²

但是，我并不认同这篇文章的普遍的、悲观的观点。依我之见，病人的情况是很危险，但是他还没有死。Sternefeld 和 Richter 所著的一篇评论指出，语言学领域的变化太快了，以后拥有 MIT 毕业的头衔已经不能保证能找到工作了（脚注 16）。可见，一些科学家在某些实证方面的问题上在重新寻找方向，以合适的方式来处理数据并在不同理论阵营中提高交流水平是走出这场危机的一个途径。

自从上世纪九十年代以来，可以看见更多的注意力放在了基于经验的研究上（特别是在德国），比如说，在语言合作研究中心（linguistic Collaborative Research Centers, SFBs）或年度语言事实（Linguistic Evidence）大会上进行的工作就是如此。正如上面引用的评论者所说的，在未来，只关注 Chomsky 决定的句法范畴已经是不够的，如“He left”这个句子（请看第 4.6.2 节）。语言学博士论文需要有经验的事实，这些事实能够证明作者真正理解语言。进而，博士论文以及其他出版作品应该能够说明作者不仅仅了解一种框架，而是了解较大范围的相关描写性与理论性文献。

正如我在第 4.6.4 节和Müller (2013c)中所讨论的，我也将在下面的章节和特别讨论的内容中来说明，在不同分析之间有很多相似性，而且他们在某些方面是一致的。走出当前的危机的方式就在于基于实证的原则、多理论的背景教育以及对于后代的培养。

简言之：所有的老师和学生都应该阅读 Sternefeld 和 Richter 所著的医疗报告。我恳求学生们不要在读完之后直接放弃阅读他们的研究，而是至少在阅读完本书的其他章节之后再做出这一决定。

⁴² 模糊性：在这篇文章中，perhaps 出现了 19 次，may 出现了 17 次，还有许多 if。一致性：结论并不一致。请看本书第 141 页的脚注 16。论证方式：术语限定语被弃用了，并声称与该术语有关的问题不存在了。由此，他们不属于这个世界。请看本书第 145 页的脚注 26。免疫性：Chomsky 这样论述空范畴原则：“明显的例外不能解释放弃一般性所到达的程度，而是探求能够解释在哪儿和为什么这样的深层原因”（第 9 页）。这一说法显然不是正确的，但是人们会想为了驳斥一个现有的分析需要多少证据。特别是在《投射的问题》(*Problems of Projection*)这篇文章中，我们必须要知道为什么这篇文章在《关于短语》(*On phrases*)发表之后的五年就发表了。反对最初设想的证据是强有力的，而且好几个点都是Chomsky (2013)自己提出来的。如果 Chomsky 想要应用他自己的标准（1957 年他的引用，请看第 5 页），以及普遍的科学方法（奥卡姆的剪刀），结果自然是回到基于中心语的标注的分析中了。

有关这篇文章的更多评论，请看第 4.6.2 节和第 4.6.3 节。

第五章 广义短语结构语法

广义短语结构语法 (Generalized Phrase Structure Grammar, GPSG) 是在 1970 年代末从转换语法的争论中发展而来。Gazdar, Klein, Pullum & Sag (1985) 所著的这本书是该理论框架下的主要文献。Hans Uszkoreit 开发了德语的大型 GPSG 语法片段 (1987)。GPSG 的分析非常准确, 由此我们可以用它们作为计算应用的基础。下表是已经应用了 GPSG 语法片段开发的一些语言:

- 德语 (Weisweber 1987; Weisweber & Preuss 1992; Naumann 1987, 1988; Volk 1988)
- 英语 (Evans 1985; Phillips & Thompson 1985; Phillips 1992; Grover, Carroll & Briscoe 1993)
- 法语 (Emirikian, Da Silva & Bouchard 1996)
- 波斯语 (Bahrani, Sameti & Manshadi 2011)

正如在第 3.1.1 节所讨论的, Chomsky (1957) 认为简单的短语结构语法已经不能很好地描述语言结构之间的关系了, 并提出我们需要转换来解释这些关系。这些观点在之后的二十年间都没有受到挑战 (除了 Harman (1963) 和 Freidin (1975) 发表的文章), 直到其他的理论出现, 如 LFG 和 GPSG。这些理论针对 Chomsky 提出的问题并针对以前只有转换分析或者尚未分析的现象的非转换的解释。论元的重新排序分析、被动以及长距离依存是在该理论框架下讨论的几个最为重要的现象。在介绍完第 5.1 节有关 GPSG 的表示形式后, 我将针对这些现象的 GPSG 分析进行详细的说明。

5.1 表示形式概说

本节分为五个部分。第 5.1.1 节说明有关特征和复杂范畴的表示的假说。第 5.1.2 节说明短语结构中子节点的线性顺序的假说。第 5.1.3 节介绍元规则。第 5.1.4 节分析语义, 以及第 5.1.5 节分析状语。

5.1.1 复杂范畴、中心语特征原则, 以及 \bar{X} 规则

在第 2.2 节, 我们在短语结构语法中增加了特征。GPSG 更进一步, 它将词类范畴描述为特征值偶对的集合。例 (1a) 中的词类范畴可以表示为 (1b):

5 广义短语结构语法

- (1) a. NP(3,sg,nom)
b. { CAT n, BAR 2, PER 3, NUM sg, CASE nom }

显然, (1b) 对应于 (1a)。(1a) 与 (1b) 的不同之处在于词类和 \bar{X} 层(在符号 NP 内)的信息是突显的, 而 (1b) 中这些信息就像格、数和人称等来处理。

词汇项具有 SUBCAT 特征。特征值用数字来表示该词能够应用的某种短语结构规则。(2) 给出了短语规则的例子并列出了能够在这些规则中出现的一些动词。¹

- (2) V2 → H[5] (kommen 'come', schlafen 'sleep')
V2 → H[6], N2[CASE acc] (kennen 'know', suchen 'search')
V2 → H[7], N2[CASE dat] (helfen 'help', vertrauen 'trust')
V2 → H[8], N2[CASE dat], N2[CASE acc] (geben 'give', zeigen 'show')
V2 → H[9], V3[+dass] (wissen 'know', glauben 'believe')

这些规则允准了 VPs, 即动词与它的补足语相组合, 而不是主语。词类符号 (V 或 N) 后的数字暗示了 \bar{X} 投射层。Uszkoreit 认为, 动词投射的最大投射数量是三, 而不是经常认为的二。

规则右边的 H 表示中心语 (head)。中心语特征原则 (Head Feature Convention, HFC) 确保了母节点中的特定特征也在标记有 H 的节点上出现 (详细情况请看 Gazdar, Klein, Pullum & Sag 1985: Section 5.4 和 Uszkoreit 1987: 67):

Principle 1 (Head Feature Convention)

母节点与子节点必须具有相同的中心语特征, 除非它们另有用途。

在 (2) 中, 可以用于规则的动词用方括号表示。与普通的短语结构规则相同的是, 在 GPSG 中, 我们也需要动词有相应的词汇项。(3) 给出了两个例子:

- (3) V[5, VFORM inf] → einzuschlafen
V[6, VFORM inf] → aufzuessen

第一条规则说明 *einzuschlafen* (睡觉) 的 SUBCAT 值为 5, 而且第二条表示 *aufzuessen* (吃完) 的 SUBCAT 值为 6。接着, *einzuschlafen* 只能用在第一条规则 (2) 中, 而 *aufzuessen* 只能用在第二条规则中。进而, (3) 包括了动词的形式信息 (*inf* 表示带 *zu* 的不定式)。

如果我们用 (2) 中的第二条规则和 (3) 中的第二条规则分析 (4) 中的句子, 我们就会得到图 5.1 在下一页中的结构。

- (4) Karl hat versucht, [den Kuchen aufzuessen].
Karl AUX 尝试 DET 蛋糕 吃完
'Karl 试着把蛋糕吃完.'

(2) 中的规则没有涉及子节点的顺序, 这就是为什么动词 (H[6]) 也可以位于末位。这一方面将在第 5.1.2 节详细讨论。关于 HFC, 我们要记住的重要一点是不定式动词的形式也在母节点上显示。与第 2 章讨论的简单短语结构规则不同是, 这一观点

¹ 下面讨论的例子选自 Uszkoreit (1987)。

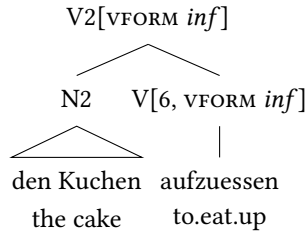


图 5.1: GPSG 中中心语特征的投射

自动承袭了 GPSG 中的中心语特征原则。在 (3) 中, *vFORM* 是给定的, 而且 HFC 确保了, 当 (2) 中的规则被应用后, 相应的信息表现在母节点上。对于例 (4) 中的短语, 我们得到了词类 $V2[vFORM\ inf]$, 而且这保证了这一短语只在它应该出现的地方出现:

- (5) a. [Den Kuchen aufzuessen] hat er nicht gewagt.
 DET 蛋糕 吃完 AUX 他不 敢
 ‘他不敢把蛋糕吃完。’
- b. * [Den Kuchen aufzuessen] darf er nicht.
 DET 蛋糕 吃完 允许他不
 想说: “不让他把蛋糕吃完。”
- c. * [Den Kuchen aufessen] hat er nicht gewagt.
 DET 蛋糕 吃完 AUX 他不 敢
 想说: “他不敢吃完蛋糕。”
- d. [Den Kuchen aufessen] darf er nicht.
 DET 蛋糕 吃完 允许他不
 ‘不让他把蛋糕吃完。’

gewagt (敢) 选择一个带 *zu* 不定式的动词和动词短语, 而不是一个单纯的不定式, 而 *darf* (允许) 带有一个单纯的不定式。

这在名词短语方面也是类似的: 有针对不带论元的名词的规则, 也有带特定论元的名词的规则。例 (6) 中给出了要么不带论元, 要么带两个 PPs 的名词的规则 (Gazdar, Klein, Pullum & Sag 1985: 127):

- (6) $N1 \rightarrow H[30]$ (*Haus* (房子), *Blume* (花))
 $N1 \rightarrow H[31], PP[mit], PP[über]$ (*Gespräch* (说话), *Streit* (争论))

\bar{N} 和限定词的组合规则如下所示:

- (7) $N2 \rightarrow Det, H1$

$N2$ 表示 NP, 即在杠第二层的名词短语投射, 而 $H1$ 表示在杠第一层的中心语子节点的投射。中心语特征规定确保了中心语子节点也是一个名词性投射, 因为在 \bar{X} 层上

区分出的中心语子节点上的所有特征与整个 NP 上的特征相同。当分析 (8) 时, (6) 中的第二条规则允准了 \bar{N} *Gesprächs mit Maria über Klaus*。事实上, *Gesprächs* (对话) 是第二个, 并在 *Gesprächs* 的词汇项中表示, 而且因为 *Gesprächs* 是中心语, 它也按照中心语特征原则在 \bar{N} 上显示。

- (8) *des Gespräch-s mit Maria über Klaus*
 DET.GEN 对话-GEN 跟 Maria 关于 Klaus
 ‘跟 Maria 有关 Klaus 的对话’

对于 \bar{N} 与限定词的组合, 我们应用 (7) 中的规则。中心语的词类范畴限定了规则左边的要素的词性, 这就是为什么 (7) 中的规则对应于我们在第 68 页 (65c) 中遇到的经典的 \bar{X} 规则。由于 *Gesprächs mit Maria über Klaus* 是中心语子节点, 有关 \bar{N} 的属格信息也在 NP 节点上显示。

5.1.2 局部重新排序

我们要讨论的第一个现象是论元的重新排序。正如在第 3.5 节已经讨论过的, 中场的论元可以按照几乎任意顺序排列。(9) 给出了一些例子:

- (9) a. [*weil*] *der Mann der Frau das Buch gibt*
 因为 the.NOM 男人 the.DAT 女人 the.ACC 书 给
 ‘因为这个男人把这本书给这个女人’
 b. [*weil*] *der Mann das Buch der Frau gibt*
 因为 DET.NOM 男人 DET.ACC 书 DET.DAT 女人 给
 c. [*weil*] *das Buch der Mann der Frau gibt*
 因为 DET.ACC 书 DET.NOM 男人 DET.DAT 女人 给
 d. [*weil*] *das Buch der Frau der Mann gibt*
 因为 DET.ACC 书 DET.DAT 女人 DET.NOM 男人 给
 e. [*weil*] *der Frau der Mann das Buch gibt*
 因为 DET.DAT 女人 DET.NOM 男人 DET.ACC 书 给
 f. [*weil*] *der Frau das Buch der Mann gibt*
 因为 DET.DAT 女人 DET.ACC 书 DET.NOM 男人 给

在第 2 章的短语结构语法中, 我们应用特征来确保动词与正确数量的论元共现。下面 (10) 中规则被用来分析 (9a) 中的句子:

- (10) $S \rightarrow NP[nom] NP[dat] NP[acc] V_nom_dat_acc$

如果有人想分析 (9) 中的其他语序, 那么我们就需要增加五条规则, 即一共有六个:

- (11) $S \rightarrow NP[nom] NP[dat] NP[acc] V_nom_dat_acc$
 $S \rightarrow NP[nom] NP[acc] NP[dat] V_nom_dat_acc$
 $S \rightarrow NP[acc] NP[nom] NP[dat] V_nom_dat_acc$
 $S \rightarrow NP[acc] NP[dat] NP[nom] V_nom_dat_acc$
 $S \rightarrow NP[dat] NP[nom] NP[acc] V_nom_dat_acc$
 $S \rightarrow NP[dat] NP[acc] NP[nom] V_nom_dat_acc$

另外，针对动词位于首位的顺序，我们有必要假定另外六条规则：

- (12) $S \rightarrow V_nom_dat_acc NP[nom] NP[dat] NP[acc]$
 $S \rightarrow V_nom_dat_acc NP[nom] NP[acc] NP[dat]$
 $S \rightarrow V_nom_dat_acc NP[acc] NP[nom] NP[dat]$
 $S \rightarrow V_nom_dat_acc NP[acc] NP[dat] NP[nom]$
 $S \rightarrow V_nom_dat_acc NP[dat] NP[nom] NP[acc]$
 $S \rightarrow V_nom_dat_acc NP[dat] NP[acc] NP[nom]$

进而，我们也将需要及物动词和不及物动词的所有可能配价的平行规则。显然，这些规则的普遍性和他们的一般性并没有被捕捉到。关键是我们有同样数量的论元，他们可以按照任意顺序来实现，而且动词可以被放在首位或末位。作为语言学家，我们认为有必要捕捉德语语言中的这一属性，并在短语结构规则之外表示出来。在转换语法中，语序之间的关系通过移位来表示：深层结构对应于带有一定顺序论元的动词末位语序，而表层语序通过 Move- α 生成。由于 GPSG 是一个非转换的理论，这种解释就是不可能的了。相反，GPSG 在直接统制 (immediate dominance) 上面添加限制，这与线性优先性 (linear precedence, LP) 是不同的：像 (13) 的规则要被理解为统制规则，这些规则没有说明子节点的顺序 (Pullum 1982)。

- (13) $S \rightarrow V, NP[nom], NP[acc], NP[dat]$

(13) 中的规则简单证明了 S 统制所有其他节点。由于弃用了规则右边的语序限制，我们只能应用一条规则，而不是十二条。

无论如何，没有规则右边的任何一种限制的话，就有太多自由了。比如说，下面的语序也是被允许的：

- (14) *Der Frau der Mann gibt ein Buch.
 DET 女人.DAT DET.NOM 男人 给 DET.ACC 书

这样的语序被所谓的线性优先规则 (Linear Precedence Rules) 或 LP-规则排除了。LP-限制是在局部树上的限制，即深度为一的树。例如，有可能应用线性规则来证明图 5.2 中 “V、NP[nom]、NP[acc] 和 NP[dat]” 的语序。

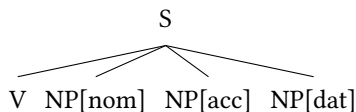


图 5.2: 局部树的例子

5 广义短语结构语法

下面的线性规则用来排除 (14) 中的那些语序: The following linearization rules serve to exclude orders such as those in (14):

- (15) $V[+MC] < X$
 $X < V[-MC]$

MC表示主句 (main clause)。LP-规则确保在主句中 (+MC)，动词位于所有其他成分前面，并在从句 (-MC) 中位于他们后面。有一个条件限制说所有带有 MC-值 ‘+’ 的动词也必须是 (+FIN)。这会排除位于首位的不定式形式。

这些 LP 规则不允许在局部树中被占有前场或后场的语序。这是想要的情况。我们将看到前置是如何在第 5.4 节进行说明的。

5.1.3 元规则

我们在之前就用过线性规则来处理带主语的句子，但是我们的规则具有例 (16) 的形式，即他们不包括主语：

- (16) $V2 \rightarrow H[7], N2[CASE\ dat]$
 $V2 \rightarrow H[8], N2[CASE\ dat], N2[CASE\ acc]$

这些规则可以用来分析 (17) 中出现的动词短语 *dem Mann das Buch zu geben* (给这个人这本书) 和 *das Buch dem Mann zu geben* (把这本书给这个人)，但是我们不能分析 (9) 这类句子，因为主语不在 (16) 的规则右边出现。

- (17) a. Er verspricht, [dem Mann das Buch zu geben].
他承诺 the.DAT 人 the.ACC 书 to 给
‘他承诺给这个人这本书。’
b. Er verspricht, [das Buch dem Mann zu geben].
他承诺 the.ACC 书 the.DAT 人 to 给
‘他承诺把这本书给这个人。’

带有 (18) 形式的规则不能用来分析德语的 GPSG 语法，因为它不能推导出 (9) 中的所有语序，因为主语可以出现在 (9c) 所示的 VP 的元素之中。

- (18) $S \rightarrow N2\ V2$

根据 (18) 中的规则，有可能分析如图 5.3 在下一页中的 (9a)，而且也可以分析在 VP 内部具有不同 NP 顺序的 (9b)。但是，(9) 中剩下的例子不能由 (18) 中的规则来分析。这与在同一局部树中的唯一元素有关系，也就是说，在规则右边出现的元素可以被重新排序。当我们将 VP 的组成部分重新排序的时候，我们可以推导 (9b)，不可能把主语放在宾语之间的低位上。相反，我们可以运用元规则来分析句子，其中主语出现在动词的其他论元之间。这一规则将短语结构规则联系到其他短语结构规则上。元规则可以被理解作为一种指导原则，该原则为带有一定形式的每条规则创造出其他规则，而且这些新造的规则可以反过来允准局部树。

对于眼前这个例子，我们可以构成一个元规则，如下所示：如果在语法中有一条带有 “V2 包括某些成分” 形式的规则，那么也有另一条规则 “V3 包括 V2 所包括的 + 主格中的 NP”。在形式化术语中，这种情况如下所示：

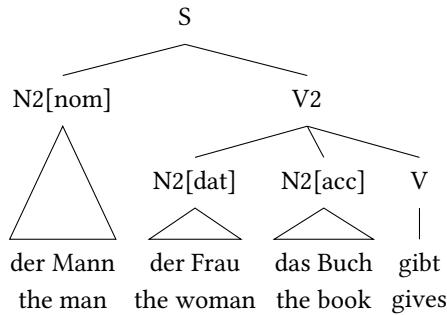


图 5.3: 德语的 VP 分析 (在 GPSG 框架下并不合适)

- (19) $V2 \rightarrow W \mapsto$
 $V3 \rightarrow W, N2[\text{CASE nom}]$

W 是一个变量, 它表示任意数量的范畴 ($W = \textit{whatever}$)。元规则创造了从 (16) 中的规则来的 (20) 中的规则:

- (20) $V3 \rightarrow H[7], N2[\text{CASE dat}], N2[\text{CASE nom}]$
 $V3 \rightarrow H[8], N2[\text{CASE dat}], N2[\text{CASE acc}], N2[\text{CASE nom}]$

现在, 主语和其他论元都出现在规则的右边, 而且只要没有 LP 规则被违反了, 就可以自由排序。

5.1.4 语义

Gazdar, Klein, Pullum & Sag (1985: Chapter 9–10) 所谓的语义学源于 Richard Montague (1974)。与规定每条规则的组合可能性的语义理论不同的是 (请看第 2.3 节), GPSG 应用更为普遍性的规则。这有可能是基于这样的事实, 每个需要组合的表达式都有一个语义类型。我们惯于区分实体 (e) 与真值 (t)。实体指称真实世界 (或可能世界) 中的对象, 而整个句子可以为真, 也可以为假, 也就是说, 他们有一个真值。我们有可能从类型 e 和 t 中创造出更为复杂的类型。通常来说, 下面的情况是可能的: 如果 a 和 b 是类型, 那么 $\langle a, b \rangle$ 也是一个类型。复杂类型的例子有 $\langle e, t \rangle$ 和 $\langle e, \langle e, t \rangle \rangle$ 。我们可以针对这类类型表达式界定如下的组合性规则:

- (21) 如果 α 属于类型 $\langle b, a \rangle$ 和 β of type b , 那么 $\alpha(\beta)$ 属于类型 a 。

这类组合也叫做功能性应用 (functional application)。根据 (21) 中的规则, 有可能类型 $\langle e, \langle e, t \rangle \rangle$ 对应于这样一个表达式, 它需要与类型 e 的两个表达式相组合, 以得到 t 的表达式。第一步与 e 的组合会产出 $\langle e, t \rangle$, 而且第二步与深层 e 的组合会得到 t 。这与我们在第 57 页看到的 λ -表达式相似: $\lambda y \lambda x \textit{like}'(x, y)$ 需要将 y 和 x 相组合。这一例子的结果是 $\textit{mögen}'(\textit{max}', \textit{lotte}')$, 即在相关世界中要么为真, 要么为假的表达式。

在Gazdar et al. (1985)中, 另一个类型是指其中表达式为真或为假的世界。为了表述的简单, 我在这里就不展开了。我们在(22)中给出了句子、NPs、N's、限定词和VPs的类型:

- (22) a. $TYP(S) = t$
 b. $TYP(NP) = \langle \langle e, t \rangle, t \rangle$
 c. $TYP(N') = \langle e, t \rangle$
 d. $TYP(Det) = \langle TYP(N'), TYP(NP) \rangle$
 e. $TYP(VP) = \langle e, t \rangle$

句子属于类型 t , 因为它可能为真, 也可能为假。一个VP需要类型 e 的表达式来产出类型 t 的句子。NP的类型也许乍看起来很奇怪, 但是, 如果我们考虑带有数量词的NPs的意义, 我们就有能理解它。对于诸如(23a)的句子, 我们一般会假定如(23b)的表达式:

- (23) a. All children laugh.
 所有孩子 笑
 ‘所有的孩子都在笑。’
 b. $\forall x \text{ child}'(x) \rightarrow \text{laugh}'(x)$

符号 \forall 表示全称量词。该公式可以这样来解读。对于每一个对象而言, 它具有孩子这样的属性, 而且它也在笑。如果我们考虑NP的贡献, 那么我们会看到全称量词对孩子的限制以及从NP而来的逻辑蕴涵:

- (24) $\forall x \text{ child}'(x) \rightarrow P(x)$

这就意味着NP是需要与表达式相组合的, 该表达式只有一个公开的槽, 对应于(24)中的 x 。这一过程在(22b)中形成: 一个NP对应于需要某种以构成真或假的表达式的类型 $\langle e, t \rangle$ 的语义表达式(即, 属于类型 t)。

N' 表示类型 $\lambda x \text{ child}(x)$ 的名词性表达式。这意味着, 如果有一个具体的个体可以插入到 x 的位置上, 那么我们将得到一个要么为真要么为假的表达式。对于一个给定的条件, 要么 John 具有称为孩子的属性, 要么没有。 N' 具有与VP相同的类型。

(22d)中的 $TYP(N')$ 和 $TYP(NP)$ 表示(22c)和(22b)中的类型, 即, 限定词在语义上需要与 N' 的意义相组合, 以给出NP的含义。

Gazdar, Klein, Pullum & Sag (1985: 209)指出了遵守规则到规则假说的语法的语义具体化是冗余的(请看第2.3节), 因为, 不管有关组合的规则与规则的要求, 它在很多方面都是足够的, 简单来说函子被应用于论元。如果我们使用(22)中那些类型, 也很清楚哪个成分是函子, 哪个是论元。按照这一方式, 名词不能被应用于限定词, 而相反是可以的。(25a)中的组合得不到合乎语法的结果, 而(25b)被规则排除了。

- (25) a. $Det'(N')$
 b. $N'(Det')$

普遍的组合原则如下所示:

- (26) 针对子节点的语义贡献的组合的功能性应用会产出对应于母节点类型的合乎语法的表达。

GPSG 著作的作者们认为这一原则可以应用于大部分的 GPSG 规则，这样只有少部分情况需要用外在的规则来处理。

5.1.5 状语

对于英语中的名词性结构，Gazdar et al. (1985: 126) 认为 \bar{X} 分析以及我们在第 2.4.1 节看到的分析适用于德语的名词结构。尽管如此，如果我们假定平面分支结构的话，在动词域内对状语的分析会有问题，因为状语可以自由出现在论元之间：

- (27) a. weil der Mann der Frau das Buch *gestern* gab
 因为DET 男人 DET 女人DET 书 昨天 给
 ‘因为这个男人昨天把这本书给这个女人了’
 b. weil der Mann der Frau *gestern* das Buch gab
 因为DET 男人 DET 女人昨天 DET 书 给
 c. weil der Mann *gestern* der Frau das Buch gab
 因为DET 男人 昨天 DET 女人DET 书 给
 d. weil *gestern* der Mann der Frau das Buch gab
 因为昨天 DET 男人 DET 女人DET 书 给

对于 (27) 来说，我们需要如下的规则：

- (28) $V3 \rightarrow H[8], N2[CASE\ dat], N2[CASE\ acc], N2[CASE\ nom], AdvP$

当然，状语也可以出现在其他配价类型的动词的论元之间：

- (29) weil (oft) die Frau (oft) dem Mann (oft) hilft
 因为 经常DET 女人 经常DET 男人 经常帮助
 ‘因为这个女人经常帮助这个男人’

进而，副词可以出现在 VP 的论元之间：

- (30) Der Mann hat versucht, der Frau heimlich das Buch zu geben.
 DET 男人 AUX 尝试 DET 女人秘密地 DET 书 INF 给
 ‘这个男人试着秘密地把这本书给这个女人。’

为了分析这些句子，我们可以使用一条元规则，其中在 V2 的右边加上一个状语 (Uszkoreit 1987: 146)。

- (31) $V2 \rightarrow W \mapsto$
 $V2 \rightarrow W, AdvP$

通过在 (19) 中主语引入的元规则，(28) 中的 V3-规则由 V2-规则推导出来。由于在一个句子中有几个状语，像 (31) 的元规则必须被允许应用多次。元规则的循

5 广义短语结构语法

环应用在文献中经常因为生成能力而被规则排除（请看第 17 章）（Thompson 1982、Uszkoreit 1987: 146）。如果我们用 Kleene 星号表示，那么就可以构成状语的元规则，这样就不用重复应用了（Uszkoreit 1987: 146）：

- (32) $V2 \rightarrow W \mapsto$
 $V2 \rightarrow W, AdvP^*$

如果我们采用 (32) 中的规则，那么如何决定状语的语义贡献就不够清楚了。² 对于 (31) 中的规则，我们可以在输入规则中将 AdvP 的语义贡献与 V2 的语义贡献组合起来。如果元规则被应用多次的话，这自然也是可能的。如果元规则被应用于 (33a)，比如说，(33a) 中的 V2-节点包括第一个副词的语义贡献。

- (33) a. $V2 \rightarrow V, NP, AdvP$
b. $V2 \rightarrow V, NP, AdvP, AdvP$

(33b) 中的 V2-节点接收到在 (33a) 中应用到 V2-节点上的副词的语义表示。

Weisweber & Preuss (1992) 指出，如果不想应用元规则来计算短语结构规则的集合，而是在句子分析的过程中直接应用元规则的话，可以应用像 (31) 一样的元规则。因为句子总是有限长度的，而且元规则在新允准的规则右边引入一个额外的 AdvP，元规则只能被应用于有限的次数。

5.2 作为元规则的被动

德语被动式可以按照下面保持整个理论中立的方式来描述：³

- 主语受到抑制。
- 如果有宾格宾语，那么该宾语变为主语。

这对可以构成被动的所有动词来说都是正确的。不管动词是带一个、两个还是三个论元，都没有什么区别：

- (34) a. weil er noch gearbeitet hat
因为他.NOM 仍然 工作 AUX
'因为他仍然在工作'
b. weil noch gearbeitet wurde
因为仍然 工作 AUX
'因为那儿仍然有工作'

² 在 LFG 中，状语被安排在功能性结构的集合中（请看第 7.1.6 节）。这也适用于 Kleene 星号标记的使用。从 f-结构来看，有可能通过指向 c-结构来计算出相应域的语义指称。在 HPSG 中，Kasper (1994) 提出了一个与 GPSG 相关的方法，该方法是关于平面分支结构与任意数量的状语的。然而，在 HPSG 中，我们可以利用所谓的关系限制。这与可以在复杂结构中创造关系的小项目是类似的。应用这些关系限制，就可能计算出平面分支结构中未受到数量限制的状语的意义。

³ 这一描述不能适用于其他语言。比如说，冰岛语允许与格宾语。请看 Zaenen, Maling & Thráinsson (1985)。

- (35) a. weil er an Maria gedacht hat
 因为他.NOM PREP Maria 想 AUX
 ‘因为他想起了 Maria’
 b. weil an Maria gedacht wurde
 因为PREP Maria 想 AUX
 ‘因为 Maria 被想起了’
- (36) a. weil sie ihn geschlagen hat
 因为她.NOM 他.ACC 打 AUX
 ‘因为她打了他’
 b. weil er geschlagen wurde
 因为他.NOM 打 AUX
 ‘因为他被打了’
- (37) a. weil er ihm den Aufsatz gegeben hat
 因为他.NOM 他.DAT DET.ACC 文章 给 AUX
 ‘因为他把这篇文章给他了’
 b. weil ihm der Aufsatz gegeben wurde
 因为他.DAT DET.NOM 文章 给 AUX
 ‘因为他被给了这篇文章’

在简单的短语结构语法中，我们需要针对每对句子列出两条独立的规则，指向要研究动词的配价类型。上面讨论的被动的特征就不会在规则集合中被明确地证明出来。在 GPSG 中，有可能应用元规则来解释主动和被动规则之间的关系：对于每条主动规则，带有受到抑制的主语的对应被动原则被允许了。主动句和被动句之间的联系可以这样来描述。

转换语法/GB 的一个重要区别是我们没有在两棵树之间创造关系，而是在主动与被动规则之间创造关系。这两条规则允许了两个没有联系的结构，即 (38b) 中的结构并不是从 (38a) 的结构中推导而来的。

- (38) a. weil sie ihn geschlagen hat
 因为她.NOM 他.ACC 打 AUX
 ‘因为她打了他’
 b. weil er geschlagen wurde
 因为他.NOM 打 AUX
 ‘因为他被打了’

尽管如此，主动/被动之间的普遍性还是被捕捉到了。

接下来，我将详细讨论 Gazdar, Klein, Pullum & Sag (1985) 中给出的被动的分析。作者认为下面的元规则适用于英语（第 59 页）：⁴

⁴ 请看 Weisweber & Preuss (1992: 1114) 针对德语的平行规则，该规则是指元规则左边的宾格。

5 广义短语结构语法

- (39) $VP \rightarrow W, NP \mapsto$
 $VP[PAS] \rightarrow W, (PP[by])$

该规则说明带一个宾语的动词可以出现在不带该宾语的被动式 VP 中。进而，可以加上 *by*-PP。如果我们将这一元规则应用到 (40) 的规则中，那么就会得到 (41) 中的规则：

- (40) $VP \rightarrow H[2], NP$
 $VP \rightarrow H[3], NP, PP[to]$
- (41) $VP[PAS] \rightarrow H[2], (PP[by])$
 $VP[PAS] \rightarrow H[3], PP[to], (PP[by])$

可以用 (40) 中的规则来分析主动句中的动词短语：

- (42) a. $[_S \text{ The man } [_{VP} \text{ devoured the carcass}]]$.
b. $[_S \text{ The man } [_{VP} \text{ handed the sword to Tracy}]]$.

VP 与主语的组合由额外的规则允准 ($S \rightarrow NP, VP$)。

根据 (41) 中的规则，我们可以分析 (43) 中的相应被动句的 VPs：

- (43) a. $[_S \text{ The carcass was } [_{VP[PAS]} \text{ devoured (by the man)}]]$.
b. $[_S \text{ The sword was } [_{VP[PAS]} \text{ handed to Tracy (by the man)}]]$.

初看上去，这一分析看起来很奇怪，因为宾语在 VP 内部替换为一个 PP，该 PP 在主动句中是主语。尽管该分析针对句法上具有合乎语法的结构做出了正确的判断，但是语义分析并不清楚。尽管如此，我们也是可以使用词汇规则来允准被动分词，并按照某种方式来操控输出词汇项的语义，其中 *by*-PP 在语义上被正确地整合进来 (Gazdar et al. 1985: 219)。

但是，如果我们试着将这一分析应用到德语，我们就会遇到问题。因为非人称被动不能简单地通过抑制宾语而得到。诸如 *arbeiten*（工作）和 *denken*（想）的动词的 V2-规则被用在下面 (34a) 和 (35a) 的分析中：

- (44) $V2 \rightarrow H[5]$
 $V2 \rightarrow H[13], PP[an]$

在这些规则的右边没有 NP 能够变成 *von*-PP。如果被动在规则中被分析为名词论元的抑制，那么它就应该从非人称被动的存在而来，被动的元规则必须应用到可以允准定式小句的规则上，因为关于是否有主语的信息只出现在定式小句的规则中。⁵在这类系统中，定式句子 (V3) 的规则是基本规则，而 V2 的规则可以从中推导出来。只有能够应用到 V3 的是适用于德语的，因为英语没有在规则右边既包括主语也包括宾语的 V3 规则。⁶对于英语来说，一般认为句子包括一个主语和一

⁵ GPSG 与 GB 的不同之处在于，非定式动词投射不包括空主语的节点。这在本书讨论的所有其他理论中也是真的，除了树邻接语法。

⁶ Gazdar et al. (1985: 62) 提出了一条元规则与我们在第 167 页提出的主语引入的元规则相似。由他们的元规则允准的规则被用来分析英语中助词的位置，并只允准 AUX NP VP 的形式序列。在这类结构中，主语和宾语不在同一棵局部树内。

个 VP (请看Gazdar et al. 1985: 139)。这就意味着我们对英语和德语的被动式会得到两个非常不同的分析结果, 该结果并不能得到同样的描述性结论, 即被动是主语的抑制与宾语的提升。德语和英语的核心区别看上去是英语强制要求有一个主语,⁷这就是为什么英语没有非人称被动。这是一个独立于被动的属性, 但是, 它会影响到是否有被动结构的可能性。

GPSG 分析的问题在于配价信息在短语结构规则中编码, 而主语在动词短语的规则中是没有的。在下面的章节中, 我们将看到 LFG、范畴语法、HPSG、构式语法和依存语法的理论方法, 其中配价是与短语结构规则分开编码的, 这样就不会在非人称被动上出现原则性的问题了。

请看Jacobson (1987b: 394–396)在 GPSG 中有关被动分析的更多有问题的地方, 而且可以看到在范畴语法、GB、LFG 和 HPSG 中配价的词汇表征可以实现现象的词汇分析, 在 GPSG 中配价表征的基础假设的原则问题是不可能实现的。

5.3 动词位置

Uszkoreit (1987)将动词位于首位与位于末尾的语序分析为平面树的线性顺序变体。这一分析的细节已经在第 5.1.2 节进行了讨论。

GPSG 版本的另一种分析来自于Jacobs (1986: 110): Jacobs 的分析是在 GB 中提出动词移位分析。他认为在末位有一个空动词, 并将其应用技术手段联系到位于首位的动词。我们将在下一节讲到更多细节。

5.4 作为局部依存结果的长距离依存

GPSG 的一个创新点是它将长距离依存看作是局部依存的序列 (Gazdar 1981b)。以德语为例, 该方法可以解释为将成分提前到前场。截至目前, 我们只看到动词首位与动词末位的 GPSG 分析: (45) 中的序列只是简单的线性变体。

- (45) a. [dass] der Mann der Frau das Buch gibt
 CONJ DET 男人 DET 女人 DET 书 给
 ‘这个男人把这本书给这个女人’
 b. Gibt der Mann der Frau das Buch?
 给 DET 男人 DET 女人 DET 书
 ‘这个男人把这本书给这个女人了吗?’

我们想要的是从 (45b) 中的 V1 语序得到 (46) 中的动词位于第二位语序的例子。

- (46) a. Der Mann gibt der Frau das Buch.
 DET 男人 给 DET 女人 DET 书
 ‘这个男人给这个女人这本书。’
 b. Der Frau gibt der Mann das Buch.
 DET 女人 给 DET 男人 DET 书
 ‘这个女人给这个男人这本书。’

⁷ 在某些条件下, 主语在英语中也可以被省略。更多有关命令式和其他无主句的例子, 请看第 488 页。

5 广义短语结构语法

对此, 必须使用 (47) 中的元规则。该元规则从规则右边的范畴集合中去除了一个任意的范畴 X, 并在左边用斜杠 (‘/’) 来表示:⁸

- (47) $V3 \rightarrow W, X \mapsto$
 $V3/X \rightarrow W$

该规则从 (48) 中创造了 (49) 的规则:

- (48) $V3 \rightarrow H[8], N2[CASE\ dat], N2[CASE\ acc], N2[CASE\ nom]$
(49) $V3/N2[CASE\ nom] \rightarrow H[8], N2[CASE\ dat], N2[CASE\ acc]$
 $V3/N2[CASE\ dat] \rightarrow H[8], N2[CASE\ acc], N2[CASE\ nom]$
 $V3/N2[CASE\ acc] \rightarrow H[8], N2[CASE\ dat], N2[CASE\ nom]$

(50) 中的规则将动词位于首位的语序与句中缺失的成分联系起来:

- (50) $V3[+FIN] \rightarrow X[+TOP], V3[+MC]/X$

在 (50) 中, X 代表一个任意范畴, 在 V3 中用 “/” 标记为空。X 指填充语 (filler)。在我们的例子中, X 的值的格在 (51) 中被详细列出来:

- (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) 并没有显示实际的规则。相反, (51) 展示了将具体范畴插入 X-位置的例子, 即规则的不同实例。

下列线性规则保证了 (50) 中标记为 [+TOP] 的成分位于句子其他成分之前:

- (52) $[+TOP] < X$

TOP 表示话题化 (topicalized)。正如我们在第 98 页提到的, 前场并没有限制为话题。焦点元素与虚位成分可以在前场出现, 这就是为什么特征名称并不理想的原因。但是, 有可能将其替换为其他名称, 如前场 (prefield)。这不会影响我们的分析。(52) 中的 X 表示一个任意范畴。这是一个新的 X, 与 (50) 中的那个是不同的。

图 5.4 在下一页显示了 (53) 中分析的规则的互动关系。⁹

- (53) Dem Mann gibt er das Buch.
DET.DAT 男人 给 他.NOM DET.ACC 书
‘他给这个男人这本书。’

(47) 中的元规则允准了将与格宾语加到斜杠上的规则。该规则现在允准了 “gibt er das Buch” (给他这本书) 的子树。V[+MC] < X 的线性规则规定了动词位于 V3 局部树的最左边。下一步, 斜杠后面的成分被绑定了。在 LP-规则 [+TOP] < X 之后, 绑定的成分必须位于 V3 节点的左边。

⁸ 这里解释的另一种 Uszkoreit 的无语迹的分析 (1987: 77), 包括在 GB 中对提取的元素使用语迹来表示。

⁹ FIN 特征在某些节点上被省略了, 因为它是冗余的: +MC-动词总是需要 FIN 的值为 “+”。

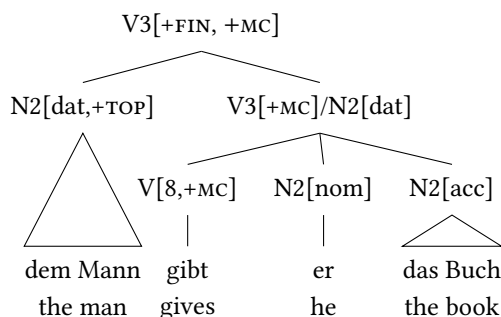


图 5.4: GPSG 中前置的分析

图 5.4 中的分析看起来有点过于复杂，因为 (53) 中的名词短语全都依存于同一个动词。我们可以创建一个线性规则系统来允许我们分析带有整个平铺结构的 (53)。尽管如此，我们仍然需要在第 98 页的 (37) 中的句子 – 为了方便，这里重复为 (54) – 的分析：

- (54) a. [Um zwei Millionen Mark]_i soll er versucht haben, [eine Versicherung _{-i} zu betrügen].¹⁰
 大约两 百万 马克 应该他试图 AUX ART 保险公司
 INF 欺骗
 ‘他显然试图从保险公司骗取两百万德国马克。’
- b. „Wer_i glaubt er, daß er _{-i} ist?“ erregte sich ein Politiker vom Nil.¹¹
 谁 相信 他 CONJ 他 AUX 反驳 REFL ART 政客 PREP 尼罗河
 ‘“他认为他是谁呀？”一位来自尼罗河的政客声称道。’
- c. Wen_i glaubst du, daß ich _{-i} gesehen habe?¹²
 谁 认为 你 CONJ 我 看见 AUX
 ‘你认为我看到谁了？’
- d. [Gegen ihn]_i falle es den Republikanern hingegen schwerer,
 反对 他 陷阱它 DET 共和党人闷 但是 更难
 [[Angriffe _{-i}] zu lancieren].¹³
 攻击 INF 发起
 ‘但是，共和党们更难对他发起攻击。’

(54) 中的句子不能按照局部重新排序来进行解释，因为前场中的元素并不依存于最高的动词，但是相反是从低层的小句而来的。因为只有从同一局部树中而来的

¹⁰ taz, 2001 年 5 月 4 日，第 20 页。

¹¹ Spiegel, 1999 年 8 月，第 18 页。

¹² Scherpenisse (1986: 84).

¹³ taz, 2008 年 2 月 8 日，第 9 页。

元素才能重新排序，(54)中的句子不能在没有针对长距离依存安排其他机制的情况下得到分析。¹⁴

在我对这章进行总结之前，我还将讨论另一个前置的例子，即(54)中更为复杂的例子。(54)中的分析包括几个步骤：导入、渗透和最后绑定长距离依存信息。这些步骤如图 5.5 所示。简单来说，我认为 *gesehen habe* (已经看见) 就像一个正

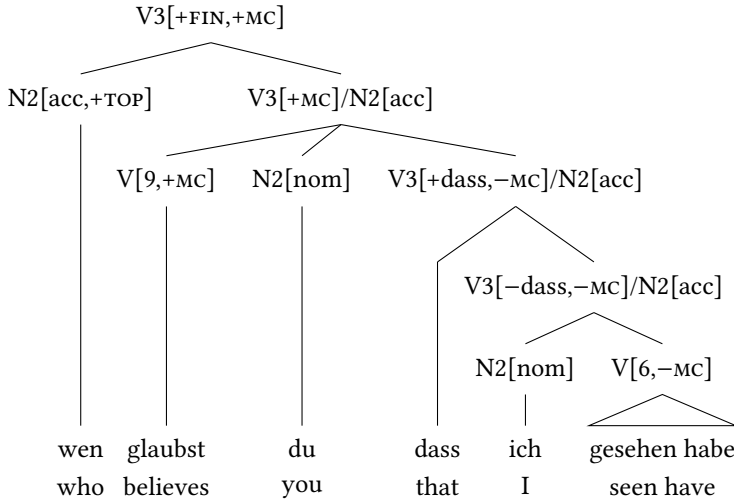


图 5.5: GPSG 中的长距离依存分析

常的及物动词。¹⁵ 由(47)中的元规则允准的短语结构规则允准了“ich”(我)和“gesehen habe”(已经看见)的组合，并在 V3 节点上表示为缺失的宾格宾语。标句词“dass”(那个)与“ich gesehen habe”(我已经看见)相组合，而且有关宾格 NP 缺失的信息在树中具有渗透性。这一渗透过程由所谓的次特征原则 (Foot Feature Principle) 控制，它是指所有的子节点的次特征也在母节点上出现。因为 SLASH 特征 (空位特征) 是一个次特征，如果他们没有在局部树内绑定的话，在“/”后的范畴会在树中自下而上地渗透。最后一步，V3/N2[acc] 与缺失的 N2[acc] 相组合。结果是一个在最高投射层的完整的定式陈述句。

5.5 总结

在 Chomsky 针对短语结构语法的批评二十年后，GPSG 框架下的大规模语法片段做成了，而且针对简单的短语结构规则无法解释的现象都提供了解释。尽管

¹⁴ 我们可以想象，针对非局部依存的特征机制只针对那些真正包括非局部依存的句子。这在 HPSG 中，由 Kathol (1995) 和 Wetta (2011) 做到了，在依存语法中由 Groß & Osborne (2009) 做到了。我将在第 11.7.1 节详细讨论依存语法分析，并且指出将简单的 V2 句子处理为非 V2 句子的变体的分析在下面几个方面是有问题的，包括前置状语的范围、简单句的和非局部依存句的并列，以及所谓的多重前置。

¹⁵ 请看 Nerbonne (1986a) 和 Johnson (1986) 有关 GPSG 中动词复杂式的分析。

GPSG 的工作基本上是基于 Harman1963 年的无转换的语法，他们也已经远远超越了这一层次。特别是，GPSG 的一个特殊成就是 Gazdar (1981b) 对长距离依存的处理。通过应用 SLASH-机制，可以解释同时从连词中提取元素（跨域提取，Ross 1967）下面从 Gazdar (1981b: 173) 选取的例子表明了连词中的空位必须是一致的，即某范畴的填充语必须对应于每个连词的空位：Some

- (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)
- c. *The kennel (in) which Mary made and Fido sleeps has been stolen.
(= S/NP & S/PP)

GPSG 貌似可以处理这类带有空位信息传送的机制。在对称并列结构中，每个连词中的 SLASH 成分必须是相同的。一方面，转换的方法并不直接，因为我们通常认为这类分析中有一棵树，其中树中的每个成分移到其他位置上而留下了一个语迹。但是，在并列结构中，填充语对应于两个或更多个语迹，而且它不能解释为什么填充语来自不只一个地方。

但是跨域提取分析是 GPSG 中的一个重中之重，其中有一些问题，我将在下面指出：配价与形态的互动、配价与部分动词短语前置的表征，以及 GPSG 形式化的表现力。

5.5.1 配价与形态

GPSG 中配价的编码在几个方面是有问题的。例如，形态化进程将词的配价属性考虑进来。带有后缀 *-bar* 的形容词只跟及物动词是可以生成的，即带有宾格宾语的动词可以进行被动化：

- (56) a. lös-bar (主格, 宾格)
解决-able
- b. vergleich-bar (主格, 宾格, PP[mit])
比较-able
- c. *schlaf-bar (主格)
睡觉-able
- d. *helf-bar (主格, 与格)
帮助-able

带有 *-bar* 的推导规则必须指向配价信息。这在 GPSG 语法中是不可能的，因为每个词条都只能指派给一个数字，该数字说明可以使用的这个词条的规则相关的信息。对于 *-bar-derivations* 来说，我们必须在推导规则中列出与宾格宾语的规则相对应的所有数字，他们当然无法充分地描述这些现象。进而，得到的形容词的配价也受限于动词的配价。例如，*vergleichen*（比较）这类动词需要一个 *mit* (with)-PP，而且 *vergleichbar* 也是一样的 (Riehmann 1993: 7, 54、1998: 68)。在下面的章节中，

我们将会遇到这样的模型，其中词汇项包括动词是否选择了宾格宾语的信息。在这类模型中，需要语言对象的配价属性的形态规则可以被较好地形式化表示出来。

关于配价和派生形态的问题还会在第 21.2.2 节有所讨论，我们会讨论 LFG 和构式语法中用到的方法，这些方法都共享了 GPSG 中有关配价的编码的思想。

5.5.2 配价与部分动词短语前置

Nerbonne (1986a) 和 Johnson (1986) 在 GPSG 框架下研究了部分 VPs 的前置问题。如 (57) 所示：在 (57a) 中，光杆动词前置，而其论元在中场被实现，在 (57b) 中，一个宾语与动词一起前置，而且在 (57c) 中所有的宾语都与动词一起前置。

- (57) a. Erzählen wird er seiner Tochter ein Märchen können.
告诉 AUX 他他的 女儿 一 童话故事AUX
b. Ein Märchen erzählen wird er seiner Tochter können.
一 童话故事告诉 AUX 他他的 女儿 AUX
c. Seiner Tochter ein Märchen erzählen wird er können.
他的 女儿 一 童话故事告诉 AUX 他AUX
‘他能够给他的女儿讲一个童话故事。’

(57) 中的句子的问题是动词 *erzählen* (告诉) 的配价要求在句子中的多个位置上被满足。对于前置的成分，我们需要一条规则来允许双及物动词实现为没有论元或者带有一个或两个宾语。进而，需要确保的是，在前场缺失的论元在小句的剩余部分被实现。省略必有论元或者将论元实现为不同的格属性是不合法的，如 (58) 中的例子所示：

- (58) a. Verschlungen hat er es nicht.
吞噬 AUX 他.NOM 它.ACC 不
‘他没有把它吞噬。’
b. *Verschlungen hat er nicht.
吞噬 AUX 他.NOM 不
c. *Verschlungen hat er ihm nicht.
吞噬 AUX 他.NOM 他.DAT 不

明显的结论是前置和非前置的论元必须加在属于动词的所有集合中。这在 GPSG 中基于规则的配价表征中是极为少见的。在诸如范畴语法的理论中（请看第 8 章），有可能构成 (58) 中较好的分析 (Geach 1970)。Nerbonne 和 Johnson 都提出了诸如 (58) 的句子分析，其中最终在范畴语法的导向下改变了配价信息的表征。

在我转向 GPSG 的形式化的表示力的问题之前，我要指出的是，我们在前面几节中讨论的问题都与 GPSG 的配价表征是相关的。我们在讨论第 5.2 节的被动的时候就已经发现了与配价相关的问题：由于主语和宾语由短语结构规则导入，而且由于在有些语言中，主语和宾语不在同一棵局部树内，看起来无法描述 GPSG 中主语受到抑制的被动式。

5.5.3 普遍能力

在 GPSG 中，线性系统、统制和元规则通常由条件所限制。这些条件我们不会在这里按照下面的方式来讨论，即我们可以从 GPSG 语法的具体化中创造出我们在第 2 章看到的这类短语结构语法。这类语法也叫做上下文无关文法。在上世纪 80 年代中期，一般认为上下文无关文法无法描述自然语言，语言需要比上下文无关文法更为强有力的语法形式化系统（Shieber 1985; Culy 1985，历史文献请看 Pullum (1986)）。所谓的语法形式化的生成能力（generative capacity）在第 17 章有所讨论。

随着 HPSG 这类基于约束的模型（请看第 9 章）和范畴语法这类基于约束的模型变体（请看第 8 章和 Uszkoreit 1986a）的出现，大部分之前在 GPSG 框架下工作的学者转向了其他理论框架。GPSG 对长距离依存的分析以及直接统制和线性优先的区别仍然在如今的 HPSG 和构式语法的变体中使用。请看第 12.2 节有关树邻接语法的变体将优先与统制区别开来的分析。

思考题

1. 在 ID/LP 形式下，语法是指什么？
2. 中场的成分变体的顺序在 GPSG 中是如何分析的？
3. 请找出一些由转换语法描述的现象，并思考 GPSG 是如何运用其他手段来分析这些现象的。

Exercises

1. 请写出能够分析下列句子的一个小型的 GPSG 语法：

- (59) a. [dass] der Mann ihn liest
 CONJ DET.NOM 男人 它.ACC 读
 ‘这个男人在读它’
- b. [dass] ihn der Mann liest
 CONJ 它.ACC DET..NOM 男人 读
 ‘这个男人在读它’
- c. Der Mann liest ihn.
 DET..NOM 男人 读 它.ACC
 ‘这个男人在读它。’

注意要在每条规则中包括所有的论元，而不是应用一条引入主语的元规则。

延伸阅读

GPSG 的主要文献是 Gazdar, Klein, Pullum & Sag (1985)。该书被 Jacobson (1987b) 进行了严肃的评判。有些有问题的分析与范畴语法中的方法进行了对比，而且参考

5 广义短语结构语法

了受到范畴语法重大影响的Pollard (1984)，这篇文献被认为是 HPSG 的先驱之作。Jacobson 的一些设想可以在 HPSG 晚一些的文献中找到。

德语语法可以在Uszkoreit (1987)和Busemann (1992)中找到。Gazdar (1981b)提出了长距离依存分析，该方法如今仍在 HPSG 等理论中所使用。

GPSG 的历史起源可以在Pullum (1989b)中找到相关的信息。

第六章 特征描写

在上一章，我们谈到了可以用来描写语言对象的特征-值偶对。本章中，我们将介绍在 LFG、HPSG、构式语法、范畴语法、TAG（甚至是最简方案理论的某些形式化 (Veenstra 1998)）等理论中发挥重要作用的特征描写。由此，本章将为后面的章节打下一些基础。

特征结构是可以模拟语言对象属性的复杂实体。大部分使用特征描写的语言学家只描写指定的特征结构中的部分属性。我们将在第 6.7 节详细解释模型与描写之间的差异。

表示特征结构的其他术语有：

- 特征-值结构
- 属性-值结构

其他有关特征描写的术语有：

- 特征-值矩阵 (AVM)
- 特征矩阵

接下来，为了保证本书的形式部分尽可能地简短，我将只讨论必要的细节。我推荐感兴趣的读者参考 Shieber (1986)、Pollard & Sag (1987: Chapter 2)、Johnson (1988)、Carpenter (1992)、King (1994) 和 Richter (2004)。Shieber 的著作对合一语法的介绍浅显易懂。King 和 Richter 介绍了 HPSG 的重要理论基础，这对于在数学方面没有打下良好基础的读者来说是较为容易理解的。尽管如此，要知道这些文献以及相应的理论是建立在坚实的基础上的是非常重要的。

6.1 特征描写

当描写语言符号时，我们必须说明属性的信息。对于名词而言，我们会说它有格、性、数和人称的特征。对于 *Mannes*（男人）这个词来说，我们可以说这些特征的值为 *genitive*（属格）、*masculine*（阳性）、*singular*（单数）及 3（第三人称）。如果我们要将这些信息写成一个特征-值偶对的列表的话，我们就会得到下面的特征描写形式：

6 特征描写

- (1) *Mannes* (男人) 的特征-值偶对:
- | | |
|-------------|-----------------------|
| CASE (格) | <i>genitive</i> (属格) |
| GENDER (性) | <i>masculine</i> (阳性) |
| NUMBER (数) | <i>singular</i> (单数) |
| PERSON (人称) | 3 (第 3 人称) |

我们也可以用特征描写来描述不同的事物。例如, 我们可以像例 (2) 那样来描述一个人:

- (2)
- | | |
|----------------------|--------------------------------------|
| FIRSTNAME (名) | <i>max</i> |
| LASTNAME (姓) | <i>meier</i> |
| DATE-OF-BIRTH (出生日期) | <i>10.10.1985</i> (1985 年 10 月 10 日) |

人与人之间的关系也可以在特征值偶对中表示。例如, Max Meier 的父亲叫做 Peter Meier 这样的事实可以通过对 (2) 的扩展来表示, 如下所示:

- (3)
- | | | | | | | | | | | | |
|---------------|---|-----------|--------------|----------|--------------|---------------|-------------------------------------|--------|-----|--------|-----|
| FIRSTNAME | <i>max</i> | | | | | | | | | | |
| LASTNAME | <i>meier</i> | | | | | | | | | | |
| DATE-OF-BIRTH | <i>10.10.1985</i> (1985 年 10 月 10 日) | | | | | | | | | | |
| 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> (1960 年 5 月 10 日)</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> (1960 年 5 月 10 日) | FATHER | ... | MOTHER | ... |
| FIRSTNAME | <i>peter</i> | | | | | | | | | | |
| LASTNAME | <i>meier</i> | | | | | | | | | | |
| DATE-OF-BIRTH | <i>10.05.1960</i> (1960 年 5 月 10 日) | | | | | | | | | | |
| FATHER | ... | | | | | | | | | | |
| MOTHER | ... | | | | | | | | | | |
| MOTHER | ... | | | | | | | | | | |

FATHER (父亲) 特征的值是包含 (2) 中同样特征的特征描写。

在特征描写中, 路径 (path) 是一个直接彼此承袭的序列特征。路径值 (value of a path) 是在路径末端的特征描写。由此, FATHER|DATE-OF-BIRTH 的值为 *10.05.1960*。

我们设想出在诸如 (3) 的表示中包括很多不同的特征。有人可能会问, 如何在 (3) 中加入后代的信息呢?

一个显而易见的解决方案就是加入 DAUGHTER (女儿) 和 SON (儿子) 这两个特征:

- (4)
- | | |
|---------------|-------------------|
| FIRSTNAME | <i>max</i> |
| LASTNAME | <i>meier</i> |
| DATE-OF-BIRTH | <i>10.10.1985</i> |
| FATHER | ... |
| MOTHER | ... |
| DAUGHTER | ... |

但是, 这一方法也有让人不满意的地方, 比如它无法直接清晰地说明如何来描写有几个女儿的人。是否应该引入诸如 DAUGHTER-1 或 DAUGHTER-3 这样的特征呢?

(5)

FIRSTNAME	<i>max</i>
LASTNAME	<i>meier</i>
DATE-OF-BIRTH	<i>10.10.1985</i>
FATHER	...
MOTHER	...
DAUGHTER-1	...
DAUGHTER-2	...
DAUGHTER-3	...

我们想要设定多少个特征呢？限制是什么？DAUGHTER-32 的值会是什么呢？

在这一点上，我们有必要使用列表。列表用尖括号表示。任意数量的元素可以出现在这些尖括号中。特殊的情况是在这些括号中没有元素。一个没有元素的列表被叫做空列表（empty list）。下例中，Max Meier 有一个女儿叫做 Clara，而她没有女儿。

(6)

FIRSTNAME	<i>max</i>												
LASTNAME	<i>meier</i>												
DATE-OF-BIRTH	<i>10.10.1985</i>												
FATHER	...												
MOTHER	...												
DAUGHTER	<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>10.10.2014 (2004 年 10 月 10 日)</i></td></tr> <tr><td>FATHER</td><td>...</td></tr> <tr><td>MOTHER</td><td>...</td></tr> <tr><td>DAUGHTER</td><td><></td></tr> </table>	FIRSTNAME	<i>clara</i>	LASTNAME	<i>meier</i>	DATE-OF-BIRTH	<i>10.10.2014 (2004 年 10 月 10 日)</i>	FATHER	...	MOTHER	...	DAUGHTER	<>
FIRSTNAME	<i>clara</i>												
LASTNAME	<i>meier</i>												
DATE-OF-BIRTH	<i>10.10.2014 (2004 年 10 月 10 日)</i>												
FATHER	...												
MOTHER	...												
DAUGHTER	<>												

现在，我们还剩下与儿子有关的问题。是否应该加上一个儿子的列表？我们希望区分儿子和女儿吗？显然，孩子的性别是一个重要的属性，但是宾语本身有属性，因为每个人都有性别。由此，(7) 中的描述更为合适。

到这儿，有人可能会问为什么父母没有用列表表示。事实上，我们在语言学的研究中也发现了类似的问题：如何能最好地整理手头工作的信息？有人可能会提出将父母的描写按照不同的特征来表示，并指出这样的表示可以说明一个母亲或者父亲，而不用在列表中搜索各自的描写。

如果元素的序列是无关的，那么我们可以用集合，而不是列表。集合用弧形括号（curly brackets）表示。¹

6.2 类型

在上一节中，我们介绍了包括特征-值偶对的特征描述，并且说明了有必要在特征上赋予复杂的值。在本节，特征描写将扩大到包括类型。赋予了类型的特征

¹ 集合的定义需要很多技术指标。本书中，我只用集合来表示语义信息。这点用列表也可以做到，这就是为什么我在这里没有引入集合，而是使用了列表。

6 特征描写

(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>10.10.2014</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>10.10.2014</i>	GENDER	<i>female</i>	FATHER	...	MOTHER	...	CHILDREN	<>
FIRSTNAME	<i>clara</i>														
LASTNAME	<i>meier</i>														
DATE-OF-BIRTH	<i>10.10.2014</i>														
GENDER	<i>female</i>														
FATHER	...														
MOTHER	...														
CHILDREN	<>														

描写也叫做类型特征描写。类型是说的哪些特征可以或者必须属于一个具体的结构。前面讲到的描写表示了类型 *person*（人）的对象。

(8)

<i>person</i>	
FIRSTNAME	<i>max</i>
LASTNAME	<i>meier</i>
DATE-OF-BIRTH	<i>10.10.1985</i>
GENDER	<i>male</i>
FATHER	...
MOTHER	...
CHILDREN	$\langle \dots, \dots \rangle$

类型用斜体表示。

类型的具体化决定了所模拟的对象具有什么样的属性。这样一个理论才能描写这些属性。诸如 OPERATING VOLTAGE（工作电压）这样的属性与类型 *person*（人）是无关的。如果我们知道一个给定对象的类型，那么我们也会知道该对象一定具有一定的属性，即使我们还不知道他们具体的值。这样，(9) 仍是对 Max Meier 的描写，即使它并不包括任何有关 Max 的生日信息：

(9) $\begin{bmatrix} person \\ \text{FIRSTNAME} & max \\ \text{LASTNAME} & meier \\ \text{GENDER} & male \end{bmatrix}$

但是，我们知道 Max Meier 一定是在某天出生的，因为这是对类型 *person*（人）的描述。对于 (9) 这类结构来说，“Max 的生日是什么？”这个问题是有意义的，而“Max 有哪种工作电压？”这个问题就是无意义的。如果我们知道一个对象是属于类型 *person*（人）的，那么就会有如下的基本结构：

(10)

<i>person</i>	
FIRSTNAME	<i>firstname</i>
LASTNAME	<i>lastname</i>
DATE-OF-BIRTH	<i>date</i>
GENDER	<i>gender</i>
FATHER	<i>person</i>
MOTHER	<i>person</i>
CHILDREN	<i>list of person</i>

在 (10) 和 (9) 中, FIRSTNAME (名) 这类特征的值用斜体表示。这些值也是类型。但是, 它们与 *person* (人) 这类特征不同, 因为它们没有特征。这些特征叫做原子式 (atomic)。

特征按照层级来进行组织。对于 *person* (人) 来说, 可以界定次类型 *woman* (女人) 和 *man* (男人)。这会决定给定对象的性别。(11) 显示了类型 *woman* (女人) 的特征结构, 这与类型 *man* (男人) 的特征结构是类似的。

(11)

<i>female person</i>	
FIRSTNAME	<i>firstname</i>
LASTNAME	<i>lastname</i>
DATE-OF-BIRTH	<i>date</i>
GENDER	<i>female</i>
FATHER	<i>person</i>
MOTHER	<i>person</i>
CHILDREN	<i>list of person</i>

在这点上, 我们应该自问是否需要 GENDER (性别) 这个特征。必要信息已经在 *woman* (女人) 这个类型中显示出来了。具体信息是由具体特征所表示的还是它是储存在没有对应个人特征的类型中这一问题会在语言学的分析中浮现出来。这两个方面的差异是巨大的, 因为通过类型模拟的信息的事实没有直接通过结构共享而获得, 这点在第 6.4 节有所讨论。

类型层级体系在捕捉语言学的一般性特征方面发挥了重要的作用, 这就是为什么类型层级体系和约束与信息的承继需要在后面的例子中进行解释的原因。我们可以将类型层级体系看成是一种有效的组织信息的方式。在百科辞典中, 个体之间是相互联系的, 比如说猴子和老师这两个词条的联系在于二者都指向哺乳动物。针对哺乳动物的描写同样也适用于从属于它的概念中。同样, 如果我们希望描写不同的电子设备, 我们可以应用图 6.1 on the next page 中的层级体系。图中最高点是最为普遍的类型“电子设备”(electrical device)。电子设备具有一定的属性, 比如说带有特定能量消耗的能量供给。“电子设备”的所有此类型都“承继”了这一属性。这样, “打印设备”和“扫描设备”也具有可供特定能量消耗的能力供给。“打印设备”可以制造信息, 而“扫描设备”可以阅读信息。“复印机”既可以制造信息也可以阅读信息。复印机具有扫描机和打印机的属性。这在图 6.1 中通过两个上位类型和“复印机”之间的联系来表示。如果一个类型同时也是几个上位类型的子类型, 那么我们就可以说这是多重承继 (multiple inheritance)。如果一个设置可以打印, 但是不能扫描, 它们就属于类型打印机 (printer)。该类型有

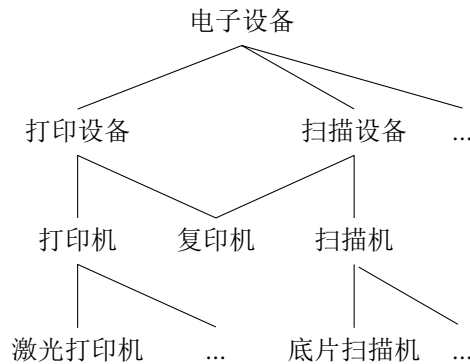


图 6.1: 多重承继的非语言学例子

更多的具体次类型，相应地可以有具体的属性，比如说“激光打印机”。新的特征可以被加进次类型中，但是也可以将承继的特征的值做得更具体。比如说，可以用“底片扫描机”扫描的材料比它的上位类型“扫描机”具有更多的限制，因为底片扫描机只能扫描底片。

模拟的对象都有一个最大的具体类型。在上例中，这是指我们可以有类型“激光打印机”和“底片打印机”的对象，而不是类型“打印设备”。这是因为，“打印设备”并不具有最大的具体化，因为该类型有两个次类型。

带有多重承继的类型层级体系是表达语言一般化的重要手段。在这些层级的最高点出现的词或短语的类型对应于语言学对象的约束条件，这对于所有语言中的语言对象来说都是合理的。这种一般类型的子类型可以具体到某些语言或语言类型。

6.3 析取

如果有人想表达一个具体物体具有不同属性的事实可以用析取来表示。如果有人想组织一场毕业二十年的聚会，但是不记得一些老同学的名字了，可以在网络中搜索“Julia (Warbanow 还是 Barbanow)”。在特征描写中，这个“还是”表示为‘ \vee ’。

(12)
$$\left[\begin{array}{l} person \\ \text{FIRSTNAME } julia \\ \text{LASTNAME } warbanow \vee barbanow \end{array} \right]$$

一些网络的搜索引擎不允许使用带有“或者”的搜索。这种情况下，我们需要给出两个不同的搜索选项：一个是“Julia Warbanow”，而另一个是“Julia Barbanow”。这就对应于下面用析取连接的描写形式：

$$(13) \quad \left[\begin{array}{l} \textit{person} \\ \text{FIRSTNAME } \textit{julia} \\ \text{LASTNAME } \textit{warbanow} \end{array} \right] \vee \left[\begin{array}{l} \textit{person} \\ \text{FIRSTNAME } \textit{julia} \\ \text{LASTNAME } \textit{barbanow} \end{array} \right]$$

因为我们将类型层级看作是一种表达的手段，我们有时可以不做具体的析取，而是用上级类型：以“打印机” \vee “复印机”为例，我们可以简单地写“打印设备”，如果我们按照图 6.1 在上一页所示的类型层级的话。

6.4 结构共享

结构共享是形式化中重要的一部分。它用来表示结构中某些相同的部分。关于值的统一性的语言学方面的例子就是一致关系。在例 (14) 的句子中，名词短语的数的值必须与动词保持一致：

- (14) a. Der Mann schläft.
DET 男人 睡觉
‘这个男人正在睡觉。’
b. Die Männer schlafen.
DET 男人 睡觉
‘这些男人正在睡觉。’
c. *Der Mann schlafen.
DET 男人 睡觉
想说“这些男人正在睡觉。”

相同的值通过带有数字的盒子来表示。这些盒子可以看作是变量。

当我们描述对象时，我们可以说明相等或相同的值。关于值的同一性的说明是更为强势的。让我们用下面的 Max 的父亲和母亲的孩子们的信息的例子来进行说明：

6 特征描写

(15)	<i>person</i>	
	FIRSTNAME	<i>max</i>
	LASTNAME	<i>meier</i>
	DATE-OF-BIRTH	<i>10.10.1985</i>
	FATHER	<i>person</i>
		FIRSTNAME <i>peter</i>
		LASTNAME <i>meier</i>
		CHILDREN $\left\langle \left[\begin{array}{l} \textit{person} \\ \text{FIRSTNAME } \textit{klaus} \end{array} \right], \dots \right\rangle$
	MOTHER	<i>person</i>
		FIRSTNAME <i>anna</i>
		LASTNAME <i>meier</i>
		CHILDREN $\left\langle \left[\begin{array}{l} \textit{person} \\ \text{FIRSTNAME } \textit{klaus} \end{array} \right], \dots \right\rangle$

请注意 FATHER|CHILDREN 和 MOTHER|CHILDREN 的路径下面，我们找到一个包括名为 Klaus 的人的描述。有关该特征描述是 Peter 和 Anna 的一个孩子还是两个孩子的问题是无法回答的。当然，我们是在分析前面关系中的两个不同的孩子，只不过他们碰巧都叫做 Klaus。

通过结构共享，我们可以确定出例（16）中两个值的同一性。在例（16）中，

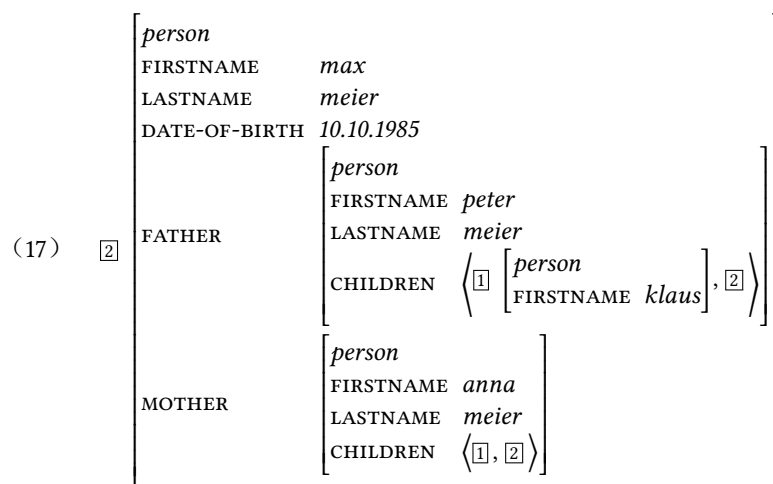
(16)	<i>person</i>	
	FIRSTNAME	<i>max</i>
	LASTNAME	<i>meier</i>
	DATE-OF-BIRTH	<i>10.10.1985</i>
	FATHER	<i>person</i>
		FIRSTNAME <i>peter</i>
		LASTNAME <i>meier</i>
		CHILDREN $\left\langle \left[\text{①} \left[\begin{array}{l} \textit{person} \\ \text{FIRSTNAME } \textit{klaus} \end{array} \right], \dots \right] \right\rangle$
	MOTHER	<i>person</i>
		FIRSTNAME <i>anna</i>
		LASTNAME <i>meier</i>
		CHILDREN $\left\langle \left[\text{①}, \dots \right] \right\rangle$

Klaus 是父母双亲的独生子女。出现在①之后的所有括号内的信息在所有的位置上都是等同出现的。我们可以把①看作是一个指针或坐标，它指向只被描写过一次的结构。还有一个问题没有得到回答：Max 呢？Max 也是他父母的一个孩子，也应该出现在他父母所有的孩子的列表之中。在（16）中有两个点所在的地方有三个点。这些省略符号表示 Peter Meier 和 Anna Meier 的其他孩子的信息。我们的世界知识

告诉我们他们必须有同一个叫做 Max Meier 的孩子。在下一节，我们会看到这在形式表达中是如何表示的。

6.5 循环结构

我们引入了结构共享来表示 Max 的父母都有一个叫做 Klaus 儿子的事实。但是，将 Max 分别放在他的父母的孩子列表之中是不够的。我们还想捕捉到这样的事实，即在这些列表中的每一个表中出现的都是同一个 Max，而且，我们想要确保被描写的孩子与描写的整体对象是一致的。否则，描写就会允许这样的情况，Max 的父母可以有第二个叫做 Max 的孩子。例 (17) 中给出的描述可以正确地捕捉所有的事实。例 (17) 中所描写的结构叫做循环结构，因为如果我们按照一个具体的路



径就会进入一个循环：如路径 FATHER|CHILDREN|...|FATHER|CHILDREN|...²可以重复无限的次数。

6.6 合一

语言规则在特征描写的帮助下可以写成与 HPSG 和构式语法中的词汇项完全相同的格式。对于短语中可以用作节点的一个词或者较大的短语实体来说，这个词或者短语必须具有语法规则中与子节点们的描写相兼容的属性。如果这种兼容性存在的话，那么我们可以说各自的对象是可以合一的 (unifiable)。³ 如果我们

² 这里的点是指到列表中值为 CHILDREN (孩子) [2] 的路径。请看练习 3。

³ 合一 (unification) 这个概念需要小心使用，它只在语言理论的形式化基础的某些假说下才是合适的。非正式的情况下，这一术语经常用于没有在技术上界定合一的形式化系统中。在 HPSG 中，它大部分是指对两个描写的约束导向一个单一的描写。我们在这里想直观说明的是，所描写的对象需要同时满足所有描写的约束条件 (约束满足)。因为合一 (unification) 这个术语应用范围较广，本节也采用这一概念。在后面的理论讨论中，除了基于合一的方法，我们不会再使用该术语。相反，这里给出的约束满足 (constraint satisfaction) 这个概念在后面的章节中发挥了重要的作用。

6 特征描写

将两个描写合一，可以得到包含这两个描写的信息，而没有额外信息的描写结果。

合一的工作原理可以通过描写人的特征描写来进行说明。我们可以想象 Bettina Kant 去私家侦探 Max Müller 那里，她想找一个人。通常来说，到私家侦探的办公室的人只有他们想找的人的一部分描写信息，比如说性别、头发颜色或者出生日期。也许还会知道那个人的汽车的登记号码。

这样，侦探就会期望他或者她能够提供出符合描写的信息。如果我们要找一位金发的叫 Meier 的女性 (18a)，那么我们就不想得到一位有红色头发的男性的描写 (18b)。例 (18) 中的描写就是不兼容的，而且不能合一：

- (18) a.
$$\left[\begin{array}{ll} \textit{person} & \\ \text{LASTNAME} & \textit{meier} \\ \text{GENDER} & \textit{female} \\ \text{HAIRCOLOR} & \textit{blonde} \end{array} \right]$$
- b.
$$\left[\begin{array}{ll} \textit{person} & \\ \text{LASTNAME} & \textit{meier} \\ \text{GENDER} & \textit{male} \\ \text{HAIRCOLOR} & \textit{red} \end{array} \right]$$

例 (19) 中描写的结果可能是寻找一位金发的、叫 Meier 的女性个体：

- (19)
$$\left[\begin{array}{ll} \textit{person} & \\ \text{FIRSTNAME} & \textit{katharina} \\ \text{LASTNAME} & \textit{meier} \\ \text{GENDER} & \textit{female} \\ \text{DATE-OF-BIRTH} & \textit{15.10.1965} \\ \text{HAIRCOLOR} & \textit{blonde} \end{array} \right]$$

Katharina Meier 还可以有其他侦探不知道的属性。重要的是侦探所知的属性与委托人要寻找的属性是一致的。进而，侦探使用可靠的信息而不是制造出有关寻找对象的任何信息是非常重要的。(18a) 中搜查的合一与 (19) 中侦探可获得的信息实际上是 (19)，而不是 (20)，如下所示：

- (20)
$$\left[\begin{array}{ll} \textit{person} & \\ \text{FIRSTNAME} & \textit{katharina} \\ \text{LASTNAME} & \textit{meier} \\ \text{GENDER} & \textit{female} \\ \text{DATE-OF-BIRTH} & \textit{15.10.1965} \\ \text{HAIRCOLOR} & \textit{blond} \\ \text{CHILDREN} & \langle \rangle \end{array} \right]$$

(20) 包括了孩子的信息，它既不属于 (18a)，也不属于 (19)。事实上，Katharina Meier 可能没有孩子，但是也有可能其他叫做 Katharina Meier 的人具有其他相同的属性。根据这一新创的信息，我们可以排除一个或者多个可能的候选人。

也有可能，我们的侦探 Max Müller 在他的文件中没有发色的信息。他的文件可以包括如下这些信息：

(21)	<i>person</i>	
	FIRSTNAME	<i>katharina</i>
	LASTNAME	<i>meier</i>
	GENDER	<i>female</i>
	DATE-OF-BIRTH	<i>15.10.1965</i>

这些数据与搜索的标准是相容的。如果我们要将 (18a) 和 (21) 进行合一，我们可以得到 (19)。如果我们假设侦探做得不错，那么现在就会知道她最初要找的人的属性，还有一些新发现的属性。

6.7 现象、模型和形式化理论

在前面的章节中，我们介绍了带有类型的特征描写。这些特征描写描述了类型特征结构，这些结构模拟了观察到的语言结构。在类型的定义中，我们决定了应该被描写的语言对象的属性。类型层级体系与类型的定义一起也叫做符号形式 (signature)。语法学家在特征描写中典型地使用了类型。这些描写包括语言对象必须具有的约束条件。如果没有约束，那么所有的符号形式中与具体化相容的值都是可能的值。例如，我们可以省略诸如 *Frau* (女人) 这一语言对象的格的描写，因为 *Frau* 可以出现在四种格之中，如例 (22) 所示：

- (22)
- | | | | |
|----|------------------|---------------|------|
| a. | Die | Frau schläft. | (主格) |
| | DET.NOM | 女人睡觉 | |
| b. | Wir gedenken der | Frau. | (属格) |
| | 我们想念 | DET.GEN 女人 | |
| c. | Er hilft der | Frau. | (与格) |
| | 他 帮助 | DET.DAT 女人 | |
| d. | Er liebt die | Frau. | (宾格) |
| | 他 爱 | DET.ACC 女人 | |

在给定的模型中，只有全部明确的表达式，即模型包括四种形式的 *Frau*，每种形式带有一种不同的格。对于阳性名词 *Mann* (男人) 来说，我们可以在描写中给出格的信息，因为属格-单数形式 *Mann-es* 与其他的单数形式是不同的，这可以在 (22) 的例子中加入 *Mann* 来看到。例 (23) 给出了 *Frau* (女人) 和 *Mann* (男人) 的特征描写：

- (23)
- | | |
|----|--|
| a. | Frau (女人) : |
| | $\left[\begin{array}{l} \text{GENDER } fem \end{array} \right]$ |
| b. | Mann (男人) : |
| | $\left[\begin{array}{l} \text{GENDER } mas \\ \text{CASE } \textit{nominative} \vee \textit{dative} \vee \textit{accusative} \end{array} \right]$ |

与 (23b) 不同的是，(23a) 并不包括格属性，这是因为我们不需要说明 *Frau* 的描写中有任何格的信息。由于所有的名词性对象都需要一个格属性，很清楚的是

Frau 的结构必须也有一个格特征。格特征的值属于类型 *case* (格)。*case* 是一个一般类型, 它包括次级类型 *nominative* (主格)、*genitive* (属格)、*dative* (与格) 和 *accusative* (宾格)。具体的语言对象总是将这些最大限度上确定的类型作为他们的格的值。(23) 的特征结构如图 6.2 和图 6.3 所示。

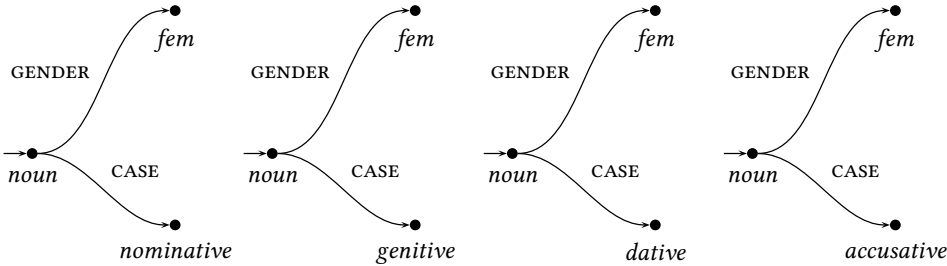


图 6.2: (23a) 中的 *Frau* (女人) 的特征结构描写

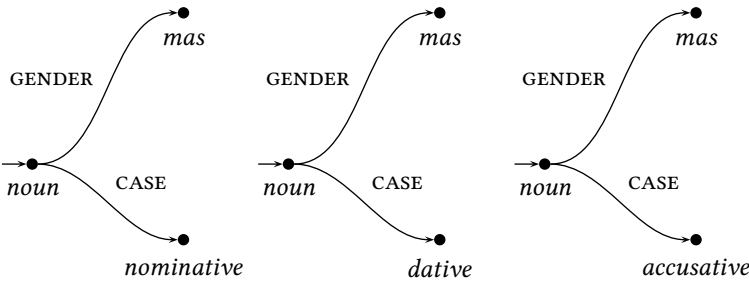


图 6.3: (23b) 中的 *Mann* (男人) 的特征结构描写

在这些表达式中, 每一个节点都有一定的类型 (*noun*、*fem*、*nominative* ...), 并且特征结构中的类型总是最大限度地具体化, 即他们没有深层的次类型。总有一个进入节点 (上例中的 *noun*), 以及其他用特征标签标注的用箭头连接起来的节点 (*GENDER*、*CASE*)。

如果我们回到上面章节中人的例子, 我们可以发现模型与描写之间的差异, 如下所示: 如果我们有一个人的模型, 它包括名、姓、出生日期、性别和发色, 那么它自然得到的结果是我们模拟的每个对象都会有生日。但是, 如果这些信息在表示约束或构成搜索时没有发挥重要的作用时, 我们可以在描写中决定省略这些细节。

语言现象、模型和形式化理论之间的联系如图 6.4 在下一页所示。模型是用来模拟语言现象的。进而, 它必须由我们的理论所允准。理论决定了模型并且对可能的现象进行预测。

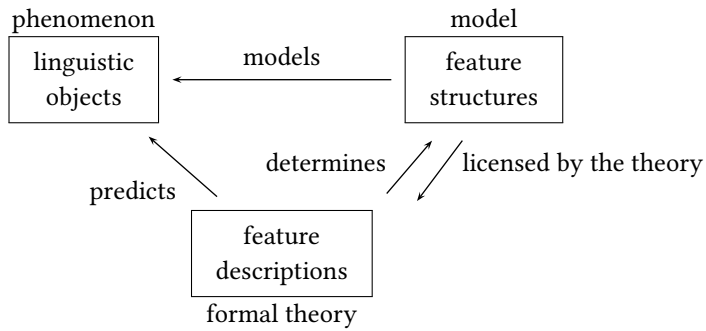


图 6.4: 现象、模型和形式化理论

思考题

1. 使用类型的原因是什么？
2. 什么是承继？多重承继有何特殊之处？
3. 下面的结构是相容的吗？也就是说，他们能用来描述相同的对象吗？

$$(24) \quad \left[\begin{array}{l} \text{FIRSTNAME } \textit{max} \\ \text{LASTNAME } \textit{meier} \\ \text{FATHER } \left[\begin{array}{l} \textit{person} \\ \text{FIRSTNAME } \textit{peter} \\ \text{LASTNAME } \textit{meier} \end{array} \right] \end{array} \right] \quad \left[\begin{array}{l} \text{FIRSTNAME } \textit{max} \\ \text{LASTNAME } \textit{meier} \\ \text{FATHER } \left[\begin{array}{l} \textit{person} \\ \text{FIRSTNAME } \textit{peter} \\ \text{LASTNAME } \textit{müller} \end{array} \right] \end{array} \right]$$

$$(25) \quad \left[\begin{array}{l} \text{FIRSTNAME } \textit{max} \\ \text{LASTNAME } \textit{meier} \\ \text{FATHER } \left[\begin{array}{l} \textit{person} \\ \text{FIRSTNAME } \textit{peter} \\ \text{LASTNAME } \textit{meier} \end{array} \right] \end{array} \right] \quad \left[\begin{array}{l} \text{FIRSTNAME } \textit{max} \\ \text{LASTNAME } \textit{meier} \\ \text{MOTHER } \left[\begin{array}{l} \textit{person} \\ \text{FIRSTNAME } \textit{ursula} \\ \text{LASTNAME } \textit{müller} \end{array} \right] \end{array} \right]$$

练习题

1. 请设想如何通过特征描写来描写乐器。
2. 请设计一个词类 (*det*、*comp*、*noun*、*verb*、*adj*、*prep*) 的类型层级体系。请设想可以组织类型层级的方式，这样我们可以对第87页的图 3.1中的二元特征的概括进行表示。

6 特征描写

3. 在本章，我们介绍了列表。这看起来好像是形式化的扩展，但是它并不是，因为可以将列表标记转化为只需要特征-值偶对的标记。请思考如何做到这一点。
4. (附加练习) 附加关系 (append) 将在第 9 章发挥作用。该关系用来将两个列表组合成第三个列表。诸如附加 (append) 的关系约束实际上构成了形式化的一种扩展。使用关系约束可以将任意数量的特征值与其他值联系起来，即我们可以写出这样的程序，它根据其他值计算出具体的值。这就导致了一个问题，在语言学理论中我们是否需要如此强有力的描写工具，并且如果我们允许使用它们，我们将会承担什么样的复杂度的代价。可见，我们最好选择那些不需要关系约束的理论，而不是那些需要关系约束的理论（请看Müller 2007b: Chapter 20对相关理论的比较）。

列表的串联可以在没有关系约束要求的特征结构中实现。请找到这是如何做到的，并提供你的数据来源，记录下你找到解决方案的途径。

延伸阅读

本章为读者设计了简单易懂的有关类型特征结构的介绍。结构的数学属性、类型层级体系以及这些结构的组合性概率不在这里详细说明，但是至少这些属性的一部分对于计算语言学的工作以及开发个人自己的分析来说都是非常重要的。更多的内容，我推荐感兴趣的读者阅读以下文献：Shieber (1986)是对合一语法理论的简短介绍。它针对重要的语法类型给出了相对全面的综述，如 DCG、LFG、GPSG、HPSG、PATR-II。Johnson (1988)按照数学的精确形式描述了非类型特征结构的形式化。Carpenter (1992) 重点分析了类型特征结构在数学上的表示。由King (1999)开发的 HPSG-语法构成了Richter (2004)的形式化的基础，该语法目前被看作是 HPSG 的标准形式化语法。

第七章 Lexical Functional Grammar

Lexical Functional Grammar (LFG) was developed in the 80s by Joan Bresnan and Ron Kaplan (Bresnan & Kaplan 1982). LFG forms part of so-called West-Coast linguistics: unlike MIT, where Chomsky works and teaches, the institutes of researchers such as Joan Bresnan and Ron Kaplan are on the west coast of the USA (Joan Bresnan in Stanford and Ron Kaplan at Xerox in Palo Alto and now at the language technology firm Nuance Communications in the Bay Area in California).

Bresnan & Kaplan (1982) view LFG explicitly as a psycholinguistically plausible alternative to transformation-based approaches. For a discussion of the requirements regarding the psycholinguistic plausibility of linguistics theories, see Chapter 15.

The more in-depth works on German are Berman (1996, 2003a) and Cook (2001).

LFG has well-designed formal foundations (Kaplan & Bresnan 1982; Kaplan 1995), and hence first implementations were available rather quickly (Frey & Reyle 1983a,b; Yasukawa 1984; Block & Hunze 1986; Eisele & Dorre 1986; Wada & Asher 1986; Delmonte 1990; Her, Higinbotham & Pentheroudakis 1991; Kohl 1992; Kohl, Gardent, Plainfossé, Reape & Momma 1992; Kaplan & Maxwell III 1996; Mayo 1997, 1999; Boullier & Sagot 2005a,b; Clément 2009; Clément & Kinyon 2001).

The following is a list of languages with implemented LFG fragments, probably incomplete:

- Arabic (Attia 2008),
- Arrernte (Dras, Lareau, Börschinger, Dale, Motazedi, Rambow, Turpin & Ulinski 2012),
- Bengali (Sengupta & Chaudhuri 1997),
- Danish (Ørsnes 2002; Ørsnes & Wedekind 2003, 2004),
- English (Her, Higinbotham & Pentheroudakis 1991; Butt, Dipper, Frank & King 1999a; Riezler, King, Kaplan, Crouch, Maxwell III & Johnson 2002; King & Maxwell III 2007),
- French (Zweigenbaum 1991; Frank 1996; Frank & Zaenen 2002; Butt, Dipper, Frank & King 1999a; Clément & Kinyon 2001; Boullier, Sagot & Clément 2005; Schwarze & de Alencar 2016),

- Georgian (Meurer 2009),
- German (Rohrer 1996; Berman 1996; Kuhn & Rohrer 1997; Butt et al. 1999a; Dipper 2003; Rohrer & Forst 2006; Forst 2006; Frank 2006; Forst & Rohrer 2009),
- Hungarian (Laczkó et al. 2010),
- Indonesian (Arka, Andrews, Dalrymple, Mistica & Simpson 2009),
- Italian (Delmonte 1990; Mayo 1999; Quaglia 2014),
- Irish (Sulger 2009, 2010),
- Japanese (Her, Higinbotham & Pentheroudakis 1991; Masuichi & Ohkuma 2003; Umemoto 2006),
- Korean (Her, Higinbotham & Pentheroudakis 1991),
- Malagasy (Randriamasimanana 2006; Dalrymple, Liakata & Mackie 2006),
- Mandarin Chinese (Her, Higinbotham & Pentheroudakis 1991; Fang & King 2007),
- Murrinh-Patha (Seiss & Nordlinger 2012),
- Norwegian (Dyvik, Meurer & Rosén 2005),
- Polish (Patejuk & Przepiórkowski 2012),
- Portuguese (de Alencar 2004, 2013),
- Spanish (Mayo 1999),
- Tigrinya (Kifle 2012),
- Turkish (Çetinoğlu & Oflazer 2006),
- Hungarian (Laczkó, Rákosi & Tóth 2010; Rákosi, Laczkó & Csernyi 2011),
- Urdu/Hindi (Butt, King & Roth 2007; Bögel, Butt & Sulger 2008),
- Welsh (Mittendorf & Sadler 2005) and
- Wolof (Dione 2014, 2013).

Many of these grammars were developed in the ParGram consortium¹ (Butt, King, Niño & Segond 1999b; Butt, Dyvik, King, Masuichi & Rohrer 2002). Apart from these grammars there is a small fragment of Northern Sotho, which is currently being expanded (Faaß 2010).

¹ <http://pargram.b.uib.no/research-groups/>. 01.10.2015.

Many of the LFG systems combine linguistically motivated grammars with a statistical component. Such a component can help to find preferred readings of a sentence first, it can increase the efficiency of processing and make the complete processing robust (for instance Kaplan et al. 2004; Riezler et al. 2002). Josef van Genabith's group in Dublin is working on the induction of LFG grammars from corpora (e.g., Johnson et al. 1999; O'Donovan et al. 2005; Cahill et al. 2005; Chrupala & van Genabith 2006; Guo et al. 2007; Cahill et al. 2008; Schluter & van Genabith 2009).

Some of the systems can be tested online:

- <http://iness.uib.no/xle-web/xle-web>
- <http://lfg-demo.computing.dcu.ie/lfgparser.html>
- <http://www.xlfg.org/>

7.1 General remarks on the representational format

LFG assumes multiple levels of representation.² The most important are c-structure and f-structure. c-structure is the constituent structure and it is licensed by a phrase structure grammar. This phrase structure grammar uses \bar{X} structures for languages for which this is appropriate. f-structure stands for functional structure. Functional structure contains information about the predicates involved and about the grammatical functions (subject, object, ...) which occur in a constituent. Mappings mediate between these representational levels.

7.1.1 Functional structure

In LFG, grammatical functions such as subject and object play a very important role. Unlike in most other theories discussed in this book, they are primitives of the theory. A sentence such as (1a) will be assigned a functional structure as in (1b):

- (1) a. David devoured a sandwich.
- b.
$$\left[\begin{array}{l} \text{PRED} \text{ 'DEVOUR(SUBJ, OBJ)'} \\ \text{SUBJ} \left[\begin{array}{l} \text{PRED} \text{ 'DAVID'} \end{array} \right] \\ \text{OBJ} \left[\begin{array}{l} \text{SPEC} \text{ A} \\ \text{PRED} \text{ 'SANDWICH'} \end{array} \right] \end{array} \right]$$

All lexical items that have a meaning (e.g., nouns, verbs, adjectives) contribute a PRED feature with a corresponding value. The grammatical functions governed by a head (government = subcategorization) are determined in the specification of PRED.³ Corresponding functions are called *governable grammatical functions*. Examples of this are shown in Table 7.1 在下一页 (Dalrymple 2006). The PRED specification corresponds to the theta grid in GB theory. The valence of a head is specified by the PRED value.

The non-governable grammatical functions are given in Table 7.2 在下一页. Topic and focus are information-structural terms. There are a number of works on their exact

² The English examples and their analyses discussed in this section are taken from Dalrymple (2001) and Dalrymple (2006).

³ In the structure in (1b), the SUBJ and OBJ in the list following *devour* are identical to the values of SUBJ and OBJ in the structure. For reasons of presentation, this will not be explicitly indicated in this structure and following structures.

表 7.1: Governable grammatical functions

SUBJ:	subject
OBJ:	object
COMP:	sentential complement or closed (non-predicative) infinitival complement
XCOMP:	open (predicative) complement, often infinitival, the SUBJ function is externally controlled
OBJ _θ :	secondary OBJ functions that are related to a special, language specific set of grammatical roles; English has OBJ _{THEME} only.
OBL _θ :	a group of thematically restricted oblique functions, as for instance OBL _{GOAL} or OBL _{AGENT} . These often correspond to adpositional phrases in c-structure.

表 7.2: Non-governable grammatical functions

ADJ:	adjuncts
TOPIC:	the topic of an utterance
FOCUS:	the focus of an utterance

definition, which differ to varying degrees (Kruijff-Korbayová & Steedman 2003: 253–254), but broadly speaking, one can say that the focus of an utterance constitutes new information and that the topic is old or given information. Bresnan (2001: 97) uses the following question tests in order to determine topic and focus:

- (2) Q: What did you name your cat?
A: Rosie I named her. (*Rosie* = FOCUS)
- (3) Q: What did you name your pets?
A: My dog, I named Harold. My cat, I named Rosie. (*my dog, my cat* = TOPIC)

f-structures are characterized using functional descriptions, for example, one can refer to a value of the feature TENSE in the functional structure f using the following expression:

- (4) (f TENSE)

It is possible to say something about the value which this feature should have in the feature description. The following descriptions express the fact that in the structure f , the feature TENSE must have the value PAST.

- (5) (f TENSE) = PAST

The value of a feature may also be a specific f-structure. The expression in (6) ensures that the SUBJ feature in f is the f-structure g :

7 Lexical Functional Grammar

$$(6) \quad (f \text{ SUBJ}) = g$$

For the analysis of (7a), we get the constraints in (7b):

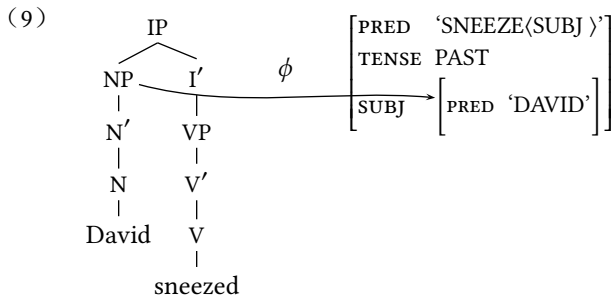
- (7) a. David sneezed.
 b. $(f \text{ PRED}) = \text{'SNEEZE(SUBJ)'}'$
 $(f \text{ TENSE}) = \text{PAST}$
 $(f \text{ SUBJ}) = g$
 $(g \text{ PRED}) = \text{'DAVID'}$

The description in (7b) describes the following structure:

$$(8) \quad f: \begin{bmatrix} \text{PRED} & \text{'SNEEZE(SUBJ)'} \\ \text{TENSE} & \text{PAST} \\ \text{SUBJ} & g: \begin{bmatrix} \text{PRED} & \text{'DAVID'} \end{bmatrix} \end{bmatrix}$$

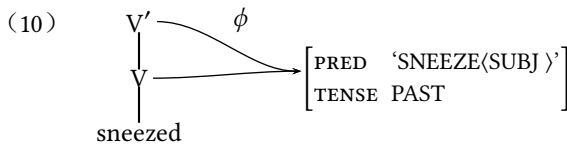
But (7b) also describes many other structures which contain further features. We are only interested in minimal structures that contain the information provided in the description.

(9) shows how a node in the c-structure can be connected to the f-structure for the entire sentence:



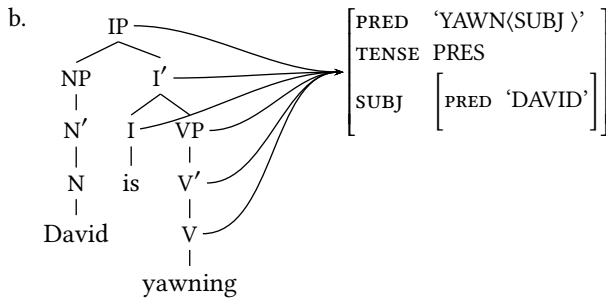
The function ϕ from the NP-node to the f-structure corresponding to the NP is depicted with an arrow marked ϕ .

A phrase and its head always correspond to the same f-structure:



In LFG grammars of English, the CP/IP system is assumed as in GB theory (see Section 3.1.5). IP, I' and I (and also VP) are mapped onto the same f-structure.

- (11) a. David is yawning.



f-structures have to fulfill two well-formedness conditions: they have to be both *complete* and *coherent*. Both these conditions will be discussed in the following sections.

7.1.2 Completeness

Every head adds a constraint of the PRED value of the corresponding f-structure. In determining completeness, one has to check that the elements required in the PRED value are actually realized. In (12b), OBJ is missing a value, which is why (12a) is ruled out by the theory.

- (12) a. * David devoured.
- b.
$$\left[\begin{array}{l} \text{PRED} \text{ 'DEVOUR(SUBJ,OBJ)'} \\ \text{SUBJ} \left[\begin{array}{l} \text{PRED} \text{ 'DAVID'} \end{array} \right] \end{array} \right]$$

7.1.3 Coherence

The Coherence Condition requires that all argument functions in a given f-structure have to be selected in the value of the local PRED attribute. (13a) is ruled out because COMP does not appear under the arguments of *devour*.

- (13) a. * David devoured a sandwich that Peter sleeps.
- b.
$$\left[\begin{array}{l} \text{PRED} \text{ 'DEVOUR(SUBJ,OBJ)'} \\ \text{SUBJ} \left[\begin{array}{l} \text{PRED} \text{ 'DAVID'} \end{array} \right] \\ \text{OBJ} \left[\begin{array}{l} \text{SPEC} \text{ A} \\ \text{PRED} \text{ 'SANDWICH'} \end{array} \right] \\ \text{COMP} \left[\begin{array}{l} \text{PRED} \text{ 'SLEEP(SUBJ)'} \\ \text{SUBJ} \left[\begin{array}{l} \text{PRED} \text{ 'PETER'} \end{array} \right] \end{array} \right] \end{array} \right]$$

The constraints on completeness and coherence together ensure that all and only those arguments required in the PRED specification are actually realized. Both of those constraints taken together correspond to the Theta-Criterion in GB theory (see page 85).⁴

⁴ For the differences between predicate-argument structures in LFG and the Deep Structure oriented Theta

7.1.4 Restrictions on the c-structure/f-structure relation

Symbols in c-structures are assigned restrictions for f-structures. The following symbols are used: ‘ \uparrow ’ refers to the f-structure of the immediately dominating node and ‘ \downarrow ’ refers to the f-structure of the c-structure node bearing the annotation. A common annotation is ‘ $\uparrow = \downarrow$ ’. This constraint states that the f-structure of the mother node is identical to that of the annotated category:

$$(14) \quad V' \rightarrow \begin{array}{c} V \\ \uparrow = \downarrow \\ \text{f-structure of the mother} = \text{own f-structure} \end{array}$$

The annotation ‘ $\uparrow = \downarrow$ ’ is below the head of a structure.

Phrases which are licensed by the annotated c-structure in (14) can be visualized as follows:

$$(15) \quad \begin{array}{c} V' \\ | \\ V \end{array} \rightarrow []$$

(16) shows a V' rule with an object:

$$(16) \quad V' \rightarrow \begin{array}{cc} V & NP \\ \uparrow = \downarrow & (\uparrow \text{OBJ}) = \downarrow \end{array}$$

The annotation on the NP signals that the OBJ value in the f-structure of the mother ($\uparrow \text{OBJ}$) is identical to the f-structure of the NP node, that is, to everything that is contributed from the material below the NP node (\downarrow). This is shown in the figure in (17):

$$(17) \quad \begin{array}{c} V' \\ / \quad \backslash \\ V \quad NP \end{array} \rightarrow [\text{OBJ} \rightarrow []]$$

In the equation $(\uparrow \text{OBJ}) = \downarrow$, the arrows ‘ \uparrow ’ and ‘ \downarrow ’ correspond to feature structures. ‘ \uparrow ’ and ‘ \downarrow ’ stand for the f and g in equations such as (6).

(18) is an example with an intransitive verb and (19) is the corresponding visualization:

$$(18) \quad \text{sneezed} \quad V \quad \begin{array}{l} (\uparrow \text{PRED}) = \text{'SNEEZE(SUBJ)'} \\ (\uparrow \text{TENSE}) = \text{PAST} \end{array}$$

$$(19) \quad \begin{array}{c} V \\ | \\ \text{sneezed} \end{array} \rightarrow \begin{bmatrix} \text{PRED} & \text{'SNEEZE(SUBJ)'} \\ \text{TENSE} & \text{PAST} \end{bmatrix}$$

Criterion, see Bresnan & Kaplan (1982: xxvi–xxviii).

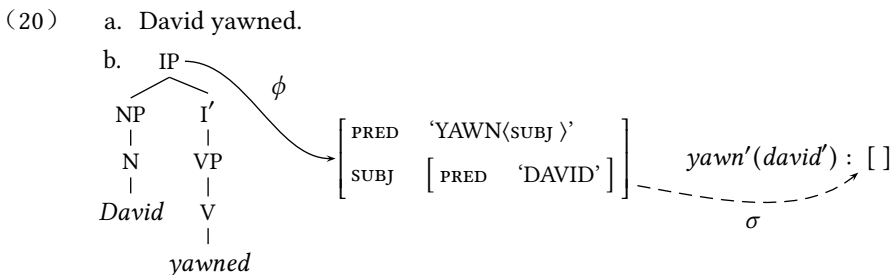
7.1.5 Semantics

Following Dalrymple (2006: 90–92), *glue semantics* is the dominant approach to semantic interpretation in LFG (Dalrymple, Lamping & Saraswat 1993; Dalrymple 2001: Chapter 8). There are, however, other variants where Kamp’s discourse representation structures (Kamp & Reyle 1993) are used (Frey & Reyle 1983a,b).

In the following, glue semantics will be presented in more detail.⁵ Under a glue-based approach, it is assumed that f-structure is the level of syntactic representation which is crucial for the semantic interpretation of a phrase, that is, unlike GB theory, it is not the position of arguments in the tree which play a role in the composition of meaning, but rather functional relations such as SUBJ and OBJ. Glue semantics assumes that each substructure of the f-structure corresponds to a semantic resource connected to a meaning and furthermore, that the meaning of a given f-structure comes from the sum of these parts. The way the meaning is assembled is regulated by certain instructions for the combination of semantic resources. These instructions are given as a set of logic premises written in linear logic as *glue language*. The computation of the meaning of an utterance corresponds to a logical conclusion.

This conclusion is reached on the basis of logical premises contributed by the words in an expression or possibly even by a syntactic construction itself. The requirements on how the meaning of the parts can be combined to yield the full meaning are expressed in linear logic, a resource-based logic. Linear logic is different from classic logic in that it does not allow that premises of conclusions are not used at all or more than once in a derivation. Hence, in linear logic, premises are resources which have to be used. This corresponds directly to the use of words in an expression: words contribute to the entire meaning exactly once. It is not possible to ignore them or to use their meaning more than once. A sentence such as *Peter knocked twice*, does not mean the same as *Peter knocked*. The meaning of *twice* must be included in the full meaning of the sentences. Similarly, the sentence cannot mean the same as *Peter knocked twice twice*, since the semantic contribution of a given word cannot be used twice.

The syntactic structure for the sentence in (20a) together with its semantic representation is given in (20b):



The semantic structure of this sentence is connected to the f-structure via the correspon-

⁵ The following discussion heavily draws from the corresponding section of Dalrymple (2006). (It is a translation of my translation of the original material into German.)

dence function σ (depicted here as a dashed line). The semantic representation is derived from the lexical information for the verb *yawned*, which is given in (21).

$$(21) \quad \lambda x. \text{yawn}'(x) : (\uparrow \text{SUBJ})_\sigma \multimap \uparrow_\sigma$$

This formula is referred to as the *meaning constructor*. Its job is to combine the meaning of *yawned* – a one place predicate $\lambda x. \text{yawn}'(x)$ – with the formula $(\uparrow \text{SUBJ})_\sigma \multimap \uparrow_\sigma$ in linear logic. Here, the connective \multimap is the *linear implication* symbol of linear logic. The symbol contains the meaning that *if* a semantic resource $(\uparrow \text{SUBJ})_\sigma$ for the meaning of the subject is available, *then* a semantic resource for \uparrow_σ must be created which will stand for the entire meaning of the sentence. Unlike the implication operator of classic logic, the linear implication must consume and produce semantic resources: the formula $(\uparrow \text{SUBJ})_\sigma \multimap \uparrow_\sigma$ states that if a semantic resource $(\uparrow \text{SUBJ})_\sigma$ is found, it is consumed and the semantic resource \uparrow_σ is produced.

Furthermore, it is assumed that a proper name such as *David* contributes its own semantic structure as a semantic resource. In an utterance such as *David yawned*, this resource is consumed by the verb *yawned*, which requires a resource for its SUBJ in order to produce the resource for the entire sentence. This corresponds to the intuition that a verb in any given sentence requires the meaning of its arguments in order for the entire sentence to be understood.

The f-structure of *David yawned* with the instantiated meaning construction contributed by *David* and *yawned* is given in (22):

$$(22) \quad y : \left[\begin{array}{cc} \text{PRED} & \text{'YAWN(SUBJ)'} \\ \text{SUBJ} & d : \left[\begin{array}{cc} \text{PRED} & \text{'DAVID'} \end{array} \right] \end{array} \right]$$

$$[\text{David}] \quad \text{david}' : d_\sigma$$

$$[\text{yawn}] \quad \lambda x. \text{yawn}'(x) : d_\sigma \multimap y_\sigma$$

The left side of the meaning constructor marked by **[David]** is the meaning of the proper name *David*, *david'* to be precise. The left-hand side of the meaning constructor **[yawn]** is the meaning of the intransitive verb – a one-place predicate $\lambda x. \text{yawn}'(x)$.

Furthermore, one must still postulate further rules to determine the exact relation between the right-hand side (the glue) of the meaning constructors in (22) and the left-hand side (the meaning). For simple, non-implicational meaning constructors such as **[David]** in (22), the meaning on the left is the same as the meaning of the semantic structure on the right. Meaning constructors such as **[yawn]** have a λ -expression on the left, which has to be combined with another expression via functional application (see Section 2.3). The linear implication on the right-hand side must be applied in parallel. This combined process is shown in (23).

$$(23) \quad \frac{x : f_\sigma \quad P : f_\sigma \multimap g_\sigma}{P(x) : g_\sigma}$$

The right-hand side of the rule corresponds to a logical conclusion following the *modus ponens* rule. With these correspondences between expressions in linear logic and the meanings themselves, we can proceed as shown in (24), which is based on Dalrymple (2006: 92). After combining the respective meanings of *yawned* and *David* and then

(24)	$david' : d_\sigma$	The meaning <i>david'</i> is associated with the semantic structure of SUBJ d_σ .
	$\lambda x.yawn'(x) : d_\sigma \multimap y_\sigma$	If we find the semantic resource for the SUBJ d_σ on the glue side, this resource is consumed and the semantic resource for the entire sentence y_σ is produced. On the meaning side, we apply the function $\lambda x.yawn'(x)$ to the meaning associated with d_σ .
<hr/>		
	$yawn'(david') : y_\sigma$	We have created the semantic structure y_σ for the entire sentence, associated with the meaning of $yawn'(david')$.

carrying out β -reduction, we arrive at the desired result of $yawn'(david')$ as the meaning of *David yawned*.

Glue analyses of quantification, modification and other phenomena have been investigated in a volume on glue semantics (Dalrymple 1999). Particularly problematic for these approaches are cases where there appear to be too many or too few resources for the production of utterances. These kinds of cases have been discussed by Asudeh (2004).

7.1.6 Adjuncts

Adjuncts are not selected by their head. The grammatical function ADJ is a non-governable grammatical function. Unlike arguments, where every grammatical function can only be realized once, a sentence can contain multiple adjuncts. The value of ADJ in the f-structure is therefore not a simple structure as with the other grammatical functions, but rather a set. For example, the f structure for the sentence in (25a) contains an ADJ set with two elements: one for *yesterday* and one for *at noon*.

- (25) a. David devoured a sandwich at noon yesterday.

b.	PRED	'DEVOUR(SUBJ,OBJ)'	
	SUBJ	$\left[\begin{array}{c} \text{PRED} \\ \text{'DAVID'} \end{array} \right]$	
	OBJ	$\left[\begin{array}{c} \text{SPEC} \quad \text{A} \\ \text{PRED} \quad \text{'SANDWICH'} \end{array} \right]$	
	ADJ	$\left\{ \left[\begin{array}{c} \text{PRED} \\ \text{'YESTERDAY'} \end{array} \right], \left[\begin{array}{c} \text{PRED} \quad \text{'AT(OBJ)'} \\ \text{OBJ} \quad \left[\begin{array}{c} \text{PRED} \\ \text{'NOON'} \end{array} \right] \end{array} \right] \right\}$	

The annotation on the c-structure rule for adjuncts requires that the f-structure of the adjuncts be part of the ADJ set of the mother's f-structure:

$$(26) \quad V' \rightarrow V' \quad PP \\ \uparrow = \downarrow \quad \downarrow \in (\uparrow \text{ ADJ})$$

The representation of adjuncts in a set is not sufficient to characterize the meaning of an utterance containing scope-bearing adjuncts (as for instance the negation in sentences like (31) on page 95). In order to determine scopal relations, one has to refer to the linear order of the adjuncts, that is, the c-structure. For linearization restrictions in LFG, see Zaenen & Kaplan (1995).

7.2 Passive

Bresnan & Mchombo (1995) argue that one should view words as “atoms” of which syntactic structure is comprised (*lexical integrity*⁶).

Syntactic rules cannot create new words or make reference to the internal structure of words. Every terminal node (each “leaf” of the tree) is a word. It follows from this that analyses such as the GB analysis of Pollock (1989) in Figure 7.1 在下一页 for the French example in (27) are ruled out (the figure is taken from Kuhn 2007: 617):

- (27) Marie ne parlerait pas
 Marie NEG speak.COND.3SG NEG
 ‘Marie would not speak.’

In Pollock's analysis, the various morphemes are in specific positions in the tree and are combined only after certain movements have been carried out.

The assumption of lexical integrity is made by all theories discussed in this book with the exception of GB and Minimalism. However, formally, this is not a must as it is also possible to connect morphemes to complex syntactic structures in theories such as Categorical Grammar, GPSG, HPSG, CxG, DG and TAG. As far as I know, this kind of analysis has never been proposed.

Bresnan noticed that, as well as passivized verbs, there are passivized adjectives which show the same morphological idiosyncrasies as the corresponding participles (Bresnan 1982c: 21; Bresnan 2001: 31). Some examples are given in (28):

- (28) a. a well-written novel (write – written)
 b. a recently given talk (give – given)
 c. my broken heart (break – broken)
 d. an uninhabited island (inhabit – inhabited)
 e. split wood (split – split)

⁶ See Anderson (1992: 84) for more on lexical integrity.

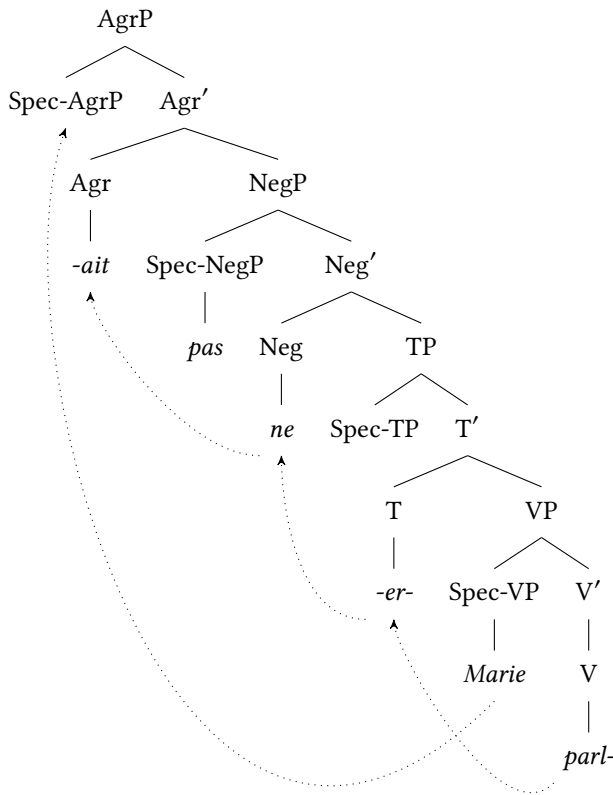


图 7.1: Pollock's analysis of *Marie ne parlerait pas* 'Marie would not speak.' according to Kuhn (2007: 617)

If one assumes lexical integrity, then adjectives would have to be derived in the lexicon. If the verbal passive were not a lexical process, but rather a phrase-structural one, then the form identity would remain unexplained.

In LFG, grammatical functions are primitives, that is, they are not derived from a position in the tree (e.g., Subject = SpecIP). Words (fully inflected word-forms) determine the grammatical function of their arguments. Furthermore, there is a hierarchy of grammatical functions. During participle formation in morphology, the highest verbal argument is suppressed. The next highest argument moves up and is not realized as the OBJECT but rather as the SUBJECT. This was explicitly encoded in earlier work (Bresnan 1982c: 8):

- (29) Passivization rule:
 (SUBJ) $\mapsto \emptyset / (\text{OBL})$
 (OBJ) $\mapsto (\text{SUBJ})$

The first rule states that the subject is either not realized (\emptyset) or it is realized as an oblique

element (the *by*-PP in English). The second rule states that if there is an accusative object, this becomes the subject.

In later work, the assignment of grammatical functions was taken over by Lexical Mapping Theory (Bresnan & Kanerva 1989). It is assumed that thematic roles are ordered in a universally valid hierarchy (Bresnan & Kanerva 1989; Bresnan 2001: 307): agent > beneficiary > experiencer/goal > instrument > patient/theme > locative. Patient-like roles are marked as unrestricted ($[-r]$) in a corresponding representation, the so-called a-structure. Secondary patient-like roles are marked as *objective* ($[+o]$) and all other roles are marked as non-objective ($[-o]$). For the transitive verb *schlagen* ‘to beat’, we have the following:

(30)		Agent	Patient
	a-structure	<i>schlagen</i> ‘beat’	$\langle x \quad y \rangle$
		$[-o]$	$[-r]$

The mapping of a-structure to f-structure is governed by the following restrictions:

- (31)
- Subject-Mapping-Principle: The most prominent role marked with $[-o]$ is mapped to SUBJ if it is initial in the a-structure. Otherwise, the role marked with $[-r]$ is mapped to SUBJ.
 - The argument roles are connected to grammatical functions as shown in the following table. Non-specified values for o and r are to be understood as ‘+’:
- | | | |
|--------|--------|------------------------------------|
| | $[-r]$ | $[+r]$ |
| $[-o]$ | SUBJ | OBL _{θ} |
| $[+o]$ | OBJ | OBJ _{θ} |
- Function-Argument Biuniqueness: Every a-structure role must be associated to exactly one function and vice versa.

For the argument structure in (30), the principle in (31a) ensures that the agent *x* receives the grammatical function SUBJ. (31b) adds an o-feature with the value ‘+’ so that the patient *y* is associated with OBJ:

(32)		Agent	Patient
	a-structure	<i>schlagen</i> ‘beat’	$\langle x \quad y \rangle$
		$[-o]$	$[-r]$
		SUBJ	OBJ

Under passivization, the most prominent role is suppressed so that only the $[-r]$ marked patient role remains. Following (31a), this role will then be mapped to the subject.

(33)		Agent	Patient
	a-structure	<i>schlagen</i> ‘beat’	$\langle x \quad y \rangle$
		$[-o]$	$[-r]$
		\emptyset	SUBJ

Unlike the objects of transitive verbs, the objects of verbs such as *helfen* ‘help’ are marked as $[+o]$ (Berman 1999). The lexical case of the objects is given in the a-structure, since

this case (dative) is linked to a semantic role (Zaenen, Maling & Thráinsson 1985: 465). The corresponding semantic roles are obligatorily mapped to the grammatical function OBJ_θ .

(34)			Agent	Beneficiary
	a-structure	<i>helfen</i> ‘help’	$\langle x \quad y \rangle$	
			$[-o]$	$[+o]/\text{DAT}$
			SUBJ	OBJ_θ

Passivization will yield the following:

(35)			Agent	Beneficiary
	a-structure	<i>helfen</i> ‘help’	$\langle x \quad y \rangle$	
			$[-o]$	$[+o]/\text{DAT}$
			\emptyset	OBJ_θ

Since there is neither a $[-o]$ nor a $[-r]$ argument, no argument is connected to the subject function. The result is an association of arguments and grammatical functions that corresponds to the one found in impersonal passives.

These mapping principles may seem complex at first glance, but they play a role in analyzing an entire range of phenomena, e.g., the analysis of unaccusative verbs (Bresnan & Zaenen 1990). For the analysis of the passive, we can now say that the passive suppresses the highest $[-o]$ role. Mentioning an eventual object in the passive rule is no longer necessary.

7.3 Verb position

There are two possibilities for the analysis of verb placement in German.

- a trace in verb-final position (as in GB) (see Choi 1999, Berman 1996: Section 2.1.4) and
- so-called *extended head domains* (see Berman 2003a).

In the analysis of extended head domains, the verb is simply omitted from the verb phrase. The following preliminary variant of the VP rule is used:⁷

$$(36) \quad \text{VP} \rightarrow (\text{NP}) (\text{NP}) (\text{NP}) (\text{V})$$

All components of the VP are optional as indicated by the brackets. As in GB analyses, the verb in verb-first clauses is in C. No I projection is assumed – as in a number of GB works (Haider 1993, 1995, 1997a; Sternefeld 2006: Section IV.3), since it is difficult to motivate its existence for German (Berman 2003a: Section 3.2.2). The verb contributes its f-structure information from the C position. Figure 7.2 on the following page contains a simplified version of the analysis proposed by Berman (2003a: 41).

⁷ See Bresnan (2001: 110) and Dalrymple (2006: Section 2.2) for a corresponding rule with optional constituents on the right-hand side of the rule.

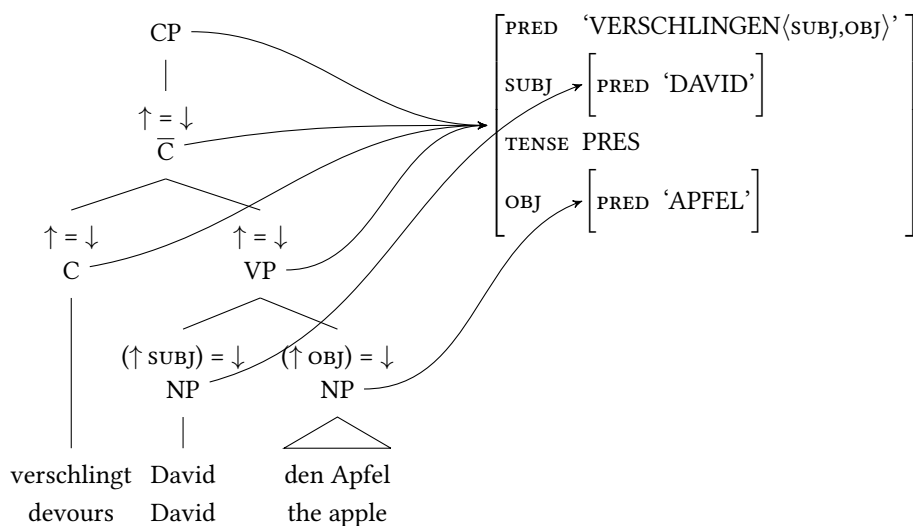


图 7.2: Analysis of verb placement following Berman (2003a: 41)

After what we learned about phrase structure rules in Chapters 2 and 5, it may seem strange to allow VPs without V. This is not a problem in LFG, however, since for the analysis of a given sentence, it only has to be ensured that all the necessary parts (and only these) are present. This is ensured by the constraints on completeness and coherence. Where exactly the information comes from is not important. In Figure 7.2, the verb information does not come from the VP, but rather from the C node. C' is licensed by a special rule:

$$(37) \quad C' \rightarrow \begin{array}{cc} C & VP \\ \uparrow = \downarrow & \uparrow = \downarrow \end{array}$$

In LFG rules, there is normally only one element annotated with ' $\uparrow = \downarrow$ ', namely the head. In (37), there are two such elements, which is why both equally contribute to the f-structure of the mother. The head domain of V has been extended to C. The information about SUBJ and OBJ comes from the VP and the information about PRED from C.

7.4 Local reordering

Two possibilities for treating local reordering have been discussed in the literature:⁸

- movement of arguments from a base configuration as in GB (see Choi 1999)
- direct licensing by phrase structure rules (see Berman 1996: Section 2.1.3.1; 2003a)

⁸ Kaplan (1995: 20–21) shows how one can write grammars in the ID/LP format in LFG. A GPSG-like analysis of German constituent order has not been proposed in the LFG framework.

If one assumes that traces are relevant for the semantic interpretation of a given structure, then the first option has the same problems as movement-based GB analyses. These have already been discussed in Section 3.5.

In what follows, I will present the analysis proposed by Berman (1996: Section 2.1.3) in a somewhat simplified form. Case and grammatical functions of verbal arguments are determined in the lexicon (Berman 1996: 22). (38) shows the lexical entry for the verb *verschlingen* ‘devour’:^{9,10}

- (38) *verschlingt* V (↑ PRED) = ‘VERSCHLINGEN(SUBJ, OBJ)’
 (↑ SUBJ AGR CAS) = NOM
 (↑ OBJ AGR CAS) = ACC
 (↑ TENSE) = PRES

Berman proposes an analysis that does not combine the verb with all its arguments and adjuncts at the same time, as was the case in GPSG. Instead, she chooses the other extreme and assumes that the verb is not combined with an adjunct or an argument, but

⁹ The four cases in German can be represented using two binary features (GOV, OBL) (Berman 1996: 22). Nominative corresponds to GOV– and OBL– and accusative to GOV+ and OBL–. This kind of encoding allows one to leave case partially underspecified. If one does not provide a value for GOV, then an element with OBL– is compatible with both nominative and accusative. Since this underspecification is not needed in the following discussion, I will omit this feature decomposition and insert the case values directly.

¹⁰ Alternative analyses derive the grammatical function of an NP from its case (Berman 2003a: 37 for German; Bresnan 2001: 187, 201 for German and Russian).

- (i) (↓ CASE) = ACC ⇒ (↑ OBJ) = ↓

Karttunen (1989: Section 2.1) makes a similar suggestion for Finnish in the framework of Categorical Grammar. Such analyses are not entirely unproblematic as case cannot always be reliably paired with grammatical functions. In German, as well as temporal accusatives (ii.a), there are also verbs with two accusative objects (ii.b–c) and predicative accusatives (ii.d).

- (ii) a. Er arbeitete den ganzen Tag.
 he worked the.ACC whole.ACC day
 b. Er lehrte ihn den Ententanz.
 he taught him.ACC the.ACC duck.dance
 c. Das kostet ihn einen Taler.
 that costs him.ACC a.ACC taler
 d. Sie nannte ihn einen Lügner.
 she called him.ACC a.ACC liar

All of these accusatives can occur in long-distance dependencies (see Section 7.5):

- (iii) Wen glaubst du, dass ich getroffen habe.
 who believe you that I met have
 ‘Who do you think I met?’

wen is not the object of *glauben* ‘believe’ and as such cannot be included in the f-structure of *glauben* ‘believe’. One would have to reformulate the implication in (i) as a disjunction of all possible grammatical functions of the accusative and in addition account for the fact that accusatives can come from a more deeply embedded f-structure.

rather forms a VP directly. The rule for this is shown in (39):

(39) $VP \rightarrow (V)$
 $\uparrow = \downarrow$

At first sight, this may seem odd since a V such as *verschlingen* ‘devour’ does not have the same distribution as a verb with its arguments. However, one should recall that the constraints pertaining to coherence and completeness of f-structures play an important role so that the theory does not make incorrect predictions.

Since the verb can occur in initial position, it is marked as optional in the rule in (39) (see Section 7.3).

The following rule can be used additionally to combine the verb with its subject or object.

(40) $VP \rightarrow \begin{matrix} NP & VP \\ (\uparrow \text{SUBJ} \mid \text{OBJ} \mid \text{OBJ}_\theta) = \downarrow & \uparrow = \downarrow \end{matrix}$

The ‘|’ here stands for a disjunction, that is, the NP can either be the subject or the object of the superordinate f-structure. Since VP occurs both on the left and right-hand side of the rule in (40), it can be applied multiple times. The rule is not complete, however. For instance, one has to account for prepositional objects, for clausal arguments, for adjectival arguments and for adjuncts. See footnote 12 on page 216.

Figure 7.3 shows the analysis for (41a).

- (41) a. [dass] David den Apfel verschlingt
 that David the apple devours
 ‘that David is devouring the apple’
 b. [dass] den Apfel David verschlingt
 that the apple David devours

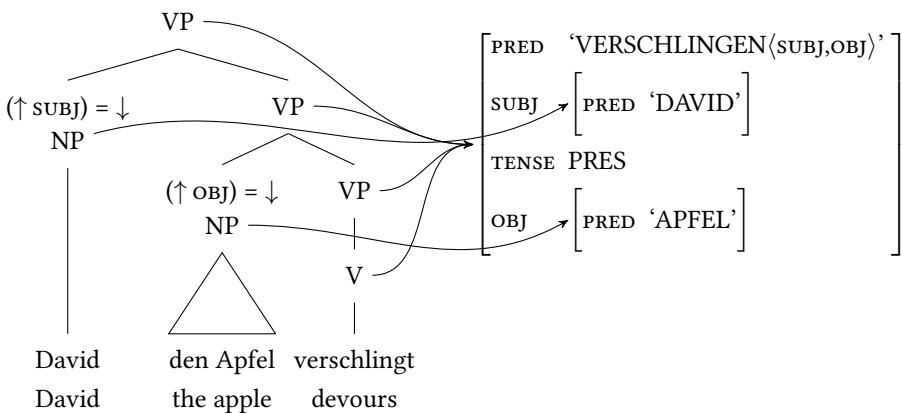


图 7.3: Analysis of SOV order following Berman (1996)

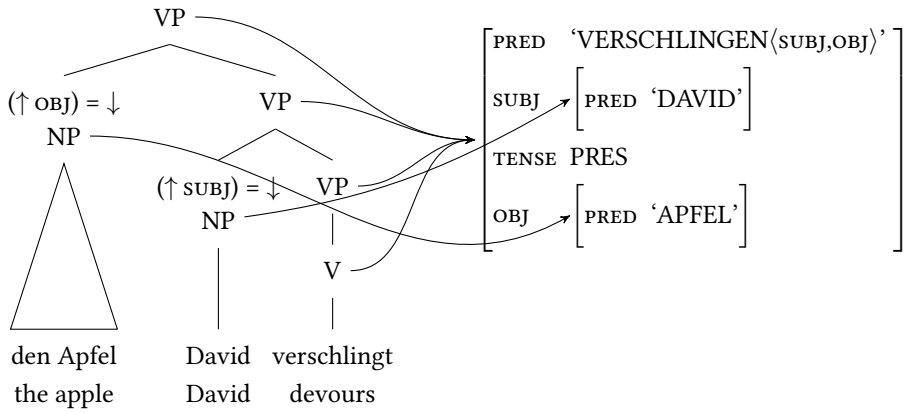


图 7.4: Analysis of OSV order following Berman (1996)

The analysis of (41b) is shown in Figure 7.4. The analysis of (41b) differs from the one of (41a) only in the order of the replacement of the NP node by the subject or object.

One further fact must be discussed: in the rule (39), the verb is optional. If it is omitted, the VP is empty. In this way, the VP rule in (40) can have an empty VP on the right-hand side of the rule. This VP is also simply omitted even though the VP symbol in the right-hand side of rule (40) is not marked as optional. That is, the corresponding symbol then also becomes optional as a result of taking the rest of the grammar into consideration as well as possible interactions with other rules.

7.5 Long-distance dependencies and functional uncertainty

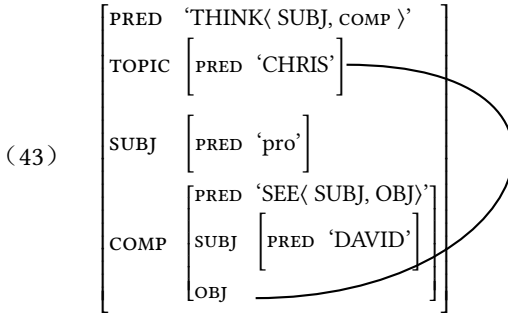
We have seen that LFG can explain phenomena such as passivization, local reordering as well as verb placement without transformations. In Chapter 5 on GPSG, we already saw that the development of a transformation-less analysis for long-distance dependencies constitutes a real achievement. In LFG, Kaplan & Zaenen (1989) proposed another transformation-less analysis of long-distance dependencies, which we will consider in further detail in what follows.

In example (42), the displaced constituent *Chris* is characterized by two functions:

(42) Chris, we think that David saw.

For one, it has an argument function which is normally realized in a different position (the OBJ function of *saw* in the above example) and additionally it has a discourse function: a certain emphasis of the information-structural status in this construction (TOPIC in the matrix clause). In LFG, TOPIC and FOCUS are assumed to be grammaticalized discourse functions (furthermore, SUBJ is classified as the default discourse function). Only grammaticalized discourse functions are represented on the level of f-structure, that is, those that are created by a fixed syntactic mechanism and that interact with the rest of the syntax.

Unlike argument functions, the discourse functions TOPIC and FOCUS are not lexically subcategorized and are therefore not subject to the completeness and coherence conditions. The values of discourse function features like TOPIC and FOCUS are identified with an f-structure that bears an argument function. (43) gives the f-structure for the sentence in (42):

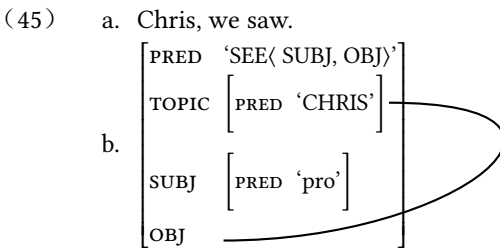


The connecting line means that the value of TOPIC is identical to the value of COMP|OBJ. In Chapter 6 on feature descriptions, I used boxes for structure sharing rather than connecting lines, since boxes are more common across frameworks. It is possible to formulate the structure sharing in (43) as an f-structure constraint as in (44):

$$(44) \quad (\uparrow \text{TOPIC}) = (\uparrow \text{COMP OBJ})$$

Fronting operations such as (42) are possible from various levels of embedding: for instance, (45a) shows an example with less embedding. The object is located in the same f-structure as the topic. However, the object in (42) comes from a clause embedded under *think*.

The f-structure corresponding to (45a) is given in (45b):

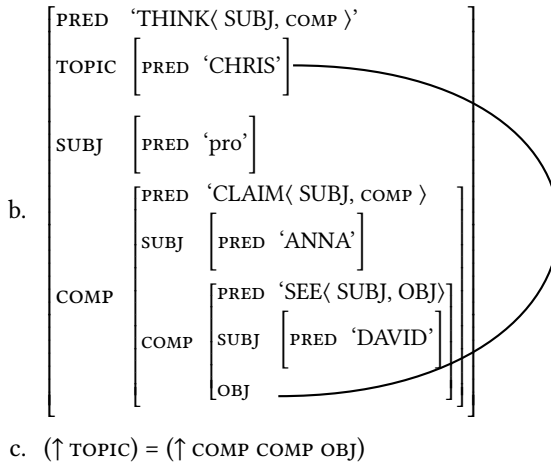


The identity restriction for TOPIC and object can be formulated in this case as in (46):

$$(46) \quad (\uparrow \text{TOPIC}) = (\uparrow \text{OBJ})$$

Example (47a) shows a case of even deeper embedding than in (42) and (47b,c) show the corresponding f-structure and the respective restriction.

(47) a. Chris, we think Anna claims that David saw.



The restrictions in (44), (46) and (47c) are c-structure constraints. The combination of a c-structure with (44) is given in (48):

$$\begin{array}{lcl}
 (48) & CP & \rightarrow \quad \quad \quad XP \quad \quad \quad C' \\
 & & (\uparrow \text{TOPIC}) = \downarrow \quad \quad \quad \uparrow = \downarrow \\
 & & (\uparrow \text{TOPIC}) = (\uparrow \text{COMP OBJ})
 \end{array}$$

(48) states that the first constituent contributes to the TOPIC value in the f-structure of the mother and furthermore that this topic value has to be identical to that of the object in the complement clause. We have also seen examples of other embeddings of various depths. We therefore need restrictions of the following kind as in (49):

- (49)
- a. $(\uparrow \text{TOPIC}) = (\uparrow \text{OBJ})$
 - b. $(\uparrow \text{TOPIC}) = (\uparrow \text{COMP OBJ})$
 - c. $(\uparrow \text{TOPIC}) = (\uparrow \text{COMP COMP OBJ})$
 - d. ...

The generalization emerging from these equations is given in (50):

$$(50) \quad (\uparrow \text{TOPIC}) = (\uparrow \text{COMP}^* \text{OBJ})$$

Here, ‘*’ stands for an unrestricted number of occurrences of COMP. This means of leaving the possible identification of discourse and grammatical function open is known as *functional uncertainty*, see Kaplan & Zaenen (1989).

As was shown in the discussion of examples (2) and (3) on page 199, it is not the case that only a TOPIC can be placed in the specifier position of CP in English as FOCUS can occur there too. One can use disjunctions in LFG equations and express the corresponding condition as follows:

$$(51) \quad (\uparrow \text{TOPIC} | \text{FOCUS}) = (\uparrow \text{COMP}^* \text{OBJ})$$

One can introduce a special symbol for TOPIC|FOCUS, which stands for a disjunction of discourse functions: DF. (51) can then be abbreviated as in (52):

$$(52) \quad (\uparrow \text{DF}) = (\uparrow \text{COMP}^* \text{OBJ})$$

The final version of the c-structure rule for fronting in English will therefore have the form of (53):¹¹

$$(53) \quad \text{CP} \rightarrow \begin{array}{cc} \text{XP} & \text{C}' \\ (\uparrow \text{DF}) = \downarrow & \uparrow = \downarrow \\ (\uparrow \text{DF}) = (\uparrow \text{COMP}^* \text{OBJ}) & \end{array}$$

In German, as well as objects, nearly any other constituent (e.g., subjects, sentential complements, adjuncts) can be fronted. The c-structure rule for this is shown in (54):¹²

$$(54) \quad \text{CP} \rightarrow \begin{array}{cc} \text{XP} & \text{C}' \\ (\uparrow \text{DF}) = \downarrow & \uparrow = \downarrow \\ (\uparrow \text{DF}) = (\uparrow \text{COMP}^* \text{GF}) & \end{array}$$

Here, GF is an abbreviation for a disjunction of grammatical functions which can occur in the prefield.

7.6 Summary and classification

LFG is a constraint-based theory and utilizes feature descriptions and PSG rules. Grammatical functions are treated as primitives of the theory, which sets LFG apart from most of the other theories covered in this book. They are not defined structurally (as in GB). LFG is a lexicalist theory. Like GPSG, LFG can do without transformations. Processes affecting argument structure such as passivization are analyzed by means of lexical rules. Whereas GPSG treated long-distance dependencies using the percolation of information in trees, LFG uses functional uncertainty: a part of the f-structure is identified with another f-structure that can be embedded to an arbitrary depth. Coherence and completeness ensure that the long-distance dependency can be correctly resolved, that is, it ensures that a fronted object is not assigned to an f-structure which already contains an object or one in which no object may occur.

While LFG does contain a phrase-structural component, this plays a significantly less important role compared to other models of grammar. There are rules in which all constituents are optional and it has even been proposed for some languages that there

¹¹ Note that the two disjunctions that are abbreviated by the respective occurrences of DF are independent in principle. This is unwanted. We want to talk about either a topic or a focus not about a topic and a focus in the mother f-structure. So additional machinery is needed to ensure that both occurrences of DF refer to the same discourse function.

¹² Berman (1996) uses the symbol ZP for symbols in the prefield rather than XP in (54). She formulates various phrase structure rules for ZPs, which replace ZP with NP, PP, AP and various adjuncts. Following Berman, ZPs can also be combined with the verb in the middle field. For reasons of exposition, I refrained from using ZP symbols in the formulation of the VP rule (40) in Section 7.4 and instead used NP directly.

are rules where the part of speech of the constituents is not specified (see Section 13.1.2). In these kinds of grammars, f-structure, coherence and completeness work together to ensure that the grammar only allows well-formed structures.

LFG differs from other theories such as HPSG and variants of Construction Grammar in that feature structures are untyped. Generalizations can therefore not be represented in type hierarchies. Until a few years ago, the hierarchical organization of knowledge in inheritance hierarchies did not form part of theoretical analyses. In computer implementations, there were macros but these were viewed as abbreviations without any theoretical status. It is possible to organize macros into hierarchies and macros were discussed explicitly in Dalrymple, Kaplan & King (2004) with reference to capturing linguistic generalizations. Asudeh, Dalrymple & Toivonen (2008) suggest using macros not only for the organization of lexical items but also for capturing generalizations regarding c-structure annotations. Because of these developments, there was a greater convergence between LFG and other theories such as HPSG and CxG.

Williams (1984) compares analyses in LFG with GB. He shows that many analyses are in fact transferable: the function that f-structure has in LFG is handled by the Theta-Criterion and Case Theory in GB. LFG can explicitly differentiate between subjects and non-subjects. In GB, on the other hand, a clear distinction is made between external and internal arguments (see Williams 1984: Section 1.2). In some variants of GB, as well as in HPSG and CxG, the argument with subject properties (if there is one) is marked explicitly (Haider 1986a; Heinz & Matiassek 1994; Müller 2003b; Michaelis & Ruppenhofer 2001). This special argument is referred to as the *designated argument*. In infinitival constructions, subjects are often not expressed inside the infinitival phrase. Nevertheless, the unexpressed subject is usually coreferential with an argument of the matrix verb:

- (55) a. Er versucht, [das Buch zu lesen].
 he tries the book to read
 ‘He is trying to read the book.’
 b. Er zwingt ihn, [das Buch zu lesen].
 he forces him the book to read
 ‘He is forcing him to read the book.’

This is a fact that every theory needs to be able to capture, that is, every theory must be able to differentiate between subjects and non-subjects.

For a comparison of GB/Minimalism and LFG/HPSG, see Kuhn (2007).

Comprehension questions

1. What do the terms *coherence* and *completeness* mean?
2. What are extended head domains?
3. What does lexical integrity mean?

Exercises

1. Give the lexical entry for *kannte* ‘knew’.
2. How could one analyze the following sentence?

(56) Den Apfel verschlingt David.
the apple devours David
‘David devours the apple.’

Provide the necessary c-structure rules. What kind of f-structure is licensed? Draw a syntactic tree with corresponding references to the f-structure. For fronted constituents, simply write NP rather than expanding the XP node. The c-structure rule for the NP can also be omitted and a triangle can be drawn in the tree.

Further reading

Section 7.1 was based extensively on the textbook and introductory article of Dalrymple (2001, 2006). Additionally, I have drawn from teaching materials of Jonas Kuhn from 2007. Bresnan (2001) is a comprehensive textbook in English for the advanced reader. Some of the more in-depth analyses of German in LFG are Berman (1996, 2003a). Schwarze & de Alencar (2016) is an introduction to LFG that uses French examples. The authors demonstrate how the XLE system can be used for the development of a French LFG grammar. The textbook also discusses the Finite State Morphology component that comes with the XLE system.

Levelt (1989) developed a model of language production based on LFG. Pinker (1984) – one of the best-known researchers on language acquisition – used LFG as the model for his theory of acquisition. For another theory on first and second language acquisition that uses LFG, see Pienemann (2005).

第八章 **Categorial Grammar**

Categorial Grammar is the second oldest of the approaches discussed in this book. It was developed in the 30s by the Polish logician Kazimierz Ajdukiewicz (Ajdukiewicz 1935). Since syntactic and semantic descriptions are tightly connected and all syntactic combinations correspond to semantic ones, Categorial Grammar is popular amongst logicians and semanticists. Some stellar works in the field of semantics making use of Categorial Grammar are those of Richard Montague (1974). Other important works come from David Dowty in Columbus, Ohio (1979), Michael Moortgat in Utrecht (1989), Glyn Morrill in Barcelona (1994), Bob Carpenter in New York (1998) and Mark Steedman in Edinburgh (1991; 1997; 2000). A large fragment for German using Montague Grammar has been developed by von Stechow (1979). The 2569-page grammar of the *Institut für Deutsche Sprache* in Mannheim (Eroms, Stickel & Zifonun 1997) contains Categorial Grammar analyses in the relevant chapters. Fanselow (1981) worked on morphology in the framework of Montague Grammar. Uszkoreit (1986a), Karttunen (1986, 1989) and Calder, Klein & Zeevat (1988) developed combinations of unification-based approaches and Categorial Grammar.

The basic operations for combining linguistic objects are rather simple and well-understood so that it is no surprise that there are many systems for the development and processing of Categorial Grammars (Yampol & Karttunen 1990; Carpenter 1994; Bouma & van Noord 1994; Lloré 1995; König 1999; Moot 2002; White & Baldridge 2003; Baldridge, Chatterjee, Palmer & Wing 2007; Morrill 2012). An important contribution has been made by Mark Steedman's group (see for instance Clark, Hockenmaier & Steedman 2002; Clark & Curran 2007).

Implemented fragments exist for the following languages:

- German (Uszkoreit 1986a; König 1999; Vierhuff, Hildebrandt & Eikmeyer 2003; Vancoppenolle, Tabbert, Bouma & Stede 2011)
- English (Villavicencio 2002; Baldridge 2002; Beavers 2003, 2004)
- Finish (Karttunen 1989)
- French (Baschung, Bes, Corluy & Guillotin 1987)
- Dutch (Bouma & van Noord 1994; Baldridge 2002)

- Tagalog (Baldridge 2002)
- Turkish (Hoffman 1995; Baldridge 2002)

In addition, Baldridge, Chatterjee, Palmer & Wing (2007: 15) mention an implementation for Classical Arabic.

Some of the systems for the processing of Categorical Grammars have been augmented by probabilistic components so that the processing is robust (Osborne & Briscoe 1997; Clark, Hockenmaier & Steedman 2002). Some systems can derive lexical items from corpora, and Briscoe (2000) and Villavicencio (2002) use statistical information in their UG-based language acquisition models.

8.1 General remarks on the representational format

In what follows I introduce some basic assumptions of Categorical Grammar. After these introductory remarks, I will discuss specific analyses that were developed by Steedman (1997) in the framework of Combinatory Categorical Grammar. There are other variants of Categorical Grammar as for instance type-logical CG, the variety espoused by Morrill (1994), Dowty (1997), Moortgat (2011), and others, which cannot be discussed here.

8.1.1 Representation of valence information

In Categorical Grammar, complex categories replace the SUBCAT feature that is used in GPSG to ensure that a head can only be used with suitable grammatical rules. Simple phrase structure rules can be replaced with complex categories as follows:

(1)	Rule	Category in the lexicon
	$vp \rightarrow v(\text{ditrans})\ np\ np$	$(vp/np)/np$
	$vp \rightarrow v(\text{trans})\ np$	vp/np
	$vp \rightarrow v(np_and_pp)\ np\ pp(\text{to})$	$(vp/pp)/np$

vp/np stands for something that needs an np in order for it to form a vp .

In Categorical Grammar, there are only a few very abstract rules. One of these is forward application, also referred to as the multiplication rule:

- (2) forward application:
 $X/Y * Y = X$

This rule combines an X looking for a Y with a Y and requires that Y occurs to the right of X/Y . The result of this combination is an X that no longer requires a Y . X/Y is called the *functor* and Y is the *argument* of the functor.

As in GB theory, valence is encoded only once in Categorical Grammar, in the lexicon. In GPSG, valence information was present in grammatical rules and in the SUBCAT feature of the lexical entry.

8.1 General remarks on the representational format

$$\frac{\frac{\text{chased}}{vp/np} \quad \frac{\text{Mary}}{np}}{vp} \rightarrow$$

图 8.1: Combination of a verb and its object (preliminary)

Figure 8.1 shows how a lexical entry for a transitive verb is combined with its object. A derivation in CG is basically a binary branching tree; it is, however, mostly represented as follows: an arrow under a pair of categories indicates that these have been combined via a combinatorial rule. The direction of this arrow indicates the direction of this combination. The result is given beneath the arrow. Figure 8.2 shows the tree corresponding to Figure 8.1.

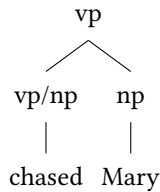


图 8.2: Derivation in Figure 8.1 as a tree diagram

One usually assumes left associativity for ‘/’; that is, $(vp/pp)/np = vp/pp/np$.

If we look at the lexical entries in (1), it becomes apparent that the category v does not appear. The lexicon only determines what the product of combination of a lexical entry with its arguments is. The symbol for vp can also be eliminated: an (English) vp is something that requires an NP to its left in order to form a complete sentence. This can be represented as $s \backslash np$. Using the rule for backward application, it is possible to compute derivations such as the one in Figure 8.3.

(3) Backward application:

$$Y * X \backslash Y = X$$

$$\frac{\frac{\frac{\text{the}}{np/n} \quad \frac{\text{cat}}{n}}{np} \rightarrow \quad \frac{\frac{\frac{\text{chased}}{(s \backslash np)/np} \quad \frac{\text{Mary}}{np}}{s \backslash np} \rightarrow}{s} <$$

图 8.3: Analysis of a sentence with a transitive verb

In Categorical Grammar, there is no explicit difference made between phrases and words: an intransitive verb is described in the same way as a verb phrase with an object: $s \backslash np$. Equally, proper nouns are complete noun phrases, which are assigned the symbol np .

8.1.2 Semantics

As already mentioned, Categorical Grammar is particularly popular among semanticists as syntactic combinations always result in parallel semantic combinations and even for complex combinations such as those we will discuss in more detail in the following sections, there is a precise definition of meaning composition. In the following, we will take a closer look at the representational format discussed in Steedman (1997: Section 2.1.2).

Steedman proposes the following lexical entry for the verb *eats*:¹

$$(4) \quad \text{eats} := (s : \text{eat}'(x, y) \backslash \text{np}_{3S} : x) / \text{np} : y$$

In (4), the meaning of each category is given after the colon. Since nothing is known about the meaning of the arguments in the lexical entry of *eat*, the meaning is represented by the variables x and y . When the verb combines with an NP, the denotation of the NP is inserted. An example is given in (5):²

$$(5) \quad \frac{(s : \text{eat}'(x, y) \backslash \text{np}_{3S} : x) / \text{np} : y \quad \text{np} : \text{apples}'}{s : \text{eat}'(x, \text{apples}') \backslash \text{np}_{3S} : x} \rightarrow$$

When combining a functor with an argument, it must be ensured that the argument fits the functor, that is, it must be unifiable with it (for more on unification see Section 6.6). The unification of $\text{np} : y$ with $\text{np} : \text{apples}'$ results in $\text{np} : \text{apples}'$ since *apples'* is more specific than the variable y . Apart from its occurrence in the term $\text{np} : y$, y occurs in the description of the verb in another position ($s : \text{eat}'(x, y) \backslash \text{np}_{3S} : x$) and therefore also receives the value *apples'* there. Thus, the result of this combination is $s : \text{eat}'(x, \text{apples}') \backslash \text{np}_{3S} : x$ as shown in (5).

Steedman notes that this notation becomes less readable with more complex derivations and instead uses the more standard λ -notation:

$$(6) \quad \text{eats} := (s \backslash \text{np}_{3S}) / \text{np} : \lambda y. \lambda x. \text{eat}'(x, y)$$

Lambdas are used to allow access to open positions in complex semantic representations (see Section 2.3). A semantic representation such as $\lambda y. \lambda x. \text{eat}'(x, y)$ can be combined with the representation of *apples* by removing the first lambda expression and inserting the denotation of *apples* in all the positions where the corresponding variable (in this case, y) appears (see Section 2.3 for more on this point):

$$(7) \quad \begin{array}{l} \lambda y. \lambda x. \text{eat}'(x, y) \text{ apples}' \\ \lambda x. \text{eat}'(x, \text{apples}') \end{array}$$

This removal of lambda expressions is called β -reduction.

If we use the notation in (6), the combinatorial rules must be modified as follows:

¹ I have adapted his notation to correspond to the one used in this book.

² The assumption that *apples* means *apples'* and not *apples'(z)* minus the quantifier contribution is a simplification here.

- (8) $X/Y:f * Y:a = X: f a$
 $Y:a * X/Y:f = X: f a$

In such rules, the semantic contribution of the argument (a) is written after the semantic denotation of the functor (f). The open positions in the denotation of the functor are represented using lambdas. The argument can be combined with the first lambda expression using β -reduction.

Figure 8.4 shows the derivation of a simple sentence with a transitive verb. After forward and backward application, β -reduction is immediately applied.

$$\begin{array}{c}
 \frac{Jacob}{np : jacob'} \quad \frac{eats}{(s \backslash np) / np : \lambda y. \lambda x. eat'(x, y)} \quad \frac{apples}{np : apples'} \\
 \hline
 \frac{s \backslash np : \lambda y. \lambda x. eat'(x, y) \quad apples'}{s \backslash np : \lambda y. \lambda x. eat'(x, y) apples'} > \\
 \hline
 \frac{s : \lambda x. eat'(x, apples') \quad jacob'}{s : \lambda x. eat'(x, apples') jacob'} < \\
 \hline
 = eat'(jacob', apples')
 \end{array}$$

图 8.4: Meaning composition in Categorical Grammar

8.1.3 Adjuncts

As noted in Section 1.6, adjuncts are optional. In phrase structure grammars, this can be captured, for example, by rules that have a certain element (for instance a VP) on the left-hand side of the rule and the same element and an adjunct on the right-hand side of the rule. Since the symbol on the left is the same as the one on the right, this rule can be applied arbitrarily many times. (9) shows some examples of this:

- (9) a. $VP \rightarrow VP PP$
 b. $Noun \rightarrow Noun PP$

One can analyze an arbitrary amount of PPs following a VP or noun using these rules.

In Categorical Grammar, adjuncts have the following general form: $X \backslash X$ or X/X . Adjectives are modifiers, which must occur before the noun. They have the category n/n . Modifiers occurring after nouns (prepositional phrases and relative clauses) have the category $n \backslash n$ instead.³ For VP-modifiers, X is replaced by the symbol for the VP ($s \backslash np$) and this yields the relatively complex expression $(s \backslash np) \backslash (s \backslash np)$. Adverbials in English are VP-modifiers and have this category. Prepositions that can be used in a PP modifying a verb require an NP in order to form a complete PP and therefore have the category $((s \backslash np) \backslash (s \backslash np)) / np$. Figure 8.5 on the next page gives an example of an adverb (*quickly*) and a preposition (*round*). Note that the result of the combination of *round* and *the*

³ In Categorical Grammar, there is no category symbol like \overline{X} for intermediate projections of \overline{X} theory. So rather than assuming $\overline{N}/\overline{N}$, CG uses n/n . See Exercise 2.

8 Categorical Grammar

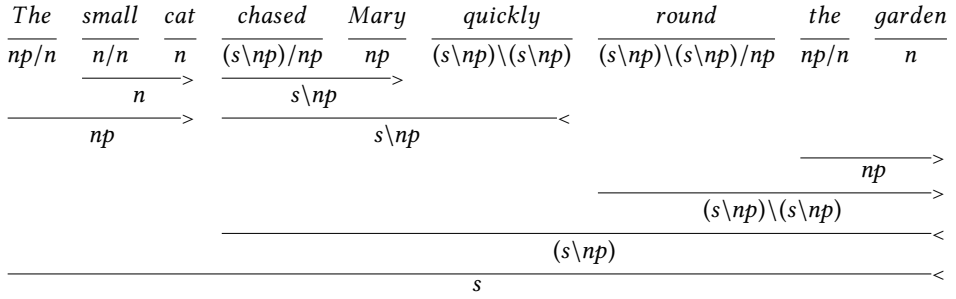


图 8.5: Example of an analysis with adjuncts in Categorical Grammar

garden corresponds to the category of the adverb $((s\ np)\ (s\ np))$. In GB theory, adverbs and prepositions were also placed into a single class (see page 87). This overarching class was then divided into subclasses based on the valence of the elements in question.

8.2 Passive

In Categorical Grammar, the passive is analyzed by means of lexical rule (Dowty 1978: 412; Dowty 2003: Section 3.4). (10) shows the rule in Dowty (2003: 49).

- (10) Syntax: $\alpha \in (s\ np)/ np \rightarrow \text{PST-PART}(\alpha) \in \text{PstP}/ np_{by}$
 Semantics: $\alpha' \rightarrow \lambda y \lambda x \alpha'(y)(x)$

Here, PstP stands for past participle and np_{by} is an abbreviation for a verb phrase modifier of the form $vp\backslash vp$ or rather $(s\ np)\ (s\ np)$. The rule says the following: if a word belongs to the set of words with the category $(s\ np)/ np$, then the word with past participle morphology also belongs in the set of words with the category PstP/ np_{by} .

(11a) shows the lexical entry for the transitive verb *touch* and (11b) the result of rule application:

- (11) a. *touch*: $(s\ np)/ np$
 b. *touched*: PstP/ np_{by}

The auxiliary *was* has the category $(s\ np)/ \text{PstP}$ and the preposition *by* has the category np_{by}/ np , or its unabbreviated form $((s\ np)\ (s\ np))/ np$. In this way, (12) can be analyzed as in Figure 8.6 在下一页.

- (12) John was touched by Mary.

The question as to how to analyze the pair of sentences in (13) still remains unanswered.⁴

- (13) a. He gave the book to Mary.
 b. The book was given to Mary.

⁴ Thanks to Roland Schäfer (p. m., 2009) for pointing out these data to me.

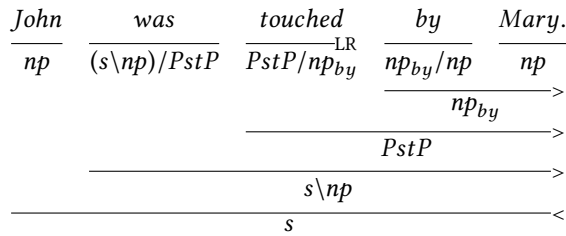


图 8.6: Analysis of the passive using a lexical rule

gave has the category $((s\ np)/pp)/np$, that is, the verb must first combine with an NP (*the book*) and a PP (*to Mary*) before it can be combined with the subject. The problem is that the rule in (10) cannot be applied to *gave* with a *to*-PP since the pp argument is sandwiched between both np arguments in $((s\ np)/pp)/np$. One would have to generalize the rule in (10) somehow by introducing new technical means⁵ or assume additional rules for cases such as (13b).

8.3 Verb position

Steedman (2000: 159) proposed an analysis with variable branching for Dutch, that is, there are two lexical entries for *at* ‘eat’: an initial one with its arguments to the right, and another occupying final position with its arguments to its left.

- (14) a. *at* ‘eat’ in verb-final position: $(s_{+SUB}\ np)\ np$
 b. *at* ‘eat’ in verb-initial position: $(s_{-SUB}/np)/np$

Steedman uses the feature SUB to differentiate between subordinate and non-subordinate sentences. Both lexical items are related via lexical rules.

One should note here that the NPs are combined with the verb in different orders. The normal order is:

- (15) a. in verb-final position: $(s_{+SUB}\ np[nom])\ np[acc]$
 b. in verb-initial position: $(s_{-SUB}/np[acc])\ np[nom]$

The corresponding derivations for German sentences with a bivalent verb are shown in Figures 8.7 and 8.8.

In Figure 8.7, the verb is first combined with an accusative object, whereas in Figure 8.8, the verb is first combined with the subject. For criticism of these kinds of analyses with variable branching, see Netter (1992) and Müller (2005b, 2015b).

Jacobs (1991) developed an analysis which corresponds to the verb movement analysis in GB. He assumes verb-final structures, that is, there is a lexical entry for verbs where arguments are selected to the left of the verb. A transitive verb would therefore have the

⁵ Baldridge (p. M. 2010) suggests using regular expressions in a general lexical rule for passive.

$$\begin{array}{c}
\frac{er}{np[nom]} \quad \frac{ihn}{np[acc]} \quad \frac{isst}{(s_{+SUB} \backslash np[nom]) \backslash np[acc]} \\
\hline
\frac{s_{+SUB} \backslash np[nom]}{s_{+SUB}} <
\end{array}$$

图 8.7: Analysis of verb-final sentences following Steedman

$$\begin{array}{c}
\frac{isst}{((s_{-SUB} / np[acc]) / np[nom])} \quad \frac{er}{np[nom]} \quad \frac{ihn}{np[acc]} \\
\hline
\frac{s_{-SUB} / np[acc]}{s_{-SUB}} >
\end{array}$$

图 8.8: Analysis of verb-initial sentences following Steedman

entry in (16a). Additionally, there is a trace in verb-final position that requires the arguments of the verb and the verb itself in initial position. (16b) shows what the verb trace looks like for a transitive verb in initial position:

- (16) a. Verb in final position:
 $(s \backslash np[nom]) \backslash np[acc]$
 b. Verb trace for the analysis of verb-first:
 $((s \backslash ((s \backslash np[nom]) \backslash np[acc])) \backslash np[nom]) \backslash np[acc]$

The entry for the verb trace is very complex. It is probably simpler to examine the analysis in Figure 8.9.

$$\begin{array}{c}
\frac{isst}{(s \backslash np[nom]) \backslash np[acc]} \quad \frac{er}{np[nom]} \quad \frac{ihn}{np[acc]} \quad \frac{-}{(((s \backslash (s \backslash np[nom]) \backslash np[acc]) \backslash np[nom]) \backslash np[acc])} \\
\hline
\frac{(s \backslash (s \backslash np[nom]) \backslash np[acc]) \backslash np[nom]}{s \backslash ((s \backslash np[nom]) \backslash np[acc])} < \\
\hline
s <
\end{array}$$

图 8.9: Analysis of verb-initial sentences following Jacobs (1991)

The trace is the head in the entire analysis: it is first combined with the accusative object and then with the subject. In a final step, it is combined with the transitive verb in initial-position.⁶ A problem with this kind of analysis is that the verb *isst* ‘eats’, as well as *er* ‘he’ and *ihn* ‘him’/‘it’, are arguments of the verb trace in (17).

- (17) Morgen [isst [er [ihn _]]]
 tomorrow eats he him
 ‘He will eat it/him tomorrow.’

⁶ See Netter (1992) for a similar analysis in HPSG.

Since adjuncts can occur before, after or between arguments of the verb in German, one would expect that *morgen* ‘tomorrow’ can occur before the verb *isst*, since *isst* is just a normal argument of the verbal trace in final position. As adjuncts do not change the categorial status of a projection, the phrase *morgen isst er ihn* ‘tomorrow he eats him’ should be able to occur in the same positions as *isst er ihn*. This is not the case, however. If we replace *isst er ihn* by *morgen isst er ihn* in (18a), the result is (18b), which is ungrammatical.

- (18) a. Deshalb isst er ihn.
 therefore eats he him
 ‘Therefore he eats it/him.’
 b. *Deshalb morgen isst er ihn.
 therefore tomorrow eats he him

An approach which avoids this problem comes from Kiss & Wesche (1991) (see Section 9.3). Here, the authors assume that there is a verb in initial position which selects a projection of the verb trace. If adverbials are only combined with verbs in final-position, then a direct combination of *morgen* ‘tomorrow’ and *isst er ihn* ‘eats he it’ is ruled out. If one assumes that the verb in first-position is the functor, then it is possible to capture the parallels between complementizers and verbs in initial position (Höhle 1997): finite verbs in initial position differ from complementizers only in requiring a projection of a verb trace, whereas complementizers require projections of overt verbs:

- (19) a. dass [er ihn isst]
 that he it eats
 b. Isst [er ihn _]
 eats he it

This description of verb position in German captures the central insights of the GB analysis in Section 3.2.

8.4 Local reordering

Up to now, we have seen combinations of functors and arguments where the arguments were either to the left or to the right of the functor. The saturation of arguments always took place in a fixed order: the argument furthest to the right was combined first with the functor, e.g., (s\np)/pp first combined with the PP, and the result of this combination was combined with the NP.

There are a number of possibilities to analyze ordering variants in German: Uszko-reit (1986b) suggests accounting for possible orders lexically; that is, that each possible order corresponds to a lexical item. One would therefore have at least six lexical items for a ditransitive verb. Briscoe (2000: 257) and Villavicencio (2002: 96–98) propose a variant of this analysis where the order of arguments is modified in the syntax: a syntactic rule can, for example, change the order (S/PRT)/NP into (S/NP)/PRT.

A different approach is suggested by Steedman & Baldridge (2006). They discuss various options for ordering arguments attested in the languages of the world. This includes languages in which the order of combination is free, as well as languages where the direction of combination is free. Steedman and Baldridge introduce the following convention for representing categories: elements in curly brackets can be discharged in any order. ‘|’ in place of ‘\’ or ‘/’ serves to indicate that the direction of combination is free. Some prototypical examples are shown in (20):

(20)	English	$(S \backslash NP) / NP$	$S(VO)$
	Latin	$S\{ NP[nom], NP[acc]\}$	free order
	Tagalog	$S\{ /NP[nom], /NP[acc]\}$	free order, verb-initial
	Japanese	$S\{ \backslash NP[nom], \backslash NP[acc]\}$	free order, verb-final

Hoffman (1995: Section 3.1) has proposed an analysis analogous to that of Japanese for Turkish and this could also be used in conjunction with an analysis of verb position for German. This would correspond to the GB/MP analysis of Fanselow (2001) or the HPSG analysis presented in Section 9.4.

8.5 Long-distance dependencies

Steedman (1989: Section 1.2.4) proposes an analysis of long-distance dependencies without movement or empty elements. For examples such as (21), he assumes that the category of *Harry must have been eating* or *Harry devours* is s/np .

- (21)
- a. These apples, Harry must have been eating.
 - b. apples which Harry devours

The fronted NP *these apples* and the relative pronoun *which* are both functors in the analysis of (21) which take s/np as their argument. Using the machinery introduced up to now, we cannot assign the category s/np to the strings *Harry must have been eating* and *Harry devours* in (21) although it is intuitively the case that *Harry devours* is a sentence missing an NP. We still require two further extensions of Categorical Grammar: type raising and forward and backward composition. Both of these operations will be introduced in the following sections.

8.5.1 Type Raising

The category np can be transformed into the category $(s/(s\ np))$ by *type raising*. If we combine this category with $(s\ np)$, then we get the same result as if we had combined np and $(s\ np)$ with the forward application rule in (2). (22a) shows the combination of an NP with a VP (a sentence missing an NP to its left). The combination of the type-raised NP with the VP is given in (22b).

- (22)
- a. $np * s\ np = s$
 - b. $s/(s\ np) * s\ np = s$

In (22a), a verb or verb phrase selects an NP to its left ($s\backslash np$). In (22b), an NP having undergone type raising selects a verb or verb phrase to its right which requires an NP to its left ($s\backslash np$).

Type raising simply reverses the direction of selection: the VP in (22a) is the functor and the NP is the argument, whereas in (22b), it is the type raised NP which acts as the functor, and the VP is the argument. In each case, the result of the combination is the same. This change of selectional direction may just seem like a trick at first glance, but as we will see, this trick can be extremely useful. First, however, we will introduce forward and backward composition.

8.5.2 Forward and backward composition

(23) shows the rules for forward and backward composition.

- (23) a. Forward composition ($> B$)
 $X/Y * Y/Z = X/Z$
 b. Backward composition ($< B$)
 $Y/Z * X/Y = X\backslash Z$

These rules will be explained using forward composition as an example. (23a) can be understood as follows: X/Y more or less means; if I find a Y, then I am a complete X. In the combinatorial rule, X/Y is combined with Y/Z . Y/Z stands for a Y that is not yet complete and is still missing a Z. The requirement that Y must find a Z in order to be complete is postponed: we pretend that Y is complete and use it anyway, but we still bear in mind that something is actually still missing. Hence, if we combine X/Y with Y/Z , we get something which becomes an X when combined with a Z.

8.5.3 Analysis of long-distance dependencies

By using forward composition, we can assign *Harry must have been eating* the category $s\backslash np$. Figure 8.10 shows how this works. *must* is a verb which requires an unmarked

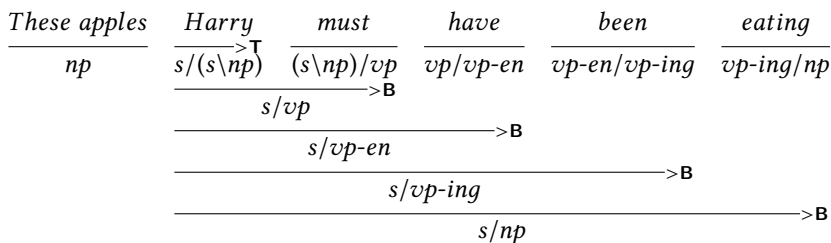


图 8.10: Application of forward composition to VP-chains

infinitive form, *have* requires a participle and *been* must combine with a gerund. In the above figure, the arrow with a small 'T' stands for type raising, whereas the arrows with

a ‘B’ indicate composition. The direction of composition is shown by the direction of the arrow.

For the analysis of (21a), we are still missing one small detail, a rule that turns the NP at the beginning of the sentence into a functor which can be combined with s/np . Normal type raising cannot handle this because it would produce $s/(s/np)$ when $s/(s/np)$ is required.

Steedman (1989: 217) suggests the rule in (24):

- (24) Topicalization (\uparrow):
 $X \Rightarrow st/(s/X)$
 where $X \in \{ np, pp, vp, ap, s' \}$

st stands for a particular type of sentence (s), namely one with topicalization (t). The \Rightarrow expresses that one can type raise any X into an $st/(s/X)$.

If we replace X with np , we can turn *these apples* into $st/(s/np)$ and complete the analysis of (21a) as shown in Figure 8.11.

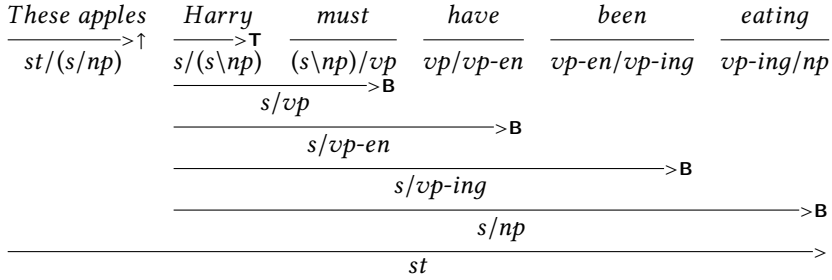


图 8.11: Analysis of long-distance dependencies in Categorical Grammar

The mechanism presented here will of course also work for dependencies that cross sentence boundaries. Figure 8.12 shows the analysis for (25):

- (25) Apples, I believe that Harry eats.

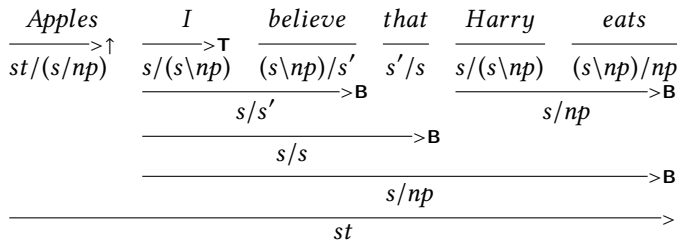


图 8.12: Analysis of long-distance dependencies across sentence boundaries

Using the previously described tools, it is, however, only possible to describe extractions where the fronted element in the sentence would have occurred at the right edge of the

phrase without fronting. This means it is not possible to analyze sentences where the middle argument of a ditransitive verb has been extracted (Steedman 1985: 532). Pollard (1988: 406) provides the derivation in Figure 8.13 for (26).

(26) Fido we put downstairs.

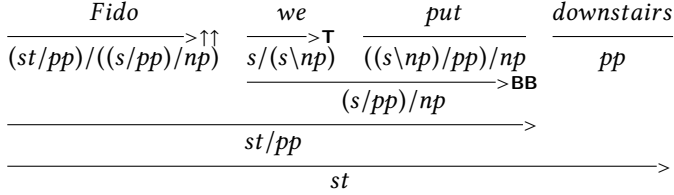


图 8.13: Analysis of long-distance dependencies across sentence boundaries

In this analysis, it is not possible to combine *we* and *put* using the rule in (23a) since $(s\backslash np)$ is not directly accessible: breaking down $((s\backslash np)/pp)/np$ into functor and argument gives us $((s\backslash np)/pp)$ and np . In order to deal with such cases, we need another variant of composition:

(27) Forward composition for $n=2$ ($> \text{BB}$)
 $X/Y * (Y/Z1)/Z2 = (X/Z1)/Z2$

With this addition, it is now possible to combine the type-raised *we* with *put*. The result is $(s/pp)/np$. The topicalization rule in (24), however, requires an element to the right of *st* with the form (s/X) . This is not the case in Figure 8.13. For the NP *Fido*, we need a functor category which allows that the argument itself is complex. The rule which is needed for the case in (26) is given in (28).

(28) Topicalization for $n=2$ ($\uparrow\uparrow$):
 $X2 \Rightarrow (st/X1)/((s/X1)/X2)$
 where $X1$ and $X2 \in \{ NP, PP, VP, AP, S' \}$

If we assume that verbs can have up to four arguments (z. B. *buy*: buyer, seller, goods, price), then it would be necessary to assume a further rule for composition as well as another topicalization rule. Furthermore, one requires a topicalization rule for subject extraction (Pollard 1988: 405). Steedman has developed a notation which provides a compact notation of the previously discussed rules, but if one considers what exactly these representations stand for, one still arrives at the same number of rules that have been discussed here.

8.6 Summary and classification

The operations of Combinatory Categorical Grammar, which go beyond those of standard Categorical Grammar, allow for so much flexibility that it is even possible to

assign a category to sequences of words that would not normally be treated as a constituent. This is an advantage for the analysis of coordination (see Section 21.6.2) and furthermore, Steedman (1991) has argued that intonation data support the constituent status of these strings. See also Section 15.2 for a direct model of incremental language processing in Categorical Grammar. In phrase structure grammars, it is possible to use GPSG mechanisms to pass information about relative pronouns contained in a phrase up the tree. These techniques are not used in CG and this leads to a large number of recategorization rules for topicalization and furthermore leads to inadequate analyses of pied-piping constructions in relative clauses. As the topicalization analysis was already discussed in Section 8.5, I will briefly elaborate on relative clauses here.

Steedman & Baldridge (2006: 614) present an analysis of long-distance dependencies using the following relative clause in (29):

(29) the man that Manny says Anna married

The relative pronoun is the object of *married* but occurs outside the clause *Anna married*.

Steedman assumes the lexical entry in (30) for relative pronouns:

(30) $(n \backslash n) / (s / np)$

This means the following: if there is a sentence missing an NP to the right of a relative pronoun, then the relative pronoun can form an N-modifier $(n \backslash n)$ with this sentence. The relative pronoun is the head (functor) in this analysis.

Utilizing both additional operations of type raising and composition, the examples with relative clauses can be analyzed as shown in Figure 8.14. The lexical entry for the

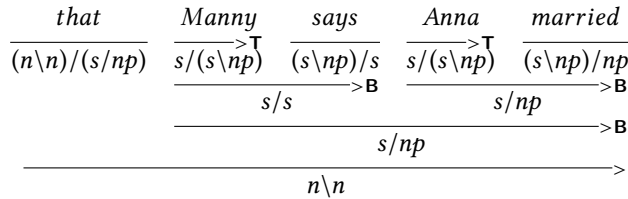


图 8.14: Categorical Grammar analysis of a relative clause with long-distance dependency

verbs corresponds to what was discussed in the preceding sections: *married* is a normal transitive verb and *says* is a verb that requires a sentential complement and forms a VP $(s \backslash np)$ with it. This VP yields a sentence when combined with an NP. The noun phrases in Figure 8.14 have been type raised. Using forward composition, it is possible to combine *Anna* and *married* to yield s / np . This is the desired result: a sentence missing an NP to its right. *Manny* and *says* and then *Manny says* and *Anna married* can also be combined via forward composition and we then have the category s / np for *Manny says Anna married*. This category can be combined with the relative pronoun using forward application and we then arrive at $n \backslash n$, which is exactly the category for postnominal modifiers.

However, the assumption that the relative pronoun constitutes the head is problematic since one has to then go to some lengths to explain pied-piping constructions such as those in (31).

- (31) a. Here's the minister [[in [the middle [of [whose sermon]]]] the dog barked].⁷
 b. Reports [the height of the lettering on the covers of which] the government prescribes should be abolished.⁸

In (31), the relative pronoun is embedded in a phrase that has been extracted from the rest of the relative clause. The relative pronoun in (31a) is the determiner of *sermon*. Depending on the analysis, *whose* is the head of the phrase *whose sermon*. The NP is embedded under *of* and the phrase *of whose sermon* depends on *middle*. The entire NP *the middle of the sermon* is a complement of the preposition *in*. It would be quite a stretch to claim that *whose* is the head of the relative clause in (31a). The relative pronoun in (31b) is even more deeply embedded. Steedman (1997: 50) gives the following lexical entries for *who*, *whom* and *which*:

- (32) a. $((n \setminus n) / (s \setminus np)) \setminus (np / np)$ (complex subject-relative phrase)
 b. $((n \setminus n) / (s \setminus pp)) \setminus (pp / np)$ (complex extracted PP-relative phrase)
 c. $((n \setminus n) / (s \setminus np)) \setminus (np / np)$ (complex extracted NP-relative phrase)

Using (32b) and (32c), it is possible to analyze (33a) and (33b):

- (33) a. a report the cover of which Keats (expects that Chapman) will design
 b. a subject on which Keats (expects that Chapman) will speak

In the analysis of (33b), *which* requires a preposition to its left (*pp/np*) so it can form the category $(n \setminus n) / (s \setminus pp)$. This category needs a sentence lacking a PP to its right in order to form a post-nominal modifier $(n \setminus n)$. In the analysis of (33a), *the cover of* becomes *np/np* by means of composition and *which* with the lexical entry (32c) can combine with *the cover of* to its left. The result is the category $(n \setminus n) / (s \setminus np)$, that is, something that requires a sentence missing an NP.

Ross' examples (31b) can also be analyzed as follows (32c):

- (34) reports [the height of the lettering on the covers of]_{np/np} which]_{(n \setminus n) / (s \setminus np)} the government prescribes

The complex expression *the height of the lettering on the covers of* becomes *np/np* after composition and the rest of the analysis proceeds as that of (33a).

In addition to entries such as those in (32), we also need further entries to analyze sentences such as (35), where the relative phrase has been extracted from the middle of the clause (see Pollard 1988: 410):

- (35) Fido is the dog which we put downstairs.

⁷ Pollard & Sag (1994: 212).

⁸ Ross (1967: 109).

8 Categorical Grammar

The problem here is similar to what we saw with topicalization: *we put* does not have the category s/np but rather $(s/pp)/np$ and as such, cannot be directly combined with the relative pronoun in (30).

Morrill (1995: 204) discusses the lexical entry in (32b) for the relative pronoun in (36):

(36) about which John talked

In the lexical entry (32b), *which* requires something to the left of it, which requires a noun phrase in order to form a complete prepositional phrase; that is, *which* selects a preposition. Morrill noted that there is a need to postulate further lexical items for cases like (37) in which the relative pronoun occurs in the middle of the fronted phrase.

(37) the contract [the loss of which after so much wrangling] John would finally have to pay for

These and other cases could be handled by additional lexical stipulations. Morrill instead proposes additional types of the combination of functors and arguments, which allow a functor $B \uparrow A$ to enclose its argument A and produce B , or a functor $A \downarrow B$ to enclose its argument to then yield B (p. 190). Even with these additional operations, he still needs the two lexical items in (38) for the derivation of a pied-piping construction with an argument NP or a PP:

- (38) a. $(NP \uparrow NP) \downarrow (N \setminus N)/(S/NP)$
 b. $(PP \uparrow NP) \downarrow (N \setminus N)/(S/PP)$

These lexical items are still not enough, however, as (38b) contains a PP but this PP corresponds to an argument PP, which is required for (36). To analyze (31a), which involves a PP adjunct, we need to assume the category $(s \setminus np)/(s \setminus np)$ for the prepositional phrase *in the middle of whose sermon*. We, therefore, also require at least three additional items for relative pronouns.

By introducing new operations, Morrill manages to reduce the number of lexical entries for *which*; however, the fact remains that he has to mention the categories which can occur in pied-piping constructions in the lexical entry of the relative pronoun.

Furthermore, the observation that relative clauses consist of a phrase with a relative pronoun plus a sentence missing a relative phrase is lost. This insight can be kept if one assumes a GPSG-style analysis where information about whether there is a relative pronoun in the relative phrase can be passed up to the highest node of the relative phrase. The relative clause can then be analyzed as the combination of a sentence with a gap and an appropriately marked relative phrase. For the discussion of such analyses in the framework of GB theory and HPSG/CxG, see Section 21.10.3.

Comprehension questions

1. Identify the functors and arguments in Figures 8.1 and 8.3.
2. Which combination operations do you know?
3. What is composition used for?

Exercises

1. Analyze the following sentence:

(39) The children in the room laugh loudly.

2. Analyze the noun phrase in (40):

(40) the picture of Mary

Compare the resulting analysis with the structure given in Figure 2.4 on page 62 and think about which categories of \bar{X} syntax the categories in Categorical Grammar correspond to.

Further reading

Mark Steedman discusses a variant of Categorical Grammar, *Combinatory Categorical Grammar*, in a series of books and articles: Steedman (1991, 2000); Steedman & Baldridge (2006).

Lobin (2003) compares Categorical Grammar with Dependency Grammar and Pickering & Barry (1993) suggest a combination of Dependency Grammar and Categorical Grammar, which they call Dependency Categorical Grammar.

Briscoe (2000) and Villavicencio (2002) discuss UG-based acquisition models in the framework of Categorical Grammar.

第九章 中心语驱动的短语结构语法

中心语驱动的短语结构语法 (Head-Driven Phrase Structure Grammar, HPSG) 是由 Carl Pollard 和 Ivan Sag 在上世纪八十年代在斯坦福和 Palo Alto 的惠普研究实验室开发出来的 (Pollard & Sag 1987, 1994)。与 LFG 一样, HPSG 隶属于西海岸语言学。另一个与 LFG 的相似之处在于 HPSG 旨于提供一个语言运用与语言能力相互兼容的理论 (Sag & Wasow 2011, 2015, 以及第 15 章)。

HPSG 语法的形式化描述语言是很好理解的, 而且我们有很多可以处理这种语法的系统 (Dörre & Seiffert 1991; Dörre & Dorna 1993; Popowich & Vogel 1991; Uszkoreit, Backofen, Busemann, Diagne, Hinkelman, Kasper, Kiefer, Krieger, Netter, Neumann, Oepen & Spackman 1994; Erbach 1995; Schütz 1996; Schmidt, Theofilidis, Rieder & Declerck 1996b; Schmidt, Rieder & Theofilidis 1996a; Uszkoreit, Backofen, Calder, Capstick, Dini, Dörre, Erbach, Estival, Manandhar, Mineur & Oepen 1996; Müller 1996c, 2004c; Carpenter & Penn 1996; Penn & Carpenter 1999; Götz, Meurers & Gerdemann 1997; Copestake 2002; Callmeier 2000; Dahllöf 2003; Meurers, Penn & Richter 2002; Penn 2004; Müller 2007a; Sato 2008; Kaufmann 2009)。¹ 目前, Ann Copestake 开发的 LKB 系统和 Gerald Penn (Meurers, Penn & Richter 2002; Penn 2004) 开发的 TRALE 系统拥有最多的用户。DELPH-IN 联盟 (它的语法片段是基于 LKB 的) 和许多的不同的 TRALE 用户已经开发了许多不同语言的小型语法和一些大规模的语法片段。下面列出的就是不同系统中实现的语言:

- 阿拉伯语 (Arabic) (Haddar, Boukedi & Zalila 2010; Hahn 2011; Masum, Islam, Rahman & Ahmed 2012; Boukedi & Haddar 2014; Loukam, Balla & Laskri 2015; Arad Greshler, Herzig Sheinfux, Melnik & Wintner 2015),
- 孟加拉语 (Bengali) (Paul 2004; Islam, Hasan & Rahman 2012),
- 保加利亚语 (Bulgarian) (Simov, Osenova, Simov & Kouylekov 2004; Osenova 2010a,b, 2011),
- 粤语 (Cantonese) (Fan, Song & Bond 2015),
- 丹麦语 (Danish) (Ørsnes 1995, 2009b; Neville & Paggio 2004; Müller 2009c; Müller & Ørsnes 2011; Müller 2012; Müller & Ørsnes 2015),

¹ Uszkoreit et al. (1996) 和 Bolc et al. (1996) 比较了已有的和 1990 年代初开发的系统。Melnik (2007) 比较了 LKB 和 TRALE。也请看 Müller (2015a: Section 5.1)。

- 德语 (German) (Kiss 1991; Netter 1993, 1996; Meurers 1994; Hinrichs et al. 1997; Kordoni 1999a; Tseng 2000; Geißler & Kiss 1994; Keller 1994; Müller 1996c, 1999a; Müller & Kasper 2000; Crysmann 2003, 2005b,c; Müller 2007b; Kaufmann & Pfister 2007, 2008; Kaufmann 2009; Fokkens 2011),
- 英语 (English) (Copestake & Flickinger 2000; Flickinger, Copestake & Sag 2000; Flickinger 2000; Dahllöf 2002, 2003; De Kuthy & Meurers 2003a; Meurers, De Kuthy & Metcalf 2003; De Kuthy, Metcalf & Meurers 2004),
- 世界语 (Esperanto) (Li 1996),
- 法语 (French) (Tseng 2003),
- Ga 语 (Ga) (Kropp Dakubu, Hellan & Beermann 2007; Hellan 2007),
- 格鲁吉亚语 (Georgian) (Abzianidze 2011),
- 希腊语 (Greek) (Kordoni & Neu 2005),
- 豪萨语 (Hausa) (Crysmann 2005a, 2009, 2011, 2012, 2016),
- 希伯来语 (Hebrew) (Melnik 2007; Haugereid, Melnik & Wintner 2013; Arad Greshler, Herzig Sheinflux, Melnik & Wintner 2015),
- 印尼语 (Indonesian) (Moeljadi, Bond & Song 2015)
- 日语 (Japanese) (Siegel 2000; Siegel & Bender 2002; Bender & Siegel 2005),
- 韩语 (Korean) (Kim & Yang 2003, 2004, 2006, 2009; Kim, Sells & Yang 2007; Song, Kim, Bond & Yang 2010; Kim, Yang, Song & Bond 2011),
- 马耳他语 (Maltese) (Müller 2009b),
- 现代汉语 (Mandarin Chinese) (Liu 1997; Ng 1997; Müller & Lipenkova 2009, 2013; Fan, Song & Bond 2015),
- 荷兰语 (Dutch) (van Noord & Bouma 1994; Bouma, van Noord & Malouf 2001b; Fokkens 2011),
- 挪威语 (Norwegian) (Hellan & Haugereid 2003; Beermann & Hellan 2004; Hellan & Beermann 2006),
- 波斯语 (Persian) (Müller 2010b; Müller & Ghayoomi 2010),
- 波兰语 (Polish) (Przepiórkowski, Kupść, Marciniak & Mykowiecka 2002; Mykowiecka, Marciniak, Przepiórkowski & Kupść 2003),
- 葡萄牙语 (Portuguese) (Branco & Costa 2008a,b; Costa & Branco 2010),
- 俄语 (Russian) (Avgustinova & Zhang 2009),

- 萨哈泼丁语 (Sahaptin) (Drellishak 2009),
- 西班牙语 (Spanish) (Pineda & Meza 2005a,b; Bildhauer 2008; Marimon 2013),
- 手语 (Sign Language) (德语、法语、英式英语、希腊语)
- 南美手语 (South African Sign Language) (Bungeroth 2002),
- 土耳其语 (Turkish) (Fokkens, Poulson & Bender 2009),
- Wambaya 语 (Wambaya) (Bender 2008a,c, 2010).
- 依地语 (Yiddish) (Müller & Ørsnes 2011),

第一个应用 HPSG 理论实现的语法是在 Palo Alto 的惠普实验室开发的英语语法 (Flickinger, Pollard & Wasow 1985; Flickinger 1987)。德语语法是由海德堡、斯图加特和萨尔布吕肯的 LILOG 项目开发的。随后, 在海德堡、萨尔布吕肯和斯坦福的 *Verbmobil* 项目共同开发了德语、英语和日语的语法。*Verbmobil* 是在德国历时最长的人工智能项目。它是针对旅游计划和日程安排领域的口语的翻译项目 (Wahlster 2000a)。

目前, 在语法开发方面有两团体: DELPH-IN 联盟 (应用 HPSG 的深层语言处理)² 和 CoGETI 网络 (基于约束的语法: 经验、理论与实现)³。上面列出的大部分语法片段是由 DELPH-IN 的成员开发的, 而有一些则来自于语法矩阵 (Grammar Matrix)。语法矩阵是为 LKB 开发的一个平台, 它为语法编写者提供了一个类型驱动的原始语法, 该语法对应于所开发的语言的属性 (Bender, Flickinger & Oepen 2002)。核心语法工程 (the CoreGram project)⁴ 是一个在柏林自由大学运行的类似项目。它是针对德语、丹麦语、波斯语、马耳他语、现代汉语、西班牙语、法语和印地语的语法开发项目, 这些语法都共享一个核心语法。针对所有语言的约束条件在一个地方表示, 并且应用到所有语法中。而且, 还有针对特定语言类型的限制, 这些限制被一起表示并且根据各自的语法来应用。所以, 语法矩阵是用来供个人语法编写者可以使用、调试以适应其开发的语法所需要的平台, 而核心语法是真正地针对同时使用的不同语言的, 并且它们需要保持同步。有关核心语法的介绍可以参考 Müller (2013a, 2015a)。

还有的系统是将语言学驱动的分析与统计成分结合起来 (Brew 1995; Miyao et al. 2005; Miyao & Tsujii 2008), 或者从语料库中学习语法和辞典 (Fouvy 2003; Cramer & Zhang 2009)。

下面的网址指向可以测试的语法页面:

- <http://www.delph-in.net/erg/>
- <http://hpsg.fu-berlin.de/Demos/>

² <http://www.delph-in.net/>. 2015 年 11 月 13 日

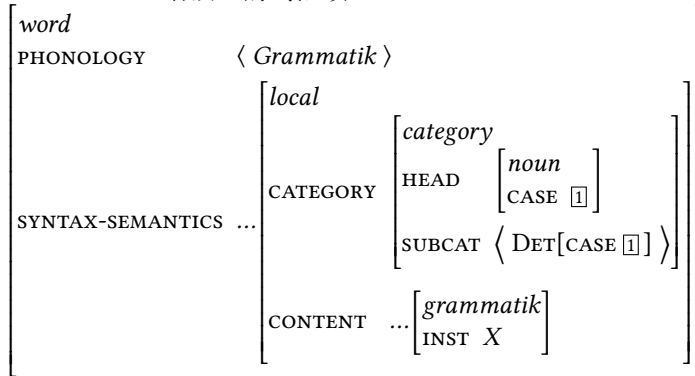
³ <http://wwwuser.gwdg.de/~cogeti/>. 2015 年 11 月 13 日. 由 DFG 基金 (基金编号: HO3279/3-1) 资助。

⁴ <http://hpsg.fu-berlin.de/Projects/CoreGram.html>. 20.03.2017.

9.1 有关表示形式的一般说明

HPSG 具有下述特征：它是基于词汇的理论，即大部分的语言约束位于词或根节点的描写中。HPSG 基于索绪尔的符号论：语言符号的形式与意义总是一起表示的。类型特征结构被用于模拟所有相关的信息。⁵这些结构可以跟 (1) 中的特征描写一起描述。词汇项、短语和原则都按照同样的形式化方法来模拟和描述。有关词类和规则模式的概括由承继层级体系来表示（请看第 6.2 节）。语音、句法和语义在单一结构中表示。没有像管辖与约束理论中 PF 或 LF 这样单独的表示层次。(1) 节录了一部分 *Grammatik*（语法）这个词的表示形式。

(1) *Grammatik*（语法）的词汇项：



我们可以看到这一特征描写包括词的语音、句法范畴和语义。为了简便，PHONOLOGY (PHON) 的值大部分按照正字法来表示。在完整的理论中，PHON value 是一个包括节律栅和轻重音信息的复杂结构。请看 Bird & Klein (1994)、Orgun (1996)、Höhle (1999)、Walther (1999)、Crysmann (2002: Chapter 6) 和 Bildhauer (2008) 在 HPSG 框架下关于语音学的分析。针对 (1) 中表示的详细信息将在下面的章节中给予解释。

HPSG 从其他理论中借鉴了许多不同的思路，而且近期的分析受到了其他理论框架下理论发展的影响。针对配价信息和功能组合的处理的函子参数结构被范畴语法所采用。函子构成在德语和韩语这类语言中的动词复杂结构的分析中起到了重要的作用。直接统制/线性优先模式 (ID/LP 模式，请看第 5.1.2 节) 和长距离依存（请看第 5.4 节）的 Slash 机制都来源于 GPSG。这里针对德语动词位置的分析受到管辖与约束理论框架下开发的语法的启发（请看第 3.2 节）。

9.1.1 配价信息的表示

第 2 章讨论的短语结构语法的缺点是需要大量不同的规则来表示不同的配价类型。(2) 给出了这类规则的一些例子以及相应的动词。

⁵ 没有按照顺序阅读的读者和对类型特征描写不太熟悉的读者可以先参考第 6 章。

- (2)
- | | |
|--|---|
| $S \rightarrow \text{NP}[\text{nom}], V$ | $X \text{ schläft}$ (X 正在睡觉) |
| $S \rightarrow \text{NP}[\text{nom}], \text{NP}[\text{acc}], V$ | $X Y \text{ erwartet}$ (X 在等 Y) |
| $S \rightarrow \text{NP}[\text{nom}], \text{PP}[\text{über}], V$ | $X \text{ über } Y \text{ spricht}$ (X 在谈论 Y) |
| $S \rightarrow \text{NP}[\text{nom}], \text{NP}[\text{dat}], \text{NP}[\text{acc}], V$ | $X Y Z \text{ gibt}$ (X 把 Z 给 Y) |
| $S \rightarrow \text{NP}[\text{nom}], \text{NP}[\text{dat}], \text{PP}[\text{mit}], V$ | $X Y \text{ mit } Z \text{ dient}$ (X 与 Z 一起服务 Y) |

为了保证语法不制造出不正确的句子，我们必须确保动词只根据合适的规则来使用。

- (3)
- * dass Peter das Buch schläft
CONJ Peter DET 书 睡觉
 - * dass Peter erwartet
CONJ Peter 等待
 - * dass Peter über den Mann erwartet
CONJ Peter 关于 DET 人 等待

所以说，动词（和通常所说的中心语）必须分成不同的配价类型。然后，这些配价类型必须被分配给语法规则。然后，我们必须要进一步明确（2）中及物动词的规则，如下所示：

- (4) $S \rightarrow \text{NP}[\text{nom}], \text{NP}[\text{acc}], V[\text{nom_acc}]$

这里，配价被编码两次。首先，我们说明了哪类成分可以或者必须发生，然后我们在词汇中说明动词所属的配价类型。在第 5.5 节，我们指出了屈折变化过程与配价信息是相关的。所以，我们需要从语法规则中去除多余的配价信息。基于这个原因，HPSG 跟范畴语法一样包括了在中心语所在的词汇项的中心语论元的表示。SUBCAT feature 这个特征带有列表值，它包括为了得到一个完整短语必须要跟中心语组合的宾语的表示。（5）给出了（2）中的动词的例子：

- (5)
- | | |
|----------------------|---|
| Verb | SUBCAT |
| <i>schlafen</i> (睡觉) | $\langle \text{NP}[\text{nom}] \rangle$ |
| <i>erwarten</i> (等待) | $\langle \text{NP}[\text{nom}], \text{NP}[\text{acc}] \rangle$ |
| <i>sprechen</i> (说话) | $\langle \text{NP}[\text{nom}], \text{PP}[\text{über}] \rangle$ |
| <i>geben</i> (给) | $\langle \text{NP}[\text{nom}], \text{NP}[\text{dat}], \text{NP}[\text{acc}] \rangle$ |
| <i>dienen</i> (服务) | $\langle \text{NP}[\text{nom}], \text{NP}[\text{dat}], \text{PP}[\text{mit}] \rangle$ |

SUBCAT 是次范畴化的缩写。通常来说，中心语需要次范畴的论元。请看第 84 页更多关于次范畴化 (subcategorization) 这个术语的内容。

图 9.1 on the following page 给出了 (6a) 的分析，而且图 9.2 on the next page 给出了 (6b) 的分析：

- (6)
- [dass] Peter schläft
CONJ Peter 睡觉
 - [dass] Peter Maria erwartet
CONJ Peter Maria 等待
'Peter 在等 Maria'

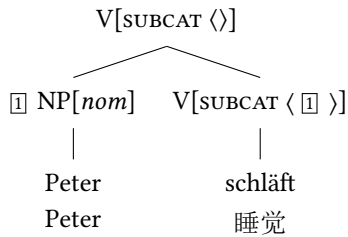


图 9.1: 小句 *dass Peter schläft* (Peter 在睡觉) 内 *Peter schläft* (Peter 睡觉) 的分析

在图 9.1和 9.2中, `SUBCAT list` 的一个元素在每个局部树中与其中心语相组合。与所选择的中心语相组合的元素不再出现在母节点的 `SUBCAT list` 中。`V[SUBCAT < >]` 对应于一个完整的短语 (VP 或 S)。带有数字的盒子表示结构共享 (请看第 6.4节)。结构共享是 HPSG 中最为重要的表达手段。它在诸如配价、一致和长距离依存中发挥着重要的作用。在上面的例子中, `[1]` 表示 `SUBCAT list` 中的描写与树中的另一个子节点是相同的。在配价列表中的描写通常是部分描写, 也就是说, 不是论元的所有属性都被穷尽地描写出来。所以说, 有可能像 *schläft* (睡觉) 这样的动词可以跟不同种类的语言对象相组合: 主语可以是一个代词、一个专有名词或是一个复杂的名词短语, 唯一关键的地方是我们所说的语言对象要有一个空的 `SUBCAT list`, 并且具有正确的格属性。⁶

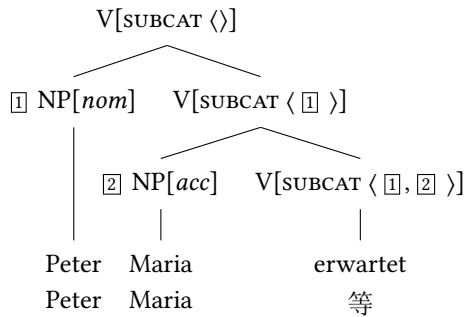


图 9.2: *Peter Maria erwartet* (Peter 在等 Maria) 的分析

9.1.2 组成成分结构的表示

正如我们已经指出的, HPSG 中的特征表示是形态规则、词汇项和句法规则的唯一描写机制。我们目前已有的树只是组成成分结构的可视化结果, 他们并不具有任何理论地位。在 HPSG 中也有重写规则。⁷ 短语结构规则的工作由特征描写来

⁶ 而且, 它必须与动词保持一致。这里并没有显示出这一点。

⁷ 但是, 在某些 HPSG 的计算实现中应用了短语结构规则, 这是为了提高处理的效率。

处理。有关统制的信息通过 **DTR** 特征（中心语子节点和非中心语子节点）表示出来，有关优先顺序的信息在 **PHON** 中表示。(7) 展示了特征表示中 **PHON values** 是如何表示的，该特征表示对应于图 9.3 中的树。

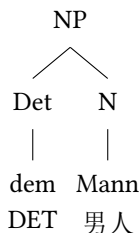
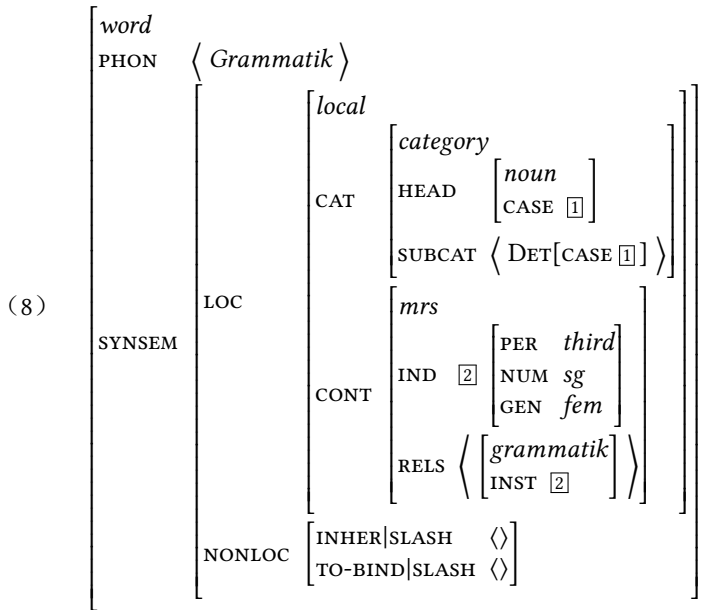


图 9.3: *dem Mann*（这个男人）的分析

$$(7) \left[\begin{array}{ll} \text{PHON} & \langle \text{dem Mann} \rangle \\ \text{HEAD-DTR} & \left[\text{PHON} \langle \text{Mann} \rangle \right] \\ \text{NON-HEAD-DTRS} & \left\langle \left[\text{PHON} \langle \text{dem} \rangle \right] \right\rangle \end{array} \right]$$

在 (7) 中，只有一个中心语子节点 (**HEAD-DTR**)。中心语子节点是包括中心语的子节点。在带有子节点 *das* (这) 和 *Bild von Maria* (Maria 的照片) 的结构中，后者是中心语子节点。原则上，可以有多个非中心语子节点。如果我们假设带有及物动词的句子是一个平铺结构的话，如图 2.1 on page 50 所示，那么我们就有三个非中心语子节点。我们也可以假设一个没有中心语的二叉结构（请看 Müller 2007b: Chapter 11 关于关系小句的分析）。在这类结构中，我们可以有不只一个非中心语子节点，具体来说是两个。

在我们展示只有那些论元与中心语的要求相匹配的中心语-论元结构是如何被允准的之前，我将先说明 HPSG 中特征描写的一般结构。在本章开头给出的结构在这里的 (8) 重复显示出来，并且加上了目前讨论相关的所有细节：



外层有特征 **PHON** 和 **SYNSEM**。正如前面提到的，**PHON** 包括语言对象的语音表示。**SYNSEM** 的值是包括可以被其他中心语所选择的句法和语义信息的特征结构。短语符号的子节点在 **SYNSEM** 之外表示。这就确保了在选择中具有一定程度的局部性：中心语不能进入它所选择的元素的内部结构（Pollard 和 Sag 1987: 143–145、1994: 23）。也可以参考第 10.6.2.1 节和第 18.2 节关于局部性的讨论。在 **SYNSEM** 内部，有关于局部语境的信息（**LOCAL**，简写为 **LOC**），以及长距离依存的信息（**NONLOCAL** 或简写为 **NONLOC**）。局部相关的信息包括句法（**CATEGORY** 或 **CAT**）和语义（**CONTENT** 或 **CONT**）信息。句法信息包括决定短语的核心属性的信息，即中心语信息。这在 **HEAD** 下面有所表示。更多细节将在第 9.1.4 节讨论。此外，语言对象的词类属于短语的中心语属性。同样，**HEAD**、**SUBCAT** 属于 **CAT** 内部的信息。符号的语义内容用 **CONT** 来表示。**CONT** value 的类型是 *mrs*，它表示最小递归语义学（Minimal Recursion Semantics）（Copestake, Flickinger, Pollard & Sag 2005）。一个 *MRS* 结构包括一个索引和限制该索引的关系列表。在 **NONLOCAL** 特征中，这里只给出了 **SLASH**。还有处理关系小句和疑问小句的特征（citealpps2、Sag 1997、Ginzburg & Sag 2000、Holler 2005），不过不在这里讨论。

正如我们看到的，*Grammatik*（语法）这个词的描写相对复杂。理论上，我们可以列出在一个单独的特征-值偶对的列表中直接列出给定对象的所有属性。但是，这会带来一些问题，这些特征-值偶对组的认定很难表示。应用（8）中的特征向量，我们表示这样的事实，在诸如（9）中的那些对称并列结构中，所有连词的 **CAT** values 是相同的。

- (9) a. [der Mann] und [die Frau]
DET 男人 和 DET 女人

- b. Er [kennt] und [liebt] diese Schallplatte.
他.NOM 认识 和 爱 这.ACC 专辑
- c. Er ist [dumm] und [arrogant].
他 COP 哑的 和 傲慢

(9b) 应该跟 (10) 中的例子相比较。在 (10a) 中，动词分别选择了一个宾格宾语和一个与格宾语，并且在 (10b) 中，动词选择了一个宾格宾语和一个介词宾语：

- (10) a. *Er kennt und hilft dieser Frau / diese Frau.
他.NOM 认识 和 帮助这.DAT 女人 这.ACC 女人
想说：“他认识并帮助这个女人。”
- b. *weil er auf Maria kennt und wartet
因为他PREP Maria 认识 和 等待
想说：“因为他认识 Maria，并等她”

不过，(10a) 的英文译文是合适的，因为 *knows* 和 *helps* 都带一个宾格宾语。而 (10a) 是不合格的，因为 *kennt* 带一个宾格宾语，而 *hilft* 带一个与格宾语。相似地，(10b) 也是不合格的，因为 *kennt* 带一个宾格宾语，而 *wartet* 带一个包括介词 *auf* 的介词短语。

如果配价和词类信息没有在一个共同的子结构中表示的话，我们就需要分别说明，诸如 (9) 的话语需要所有的连词具有相同的配价和词类信息。

在介绍了特征向量之后，我们现在可以转向中心语-论元模式的内容了：

Schema 1 (中心语-论元模式 (二叉结构, 初级版本))

$$\text{head-argument-phrase} \Rightarrow \left[\begin{array}{l} \text{SYNSEM|LOC|CAT|SUBCAT } \boxed{1} \\ \text{HEAD-DTR|SYNSEM|LOC|CAT|SUBCAT } \boxed{1} \oplus \langle \boxed{2} \rangle \\ \text{NON-HEAD-DTRS } \langle [\text{SYNSEM } \boxed{2}] \rangle \end{array} \right]$$

模式 1 表明了具有类型 *head-argument-phrase* 的语言对象必须具有的属性。模式 1 中的箭头表示逻辑蕴含，并不是我们从短语结构语法中所知的重写规则的箭头。‘ \oplus ’ (附加关系) 是包括两个列表的关系。(11) 显示了包括两个元素的列表的可能分叉结构：

$$(11) \quad \langle x, y \rangle = \langle x \rangle \oplus \langle y \rangle \text{ or } \langle \rangle \oplus \langle x, y \rangle \text{ or } \langle x, y \rangle \oplus \langle \rangle$$

列表 $\langle x, y \rangle$ 可以进一步被划分为两个列表，每个列表包括一个元素，或者相反地被划分为空列表和 $\langle x, y \rangle$ 。

模式 1 可以这样来解读：如果一个对象属于类型 *head-argument-phrase*，那么它必须具有蕴含右手边的属性。在具体的术语中，这意味着这些对象总是具有对应于 $\boxed{1}$ 的配价列表，他们具有一个中心语子节点，该子节点具有一个可以划分为两个

子列表①和(②)的配价列表, 而且他们具有一个非中心语子节点, 它的句法和语义属性 (SYNSEM value) 与中心语子节点 (②) 的 SUBCAT list 的最后一个元素是兼容的。(12) 提供了与 (6a) 中的例子相对应的特征描写。

$$(12) \left[\begin{array}{l} \text{head-argument-phrase} \\ \text{PHON} \langle \text{Peter schläft} \rangle \\ \text{SYNSEM|LOC|CAT|SUBCAT} \langle \rangle \\ \text{HEAD-DTR} \left[\begin{array}{l} \text{PHON} \langle \text{schläft} \rangle \\ \text{SYNSEM|LOC|CAT|SUBCAT} \langle \text{① NP[nom]} \rangle \end{array} \right] \\ \text{NON-HEAD-DTRS} \left\langle \left[\begin{array}{l} \text{PHON} \langle \text{Peter} \rangle \\ \text{SYNSEM} \text{①} \end{array} \right] \right\rangle \end{array} \right]$$

NP[nom] 是复杂特征描写的缩写。模式 1 将中心语子节点的 SUBCAT list 划分为一个单一元素列表和其他部分。由于 *schläft* (睡觉) 在它的 SUBCAT list 中只有一个元素, 剩余的是空列表。这个剩余部分也是母节点的 SUBCAT value。

9.1.3 线性顺序规则

统制模式并没有说有关子节点顺序的问题。正如在 GPSG 中, 线性顺序规则被分开处理。线性顺序规则可以指定子节点的属性, 他们在模式中的功能 (中心语、论元、说明语、...) 或者两者都有。如果我们假定所有的中心语具有特征 INITIAL, 那么位于他们所带论元之前的中心语将的 INITIAL 值为 ‘+’, 而位于他们所带论元之后的中心语的值为 ‘-’。(13) 中的线性顺序规则确保了诸如 (14b,d) 的不合乎语法的顺序被规则排除了。⁸

- (13) a. Head[INITIAL +] < Argument
b. Argument < Head[INITIAL -]

介词的 INITIAL 值为 ‘+’, 而且必须位于论元的前面。位于末尾的动词的值为 ‘-’, 由此他们必须位于他们所带的论元后面。

- (14) a. [in [den Schrank]]
PREP DET 壁橱
b. *[[den Schrank] in]
DET 壁橱 PREP
c. dass [er [ihn umfüllt]]
CONJ 他 它 倒
d. *dass [er [umfüllt ihn]]
CONJ 他 倒 它

⁸ 名词短语会给 (13) 带来问题: 截至目前, 限定词被看作是论元, 并且包括在中心语名词的 SUBCAT list 中。限定词在名词的左边出现, 而名词的所有其他论元在右边出现。这个问题可以通过重新界定线性顺序规则 (Müller 1999a: 164–165) 来解决, 或者通过为限定词引入一个特殊的配价属性来解决 (Pollard & Sag 1994: Section 9.4)。有关使用这一特征的方法, 请看第 9.6.1 节。

9.1.4 中心语属性的投射

正如在第 1.5 节所讨论的，中心语的某些属性对于整个短语的分布是非常重要的。比如说，动词形式是对于动词的投射分布重要的那些特征。某些动词要求带有特殊形式的动词性论元：

- (15) a. [Dem Mann helfen] will er nicht.
 DET 男人 帮助 想 他不
 ‘他不想帮助这个男人。’
 b. [Dem Mann geholfen] hat er nicht.
 DET 男人 帮助 AUX 他不
 ‘他没帮助这个男人。’
 c. * [Dem Mann geholfen] will er nicht.
 DET 男人 帮助 想 他不
 d. * [Dem Mann helfen] hat er nicht.
 DET 男人 帮助 AUX 他不

wollen (想) 总是带一个不带 *zu* 的不定式，而 *haben* 则要求分词形式的动词。*glauben* (认为) 可以与定式小句共现，但是不能跟不带 *zu* 的不定式共现：

- (16) a. Ich glaube, Peter kommt morgen.
 我 认为 Peter 来 明天
 ‘我认为 Peter 明天会来。’
 b. * Ich glaube, Peter morgen kommen.
 我 认为 Peter 明天 来
 c. * Ich glaube, morgen kommen.
 我 认为 明天 来

这说明动词的投射不能只包括词类的信息，也要包括动词形式的信息。图 9.4 on the following page 在定式动词 *gibt* (给) 的基础上说明了这一点。

GPSG 的中心语特征规约确保了子节点的中心语特征与那些中心语子节点上的特征是一致的。HPSG 语法有一个类似的规则。与 GPSG 不同的是，中心语特征明确地被包含于特征结构中的一组特征。他们被列于路径 *SYNSEM|LOC|CAT|HEAD* 下。(17) 说明了词汇项 *gibt* (给) 的信息：

- (17) *gibt* (给):
- $$\left[\begin{array}{l} \text{word} \\ \text{PHON} \langle \text{gibt} \rangle \\ \text{SYNSEM|LOC|CAT} \left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \text{verb} \\ \text{VFORM } \textit{fin} \end{array} \right] \\ \text{SUBCAT} \langle \text{NP}[\textit{nom}], \text{NP}[\textit{dat}], \text{NP}[\textit{acc}] \rangle \end{array} \right] \end{array} \right]$$

中心语特征原则 (Head Feature Principle) 具有如下的形式：

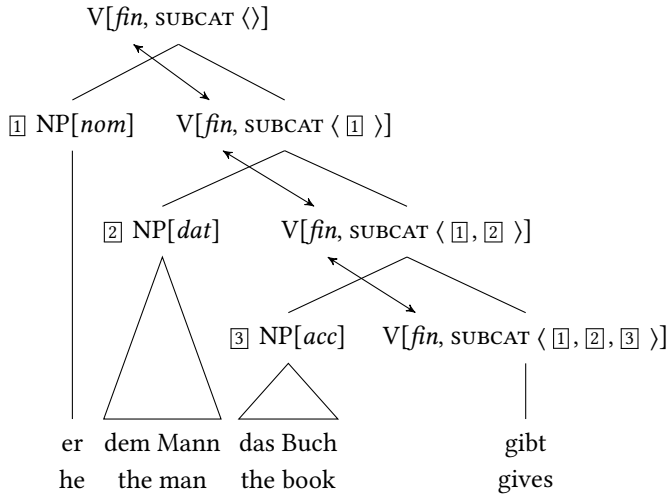


图 9.4: 动词的中心语特征的投射

Principle 1 (中心语特征原则)

任何中心语短语的 **HEAD** 值与其中中心语子节点的 **HEAD** 值是结构共享的。

图 9.5 在下一页是具有结构共享的图 9.4 的变体。

下一节将说明该原则是如何形式化的，以及它是如何整合进 HPSG 的理论框架之中的。

9.1.5 承继层级体系与概说

截至目前，我们已经看到了承继范式的一个例子，在接下来的章节中会有更多的内容，比如说中心语-说明语结构的范式，以及长距离依存问题的解决。中心语特征原则是一个普遍性的原则，它必须满足所有这些模式所允准的结构的要求。正如上面所提及的，所有的结构都必须有一个中心语。在形式上，这可以通过将句法结构分成带有中心语和不带中心语两类来进行区分，并且将类型 *headed-phrase* 赋予到那些具有中心语的结构上。类型 *head-argument-phrase*（第 245 页上的模式 1 的描写类型）是 *headed-phrase* 的一个子类型。某个类型 *x* 的宾语总是具有 *x* 的上位类型的宾语的所有属性。回想第 6.2 节所举的例子：类型 *female person* 的宾语具有类型 *person* 的所有属性。进而，类型 *female person* 的宾语具有额外的、不跟 *person* 的其他子类型共享的更为具体的属性。

如果我们在上位类型上进行约束，那么这会自动影响到它的所有下位类型。由此，中心语特征原则就按照下面的内容来表示：

$$(18) \quad \textit{headed-phrase} \Rightarrow \left[\begin{array}{l} \text{SYNSEM|LOC|CAT|HEAD } [1] \\ \text{HEAD-DTR|SYNSEM|LOC|CAT|HEAD } [1] \end{array} \right]$$

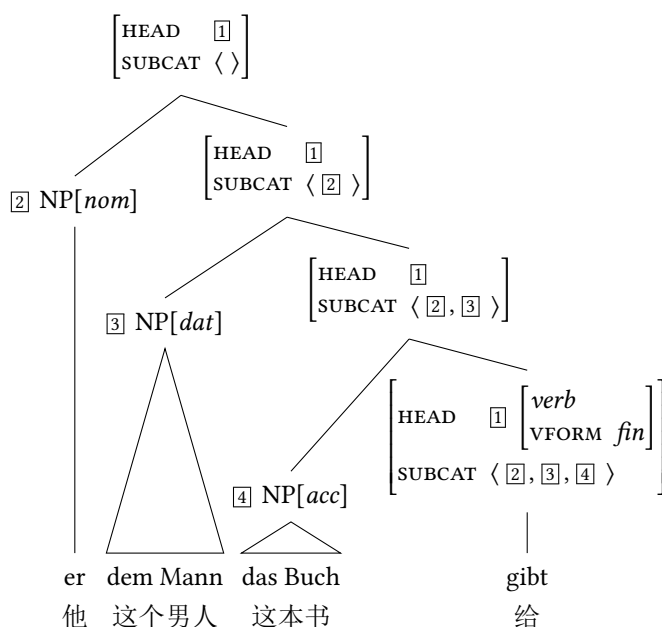
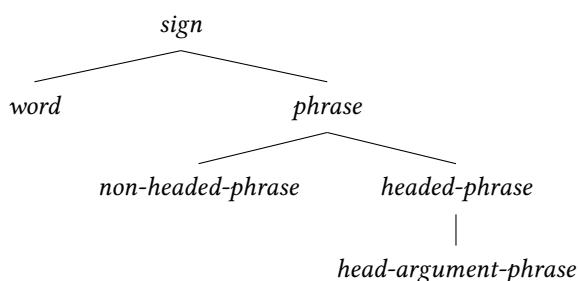


图 9.5: 具有结构共享的动词的中心语特征投射

箭头对应于上面提到的逻辑蕴涵。所以说, (18) 可以这样来解读: 如果一个结构属于类型 *headed-phrase*, 那么它必须满足这样的条件, $\text{SYNSEM|LOC|CAT|HEAD}$ 的值与 $\text{HEAD-DTR|SYNSEM|LOC|CAT|HEAD}$ 的值是相同的。

在 *sign* 下的类型层级的表示如图 9.6 所示。

图 9.6: *sign* 的类型层级: *headed-phrase* 的所有子类型都承袭了约束条件

word 和 *phrase* 是语言符号的子类型。短语可以划分为带有中心语的短语 (*headed-phrase*) 和不带中心语的短语 (*non-headed-phrase*)。还有短语的子类型 *non-headed-phrase* 和 *headed-phrase*。我们已经讨论了 *head-argument-phrase*, 还有 *headed-phrase*

的其他子类型将在后面的章节中详细讨论。与 *word* 和 *phrase* 相似的是，类型 *root* 和 *stem* 也在词汇和形态的结构中起到了重要的作用。限于本书的篇幅，我们不可能在这里深入讨论这些类型，但是可以参考第 22 章的内容。

(19) 中的描写显示了第 245 页的中心语论元模式，还有从 *headed-phrase* 承继而来的类型 *head-argument-phrase* 的限制。

(19) 中心语-论元模式 + 中心语特征原则：

$$\left[\begin{array}{l} \text{head-argument-phrase} \\ \text{SYNSEM|LOC|CAT} \left[\begin{array}{l} \text{HEAD} \quad [1] \\ \text{SUBCAT} \quad [2] \end{array} \right] \\ \text{HEAD-DTR|SYNSEM|LOC|CAT} \left[\begin{array}{l} \text{HEAD} \quad [1] \\ \text{SUBCAT} \quad [2] \oplus \langle [3] \rangle \end{array} \right] \\ \text{NON-HEAD-DTRS} \left\langle [\text{SYNSEM} [3]] \right\rangle \end{array} \right]$$

(20) 给出了由模式 1 允准的结构的描写。与配价信息一样的是，中心语信息在 (20) 中得到了确认，而且中心语特征原则是如何确保特征的投射也是比较明显的：整个结构 (19) 的中心语的值对应于动词 *gibt* (给) 的中心语的值。

$$(20) \left[\begin{array}{l} \text{head-argument-phrase} \\ \text{PHON} \langle \text{das Buch gibt} \rangle \\ \text{SYNSEM|LOC|CAT} \left[\begin{array}{l} \text{HEAD} \quad [1] \\ \text{SUBCAT} \quad [2] \langle \text{NP[nom], NP[dat]} \rangle \end{array} \right] \\ \text{HEAD-DTR} \left[\begin{array}{l} \text{word} \\ \text{PHON} \langle \text{gibt} \rangle \\ \text{SYNSEM|LOC|CAT} \left[\begin{array}{l} \text{HEAD} \quad [1] \left[\begin{array}{l} \text{verb} \\ \text{VFORM fin} \end{array} \right] \\ \text{SUBCAT} \quad [2] \oplus \langle [3] \rangle \end{array} \right] \end{array} \right] \\ \text{NON-HEAD-DTRS} \left\langle \begin{array}{l} \text{PHON} \langle \text{das Buch} \rangle \\ \text{SYNSEM} [3] \left[\begin{array}{l} \text{LOC|CAT} \left[\begin{array}{l} \text{HEAD} \quad \left[\begin{array}{l} \text{noun} \\ \text{CAS acc} \end{array} \right] \\ \text{SUBCAT} \langle \rangle \end{array} \right] \end{array} \right] \end{array} \right\rangle \\ \text{HEAD-DTR ...} \\ \text{NON-HEAD-DTRS ...} \end{array} \right]$$

对于整个句子 *er das Buch dem Mann gibt* (他把这本书给这个男人) 来说，我们得到了由例 (21) 描述的结构 (在图 9.5 中已经显示过了)：

$$(21) \left[\begin{array}{l} \text{SYNSEM|LOC|CAT} \left[\begin{array}{l} \text{HEAD} \quad \left[\begin{array}{l} \text{verb} \\ \text{VFORM fin} \end{array} \right] \\ \text{SUBCAT} \langle \rangle \end{array} \right] \end{array} \right]$$

该描写对应于第 49 页的短语结构语法中的句子符号 S，但是 (21) 还额外地包括了动词形式的信息。

我们将承继模式作为例子来说明我们是如何对语言对象进行概括的，但是，我们也想在理论的其他方面来捕捉这些信息：如范畴语法、HPSG 的词库包括大量的信息。词汇项（根与词）可以被分成不同的类别，进而被赋予不同的类型。按照这一方式，我们可以描写所有的动词、不及物动词和及物动词具有的共同信息。请看第 626 页的图 22.1。

这里我们介绍了 HPSG 理论的一些基本概念，在下面的章节中，我们将分析词的语义是如何表示的，以及短语的意义是如何通过组合性原则来表示的。

9.1.6 语义

在 GB、LFG 和 TAG 这些理论和 HPSG 与 CxG 这些理论中的一个重要的差别在于语言对象的语义内容是按照特征结构来模拟的，这跟所有其他的属性是一样的。正如前面提到的，语义信息在路径 SYNSEM|LOC|CONT 下。(22) 给出了 *Buch* (书) 的 CONT value 的例子。这一表示是基于最简递归语义 (MRS) 的：⁹

$$(22) \quad \left[\begin{array}{l} mrs \\ IND \quad [1] \quad \left[\begin{array}{l} PER \quad 3 \\ NUM \quad sg \\ GEN \quad neu \end{array} \right] \\ RELS \quad \left\langle \left[\begin{array}{l} buch \\ INST \quad [1] \end{array} \right] \right\rangle \end{array} \right]$$

IND 表示标引，RELS 是关系的列表。诸如人称、数和性的特征是名词性标引的一部分。iscale 这些指标在决定指称或共指关系中是非常重要的。比如说，(23) 中的 *sie* (她) 指称 *Frau* (女人)，但是不指 *Buch* (书)。另一方面，*es* (它) 不能指称 *Frau* (女人)。

- (23) Die Frau_i kauft ein Buch_j. Sie_i liest es_j.
 DET 女人 买 一 书 她 读 它
 ‘这个女人买了一本书。她在读它。’

通常来说，代词必须在人称、数和性上与其所指代的成分相一致。相应的标引需要保持一致。在 HPSG 中，这点通过结构共享来实现。也可以说是共指关系 (coindexation)。(24) 给出了反身代词的共指关系的一些例子：

- (24) a. Ich_i sehe mich_i.
 我 看见我自己
 b. Du_i siehst dich_i.
 你 看见 你自己

⁹ Pollard & Sag (1994) 和 Ginzburg & Sag (2000) 利用了情景语义学 (Barwise & Perry 1983; Cooper, Mukai & Perry 1990; Devlin 1992)。另一种已经在 HPSG 理论中应用的方法是词汇资源语法 (Richter & Sailer 2004)。有关 HPSG 理论中早期的未充分分析请看 Nerbonne (1993)。

- c. Er_i sieht sich_i.
他 看见 他自己
- d. Wir_i sehen uns_i.
我们看见 我们自己
- e. Ihr_i seht euch_i.
你们看见你们自己
- f. Sie_i sehen sich_i.
他们看见 他们自己

共指的哪个部分是可能的还是必须的这个问题是由约束理论决定的。Pollard & Sag (1992, 1994)指出, HPSG 中的约束理论在实现约束关系时并没有像 GB 理论中关于树的构型问题那样引起许多的问题。但是, HPSG 理论中的约束理论还是有一些公开的问题的 (Müller 1999a: Section 20.4)。

(25) 给出了动词 *geben* (给) 的 CONT value 信息:

$$(25) \left[\begin{array}{l} mrs \\ IND \quad [1] \quad event \\ RELS \quad \left\langle \begin{array}{l} geben \\ EVENT \quad [1] \\ AGENT \quad index \\ GOAL \quad index \\ THEME \quad index \end{array} \right\rangle \end{array} \right]$$

一般认为, 带有 *event* 类型的事件变量的动词是在 IND 下表示的, 这跟名词对象的标引是一样的。截至目前, 我们没有将配价列表中的元素指派给语义表示中的论元角色。这一联系叫做联接 (linking)。(26) 说明了 HPSG 理论中, 联接是如何运作的。名词短语的论元的指称标引与中心语决定的关系的语义角色中的一种情况是结构共享的。

(26) *gibt* (给):

$$\left[\begin{array}{l} CAT \quad \left[\begin{array}{l} HEAD \quad \left[\begin{array}{l} verb \\ VFORM \quad fin \end{array} \right] \\ SUBCAT \quad \left\langle NP[nom]_{[1]}, NP[dat]_{[2]}, NP[acc]_{[3]} \right\rangle \end{array} \right] \\ CONT \quad \left[\begin{array}{l} mrs \\ IND \quad [4] \quad event \\ RELS \quad \left\langle \begin{array}{l} geben \\ EVENT \quad [4] \\ AGENT \quad [1] \\ GOAL \quad [2] \\ THEME \quad [3] \end{array} \right\rangle \end{array} \right] \end{array} \right]$$

因为我们使用诸如施事 (agent) 和受事 (patient) 这样的术语来表示论元角色, 我们就可以说明配价类型的概况和论元角色的实现。比如说, 我们将动词分成带有一个施事的动词、带有一个施事和主题的动词, 以及带有施事和受事的动词等。这些不同的配价或联接模式可以在类型层级体系中表示, 而且这些类别可以被指派到具体的词汇项上, 即我们可以让他们继承各自类型的约束条件。带有施事、主题和目标的动词的类型的约束条件可以按照 (27) 中的形式来表示:

$$(27) \left[\begin{array}{c} \text{CAT|SUBCAT} \left\langle \boxed{}, \boxed{}, \boxed{} \right\rangle \\ \left[\begin{array}{c} \text{mrs} \\ \text{IND} \quad \boxed{4} \text{ event} \\ \text{CONT} \left[\begin{array}{c} \text{RELS} \left\langle \begin{array}{c} \text{agent-goal-theme-rel} \\ \text{EVENT} \quad \boxed{4} \\ \text{AGENT} \quad \boxed{1} \\ \text{GOAL} \quad \boxed{2} \\ \text{THEME} \quad \boxed{3} \end{array} \right\rangle \end{array} \right] \end{array} \right] \end{array} \right]$$

$\boxed{1}$ 表示带有标引 $\boxed{1}$ 的未被指定的句法范畴的对象。具有 *geben'* 关系的类型是 *agent-goal-theme-rel* 的一个子类型。在 (27) 中, 是词 *geben* (给) 的词汇项, 而不是词根 *geb-* 具有联接的范式。更多有关 HPSG 理论中的联接理论, 请看 Davis (1996)、Wechsler (1991) 和 Davis & Koenig (2000)。

目前, 我们只看到了词汇项的意义是如何表示的。语义原则决定了短语的语义贡献的计算: 整个表达式的标引对英语中心语子节点的标引, 而且整个符号的 **RELS value** 对应于子节点的 **RELS values** 的加合, 以及由统制模式引入的任何关系。最后一点很重要, 这是因为模式可以对语义有所贡献的假设可以说明有些情况下一个短语的整体意义不仅仅只是其组成成分的简单加合。与之相关的例子经常在构式语法中进行讨论。HPSG 理论主张语义的组合性, 即表示特定模式的语义部分可以整合进一段话语的整体意义之中。例子请参考第 21.10 节。

动词及其论元的语义贡献之间的联系是在词汇项中建立起来的。这样, 我们可以确保动词的论元角色被指派到句中正确的论元上面。但是, 这并不是语义所担负的唯一责任。它必须能够生成不同的意义解读, 这与量词辖域的歧义 (请看第 84 页), 以及其他谓词下面谓词的语义嵌套的处理是有关系的。所有这些要求都在 MRS 中得到满足。受限于篇幅, 我们不展开说明。读者可以参考 Copestake, Flickinger, Pollard & Sag (2005) 的文章, 以及讨论部分中第 19.3 节的内容。

9.1.7 说明语

与中心语通过 SUBCAT 来选择论元类似的是, 说明语也可以通过使用特征 (MODIFIED) 来选择他们的中心语。修饰名词和关系小句的形容词和介词短语选择一个几乎完整的名词性投射, 即一个名词只需要与限定词相组合以构成一个完整的 NP。(28) 显示了每个 *synsem* 对象的描写。与 \bar{X} 理论中 (请看第 2.5 节) 类似的符号 \bar{N} 表示这个特征描写的缩写形式。

(28) AVM 被简写为 \bar{N} :

$$\left[\text{CAT} \left[\begin{array}{ll} \text{HEAD} & \textit{noun} \\ \text{SUBCAT} & \langle \textit{DET} \rangle \end{array} \right] \right]$$

(29) 显示了 *interessantes* (有趣的) 的部分词汇项:

(29) *interessantes* (有趣的) CAT value :

$$\left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{l} \textit{adj} \\ \text{MOD } \bar{\text{N}} \end{array} \right] \\ \text{SUBCAT} & \langle \rangle \end{array} \right]$$

interessantes (有趣的) 是一个形容词, 它不带任何论元成分, 由此它有一个空的 SUBCAT list。诸如 *treu* (忠诚的) 这样的形容词在他们的 SUBCAT list 中会有一个与格 NP。

(30) *ein dem König treues Mädchen*
 一 DET.DAT 国王 忠诚的女孩
 ‘对国王忠诚的女孩儿’

在 (31) 中可以看到 CAT value :

(31) *treues* (忠诚的) 的 CAT value :

$$\left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{l} \textit{adj} \\ \text{MOD } \bar{\text{N}} \end{array} \right] \\ \text{SUBCAT} & \langle \text{NP}[\textit{dat}] \rangle \end{array} \right]$$

dem König treues (对国王的忠诚) 构成了一个形容词短语, 它修饰 *Mädchen* (女孩儿) 这个词。

与属于 CAT 的选择性特征 SUBCAT 不同的是, MOD 是一个中心语特征。原因是选择修饰中心语的特征必须在说明语的最大投射中出现。 $\bar{\text{N}}$ -短语 *dem König treues* (对国王的忠诚) 必须包括在整个 AP 的表达式中, 就像它在词汇层面的 (29) 中的形容词的词汇项中出现一样。形容词短语 *dem König treues* 具有跟基本形容词 *interessantes* (有趣的) 一样的句法属性。

(32) *dem König treues* 的 CAT value :

$$\left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{l} \textit{adj} \\ \text{MOD } \bar{\text{N}} \end{array} \right] \\ \text{SUBCAT} & \langle \rangle \end{array} \right]$$

因为 MOD 是一个中心语特征, 中心语特征原则会保证整个投射的 MOD value 与 *treues* (忠诚的) 的词汇项的 MOD value 是一致的。

作为修饰语所选择的中心语的另一种情况是，我们可以假定在中心语自身上的所有可能的说明语。这一观点由Pollard & Sag (1987: 161)提出。Pollard & Sag (1994: Section 1.9) 对前面的分析进行了修订，因为不能说明修饰语的语义。¹⁰

图 9.7表示了中心语-说明语结构中的选择信息。

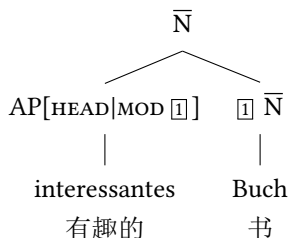


图 9.7: 中心语-说明语结构（选择）

中心语-说明语结构是由模式 2 所允准的。

Schema 2 (中心语-说明语模式（初级版本）)

head-adjunct-phrase \Rightarrow

$$\left[\begin{array}{l} \text{HEAD-DTR} | \text{SYNSEM } \boxed{1} \\ \text{NON-HEAD-DTRS} \left\langle \left[\begin{array}{l} \text{SYNSEM} | \text{LOC} | \text{CAT} \left[\begin{array}{l} \text{HEAD} | \text{MOD } \boxed{1} \\ \text{SUBCAT} \langle \rangle \end{array} \right] \end{array} \right] \right\rangle \end{array} \right]$$

说明语 ($\boxed{1}$) 的选择性特征的值与中心语子节点的 **SYNSEM value** 是相同的。这样就可以确保中心语子节点具有说明语所确定的属性。非中心语子节点的 **SUBCAT value** 是空列表，这也就是为什么只有完全饱和的说明语允许出现在中心语-说明语结构中。这样，诸如 (33b) 的短语就被规则排除出去了：

- (33) a. die Wurst in der Speisekammer
 DET 香肠 PREP DET 食品箱
 b. *die Wurst in
 DET 香肠 PREP

例 (33a) 需要进一步的解释。如 (33a) 中使用的介词 *in* 具有下面的 **CAT value**：

(34) *in* 的 **CAT value** :

$$\left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \text{prep} \\ \text{MOD } \bar{N} \end{array} \right] \\ \text{SUBCAT} \langle \text{NP}[\text{dat}] \rangle \end{array} \right]$$

¹⁰ 不过，可以参考 Bouma, Malouf & Sag (2001a)。Bouma, Malouf & Sag (2001a) 提出了一个整合的分析，其中说明语可以选择中心语，说明语也可以被中心语所选择。最简递归语义是支持这一分析的语义理论。应用这种语义分析方法，就可以避免由 Pollard & Sag (1987) 带来的修饰语的语义问题。

在将 *in* 和名词短语 *der Speisekammer*（食品箱）相组合后，我们会得到：

(35) *in der Speisekammer*（在食品箱里）的 CAT value：

$$\left[\begin{array}{c} \text{HEAD} \left[\begin{array}{c} \text{prep} \\ \text{MOD } \bar{N} \end{array} \right] \\ \text{SUBCAT } \langle \rangle \end{array} \right]$$

该表达式对应于形容词 *interessantes*（有趣的），而且也可按照同样的方式来使用，如忽略 PP 的位置：PP 修饰 \bar{N} 。

那些只能用作论元，不能修饰任何成分的中心语具有 *none* 的 MOD value。这样，他们就不会出现在中心语说明语结构中的非中心语子节点的位置上了，因为中心语子节点的 MOD value 需要与中心语子节点的 SYNSEM value 兼容。

9.2 被动

HPSG 理论 遵循 Bresnan 的思想（请看第 7.2 节），并将被动放在词汇中处理。¹¹一条词汇规则将词根作为输入，并允准了分词形式并且最凸显的论元（所谓的指定论元）受到了抑制。¹²因为语法功能并不属于 HPSG 理论中的一部分，我们不需要任何映射的原则来将宾语映射到主语上。无论如何，我们还是要解释被动下格的变化。如果有人在词汇项中完整地区分了指定论元的格，那么这个人就需要确保及物动词的宾格论元在被动式中被实现为主格。(36) 展示了这样的词汇规则是什么样子的：

(36) 从 Kiss (1992) 而来的人称被动的词汇规则：

$$\left[\begin{array}{c} \text{stem} \\ \text{PHON } \boxed{1} \\ \text{SYNSEM|LOC|CAT} \left[\begin{array}{c} \text{HEAD } \text{verb} \\ \text{SUBCAT } \langle \text{NP}[\text{nom}], \text{NP}[\text{acc}]_{\boxed{2}} \rangle \oplus \boxed{3} \end{array} \right] \end{array} \right] \mapsto \left[\begin{array}{c} \text{word} \\ \text{PHON } f(\boxed{1}) \\ \text{SYNSEM|LOC|CAT} \left[\begin{array}{c} \text{HEAD } [\text{VFORM } \text{passive-part}] \\ \text{SUBCAT } \langle \text{NP}[\text{nom}]_{\boxed{2}} \rangle \oplus \boxed{3} \end{array} \right] \end{array} \right]$$

¹¹ 有些例外是受到构式语法影响的一些分析，如 Tseng (2007) 和 Haugereid (2007)。但是，这些方法是有问题的，因为他们无法解释 Bresnan 的形容词性被动。对于 Haugereid 的分析的其他问题，请看 Müller (2007c) 和第 21.3.6 节。

¹² 更多有关指定论元的内容请看 Haider (1986a)。德语中被动的 HPSG 分析相当程度上受到了 Haider 的影响。Haider 使用指定论元来模拟所谓的非宾格和非作格动词之间的区别 (Perlmutter 1978)：非宾格动词与非作格动词和及物动词的区别在于他们没有指定论元。我们在这里不列出非宾格方面的文献。读者可以去看 Haider 的原始研究以及 Müller (2007b) 中与被动有关的内容。

词汇规则将动词词根¹³ 作为它的输入，这就要求有一个主格论元、一个宾格论元以及其他可能的论元成分（如果③不是一个空列表的话），并且允准一个需要带有主格论元和③中可能论元成分的词汇项。¹⁴ 该词汇规则的输入指定了输出词的 *VFORM* value。这是非常重要的，因为助动词和核心动词必须一起出现。比如说，不能用完成分词来取代被动分词，因为在 Kiss 的理论中，他们的格是不同的。

- (37) a. Der Mann hat den Weltmeister geschlagen.
 DET 男人 AUX DET 世界. 冠军 打
 ‘这个男人打了世界冠军。’
 b. *Der Mann wird den Weltmeister geschlagen.
 DET 男人 AUX DET 世界. 冠军 打
 c. Der Weltmeister wird geschlagen.
 DET 世界. 冠军 AUX 打
 ‘世界冠军被人打了。’

词汇规则的解释有一些规定：在输入符号中没有提及的所有信息都被输入符号替代了。这样，动词的意义在被动规则中没有被提及，这就使得被动规则是一个保留意义的规则。输入和输出的 *CONT* values 没有在规则中提及，所以是相同的。这里重要的是它们保留的联接信息。比如说，以动词词根 *schlag-*（打）所应用的规则为例：

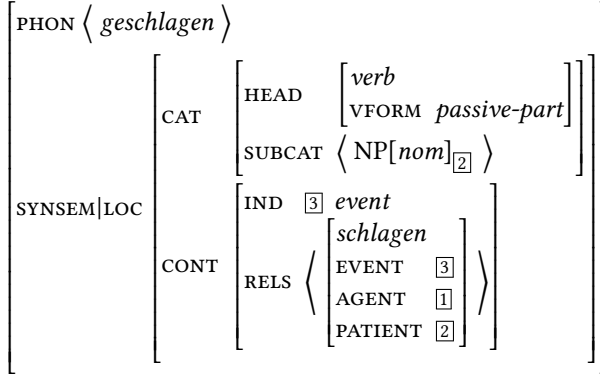
- (38) a. *schlag-*（打）的输入：

SYNSEM LOC	CAT	PHON \langle <i>schlag</i> \rangle	
		HEAD <i>verb</i>	
		SUBCAT \langle NP[<i>nom</i>] _① , NP[<i>acc</i>] _② \rangle	
		IND ③ <i>event</i>	
CONT	RELS	<i>schlagen</i>	
		EVENT ③	
		AGENT ①	
		PATIENT ②	

¹³ 术语 *stem* 包括词根 (*help-*，“帮助”)、派生词 (*besing-*，“唱”) 和复合词。这样词汇规则就可以用在像 *help-* 的词根和诸如 *besing-* 的派生形式之中了。

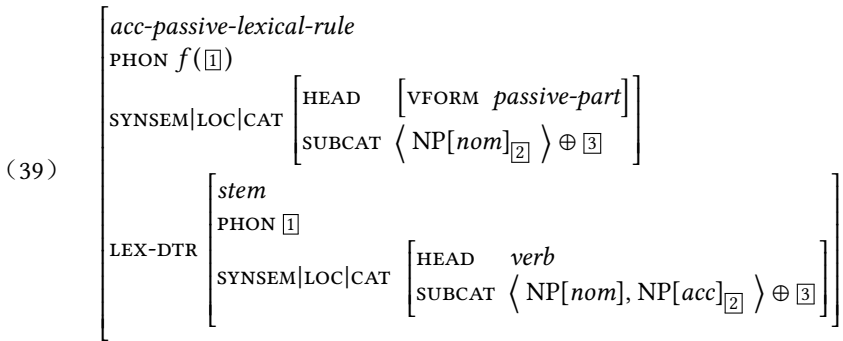
¹⁴ 该规则假定了双及物动词的论元是按照主格、宾语和与格的顺序排列的。在本章中，我假设了主格、与格和宾格的顺序，这对应于德语论元的未标记语序。Kiss (2001)指出，未标记语序的表示可以用来说明德语的辖域事实。而且，论元的顺序对应于英语的顺序，这在捕捉跨语言的共性方面是具有优势的。在早先的工作中，我认为语序是主格、宾格和与格这样排列的，因为这个顺序表示了凸现的层级体系，而这在德语语法的大部分方面都是具有相关性的。例子有：省略 (Klein 1985)、话题省略 (Fries 1988)、自由关系小句 (Bausewein 1990; Pittner 1995; Müller 1999b)、描述性次级谓词 (Müller 2004d, 2002a, 2008a)、约束理论 (Grewendorf 1985; Pollard 和 Sag: 1992, 1994: Chapter 6)。这一语序也对应于 Keenan & Comrie (1977) 和 Pullum (1977) 提出的旁格的层级体系。为了说明这一层级体系，需要提出一个带有主格、宾格和与格语序的列表。下面将要提出的被动规则与这两种论元语序都是相容的。

b. *geschlagen* (被打) 的输出:



施事的角色联接到 *schlag-* 的主语。在被动之后, 主语受到抑制, 联接到 *schlag-* 的受事角色上的论元成为分词的主语。论元联接没有受到这个影响, 并且名词性论元被正确地指派到受事角色上。

正如Meurers (2001)所指出的, 词汇规则也可以捕捉到特征描写。(39) 给出了(36)的特征描写表示。在(36)的规则的手左边囊括进了(39)中的 **LEX-DTR**



值。因为这类词汇规则被完整地整合到形式化系统中, 对应于这些词汇规则的特征结构也有他们自己的类型。如果一个给定规则的应用结果是一个屈折变化的词, 那么这个词规则的类型(我们所举的例子中的 *acc-passive-lexical-rule*)就是 *word* 的次类型。由于词汇规则具有类型, 就可以对词汇规则进行概括。

目前我们讨论的词汇规则适用于人称被动。但是, 对于非人称被动, 我们就需要第二条词汇规则了。而且, 我们需要针对被动和完成时准备两条不同的词汇项, 尽管在德语中他们的形式是一样的。在下面, 我将讨论被动的理论所需的基本假设, 该假设可以充分地解释人称被动和非人称被动, 并且只需要用分词形式的一个词汇项来说明就足够了。

9.2.1 配价信息与格原则

在第 3.4.1 节, 结构格与结构之间是有差异的。在 HPSG 的文献中, 一般按照 Haider (1986a) 的观点, 认为与格是一个词汇格。对于带有词汇格标记的论元来说, 他们的格的值直接在论元的描写中有所表示。带有结构格的论元也在词汇中说明了, 他们带有结构格, 只不过没有给出真正的格的值。为了保证语法不会得到任何错误的结论, 它必须要确保结构格根据他们出现的语境而得到一个独一无二的值。这点由格原则来处理:¹⁵

Principle 2 (格原则)

- 在一个包括主语和补足语的动词中心语的列表中, 第一个带有结构格的成分是主格。
- 该列表中的所有其他带有结构格的成分是宾格。
- 在主格的上下文中, 带有结构格的成分被赋予了属格。

例 (40) 给出了定式动词的原型配价列表:

- (40) a. *schläft* (睡觉): SUBCAT $\langle \text{NP}[\text{str}]_j \rangle$
 b. *unterstützt* (支持): SUBCAT $\langle \text{NP}[\text{str}]_j, \text{NP}[\text{str}]_k \rangle$
 c. *hilft* (帮助): SUBCAT $\langle \text{NP}[\text{str}]_j, \text{NP}[\text{ldat}]_k \rangle$
 d. *schenkt* (给): SUBCAT $\langle \text{NP}[\text{str}]_j, \text{NP}[\text{ldat}]_k, \text{NP}[\text{str}]_l \rangle$

str 表示词汇与格的 *structural* 和 *ldat*。格原则确保了上面所列动词的主语实现为主格, 而带有结构格的宾语被赋予了宾格。

对于结构格和词汇格之间的区别, 我们可以构造一个被动—词汇规则来表示人称被动和非人称被动:

- (41) 人称被动和非人称被动的词汇规则 (简化版):

$$\left[\begin{array}{l} \text{stem} \\ \text{PHON } \boxed{1} \\ \text{SYNSEM|LOC|CAT } \left[\begin{array}{ll} \text{HEAD} & \text{verb} \\ \text{SUBCAT} & \langle \text{NP}[\text{str}] \rangle \oplus \boxed{2} \end{array} \right] \end{array} \right] \mapsto \left[\begin{array}{l} \text{word} \\ \text{PHON } f(\boxed{1}) \\ \text{SYNSEM|LOC|CAT } \left[\begin{array}{ll} \text{HEAD} & \left[\text{VFORM } \textit{ppp} \right] \\ \text{SUBCAT} & \boxed{2} \end{array} \right] \end{array} \right]$$

¹⁵ 这里的格原则被简化了。所谓的“升”格需要特殊的处理。更多细节可以参考 Meurers (1999c)、Przepiórkowski (1999a) 和 Müller (2007b: Chapter 14, Chapter 17)。这些著作中给出的格原则与 Yip, Maling & Jackendoff (1987) 提出的理论非常相似, 由此该原则也可以解释他们的工作中所讨论的语言的格系统, 尤其是爱尔兰语复杂的格系统。

这条词汇规则真正做到了我们在被动的预备理论的方面上所期待的功能：它抑制了带有结构格的最凸显的论元，即对应于主动句的主语的论元。动词—助动词结构的标准分析认为主动词和助动词构成了一个动词性复杂结构 (Hinrichs & Nakazawa 1994; Pollard 1994; Müller 1999a, 2002a; Meurers 2000; Kathol 2000)。嵌套的论元被助动词替代。在加入被动助动词分词后，我们可以得到如下的 SUBCAT 列表：

- (42) a. *geschlafen wird* (被睡觉) : SUBCAT $\langle \rangle$
 b. *unterstützt wird* (被支持) : SUBCAT $\langle \text{NP}[\text{str}]_k \rangle$
 c. *geholfen wird* (被帮助) : SUBCAT $\langle \text{NP}[\text{ldat}]_k \rangle$
 d. *geschenkt wird* (被给) : SUBCAT $\langle \text{NP}[\text{ldat}]_k, \text{NP}[\text{str}]_l \rangle$

(42) 与 (40) 是不同的，因为首位的 NP 是不同的。如果 NP 具有结构格，它就会得到主格。如果没有带结构格的 NP，如例 (42c) 所示，那么格不会变化，即由词汇确定的。

我们在这儿无法得到完美的分析。不过，需要指出的是，对于与分词相同的词汇项被用于 (43)。

- (43) a. Er hat den Weltmeister geschlagen.
 他 AUX DET 世界. 冠军 打
 ‘他把世界冠军打了。’
 b. Der Weltmeister wurde geschlagen.
 DET 世界. 冠军 AUX 打
 ‘世界冠军被打了。’

助动词决定了哪些论元被实现了 (Haider 1986a, Müller 2007b: Chapter 17)。(41) 中的词汇规则允准了可以用在被动和完成式中的形式。这样，vFORM value 属于 *ppp*，它表示完成式被动分词 (participle perfect passive)。

我们应该注意到，该分析适用于没有成分移动的被动。这里没有涉及 GB 分析中的问题。论元的重新排序 (请看第 9.4 节) 是独立于被动化的。与 GPSG、Categorial Grammar 或 Bresnan 的 LFG 分析不同的是，在词汇映射理论 (请看第 208 页) 引入之前完全没有提及宾格宾语。被动可以直接分析为主语的抑制。任何别的成分都与语法的其他原则具有互动关系。

9.3 动词位置

我这里要说明的动词位置分析是基于 GB 分析的。在 HPSG 中，有许多不同的方法来描述动词的位置，但是，依我看，GB 分析的 HPSG 变体是唯一合适的 (Müller 2005b,c, 2015b)。(44) 的分析可以总结如下：在动词首位的小句中，动词末位上有一个语迹。在首位的位置上有一个动词的特殊形式选择了动词语迹的投射。特殊的词汇项由词汇规则允准。动词和语迹之间的连接被看作是 GPSG 中的长距离依存问题，并通过树中的信息或特征结构 (结构共享) 的识别来实现。

- (44) Kennt_k jeder diesen Mann _{-k}?
 认识 每人 这 男人
 ‘每个人都认识这个男人吗?’

图 9.8给出了这一问题的整体情况。位于动词末位的动词语迹跟动词在句法和语义

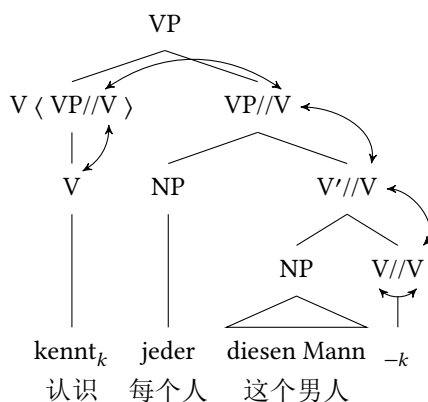


图 9.8: HPSG 中动词位置的分析

上非常相似。缺失动词的信息表示为特征 DOUBLE SLASH (缩写为: DSL) 的值。这是一个中心语特征, 并且传递到了最大投射上 (VP)。位于首位的动词在它的缺失了动词 (VP//V) 的 SUBCAT list 中包含一个 VP。这是跟规则的输入中的同样的动词, 它正常出现在末位上。在图 9.8中, 由两个最大的动词投射: 作为中心语带有语迹的 *jeder diesen Mann*_{-k} 和作为中心语带有 *kennt* 的 *kennt jeder diesen Mann*_{-k}。

该分析将在下面的内容中得到更多细节上的解释。对于图 9.8中的语迹来说, 我们需要假定 (45) 中的词汇项。

(45) *kennt* (认识) 的动词语迹:

PHON <>	CAT	HEAD	[<i>verb</i> <i>VFORM fin</i>]	SUBCAT < NP[nom] _[1] , NP[acc] _[2] >
SYNSEM LOC	CONT	IND	[3]	RELS < [<i>kennen</i> EVENT [3] EXPERIENCER [1] THEME [2]] >

这个词汇项与普通动词 *kennt* 的区别只在于它的 PHON value。带有语迹的分析的句法过程如图 9.9 on the next page所示。

将语迹与 *diesen Mann* (这个男人) 和 *jeder* (每个人) 的组合遵守了我们目前提到的规则与原则。这就要求我们立刻回答是什么允准了图 9.9中的动词以及它具有的地位。

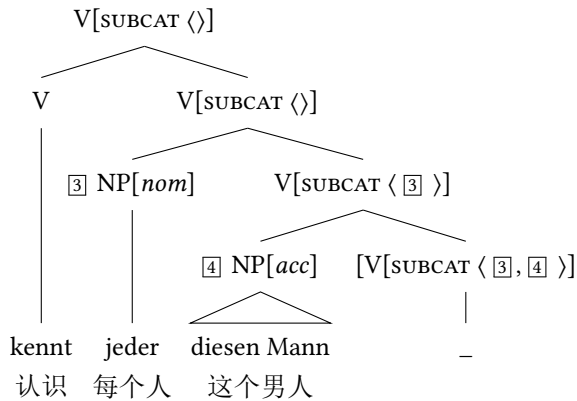


图 9.9: *Kennt jeder diesen Mann?* (每个人都认识这个男人吗?) 的分析

如果我们想要捕捉到这样的事实，位于首位的定式动词像一个补足语的话 (Höhle 1997)，那么就可以给予图 9.9 中 *kennt* 以中心语地位，而且允许 *kennt* 选择一个饱和的、动词位于末位的动词投射。位于首位的定式动词与补足语是不同的，因为他们需要一个动词语迹的投射，而补足语需要显性动词的投射。

- (46) a. *dass* [*jeder diesen Mann kennt*]
 CONJ 每人 这 男人 认识
 ‘每个人都认识这个男人’
 b. *Kennt* [*jeder diesen Mann _*]
 认识 每人 这 男人
 ‘每个人都认识这个男人吗?’

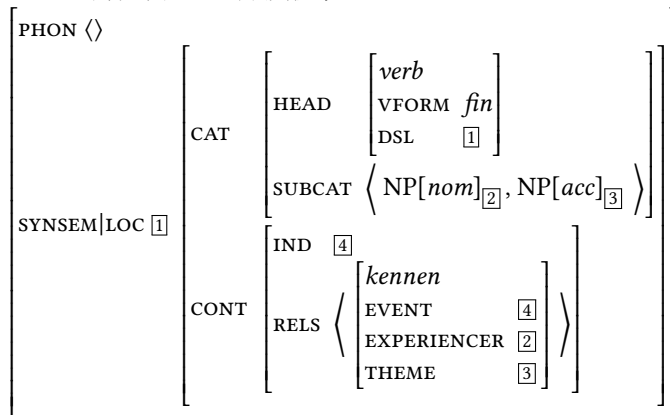
通常来说，并不是 *kennen* (认识) 选择一个完整的句子，而且没有其他任何东西对于分析 (46b) 中的作为中心语的 *kennt* 来说是必要的。进而，我们必须确保带有 *kennt* 的动词投射与包括属于 *kennt* 的动词投射相组合。比如说，如果它能包括一个属于 *gibt* (给) 的语迹，那么我们就可以分析 (47b) 中的句子了：

- (47) a. *Gibt* [*der Mann der Frau das Buch _gibt*]?
 给 DET 男人 DET 女人 DET 书
 ‘这个男人给这个女人书了吗?’
 b. **Kennt* [*der Mann der Frau das Buch _gibt*]?
 认识 DET 男人 DET 女人 DET 书

在前面的讨论中，前置动词和动词语迹的依存关系通过同指来表示。在 HPSG 中，同指总是由结构共享来实现的。位于首位的动词必须要求语迹具有确切地具有动词应该具有的那些属性，即它位于末尾。由此，必须共享的信息都是局部相关的句法和语义信息，即在 *LOCAL* 下的所有信息。因为 *PHON* 不是 *LOCAL* 特征的一部分，它没有被共享，这就是为什么语迹的 *PHON* 值和动词的值可以不同的原因。截至目前，该分析中有一个重要的细节缺失了：语迹的 *LOCAL* 值不能与首位动词的

要求直接结构共享, 因为动词 *kennt* 只能选择语迹的投射的属性, 而被选择的投射的 SUBCAT list 是空列表。这就导致 (47b) 的讨论中所指出的问题。由此, 必须保证动词语迹的所有信息在它的投射的最高点上可获得。这可以通过中心语特征的引入来获得, 它的值与语迹的 LOCAL value 值是相同的。这个特征被称为 DSL。正如在上面已经提及的, DSL 表示双重斜杠 (double slash)。它被这样称呼的原因是, 它有一个跟 SLASH feature 相似的功能, 我们可以在后面的章节中讲到这个功能。¹⁶ (48) 显示了动词语迹的修订版本:

(48) *kennt* 的动词语迹 (初级版本):



通过对 (48) 中的 LOCAL value 和 DSL 的值的共享, 动词语迹的句法和语义信息在它的最大投射上表示出来, 而且位于首位的动词可以核查这个语迹的投射是否是兼容的。¹⁷

对于动词位于首位的具体的词汇项由下面的词汇规则允准:¹⁸

¹⁶ 特征 DSL 是由 Jacobson (1987a) 在范畴语法的框架下提出用来描写英语倒装的中心语移动的。Borsley (1989) 采用了这一观点, 并将其译为 HPSG 的术语, 这样就可以看到, 在 CP/IP 系统的 HPSG 变体中, 中心语移位是如何用 DSL 模拟的。DSL 特征在 HPSG 的中心语移位过程的引入是由这样的事实驱动的, 与第 9.5 节讨论的长距离依存不同的是, 这类移位是局部的。这种将动词语迹作为中心语的部分的信息的渗透来自于 Oliva (1992)。

¹⁷ 需要注意的是, (48) 中的描写是循环的, 因为标签 \square 用于它自己内部。请看第 6.5 节关于循环特征描写的内容。循环描写是用来表示语言对象的最为直接的方式, 这个语言对象带有缺失的局部属性, 并且将这些信息作为 DSL feature 的值按照中心语路径传递上去。当我们看到第 265 页上的 (50) 的动词语迹的最终版本时, 这种思路会显得更为清晰。

¹⁸ 这个词汇规则分析不能解释 (i) 这类的句子:

- (i) Karl kennt und liebt diese Schallplatte.
Karl 认识 和 爱 这 专辑

这与词汇规则不能用于并列的结果是有关联的, 它构成了一个复杂的句法对象。如果我们将词汇规则分别应用到每个动词上的话, 那么我们会得到动词的不同类型, 他们分别选择 *kennen* (认识) 和 *lieben* (爱) 的动词语迹。由于连词的 CAT values 在并列中互相指认, 包括 *kennt* 和 *liebt* 的 V1 的变体的并列可以被排除出去, 因为被选择的 VPs 的 DSL values 包括各自动词的意义, 而且是不兼容的 (Müller 2005b: 13)。除了词汇规则, 我们需要假定一个一元的句法规则, 它应用到短语 *kennt und liebt* (认识和爱) 上。正如我们看到的, 这里假定的 HPSG 形式化的词汇规则对应于一元规则, 这样 (49) 和相应的句法规则的区别很大程度上是表达上的差异。

(49) 位于首位的动词的词汇规则:

$$\left[\begin{array}{c} \text{SYNSEM|LOC } \boxed{1} \\ \text{CAT|HEAD} \left[\begin{array}{c} \text{verb} \\ \text{VFORM } \textit{fin} \\ \text{INITIAL } - \end{array} \right] \end{array} \right] \mapsto \left[\begin{array}{c} \text{SYNSEM|LOC|CAT} \\ \text{HEAD} \left[\begin{array}{c} \text{verb} \\ \text{VFORM } \textit{fin} \\ \text{INITIAL } + \\ \text{DSL } \textit{none} \end{array} \right] \\ \text{SUBCAT} \left\langle \left[\begin{array}{c} \text{LOC|CAT} \\ \text{HEAD} \left[\begin{array}{c} \text{verb} \\ \text{DSL } \boxed{1} \end{array} \right] \\ \text{SUBCAT } \langle \rangle \end{array} \right] \right\rangle \end{array} \right]$$

由这个词汇规则允许的动词选择了动词语迹的最大投射，它跟输入动词具有相同的局部属性。这是由输入动词的 LOCAL 和选择的动词投射的 DSL 值的共指来实现的。只有末位 (INITIAL-) 中的定式动词可以作为这一规则的输入。输出是位于首位 (INITIAL+) 的动词。相应的扩展分析如图 9.10 所示。V1-LR 表示动词首位的词汇规则。

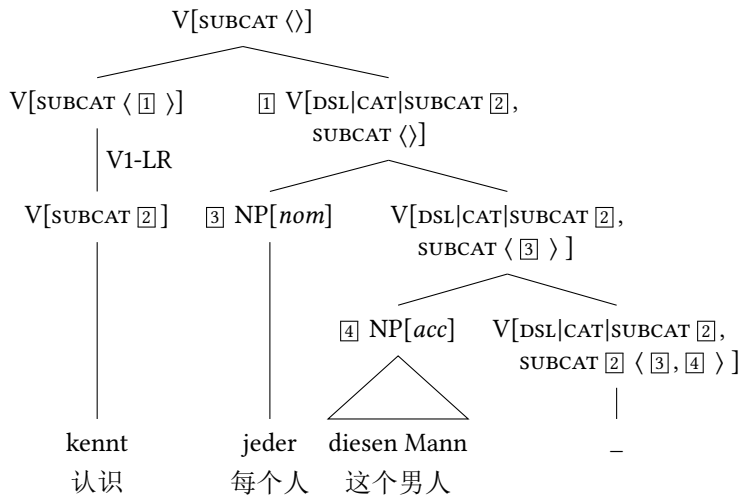


图 9.10: *Kennt jeder diesen Mann?* (每个人都认识这个男人吗?) 的分析的可视化

(49) 中的词汇规则允许了选择 VP (图 9.10 中的 1) 的动词。这个 VP 的 DSL value 对应于动词的 LOCAL value，它是词汇规则的输入。DSL value 的部分也是图 9.10 (2) 中表示的定价信息。因为 DSL 是一个中心语特征，VP 的 DSL value 与动词语迹是相同的，而且由于动词语迹的 LOCAL value 与 DSL value 是相同的，动词 *kennen* 的

SUBCAT 信息在语迹中也是可获得的。语迹与其论元的组合跟普通动词相比是完全一样的。

如果我们必须为每个动词假定一个具体的语迹，这样就不令人满意了。幸运的是，这是不必要的，因为 (50) 中的一个普通的语迹对于带有动词移位的句子的分析来说是足够的了。

(50) 按照 Meurers (2000: 206–208) 观点的一般动词语迹：

$$\left[\begin{array}{l} \text{PHON } \langle \rangle \\ \text{SYNSEM} | \text{LOC } [1] \left[\text{CAT} | \text{HEAD} | \text{DSL } [1] \right] \end{array} \right]$$

这刚看上去可能有些出人意料，但是如果我们仔细看词汇规则 (49) 和树中 DSL 特征的渗透的互动关系，那么动词投射的 DSL value 就更加清晰了，由此，动词语迹的 LOCAL value 由输入动词的 LOCAL value 决定。在图 9.10 中，*kennt* 是动词移位词汇规则的输入。相应的结构共享可以确保，在 (44) 的分析中，动词语迹的 LOCAL value 确切地对应于 (48) 中给出的内容。

动词位置分析的最为重要的内容的总结如下所示：

- 词汇规则允准了每个定式动词的具体的词汇项。
- 该词汇项占据了首位，并且要求它的论元是动词语迹的完整投射。
- 动词语迹的投射必须具有 DSL value，以对应于词汇规则的输入动词的 LOCAL value。
- 由于 DSL 是一个中心语特征，被选择的 DSL value 也在语迹中表示。
- 正如语迹的 DSL value 与它的 LOCAL value 是相同的，语迹的 LOCAL value 与词汇规则的输入动词的 LOCAL value 是一致的。

在讨论完动词开头的句子之后，我们现在来看局部重新排序的内容。

9.4 局部重新排序

针对中场的语序分析有许多种可能性：我们可以假定一个 GPSG 中的平铺结构 (Kasper 1994)、或者假定一个二叉结构，并且允许论元按照任意语序来填充。Kathol (2001) 和 Müller (1999a, 2002a, 2004c) 提出了一个折衷的看法：带有具体列表的二叉结构，这个列表包括属于一个中心语的论元和说明语。论元和说明语在这些列表内部按照自由语序排列。请看 Reape (1994) 和本书的第 11.7.2.2 节有关这些方法的形式化表示。所有这些平铺分析和折衷分析都被证明是错误的 (请看 Müller 2005b, 2014c 和 Müller 2007b: Section 9.5.1)，由此，我只讨论二叉结构的分析。

图 9.11 on the next page 表示了 (51a) 的分析。

- (51) a. [weil] jeder diesen Mann kennt
因为 每人 这 人 认识

- b. [weil] diesen Mann jeder kennt
 因为 这 人 每人 认识
 ‘因为每个人都认识这个男人’

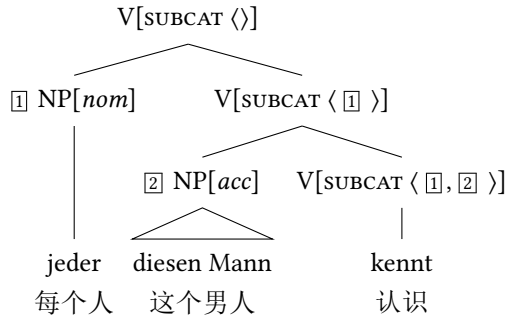


图 9.11: HPSG 中的成分序列分析：无标记语序

正如第 9.1.2 节所解释的，动词的论元与 SUBCAT list 的最后一个成分开头的动词相组合。有标记语序的分析如图 9.12 所示。这两棵树的区别只在于从 SUBCAT list 中取

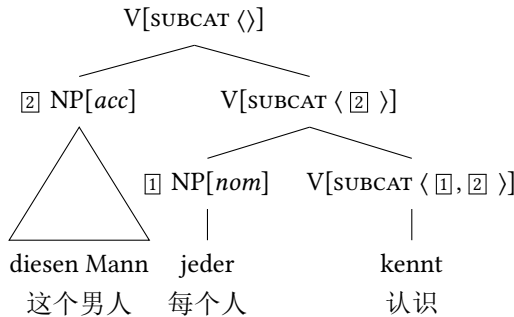


图 9.12: HPSG 中成分序列的分析：有标记语序

走的成分的顺序：在图 9.11 中，SUBCAT list 的最后一个成分先被释放，而图 9.12 中是第一个成分。

下面的模式是中心语-论元模式的修订版：

Schema 3 (中心语-论元模式 (二叉结构))

$$\text{head-argument-phrase} \Rightarrow \left[\begin{array}{l} \text{SYNSEM|LOC|CAT|SUBCAT } [1] \oplus [3] \\ \text{HEAD-DTR|SYNSEM|LOC|CAT|SUBCAT } [1] \oplus [2] \oplus [3] \\ \text{NON-HEAD-DTRS } \left([\text{SYNSEM } [2]] \right) \end{array} \right]$$

在第一版的中心语-论元模式中，总是 SUBCAT list 中的最后一个成分与中心语相组合，SUBCAT list 通过附加 (append) 被分成了三个部分：一个任意长度的列表 ([1])、一个只包括一个成分 ([2]) 的列表，以及一个更长的任意长度的列表 ([3])。列表 [1] 和 [3] 被组合起来，并且结果是母节点的 SUBCAT value。具有固定语序的语言（如英语）与德语这类语言是不同的，因为他们从一个方向开始释放论元（更多有关英语的主语内容，请看第 9.6.1 节），而自由语序的语言可以按照任意顺序将动词与论元相组合。在固定语序的语言中，要么是 [1] 要么是 [3] 总是空列表。由于德语结构没有受到与 [1] 或 [3] 相关的限制，也就是 [1] 和 [3] 要么是空列表，要么包括某些成分，这一直觉是指自由语序的语言比固定语序的语言具有更少的限制。我们可以将这个与第 4.6.1 节的 Kayne 式的分析相比较，这里它被认为是所有语言都是从基础语序 [specifier [head complement]] 推导而来的（请看图 4.20 on page 135 关于德语作为 SVO 语言的分析 (Laenzlinger 2004)）。在这些分析中，诸如英语的语言包括最为基本的格，而自由语序的语言需要花费一些理论上的努力来得到正确的语序。与之相对比的是，这里提出的分析需要更多理论上的限制，如果这个语言在它的成分排列上具有更多的限制的话。被允许结构的复杂度在 HPSG 理论的方法下并没有语言与语言之间的区别。语言只是在他们所属的分支类型上是不同的。^{19, 20}

这里展示的分析应用了任意顺序的论元的组合，这与 GB/MP 框架下的 Fanselow (2001) 的分析，以及 Hoffman (1995: Section 3.1) 和 Steedman & Baldridge (2006) 的范畴语法的分析是很相似的。Gunji 早在 1986 年就针对日语提出了类似的 HPSG 分析。

9.5 长距离依存

长距离依存分析应用了最初由 GPSG 中发展而来的技术：关于缺失成分的信息被传递到树上（或者特征结构中）。²¹ 在前置的成分的位置上一般会有一个语迹。图 9.13 on the following page 表示了 (52) 的分析。

- (52) [Diesen Mann]_j kennt_i _{-j} jeder _{-i}.
 这 男人 认识 每人
 ‘每个人都认识这个男人。’

原则上，我们也可以假定宾语是从未标记的位置上提取出来的（请看第 3.5 节关于未标记位置的内容）。提取的语迹可以在主语后面：

¹⁹ 这并没有排除这样的事实，我们讨论的这个结构具有不同的属性，只要把他们的处理度考虑进来的话。请看 Gibson (1998); Hawkins (1999) 和第 15 章。

²⁰ Haider (1997b: 18) 指出，这里提出的这类分析中，VX 语言的分支类型与 XV 语言的分支类型不同。这影响了 c-统制关系，并且这样对 GB/MP 下的约束理论产生了影响。但是，分支的方向与 HPSG 的分析是无关的，因为约束原则被界定为使用了 o-统制 (Pollard & Sag 1994: Chapter 6)，而 o-统制对应于旁格层级，即 SUBCAT list 中元素的语序，而不是这些元素与中心语相组合的语序。

²¹ 在 HPSG 中，没有什么真正地在特征结构和树中按照字面意义“向上传递”了。这可以看作是确定性理论（如 HPSG 理论）与像转换语法一样的推导性理论之间的最为重要的差别（请看第 15.1 节）。不过，它对于解释性的目的是有意义的，它为了解释这个分析就好像这个结构是自底向上构建的，但是语言知识独立于处理的方向。在最近的计算实现中，结构的构建更多是自底向上的，但是还有其他自顶向下工作的系统。非局部依存的分析中唯一重要的事情是缺失成分的信息，这个成分在所有的中间节点上与填充语和空位的信息是一致的。

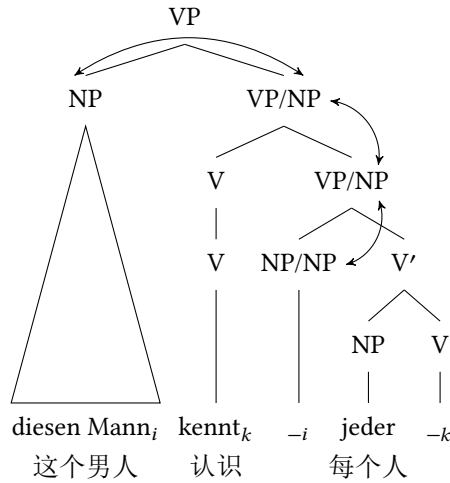


图 9.13: HPSG 中的长距离依存分析

- (53) [Diesen Mann]_j kennt_i jeder_{-j -i}.
这 男人 认识 每人
‘每个人都认识这个男人。’

Fanselow (2004b)认为, 某些特定的短语可以不具有特殊的语用功能而在前场被替换。比如说, 主动句中的(虚位)主语(54a)、时间副词(54b)、句子副词(54c)、心理动词的与格宾语(54d), 以及被动中的宾语(54e)可以在前场被替换, 即使他们既不是话题, 也不是焦点。

- (54) a. Es regnet.
它 下雨
‘下雨了。’
b. Am Sonntag hat ein Eisbär einen Mann gefressen.
PREP 星期天 AUX 一 北极熊 一 人 吃
‘在星期天, 有一头北极熊吃了一个人。’
c. Vielleicht hat der Schauspieler seinen Text vergessen.
也许 AUX DET 男演员 他的 台词 忘记
‘也许, 这个男演员已经忘记他的台词了。’
d. Einem Schauspieler ist der Text entfallen.
一.DAT 男演员 AUX DET.NOM 台词 忘记
‘一位男演员忘记台词了。’
e. Einem Kind wurde das Fahrrad gestohlen.
一.DAT 孩子 AUX DET.NOM 自行车 偷
‘一辆自行车从一个孩子那里被偷走了。’

Fanselow 认为信息结构的影响与中场的重新排序有关。所以通过 (55) 中宾格宾语的排序, 我们可以得到特定的效果:

- (55) Kennt diesen Mann jeder?
 认识 这 男人 每人
 ‘每个人都认识这个男人吗?’

如果有人认为有前置成分移到了 *Vorfeld*, 而且他们不具有连接到他们上的结构限制的信息, 而且这些结构限制的信息与中场的重新排序是有联系的, 那么这个在中场的首位成分被前置的假设解释了为什么 (54) 中的例子没有标记的结构信息。前场的成分在中场的首位也是没有标记的:

- (56) a. Regnet es?
 下雨 它
 ‘下雨了吗?’
- b. Hat am Sonntag ein Eisbär einen Mann gefressen?
 AUX PREP 星期天 一 北极熊 一 人 吃
 ‘一个北极熊在星期天吃了一个人吗?’
- c. Hat vielleicht der Schauspieler seinen Text vergessen?
 AUX 也许 DET 男演员 他的 台词忘记
 ‘这位男演员忘记他的台词了吗?’
- d. Ist einem Schauspieler der Text entfallen?
 AUX 一.DAT 男演员 DET.NOM 台词忘记
 ‘一个男演员忘记台词了吗?’
- e. Wurde einem Kind das Fahrrad gestohlen?
 AUX 一.DAT 孩子 DET.NOM 自行车 偷
 ‘有辆自行车从孩子那里偷走了吗?’

所以, 我认为前置论元的语迹在未标记的语序中不是中场-首位的, 而是最后与中心语相组合, 正如在第 9.4 节所描述的那样。当然, 这也同样适用于那些在未标记语序的中场-首位的所有提取的论元: 以 (57) 为例, 语迹最后与中心语组合:

- (57) [Jeder]_j kennt_i _{-j} diesen Mann _{-i}.
 每人 认识 这 男人
 ‘每个人都认识这个男人。’

在介绍完基本思想之后, 我们现在来看技术上的细节: 与我们在第 9.3 节讨论的动词移位不同的是, 成分移位是非局部的, 这就是为什么两个移位类型按照不同的特征 (SLASH vs. DSL) 来模拟的原因。DSL 是一个中心语特征, 而且跟所有其他中心语特征一样, 投射到投射层的最高点 (更多有关中心语特征原则的内容, 请看第 248 页)。另一方面, SLASH 是一个属于 SYNSEM|NONLOC 下表示的 NONLOC 特征的特征。NONLOC 特征的值是带有特征 INHERITED (或者简写为 INHER) 和 TO-BIND

的结构。*INHER* 的值是一个包括长距离依存中的成分的信息。(58) 给出了 Pollard & Sag (1994: 163) 提出的结构：²²

$$(58) \left[\begin{array}{ll} \text{nonloc} & \\ \text{QUE} & \text{list of } n\text{pros} \\ \text{REL} & \text{list of indices} \\ \text{SLASH} & \text{list of local structures} \end{array} \right]$$

QUE 对于疑问句的分析是很重要的，就像 *REL* 对于关系小句的分析是非常重要的。由于这些内容不在本书的范围内，所以后面我们会省略这些内容。*SLASH* 的值是 *local* 对象的一个列表。

正如动词移位的分析中，我们假定在宾格宾格通常出现的位置上有一个语迹，而且这个语迹共享了那个宾语的属性。由此，动词可以在局部满足它的配价要求。对于语迹是否与真正的论元相组合的信息在复杂的符号内部表示，并且在树上传递。这样，长距离依存的问题就可以通过树中更高的位于前场的成分得到解决。

长距离依存通过语迹而引入，它在其 *SLASH* list 中有一个对应于必有论元的 *LOCAL* value 的特征。(59) 显示了对于 (52) 的分析必需的语迹的描述：

(59) *kennen* 的宾格宾语的语迹（初级版本）：

$$\left[\begin{array}{ll} \text{word} & \\ \text{PHON} & \langle \rangle \\ \text{SYNSEM} & \left[\begin{array}{ll} \text{LOC} & \boxed{1} \left[\begin{array}{ll} \text{CAT} & \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{l} \text{noun} \\ \text{CAS } \text{acc} \end{array} \end{array} \right] \\ \text{SUBCAT} & \langle \rangle \end{array} \right] \end{array} \right] \\ \text{NONLOC} & \left[\begin{array}{ll} \text{INHER|SLASH} & \langle \boxed{1} \rangle \\ \text{TO-BIND|SLASH} & \langle \rangle \end{array} \right] \end{array} \right] \end{array} \right]$$

由于语迹没有内部结构（没有子节点），他们属于类型 *word*。语迹跟宾格宾语具有相同的属性。宾格宾语没在语迹占据的位置上出现的事实通过 *SLASH* 的值来表示。

下面的原则管的是 *NONLOC* 的信息被传递到了树上。

Principle 3 (非局部特征原则)

在中心语短语中，对于每个非局部特征来说，母节点的 *INHERITED* 值是一个列表，该列表是子节点的 *INHERITED* 值减去中心语子节点的 *TO-BIND* 列表中的成分的连接。

中心语-填充语模式（模式 4）允准了图 9.14 在下一页中的最高节点。该模式组合了一个定式的、动词首位的小句（*INITIAL+*），这个小句在带有非中心语的子节点的 *SLASH* 中有一个成分，它的 *LOCAL* 值与 *SLASH* 成分的值是相同的。在这个结构中，没有论元被满足了。没有任何动词可以从填充语子节点本身提取出来，这通过非

²² Pollard & Sag 认为，*QUE*、*REL* 和 *SLASH* 的值是集合，而不是列表。集合背后的数学原理更为复杂，这就是为什么我在这里认为是列表。

Schema 4 (中心语-填充语模式)

head-filler-phrase \Rightarrow

$$\left[\begin{array}{l} \text{HEAD-DTR|SYNSEM} \\ \text{NON-HEAD-DTRS} \left\langle \begin{array}{l} \text{SYNSEM} \left[\begin{array}{l} \text{LOC } \boxed{1} \\ \text{NONLOC|INHER|SLASH } \langle \rangle \end{array} \right] \end{array} \right\rangle \end{array} \right. \left. \begin{array}{l} \text{LOC|CAT} \left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \text{verb} \\ \text{VFORM } fin \\ \text{INITIAL } + \end{array} \right] \\ \text{SUBCAT } \langle \rangle \end{array} \right] \\ \text{NONLOC} \left[\begin{array}{l} \text{INHER|SLASH } \langle \boxed{1} \rangle \\ \text{TO-BIND|SLASH } \langle \boxed{1} \rangle \end{array} \right] \end{array} \right] \end{array} \right]$$

中心语子节点的 SLASH 值的确定而实现的。图 9.14 给出了前置到前场的分析的具体变体。*kennt* (认识) 的动词移位语迹跟一个名词性 NP 和一个提取的语迹相组合。

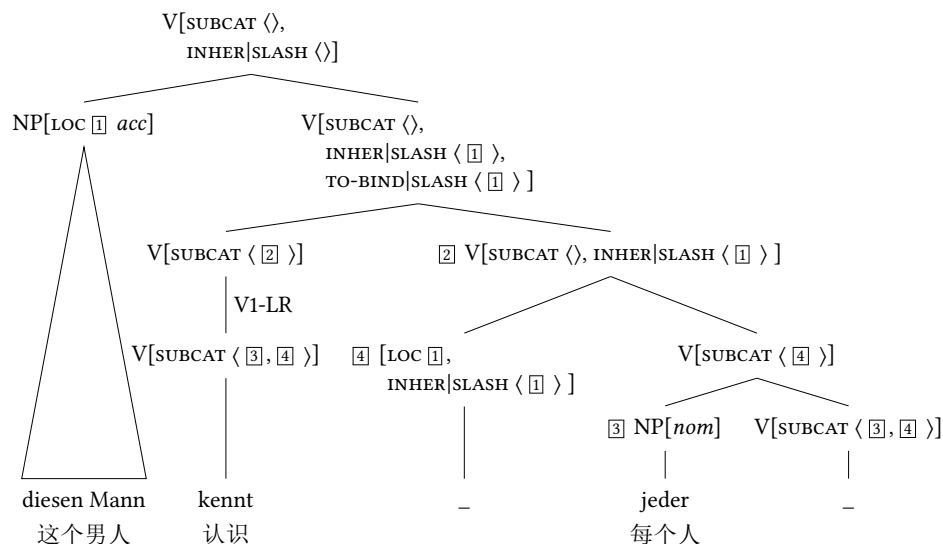
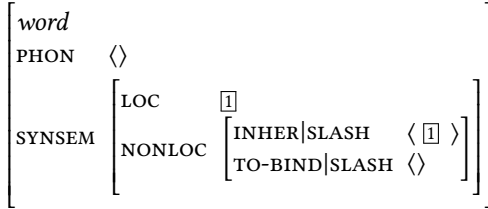


图 9.14: 针对动词首位语的结合了动词移位分析的 *Diesen Mann kennt jeder.* (每个人都认识这个男人。) 的分析

提取的语迹表示我们例子中的宾格宾语。宾格宾语在动词 (4) 的 SUBCAT list 中有所描述。按照动词移位的机制, 在 *kennt* 的词汇项最初包括的配价信息 ($\langle \boxed{3}, \boxed{4} \rangle$) 在动词语迹中有所表示。动词语迹的投射与提取语迹的组合跟非前置的论元具有相同的方式。提取语迹的 SLASH value 被传递到树上, 并且通过中心语-补足语模式而完成。

(59) 为语迹提供了词汇项，它可以用作 *kennen* (认识) 的宾格宾语。正如动词移位分析，没有必要在词库中包括许多具有不同属性的提取语迹。一个更为普遍的词汇项将满足如下条件，如 (60) 中的例子所示：

(60) 提取语迹：



这与这样的事实是有关系的，中心语可以令人满意地决定它所带论元的 LOCAL 属性，而且也可以决定它需组合的语迹的局部属性。在带有语迹 SYNSEM value 的中心语的 SUBCAT list 中的宾语的指认与这样的信息的指认是相关的，这些信息是关于前置的成分用来确保那些只能在前场实现的成分是哪符合中心语的 SUBCAT list 的描述的成分。同样的要求也适用于前置的状态语：因为通过 SLASH 特征前场内成分的 LOCAL value 与语迹的 LOCAL value 是一致的，那么就有足够的关于语迹的属性信息。

上述分析的核心观点可以总结如下：关于语迹的局部属性的信息属于语迹本身，然后出现在所有统制它的节点上，直到它到达了填充语。这一分析可以为所谓的提取路径标记语言提供解释，其中某些成分表现出了屈折变化，这取决于他们是否是跟一个从长距离依存中提取出的某个成分相组合。Bouma, Malouf & Sag (2001a) 将爱尔兰语、Chamorro、Palauan、冰岛语、Kikuyu、Ewe、Thompson Salish、Moore、法语、西班牙语和依地语这类语言作为例子，并且提供了相应的参考信息。由于在 HPSG 的分析中，信息是逐步传递的，所有参与到长距离依存的节点可以接触到那个依存关系里的成分。

9.6 新的进展与理论变体

本节讨论第 9.6.1 节的配价信息表示的修订与完善，并且简短地提到了 HPSG 理论的重要变体，即第 9.6.2 节的基于线性的 HPSG。

9.6.1 限定语、补足语与论元结构

在本章，SUBCAT 被认为是唯一的配价特征。这对应于 Pollard & Sag (1994: Chapter 1-8) 的理论地位。对于组成成分的组合，还需要至少一个额外的配价特征和一个对应的模式。这个额外的特征叫做限定语 (SPECIFIER, SPR)，它被用来表示英语 (Pollard & Sag 1994: Chapter 9) 和德语 (Müller 2007b: Section 9.3) 语法中限定词与名词的组合。一般认为，名词选择它的限定词。对于名词 *Zerstörung* (毁坏) 来说，我们有如下的 cat value：

(61)
$$\left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \textit{noun} \\ \text{INITIAL } + \end{array} \right] \\ \text{SPR} \quad \langle \text{DET} \rangle \\ \text{SUBCAT} \langle \text{NP}[\text{GEN}], \text{PP}[\textit{durch}] \rangle \end{array} \right]$$

模式 5 可以像中心语-论元模式那样来将名词和限定语相组合。

Schema 5 (限定语-中心语模式)

head-specifier-phrase \Rightarrow

$$\left[\begin{array}{l} \text{SYNSEM|LOC|CAT|SPR } [1] \\ \text{HEAD-DTR|SYNSEM|LOC|CAT} \left[\begin{array}{l} \text{SPR} \quad [1] \oplus \langle [2] \rangle \\ \text{SUBCAT} \quad \langle \rangle \end{array} \right] \\ \text{NON-HEAD-DTRS} \langle [\text{SYNSEM } [2]] \rangle \end{array} \right]$$

应用限定语模式的对 (62) 中的 NP 的分析如图 9.15 所示。

(62) die Zerstörung der Stadt durch die Soldaten
 DET 毁灭 PREP.DET 城市 PREP DET 士兵

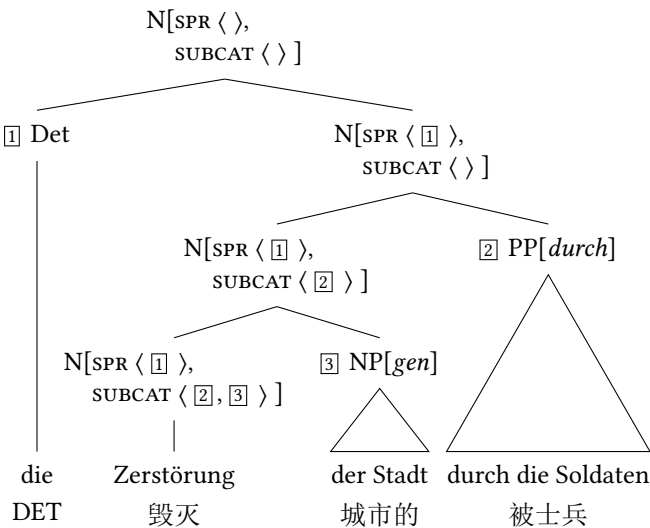


图 9.15: 带有配价特征 SPR 的 NP 分析

根据第 9.1.3 节讨论的线性规则，可以保证名词在补足语之前，因为名词的 INITIAL value 是 ‘+’。(63) 中的 LP-规则规定了限定词位于名词的左边。

(63) 限定语 < 中心语

在英语语法中, *SPR feature* 也用来表示动词对主语的选择 (Sag, Wasow & Bender 2003: Section 4.3)。在 (64) 这样的句子中, 动词首先与它的所有补足语相组合 (在较新的工作中 *SUBCAT* 和 *COMPS* 中的成分), 然后在第二步应用 5 模式来将主语组合进来。

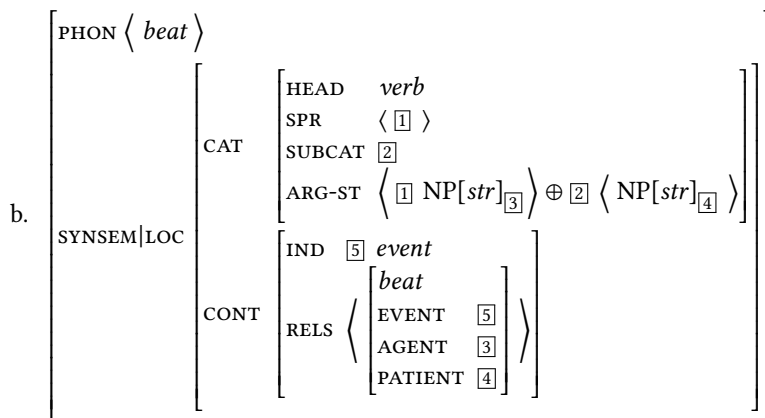
- (64) Max likes ice cream.
Max 喜欢冰 奶油
'Max 喜欢冰淇淋.'

正如我们在第 9.4 节看到的, 在定式句子的分析中按照相同的配价列表来表示主语和论元是有意义的。按照这种方式, 我们可以捕捉到这样的事实, 其中与论元相组合的动词不是固定的。尽管我们可以通过假定主语通过 *SPR* 来选择可以捕捉到不同的语序, 事实上置换是一个影响到所有论元的现象, 这一方式是没有覆盖在基于 *SPR* 的分析中的。进而, 主语的提取在英语这类语言中是不可能的, 但是在德语中是可能的 (相关参考资料和测试例子请看第 502 页)。我们可以通过假定英语中主语是通过 *SPR* 来选择的, 而 *SPR list* 中成分的提取是被禁止的这样的观点来说明他们的不同之处。因为德语中主语是表示在 *COMPS list* 上的, 这样就可以捕捉到他们与带有可能提取的宾语共存的事实。

Pollard & Sag (1994: Chapter 9) 提出的进一步的扩展是引入一个额外的列表, 它在较新的研究中叫做 *ARG-ST*。*ARG-ST* 表示论元结构。*ARG-ST list* 对应于我们在本章遇到的 *SUBCAT list*。它包括中心语的论元, 它们按照旁格等级来确定顺序。这个列表中的成分连接到中心语的语义内容中的论元角色 (请看第 9.1.6 节)。约束理论应用于 *ARG-ST list*。这一层次的表达可能对大部分语言来说都是一样的: 每一种语言都有语义谓词和语义论元。大多数语言利用在选择中发挥作用的句法范畴, 所以既有句法选择, 也有语义选择。²³ 语言之间的区别在于这些论元是如何实现的。在英语中, 配价列表中的第一个元素匹配到 *SPR list* 上, 而剩余的论元匹配到 *SUBCAT* (和近期工作中所说的 *COMPS list*) 上。在德语中, 动词的 *SPR list* 一直是空的。(65) 表示了德语和英语中相关的例子。

- (65) a.
$$\left[\begin{array}{c} \text{PHON} \langle \text{schlag} \rangle \\ \left[\begin{array}{c} \text{CAT} \left[\begin{array}{c} \text{HEAD} \quad \text{verb} \\ \text{SPR} \quad \langle \rangle \\ \text{SUBCAT} \quad [1] \\ \text{ARG-ST} \quad [1] \langle \text{NP}[\text{str}]_2, \text{NP}[\text{str}]_3 \rangle \end{array} \right] \\ \text{SYNSEM|LOC} \left[\begin{array}{c} \text{CONT} \left[\begin{array}{c} \text{IND} \quad [4] \text{event} \\ \text{RELS} \left\langle \begin{array}{c} \text{[schlagen} \\ \text{EVENT} \quad [4] \\ \text{AGENT} \quad [2] \\ \text{PATIENT} \quad [3] \end{array} \right\rangle \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right]$$

²³ Koenig & Michelson (2012) 指出, 奥奈达语 (北易洛魁语) 的分析没有囊括句法配价的表示。如果这个分析是正确的, 句法论元结构就不具有语言共性了, 而是大部分语言的特征而已。



我们可以将 ARG-ST list 看作是等同于 GB 理论的深层结构：语义角色按照这个列表来指派的。区别在于这里没有经历转换过程的有序的树。这样，有关所有的语言是从 VO 还是 OV 的语序生成而来的问题就变成无关的了。

9.6.2 基于线性顺序的 HPSG 理论

本章介绍的模式将邻接的成分组合起来。这里，关于邻接的假设可以忽略了，而非连续成分可以被允准。允许非连续成分的 HPSG 的变体通常被叫做基于线性的 HPSG 理论（Linearization-based HPSG）。最早的形式化体系是由 MikeReape (1991, 1992, 1994)开发的。支持线性方法的学者有Kathol (1995, 2000); Donohue & Sag (1999); Richter & Sailer (1999b); Crysmann (2008); Beavers & Sag (2004); Sato (2006); Wetta (2011)。我也提出了基于线性的分析 (Müller 1999a, 2002a)，并在 Reape 的思想的基础上实现了大规模的语法片段 (Müller 1996c)。基于限定的方法对于德语句子结构的分析与 GPSG 采用的方法是十分相似的，因为它认为动词、论元和说明语是相同线性范畴的成员，由此可以按照任意顺序来排列。比如说，动词可以位于论元和说明语的前面或后面。所以说，在动词位于末位的位置上没有空成分是十分必要的。如果我们允许在动词位置的分析中不带空成分的话，那么就不清楚明显的多重前置该怎么处理了，尽管这些数据可以在本章提出的方法中被直接地获得。整个问题在Müller (2015b)中有更为详细的讨论。我在这里不对 Reape 的形式化进行解释，但是会在第 11.7.2.2节来讨论，那里我们将对依存语法中的非连续、非投射的结构与基于线性的 HPSG 理论进行对比。明显的多重前置以及他们对简单的基于线性的方法提出的挑战将在第 11.7.1节进行讨论。

9.7 总结

在 HPSG 中，特征描写被用于模拟语言对象的所有属性：根、词、词汇规则和统制模式都用相同的形式工具来描写。与 GPSG 和 LFG 不同的是，这里没有独立的短语结构规则。由此，尽管 HPSG 代表中心语驱动的短语结构语法，这里并没有短语结构语法。在 HPSG 的实现中，短语结构的支撑通常用来提高处理的效率。但是，这并不属于理论的一部分，而且在语言学上并不是必须的。

HPSG 与范畴语法 的不同之处在于，它假定了更多的特征，也在于被组合在一起的特征的方式在理论中起到了重要的作用。

在 HPSG 中，长距离依存并没有像范畴语法那样应用函数组合来分析，而是相反地跟 GPSG 一样利用树间的信息渗透。按照这样的方式，我们可以分析第 8.6 节讨论的那些随伴结构，这些结构中每个关系连词只有一个词汇项，而且相关的局部属性与指示代词的属性是相同的。(66) 中的关系小句被分析为一个定式小句，其中 PP 被提取出来：

- (66) *der Mann*, [_{RS} [_{PP} *an den*] [_{S/PP} *wir gedacht haben*]]
 DET 男人 PREP 谁 我们想 AUX
 ‘我们想起的那个男人’

对于关系小句来说，我们要求第一个子节点包括一个关系代词。正如第 233 页的英语例子，这个代词事实上可以嵌套的非常深。关于 *an den* (谁的) 包括一个关系代词的信息的事实通过明确 NONLOC|INHER|REL 的值而表现在关系代词 *den* 的词汇项中。非局部特征原则将信息向上传递，这样有关关系代词的信息就包括在短语 *an den* 的表示中了。当这个关系小句被放在一起的时候 (Pollard & Sag 1994: Chapter 5, Sag 1997)，这个信息就完成使命了。我们不可能在 (66) 和 (67) 的分析中对 *den* 使用相同的词汇项，这与范畴语法是不同的，关系代词并不需要知道它可以使用的语境的任何信息。

- (67) *der Mann*, [_{RS} [_{NP} *den*] [_{S/NP} *wir kennen*]]
 DET 男人 CONJ 我们认识
 ‘我们认识的那个男人’

任何想要表示这里所述的分析的理论必须要有这样的机制，要使得在复杂短语中有关关系代词的信息是可获取的。如果我们在我们的理论中有这样一个机制，正如 LFG 和 HPSG 中的那样，那么我们也可以将之用于长距离依存的分析。这样，诸如 LFG 和 HPSG 的理论在描述工具方面与其他理论相比就会显得有些吝啬，尤其是在针对关系短语的分析之中。

在 HPSG 历史的第一个十年中 (Pollard & Sag 1987, 1994; Nerbonne, Netter & Pollard 1994a)，尽管这里已经提到了一些区别，HPSG 与范畴语法仍是非常相似的，因为它是一个强势的基于词汇的理论。短语的句法构造与语义内容都是由中心语决定的 (所以叫做中心语驱动的)。一旦遇到无法直接进行中心语驱动的分析，因为在讨论的短语中没有中心语，那么通常的做法就是假定空中心语。一个例子就是 Pollard & Sag (1994: Chapter 5) 中的关系小句的分析。由于空中心语可以被指派给任意句法配价和任意的语义 (关于这一点的讨论，请看第 19 章)，我们并没有好的理由来解释人们为什么要假定空的中心语，比如说这个空位置可以在其他语境中得到实现。但是，空中心语的提出并不是为了维护理论的假设。基于此，Sag (1997) 提出了不用任何空成分的关系小句的分析。正如 (66) 和 (67) 草拟的分析一样，关系小句是直接由分句组合而成的，以构成关系小句。对于英语中可观察到的不同类型的关系小句，Sag 提出了不同的统制规则。他的分析偏离了强势的词汇主义：在 Pollard & Sag (1994) 中，只有六条统制模式，而在 Ginzburg & Sag (2000) 中有 23 条。

在最近的会议论文集中，也可以看到对于短语模式的区分的倾向。提出的观点从对空元素的删除到激进地采取短语的分析都有。²⁴

即使向短语分析的倾斜会导致一些有问题的分析，事实上仍有一些语法的部分是需要短语分析的（请看第 21.10 节）。对于 HPSG 来说，这意味着它不再是中心语驱动的，而且既不是中心语驱动的，也不是短语结构语法。

HPSG 利用了类型特征描写来描述语言对象。概化可以通过带有多重承继的体系来表示。承继关系在构式语法中也起到了重要的作用。在诸如 GPSG、范畴语法和 TAG 中，它并不是理论解释的一部分。在实现中，宏语（macros）通常用来表示共现的特征值偶对（Dalrymple, Kaplan & King 2004）。按照假定的构架，这类宏语不适合短语的描写，因为，在诸如 GPSG 和 LFG 的理论中，短语结构规则的表达是不同于其他特征值偶对的（但是，请看 Asudeh, Dalrymple & Toivonen (2008, 2013) 用于 c-结构标记的宏和承继关系）。进而，在类型和宏之间还有更深的区别，这些区别具有更为正式的本质：在类型系统的，可以在一定条件下从具体特征和具体值的存在中推导出具体结构的类型。对于宏来说，并不是这样，因为它们只是简称。不过，由这个区别引起的语法分析的后果是微不足道的。

HPSG 理论不同于 GB 理论及其后续的变体，因为它并没有假定转换关系。在上世纪 80 年代，有一些 GB 的表示变体被提出来，即他们认为没有 D-结构，S-结构是在自然对移位成分的原始位置的标记中创造出来的。相反，有人直接假定带有语迹的 S-结构，而且连 S-结构到逻辑形式的映射也被放弃了（Koster 1978、Haider 1993: Section 1.4、Frey 1993: 14）。这个观点对应于 HPSG 理论中的观点，而且一个框架内的许多分析都可以翻译到对方的理论中。

在 GB 理论中，术语主语和宾语没有起到直接的作用：我们可以用这些术语来进行描述，但是主语和宾语并没有根据特征或相似的机制来标记。无论如何，我们也是可以做出区分的，因为主语和宾语通常都实现在树中不同的位置上（主语位于 IP 的限定语位置上，而宾语作为动词的补足语）。在 HPSG 理论中，主语和宾语也不是理论的原始对象。因为配价列表是有序的，然后，这就意味着可以将 ARG-ST 成分与语法功能联系起来：如果有一个主语，这发生在配价列表的第一个位置上，然后宾语紧随其后。²⁵ 对于基于转换语法的（68b）的分析来说，目标是为了连接（68a）中的基本语序和（68b）中的派生语序。一旦我们构造出了基本语法，那么什么是主语什么是宾语就非常清楚了。所以说，应用到（68a）中的基础结构的转换是需要被反转的。

- (68) a. [weil] jeder diesen Mann kennt
 因为 每人 这 男人 认识
 ‘因为每个人都认识这个男人’
 b. [weil] diesen Mann jeder kennt
 因为 这 男人 每人 认识

在 HPSG 和其他无转换的模型中，目标是为了将按照（68b）中的顺序排列的论元指派到配价列表中的描写上。配价列表（或者新方法中的 ARG-ST）对应于 GB 的

²⁴ 更多讨论，请看 Müller (2007c) 和第 21.3.6 节。

²⁵ 当构成复杂谓词时，宾语出现在第一个位置上。请看 Müller (2002a: 157) 关于带有 *erlauben*（允许）这类动词的长被动。通常来说具有下面的条件：主语是第一个带有结构格的论元。

深层结构 (Deep Structure)。不同之处在于, 中心语本身没有被包括进论元结构中, 而这就是 D-结构的情况。

Bender (2008c)已经说明了我们可以如何从 Wambaya 这类非构型语言中通过指向中心语的论元结构来分析现象。在 Wambaya 中, 通常在英语或德语中算作是成分的词语可以非连续地出现, 它是一个形容词, 并在语义上属于名词短语, 而跟名词短语中其他部分共享相同的格、数和性的值。它可以出现在那些没有连接到剩余的名词短语的句子的位置上。Nordlinger (1998)分析了 LFG 中的相关数据。在她的分析中, 成分的不同部分指向句子的 f-结构, 并且直接保证了名词短语的所有部分都具有相同的格。Bender 采用了 HPSG 的一个变体, 其中在论元与其中心语组合后, 其配价信息没有从配价列表中移除, 而这个信息仍在配价列表中, 并且朝向中心语的最大投射向上传递 (Meurers 1999c、Przepiórkowski 1999b、Müller 2007b: Section 17.4)。Higginbotham (1985: 560)和Winkler (1997)提出了 GB 理论相似的观点。通过对完整配价信息的投射, 它在整个句子中都是可获得的, 并且非连续的成分可以指向它 (如通过 MOD), 而且可以构建出各自的限制。²⁶ 在这个分析中, HPSG 中的论元结构对应于 LFG 中的 f-结构。LFG 的扩展的中心语范畴也可以在 HPSG 中来模拟, 其中多重中心语可以共享相同的 f-结构。为此, 我们可以利用函数组合, 因为它在有关范畴语法那一章的内容上表示出来了 (请看第 8.5.2 章)。这点被译成 HPSG 理论的确切方式限于篇幅就不在这里解释了。读者可以参考Hinrichs & Nakazawa (1994)的原始作品, 以及Müller (2007b: Chapter 15)中的解释。

配价信息在 HPSG 理论中发挥了重要的作用。动词的词汇项在原则上预先裁定了该项目可以出现的结构的集合。应用词汇规则, 有可能将一个词汇项与其他词汇项联系起来。这些可以用在结构的其他集合中。所以我们可以看到在可能的结构的集合中建立联系的词汇规则的功能。词汇规则对应于转换语法中的转换。这点在第 19.5 节有更为详细的讨论。词汇规则的效果也可以通过空成分来取得。这也将成为第 19.5 节要讨论的内容。

在 GPSG 中, 元规则被用来允准那些为词汇中心语创造额外配价模式的规则。原则上, 元规则也可以用于没有词汇中心语的规则。这点被Flickinger (1983)和Gazdar et al. (1985: 59)通过特殊的限制而排除了。Flickinger, Pollard & Wasow (1985: 265)指出这类限制是不必要的, 如果有人应用词汇规则而不是元规则的话, 因为前者只能用于词汇中心语。

对于 HPSG 和 Stabler 的最简语法的比较, 请看第 4.6.4 节。

思考题

1. 在 HPSG 中, 句法树的地位是什么?
2. 在例子 (69) 的分析中, 格指派是如何发生的? ?

(69) Dem Mann wurde ein Buch geschenkt.
DET.DAT 男人 AUX 一.NOM 书 给
'这个男人被给了一本书。'

²⁶ 也请看Müller (2008a)关于德语和英语中描写性谓词的分析, 他们分别指向中心语的实现的和未实现的论元的列表。这一分析也可以在第 18.2 节得到解释。

3. 什么是联接 (linking)，它在 HPSG 中是如何表示的？

练习题

1. 请给出 (70) 的特征描写，*dass* 不用分析。

(70) [dass] Max lacht
CONJ Max 笑

2. 第 9.1.7 节中名词和所修饰形容词的组合的分析只是分析的一个草稿。比如说，没有解释我们如何能够确定形容词和名词在格上保持一致。请思考一下如何可以扩展这个分析，这样就可以分析 (71a) 中的形容词名词组合了，而不是 (71b) 中的对象：

(71) a. eines interessanten Mannes
一.GEN 有趣的.GEN 男人.GEN
b. *eines interessanter Mannes
一.GEN 有趣的.NOM 男人.GEN

延伸阅读

这里，理论各部分的表示跟其他理论一样都是相对来说比较简短的。对于 HPSG 理论更为全面的介绍，包括特征构架的动机，请看 Müller (2007b)。特别是，这里简略说明了被动的分析。更为全面的分析包括非宾格动词、形容词分词、情态不及物动词、不同的被动变体以及长被动 的内容，这些可以参考 Müller (2002a: Chapter 3) 和 Müller (2007b: Chapter 17)。

HPSG 理论的综述可以参考 Levine & Meurers (2006)、Przepiórkowski & Kupść (2006)、Bildhauer (2014) 和 Müller (2015c)。

第十章 Construction Grammar

Like LFG and HPSG, *Construction Grammar* (CxG) forms part of West Coast linguistics. It has been influenced considerably by Charles Fillmore, Paul Kay and George Lakoff (all three at Berkeley) and Adele Goldberg (who completed her PhD in Berkeley and is now in Princeton) (Fillmore 1988; Fillmore, Kay & O'Connor 1988; Kay & Fillmore 1999; Kay 2002, 2005; Goldberg 1995, 2006).

Fillmore, Kay, Jackendoff and others have pointed out the fact that, to a large extent, languages consist of complex units that cannot straightforwardly be described with the tools that we have seen thus far. In frameworks such as GB, an explicit distinction is made between core grammar and the periphery (Chomsky 1981a: 8), whereby the periphery is mostly disregarded as uninteresting when formulating a theory of Universal Grammar. The criticism leveled at such practices by CxG is justified since what counts as the 'periphery' sometimes seems completely arbitrary (Müller 2014d) and no progress is made by excluding large parts of the language from the theory just because they are irregular to a certain extent.

In Construction Grammar, idiomatic expressions are often discussed with regard to their interaction with regular areas of grammar. Kay & Fillmore (1999) studied the *What's X doing Y?*-construction in their classic essay. (1) contains some examples of this construction:

- (1) a. What is this scratch doing on the table?
- b. What do you think your name is doing in my book?

The examples show that we are clearly not dealing with the normal meaning of the verb *do*. As well as the semantic bleaching here, there are particular morphosyntactic properties that have to be satisfied in this construction. The verb *do* must always be present and also in the form of the present participle. Kay and Fillmore develop an analysis explaining this construction and also capturing some of the similarities between the WXDY-construction and the rest of the grammar.

There are a number of variants of Construction Grammar:

- Berkeley Construction Grammar (Fillmore 1988; Kay & Fillmore 1999; Fried 2015)
- Goldbergian/Lakovian Construction Grammar (Lakoff 1987; Goldberg 1995, 2006)
- Cognitive Grammar (Langacker 1987, 2000, 2008; Dąbrowska 2004)

- Radical Construction Grammar (Croft 2001)
- Embodied Construction Grammar (Bergen & Chang 2005)
- Fluid Construction Grammar (Steels & De Beule 2006; Steels 2011a)
- Sign-Based Construction Grammar (Sag 2010, 2012)

The aim of Construction Grammar is to both describe and theoretically explore language in its entirety. In practice, however, irregularities in language are often given far more importance than the phenomena described as ‘core grammar’ in GB. Construction Grammar analyses usually analyze phenomena as phrasal patterns. These phrasal patterns are represented in inheritance hierarchies (e.g., Croft 2001; Goldberg 2003b). An example for the assumption of a phrasal construction is Goldberg’s analysis of resultative constructions. Goldberg (1995) and Goldberg & Jackendoff (2004) argue for the construction status of resultatives. In their view, there is no head in (2) that determines the number of arguments.

(2) Willy watered the plants flat.

The number of arguments is determined by the construction instead, that is, by a rule or schema saying that the subject, verb, object and a predicative element must occur together and that the entire complex has a particular meaning. This view is fundamentally different from analyses in GB, Categorical Grammar, LFG¹ and HPSG. In the aforementioned theories, it is commonly assumed that arguments are always selected by lexical heads and not independently licensed by phrasal rules. See Simpson (1983), Neeleman (1994), Wunderlich (1997), Wechsler (1997), and Müller (2002a) for corresponding work in LFG, GB, Wunderlich’s Lexical Decomposition Grammar and HPSG.

Like the theories discussed in Chapters 5–9, CxG is also a non-transformational theory. Furthermore, no empty elements are assumed in most variants of the theory and the assumption of lexical integrity is maintained as in LFG and HPSG. It can be shown that these assumptions are incompatible with phrasal analyses of resultative constructions (see Section 21.2.2 and Müller 2006, 2007c). This point will not be explained further here. Instead, I will discuss the work of Fillmore and Kay to prepare the reader to be able to read the original articles and subsequent publications. Although the literature on Construction Grammar is now relatively vast, there is very little work on the basic formal assumptions or analyses that have been formalized precisely. Examples of more formal works are Kay & Fillmore (1999), Kay (2002), Michaelis & Ruppenhofer (2001), and Goldberg (2003b). Another formal proposal was developed by Jean-Pierre Koenig (1999) (formerly Berkeley). This work is coached in the framework of HPSG, but it has been heavily influenced by CxG. Fillmore and Kay’s revisions of their earlier work took place in close collaboration with Ivan Sag. The result was a variant of HPSG known as Sign-Based Construction Grammar (SBCG) (Sag 2010, 2012). See Section 10.6.2 for further discussion.

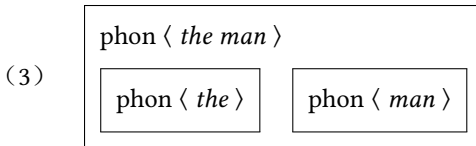
¹ See Alsina (1996) and Asudeh, Dalrymple & Toivonen (2008, 2013), however. For more discussion of this point, see Sections 21.1.3 and 21.2.2.

John Bryant, Nancy Chang, Eva Mok have developed a system for the implementation of Embodied Construction Grammar². Luc Steels is working on the simulation of language evolution and language acquisition (Steels 2003). Steels works experimentally modeling virtual communities of interacting agents. Apart from this he uses robots that interact in language games (Steels 2015). In personal communication (p.c. 2007) Steels stated that it is a long way to go until robots finally will be able to learn to speak but the current state of the art is already impressive. Steels can use robots that have a visual system (camera and image processing) and use visual information paired with audio information in simulations of language acquisition. The implementation of Fluid Construction Grammar is documented in Steels (2011a) and Steels (2012). The second book contains parts about German, in which the implementation of German declarative clauses and *w* interrogative clauses is explained with respect to topological fields (Micelli 2012). The FCG system, various publications and example analyses are available at: <http://www.fcg-net.org/>. Jurafsky (1996) developed a Construction Grammar for English that was paired with a probabilistic component. He showed that many performance phenomena discussed in the literature (see Chapter 15 on the Competence/Performance Distinction) can be explained with recourse to probabilities of phrasal constructions and valence properties of words. Bannard, Lieven & Tomasello (2009) use a probabilistic context-free grammar to model grammatical knowledge of two and three year old children.

10.1 General remarks on the representational format

In this section, I will discuss the mechanisms of Berkeley Construction Grammar (BCG). As I pointed out in Müller (2006), there are fundamental problems with the formalization of BCG. The details will be given in Section 10.6.1. While the framework was developed further into Sign-Based Construction Grammar (see Section 10.6.2) by its creators Kay and Fillmore, there are still authors working in the original framework (for instance Fried 2013). I will therefore present the basic mechanisms here to make it possible to understand the original ideas and put them into a broader context.

As we saw in Section 9.1.2, dominance relations in HPSG are modeled like other properties of linguistic objects using feature-value pairs. In general, CxG uses feature-value pairs to describe linguistic objects, but dominance relations are represented by boxes (Kay & Fillmore 1999; Goldberg 2003b):



The structure can be written using feature-value pairs as follows:

(4)

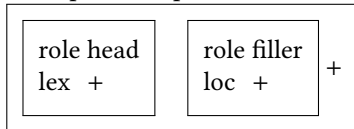
$$\left[\begin{array}{l} \text{PHON} \langle \textit{the man} \rangle \\ \text{DTRS} \langle [\text{PHON} \langle \textit{the} \rangle], [\text{PHON} \langle \textit{man} \rangle] \rangle \end{array} \right]$$

² See <http://www.icsi.berkeley.edu/~jbryant/old-analyzer.html> and Bryant (2003).

10.1.1 The head-complement construction

Kay & Fillmore (1999) assume the following construction for the combination of heads with their complements:

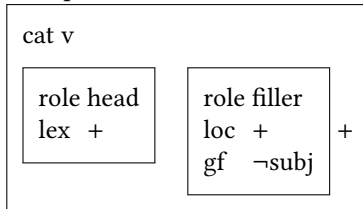
- (5) Head plus Complements Construction (HC)



A head is combined with at least one complement (the '+' following the box stands for at least one sign that fits the description in that box). LOC+ means that this element must be realized locally. The value of ROLE tells us something about the role that a particular element plays in a construction. Unfortunately, here the term *filler* is used somewhat differently than in GPSG and HPSG. Fillers are not necessarily elements that stand in a long-distance dependency to a gap. Instead, a *filler* is a term for a constituent that fills the argument slot of a head.

The verb phrase construction is a sub-construction of the head-complement construction:

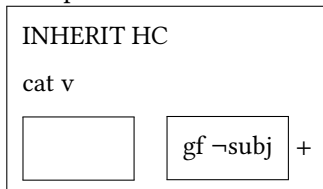
- (6) Verb phrase Construction:



The syntactic category of the entire construction is V. Its complements cannot have the grammatical function subject.

The VP construction is a particular type of head-complement construction. The fact that it has much in common with the more general head-complement construction is represented as follows:

- (7) Verb phrase Construction with inheritance statement:



This representation differs from the one in HPSG, aside from the box notation, only in the fact that feature descriptions are not typed and as such it must be explicitly stated in the representation from which superordinate construction inheritance takes place. HPSG – in addition to the schemata – has separate type hierarchies specifying the inheritance relation between types.

10.1.2 Representation of valence information

In Kay and Fillmore, valence information is represented in a set (VAL). The Valence Principle states that local filler-daughters have to be identified with an element in the valence set of the mother.³ The Subset Principle states that the set values of the head-daughter are subsets of the corresponding sets of the mother. This is the exact opposite approach to the one taken in Categorical Grammar and HPSG. In HPSG grammars, valence lists at the mother nodes are shorter, whereas in Berkeley CxG at least as many elements are present on the mother node as on the head-daughter.

10.1.3 Semantics

Semantics in CxG is handled exactly the same way as in HPSG: semantic information is contained in the same feature structure as syntactic information. The relation between syntax and semantics is captured by using the same variable in the syntactic and semantic description. (8) contains a feature description for the verb *arrive*:

(8) Lexical entry for *arrive* following Kay & Fillmore (1999: 11):

cat	v
sem	$\left\{ \left[\begin{array}{cc} \text{FRAME} & \text{ARRIVE} \\ \text{ARGS} & \{ A \} \end{array} \right] \right\}$
val	$\{ [\text{SEM } \{ A \}] \}$

Kay & Fillmore (1999: 9) refer to their semantic representations as a notational variant of the Minimal Recursion Semantics of Copestake, Flickinger, Pollard & Sag (2005). In later works, Kay (2005) explicitly uses MRS. As the fundamentals of MRS have already been discussed in Section 9.1.6, I will not repeat them here. For more on MRS, see Section 19.3.

10.1.4 Adjuncts

For the combination of heads and modifiers, Kay and Fillmore assume further phrasal constructions that are similar to the verb phrase constructions discussed above and create a relation between a head and a modifier. Kay and Fillmore assume that adjuncts also contribute something to the VAL value of the mother node. In principle, VAL is nothing more than the set of all non-head daughters in a tree.

10.2 Passive

The passive has been described in CxG by means of so-called linking constructions, which are combined with lexical entries in inheritance hierarchies. In the base lexicon, it is only listed which semantic roles a verb fulfils and the way in which these are realized is determined by the respective linking constructions with which the basic lexical entry is combined. Figure 10.1 on the next page gives an example of a relevant inheritance hierarchy. There is a linking construction for both active and passive as well as lexical

³ Sets in BCG work differently from those used in HPSG. A discussion of this is deferred to Section 10.6.1.

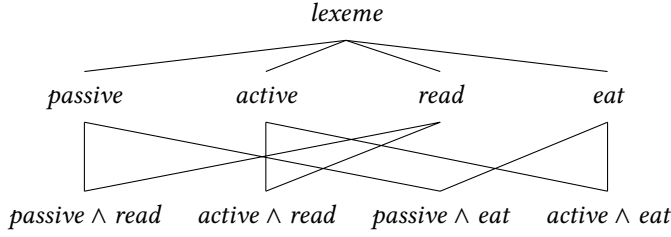


图 10.1: Passive and linking constructions

entries for *read* and *eat*. There is then a cross-classification resulting in an active and a passive variant of each verb.

The idea behind this analysis goes back to work by Fillmore and Kay between 1995 and 1997⁴, but variants of this analysis were first published in Koenig (1999: Chapter 3) and Michaelis & Ruppenhofer (2001: Chapter 4). Parallel proposals have been made in TAG (Candito 1996; Clément & Kinyon 2003: 188; Kallmeyer & Osswald 2012: 171–172) and HPSG (Koenig 1999; Davis & Koenig 2000; Kordoni 2001).

Michaelis & Ruppenhofer (2001: 55–57) provide the following linking constructions:⁵

- (9) a. *Transitive Construction*:
- $$\left[\begin{array}{l} \text{SYN} \left[\begin{array}{ll} \text{CAT} & v \\ \text{VOICE} & \text{active} \end{array} \right] \\ \text{VAL} \left\{ \left[\begin{array}{l} \text{ROLE} \left[\begin{array}{ll} \text{GF} & \text{obj} \\ \text{DA} & - \end{array} \right] \end{array} \right] \right\} \end{array} \right]$$
- b. *the Subject Construction*:
- $$\left[\begin{array}{l} \text{SYN} \left[\begin{array}{ll} \text{CAT} & v \end{array} \right] \\ \text{VAL} \left\{ \left[\begin{array}{l} \text{ROLE} \left[\begin{array}{ll} \text{GF} & \text{subj} \end{array} \right] \end{array} \right] \right\} \end{array} \right]$$
- c. *the Passive Construction*:
- $$\left[\begin{array}{l} \text{SYN} \left[\begin{array}{ll} \text{CAT} & v \\ \text{FORM} & \text{PastPart} \end{array} \right] \\ \text{VAL} \left\{ \left[\begin{array}{l} \text{ROLE} \left[\begin{array}{ll} \text{GF} & \text{obl} \\ \text{DA} & + \end{array} \right] \\ \text{SYN} & \text{P[von]/zero} \end{array} \right] \right\} \end{array} \right]$$

⁴ <http://www.icsi.berkeley.edu/~kay/bcg/ConGram.html>. 03.05.2010.

⁵ In the original version of the transitive construction in (9a), there is a feature θ that has the value DA–, however, DA is a feature itself and – is the value. I have corrected this in (9a) accordingly.

In the following structures, GF stands for *grammatical function* and DA for *distinguished argument*. The distinguished argument usually corresponds to the subject in an active clause.

The structure in (9a) says that the valence set of a linguistic object that is described by the transitive construction has to contain an element that has the grammatical function *object* and whose DA value is ‘-’. The DA value of the argument that would be the subject in an active clause is ‘+’ and ‘-’ for all other arguments. The subject construction states that an element of the valence set must have the grammatical function *subject*. In the passive construction, there has to be an element with the grammatical function *oblique* that also has the DA value ‘+’. In the passive construction the element with the DA value ‘+’ is realized either as a *by*-PP or not at all (*zero*).

The interaction of the constructions in (9) will be explained on the basis of the verb *schlagen* ‘to beat’:

(10) Lexical entry for *schlag-* ‘beat’:

$$\left[\begin{array}{l} \text{SYN} \left[\begin{array}{l} \text{CAT } v \\ \text{VOICE } \text{active} \end{array} \right] \\ \text{VAL} \left\{ \left[\begin{array}{l} \text{ROLE} \left[\begin{array}{l} \theta \text{ agent} \\ \text{DA } + \end{array} \right] \end{array} \right], \left[\begin{array}{l} \text{ROLE} \left[\begin{array}{l} \theta \text{ patient} \end{array} \right] \end{array} \right] \right\} \end{array} \right]$$

If we combine this lexical entry with the transitive and subject constructions, we arrive at (11a) following Fillmore, Kay, Michaelis, and Ruppenhofer, whereas combining it with the subject and passive construction yields (11b):⁶

(11) a. *schlag-* + Subject and Transitive Construction:

$$\left[\begin{array}{l} \text{SYN} \left[\begin{array}{l} \text{CAT } v \\ \text{VOICE } \text{active} \end{array} \right] \\ \text{VAL} \left\{ \left[\begin{array}{l} \text{ROLE} \left[\begin{array}{l} \theta \text{ agent} \\ \text{GF } \text{subj} \\ \text{DA } + \end{array} \right] \end{array} \right], \left[\begin{array}{l} \text{ROLE} \left[\begin{array}{l} \theta \text{ patient} \\ \text{GF } \text{obj} \\ \text{DA } - \end{array} \right] \end{array} \right] \right\} \end{array} \right]$$

b. *schlag-* + Subject and Passive Construction:

$$\left[\begin{array}{l} \text{SYN} \left[\begin{array}{l} \text{CAT } v \\ \text{FORM } \text{PastPart} \end{array} \right] \\ \text{VAL} \left\{ \left[\begin{array}{l} \text{ROLE} \left[\begin{array}{l} \theta \text{ agent} \\ \text{GF } \text{obl} \\ \text{DA } + \end{array} \right] \end{array} \right], \left[\begin{array}{l} \text{ROLE} \left[\begin{array}{l} \theta \text{ patient} \\ \text{GF } \text{subj} \end{array} \right] \end{array} \right] \right\} \end{array} \right]$$

Using the entries in (11), it is possible to analyze the sentences in (12):

- (12) a. Er schlägt den Weltmeister.
 he beats the world.champion
 ‘He is beating the world champion.’

⁶ This assumes a particular understanding of set unification. For criticism of this, see Section 10.6.1.

- b. Der Weltmeister wird (von ihm) geschlagen.
 the world.champion is by him beaten
 ‘The world champion is being beaten (by him).’

This analysis is formally inconsistent as set unification cannot be formalized in such a way that the aforementioned constructions can be unified (Müller 2006; Müller 2007b: Section 7.5.2, see also Section 10.6.1 below). It is, however, possible to fix this analysis by using the HPSG formalization of sets (Pollard & Sag 1987; Pollard & Moshier 1990). The Subject, Transitive and Passive Constructions must then be modified such that they can say something about what an element in *VAL* looks like, rather than specifying the *VAL* value of a singleton set.

- (13) The *Subject Construction* with Pollard & Moschier’s definition of sets:

$$\left[\begin{array}{c} \text{SYN} | \text{CAT } v \\ \text{VAL } \boxed{1} \end{array} \right] \wedge \left\{ \left[\text{ROLE } \left[\text{GF } \textit{subj} \right] \right] \right\} \subset \boxed{1}$$

The restriction in (13) states that the valence set of a head has to contain an element that has the grammatical function *subj*. By these means, it is possible to suppress arguments (by specifying *SYN* as *zero*), but it is not possible to add any additional arguments to the fixed set of arguments of *schlagen* ‘to beat’.⁷ For the analysis of Middle Constructions such as (14), inheritance-based approaches do not work as there is no satisfactory way to add the reflexive pronoun to the valence set:⁸

- (14) Das Buch liest sich gut.
 the book reads REFL good
 ‘The book reads well / is easy to read.’

If we want to introduce additional arguments, we require auxiliary features. An analysis using auxiliary features has been suggested by Koenig (1999). Since there are many argument structure changing processes that interact in various ways and are linked to particular semantic side-effects, it is inevitable that one ends up assuming a large number of syntactic and semantic auxiliary features. The interaction between the various linking constructions becomes so complex that this analysis also becomes cognitively implausible and has to be viewed as technically unusable. For a more detailed discussion of this point, see Müller (2007b: Section 7.5.2).

⁷ Rather than requiring that *schlagen* ‘to beat’ has exactly two arguments as in HPSG, one could also assume that the constraint on the main lexical item would be of the kind in (11a). One would then require that *schlagen* has at least the two members in its valence set. This would complicate everything considerably and furthermore it would not be clear that the subject referred to in (13) would be one of the arguments that are referred to in the description of the lexical item for *schlagen* in (11a).

⁸ One technically possible solution would be the following: one could assume that verbs that occur in middle constructions always have a description of a reflexive pronoun in their valence set. The Transitive Construction would then have to specify the *SYN* value of the reflexive pronoun as *zero* so that the additional reflexive pronoun is not realized in the Transitive Construction. The middle construction would suppress the subject, but realizes the object and the reflexive.

This solution cannot be applied to the recursive processes we will encounter in a moment such as causativization in Turkish, unless one wishes to assume infinite valence sets.

The following empirical problem is much more serious: some processes like passivization, impersonalization and causativization can be applied in combination or even allow for multiple application, but if the grammatical function of a particular argument is determined once and for all by unification, additional unifications cannot change the initial assignment. We will first look at languages which allow for a combination of passivization and impersonalization, such as Lithuanian (Timberlake 1982: Section 5), Irish (Noonan 1994), and Turkish (Özkaragöz 1986; Knecht 1985: Section 2.3.3). I will use Özkaragöz's Turkish examples in (15) for illustration (1986: 77):

- (15) a. Bu şato-da boğ-ul-un-ur.
this château-LOC strangle-PASS-PASS-AOR
'One is strangled (by one) in this château.'
- b. Bu oda-da döv-ül-ün-ür.
this room-LOC hit-PASS-PASS-AOR
'One is beaten (by one) in this room.'
- c. Harp-te vur-ul-un-ur.
war-LOC shoot-PASS-PASS-AOR
'One is shot (by one) in war.'

-In, *-n*, and *-Il* are allomorphs of the passive/impersonal morpheme.⁹

Approaches that assume that the personal passive is the unification of some general structure with some passive-specific structure will not be able to capture double passivization or passivization + impersonalization since they have committed themselves to a certain structure too early. The problem for nontransformational approaches that state syntactic structure for the passive is that such a structure, once stated, cannot be modified. That is, we said that the underlying object is the subject in the passive sentence. But in order to get the double passivization/passivization + impersonalization, we have to suppress this argument as well. What is needed is some sort of process (or description) that takes a representation and relates it to another representation with a suppressed subject. This representation is related to a third representation which again suppresses the subject resulting in an impersonal sentence. In order to do this one needs different strata as in Relational Grammar (Timberlake 1982; Özkaragöz 1986), metarules (Gazdar, Klein, Pullum & Sag 1985), lexical rules (Dowty, 1978: 412; 2003: Section 3.4; Bresnan 1982c; Pollard & Sag 1987; Blevins 2003; Müller 2003b), transformations (Chomsky 1957), or just a morpheme-based morphological analysis that results in items with different valence properties when the passivization morpheme is combined with a head (Chomsky 1981a).

The second set of problematic data that will be discussed comes from causativization in Turkish (Lewis 1967: 146):

⁹ According to Özkaragöz, the data is best captured by an analysis that assumes that the passive applies to a passivized transitive verb and hence results in an impersonal passive. The cited authors discussed their data as instances of double passivization, but it was argued by Blevins (2003) that these and similar examples from other languages are impersonal constructions that can be combined with personal passives.

- (16) öl-dür-t-tür-t-
 ‘to cause someone to cause someone to cause someone to kill someone’
 (kill = cause someone to die)

The causative morpheme *-t* is combined four times with the verb (*tür* is an allomorph of the causative morpheme). This argument structure-changing process cannot be modeled in an inheritance hierarchy since if we were to say that a word can inherit from the causative construction three times, we would still not have anything different to what we would have if the inheritance via the causative construction had applied only once. For this kind of phenomenon, we would require rules that relate a linguistic object to another, more complex object, that is, lexical rules (unary branching rules which change the phonology of a linguistic sign) or binary rules that combine a particular sign with a derivational morpheme. These rules can semantically embed the original sign (that is, add *cause* to *kill*).

The problem of repeated combination with causativization affixes is an instance of a more general problem: derivational morphology cannot be handled by inheritance as was already pointed out by Krieger & Nerbonne (1993) with respect to cases like *preprepreversion*.

If we assume that argument alternations such as passive, causativization and the Middle Construction should be described with the same means across languages, then evidence from Lithuanian and Turkish form an argument against inheritance-based analyses of the passive (Müller 2006, 2007c; Müller & Wechsler 2014a). See also Section 21.2.2 for the discussion of an inheritance-based approach to passive in LFG and Section 21.4.2 for the discussion of an inheritance-based approach in Simpler Syntax.

10.3 Verb position

At present, I only know of one article in the framework of CxG that has dealt with the sentence structure in German. This is the article by Vanessa Micelli (2012), where she describes a computer implementation of a German grammar in Fluid Construction Grammar. This fragment is restricted to declarative V2-clauses and *wh*-questions. In her analysis, the middle field forms a constituent comprising exactly two constituents (the direct and indirect object).¹⁰ The right sentence bracket and the postfield are empty. Long-distance dependencies are not discussed. It is only possible for arguments of the verb in the left sentence bracket to occur in the prefield. Micelli’s work is an interesting starting point but one has to wait and see how the analysis will be modified when the grammar fragment is expanded.

In the following, I will not discuss Micelli’s analysis further, but instead explore some of the possibilities for analyzing German sentence structure that are at least possible in principle in a CxG framework. Since there are neither empty elements nor transformations, the GB and HPSG analyses as well as their variants in Categorical Grammar are ruled out. The following options remain:

¹⁰ Note that none of the constituent tests that were discussed in Section 1.3 justifies such an analysis and that no other theory in this book assumes the *Mittelfeld* to be a constituent.

- an analysis similar to LFG with an optional verb
- an entirely flat analysis as proposed in GPSG
- an analysis with binary branching but variable verb position like that of Steedman (2000: 159)

Different variants of CxG make different assumptions about how abstract constructions can be. In Categorical Grammar, we have very general combinatorial rules which combine possibly complex signs without adding any meaning of their own (see rule (2) on page 220 for example). (17) shows an example in which the abstract rule of forward application was used:

- (17) [[[[Gibt] der Mann] der Frau] das Buch]
 give the man the woman the book
 ‘Does the man give the woman the book?’

If we do not want these kinds of abstract combinatorial rules, then this analysis must be excluded.

The LFG analysis in Section 7.3 is probably also unacceptable on a CxG view as it is assumed in this analysis that *der Mann der Frau das Buch* forms a VP although only three NPs have been combined. CxG has nothing like the theory of extended head domains that was presented in Section 7.3.

Thus, both variants with binary-branching structures are ruled out and only the analysis with flat branching structures remains. Sign-based CxG, which is a variant of HPSG (Sag 2010: 486), as well as Embodied Construction Grammar (Bergen & Chang 2005: 156) allow for a separation of immediate dominance and linear order so that it would be possible to formulate a construction which would correspond to the dominance rule in (18) for transitive verbs:¹¹

- (18) $S \rightarrow V, NP, NP$

Here, we have the problem that adjuncts in German can occur between any of the arguments. In GPSG, adjuncts are introduced by metarules. In formal variants of CxG, lexical rules, but not metarules, are used.¹² If one does not wish to expand the formalism to include metarules, then there are three options remaining:

- Adjuncts are introduced in the lexicon (van Noord & Bouma 1994; Bouma, Malouf & Sag 2001a) and treated as arguments in the syntax,
- Constructions always have slots available for an arbitrary number of adjuncts, or

¹¹ In principle, this is also Micelli’s analysis, but she assumed that the middle field forms a separate constituent.

¹² Goldberg (2014: 116) mentions metarule-like devices and refers to Cappelle (2006). The difference between metarules and their CxG variant as envisioned by Cappelle and Goldberg is that in CxG two constructions are related without one construction being basic and the other one derived. Rather there exists a mutual relation between two constructions.

- Constructions can be discontinuous

Kasper (1994) has proposed an analysis of the first type in HPSG: adjuncts and arguments are combined with the head in a flat structure. This corresponds to the dominance rule in (19), where the position of adjuncts is not stated by the dominance rule.

(19) $S \rightarrow V, NP, NP, Adj^*$

If we want to say something about the meaning of the entire construction, then one has to combine the original construction (transitive, in the above example) with the semantics contributed by each of the adjuncts. These computations are not trivial and require relational constraints (small computer programs), which should be avoided if there are conceptually simpler solutions for describing a particular phenomenon.

The alternative would be to use discontinuous constructions. Analyses with discontinuous constituents have been proposed in both HPSG (Reape 1994) and Embodied Construction Grammar (Bergen & Chang 2005). If we apply Bergen and Chang's analysis to German, the italicized words in (20) would be part of a ditransitive construction.

(20) *Gibt der Mann morgen der Frau unter der Brücke das Geld?*
 gives the man tomorrow the woman under the bridge the money
 'Is the man going to give the woman the money under the bridge tomorrow?'

The construction has been realized discontinuously and the adjuncts are inserted into the gaps. In this kind of approach, one still has to explain how the scope of quantifiers and adjuncts is determined. While this may be possible, the solution is not obvious and has not been worked out in any of the CxG approaches to date. For further discussions of approaches that allow for discontinuous constituents see Section 11.7.2.2.

10.4 Local reordering

If we assume flat branching structures, then it is possible to use the GPSG analysis for the order of arguments. However, Kay (2002) assumes a phrasal construction for so-called Heavy-NP-Shift in English, which means that there is a new rule for the reordering of heavy NPs in English rather than one rule and two different ways to linearize the daughters.

In CxG, it is often argued that the usage contexts of certain orders differ and we therefore must be dealing with different constructions. Accordingly, one would have to assume six constructions to capture the ordering variants of sentences with ditransitive verbs in final position (see also page 164). An alternative would be to assume that the ordering variants all have a similar structure and that the information-structural properties are dependent on the position of constituents in the respective structure (see De Kuthy 2000 for German and Bildhauer 2008 for Spanish).

10.5 Long-distance dependencies

Kay & Fillmore (1999: Section 3.10) discuss long-distance dependencies in their article. Since the number of arguments is not specified in the verb phrase construction, it is possible that an argument of the verb is not locally present. Like the LFG and GPSG analyses in previous chapters, there are no empty elements assumed for the analysis of long-distance dependencies. In the *Left Isolation Construction* that licenses the entire sentence, there is a left daughter and a right daughter. The left daughter corresponds to whatever was extracted from the right daughter. The connection between the fronted element and the position where it is missing is achieved by the operator VAL. VAL provides all elements of the valence set of a linguistic object as well as all elements in the valence set of these elements and so on. It is thereby possible to have unrestricted access to an argument or adjunct daughter of any depth of embedding, and then identify the fronted constituent with an open valence slot.¹³ This approach corresponds to the LFG analysis of Kaplan & Zaenen (1989) based on functional uncertainty.

10.6 New developments and theoretical variants

Berkeley Construction Grammar was already discussed in the main part of this chapter. The discussion of the formal underpinnings was deferred until the theoretical variants section, since it is more advanced. I made some comments on set unification in Müller (2006: 858), but the longer discussion is only available in Müller (2007b: Section 7.5.2), which is in German. Therefore, I include Section 10.6.1 here, which discusses the formal underpinnings of Berkeley Construction Grammar in more detail and shows that they are not suited for what they were intended to do.

Section 10.6.2 discusses Sign-Based Construction Grammar, which was developed in joint work by Charles Fillmore, Paul Kay and Ivan Sag. It embodies ideas from BCG without having its formal flaws. Section 10.6.3 deals with Embodied Construction Grammar, which is based on work by Charles Fillmore, Paul Kay and George Lakoff. Section 10.6.4 deals with Fluid Construction Grammar.

10.6.1 Berkeley Construction Grammar

Section 10.2 discussed the valence representation in BCG and linking constructions for active and passive. Kay & Fillmore (1999) represent valence information in sets and I deferred the discussion of the formal properties of sets in BCG to this section. Fillmore and Kay's assumptions regarding set unification differ fundamentally from those that are made in HPSG. Kay and Fillmore assume that the unification of the set { a } with the set { b }, where a and b do not unify, results in the union of the two sets, that is

¹³ Note again, that there are problems with the formalization of this proposal in Kay & Fillmore's paper. The formalization of VAL, which was provided by Andreas Kathol, seems to presuppose a formalization of sets as the one that is used in HPSG, but the rest of Fillmore & Kay's paper assumes a different formalization, which is inconsistent. See Section 10.6.1.

$\{a, b\}$. Due to this special understanding of sets it is possible to increase the number of elements in a set by means of unification. The unification of two sets that contain compatible elements is the disjunction of sets that contain the respective unifications of the compatible elements. This sounds complicated, but we are only interested in a specific case: the unification of an arbitrary set with a singleton set:

$$(21) \quad \{NP[nom], NP[acc]\} \wedge \{NP[nom]\} = \{NP[nom], NP[acc]\}$$

According to Fillmore & Kay the unification of a set with another set that contains a compatible element does not result in an increase of the number of list elements.

(22) illustrates another possible case:

$$(22) \quad \{ \text{NP}, \text{NP}[\text{acc}] \} \wedge \{ \text{NP}[\text{nom}] \} = \{ \text{NP}[\text{nom}], \text{NP}[\text{acc}] \}$$

The first NP in (22) is underspecified with respect to its case. The case of the NP in the second set is specified as nominative. $\text{NP}[\text{nom}]$ does not unify with $\text{NP}[\text{acc}]$ but with NP.

This particular conception of unification has consequences. Unification is usually defined as follows:

- (23) The unification of two structures FS_1 and FS_2 is the structure FS_3 that is subsumed by both FS_1 and FS_2 where there is no other structure that subsumes FS_1 and FS_2 and is subsumed by FS_3 .

A structure FS_1 is said to subsume FS_3 iff FS_3 contains all feature value pairs and structure sharings from FS_1 . FS_3 may contain additional feature value pairs or structure sharings. The consequence is that the subsumption relations in (24b,c) have to hold if unification of valence sets works as in (24a):

- (24) Properties of the set unification according to Kay & Fillmore (1999):
- a. $\{ \text{NP}[\text{nom}] \} \wedge \{ \text{NP}[\text{acc}] \} = \{ \text{NP}[\text{nom}], \text{NP}[\text{acc}] \}$
 - b. $\{ \text{NP}[\text{nom}] \} \geq \{ \text{NP}[\text{nom}], \text{NP}[\text{acc}] \}$
 - c. $\{ \text{NP}[\text{acc}] \} \geq \{ \text{NP}[\text{nom}], \text{NP}[\text{acc}] \}$

(24b) means that a feature structure with a valence set that contains just one $\text{NP}[\text{nom}]$ is more general than a feature structure that contains both an $\text{NP}[\text{nom}]$ and an $\text{NP}[\text{acc}]$. Therefore the set of transitive verbs is a subset of the intransitive verbs. This is rather unintuitive, but compatible with Fillmore & Kay's system for the licensing of arguments. However, there are problems with the interaction of valence specifications and linking constructions, which we turn to now.

We have seen the result of combining lexical items with linking constructions in (11a) and (11b), but the question of how these results are derived has not been addressed so far. Kay (2002) suggests an automatic computation of all compatible combinations of maximally specific constructions. Such a procedure could be used to compute the lexical representations we saw in Section 10.2 and these could be then used to analyze the well-formed sentences in (12).

However, problems would result for ungrammatical sentences like (25b). *grauen* 'to dread' is a subjectless verb. If one would simply combine all compatible linking constructions with *grauen*, the Kay & Fillmoreian conception of set unification would cause the introduction of a subject into the valence set of *grauen*. (25b) would be licensed by the grammar:

- (25) a. Dem Student graut vor der Prüfung.
the.DAT student dreads before the exam
'The student dreads the exam.'
- b. *Ich graue dem Student vor der Prüfung.
I dread the.DAT student before the exam

One could solve this problem by specifying an element with the grammatical function *subject* in the lexical entry of *grauen* ‘to dread’. In addition, it would have to be stipulated that this subject can only be realized as an overt or covert expletive (The covert expletive would be *syn zero*). For the covert expletive, this means it has neither a form nor a meaning. Such expletive pronouns without phonological realization are usually frowned upon in Construction Grammar and analyses that can do without such abstract entities are to be preferred.

Kay & Fillmore (1999) represent the semantic contribution of signs as sets as well. This excludes the possibility of preventing the unwanted unification of linking constructions by referring to semantic constraints since we have the same effect as we have with valence sets: if the semantic descriptions are incompatible, the set is extended. This means that in an automatic unification computation all verbs are compatible with the Transitive Construction in (9a) and this would license analyses for (26) in addition to those of (25b).

- (26) a. * Der Mann schläft das Buch.
 the man sleeps the book
 b. * Der Mann denkt an die Frau das Buch.
 the man thinks at the woman the book

An intransitive verb was unified with the Transitive Construction in the analysis of (26a) and in (26b) a verb that takes a prepositional object was combined with the Transitive Construction. This means that representations like (11) cannot be computed automatically as was intended by Kay (2002). Therefore one would have to specify subconstructions for all argument structure possibilities for every verb (active, passive, middle, ...). This does not capture the fact that speakers can form passives after acquiring new verbs without having to learn about the fact that the newly learned verb forms one.

Michaelis & Ruppenhofer (2001) do not use sets for the representation of semantic information. Therefore they could use constraints regarding the meaning of verbs in the Transitive Construction. To this end, one needs to represent semantic relations with feature descriptions as it was done in Section 9.1.6. Adopting such a representation, it is possible to talk about two-place relations in an abstract way. See for instance the discussion of (27) on page 253. However, the unification with the Subject Construction cannot be blocked with reference to semantics since there exist so-called raising verbs that take a subject without assigning a semantic role to it. As is evidenced by subject verb agreement, *du* ‘you’ is the subject in (27), but the subject does not get a semantic role. The referent of *du* is not the one who *seems*.

- (27) Du scheinst gleich einzuschlafen.
 you seem.2SG soon in.to.sleep
 ‘You seem like you will fall asleep soon.’

This means that one is forced to either assume an empty expletive subject for verbs like *grauen* or to specify explicitly which verbs may inherit from the subject construction and which may not.

In addition to (27), there exist object raising constructions with accusative objects that can be promoted to subject in passives. The subject in the passive construction does not get a semantic role from the finite verb:

- (28) a. Richard lacht ihn an.
 Richard laughs him towards
 ‘Richard smiles at him.’
 b. Richard fischt den Teich leer.
 Richard fishes the pond empty

The objects in (28) are semantic arguments of *an* ‘towards’ and *leer* ‘empty’, respectively, but not semantic arguments of the verbs *lacht* ‘laughs’ and *fischt* ‘fishes’, respectively. If one wants to explain these active forms and the corresponding passive forms via the linking constructions in (9), one cannot refer to semantic properties of the verb. Therefore, one is forced to postulate specific lexical entries for all possible verb forms in active and passive sentences.

10.6.2 Sign-Based Construction Grammar

In more recent work by Fillmore, Kay, Michaelis and Sag, the Kay & Fillmore formalization of the description of valence using the Kay & Fillmore version of sets was abandoned in favor of the HPSG formalization (Kay 2005; Michaelis 2006; Sag 2012; Sag, Boas & Kay 2012: 10–11). Sign-Based Construction Grammar was developed from the Berkeley variant of CxG. Sign-Based Construction Grammar is a variant of HPSG (Sag 2010: 486) and as such uses the formal apparatus of HPSG (typed feature structures). Valence and saturation are treated in exactly the same way as in standard HPSG. Changes in valence are also analyzed as in HPSG using lexical rules (Sag, Boas & Kay 2012: Section 2.3). The analysis of long-distance dependencies was adopted from HPSG (or rather GPSG). Minimal Recursion Semantics (MRS; Copestake, Flickinger, Pollard & Sag 2005) is used for the description of semantic content. The only difference to works in standard HPSG is the organization of the features in feature structures. A new feature geometry was introduced to rule out constructions that describe daughters of daughters and therefore have a much larger locality domain in contrast to rules in phrase structure grammars, LFG, and GPSG. I do not view this new feature geometry as particularly sensible as it can be easily circumvented and serves to complicate the theory. This will be discussed in Section 10.6.2.1. Other changes regard the omission of the *LOCAL* feature and the omission of valence features. These changes are discussed in Section 10.6.2.2 and 10.6.2.4, respectively.

10.6.2.1 Locality and *MOTHER*

Sag, Wasow & Bender (2003: 475–489) and Sag (2007, 2012) suggest using a *MOTHER* feature in addition to daughter features. The Head-Complement Construction would then have the form in (29):

(29) Head-Complement Construction following Sag, Wasow & Bender (2003: 481):

$$\text{head-comp-cx} \rightarrow \left[\begin{array}{l} \text{MOTHER|SYN|VAL|COMPS } \langle \rangle \\ \text{HEAD-DTR } \boxed{0} \left[\begin{array}{l} \text{word} \\ \text{SYN|VAL|COMPS } \boxed{A} \end{array} \right] \\ \text{DTRS } \langle \boxed{0} \rangle \oplus \boxed{A} \text{ nelist} \end{array} \right]$$

The value of COMPS is then a list of the complements of a head (see Section 9.6.1). Unlike in standard HPSG, it is not *synsem* objects that are selected with valence lists, but rather signs. The analysis of the phrase *ate a pizza* takes the form in (30).¹⁴

$$(30) \left[\begin{array}{l} \text{head-comp-cx} \\ \left[\begin{array}{l} \text{phrase} \\ \text{FORM } \langle \text{ate, a, pizza} \rangle \\ \text{MOTHER} \left[\begin{array}{l} \text{HEAD } \text{verb} \\ \text{SYN} \left[\begin{array}{l} \text{SPR } \langle \text{NP[nom]} \rangle \\ \text{COMPS } \langle \rangle \end{array} \right] \\ \text{SEM } \dots \end{array} \right] \end{array} \right] \\ \left[\begin{array}{l} \text{word} \\ \text{FORM } \langle \text{ate} \rangle \\ \text{HEAD-DTR } \boxed{1} \left[\begin{array}{l} \text{HEAD } \text{verb} \\ \text{SYN} \left[\begin{array}{l} \text{SPR } \langle \text{NP[nom]} \rangle \\ \text{COMPS } \langle \boxed{2} \text{ NP[acc]} \rangle \end{array} \right] \\ \text{SEM } \dots \end{array} \right] \end{array} \right] \\ \text{DTRS } \langle \boxed{1}, \boxed{2} \rangle \end{array} \right]$$

The difference to HPSG in the version of Pollard & Sag (1994) is that for Sag, Wasow & Bender, signs do not have daughters and this makes the selection of daughters impossible. As a result, the SYNSEM feature becomes superfluous (selection of the PHON value and of the value of the newly introduced FORM feature is allowed in Sag, Wasow & Bender (2003) and Sag (2012)). The information about the linguistic objects that contribute to a complex sign is only represented in the very outside of the structure. The sign represented under MOTHER is of the type *phrase* but does not contain any information about the daughters. The object described in (30) is of course also of another type than the phrasal or lexical signs that can occur as its daughters. We therefore need the following extension so that the grammar will work (Sag, Wasow & Bender 2003: 478):¹⁵

¹⁴ SBCG uses a FORM feature in addition to the PHON feature, which is used for phonological information as in earlier versions of HPSG (Sag 2012: Section 3.1, Section 3.6). The FORM feature is usually provided in example analyses.

¹⁵ A less formal version of this constraint is given as the Sign Principle by Sag (2012: 105): “Every sign must be listemically or constructionally licensed, where: a sign is listemically licensed only if it satisfies some listeme, and a sign is constructionally licensed if it is the mother of some well-formed construct.”

- (31) Φ is a Well-Formed Structure according to a grammar G if and only if:
1. there is a construction C in G , and
 2. there is a feature structure I that is an instantiation of C , such that Φ is the value of the MOTHER feature of I .

For comparison, a description is given in (32) with the feature geometry that was assumed in Section 9.6.1.

- (32)
- | | | |
|---|--|---|
| [| <i>head-complement-phrase</i> |] |
| | PHON $\langle \text{ate, a, pizza} \rangle$ | |
| | <div style="display: flex; align-items: center;"> <div style="margin-right: 10px;">SYNSEM LOC</div> <div style="border-left: 1px solid black; padding-left: 10px;"> <div style="display: flex; align-items: center;"> <div style="margin-right: 10px;">CAT</div> <div style="border-left: 1px solid black; padding-left: 10px;"> <div style="display: flex; align-items: center;"> <div style="margin-right: 10px;">HEAD</div> <div><i>verb</i></div> </div> <div style="display: flex; align-items: center;"> <div style="margin-right: 10px;">SPR</div> <div>$\langle \text{NP[nom]} \rangle$</div> </div> <div style="display: flex; align-items: center;"> <div style="margin-right: 10px;">COMPS</div> <div>$\langle \rangle$</div> </div> </div> </div> </div> </div> | |
| | <div style="display: flex; align-items: center;"> <div style="margin-right: 10px;">CONT</div> <div>...</div> </div> | |
| | <div style="display: flex; align-items: center;"> <div style="margin-right: 10px;">HEAD-DTR</div> <div style="border-left: 1px solid black; padding-left: 10px;"> <div style="display: flex; align-items: center;"> <div style="margin-right: 10px;">SYNSEM LOC</div> <div style="border-left: 1px solid black; padding-left: 10px;"> <div style="display: flex; align-items: center;"> <div style="margin-right: 10px;">CAT</div> <div style="border-left: 1px solid black; padding-left: 10px;"> <div style="display: flex; align-items: center;"> <div style="margin-right: 10px;">HEAD</div> <div><i>verb</i></div> </div> <div style="display: flex; align-items: center;"> <div style="margin-right: 10px;">SPR</div> <div>$\langle \text{NP[nom]} \rangle$</div> </div> <div style="display: flex; align-items: center;"> <div style="margin-right: 10px;">COMPS</div> <div>$\langle \boxed{2} \text{ NP[acc]} \rangle$</div> </div> </div> </div> </div> </div> </div></div> | |
| | <div style="display: flex; align-items: center;"> <div style="margin-right: 10px;">CONT</div> <div>...</div> </div> | |
| | NON-HEAD-DTRS $\langle [\text{SYNSEM } \boxed{2}] \rangle$ | |

In (32), the features HEAD-DTR and NON-HEAD-DTRS belong to those features that phrases of type *head-complement-phrase* have. In (30), however, the phrase corresponds only to the value of the MOTHER feature and therefore has no daughters represented in the sign itself. Using the feature geometry in (32), it is in principle possible to formulate restrictions on the daughters of the object in the NON-HEAD-DTRS list, which would be completely ruled out under the assumption of the feature geometry in (30) and the restriction in (31).

There are several arguments against this feature geometry, which will be discussed in the following subsections. The first one is an empirical one: there may be idioms that span clauses. The second argument concerns the status of the meta statement in (31) and the third one computational complexity.

10.6.2.1.1 Idioms that cross constituent boundaries

In Müller (2007b: Chapter 12) I conjectured that the locality restrictions may be too strong since there may be idioms that require one to make reference to daughters of daughters for their description. Richter & Sailer (2009) discuss the following idioms as examples:

- (33) a. nicht wissen, wo X_Dat der Kopf steht
 not know where X the head stands
 ‘to not know where x’s head is at’

- b. glauben, X_Acc tritt ein Pferd
believe X kicks a horse
'be utterly surprised'
- c. aussehen, als hätten X_Dat die Hühner das Brot weggefressen
look as.if had X the chicken the bread away.eaten
'to look confused/puzzled'
- d. look as if butter wouldn't melt [in X's mouth]
'to look absolutely innocent'

In sentences containing the idioms in (33a–c), the X-constituent has to be a pronoun that refers to the subject of the matrix clause. If this is not the case, the sentences become ungrammatical or lose their idiomatic meaning.

- (34)
- a. Ich glaube, mich / # dich tritt ein Pferd.
I believe me.ACC you.ACC kicks a horse
 - b. Jonas glaubt, ihn tritt ein Pferd.¹⁶
Jonas believes him kicks a horse
'Jonas is utterly surprised.'
 - c. # Jonas glaubt, dich tritt ein Pferd.
Jonas believes you kicks a horse
'Jonas believes that a horse kicks you.'

In order to enforce this co-reference, a restriction has to be able to refer to both the subject of *glauben* 'believe' and the object of *treten* 'kick' at the same time. In SBCG, there is the possibility of referring to the subject since the relevant information is also available on maximal projections (the value of a special feature (XARG) is identical to the subject of a head). In (33a–c), we are dealing with accusative and dative objects. Instead of only making information about one argument accessible, one could represent the complete argument structure on the maximal projection (as is done in some versions of HPSG, see page 278 and pages 514–516). This would remove locality of selection, however, since if all heads project their argument structure, then it is possible to determine the properties of arguments of arguments by looking at the elements present in the argument structure. Thus, the argument structure of *wissen* 'to know' in (35) would contain the description of a *dass* clause.

- (35) Peter weiß, dass Klaus kommt.
Peter knows that Klaus comes
'Peter knows that Klaus is coming.'

Since the description of the *dass* clause contains the argument structure of *dass*, it is possible to access the argument of *dass*. *wissen* 'to know' can therefore access *Klaus kommt*. As such, *wissen* also has access to the argument structure of *kommt* 'to come',

¹⁶ <http://www.machandel-verlag.de/der-katzenschatz.html>, 06.07.2015.

which is why *Klaus* is also accessible to *wissen*. However, the purpose of the new, more restrictive feature geometry was to rule out such nonlocal access to arguments.

An alternative to projecting the complete argument structure was suggested by Kay et al. (2015: Section 6): instead of assuming that the subject is the XARG in idiomatic constructions like those in (33), they assume that the accusative or dative argument is the XARG. This is an interesting proposal that could be used to fix the cases under discussion, but the question is whether it scales up if interaction with other phenomena are considered. For instance, Bender & Flickinger (1999) use XARG in their account of question tags in English. So, if English idioms can be found that require a non-subject XARG in embedded sentences while also admitting the idiom parts in the embedded sentence to occur as full clause with question tag, we would have conflicting demands and would have to assume different XARGS for root and embedded clauses, which would make this version of the lexical theory rather unattractive, since we would need two lexical items for the respective verb.

(33d) is especially interesting, since here the X that refers to material outside the idiom is in an adjunct. If such cases existed, the XARG mechanism would be clearly insufficient since XARG is not projected from adjuncts. However, as Kay et al. (2015) point out the X does not necessarily have to be a pronoun that is coreferent with an element in a matrix clause. They provide the following example:

- (36) Justin Bieber—Once upon a time Ø butter wouldn't melt in little Justin's mouth.
Now internationally famous for being a weapons-grade petulant brat ...

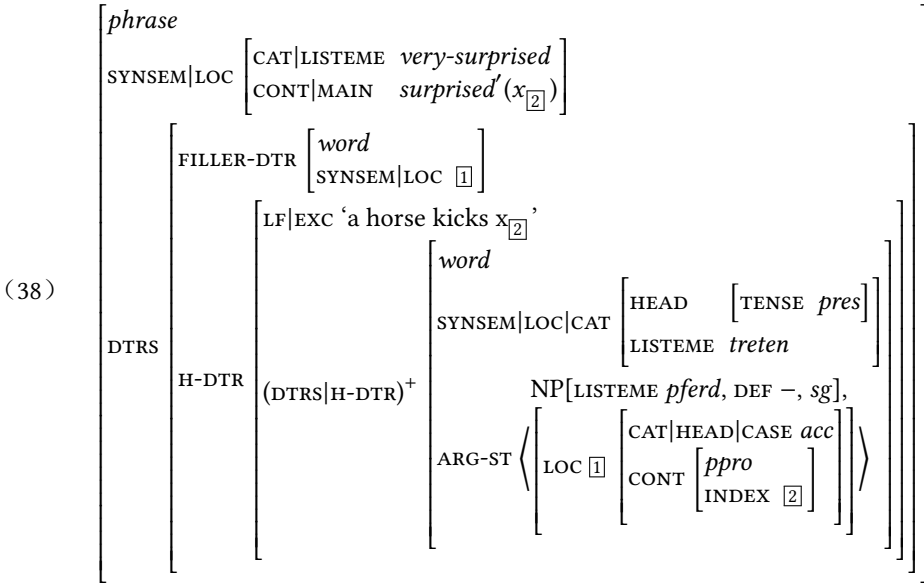
So, whether examples of the respective kind can be found is an open question.

Returning to our *horse* examples, Richter & Sailer (2009: 313) argue that the idiomatic reading is only available if the accusative pronouns is fronted and the embedded clause is V2. The examples in (37) do not have the idiomatic reading:

- (37) a. Ich glaube, dass mich ein Pferd tritt.
I believe that me a horse kicks
'I believe that a horse kicks me.'
b. Ich glaube, ein Pferd tritt mich.
I believe a horse kicks me
'I believe that a horse kicks me.'

Richter & Sailer assume a structure for *X_Acc tritt ein Pferd* in (33b) that contains, among others, the constraints in (38).

The feature geometry in (38) differs somewhat from what was presented in Chapter 9 but that is not of interest here. It is only of importance that the semantic contribution of the entire phrase is *surprised'*($x_{[2]}$). The following is said about the internal structure of the phrase: it consists of a filler-daughter (an extracted element) and also of a head daughter corresponding to a sentence from which something has been extracted. The head daughter means 'a horse kicks $x_{[2]}$ ' and has an internal head somewhere whose argument structure list contains an indefinite NP with the word *Pferd* 'horse' as its head.



The second element in the argument structure is a pronominal NP in the accusative whose LOCAL value is identical to that of the filler ([1]). The entire meaning of this part of the sentence is *surprised'*($x_{[2]}$), whereby [2] is identical to the referential index of the pronoun. In addition to the constraints in (38), there are additional ones that ensure that the partial clause appears with the relevant form of *glauben* 'to believe' or *denken* 'to think'. The exact details are not that important here. What is important is that one can specify constraints on complex syntactic elements, that is, it must be possible to refer to daughters of daughters. This is possible with the classical HPSG feature geometry, but not with the feature geometry of SBCG. For a more general discussion of locality, see Section 18.2.

The restrictions on *Pferd* clauses in (38) are too strict, however, since there are variants of the idiom that do not have the accusative pronoun in the *Vorfeld*:

- (39) a. ich glaub es tritt mich ein Pferd wenn ich einen derartigen Unsinn
 I believe EXPL kicks me a horse when I a such nonsense
 lese.¹⁷
 read
 'I am utterly surprised when I read such nonsense.'

¹⁷ <http://www.welt.de/wirtschaft/article116297208/Die-verlogene-Kritik-an-den-Steuerparadiesen.html>, commentary section, 10.12.2015.

- b. omg dieser xBluuR der nn ist wieder da ey nein ich glaub es tritt
 omg this XBluuR he is again there no I believe EXPL kicks
 mich ein Pferd!!¹⁸
 me a horse
 ‘OMG, this xBluuR, the nn, he is here again, no, I am utterly surprised.’
- c. ich glaub jetzt tritt mich ein pferd!¹⁹
 I believe now kicks me a horse
 ‘I am utterly surprised now.’

In (40a–b) the *Vorfeld* is filled by an expletive and in (40c) an adverb fills the *Vorfeld* position. While these forms of the idiom are really rare, they do exist and should be allowed for by the description of the idiom. So, one would have to make sure that *ein Pferd* ‘a horse’ is not fronted, but this can be done in the lexical item of *tritt* ‘kicks’. This shows that the cases at hand cannot be used to argue for models that allow for the representation of (underspecified) trees of depth greater one, but I still believe that such idioms can be found. Of course this is an open empirical question.

What is not an open empirical question though is whether humans store chunks with complex internal structure or not. It is clear that we do and much Construction Grammar literature emphasizes this. Constructional HPSG can represent such chunks, but SBCG cannot since linguistic signs do not have daughters. So here Constructional HPSG and TAG are the theories that can represent complex chunks of linguistic material with its internal structure, while other theories like GB, Minimalism, CG, LFG and DG can not.

10.6.2.1.2 Complicated licensing of constructions

In addition to these empirical problems, there is a conceptual problem with (31): (31) is not part of the formalism of typed feature structures but rather a meta-statement. Therefore, grammars which use (31) cannot be described with the normal formalism. The formalization given in Richter (2004) cannot be directly applied to SBCG, which means that the formal foundations of SBCG still have to be worked out.²⁰ Furthermore, the original problem that (31) was designed to solve is not solved by introducing the new feature geometry and the meta statement. Instead, the problem is just moved to another level since we now need a theory about what is a permissible meta-statement and what is not. As such, a grammarian could add a further clause to the meta statement stating that Φ is only a well-formed structure if it is true of the daughters of a relevant construction C that they are the MOTHER value of a construction C' . It would be possible to formulate

¹⁸ <http://forum.gta-life.de/index.php?user/3501-malcolm/>, 10.12.2015.

¹⁹ <http://www.castingshow-news.de/menowin-frhlich-soll-er-zum-islam-konvertieren-7228/>, 10.12.2015.

²⁰ A note of caution is necessary since there were misunderstandings in the past regarding the degree of formalizations of SBCG: in comparison to most other theories discussed in this book, SBCG is well-formalized. For instance it is easy to come up with a computer implementation of SBCG fragments. I implemented one in the TRALE system myself. The reader is referred to Richter (2004) to get an idea what kind of deeper formalization is talked about here.

constraints in the meta-statement about the construction C' or individual values inside the corresponding feature structures. In this way, locality would have been abandoned since it is possible to refer to daughters of daughters. By assuming (31), the theoretical inventory has been increased without any explanatory gain.

10.6.2.1.3 Computational complexity

One motivation behind restrictions on locality was to reduce the computational complexity of the formalism (Ivan Sag, p. c. 2011, See Chapter 17 on computational complexity and generative power). However, the locality restrictions of SBCG can be circumvented easily by structure sharing (Müller 2013b: Section 9.6.1). To see this consider a construction with the following form:

$$(40) \quad \left[\begin{array}{l} \text{MOTHER} \\ \text{DTRS} \end{array} \left[\begin{array}{l} \text{sign} \\ \text{PHON } \textit{phonological-object} \\ \text{FORM } \textit{morphological-object} \\ \text{SYN } \textit{syntactic information} \\ \text{SEM } \textit{semantic information} \\ \text{NASTY } \boxed{1} \end{array} \right] \right]$$

$\boxed{1} \textit{ list of signs}$

The feature NASTY in the MOTHER sign refers to the value of DTRS and hence all the internal structure of the sign that is licensed by the constructional schema in (40) is available. Of course one could rule out such things by stipulation – if one considered it to be empirically adequate, but then one could as well continue to use the feature geometry of Constructional HPSG (Sag 1997) and stipulate constraints like “Do not look into the daughters.” An example of such a constraint given in prose is the Locality Principle of Pollard & Sag (1987: 143–144).

10.6.2.2 Lexical extraction and the LOCAL feature

A main theme in Ivan Sag’s work was to eliminate empty elements from grammar. He co-developed a lexical analysis of extraction that did not use empty elements (van Noord & Bouma 1994; Bouma, Malouf & Sag 2001a). Rather than assuming a trace as in earlier versions of HPSG (see Section 9.5) a lexical rule (lexical construction) or a special mapping between list-valued features (for instance ARG-ST and VALENCE) is assumed so that a lexical item is licensed that has an element in the GAP list (SLASH in earlier versions of HPSG). The LOCAL feature in Standard HPSG is used to bundle all those information that is shared between a filler and the gap. The lexical entry for the trace was given as (60) on page 272 and is repeated as (41) for convenience. For the trace-based analysis of extraction it is crucial that only information under CATEGORY and CONT is shared between filler and gap. This information is bundled under LOCAL. Information about daughters and phonology of the filler is not shared. Since traces are not pronounced, their PHONOLOGY value would be incompatible with any PHONOLOGY value of a filler.

(41) Extraction trace:

word									
PHON	$\langle \rangle$								
SYNSEM	<table> <tr> <td>LOC</td><td>$\boxed{1}$</td></tr> <tr> <td>NONLOC</td><td> <table> <tr> <td>INHER SLASH</td><td>$\langle \boxed{1} \rangle$</td></tr> <tr> <td>TO-BIND SLASH</td><td>$\langle \rangle$</td></tr> </table> </td></tr> </table>	LOC	$\boxed{1}$	NONLOC	<table> <tr> <td>INHER SLASH</td><td>$\langle \boxed{1} \rangle$</td></tr> <tr> <td>TO-BIND SLASH</td><td>$\langle \rangle$</td></tr> </table>	INHER SLASH	$\langle \boxed{1} \rangle$	TO-BIND SLASH	$\langle \rangle$
LOC	$\boxed{1}$								
NONLOC	<table> <tr> <td>INHER SLASH</td><td>$\langle \boxed{1} \rangle$</td></tr> <tr> <td>TO-BIND SLASH</td><td>$\langle \rangle$</td></tr> </table>	INHER SLASH	$\langle \boxed{1} \rangle$	TO-BIND SLASH	$\langle \rangle$				
INHER SLASH	$\langle \boxed{1} \rangle$								
TO-BIND SLASH	$\langle \rangle$								

In a lexical approach, one would assume an additional lexical item with an element in SLASH or GAP.

(42) Lexical introduction of nonlocal dependencies according to Sag (2012: 163):

FORM	$\langle \textit{like} \rangle$				
ARG-ST	$\langle \boxed{1} \left[\begin{array}{c} \text{NP} \\ \text{GAP} \end{array} \right] \langle \rangle, \boxed{2} \left[\begin{array}{c} \text{NP} \\ \text{GAP} \end{array} \right] \langle \rangle \rangle$				
SYN	<table> <tr> <td>VAL</td><td>$\langle \boxed{1} \rangle$</td></tr> <tr> <td>GAP</td><td>$\langle \boxed{2} \rangle$</td></tr> </table>	VAL	$\langle \boxed{1} \rangle$	GAP	$\langle \boxed{2} \rangle$
VAL	$\langle \boxed{1} \rangle$				
GAP	$\langle \boxed{2} \rangle$				

The analysis of (43) then results in a linguistic object in which the second element of the ARG-ST list of *like* has the FORM value $\langle \textit{bagels} \rangle$.

(43) Bagels, I like.

In a trace-based account, the FORM value of the extracted element would be the empty list.

Now, the problem is that not everybody agrees with the traceless analysis of unbounded dependencies. For instance, Levine & Hukari (2006) wrote a monograph discussing various versions of traceless accounts of extraction and argue against them. Chaves (2009) suggested solutions to some of the puzzles, but does not solve them entirely. While feature geometries that include a LOCAL feature allow researchers to assume trace-based analyses, the SBCG geometry makes this impossible. So those who use traces in their theories will never adopt the SBCG geometry. See Section 19 for more on empty elements.

A further advantage of having a package of information that is shared in nonlocal dependencies is that some information can be excluded from sharing by specifying it outside of such a package. This was used by Höhle (1994), Müller (1996d, 2002a), and Meurers (1999a) to account for partial verb phrase fronting in German. The generalization about partial verb phrase fronting in German is that verbs can be fronted together with some or all of their objects even though the verb does not form a VP in other contexts. For instance, *erzählen* ‘tell’ and *wird* ‘will’ in (44a) usually form a complex that may not be separated by scrambling a projection of *erzählen* to the left.

- (44) a. dass er seiner Tochter ein Märchen erzählen wird
 that he his daughter a fairy.tale tell will
 ‘that he will tell his daughter a fairy tale’

- b. * dass er seiner Tochter ein Märchen erzählen nie wird
 that he his daughter a fairy.tale tell never will

Hinrichs & Nakazawa (1994) account for this by assuming a special schema for verbal complexes and by assuming that such verbal complexes are marked in a certain way (later this was encoded with the *LEX* feature with the value ‘+’) and that heads that form a verbal complex select for appropriately marked elements. So in the analysis of (44a) *wird* ‘will’ selects a *LEX* + element and the word *erzählen* ‘tell’ satisfies this requirement.

The problem for many accounts of nonlocal dependencies is that the fronted material is not necessarily a single verb:

- (45) a. Erzählen wird er seiner Tochter ein Märchen.
 tell will he.NOM his.DAT daughter a.ACC fairytale
 ‘He will tell his daughter a fairytale.’
 b. Ein Märchen erzählen wird er seiner Tochter.
 a fairytale.ACC tell will he.NOM his.DAT daughter
 c. Seiner Tochter erzählen wird er das Märchen.
 his.DAT daughter tell will he.NOM the.ACC fairytale
 d. Seiner Tochter ein Märchen erzählen wird er.
 his.DAT daughter a.ACC fairytale tell will he.NOM

If the local requirement of *wird* would be shared between filler and gap, the sentences in (45b)–(45d) would be ruled out, since the projections in the *Vorfeld* are not lexical elements but complex phrasal projections that are *LEX* –. Now, as pointed out by the authors mentioned above, this turns into a non-issue if not all information is shared between filler and gap. If *LEX* is outside *LOCAL*, the filler maybe *LEX* –, while the local constraints on the extracted element require the *LEX* value +. If all information is shared as in SBCG, no such solution is possible. Of course one could stick to the SBCG feature geometry and say that not the entire *ARG-ST* element is shared with the *GAP* element, but then one would end up with single structure sharings of all the features at the outermost level of a sign, that is, *PHON*, *FORM*, *SYN*, *SEM*, and *CNTXT* (see Sag 2012: 98 for these features)²¹ and the whole point of grouping features is to avoid such multiple individual structure sharings in situations in which there is a systematic relation between feature values.

10.6.2.3 Selection of *PHON* and *FORM* values

The feature geometry of constructional HPSG has the *PHON* value outside of *SYNSEM*. Therefore verbs can select for syntactic and semantic properties of their arguments but not for their phonology. For example, they can require that an object has accusative

²¹ Note that these structure sharings are necessary. It is not an option to leave the values of these features unspecified in the argument structure list. Since we have a model-theoretic view on things, an unspecified value would make the respective structures infinitely ambiguous, since infinitely many instantiation of these features are possible.

case but not that it starts with a vowel. SBCG allows for the selection of phonological information (the feature is called `FORM` here) and one example of such a selection is the indefinite article in English, which has to be either *a* or *an* depending on whether the noun or nominal projection it is combined with starts with a vowel or not (Flickinger, Mail to the HPSG mailing list, 01.03.2016):

- (46) a. an institute
 b. a house

The distinction can be modeled by assuming a selection feature for determiners.²² An alternative would be of course to capture all phonological phenomena by formulating constraints on phonology on the phrasal level (see Bird & Klein 1994 and Walther 1999 for phonology in HPSG).

Note also that the treatment of raising and nonlocal dependencies in SBCG admits nonlocal selection of phonology values, since the `FORM` value of the filler is present at the `ARG-ST` list of the head from which the argument is extracted. In earlier HPSG versions only `LOCAL` information is shared and elements in valence lists do not have a `PHON` feature. In principle, SBCG could be used to model languages in which the phonology of a filler is relevant for a head from which it is extracted. So for instance *likes* can see the phonology of *bagels* in (47):

- (47) Bagels, I think that Peter likes.

It would be possible to state constraints saying that the filler has to contain a vowel or two vowels or that it ends with a consonant. In addition all elements on the extraction path (*that* and *think*) can see the phonology of the filler as well. While there are languages that mark the extraction path, I doubt that there are languages that have phonological effects across long distances.

Similarly, the analysis of raising in SBCG assumes that the element on the valence list of the embedded verb is identical to an element in the `ARG-ST` list of the matrix verb (Sag 2012: 159). Hence, both verbs in (48) can see the phonology of the subject:

- (48) Kim can eat apples.

In principle there could be languages in which the form of the downstairs verb depends on the presence of an initial consonant in phonology of the subject. English allows for long chains of raising verbs and one could imagine languages in which all the verbs on the way are sensitive to the phonology of the subject. Such languages probably do not exist.

Now, is this a problem? Not for me, but if one develops a general setup in a way to exclude everything that is not attested in the languages of the world (as for instance the selection of arguments of arguments of arguments), then it is a problem that heads can see the phonology of elements that are far away.

²² In Standard HPSG there is mutual selection between the determiner and the noun. The noun selects the determiner via `SPR` and the determiner selects the noun via a feature called `SPECIFIED`. This feature is similar to the `MOD` feature, which was explained in Section 9.1.7.

There are two possible conclusions for practitioners of SBCG: either the *MOTHER* feature could be given up since one agrees that theories that do not make wrong predictions are sufficiently constraint and one does not have to explicitly state what cannot occur in languages or one would have to react to the problem with nonlocally selected phonology values and therefore assume a *SYNSEM* or *LOCAL* feature that bundles information that is relevant in raising and nonlocal dependencies and does not include the phonology.²³ This supports the arguments I made on *MOTHER* and *LOCAL* in the previous subsections.

10.6.2.4 The *VALENCE* list

Another change from Constructional HPSG to SBCG involves the use of a single valence feature rather than three features *SPR*, *SUBJ* and *COMPS* that were suggested by Borsley (1987) to solve problems in earlier HPSG versions that used a single valence feature (*SUBCAT*). Borseley's suggestion was taken up by Pollard & Sag (1994: Chapter 9) and has been used in some version or other in other HPSG versions since then.

Sag (2012: 85) assumes that VPs are described by the following feature description:

$$(49) \quad \left[\text{SYN} \left[\text{VAL} \langle \text{NP} \rangle \right] \right]$$

The problem with such an approach is that VPs differ from other phrasal projections in having an element on their *VALENCE* list. APs, NPs, and (some) PPs have an empty *VALENCE* list. In other versions of HPSG the complements are represented on the *COMPS* list and generalizations about phrases with fully saturated *COMPS* lists can be expressed directly. One such generalization is that projections with an empty *COMPS* list (NPs, PPs, VPs, adverbs, CPs) can be extraposed in German (Müller 1999a: Section 13.1.2).

10.6.2.5 Conclusion

Due to the conceptual problems with meta-statements and the relatively simple ways of getting around locality restrictions, the reorganization of features (*MOTHER* vs. *SYNSEM*) does not bring with it any advantages. Since the grammar becomes more complex due to the meta-constraint, we should reject this change.²⁴ Other changes in the

²³ If *SYNSEM* is reintroduced, the elements in the valence lists could be of type *synsem*. Information about phonology would not be part of the description of the selected elements. This would not solve the problem of partial verb phrase fronting though, since the *LEX* feature is selected for (hence part of the information under *SYNSEM*) but not shared with the filler. One would need a *LOCAL* feature in addition to *SYNSEM*. See Section 10.6.2.2.

²⁴ In Müller (2013b: 253) I claimed that SBCG uses a higher number of features in comparison to other variants of HPSG because of the assumption of the *MOTHER* feature. As Van Eynde (2015) points out this is not true for more recent variants of HPSG since they have the *SYNSEM* feature, which is not needed if *MOTHER* is assumed. (Van Eynde refers to the *LOCAL* feature, but the *LOCAL* feature was eliminated because it was considered superfluous because of the lexical analysis of extraction, see Section 10.6.2.2). If one simply omits the *MOTHER* feature from SBCG one is back to the 1987 version of HPSG (Pollard & Sag 1987), which also used a *SYN* and a *SEM* feature. What would be missing would be the locality of selection (Sag 2012: 149) that was enforced to some extent by the *SYNSEM* feature. Note that the locality of selection that is enforced

feature geometry (elimination of the `LOCAL` feature and use of a single valence feature) are problematic as well. However, if we do reject the revised feature geometry and revert to the feature geometry that was used before, then Sign-Based Construction Grammar and Constructional HPSG (Sag 1997) are (almost) indistinguishable.

10.6.3 Embodied Construction Grammar

Embodied Construction Grammar was developed by Bergen & Chang (2005) and there are some implementations of fragments of German that use this format (Porzel et al. 2006). In the following, I will briefly present the formalism using an example construction. (50) gives the `DetNoun` construction:²⁵

(50)

Construction <code>DetNoun</code>	
subcase of <code>RefExp</code>	
constructional	
<code>d:Determiner</code>	
<code>c:CommonNoun</code>	
self.case	\leftrightarrow <code>d.case</code>
self.number	\leftrightarrow <code>c.number</code>
<code>d.gender</code>	\leftrightarrow <code>c.gender</code>
<code>d.case</code>	\leftrightarrow <code>c.case</code>
<code>d.number</code>	\leftrightarrow <code>c.number</code>
form	
<code>d.f before c.f</code>	
meaning	
self.m	\leftrightarrow <code>c.m</code>

This representational form is reminiscent of PATR-II grammars (Shieber, Uszkoreit, Pereira, Robinson & Tyson 1983): as in PATR-II, the daughters of a construction are given names. As such, (50) contains the daughters `c` and `d`. `d` is a determiner and `c` is a common noun. It is possible to refer to the construction itself via the object **self**. Constructions (and also their daughters) are feature-value descriptions. Structure sharing is represented by path equations. For example, `d.gender` \leftrightarrow `c.gender` states that the value of the gender feature of the determiner is identical to the gender feature of the noun. As well as restrictions on features, there are restrictions on the form. `d.f before c.f` states that the form contribution of the determiner must occur before that of the noun. Bergen & Chang (2005) differentiate between immediate (**meets**) and non-immediate precedence (**before**). Part of the information represented under `f` is the orthographic form (`f.orth`). The inheritance relation is given explicitly in the construction as in Kay & Fillmore (1999).

by `SYNSEM` can be circumvented by the use of relational constraints as well (see Frank Richter and Manfred Sailer's work on collocations (Richter & Sailer 1999a; Soehn & Sailer 2008)). So in principle, we end up with style guides in this area of grammar as well.

²⁵ For a similar construction, see Bergen & Chang (2005: 162).

The construction in (50) can be represented in a similar way to the format used in Chapter 6: (51) shows how this is done. The structure in (51) corresponds to a construction

$$(51) \left[\begin{array}{l} \text{DetNoun} \\ \text{F|ORTH} \quad \boxed{1} \oplus \boxed{2} \\ \text{CASE} \quad \boxed{3} \\ \text{NUMBER} \quad \boxed{4} \\ \text{M} \quad \boxed{5} \\ \\ \text{DTRS} \quad \left\langle \begin{array}{l} \text{Determiner} \\ \text{F|ORTH} \quad \boxed{1} \\ \text{CASE} \quad \boxed{3} \\ \text{NUMBER} \quad \boxed{4} \\ \text{GENDER} \quad \boxed{6} \end{array}, \begin{array}{l} \text{CommonNoun} \\ \text{F|ORTH} \quad \boxed{2} \\ \text{CASE} \quad \boxed{3} \\ \text{NUMBER} \quad \boxed{4} \\ \text{GENDER} \quad \boxed{6} \\ \text{M} \quad \boxed{5} \end{array} \right\rangle \end{array} \right]$$

where the determiner directly precedes the noun because the form contribution of the determiner has been combined with that of the noun. This strict adjacency constraint makes sense as the claim that the determiner must precede the noun is not restrictive enough since sequences such as (52b) would be allowed:

- (52) a. [dass] die Frauen Türen öffnen
 that the women doors open
 ‘that the woman open doors’
 b. *die Türen öffnen
 Frauen

If discontinuous phrases are permitted, *die Türen* ‘the doors’ can be analyzed with the DetNoun Construction although another noun phrase intervenes between the determiner and the noun (Müller 1999a: 424; 1999c). The order in (52b) can be ruled out by linearization constraints or constraints on the continuity of arguments. If we want the construction to require that the determiner and noun be adjacent, then we would simply use **meets** instead of **before** in the specification of the construction.

This discussion has shown that (51) is more restrictive than (50). There are, however, contexts in which one could imagine using discontinuous constituents such as the deviant one in (52b). For example, discontinuous constituents have been proposed for verbal complexes, particle verbs and certain coordination data (Wells 1947). Examples for analyses with discontinuous constituents in the framework of HPSG are Reape (1994), Kathol (1995), Kathol (2000), Crysmann (2008), and Beavers & Sag (2004).²⁶ These analyses, which are discussed in Section 11.7.2.2 in more detail, differ from those previously presented in that they use a feature **DOMAIN** instead of or in addition to the daughters features. The value of the **DOMAIN** feature is a list containing the head and the elements

²⁶ Crysmann, Beaver and Sag deal with coordination phenomena. For an analysis of coordination in TAG that also makes use of discontinuous constituents, see Sarkar & Joshi (1996) and Section 21.6.2.

dependent on it. The elements do not have to necessarily be adjacent in the utterance, that is, discontinuous constituents are permitted. Which elements are entered into this list in which way is governed by the constraints that are part of the linguistic theory. This differs from the simple **before** statement in ECG in that it is much more flexible and in that one can also restrict the area in which a given element can be ordered since elements can be freely ordered inside their domain only.

There is a further difference between the representation in (50) and the general HPSG schemata: in the ECG variant, linearization requirements are linked to constructions. In HPSG and GPSG, it is assumed that linearization rules hold generally, that is, if we were to assume the rules in (53), we would not have to state for each rule explicitly that shorter NPs tend to precede longer ones and that animate nouns tend to occur before inanimate ones.

- (53)
- a. $S \rightarrow \text{NP}[\text{nom}], \text{NP}[\text{acc}], V$
 - b. $S \rightarrow \text{NP}[\text{nom}], \text{NP}[\text{dat}], V$
 - c. $S \rightarrow \text{NP}[\text{nom}], \text{NP}[\text{dat}], \text{NP}[\text{acc}], V$
 - d. $S \rightarrow \text{NP}[\text{nom}], \text{NP}[\text{acc}], \text{PP}, V$

It is possible to capture these generalizations in ECG if one specifies linearization constraints for more general constructions and more specific constructions inherit them from these. As an example, consider the Active-Ditransitive Construction discussed by Bergen & Chang (2005: 170):

- (54)
- | |
|---|
| Construction Active-Ditransitive |
| subcase of Pred-Expr |
| constructional |
| agent:Ref-Expr |
| action:Verb |
| recipient:Ref-Expr |
| theme:Ref-Expr |
| ... |
| form |
| agent.f before action.f |
| action.f meets recipient.f |
| recipient.f meets theme.f |
| meaning |
| ... |

These restrictions allow the sentences in (55a,b) and rule out those in (55c):

- (55)
- a. Mary tossed me a drink.
 - b. Mary happily tossed me a drink.
 - c. * Mary tossed happily me a drink.

The restriction *agent.f before action.f* forces an order where the subject occurs before the verb but also allows for adverbs to occur between the subject and the verb. The other constraints on form determine the order of the verb and its object: the recipient must be adjacent to the verb and the theme must be adjacent to the recipient. The requirement that an agent in the active must occur before the verb is not specific to ditransitive constructions. This restriction could therefore be factored out as follows:

- (56)
- | |
|--|
| Construction Active-Agent-Verb
subcase of Pred-Expr
constructional
agent:Ref-Expr
action:Verb
form
agent.f before action.f |
|--|

The Active-Ditransitive Construction in (54) would then inherit the relevant information from (56).

In addition to the descriptive means used in (50), there is the evokes operator (Bergen & Chang 2005: 151–152). An interesting example is the representation of the term *hypotenuse*: this concept can only be explained by making reference to a right-angled triangle (Langacker 1987: Chapter 5). Chang (2008: 67) gives the following formalization:

- (57)
- | |
|--|
| Schema <i>hypotenuse</i>
subcase of line-segment
evokes right-triangle as rt
constraints
self ↔ rt.long-side |
|--|

This states that a *hypotenuse* is a particular line segment, namely the longest side of a right-angled triangle. The concept of a right-angled triangle is activated by means of the evokes operator. Evokes creates an instance of an object of a certain type (in the example, *rt* of type *right-triangle*). It is then possible to refer to the properties of this object in a schema or in a construction.

The feature description in (58) is provided in the notation from Chapter 6. It is the equivalent to (57).

- (58)
- $$\boxed{1} \left[\begin{array}{l} \textit{hypotenuse} \\ \text{EVOKES} \left\langle \left[\begin{array}{l} \textit{right-triangle} \\ \text{LONG-SIDE} \boxed{1} \end{array} \right] \right\rangle \end{array} \right]$$

The type *hypotenuse* is a subtype of *line-segment*. The value of EVOKES is a list since a schema or construction can evoke more than one concept. The only element in this list in (58) is an object of type *right-triangle*. The value of the feature LONG-SIDE is identified

with the entire structure. This essentially means the following: I, as a hypotenuse, am the long side of a right-angled triangle.

Before turning to FCG in the next subsection, we can conclude that ECG and HPSG are notational variants.

10.6.4 Fluid Construction Grammar

Van Trijp (2013, 2014) claims that SBCG and HPSG are fundamentally different from Fluid Construction Grammar (FCG). He claims that the former approaches are generative ones while the latter is a cognitive-functional one. I think that it is not legitimate to draw these distinctions on the basis of what is done in FCG.²⁷ I will comment on this at various places in this section. I first deal with the representations that are used in FCG, talk about argument structure constructions, the combination operations fusion and merging that are used in FCG and then provide a detailed comparison of FCG and SBCG/HPSG.

10.6.4.1 General remarks on the representational format

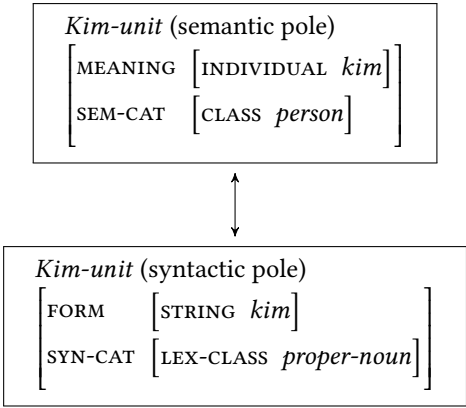
Fluid Construction Grammar (FCG, Steels 2011a) is similar to HPSG in that it uses attribute value matrices to represent linguistic objects. However, these AVMs are untyped as in LFG. Since there are no types, there are no inheritance hierarchies that can be used to capture generalizations, but one can use macros to reach similar effects. Constructions can refer to more general constructions (van Trijp 2013: 105). Every AVM comes with a name and can be depicted as follows:

(59)
$$\boxed{\begin{array}{l} \textit{unit-name} \\ \left[\begin{array}{ll} \text{FEATURE}_1 & \textit{value}_1 \\ \dots & \\ \text{FEATURE}_n & \textit{value}_n \end{array} \right] \end{array}}$$

Linguistic objects have a form and a meaning pole. The two poles could be organized into a single feature description by using a SYN and a SEM feature, but in FCG papers the two poles are presented separately and connected via a double arrow. (60) is an example:

(60) The name *Kim* according to van Trijp (2013: 99):

²⁷ Steels (2013: 153) emphasizes the point that FCG is a technical tool for implementing constructionist ideas rather than a theoretical framework of its own. However, authors working with the FCG system publish linguistic papers that share a certain formal background and certain linguistic assumptions. So this section addresses some of the key assumptions made and some of the mechanisms used.



Depending on the mode in which the lexical items are used, the syntactic pole or the semantic pole is used first. The first processing step is a matching phase in which it is checked whether the semantic pole (for generation) or the syntactic pole (for parsing) matches the structure that was build so far. After this test for unification, the actual unification, which is called merging, is carried out. After this step, the respective other pole (syntax for generation and semantics for parsing) is merged. This is illustrated in Figure 10.2.

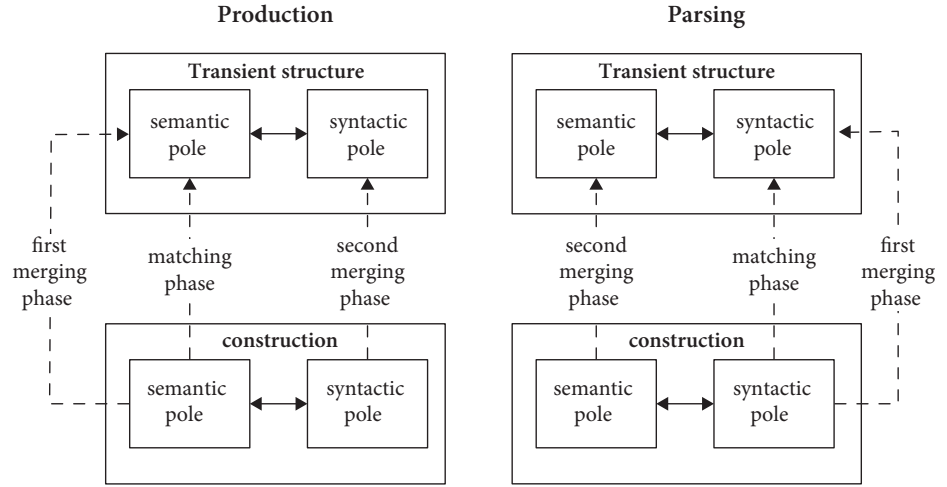


图 10.2: Generation and parsing in FCG (van Trijp 2013: 99)

10.6.4.2 Argument Structure Constructions

Fluid Construction Grammar assumes a phrasal approach to argument structure, that is, it is assumed that lexical items enter into phrasal configurations that contribute independent meaning (van Trijp 2011). The FCG approach is one version of implementing Goldberg's plugging approach to argument structure constructions (Goldberg 1995). Van Trijp suggests that every lexical item comes with a representation of potential argument roles like Agent, Patient, Recipient, and Goal. Phrasal argument structure constructions are combined with the respective lexical items and realize a subset of the argument roles, that is they assign them to grammatical functions. Figure 10.3 on the next page shows an example: the verb *sent* has the semantic roles Agent, Patient, Recipient, and Goal (upper left of the figure). Depending on the argument structure construction that is chosen, a subset of these roles is selected for realization.²⁸ The figures show the relation between sender, sent, and sendee and more the more abstract semantic roles and the relation between these roles and grammatical functions for the sentences in (61):

- (61) a. He sent her the letter.
 b. He sent the letter.
 c. The letter was sent to her.

While in (61a) the agent, the patient and the recipient are mapped to grammatical functions, only the agent and the patient are mapped to grammatical functions in (61b). The recipient is left out. (61c) shows an argument realization in which the sendee is realized as a *to* phrase. According to van Trijp this semantic role is not a recipient but a goal.

Note that under such an approach, it is necessary to have a passive variant of every active construction. For languages that allow for the combination of passive and impersonal constructions, one would be forced to assume a transitive-passive-impersonal construction. As was argued in Müller (2006: Section 2.6) free datives (*commodi/incommodi*) in German can be added to almost any construction. They interact with the dative passive and hence should be treated as arguments. So, for the resultative construction one would need an active variant, a passive variant, a variant with dative argument, a variant with dative argument and dative passive, and a middle variant. While it is technically possible to list all these patterns and it is imaginable that we store all this information in our brains, the question is whether such listings really reflect our linguistic knowledge. If a new construction comes into existence, lets say an active sentence pattern with a nominative and two datives in German, wouldn't we expect that this pattern can be used in the passive? While proposals that establish relations between active and passive con-

²⁸ It is interesting to note here that van Trijp (2011: 141) actually suggests a lexical account since every lexical item is connected to various phrasal constructions via coapplication links. So every such pair of a lexical item and a phrasal construction corresponds to a lexical item in Lexicalized Tree Adjoining Grammar (LTAG). See also Müller & Wechsler (2014a: 25) on Goldberg's assumption that every lexical item is associated with phrasal constructions.

Note that such coapplication links are needed since without them the approach cannot account for cases in which two or more argument roles can only be realized together but not in isolation or in any other combination with other listed roles.

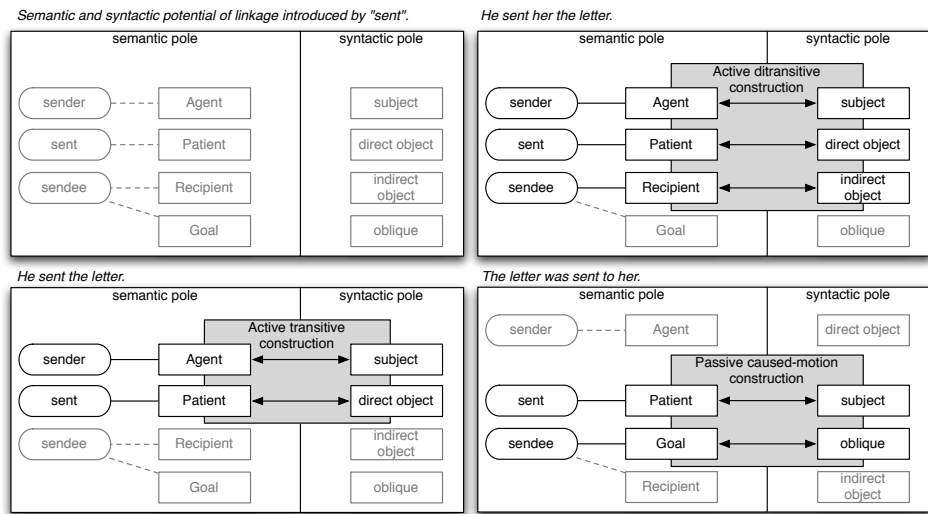


图 10.3: Lexical items and phrasal constructions. Figure from van Trijp (2011: 122)

structions would predict this, alternative proposals that just list the attested possibilities do not.

The issue of how such generalizations should be captured was discussed in connection with the organization of the lexicon in HPSG (Flickinger 1987; Meurers 2001). In the lexical world, one could simply categorize all verbs according to their valence and say that *loves* is a transitive verb and the passive variant *loved* is an intransitive verb. Similarly *gives* would be categorized as a ditransitive verb and *given* as a strictly transitive one. Obviously this misses the point that *loved* and *given* share something: they both are related to their active form in a systematic way. This kind of generalization is captured by lexical rules that relate two lexical items. The respective generalizations that are captured by lexical rules are called a horizontal generalizations as compared to vertical generalizations, which describe relations between subtypes and supertypes in an inheritance hierarchy.

The issue is independent of the lexical organization of knowledge, it can be applied to phrasal representations as well. Phrasal constructions can be organized in hierarchies (vertical), but the relation between certain variants is not covered by this. The analog to the lexical rules in a lexical approach are GPSG-like metarules in a phrasal approach. So what seems to be missing in FCG is something that relates phrasal patterns, e.g., allo-constructions (Cappelle 2006; Goldberg 2014: 116, see also footnote 12).

10.6.4.3 Fusion, matching and merging

As was pointed out by Dowty (1989: 89–90), checking for semantic compatibility is not sufficient when deciding whether a verb may enter (or be fused with) a certain con-

struction. The example is the contrast between *dine*, *eat*, and *devour*. While the thing that is eaten may not be realized with *dine*, its realization is optional with *eat* and obligatory with *devour*. So the lexical items have to come with some information about this.

Van Trijp (2011) and Steels & van Trijp (2011) make an interesting suggestion that could help here: every verb comes with a list of potential roles and argument structure constructions can pick subsets of these roles (see Figure 10.3). This is called *matching*: introducing new argument roles is not allowed. This would make it possible to account for *dine*: one could say that there is something that is eaten, but that no Theme role is made available for linking to the grammatical functions. This would be a misuse of thematic roles for syntactic purposes though since *dine* is semantically a two-place predicate. To account for the extension of argument roles as it is observed in the Caused-Motion Construction (Goldberg 1995: Chapter 7), Steels & van Trijp (2011) suggest a process called *merging*. Merging is seen as a repair strategy: if an utterance involves an intransitive verb and some other material, the utterance cannot be processed with matching alone. For example, when processing Goldberg's example in (62), *he sneezed* could be parsed, but *the foam* and *off the cappuccino* would be unintegrated (see Chapter 21 for an extended discussion of such constructions).

(62) He sneezed the foam off the cappuccino.²⁹

So, Steels & van Trijp (2011: 319–320) suggest that only if regular constructions cannot apply, merging is allowed. The problem with this is that human language is highly ambiguous and in the case at hand this could result in situations in which there is a reading for an utterance, so that the repair strategy would never kick in. Consider (63):³⁰

(63) Schlag den Mann tot!
beat the man dead
'Beat the man to death!' or 'Beat the dead man!'

(63) has two readings: the resultative reading in which *tot* 'dead' expresses the result of the beating and another reading in which *tot* is a depictive predicate. The second reading is dispreferred, since the activity of beating dead people is uncommon, but the structure is parallel to other sentences with depictive predicates:

(64) Iss den Fisch roh!
eat the fish raw

The depictive reading can be forced by coordinating *tot* with a predicate that is not a plausible result predicate:

(65) Schlag ihn tot oder lebendig!
beat him dead or alive
'Beat him when he is dead or while he is alive!'

²⁹ Goldberg (2006: 42).

³⁰ I apologize for these examples ...

So, the problem is that (63) has a reading which does not require the invocation of the repair mechanism: *schlug* ‘beat’ is used with the transitive construction and *tot* is an adjunct (see Winkler 1997). However, the more likely analysis of (63) is the one with the resultative analysis, in which the valence frame is extended by an oblique element. So this means that one has to allow the application of merging independent of other analyses that might be possible. As Steels & van Trijp (2011: 320) note, if merging is allowed to apply freely, utterances like (66a) will be allowed and of course (66b) as well.

- (66) a. *She sneezed her boyfriend.
b. *She dined a steak.

In (66) *sneeze* and *dined* are used in the transitive construction.

The way out of this dilemma is to establish information in lexical items that specifies in which syntactic environments a verb can be used. This information can be weighted and for instance the probability of *dine* to be used transitively would be extremely low. Steels and van Trijp would connect their lexical items to phrasal constructions via so-called coapplication links and the strength of the respective link would be very low for *dine* and the transitive construction and reasonably high for *sneeze* and the Caused-Motion Construction. This would explain the phenomena (and in a usage-based way), but it would be a lexical approach, as it is common in CG, HPSG, SBCG, and DG.

10.6.4.4 Long-distance dependencies

Van Trijp (2014) compares the SLASH-based approaches that are used in GPSG, HPSG, and SBCG with the approach that he suggests within the framework of FCG. He claims that there are fundamental differences between SBCG and FCG and assigns SBCG to the class of generative grammars, while placing FCG in the class of cognitive-functional approaches. He claims that his cognitive-functional approach is superior in terms of completeness, explanatory adequacy, and theoretical parsimony (p. 2). What van Trijp (2014) suggests is basically an analysis that was suggested by Reape (2000) in unpublished work (see Reape (1994) for a published version of an linearization-based approach and Kathol (2000); Müller (1996c, 1999a, 2002a) for linearization-based approaches that despite of being linearization-based assume the SLASH approach for nonlocal dependencies). Van Trijp develops a model of grammar that allows for discontinuous constituents and just treats the serialization of the object in sentences like (67) as an alternative linearization option.

- (67) a. This book, I read.
b. What did the boy hit?

Van Trijp’s analysis involves several units that do not normally exist in phrase structure grammars, but can be modeled via adjacency constraints or represent relations between items which are part of lexical representations in HPSG/SBCG anyway. An example is the subject-verb anchor that connects the subject and the verb to represent the fact that these two items play an important functional role. Figure 10.4 在下一页 shows the analysis of (68).

(68) What did the boy hit?

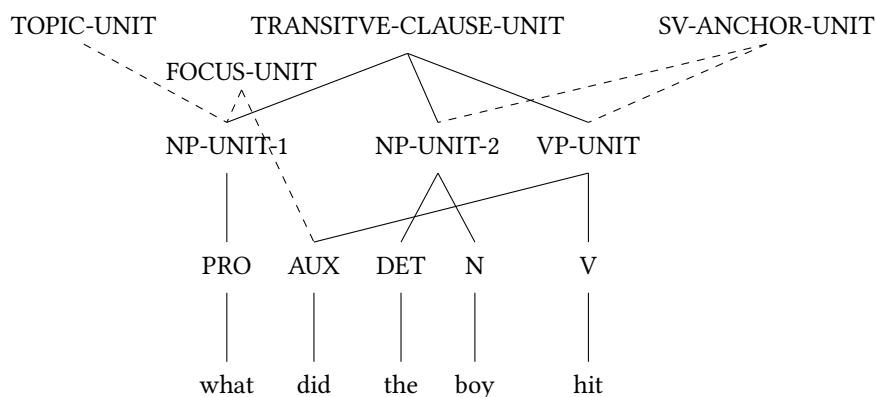


图 10.4: The analysis of *What did the boy hit?* according to van Trijp (2014: 265)

As can be seen in the figure, van Trijp also refers to information structural terms like topic and focus. It should be noted here that the analysis of information structure has quite some history in the framework of HPSG (Engdahl & Vallduví 1996; Kuhn 1995, 1996; Günther et al. 1999; Wilcock 2001; De Kuthy 2002; Paggio 2005; Bildhauer 2008; Bildhauer & Cook 2010). The fact that information structure is not talked about in syntax papers like Sag (2012) does not entail that information structure is ignored or should be ignored in theories like HPSG and SBCG. So much for completeness. The same holds of course for explanatory adequacy. This leaves us with theoretical parsimony, but before I comment on this, I want to discuss van Trijp's analysis in a little bit more detail in order to show that many of his claims are empirically problematic and that his theory therefore cannot be explanatory since empirical correctness is a precondition for explanatory adequacy.

Van Trijp claims that sentences with nonlocal dependency constructions in English start with a topic.³¹ Bresnan's sentences in (2) and (3) were discussed on page 199 (Bresnan 2001: 97) and are repeated below for convenience:

(69) Q: What did you name your cat?
A: Rosie I named her. (*Rosie* = FOCUS)

(70) Q: What did you name your pets?
A: My dog, I named Harold. My cat, I named Rosie. (*my dog, my cat* = TOPIC)

These sentences show that the pre-subject position is not unambiguously a topic or a focus position. So, a statement saying that the fronted element is a topic is empirically

³¹ Van Trijp (2014: 256) uses the following definitions for topic and focus: "Topicality is defined in terms of aboutness: the topic of an utterance is what the utterance is 'about'. Focality is defined in terms of salience: focus is used for highlighting the most important information given the current communicative setting."

not correct. If this position is to be associated with an information structural function, this association has to be a disjunction admitting both topics and focused constituents.

A further problematic aspect of van Trijp's analysis is that he assumes that the auxiliary *do* is an object marker (p. 10, 22) or a non-subject marker (p. 23). It is true that *do* support is not necessary in subject questions like (71a), but only in (71b), but this does not imply that all items that are followed by *do* are objects.

- (71) a. Who saw the man?
b. Who did John see?

First, *do* can be used to emphasize the verb:

- (72) Who *did* see the man?

Second, all types of other grammatical functions can precede the verb:

- (73) a. Where did you see the man? (adverbial)
b. How tall is the man? (predicative)
c. What did John consider Peter? (predicative)
d. What does this book cost? (adverbial)
e. About what did you talk? (prepositional object)

And finally, even a subject can appear in front of *do* if it is extracted from another clause:

- (74) Who does he think saw this man? (subject)

There is a further empirical problem: approaches that assume that a filler is related to its origin can explain scope ambiguities that only arise when an element is extracted. Compare for instance the sentence in (75a) with the sentences in (75b, c): although the order of *oft* 'often' and *nicht* 'not' in (75a) and (75c) is the same, (75a) is ambiguous but (75c) is not.

- (75) a. Oft liest er das Buch nicht.
often reads he the book not
'It is often that he does not read the book.' or 'It is not the case that he reads the book often.'
b. dass er das Buch nicht oft liest
that he the book not often reads
'that it is not the case that he reads the book often'
c. dass er das Buch oft nicht liest
that he the book often not reads
'that it is often that he does not read the book'

(75a) has the two readings that correspond to (75b) and (75c). A purely linearization-based approach probably has difficulties to explain this. A SLASH-based approach can assume that (75a) has a gap (or some similar means for the introduction of nonlocal

dependencies) at the position of *oft* in (75b) or (75c). The gap information is taken into account in the semantic composition at the site of the gap. This automatically accounts for the observed readings.

Another empirical problem that has to be solved is the existence of extraction path marking languages. Bouma, Malouf & Sag (2001a) list a number of languages in which elements vary depending on the existence or absence of a gap in a constituent they attach to. For instance, Irish has complementizers that have one form if the clause they attach to has an element extracted and another form if it does not. SLASH-based proposals can account for this in a straight-forward way: the fact that a constituent is missing in a phrase is represented in the SLASH value of the trace and this information is percolated up the tree. So even complex structures contain the information that there is a constituent missing inside them. Complementizers that are combined with sentences therefore can select sentences with SLASH values that correspond to their inflection. Van Trijp's answer to this challenge is that all languages are different (van Trijp 2014: 263) and that the evidence from one language does not necessarily mean that the analysis for that language is also appropriate for another language. While I agree with this view in principle (see Section 13.1), I do think that extraction is a rather fundamental property of languages and that nonlocal dependencies should be analyzed in parallel for those languages that have it.

10.6.5 Coordination

One of the success stories of non-transformational grammar is the SLASH-based analysis of nonlocal dependencies by Gazdar (1981b). This analysis made it possible for the first time to explain Ross's Across the Board Extraction (Ross 1967). The examples were already discussed on page 177 and are repeated here for convenience:

- (76) 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)
- c. *The kennel (in) which Mary made and Fido sleeps has been stolen.
(= S/NP & S/PP)

The generalization is that two (or more) constituents can be coordinated if they have identical syntactic categories and identical SLASH values. This explains why *which* and *in which* in (76a,b) can fill two positions in the respective clauses. Now, theories that do not use a SLASH feature for the percolation of information about missing elements have to find different ways to make sure that all argument slots are filled and that the correct correspondence between extracted elements and the respective argument role is established. Note that this is not straightforward in models like the one suggested by van Trijp, since he has to allow the preposition *in* to be combined with some material to the left of it that is simultaneously also the object of *made*. Usually an NP cannot simply be used by two different heads as their argument. As an example consider (77a):

- (77) a. *John said about the cheese that I like.
 b. John said about the cheese that I like it.

If it would be possible to use material several times, a structure for (77a) would be possible in which *the cheese* is the object of the preposition *about* and of the verb *like*. This sentence, however, is totally out: the pronoun *it* has to be used to fill the object slot.

10.6.6 Discontinuous constituents and performance models

Van Trijp points out that SBCG does not have a performance model and contrasts this with FCG. On page 252 he states:

So parsing starts by segmenting the utterance into discrete forms, which are then categorized into words by morphological and lexical constructions, and which can then be grouped together as phrases (see Steels, 2011b, for a detailed account of lexico-phasal processing in FCG). So the parser will find similar constituents for all four utterances, as shown in examples (21–24). Since auxiliary-*do* in example (24) falls outside the immediate domain of the VP, it is not yet recognized as a member of the VP.

All of these phrases are disconnected, which means that the grammar still has to identify the relations between the phrases. (van Trijp 2014: 252)

In his (21)–(24), van Trijp provides several tree fragments that contain NPs for subject and object and states that these have to be combined in order to analyze the sentences he discusses. This is empirically inadequate: if FCG does not make the competence/performance distinction, then the way utterances are analyzed should reflect the way humans process language (and this is what is usually claimed about FCG). However, all we know about human language processing points towards an incremental processing, that is, we process information as soon as it is available. We start to process the first word taking into account all of the relevant aspects (phonology, stress, part of speech, semantics, information structure) and come up with an hypothesis about how the utterance could proceed. As soon as we have two words processed (in fact even earlier: integration already happens during the processing of words) we integrate the second word into what we know already and continue to follow our hypothesis, or revise it, or simply fail. See Section 15.2 for details on processing and the discussion of experiments that show that processing is incremental. So, we have to say that van Trijp's analysis fails on empirical grounds: his modeling of performance aspects is not adequate.

The parsing scheme that van Trijp describes is pretty much similar to those of computational HPSG parsers, but these usually come without any claims about human performance. Modeling human performance is rather complex since a lot of factors play a role. It is therefore reasonable to separate competence and performance and continue to work the way it is done in HPSG and FCG. This does not mean that performance aspects should not be modeled, in fact psycholinguistic models using HPSG have been developed in the past (Konieczny 1996), but developing both a grammar with large coverage and the performance model that combines with it demands a lot of resources.

10.6.7 Discontinuity vs. Subject-Head and Head-Filler Schema

I now turn to parsimony: van Trijp uses a subject-verb anchor construction that combines the subject and the main verb. Because of examples like (78) it must be possible to have discontinuous subject-verb constructions:³²

(78) Peter often reads books.

But if such constructions can be discontinuous one has to make sure that (79b) cannot be an instantiation of the subject-verb construction:

- (79) a. The boy I think left.
b. *I the boy think left.

Here it is required to have some adjacency between the subject and the verb it belongs to, modulo some intervening adverbials. This is modelled quite nicely in phrase structure grammars that have a VP node. Whatever the internal structure of such a VP node may be, it has to be adjacent to the subject in sentences like (78) and (79a) above. The dislocated element has to be adjacent to the complex consisting of subject and VP. This is what the Filler-Head Schema does in HPSG and SBCG. Van Trijp criticizes SBCG for having to stipulate such a schema, but I cannot see how his grammar can be complete without a statement that ensures the right order of elements in sentences with fronted elements.

Van Trijp stated that FCG differs from what he calls generative approaches in that it does not want to characterize only the well-formed utterances of a language. According to him, the parsing direction is much more liberal in accepting input than other theories. So it could well be that he is happy to find a structure for (79b). Note though that this is incompatible with other claims made by van Trijp: he argued that FCG is superior to other theories in that it comes with a performance model (or rather in not separating competence from performance at all). But then (79b) should be rejected both on competence and performance grounds. It is just unacceptable and speakers reject it for whatever reasons. Any sufficiently worked out theory of language has to account for this.

10.6.8 Restricting discontinuity

There is a further problem related to discontinuity. If one does not restrict continuity, then constituent orders like (80b) are admitted by the grammar:

- (80) a. Deshalb klärt, dass Peter kommt, ob Klaus spielt.
therefore resolves that Peter comes whether Klaus plays
'Therefore that Peter comes resolves whether Klaus will play.'

³² Unless modals and tense auxiliaries are treated as main verbs (which they should not in English), constructions with modals seem to be another case where the subject and the main verb are not adjacent:

- (i) a. Peter will read the book.
b. Peter has read the book.

- b. *Deshalb klärt dass ob Peter Klaus kommt spielt.
therefore resolves that whether Peter Klaus comes plays

The interesting thing about the word salad in (80b) is that the constituent order within the *dass* clause and within the *ob* clause is correct. That is, the complementizer precedes the subject, which in turn precedes the verb. The problem is that the constituents of the two clauses are mixed.

In a model that permits discontinuous constituents, one cannot require that all parts of an argument have to be arranged after all parts that belong to another argument since discontinuity is used to account for nonlocal dependencies. So, it must be possible to have *Klaus* before other arguments (or parts of other arguments) since *Klaus* can be extracted. An example of mixing parts of phrases is given in (81):

- (81) Dieses Buch hat der Mann mir versprochen, seiner Frau zu geben, der gestern
this book has the man me promised his wife to give who yesterday
hier aufgetreten ist.
here performed is
'The man who performed here yesterday promised me to give this book to his wife.'

We see that material that refers to *der Mann* 'the man', namely the relative clause *der gestern hier aufgetreten ist* 'who performed here yesterday', appears to the right. And the object of *geben* 'to give', which would normally be part of the phrase *dieses Buch seiner Frau zu geben* 'this book his wife to give' appears to the left. So, in general it is possible to mix parts of phrases, but this is possible in a very restricted way only. Some dependencies extend all the way to the left of certain units (fronting) and others all the way to the right (extraposition). Extraposition is clause-bound, while extraction is not. In approaches like GPSG, HPSG and SBCG, the facts are covered by assuming that constituents for a complete clause are continuous apart from constituents that are fronted or extraposed. The fronted and extraposed constituents are represented in SLASH and EXTRA (Keller 1995; Müller 1999a), respectively, rather than in valence features, so that it is possible to require of constituents that have all their valents saturated to be continuous.

Summing up the discussion of parsimony, it has to be said that van Trijp has to provide the details on how continuity is ensured. The formalization of this is not trivial and only after this is done can FCG be compared with the SLASH-based approach.

In addition to all the points discussed so far, there is a logical flaw in van Trijp's argumentation. He states that:

whereas the filler-gap analysis cannot explain WHY *do*-support does not occur in *wh*-questions where the subject is assigned questioning focus, this follows naturally from the interaction of different linguistic perspectives in this paper's approach. (van Trijp 2014: 263)

The issue here is whether a filler-gap analysis or an analysis with discontinuous constituents is suited better for explaining the data. A correct argumentation against the

filler-gap analysis would require a proof that information structural or other functional constraints cannot be combined with this analysis. This proof was not provided and in fact I think it cannot be provided since there are approaches that integrate information structure. Simply pointing out that a theory is incomplete does not falsify a theory. This point was already made in my review of Boas (2003) and in a reply to Boas (2014). See Müller (2005a: 655–656), Müller (2007b: Chapter 20), and Müller & Wechsler (2014b: Footnote 15).

The conclusion about the FCG analysis of nonlocal dependencies is that there are some empirical flaws that can be easily fixed or assumptions that can simply be dropped (role of *do* as object marker, claim that the initial position in English fronting construction is the topic), some empirical shortcomings (coordination, admittance of illformed utterances with discontinuous constituents), some empirical problems when the analysis is extended to other languages (scope of adjuncts in German), and the parsimony of the analyses is not really comparable since the restrictions on continuity are not really worked out (or at least not published). If the formalization of restrictions on continuity in FCG turns out to be even half as complex as the formalization that is necessary for accounts of nonlocal dependencies (extraction and extraposition) in linearization-based HPSG that Reape (2000) suggested,³³ the SLASH-based analysis would be favorable.

In any case, I do not see how nonlocal dependencies could be used to drive a wedge between SBCG and FCG. If there are functional considerations that have to be taken into account, they should be modeled in both frameworks. In general, FCG should be more restrictive than SBCG since FCG claims to integrate a performance model, so both competence and performance constraints should be operative. I will come back to the competence/performance distinction in the following section, which is a more general comparison of SBCG and FCG.

10.6.8.1 Comparison to Sign-Based Construction Grammar/HPSG

According to van Trijp (2013), there are the differences shown in Table 10.1 on the next page. These differences will be discussed in the following subsections.

10.6.8.1.1 Competence/performance distinction

As for the linguistic approach, the use of the term *generative* is confusing. What van Trijp means – and also explains in the paper – is the idea that one should separate competence and performance. We will deal with both the generative-enumerative vs. constraint-based view and with the competence/performance distinction in more detail in the Chapters 14 and 15, respectively. Concerning the cognitive-functional approach, van Trijp writes:

³³ See Kathol & Pollard (1995) for a linearization-based account of extraposition. This account is implemented in the Babel System (Müller 1996c). See (Müller 1999d) on restricting discontinuity. Linearization-based approaches were argued to not be able to account for apparent multiple frontings in German (Müller 2005c, 2015b) and hence linearization-based approaches were replaced by more traditional variants that allow for continuous constituents only.

表 10.1: Differences between SBCG and FCG according to van Trijp (2013: 112)

Scientific model	Theoretical physics (abstract calculus)	Evolutionary theory (complex adaptive system)
Linguistic approach	Generative (competence model)	Cognitive-functional (parsing and production)
Formalization	Mathematical (amenable for implementation)	Computational (implemented)
Constructions	Static type constraints	Dynamic mappings
Constructicon	Signature and grammar	Open-ended inventory
Processing	Assumption of processing- independence	Bidirectional processing model

The goal of a cognitive-functional grammar, on the other hand, is to explain how speakers express their conceptualizations of the world through language (= *production*) and how listeners analyze utterances into meanings (= *parsing*). Cognitive-functional grammars therefore implement both a competence and a processing model. (van Trijp 2013: 90)

It is true that HPSG and SBCG make a competence/performance distinction (Sag & Wasow 2011). HPSG theories are theories about the structure of utterances that are motivated by distributional evidence. These theories do not contain any hypothesis regarding brain activation, planning of utterances, processing of utterances (garden path effects) and similar things. In fact, none of the theories that are discussed in this book contains an explicit theory that explains all these things. I think that it is perfectly legitimate to work in this way: it is legitimate to study the structure of words without studying their semantics and pragmatics, it is legitimate to study phonology without caring about syntax, it is legitimate to deal with specific semantic problems without caring about phonology and so on, provided there are ways to integrate the results of such research into a bigger picture. In comparison, it is wrong to develop models like those developed in current versions of Minimalism (called Biolinguistics), where it is assumed that utterances are derived in phases (NPs, CPs, depending on the variant of the theory) and then shipped to the interfaces (spell out and semantic interpretation). This is not what humans do (see Chapter 15).³⁴ But if we are neutral with respect towards such issues, we are fine. In fact, there is psycholinguistic work that couples HPSG grammars to performance models (Konieczny 1996) and similar work exists for TAG (Shieber & Johnson 1993; Demberg & Keller 2008).

Finally, there is also work in Construction Grammar that abstracts away from performance considerations. For instance, Adele Goldberg's book from 1995 does not contain a worked out theory of performance facts. It contains boxes in which grammatical functions are related to semantic roles. So this basically is a competence theory as well.

³⁴ Attempts to integrate current Minimalist theories with psycholinguistic findings (Phillips 2003) are incompatible with core principles of Minimalism like the *No Tampering Condition* of Chomsky (2008).

Of course there are statements about how this is connected to psycholinguistic findings, but this is also true for theories like HPSG, SBCG and Simpler Syntax (Jackendoff 2011: 600) that explicitly make the competence/performance distinction.

10.6.8.1.2 Mathematical formalization vs. implementation

The difference between mathematical and computational formalization is a rather strange distinction to make. I think that a formal and precise description is a prerequisite for implementation (see the discussion in Section 3.6.2 and Section 4.7.2). Apart from this, a computer implementation of SBCG is trivial, given the systems that we have for processing HPSG grammars. In order to show this, I want to address one issue that van Trijp discusses. He claims that SBCG cannot be directly implemented. On issues of complexity of constraint solving systems he quotes (Levine & Meurers 2006: Section 4.2.2):

Actual implementations of HPSG typically handle the problem by guiding the linguistic processor using a (rule-based) phrase structure backbone, but the disadvantage of this approach is that the “organization and formulation of the grammar is different from that of the linguistic theory” (Levine & Meurers 2006: Section 4.2.2). (van Trijp 2013: 108)

He concludes:

Applying all these observations to the operationalization of SBCG, we can conclude that an SBCG grammar is certainly amenable for computational implementation because of its formal explicitness. There are at least two computational platforms available, mostly used for implementing HPSG-based grammars, whose basic tenets are compatible with the foundations of SBCG: LKB (Copestake 2002) and TRALE (Richter 2006). However, none of these platforms supports a ‘direct’ implementation of an SBCG grammar as a general constraint system, so SBCG’s performance-independence hypothesis remains conjecture until proven otherwise.

There are two issues that should be kept apart here: efficiency and faithfulness to the theory. First, as Levine and Meurers point out, there were many constraint solving systems at the beginning of the 90’s. So there are computer systems that can and have been used to implement and process HPSG grammars. This is very valuable since they can be used for direct verification of specific theoretical proposals. As was discussed by Levine and Meurers, trying to solve constraints without any guidance is not the most efficient way to deal with the parsing/generation problem. Therefore, additional control-structure was added. This control structure is used for instance in a parser to determine the syntactic structure of a phrase and other constraints will apply as soon as there is sufficient information available for them to apply. For instance, the assignment of structural case happens once the arguments of a head are realized. Now, is it bad to have a phrase structure backbone? One can write down phrase structure grammars that use phrase structure rules that have nothing to do with what HPSG grammars usually do. The systems TRALE (Meurers, Penn & Richter 2002; Penn 2004) and LKB will process them. But

one is not forced to do this. For instance, the grammars that I developed for the Core-Gram project (Müller 2013a, 2015a) are very close to the linguistic theory. To see that this is really the case, let us look at the Head-Argument Schema. The Head-Argument Schema is basically the type *head-argument-phrase* with certain type constraints that are partly inherited from its supertypes. The type with all the constraints was given on page 250 and is repeated here as (82):

$$(82) \quad (\text{syntactic}) \text{ constraints on } \textit{head-argument-phrase}: \left[\begin{array}{l} \textit{head-argument-phrase} \\ \text{SYNSEM|LOC|CAT} \left[\begin{array}{l} \text{HEAD} \quad [1] \\ \text{SUBCAT} \quad [2] \end{array} \right] \\ \text{HEAD-DTR|SYNSEM|LOC|CAT} \left[\begin{array}{l} \text{HEAD} \quad [1] \\ \text{SUBCAT} \quad [2] \oplus \langle [3] \rangle \end{array} \right] \\ \text{NON-HEAD-DTRS} \left\langle [\text{SYNSEM} [3]] \right\rangle \end{array} \right]$$

This can be translated into a phrase structure grammar in a straight-forward way:

$$(83) \quad \begin{array}{ll} \text{a.} & \left[\begin{array}{l} \textit{head-argument-phrase} \\ \text{SYNSEM|LOC|CAT} \left[\begin{array}{l} \text{HEAD} \quad [1] \\ \text{SUBCAT} \quad [2] \end{array} \right] \\ \text{HEAD-DTR} [4] \mid \text{SYNSEM|LOC|CAT} \left[\begin{array}{l} \text{HEAD} \quad [1] \\ \text{SUBCAT} \quad [2] \oplus \langle [3] \rangle \end{array} \right] \\ \text{NON-HEAD-DTRS} \left\langle [5] [\text{SYNSEM} [3]] \right\rangle \end{array} \right] \rightarrow [4], [5] \\ \text{b.} & \left[\begin{array}{l} \textit{head-argument-phrase} \\ \text{SYNSEM|LOC|CAT} \left[\begin{array}{l} \text{HEAD} \quad [1] \\ \text{SUBCAT} \quad [2] \end{array} \right] \\ \text{HEAD-DTR} [4] \mid \text{SYNSEM|LOC|CAT} \left[\begin{array}{l} \text{HEAD} \quad [1] \\ \text{SUBCAT} \quad [2] \oplus \langle [3] \rangle \end{array} \right] \\ \text{NON-HEAD-DTRS} \left\langle [5] [\text{SYNSEM} [3]] \right\rangle \end{array} \right] \rightarrow [5], [4] \end{array}$$

The left hand side of the rule is the mother node of the tree, that is, the sign that is licensed by the schema provided that the daughters are present. The right hand side in (83a) consists of the head daughter [4] followed by the non-head daughter [5]. We have the opposite order in (83b), that is, the head daughter follows the non-head daughter. The two orders correspond to the two orders that are permitted by LP-rules: the head precedes its argument if it is marked INITIAL+ and it follows it if it is marked INITIAL−.

The following code shows how (83b) is implemented in TRALE:

```
arg_h rule (head_argument_phrase,
            synsem:loc:cat:head:initial:minus,
            head_dtr:HeadDtr,
            non_head_dtrs:[NonHeadDtr]
```

```

)
==>
cat> NonHeadDtr,
cat> HeadDtr.

```

A rule starts with an identifier that is needed for technical reasons like displaying intermediate structures in the parsing process in debugging tools. A description of the mother node follows and after the arrow we find a list of daughters, each introduced by the operator `cat>`.³⁵ Structure sharing is indicated by values with capital letters. The above TRALE rule is a computer-readable variant of (83b), but includes the explicit specification of the value of `INITIAL`.

Now, the translation of a parallel schema using a `MOTHER` feature like (84a) into a phrase structure rule is almost as simple:

$$\begin{array}{lcl}
 (84) & \text{a.} & \left[\begin{array}{l} \text{head-argument-cx} \\ \text{MOTHER} | \text{SYNSEM} | \text{LOC} | \text{CAT} \left[\begin{array}{l} \text{HEAD} \quad [1] \\ \text{SUBCAT} \quad [2] \end{array} \right] \\ \text{HEAD-DTR} [4] | \text{SYNSEM} | \text{LOC} | \text{CAT} \left[\begin{array}{l} \text{HEAD} \quad [1] \\ \text{SUBCAT} \quad [2] \oplus \langle [3] \rangle \end{array} \right] \\ \text{NON-HEAD-DTRS} \langle [5] [\text{SYNSEM} [3]] \rangle \end{array} \right] \\
 & \text{b.} & [6] \rightarrow [4], [5] \text{ where } \left[\begin{array}{l} \text{head-argument-cx} \\ \text{MOTHER} [6] | \text{SYNSEM} | \text{LOC} | \text{CAT} \left[\begin{array}{l} \text{HEAD} \quad [1] \\ \text{SUBCAT} \quad [2] \end{array} \right] \\ \text{HEAD-DTR} [4] | \text{SYNSEM} | \text{LOC} | \text{CAT} \left[\begin{array}{l} \text{HEAD} \quad [1] \\ \text{SUBCAT} \quad [2] \oplus \langle [3] \rangle \end{array} \right] \\ \text{NON-HEAD-DTRS} \langle [5] [\text{SYNSEM} [3]] \rangle \end{array} \right]
 \end{array}$$

(84b) is only one of the two phrase structure rules that correspond to (84a), but since the other one only differs from (84b) in the ordering of [4] and [5], it is not given here.

For grammars in which the order of the elements corresponds to the observable order of the daughters in a `DTRS` list, the connection to phrase structure rules is even simpler:

$$(85) \quad [1] \rightarrow [2] \text{ where } \left[\begin{array}{l} \text{construction} \\ \text{MOTHER} \quad [1] \\ \text{DTRS} \quad [2] \end{array} \right]$$

The value of `DTRS` is a list and hence [2] stands for the list of daughters on the right hand side of the phrase structure rule as well. The type *construction* is a supertype of all

³⁵ Other operators are possible in TRALE. For instance, `sem_head` can be used to guide the generator. This is control information that has nothing to do with linguistic theory and not necessarily with the way humans process natural language. There is also a `cats` operator, which precedes lists of daughters. This can be used to implement phrase structures with more than one non-head daughter.

constructions and hence (85) can be used to analyze all phrases that are licensed by the grammar. In fact, (85) is one way to put the meta constraint in (31).

So, this shows that the version of SBCG that has been developed by Sag (2012) has a straightforward implementation in TRALE.³⁶ The question remains whether *SBCG's performance-independence hypothesis remains conjecture until proven otherwise* as van Trijp sees it. The answer is: it is not a conjecture since any of the old constraint-solving systems of the nineties could be used to process SBCG. The question of whether this is efficient is an engineering problem that is entirely irrelevant for theoretical linguistics. Theoretical linguistics is concerned with human languages and how they are processed by humans. So whether some processing system that does not make any claims about human language processing is efficient or not is absolutely irrelevant. Phrase structure-based backbones are therefore irrelevant as well, provided they refer to the grammar as described in theoretical work.

Now, this begs the question whether there is a contradiction in my claims. On page 303 I pointed out that SBCG is lacking a formalization in Richter's framework (Richter 2004). Richter and also Levine & Meurers (2006) pointed out that there are problems with certain theoretically possible expressions and it is these expressions that mathematical linguists care about. So the goal is to be sure that any HPSG grammar has a meaning and that it is clear what it is. Therefore, this goal is much more foundational than writing a single grammar for a particular fragment of a language. There is no such foundational work for FCG since FCG is a specific toolkit that has been used to implement a set of grammars.

10.6.8.1.3 Static constraints vs. dynamic mappings and signature + grammar vs. open-endedness

One very interesting feature of Fluid Construction Grammar is its fluidity, that is there are certain constraints that can be adapted if there is pressure, the inventory of the theory is open-ended, so categories and features can be added if need be.

Again, this is not a fundamental difference between HPSG/SBCG and FCG. An HPSG grammar fragment of a specific language is a declarative representation of linguistic knowledge and as such it of course just represents a certain fragment and does not contain any information how this set of constraints evolved or how it is acquired by speakers. For this we need specific theories about language evolution/language change/ll-language acquisition. This is parallel to what we said about the competence/performance distinction, in order to account for language evolution we would have to have several HPSG grammars and say something about how one developed from the other. This will involve weighted constraints, it will involve recategorization of linguistic items and lots more.³⁷ So basically HPSG has to be extended, has to be paired with a model about

³⁶ A toy fragment of English using a MOTHER feature and phrase structure rules with specifications of the kind given above can be downloaded at <http://hpsg.fu-berlin.de/Fragments/SBCG-TRALE/>.

³⁷ There are systems that use weighted constraints. We had a simple version of this in the German HPSG grammar that was developed in *Verbmobil* project (Müller & Kasper 2000) already. Further theoretical approaches to integrate weighted constraints are Brew (1995) and more recently Guzmán Naranjo (2015).

language evolution in the very same way as FCG is.

10.6.8.1.4 Theoretical physics vs. Darwinian evolutionary theory

Van Trijp compares SBCG and FCG and claims that SBCG follows the model of theoretical physics – like Chomsky does –, while FCG adopts a Darwinian model of science – like Croft does –, the difference being that SBCG makes certain assumptions that are true of all languages, while FCG does not make any a priori assumptions. The fundamental assumptions made in both theories are that the objects that we model are best described by feature value pairs (a triviality). FCG assumes that there is always a syntactic and a semantic pole (fundamental assumption in the system) and researchers working in HPSG/SBCG assume that if languages have certain phenomena, they will be analyzed in similar ways. For instance, if a language has nonlocal dependencies, these will be analyzed via the *SLASH* mechanism. However, this does not entail that one believes that grammars of all languages have a *SLASH* feature. And in fact, there may even be languages that do not have valence features (Koenig & Michelson 2010), which may be a problem for FCG since it relies on the *SYN*-pole for the matching phase. So as far as SBCG is concerned, there is considerable freedom to choose features that are relevant in an analysis, and of course additional features and types can be assumed in case a language is found that provides evidence for this. The only example of a constraint provided by van Trijp that is possibly too strong is the locality constraint imposed by the *MOTHER* feature. The idea about this feature is that everything that is of relevance in a more nonlocal context has to be passed up explicitly. This is done for nonlocal dependencies (via *SLASH*) and for instance also for information concerning the form of a preposition inside of a PP (via *PFORM* or more recently via *FORM*). Certain verbs require prepositional objects and restrict the form of the preposition. For instance, *wait* has to make sure that its prepositional object has the preposition *for* in it. Since this information is usually available only at the preposition, it has to be passed up to the PP level in order to be directly selectable by the governing verb.

(86) I am waiting for my man.

So, assuming strict locality of selection requires that all phenomena that cannot be treated locally have to be analyzed by passing information up. Assuming strict locality is a design decision that does not have any empirical consequences, as far as it does not rule out any language or construction in principle. It just requires that information has to be passed up that needs to be accessed at higher nodes. As I have shown in Section 10.6.2, the locality constraint is easily circumvented even within SBCG and it makes the analysis of idioms unnecessarily complicated and unintuitive, so I suggest dropping the *MOTHER* feature. But even if *MOTHER* is kept, it is not justified to draw a distinction between SBCG and FCG along the lines suggested by van Trijp.

Independent of the *MOTHER* issue, the work done in the CoreGram project (Müller 2013a, 2015a) shows that one can derive generalizations in a bottom-up fashion rather

Usually such weighted constraints are not part of theoretical papers, but there are exceptions as for instance the paper by Briscoe and Copestake about lexical rules (Briscoe & Copestake 1999).

than imposing constraints on grammars in a top-down way. The latter paper discusses Croft's methodological considerations and shows how methodological pitfalls are circumvented in the project. HPSG/SBCG research differs from work in Chomskyan frameworks in not trying to show that all languages are like English or Romance or German or whatever, rather languages are treated on their own as it is common in the Construction Grammar community. This does not imply that there is no interest in generalizations and universals or near universals or tendencies, but again the style of working and the rhetoric in HPSG/SBCG is usually different from the ones in Mainstream Generative Grammar. Therefore, I think that the purported difference between SBCG and FCG does not exist.

10.6.8.1.5 Permissiveness of the theories

Van Trijp claims that HPSG/SBCG is a “generative grammar” since its aim is to account for and admit only grammatical sentences. FCG on the other hand is more permissive and tries to get the most out of the input even if it is fragmentary or ungrammatical (see also Steels 2013: 166). While it is an engineering decision to be able to parse ungrammatical input – and there are most certainly systems for the robust processing of HPSG grammars (Kiefer, Krieger & Nederhof 2000; Copestake 2007), it is also clear that humans cannot parse everything. There are strong constraints whose violations cause measurable effects in the brain. This is something that a model of language (that includes competence and performance factors or does not make the difference at all) has to explain. The question is what the cause of deviance is: is it processing complexity? Is it a category mismatch? A clash in information structure? So, if FCG permits structures that are not accepted by human native speakers and that do not make any sense whatsoever, additional constraints have to be added. If they are not added, the respective FCG theory is not an adequate theory of the language under consideration. Again, there is no difference between HPSG/SBCG and FCG.

10.6.8.1.6 A note on engineering

A problematic property of work done in FCG is that linguistic and engineering aspects are mixed.³⁸ Certain bookkeeping features that are needed only for technical reasons appear in linguistic papers, technical assumptions that are made to get a parser running are mixed with linguistic constraints. Bit vector encodings that are used to represent case information are part of papers about interesting case systems. There is certainly nothing wrong with bit vector encodings. They are used in HPSG implementations as well (Reape 1991: 55; Müller 1996c: 269), but this is not mixed into the theoretical papers.

³⁸ This is not a problem if all FCG papers are read as papers documenting the FCG-system (see Footnote 27 on page 313) since then it would be necessary to include these technical details. If the FCG papers are to be read as theoretical linguistics papers that document a certain Construction Grammar analysis, the Lisp statements and the implementational details are simply an obstacle.

It was a big breakthrough in the 80's when theoretical linguists and computational linguists started working together and developed declarative formalisms that were independent of specific parsers and processing systems. This made it possible to take over insights from a lot of linguists who were not concerned with the actual implementation but took care of finding linguistic generalizations and specifying constraints. Since this separation is given up in FCG, it will remain an engineering project without much appeal to the general linguist.

10.7 Summary and classification

There are currently three formalized variants of Construction Grammar: Sign-Based Construction Grammar, Embodied Construction Grammar, and Fluid Construction Grammar. The first two variants can be viewed as notational variants of (Constructional) HPSG (for SBCG with regard to this point, see Sag (2007: 411) and Sag (2010: 486)), or put differently, sister theories of HPSG. This is also true to a large extent for FCG, although van Trijp (2013) spends 25 pages working out the alleged differences. As I have shown in Section 10.6.4, HPSG and FCG are rather similar and I would say that these theories are sister theories as well.

Due to the origins of all three theories, respective analyses can differ quite considerably: HPSG is a strongly lexicalized theory, where phrasal dominance schemata have only been increasingly more used in the last ten years under the influence of Ivan Sag. The phrasal dominance schemata that Ivan Sag uses in his work are basically refinements of schemata that were present in earlier versions of HPSG. Crucially, all phenomena that interact with valence receive a lexical analysis (Sag, Boas & Kay 2012: Section 2.3). In CxG, on the other hand, predominantly phrasal analyses are adopted due to the influence of Adele Goldberg.

As already emphasized in Chapter 9, these are only tendencies that do not apply to all researchers working in the theories in question.

Exercises

1. Find three examples of utterances whose meaning cannot be derived from the meaning of the individual words. Consider how one could analyze these examples in Categorical Grammar (yes, Categorical Grammar).

Further reading

There are two volumes on Construction Grammar in German: Fischer & Stefanowitsch (2006) and Stefanowitsch & Fischer (2008). Deppermann (2006) discusses Construction Grammar from the point of view of conversational analysis. The 37(3) volume of the *Zeitschrift für germanistische Linguistik* from 2009 was also devoted to Construction Grammar. Goldberg (2003a) and Michaelis (2006) are overview articles in English. Goldberg's books constitute important contributions to Construction Grammar (1995;

2006; 2009). Goldberg (1995) has argued against lexical analyses such as those common in GB, LFG, CG, HPSG, and DG. These arguments can be invalidated, however, as will be shown in Section 21.7.1. Sag (1997), Borsley (2006), Jacobs (2008) and Müller & Lipenkova (2009) give examples of constructions that require a phrasal analysis if one wishes to avoid postulating empty elements. Jackendoff (2008) discusses the noun-preposition-noun construction that can only be properly analyzed as a phrasal construction (see Section 21.10). The discussion on whether argument structure constructions should be analyzed phrasally or lexically (Goldberg 1995, 2006; Müller 2006) culminated in a series of papers (Goldberg 2013a) and a target article by Müller & Wechsler (2014a) with several responses in the same volume.

Tomasello's publications on language acquisition (Tomasello 2000, 2003, 2005, 2006c) constitute a Construction Grammar alternative to the Principle & Parameters theory of acquisition as it does not have many of the problems that P&P analyses have (for more on language acquisition, see Chapter 16). For more on language acquisition and Construction Grammar, see Behrens (2009).

Dąbrowska (2004) looks at psycholinguistic constraints for possible grammatical theories.

第十一章 依存语法

依存语法是本书介绍的最古老的一种语法理论。它的现代版本是由法国语言学家 Lucien Tesnière (1893–1954) 进一步发展而来的。他的奠基性著作 *Eléments de syntaxe structurale* (《句法理论要素》) 是在 1938 年完成的, 这是在 Ajdukiewicz 发表范畴语法的论文的三年前 (1935), 但是真正的出版被推迟到 1959 年, 也就是在他去世五年后。因为配价是依存语法的核心, 它有时也叫做配价语法。Tesnière 的思想在今天广为流传。几乎所有的现代理论都有配价和依存的概念 (Ágel & Fischer 2010: 262–263, 284)。

尽管有一些英语方面的文献 (Anderson 1971; Hudson 1984), 依存语法在欧洲中部更为流行, 尤其是在德国 (Engel 1996: 56–57)。Ágel & Fischer (2010: 250) 指出了一个可能的原因: Tesnière 的原始著作直到最近才有英译本 (Tesnière 2015), 但是德语译本都已经有 35 年了 (Tesnière 1980)。由于依存语法的重点在于依存关系, 而不是成分的线性顺序, 通常认为它更适于分析自由语序的语言, 这也是它在斯拉夫语言的研究中比较流行的一个原因: 在 1960 年开始的以 Sgall、Hajičová 和 Panevova 代表的新布拉格学派深入发展了依存语法 (全面的信息请参考 Hajičová & Sgall 2003)。Igor A. Mel'čuk 和 A. K. Žolkovskij 在 1960 年的苏联开始研究一个叫做 Meaning-Text Theory 的模型, 它也用于机器翻译项目中 (Mel'čuk 1964, 1981, 1988; Kahane 2003)。1970 年, Mel'čuk 离开了苏联, 前往加拿大, 现在蒙特利尔工作。

依存语法在德国及世界范围内的学者中广为流传。它在德语作为外语的教学中非常成功 (Helbig & Buscha 1969, 1998)。在东德的莱比锡工作的 Helbig 和 Buscha 开始编辑配价辞典 (Helbig & Schenkel 1969), 之后的研究者在曼海姆的德语语法研究所 (the Institut für Deutsche Sprache) 工作, 他们也展开了类似的辞典编纂工程 (Schumacher et al. 2004)。

下面列出的是可能在德国工作或工作过的语言学家的不完整列表: Vilmos Ágel (2000), 卡塞尔; Klaus Baumgärtner (1965, 1970), 莱比锡之后是斯图加特; Ulrich Engel (1977, 2014), IDS 曼海姆; Hans-Werner Eroms (1985, 1987, 2000), 帕绍; Heinz Happ, 图宾根; Peter Hellwig (1978, 2003), 海德堡; Jürgen Heringer (1996), 奥格斯堡; Jürgen Kunze (1968, 1975), 柏林; Henning Lobin (1993), 基森; Klaus Schubert (1987), 希尔德斯海姆; Heinz Josef Weber (1997), 特里尔; Klaus Welke (1988, 2011), 洪堡大学; Edeltraud Werner (1993), 哈雷-维滕贝格。

尽管从 1959 年起的连续几十年间, 许多国家都有相关的研究, 但是阶段性的

全球性会议直到 2011 年才开始举办。^{1,2}

依存语法很早就被用于计算实现的工程中。Meľčuk 在苏联研究机器翻译，而 David G. Hays 在美国研究机器翻译。Jürgen Kunze 在东德的德国科学院任计算语言学主席，他从 1960 年也开始机器翻译的研究。Kunze (1975) 这本书讲述了语言学研究的公式化背景。还有许多研究者从 1973 年到 1986 年在萨尔布吕肯的合作研究中心 100 — 电子语言学研究 (SFB 100, Elektronische Sprachforschung) 工作。这个研究中心的主要内容也是机器翻译。这些工程研究了俄语到德语、法语到德语、英语到德语以及世界语到德语的翻译。对于这一范围内萨尔布吕肯的工作，请看 Klein (1971)、Rothkegel (1976) 和 Weissgerber (1983)。Muraki et al. (1985) 在一个分析日语和生成英语的工程中使用了依存语法。Richard Hudson 从 1980 年开始研究基于依存语法框架的词语法 (Hudson 1984, 2007)，而 Sleator 和 Temperley 从 1990 年开始研究链语法 (Sleator & Temperley 1991; Grinberg et al. 1995)。Fred Karlsson 的约束语法 (1990) 被应用到许多语言 (可用的大规模的语法片段有丹麦语、葡萄牙语、西班牙语、英语、瑞典语、挪威语、法语、德语、世界语、意大利语和荷兰语) 和学校的教学、语料库标注和机器翻译中。在项目网站上可以观看在线的演示视频。³

近年来，依存语法在计算语言学中越来越受到欢迎。原因是许多标注语料库 (树库) 都包括依存信息。⁴ 统计剖析器在这样的树库中进行训练 (Yamada & Matsumoto 2003; Attardi 2006; Nivre 2003; Kübler et al. 2009; Bohnet 2010)。许多句法剖析器都适用于多种语言，因为普遍的方法是独立于语言的。对依存关系进行一致的标注是较为容易的，因为可供选择的是比较少的。不过支持基于短语成分模型的句法学家会关心二义的分支结构还是平铺模型，说明语的高附加性或低附加性，空成分和无空成分这些问题，并且会对这些问题进行激烈的讨论。话语中的依存关系是比较清楚的。由此，我们可以很容易地保持标注的一致性，并在这些标注的语料上训练统计剖析器。

除了统计模型，还有所谓的深层处理系统，这些系统依赖于手工构造的、语言学驱动的语法。我已经提到了 Meľčuk 在机器翻译的内容下做了一些工作；Hays & Ziehe (1960) 有针对俄语的剖析器；Starosta & Nomura (1986) 开发了一个使用英语语法的句法剖析器，Jäppinen, Lehtola & Valkonen (1986) 开发了一个用芬兰语演示的剖析器，Hellwig (1986, 2003, 2006) 开发了英语的词语法，Covington (1990) 开发了俄语和拉丁语的剖析器，这个系统可以剖析非连续的成分，还有 Menzel (1998) 实现了一个德语依存语法的高质量的剖析器。在 Kettunen (1986); Lehtola (1986); Menzel & Schröder (1998b) 中还提及了其他计算剖析器。下面列出了有依存语法的语法片段的若干种语言：

- 丹麦语 (Danish) (Bick 2001; Bick & Nygaard 2007)
- 英语 (English) (Muraki et al. 1985; Starosta & Nomura 1986; Lavoie & Rambow 1997; Hudson 1989; Sleator & Temperley 1991; Voutilainen et al. 1992; Iordanskaja et al. 1992; Coch 1996)
- 世界语 (Esperanto) (Bick 2009)

¹ <http://depling.org/>. 2015 年 4 月 10 日。

² Meaning-Text Theory 的会议从 2003 年开始每两年举办一次。

³ http://beta.visl.sdu.dk/constraint_grammar. 2015 年 7 月 24 日。

⁴ 根据 Kay (2000)，Hays 开发的史上第一个树库就标注了依存关系。

- 爱沙尼亚语 (Estonian) (Müürisep 1999; Müürisep, Puolakainen, Muischnek, Koit, Roosmaa & Uibo 2003)
- 法罗语 (Faroese) (Trosterud 2009)
- 芬兰语 (Finnish) (Nelimarkka, Jäppinen & Lehtola 1984; Jäppinen, Lehtola & Valkonen 1986)
- 法语 (French) (Iordanskaja et al. 1992; Coch 1996; Bick 2010)
- 德语 (German) (Hellwig 1986; Coch 1996; Heinecke et al. 1998; Menzel & Schröder 1998a; Hellwig 2003, 2006; Gerdes & Kahane 2001)
- 爱尔兰语 (Irish) (Dhonnchadha & van Genabith 2006)
- 日语 (Japanese) (Muraki, Ichiyama & Fukumochi 1985)
- 拉丁语 (Latin) (Covington 1990)
- 现代汉语 (Mandarin Chinese) (Liu & Huang 2006; Liu 2009)
- 挪威语 (Norwegian) (Hagen, Johannessen & Nøklestad 2000),
- 古冰岛语 (Old Icelandic) (Maas 1977)
- 葡萄牙语 (Portuguese) (Bick 2003)
- 俄语 (Russian) (Hays & Ziehe 1960; Mel'čuk 1964; Covington 1990)
- 西班牙语 (Spanish) (Coch 1996; Bick 2006)
- 斯瓦希里语 (Swahili) (Hurskainen 2006)

约束语法的网页⁵另外列出的还有的巴斯克语 (Basque)、加泰罗尼亚语 (Catalan)、英语 (English)、芬兰语 (Finnish)、德语 (German)、意大利语 (Italian)、萨米语 (Sami) 和瑞典语 (Swedish) 的语法。

11.1 表示形式概述

11.1.1 价、结与卫星结构

依存语法的核心概念是“价”（请看第 1.6 节）。这个概念的核心隐喻是稳定分子的信息，这在化学中解释为电子的层级关系。化合物与语言结构之间的区别在于化合物是没有方向的，也就是说，我们说氧原子在构成水的过程中比氢原子更重要是没有道理的。与之相比，动词比跟它一起构成完整小句的名词短语更为重要。在诸如英语和德语的语言中，动词决定了它的依存成分的形式，比如说他们的格。

描述依存关系的一个方法如图 11.1 on the following page 所示。最高阶的节点是动词 *reads*。它的价是一个主格 NP（主语）和一个宾格 NP（宾语）。

⁵ http://beta.visl.sdu.dk/constraint_grammar_languages.html

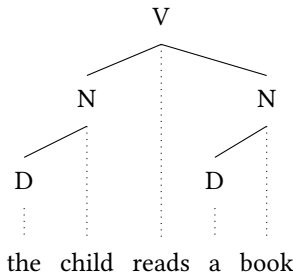


图 11.1: *The child reads a book.* 的分析

这通过在表示动词的节点和表示各自名词的节点之间的依存关系来描述。名词本身需要一个限定词，它是分别通过 *the* 和 *a* 的依存关系来表示的。需要注意的是，这里展示的分析对应于 HPSG 中假定的 NP 分析，即名词选择它的限定词（请看第 9.6.1 节）。不过，需要指出的是，关于 NP 还是 DP 的分析是合理的问题在依存语法社团中也展开了讨论（Hudson 1984: 90、Van Langendonck 1994; Hudson 2004）。请看 Engel (1977) 关于 N 作为中心语的分析以及 Welke (2011: 31) 有关限定词作为中心语的分析。

动词是小句的中心语，名词叫做从属词（dependants）。另一套针对中心语与从属语的术语是结（nucleus）与卫星结构（satellite）。

另一个用来描写依存关系的方法是在依存语法的变体词语法中所使用的（Hudson 2007），如图 11.2 所示。这张图显示了语法功能，而不是词类信息，但是除此以外，它与图 11.1 是一致的。图 11.1 中的最高节点被标记为图 11.2 中的 root 箭头。向下的关系是由箭头的方法来预测的。

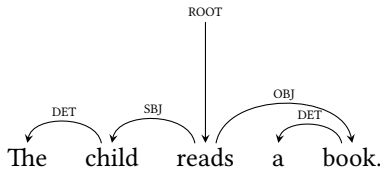
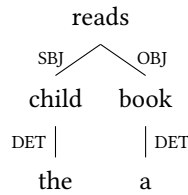


图 11.2: *The child reads a book.* 的分析的另一种表示形式

第三种表示相同的依存关系的形式如图 11.3 在下一页所示，这里又有树格式了。如果我们将图中的根节点向上表示，就得到了这棵树。由于我们对从属词上的结的依存关系有清晰的可视化表示，我们就不用箭头来表示这个信息了。但是，有些依存语法的变体，比如说 Word Grammar，使用了相互依存。由此，比如说，有些理论认为在 *his child* 的分析中，*his* 依存于 *child*，而且 *child* 依存于 *his*。如果必须要描述这种相互依存关系，则需要对所有的依存关系使用箭头或者有些层级树中的依存关系用下划线表示，而其他依存关系用箭头表示。

当然，词类信息也可以加进图 11.2 和图 11.3 中，语法功能标签可以加在图 11.1 中，

图 11.3: *The child reads a book.* 的分析的另一种表示形式

而语序可以加在图 11.3 中。

上图描述了中心语与各自的从属词之间的依存关系。这可以按照 n -ary 规则来表示得更为形式化，该规则与第 2 章讨论的短语结构规则是十分类似的（Gaifman 1965: 305、Hays 1964: 513、Baumgärtner 1970: 61、Heringer 1996: Section 4.1）。比如说，Baumgärtner 提出了 (1) 中的规则：

$$(1) \quad \chi \rightarrow \varphi_1 \dots \varphi_i * \varphi_{i+2} \dots \varphi_n, \text{ where } 0 < i \leq n$$

(1) 中的星号对应于范畴 χ 中的词。在我们的例子中， χ 可以是 V，‘*’ 的位置可以由 *reads* 表示，而且 φ_1 和 φ_3 可以是 N。与 (2b) 中关于限定词—名词组合的规则一起，(2a) 中的规则可以允准图 11.1 中的依存树。

$$(2) \quad \text{a. } V \rightarrow N * N$$

$$\text{b. } N \rightarrow D *$$

另外，许多二叉规则可以被看作是与它的主语、直接宾语或间接宾语相组合 (Kahane 2009)。我们会在第 11.7.2 节详细讨论依存规则，以及依存语法与短语结构语法的对比。

11.1.2 说明语

Tesnière 使用的另一个隐喻是“小戏”。一个事件的核心参与者是行动元 (actants)，除此之外就是背景，即舞台，一般的布置。行动元这个概念在其他理论中是论元，而描述为舞台的实体被叫做场景 (circumstants)。这些场景是修饰语和在本书介绍的其他理论中通常所说的说明语。在依存关系的表示中，依存语法的论元和说明语之间没有太大的区别。图 11.4 on the next page 表示了 (3) 的分析：

$$(3) \quad \text{The child often reads the book slowly.}$$

DET 孩子 经常 读 DET 书 慢慢地

‘孩子们常常慢慢地读书。’

依存关系的标注使用了 Engel (1977) 提出来的描写不同依存关系的技术手段：说明语用一条从说明语节点向上的额外的线来表示（也请看 Erms 2000）。另一种区分论元/说明语，或者行动元/场景元的区分的方法是作为论元和说明语的地位的明确的区别。所以我们可以使用针对说明语和论元的明确标记，就像在前面标记语法功能一样。德语语法和配价辞典经常使用标签 E 和 A 来分别表示 *Ergänzung* 和 *Angabe*。

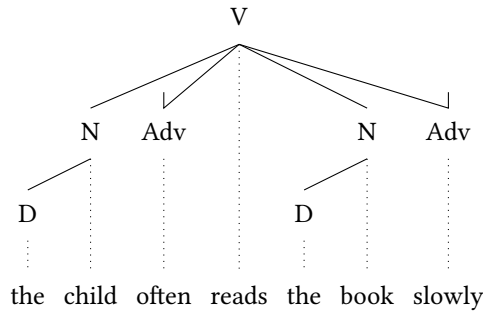


图 11.4: *The child often reads the book slowly.* 的分析

11.1.3 线性

目前，我们已经看到了依存图，它们连接了按照一些线性顺序排列的词。不过，从属词的语序在原则上不是由依存关系决定的，由此依存语法必须使用额外的手段来表示语言对象（词根、形素、词）的合理语序。Engel (2014: 50)提出了针对（4）中的句子的依存图 11.5 在下一页。⁶

- (4)
- a. Gestern war ich bei Tom.
昨天 COP I PREP Tom
‘我昨天跟 Tom 在一起。’
 - b. Ich war gestern bei Tom.
我 COP 昨天 PREP Tom
 - c. Bei Tom war ich gestern.
PREP Tom COP 我 昨天
 - d. Ich war bei Tom gestern.
我 COP PREP Tom 昨天

根据Engel (2014: 50)，正确的语序是由表层语法规则决定的，比如说，证明在陈述性主句的前场位置总有一个成分的规则，以及在第二位是定式动词的规则。^{7,8} 进而，还有关于语用属性的线性规则，比如说旧信息位于新信息之前。另一条规则保证了弱代词被置于前场或中场的开头。这一线性顺序的概念在经验与概念上都是有问题的，我们会在第 11.7.1 节再来讨论。这里需要指出的是，只处理依存关系的方法承认中心语和它的从属词的非连续的实现形式。除了任何更进一步的约束，依存语法还有一个在第 10.6.3 节第 310 页中已经讨论过的有关嵌套构式语法以及第 10.6.4.4 节关于流体构式语法的问题。其中，一个论元可以打断另一个论元，如图 11.6 在下一页所示。为了排除语言中的这种不可能的线性排列，有时会认为分析需要是投影

⁶ Engel 使用 E_{sub} 来表示主语，还有 E_{acc}、E_{dat} 和 E_{gen} 来表示具有不同格的宾语。
⁷ “Die korrekte Stellung ergibt sich dann zum Teil aus oberflächensyntaktischen Regeln (zum Beispiel: im Vorfeld des Konstativsatzes steht immer genau ein Element; das finite Verb steht an zweiter Stelle) [...]”
⁸ Engel (1970: 81)针对 *Vorfeld* 中只有一个成分的说法提出了反例。相关的例子将在第 11.7.1 节讨论。

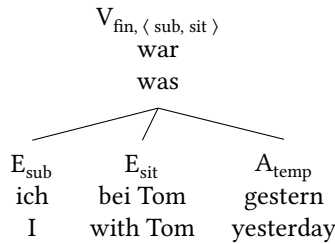


图 11.5: 根据Engel (2014: 50)的 *ich*、*war*、*bei Tom* 和 *gestern* (我昨天跟 Tom 在一起。)的不同语序的依存图

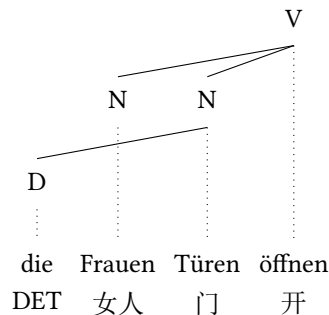


图 11.6: *dass die Frauen Türen öffnen* (这个女人开门) 的不想要的分析

的, 即如图 11.6 中交叉的分支是不被允许的。这实际上重新将直接成分的概念引入了框架中, 因为这就意味着中心语的所有从属词都必须在中心语周边实现, 除非故意使用了特殊的机制 (请看第 11.5 节关于非局部依存的例子)。⁹有些作者明确地使用短语结构成分来构成成分的线性约束 (Gerdes & Kahane 2001; Hellwig 2003)。

11.1.4 语义

Tesnière 已经按照后来理论中常见的语义角色区分了动词的参与者。他指出, 第一个行动元是施事, 第二个是受事, 第三个是受益者 (Tesnière 2015: Chapter 106)。考虑到依存语法是一个基于词汇的框架, 所有关于论元连接的词汇化方法都可以被采用。但是, 论元联接与语义角色指派只是当自然语言表达需要指派意义时亟需解决的一小部分问题。说明语和量词的辖域的问题是需要被解决的, 而且很清楚的是没有考虑到线性语序的表示依存关系的依存图是不够的。一个没有排序的依存图将语法功能指派给中心语的从属词, 并且它在很多方面都类似于 LFG 的 f-

⁹ 尽管这可以得到短语结构语法中所谓的单位 (unit), 这里也是有区别的: 在短语结构语法中, 单位具有功能标签 (如 NP), 这在依存语法是不同的。在依存语法中, 我们只指出中心语的标签 (如图 11.4 中属于 *child* 的 N) 或者直接指出中心词 (如图 11.3 中的词 *child*) 所以在依存语法表示中有更少的节点 (但是请看第 11.7.2.3 节的讨论)。

结构。¹⁰ 对于第 205 页的 (25a) 这类句子，这里重复为 (5)，我们得到第 205 页的 (25b) 中的 f-结构。这个 f-结构包括一个主语 (*David*)、一个宾语 (*a sandwich*)，以及带有两个成分的说明语的集合 (*at noon* 和 *yesterday*)。

- (5) David devoured a sandwich at noon yesterday.
David 吞光 一三明治 在中午 昨天
'David 昨天中午吞光了一个三明治。'

这就是未排序的依存图中的编码形式。由于这个平行特征，Bröker (2003: 308) 针对依存语法也提出了粘着语义学就不足为奇了 (Dalrymple, Lamping & Saraswat 1993, Dalrymple 2001: Chapter 8)。我们在第 7.1.5 节已经介绍过粘着语义学了。

依存语法的有些变体对语义有明确的处理。一个例子是意义文本理论 (Meaning-Text Theory) (Mel'čuk 1988)。词语法是另一个例子 (Hudson 1991: Chapter 7, 2007: Chapter 5)。这里不便介绍这些理论的概念。需要指出的是，像 Hudson 的词语法这类理论对线性顺序是十分严格的，而且并不认为 (4) 中的所有句子都具有相同的依存结构 (请看第 11.5 节)。词语法更接近于短语结构语法，并且具有跟基于成分的理论相同的成分序列与语义的互动关系。

11.2 被动

依存语法是一个基于词汇的理论，而且价是其核心概念。基于这个原因，被动的分析也是基于词汇的方法就不足为奇了。也就是说，一般认为有一个被动分词，它具有与主动动词不同的配价需求 (Hudson 1990: Chapter 12, Eroms 2000: Section 10.3, Engel 2014: 53–54)。

(6) 中的标准例子被分析为图 11.7 在下一页中所示的形式。

- (6) [dass] der Weltmeister geschlagen wird
CONJ DET 世界. 冠军 击败 AUX
'世界冠军被击败了'

这张图是针对被动结构的直觉上的描述。对于人称被动来说，其形式化很有可能会落实到词汇规则上。请看 Hellwig (2003: 629–630) 关于英语中被动分析的词汇规则的明确建议。

请注意，*der Weltmeister* (世界冠军) 不是 Engel 的分析中被动助词 *wird* 的论元。这意味着主语-动词的一致关系不能受限于局部，而且我们需要为了一致关系来开发一些精细化的机制。¹¹ Hudson (1990)、Eroms (2000: Section 5.3) 和 Groß & Osborne (2009) 认为主语依存于助词，而不是助动词。这需要范畴语法 (请看第 8.5.2 节) 和 HPSG 语法 (Hinrichs & Nakazawa 1994) 中较为普遍的论元转换。更合适的分析是将分词的主语当作助词的主语，如图 11.8 在下一页所示。

¹⁰ Tim Osborne (p. c. 2015) 提醒我，所有的情况下并不是这样的：比如说非谓词性介词没有在 f-结构中显示，但是他们必然出现在依存图中。

¹¹ 这个问题对于所谓的远被动来说是更为迫切的问题：

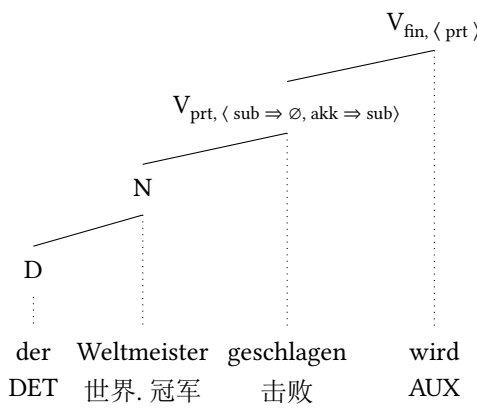


图 11.7: [dass] *der Weltmeister geschlagen wird* (世界冠军被击败了) 的分析类似于Engel (2014: 53–54)提出的分析

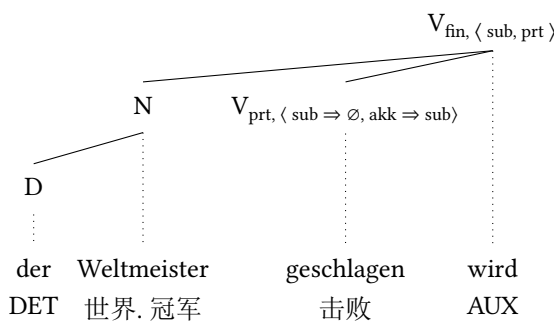


图 11.8: 带有作为助词的从属词的主语的 [dass] *der Weltmeister geschlagen wird* (世界冠军被打了) 的分析

- (i) a. weil der Wagen zu reparieren versucht wurde
 因为DET.SG.NOM 汽车 INF 修理 尝试 AUX
 ‘因为试着修理这辆汽车’
 b. weil die Wagen zu reparieren versucht wurden
 因为DET.PL.NOM 汽车 INF 修理 尝试 AUX
 ‘因为试着修理这辆汽车’

这里，*zu reparieren* 的宾语与助词 *wurde*（单数）和 *wurden*（复数）保持一致，这个宾语是嵌套在两层深的动词的宾语。但是，关于如何分析这些远被动的问题在 Engel 的系统中是开放的，而且这个问题的解决方法可能会包含 HPSG 中应用的机制：*zu reparieren* 的论元被提升到统制动词 *versucht* 上，被动应用于这个动词，并将宾语转化为主语，这是通过助词提升的。这就解释了 *zu reparieren*（修理）的隐含宾语与 *wurde*（单数）之间的一致关系。Hudson (1997) 在词语法的框架下提出了德语的动词性补足语的分析，这个分析包括了他所谓的概化提升（generalized raising）的过程。他指出，主语和补足语一起提升到了统制中心语中。请注意，这样包括概化提升的分析可以直接对 (i) 这类

11.3 动词位置

在许多德语的依存语法的著作中，并没有处理线性顺序问题，而且作者只关注于依存关系。动词及其论元之间的依存关系基本上等同于动词位于首位与动词位于末位的句子。如果我们比较图 11.9和图 11.10中给出的（7）中的例子的依存图的话，我们会看到只有动词的位置是不同的，但是依存关系就如他们应当那样是相同的。¹²

- (7) a. [dass] jeder diesen Mann kennt
CONJ 每人 这 男人 认识
‘每个人都认识这个男人’
b. Kennt jeder diesen Mann?
认识 每人 这 男人
‘每个人都认识这个男人吗？’

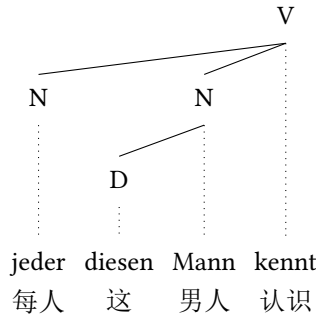


图 11.9: [dass] *jeder diesen Mann kennt*（每个人都认识这个男人）的分析

句子进行分析，因为宾语将依存于与主语相同的中心语上，即 *hat* (AUX)，由此可以放在主语前。

- (ii) Gestern hat sich der Spieler verletzt.
昨天 AUX 自己DET 选手 受伤
‘这个选手昨天伤到了自己。’

关于 Groß & Osborne 对 (ii) 的讨论，请看第 546 页。

¹² Eroms (2000) 用词性 *Pro* 来表示 *jeder* (每人) 这类代词。如果词类的信息在选择中发挥了重要的作用，这就使得管辖名词性表达的中心语的所有配价框架的析取的区分是十分必要的，因为他们要么跟带有内部结构的 NP，要么跟带有介词的 NP 相组合。通过对代词赋予范畴 *N*，我们可以避免这个析取的具体化的问题。代词跟名词的区别在于它的价（当名词需要限定词时，它是完全满足的），而不是它的词性。Eroms & Heringer (2003: 259) 使用符号 *N_{pro}* 来表示代词。如果 *pro*-部分被理解为带有词性的特殊属性对象，这就跟我们上面讲到的内容不一致了：中心语就可以选择 *Ns* 了。如果 *N_{pro}* 和 *N* 被看作是不同的原子符号，问题还是存在的。

使用 *N* 而不是 *Pron* 作为代词的词性，这在依存语法的其他版本中是标准的表示，例如词语法 (Hudson 1990: 167; Hudson 2007: 190)。也请看第 49 页的脚注 2 关于短语结构语法中代词和 NPs 的区别。

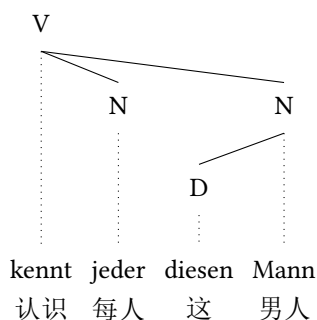


图 11.10: *Kennt jeder diesen Mann?* (每个人都认识这个男人吗?) 的分析

关于带有论元和说明语的动词的正确的语序由线性约束条件来保证，这些条件与各自的空间位置是相关的。请看第 11.1.3 节和第 11.7.1 节更多有关线性的细节问题。

11.4 局部重新排序

与局部重新排序的情况是相同的。(8b) 中句子的依存关系如图 11.11 所示。(8a) 中带有正常语序的句子的分析已经在图 11.9 中给出了。

- (8) a. [dass] jeder diesen Mann kennt
 CONJ 每人 这 男人 认识
 ‘每个人都认识这个男人’
 b. [dass] diesen Mann jeder kennt
 CONJ 这 男人 每人 认识
 ‘每个人都认识这个人’

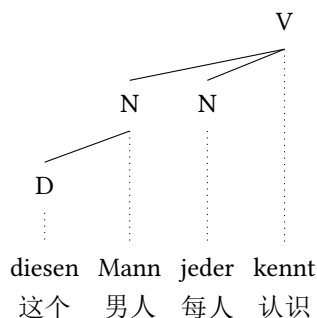


图 11.11: *[dass] diesen Mann jeder kennt* (每个人都认识这个人) 的分析

11.5 长距离依存

在依存语法中有许多方法来分析非局部依存关系。最简单的方法我们已经在前面的章节中看到了。许多分析只关注依存关系，并且认为位于第二位的动词只是一种可能的线性变体（Eroms 2000: Section 9.6.2、Groß & Osborne 2009）。图 11.12展示了（9）的分析：

- (9) [Diesen Mann] kennt jeder.
这 男人 认识 每人
‘每个人都认识这个男人。’

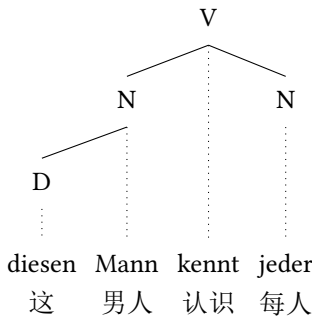


图 11.12: 没有对前置进行特殊处理的 *Diesen Mann kennt jeder.*（这个人，每个人都认识。）的分析

现在，这是最简单的情况，所以让我们看看（10）中的例子，它真正包括非局部（nonlocal）的依存关系：

- (10) Wen_i glaubst du, daß ich _{-i} gesehen habe?¹³
谁.ACC 认为.2SG 你.NOM CONJ 我.NOM 看见 AUX
‘你认为我看见了谁？’

依存关系的描写如图 11.13 在下一页所示。这张图与我们之前看到的很多图不同，它不具有投影性。这意味着有交叉：对于 *wen*（谁）的 V_{prt} 和 N 之间的联系与联接 *glaubst*（认为）和 *du*（你）以及他们的范畴符号的关系有所交叉。根据设想的依存语法的版本，这可以被认为是一个问题，或者不是。让我们来看一下两个选项：如果图 11.13 中给出的类型的非联系性被允许用在 Heringer 和 Eroms 的语法中（Heringer 1996: 261、Eroms 2000: Section 9.6.2），¹⁴这就需要语法中有一些可以排除那些不合乎语法的非连续体。比如说，图 11.14 在下一页中（11）的分析就应该被排除。

- (11) *Wen glaubst ich du, dass gesehen habe?
谁.ACC 认为.2SG 我.NOM 你.NOM CONJ 看见 AUX
想说：“你认为我看见谁了？”

¹³ Scherpenisse (1986: 84).
¹⁴ 但是，作者提到了将提取的成分向更高的节点提升的可能性。请看Eroms & Heringer (2003: 260)。

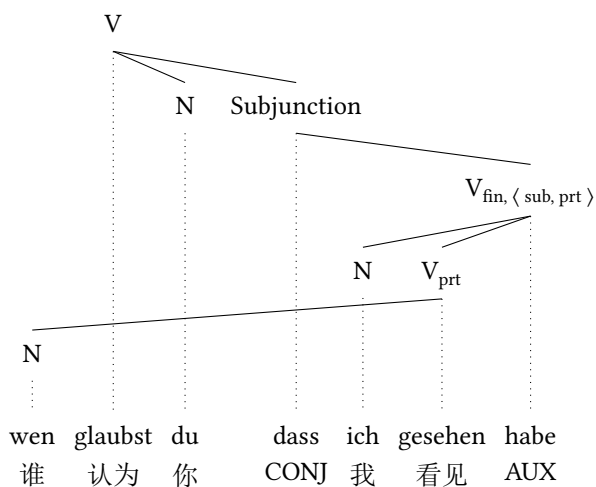


图 11.13: *Wen glaubst du, dass ich gesehen habe?* (你认为我看见了谁?) 的非投影分析

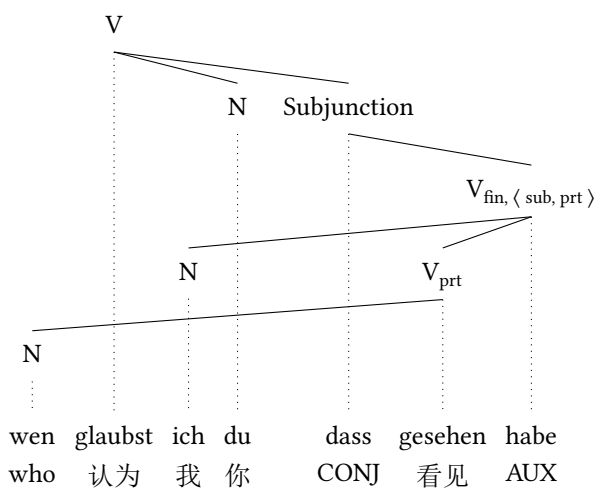


图 11.14: * *Wen glaubst ich du, dass gesehen habe?* (你认为我看见谁了) 的不想得到的依存图

需要注意的是，(11) 中成分的语序与提出的表示空间位置的观点是兼容的Engel (2014: 50): 有一个由 *wen* (谁) 填充的 *Vorfeld*，有一个由 *glaubst* (认为) 填充的句子左边界，还有由 *ich* (我)、*du* (你) 和小句论元填充的 *Mittelfeld*。在 *Mittelfeld* 有像 *ich* (我) 和 *du* (你) 这样的代词是完全正常的。问题在于这些代词来属于不同的类型: *du* 属于主动词 *glaubst* (认为)，而 *ich* (我) 依存于 (*gesehen* 看见) *habe* (AUX)。一个理论需要覆盖的事实是，前置和外置分别瞄准小句的最左边和最右边。这可以直接按照基于成分的方法来模拟，正如我们在前几章所看到的。

作为另一种非连续成分，我们可以假定一个能够在结构中将嵌套的中心语的依存关系提高到更高的节点上的机制。这一分析由Kunze (1968)、Hudson (1997, 2000)、Kahane (1997)、Kahane et al. (1998)和Groß & Osborne (2009)提出。接下来，我会用Groß & Osborne (2009)提出的有关这类分析的例子。Groß & Osborne 用图 11.15中的虚线描述了重新组织的依存关系。^{15, 16} 依存关系 (V_{prt}) 的源头被标记为一个 *g*

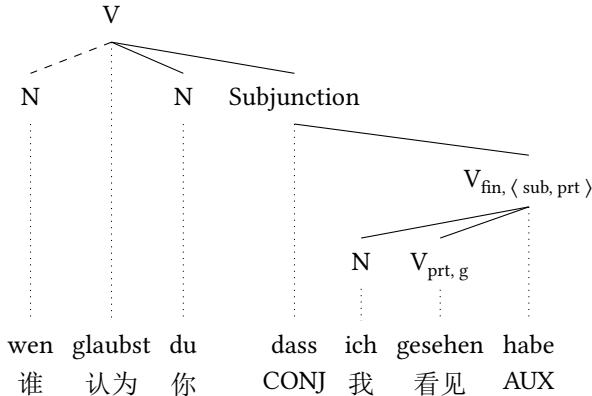


图 11.15: 包括提升的 *Wen glaubst du, dass ich gesehen habe?* (你认为我看到了什么?) 的投影分析

和一个从属词，这个从属词通过虚线连接到他提升（最高的 *V*）的节点上。除了在局部实现 *gesehen* (看见) 的宾格从属词之外，关于缺失成分的信息传递到了更高的节点上并在那里得到实现。

Groß & Osborne (2009)的分析并不是十分准确。这里有一个 *g* 和一条虚线，但是句子可能会包括多重非局部的依存关系。比如说，例 (12) 中，在关系小句 *den wir alle begrüßt haben* (我们都打过招呼的那些人) 和 *die noch niemand hier gesehen hat* (这里没有人看到的那个人) 中有一个非局部的依存关系: 关系代词在关系小句内被前置。短语 *dem Mann, den wir alle kennen* (我们都认识的那个人) 是 *gegeben* (给) 的前置的与格宾语，而且 *die noch niemand hier gesehen hat* (这里还没有人看到的那个人) 是从 *Frau* (女人) 为中心语的 NP 中外置而来的。

¹⁵ Eroms & Heringer (2003: 260)提出了类似的分析，但是没有说明任何形式化的细节。
¹⁶ 需要注意的是，Groß & Osborne (2009)并没有针对简单和复杂的动词位于第二位的句子的一个统一的分析。也就是说，对于可以解释为局部重新排序的情况，他们提出了没有提升的分析。他们对 (9) 的分析如图 11.12所示。这就导致了第 11.7.1节讨论过的问题。

- (12) Dem Mann, den wir alle begrüßt haben, hat die Frau das Buch gegeben,
 DET 男人 谁 我们都 打招呼 AUX AUX DET 女人 DET 书 给
 die noch niemand hier gesehen hat.
 谁 但 没人 这里看见 AUX
 ‘这里没有人看见过的那个女人把这本书给了这个男人，这个我们都打过招呼的男人。’

所以，这就意味着中心语和移位成分之间的联系（依存关系）已经说的很清楚的。这就是Hudson (1997, 2000)在他的词语法中对非局部依存关系的分析：除了连接词与它的主语、宾语和其他成分的依存关系，他还提出了提取成分的更深层次的依存关系。比如说，(10)中的 *wen*（谁），为了简便这里重复为(13)，是 *gesehen*（看见）的宾语以及 *glaubst*（认为）和 *dass*（CONJ）的提取。

- (13) Wen glaubst du, dass ich gesehen habe?
 谁 认为 你 CONJ 我 看见 AUX
 ‘你认为我看见了谁？’

Hudson 证明了词语法中多重依存关系的使用对应于 HPSG 中的结构共享(Hudson 1997: 15)。非局部依存关系被模拟为一系列局部依存关系，这跟 GPSG 和 HPSG 中所做的是一样的。这是重要的，因为它允许我们捕捉提取路径标记的效果 (Bouma, Malouf & Sag 2001a: 1–2, Section 3.2)：比如说，有的语言对有成分提取的句子使用一种特殊形式的补足语。图 11.16 给出了词语法中(13)的分析。词上的连线是通常对

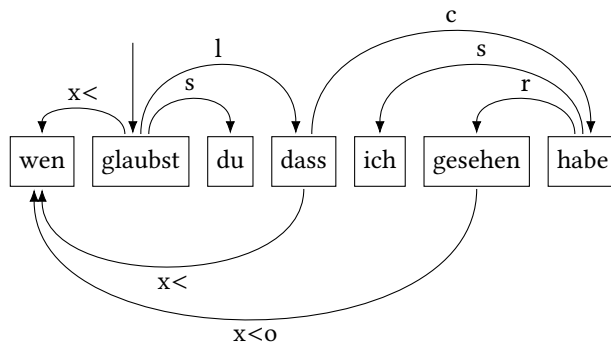


图 11.16: 词语法中带有双重依存关系的 *Wen glaubst du, dass ich gesehen habe?*（你认为我看见了谁？）的投影分析

主语、宾语和其他论元的依存连线（*r* 是 *sharer* 的缩写，它指动词性补足语，*l* 代表 *clausal complement*），在词下的连线是针对提取成分的连线（*x<*）。从 *gesehen*（看见）连到 *wen*（谁）是特殊的，因为它既是一个宾语连线也是一个提取连线（*x<0*）。这个连线是对图 11.15 中由虚线标记的小 *g* 和 *N* 的明确的说明。除了图 11.15 中的情况，图 11.16 也有一个从 *dass*（CONJ）到 *wen*（谁）的提取连线。我们可以用 Engel、Eroms 和 Gross & Osborne 的图形表示法来表示词语法的依存关系：我们可以简单地从 V_{prt} 节点和从属词节点上加虚线到统制 *wen*（谁）的 *N* 节点。

尽管这看上去比较简单，但是我想指出的是，词语法还应用了其他原则来满足合乎语法的结构。在下面，我来解释无缠绕原则（No-tangling Principle）、不孤单原则（No-dangling Principle）和句子一根原则（Sentence-root Principle）。

Principle 1 (无缠绕原则) 依存箭头不能缠绕。

Principle 2 (不孤单原则) 每个词必须有一个母节点。

Principle 3 (句子一根原则) 在每一个非复合中，只有一个词，它的母节点不是一个词，而是一个上下文的成分。

非缠绕原则保证了没有交叉的依存线，即它保证了结构是投影的 (Hudson 2000: 23)。因为非局部依存关系是通过具体的依存机制确立的，有人想排除非投影的分析。这一原则也排除了 (14b)，这里 *green* (绿色) 依存于 *peas* (豌豆)，但是并不是邻接于 *peas* (豌豆)。由于 *on* (PREP) 选择了 *peas* (豌豆)，从 *on* (PREP) 到 *peas* (豌豆) 的箭头就会与从 *peas* (豌豆) 到 *green* (绿色) 的箭头交叉。

- (14) a. He lives on green peas.
他 生存 PREP 绿色 豌豆
‘他靠绿色豌豆生存。’
b. *He lives green on peas.
他 生存 绿色 PREP 豌豆

不孤单原则确保了没有单独的词组连接到结构的主要部分上。没有这条规则，(14b) 就会被分析为孤立的 (*green* 绿色) (Hudson 2000: 23)。

句子一根原则需要用来排除那些不止有一个最高成分的结构。*glaubst* (认为) 是图 11.16 中的根。没有其他词统制并选择它了。这个原则确保了没有其他的根。所以这个原则排除了那些短语中的所有成分都是根的情况，因为如果不这样的话，不孤单原则就会失去效力，因为它就会被无能地实现 (Hudson 2000: 25)。

我这里加入了相当复杂的原则集合，是为了与基于短语结构方法的进行合理的比较。如果这里针对一般的短语提出一致性的话，就不需要这三个原则了。比如说，LFG 和 HPSG 就不需要这样的原则。

需要注意的是 Hudson (1997: 16) 认为 *Vorfeld* 中的成分被提取出来了，即使是像 (9) 这样简单的句子。我将在第 11.7.1 节说明为什么我认为这样的分析比那些分析更好，那些分析认为像 (9) 这样的简单句只是对应于动词位于首位或动词位于末位的语序变体。

11.6 新的发展与理论变体

本节主要介绍 Tesnière 的依存语法的变体。第 11.6.1 节介绍 Tesnière 的词类系统，第 11.6.2 节描述由 Tesnière 界定的语言对象的组合的模型。

11.6.1 Tesnière 的词类划分

正如在导言中提到的，Tesnière 是依存语法历史上的核心人物，因为他是第一位开发了形式化模型的 (Tesnière 1959, 1980, 2015)。现今有许多依存语法的版本，而且大部分都用其他语言中使用的词类标签 (N、P、A、V、Adv、Conj, ...)。Tesnière 的词类系统包括四个主要的范畴：名词、动词、形容词和副词。这些范畴的标签来自于用在世界语的词尾，即分别是 O、I、A 和 E。这些范畴按照语义来界定，如表 11.1 所示。¹⁷ Tesnière 认为这些范畴是普遍的，而且认为这些范畴互相依存的方式也是有限制的。

表 11.1: Tesnière 提出的语义驱动的词类类型

	物质	过程
具体	名词	动词
抽象	形容词	副词

根据 Tesnière，名词和副词可以依存于动词，形容词依存于名词，而副词依存于形容词或副词。这个情景在图 11.17 中的普通依存图中有所描述。“*”表示 Es 中的依存关系的任意数量。dependencies between Es. 当然，要找到依存于动词的形容词

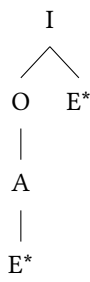


图 11.17: 根据 Tesnière 的依存关系的普遍配置
(I = verb, O = noun, A = adjective, E = adverb)

和依存于名词的句子（动词）是比较容易的。这类情况在 Tesnière 的系统中由所谓的转移 (transfers) 来处理。而且，这个范畴的集合中没有连词、限定词和介词。对于这些成分与他们的从属词的组合来说，Tesnière 使用了特殊的组合性关系：联结和转移。我们将在下面介绍这些概念。

¹⁷ 正如Weber (1997: 77)指出的，这个范畴不是没有问题的：根据什么说 *Angst*（害怕）是一个实体呢？为什么 *glauben*（认为）是一个具体的过程？也请看Klein (1971: Section 3.4)关于 *schlagen*（击打）和 *Schlag*（击打）以及类似例子的讨论。即使我们认为 *Schlag* 是通过转移到范畴 O 而从 *schlag-* 实际的过程生成的，这样的 Os 表示实体的假设也是有问题的。

11.6.2 联系、联结与转用

Tesnière (1959)提出了节点间的三个基本关系：联系、联结与转移。联系是我们在之前的章节中就介绍过的中心语与从属语之间的简单关系。联结是在并列分析中起到重要作用的特殊关系，而转用是允许我们变换词或短语的范畴的工具。

11.6.2.1 联结

图 11.18说明了联结关系：两个并列成分 *John* 和 *Mary* 通过连词 *and* 连到了一起。有趣的是，并列成分都连接到中心语 *laugh* 上。

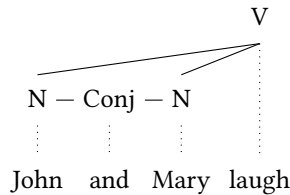


图 11.18: 对并列应用特殊的联结（junction）关系的分析

对于两个并列的名词的情况，我们得到图 11.19中的依存图。所有的名词都联

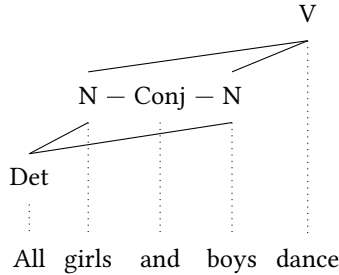


图 11.19: 对并列应用特殊的联结（junction）关系的分析

接到支配动词上，而且所有的名词都支配相同的限定词。

对并列的另一种分析方法是将连词看作是中心词，并列的成分看作是它的从属词。¹⁸ 这个方法的唯一问题是连词的范畴。它不能是 *Conj*，因为支配动词不选择 *Conj*，而是 *N*。这里可以使用的方法基本上与范畴语法中使用的是一样的（请看第 21.6.2 节）：范畴语法中连词的范畴是 $(X \setminus X) / X$ 。我们有一个函数，它带有同样范畴的两个论元，这个组合的结果是一个跟这两个论元具有相同范畴的对象。要把这个方法翻译到依存语法中，我们就会得到图 11.20 在下一页中描述的分析，而不是图 11.18 和图 11.19 中的分析。*all girls and boys* 这个句子的图看起来太奇怪了，因为限

¹⁸ 我这里不用 Tesnière 的范畴标签，这样读者就不用将 *I* 翻译成 *V*，以及 *O* 翻译成 *N*。

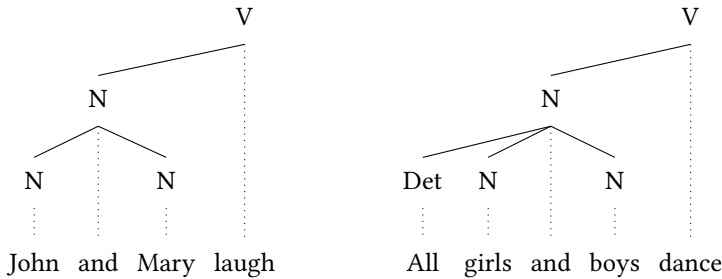


图 11.20: 无联结 (junction) 的将连词看作是中心语的并列的分析

定词和两个并列成分都依存于连词,但是由于这两个Ns选择了Det,并列的结果也是这样的。在范畴语法概念中,连词的范畴应该是 $((NP\backslash Det)(NP\backslash Det))/(NP\backslash Det)$,因为X实例为具有范畴 $(NP\backslash Det)$ 的名词,在这个分析中名词是中心语,限定词是从属语。

需要注意的是,这两种方法都需要给主语—动词的一致关系提供解释。Tesnière 最早的分析假定了动词和连词之间有两种依存关系。¹⁹因为连词是单数的,动词是复数的,在这个方法中根据依存关系得到的模型无法模拟一致关系。如果第二种分析找到了描述联结中并列的一致属性,这样一致的事实就是算数的,而且没有任何问题。

与图 11.20 中描述的中心语方法相比的另一种方法是没有中心语的。在基于短语结构的框架下工作的几个作者提出了没有中心语的并列结构的分析。这类分析也在依存语法中被提及 (Hudson 1988; Kahane 1997)。Hudson (1988) 和其他作出了相似假设的学者提出了一个针对并列的短语结构成分: 两个名词和连词被组合在一起构成一个更大的对象,它具有与任何一个组合词语不对应的属性。

相似地,基于联结的并列结构的分析为表达式的解读提出了问题。如果语义角色的指派在依存关系系统中发生,那么就会有如图 11.18 的图式的问题了。因为 *laugh* 的语义角色不能同时被 *John* 和 *Mary* 填充。相反,它由一个实体填充,即指向包括 *John* 和 *Mary* 的那个集合。这个语义表示属于短语 *John and Mary*,而且在这个并列结构中,这个最高实体的自然候选者是 *and*,因为它涵盖了 *John* 和 *Mary* 的意义: *and' (John', Mary')*。

这类联结也适用于动词的并列。但是,也不是没有问题的,因为说明语的辖域可以覆盖到离它最近的连词,或者是整个并列结构。例如下面 Levine (2003: 217) 中的句子:

- (15) Robin came in, found a chair, sat down, and whipped off her logging
Robin 进 找到一 椅子坐 下 和 拿 开 她的测井靴子

¹⁹ Eroms (2000: 467) 指出了这个一致性问题,并描述了事实。在他的分析中,他将第一个连词连接到统制中心语上,尽管看起来更合适的做法应该是假定一个内在构造的并列结构,然后连接到更高层的连词上。

boots in exactly thirty seconds flat.
在 确定地三十 秒 干脆
‘Robin 进来，找到了一把椅子，坐下来，并在三十秒内就把她的靴子脱下来了。’

说明语 *in exactly thirty seconds flat* 可以指向 (16a) 中的 *whipped off her logging boots*, 或者像 (16b) 中那样覆盖所有三个连词:

- (16) a. Robin came in, found a chair, sat down, and [[pulled off her logging boots] in exactly thirty seconds flat].
b. Robin [[came in, found a chair, sat down, and pulled off her logging boots] in exactly thirty seconds flat].

图 11.21 中的泰尼埃式的分析对应于 (17)，而如图 11.22 在下一页中将连词作为中心语的分析对应于 (16b)。

- (17) Robin came in in exactly thirty seconds flat and Robin found a chair in exactly thirty seconds flat and Robin pulled off her logging boots in exactly thirty seconds flat.

当说明语分别指向每个连词的，而不是如 (16b) 中由动词短语表示的指向一个累积的事件的时候，就会得到 (17) 中的解读。

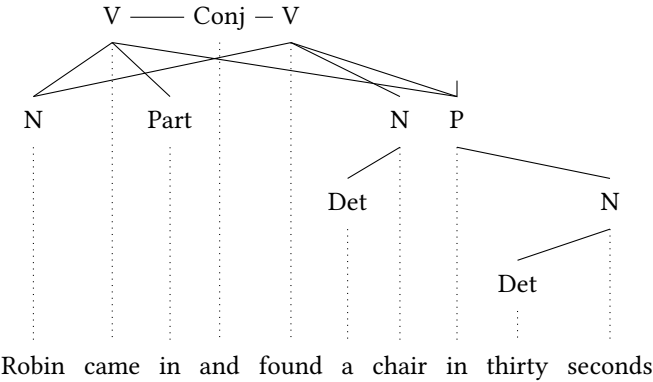


图 11.21: 带有联结关系的动词并列的分析

Levine (2003: 217) 讨论了连接到 Bouma, Malouf & Sag (2001a) 提出的提取的 HPSG 分析的这些句子。Bouma, Malouf & Sag 提出了，说明语作为中心语的从属语而从词汇的角度被引入。因为说明语按照词汇的方法来引入，并列结构基本上与泰尼埃式的分析具有相同的结构。有可能会想到正确得到语义组成成分的方式，即使句法并不对应于语义依存关系（请看 Chaves 2009 的观点），但是清楚的是，针对语义中的情况，从句法结构得到语义是更简单的。

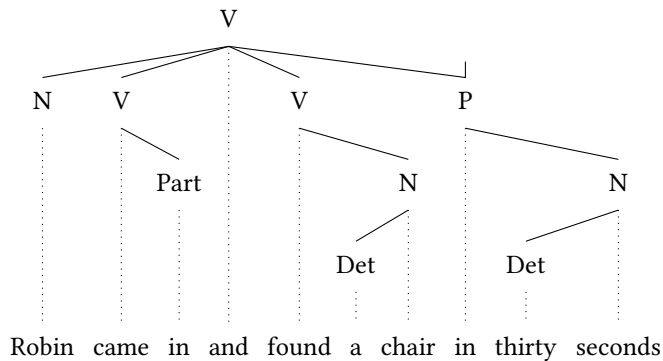


图 11.22: 带有联系关系的动词并列的分析

11.6.2.2 转用 (translation?)

在 Tesnière 的系统中, 转用被用于主要范畴 (如名词) 中的中心语与次要范畴 (如介词) 的词相组合的词或短语。另外, 转移可以在不需要任何其他词的参与下, 变换词或短语的范畴。

图 11.23 就是转用的一个例子。介词 *in* 引发了范畴的变化: 由于 *Traumboot* (梦

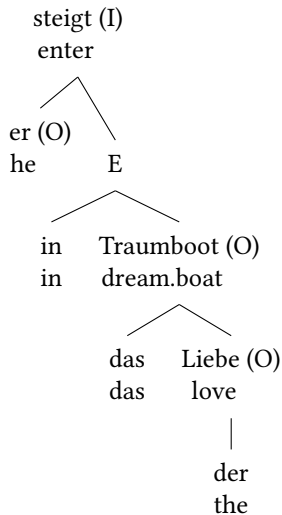


图 11.23: 从 Weber (1997: 83) 而来的转用的例子

想船) 是一个 O (名词), 介词与名词的组合是一个 E。这个例子说明了 Tesnière 使用了语法范畴来对语法功能进行编码。在 HPSG 这样的理论中, 有一个清晰的

区别：一方面，这里有关于词性的信息，另一方面，也有作为修饰语和谓语的成分的函项。修饰语函项被编码为修饰性特征 *MOD*，它是独立于词性的。由此，这就可以有修饰性和非修饰性的形容词、修饰性和非修饰性的介词短语，以及修饰性和非修饰性的名词短语等。对于手边的例子，我们可以假设一个带有方向性语义的介词选择了 NP。介词是带有填充了 *MOD value* 的 PP 的中心语。

使用转用的另一个方面是形态学。比如说，法语 *frappant*（惊人的）的派生是通过词根 *frapp* 加上后缀 *-ant* 构成的，如图 11.24 在下一页所示。这类转用可以被看作是在普通连接关系，如果词缀被看作是中心语。在实现形态学和构式形态也领域的形态学家反对这类基于语素的分析，因为他们在会话中包括了很多空成分，如动词 *play* 到名词 *play* 的转化（请看图 11.25 on page 358）。由此，词汇规则被用来表示 HPSG 这类理论中的派生与对话。HPSG 的词汇规则基本上等同于一元分支规则（请看第 258 页有关 39 的讨论和第 19.5 节）。词缀整合进词汇规则或者整合进区分分的形态形式的功能实现都是由词汇规则允准的。

总结一下，这里所说的转用是指

- 如果一个词或者短语与另一个词相组合，那么就应用二叉的短语结构规则，
- 如果一个短语没有通过任意一个额外的成分而转换到另一个范畴中，那么就应用一元的短语结构规则或者带有空中心语的二叉的短语结构规则，或者
- 如果一个词或词根匹配到一个词或者词根上，那么就应用一个（一元）的词汇规则。

关于 Tesnière 的转用规则和短语成分规则的关系的进一步讨论请看 Osborne & Kahane (2015: Section 4.9.1–4.9.2)。Osborne & Kahane 指出，转用规则可以用来模拟特异的结构，即那些没有一个单一部分可以指出中心语的结构。更多有关无中心语的构式内容请看第 11.7.2.4 节。

11.6.3 辖域

正如 Osborne & Kahane (2015: lix)所指出的，Tesnière 使用了所谓的测谎仪来表示辖域关系。所以说，由于例 (18) 中的 *that you saw yesterday* 是指 *red cars*，而不单是 *cars*，这通过 *red* 和 *cars* 的联系而不是单个成分开始的线来表示 (Tesnière 2015: 150, Stemma 149)。

- (18) *red cars that you saw yesterday*
 红色汽车 CONJ 你 看见昨天
 ‘你昨天看见的红色汽车’

Tesnière 的分析如图 11.26 on page 359 的左边表达式的描述。值得指出的是，这个表达式对英语图 11.26 右边的短语结构树。在 *red* 和 *cars* 之间的组合 B 对应于右边图中的 B 节点，而且 *red cars* 和 *that you saw yesterday* 的组合 A 对应于 A 节点。所以说，在短语结构语法中被清晰表示并指派了名称的对象在 Tesnière 的分析中是没有名字的，但是由于测谎仪的假说，可以指向这些组合。也请看图 ?? 的讨论，它显示出了 Hudson 为了模拟语义关系而提出的额外节点。

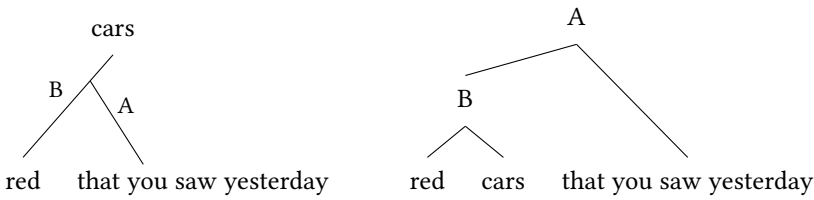


图 11.26: Tesnière 表示辖域的方式以及由Osborne & Kahane (2015: lix)提出的基于短语结构分析的比较

11.7 总结

依存语法的支持者们强调的一点是依存语法比短语结构语法简单多了，因为它们有更少的节点，而且一般的概念更容易被捕捉到（比如说Osborne 2014: Section 3.2, Section 7）。这是真的：依存语法适用于导论课程中语法的讲解。但是，正如Sternefeld & Richter (2012: 285)在一篇相当具有一般性的讨论中所指出的，简单的句法的代价是复杂语义及其他成分的缺失。所以说，除了依存句中描述的依存结构，人们还需要其他层面的信息。一个层面就是语义层，另一个是线性顺序。考虑到线性顺序，依存语法有两个选项：假设连续的成分，即投影结构或者允许非连续成分。这些选项将在后面的章节中讨论。第 11.7.2 节比较了依存语法和短语结构语法。它也指出了非投影结构可以在 HPSG 这类理论中模拟。语义的整合在第 11.7.2.3 节进行了讨论，而且逐渐清楚的是，一旦考虑到其他层面，依存语法就不必然要比短语结构语法简单了。

11.7.1 线性化

我们在这一章看到了许多线性化的方法。许多人只提出了依存图和根据拓扑模型的一些线性化方法。正如在第 11.5 节所讨论的，允许中心语及其从属语的非连续排列好像打开了潘多拉的盒子。我已经讨论了Kunze (1968)、Hudson (1997, 2000)、Kahane, Nasr & Rambow (1998)和Groß & Osborne (2009)提出的非局部依存的分析。除了 Hudson，那些作者们认为中心语的从属语只在那些条件下升到统制语的位置上，其中非连续性会在其他方面出现。但是，我们有理由认为前置应该按照特殊的机制来处理，即使在允许连续的序列化的情况下。比如说，例（19）中的例子的歧义性或歧义性的缺失不能按照直接的方式来解释：

- (19) a. Oft liest er das Buch nicht.
经常读 他DET 书 不
‘经常的情况是他不读书。’或者
‘他经常读书，事实不是这样的。’
b. dass er das Buch nicht oft liest
CONJ 他DET 书 不 经常读
‘他经常读书，事实不是这样的。’

- c. dass er das Buch oft nicht liest
CONJI 他DET 书 经常不 读
‘经常的情况是，他不读书。’

这三个例子的情况是，只有 (19a) 是有歧义的。即使 (19c) 具有相同的语序，只要考虑到 *oft* (经常) 和 *nicht* (不)，这个句子就不是歧义的。所以，说明语的前置是歧义的原因。(19a) 的依存图如图 11.27 所示。当然，(19b) 和 (19c) 的依存关

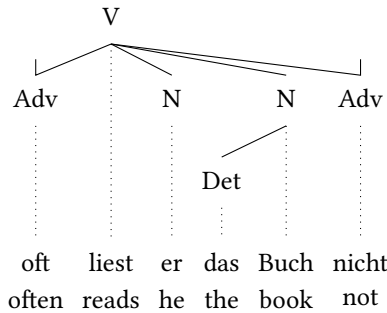


图 11.27: *Oft liest er das Buch nicht.* (他不经常读这本书) 的依存图

系不是不同的。图是相同的，只在线性序列中是不同的。所以说，辖域的不同不能从依存关系中推导出来，而且像 (20) 的复杂声明是必要的：

- (20) 如果从属语在 *Vorfeld* 中线性化了，它可以跨域，而且位于它所从属的中心语的所有其他说明语下面。

Eroms (1985: 320) 提出了否定的分析，其中否定被当作中心语来处理；也就是说，例 (21) 中的句子具有图 11.28 在下一页中的结构。²⁰

- (21) Er kommt nicht.
他来 不
‘他没来。’

这个分析对应于最简方案中的 NegP，而且它有相同的问题：整个宾语的范畴是 Adv，但是它应该是 V。这是一个问题，因为更高层的谓词可以选择 V，而不是 Adv。²¹

对于成分否定或者带有成分的其他辖域来说也是一样的。比如说，例 (22) 的分析可以如图 11.29 在下一页所示。

- (22) der angebliche Mörder
DET 所谓的 杀人犯

²⁰ 但是请看 Eroms (2000: Section 11.2.3)。

²¹ 请看下面 (23) 中嵌套句的分析。

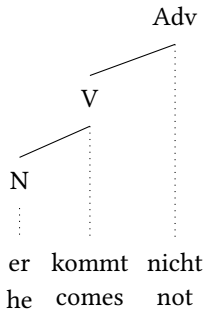


图 11.28: 根据Eroms (1985: 320)的否定的分析

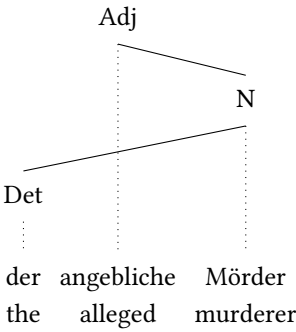


图 11.29: 覆盖全域的说明语作为中心语所得到结果的分析

这个结构可以具有非投影性的其他问题。Eroms 确实在这里所说的问题中对限定词进行了不同的处理，所以这个非投影性的类型对他来说不是一个问题。但是，否定的中心语分析会得到德语的所谓一致结构的非投影性。例（23）中的句子有两种解读：在第一种解读中，否定的辖域包括 *singen*（唱歌），而在第二种解读中包括 *singen darf*（唱歌允许）。

- (23) dass er nicht singen darf
 CONJ 他不 唱歌 允许
 ‘他不被允许唱歌’ 或 ‘他被允许不唱歌’

nicht（不）的辖域包括动词复杂形式的解读会得到图 11.30 on the following page 中给出的非投影结构。Eroms 还提出了一个分析，其中否定是一个词的部分（‘Wortteiläquivalent’）。但是，这里没有任何帮助，因为首先否定和动词在（19a）中的动词二位的语境中不是相邻的，而且即使在（23）中的动词末位的语境中也是不相邻的。Eroms 不得不假设，依附到否定的宾语是整个动词的复杂形式 *singen darf*（唱歌允许），即包括两个词的复杂宾语。

这就给我们留下了图 11.27 中的分析的一个问题，因为我们针对不同解释的两个可能的说明语实现形式只有一个结构。这不是通过将两个可能的线性顺序简单

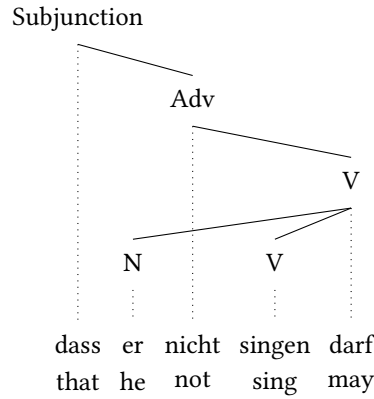


图 11.30: 否定作为中心语的假说所得到结果的分析

地作为另一种语序的分析所能得到的。

Thomas Groß (p. c. 2013) 提出了一个分析, 其中 *oft* 并不依存于动词, 而是否定。这对应于短语结构方法中的成分否定。依存图如图 11.31 的左手边所示。右手

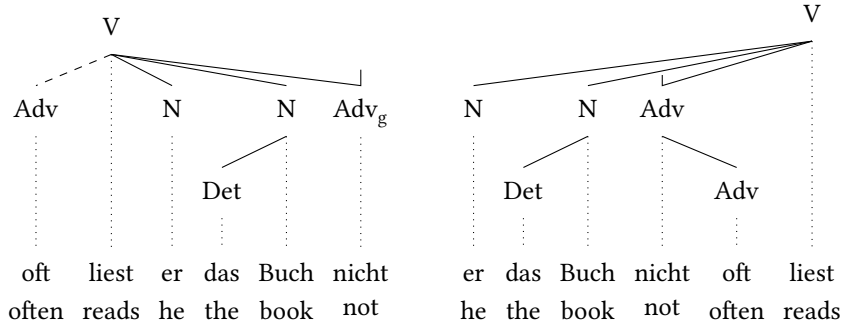


图 11.31: 根据 Groß 和动词末位变体的 *Oft liest er das Buch nicht*. (他不经常读书。) 的依存图

边的图显示了相应的动词末位句子的图。这个解释对应于能够用对比解释的成分否定。但是在 (24a) 中, 只有 *oft* (经常) 是被否定的, *oft gelesen* (经常读) 是在 (24b) 的否定辖域中。

- (24) a. Er hat das Buch nicht oft gelesen, sondern selten.
他 AUX DET 书 不 经常读 但是 很少
‘他不经常读书, 但是偶尔读。’

- b. Er hat das Buch nicht oft gelesen, sondern selten gekauft.
 他 AUX DET 书 不 经常读 但是 很少 买
 ‘他不经常读书，但是偶尔会买书。’

这两个解释对应于图 11.32 中的两个短语结构树。需要注意的是，在 HPSG 的分析

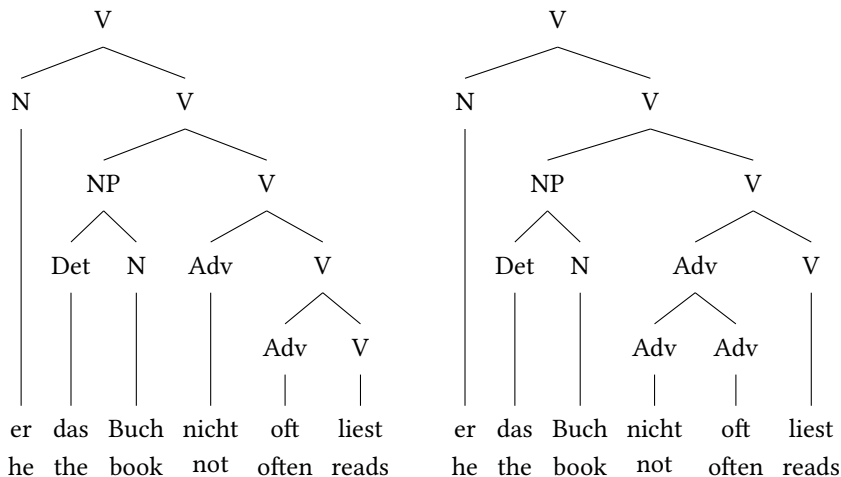


图 11.32: *er das Buch nicht oft liest* (他不经常读书) 的可能的句法分析

中，副词 *oft* 可以是短语 *nicht oft* (不经常) 的中心语。这区别于 Groß 提出的依存语法分析。进而，依存语法分析有两个结构：一个带依存于相同动词的所有副词的平铺结构，和一个依存于否定的 *oft*。基于短语结构的分析有三个结构：一个结构的语序是 *oft* 在 *nicht* 之前，一个语序是 *nicht* 在 *oft* 之前，还有一个语序带有 *nicht* 和 *oft* 的直接组合。关于 (19a) 中的例子的问题在于头两个结构的一个没有依存语法表示。这就可能使得它无法推导出语义，但是在基于成分的分析中是更困难的。

进而，需要注意的是，直接将依存图联系到拓扑场的模型不能说明例 (25) 的句子。

- (25) Dem Saft eine kräftige Farbe geben Blutorangen.²²
 DET 果汁一 强烈 颜色 给 血. 橙
 ‘血橙给果汁一个强烈的颜色。’

这个句子的依存图如图 11.33 on the next page 所示。

这类明显的多重前置并不局限于 NPs。从属词的不同类型可以在 *Vorfeld* 中被替换。数据的进一步的讨论可以参考 Müller (2003a)。还有在“多重前置和信息结构”这个研究项目中收集了很多其他的数据 (Bildhauer 2011)。任何单独基于依存关系的理论以及不允许空成分的理论被迫放弃了动词二位 (动词位于第二位) 的语言分析中较为常见的限制。相较而言，像 GB 的分析和那些假定空的动词中心语的

²² Bildhauer & Cook (2010) 在曼海姆的德语系举办的 *Deutsches Referenzkorpus* (DeReKo) 中发现了这个例子: <http://www.ids-mannheim.de/kl/projekte/korpora>

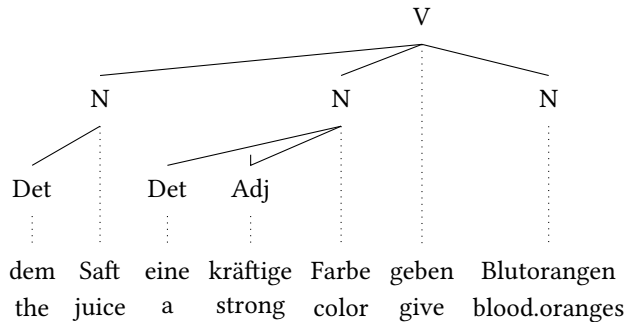


图 11.33: *Dem Saft eine kräftige Farbe geben Blutorangen.* (血橙给果汁一个强烈的颜色。)的依存图

HPSG 变体可以假设, 这一动词中心语的投影占据了 *Vorfeld* 的位置。这就解释了为什么 *Vorfeld* 中的材料看上去像包括一个可见动词的动词性中心语: 这个 *Vorfelds* 在内部结构上是拓扑显示的。它们可以有一个填充的 *Nachfeld*, 还有一个填充到句子右边界的助词。更多的数据、讨论和分析请看 Müller (2005c, 2015b)。在 Gross & Osborne 的框架 (2009) 中的同等的分析可以是图 11.34 中所示的图, 但是需要注意的是 Groß & Osborne (2009: 73) 明确地反对空成分, 而且在任何情况下就为了处理多重前置而提出一个空成分是十分特异的。²³ 需要指出的是, 这个问题并没有通

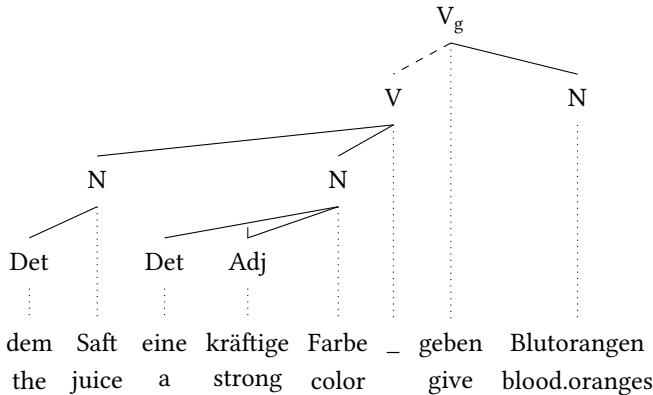


图 11.34: 在 *Vorfeld* 中带有空的动词中心语的 *Dem Saft eine kräftige Farbe geben Blutorangen.* (血橙给果汁一个强烈的颜色。)的依存图

过简单地去除 V2 限制和允许定式动词的从属词在它的左边实现而得到解决, 因为前置的成分并不必然依存于例 (26) 中所示的例子中的定式动词:

²³ 我在允许非连续成分的 HPSG 的基于线性化的变体中提出了这样一个空成分 (Müller 2002b), 但是后来进行了修改, 这样只有连续成分才是可以的, 动词位置被处理为中心语移位, 以及动词移位分析中使用的包括相同空的动词中心语的多重前置 (Müller 2005c, 2015b)。

- (26) a. [Gezielt] [Mitglieder] [im Seniorenbereich] wollen die Kendoka
 特别地 成员 PREP.DET 高级. 市民. 选区 想要.INF DET 剑道家
 allerdings nicht werben.²⁴
 但是 不 招募
 ‘但是, 剑道家的会员招募策略并不想指向高级市民的选区。’
- b. [Kurz] [die Bestzeit] hatte der Berliner Andreas Klöden [...] gehalten.²⁵
 简短地 DET 最好. 时光AUX DET 柏林人 Andreas Klöden 拥有
 ‘来自柏林的 Andreas Klöden 拥有过简短的美好时光。’

And although the respective structures are marked, such multiple frontings can even cross clause boundaries:

- (27) Der Maria einen Ring glaube ich nicht, daß er je schenken wird.²⁶
 DET.DAT Maria 一.ACC 戒指 相信 我 不 CONJ 他 曾经给 AUX
 ‘我不认为他曾经给了 Maria 一个戒指。’

如果允许了这种依存关系, 要限制他们就十分困难了。这里不便讨论细节问题, 但是读者可以参考Müller (2005c, 2015b)。

还需要指出的是, Engel 有关德语句子的线性序列的说明 (2014: 50), 即指向定式动词前的一个成分 (请看脚注 7) 是十分不准确的。我们只能猜测词 *element* 想要表达什么意思。一个解释是它是基于成分语法的经典论断的连续成分。另一种方法是, 一个中心语和一些从属语 (没有必要是所有它的从属语) 的连续实现。这一方法可以允许图 11.35 on the next page 中描述的例 (28) 的非连续成分的外置分析。

- (28) Ein junger Kerl stand da, mit langen blonden Haaren, die sein Gesicht
 一 年轻 人 站 那儿PREP 长 金色 头发 CONJ 他的脸
 einrahmten, [...] ²⁷
 装框
 ‘一位脸旁镶满金发的年轻小伙子站在那儿’

这一分析的形式化不是一个小问题, 因为我们明确知道什么可以非连续地实现, 而且依存关系的哪些部分一定可以连续地实现。Kathol & Pollard (1995) 在 HPSG 的框架下展开了外置的分析。也请看Müller (1999a: Section 13.3)。我在下一节讨论了 HPSG 中这种线性化分析的基本机制。

11.7.2 依存语法与短语结构语法

本章探讨依存语法和短语结构语法之间的关系。我首先指出, 投影的依存语法可以译成短语结构语法 (第 11.7.2.1 节)。然后, 我将讨论带有非投影的依存语

²⁴ taz, 1999 年 7 月 7 日, p. 18. 摘自Müller (2002b)。

²⁵ Märkische Oderzeitung, 28./29.07.2001, p. 28.

²⁶ Fanselow (1993: 67).

²⁷ Charles Bukowski, *Der Mann mit der Ledertasche* (《带皮包的男人》)。慕尼黑: 德国平装书出版社, 1994, p. 201, 由 Hans Hermann 翻译。

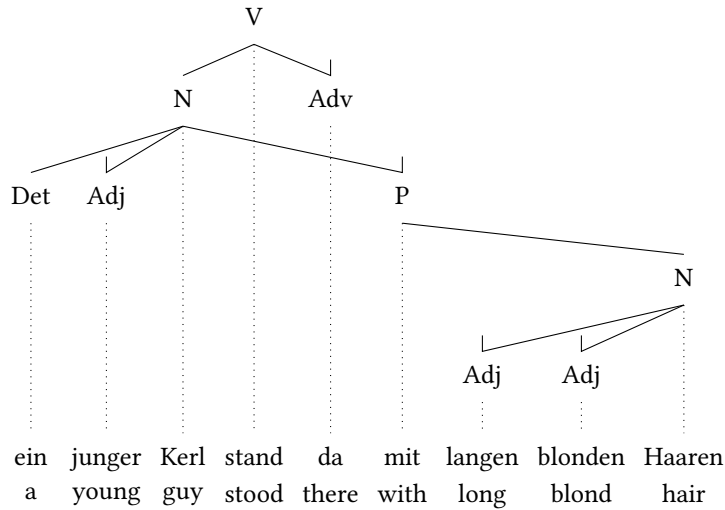


图 11.35: 在 *Vorfeld* 中带有非连续成分的 *Ein junger Kerl stand da, mit langen blonden Haaren.* (一位满头金发的小伙子站在那里。) 的依存图

法，并且说明他们是如何在基于线性化的 HPSG 理论中表示的（第 11.7.2.2 节）。第 11.7.2.3 节讨论了基于短语结构的理论中提出的额外节点，而且第 11.7.2.4 节讨论了无中心语的结构，这对于所有的依存语法都是一个问题。

11.7.2.1 将投影性依存语法变为短语结构语法

正如 Gaifman (1965)、Covington (1990: 234)、Oliva (2003) 和 Hellwig (2006: 1093) 所指出的，某些投影性中心语的短语结构语法可以通过将中心语移动到上一层来取代统制的节点以变成依存语法。所以在一个 NP 结构中，N 变换到了 NP 的位置上，而且所有其他联系保持不变。如图 11.36 所示。



图 11.36: 在短语结构语法和依存语法分析中的 *a book*

当然，这个程序不能直接应用到所有的短语结构语法中，因为有些包括了更为精细的结构。比如说，规则 $S \rightarrow NP, VP$ 不能翻译成依存规则，因为 NP 和 VP 都属于复杂的范畴。

接下来,我想说明图 11.1 on page 338 中的依存图是如何重新表示为允准了一个相似树的中心语的短语结构规则,即如图 11.37 所示。我没有使用 NP 和 VP 的标签

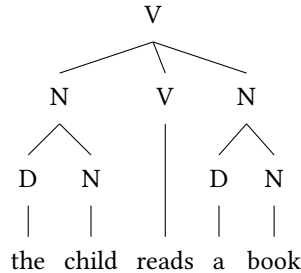


图 11.37: 带有平铺规则的短语结构的 *The child reads a book.* 的分析

来保证两张图在最大程度上是相似的。NP 和 VP 中的 P 部分是指投影的满足,并且经常在图中被忽视。请看第 9 章有关 HPSG 的内容,比如说,允准了例 (29) 中给出的树的语法再次忽视了配价信息。

- (29) $N \rightarrow D N$ $N \rightarrow \text{child}$ $D \rightarrow \text{the}$ $D \rightarrow a$
 $V \rightarrow N V N$ $N \rightarrow \text{book}$ $V \rightarrow \text{reads}$

如果我们在带有各自词汇项的例 (29) 中的两个最左边规则右手边替换 N 和 V, 我们就会去除允准了词语的规则, 我们就会得到 (30) 中给出的语法的词汇化变体:

- (30) $N \rightarrow D \text{ book}$ $D \rightarrow \text{the}$
 $N \rightarrow D \text{ child}$ $D \rightarrow a$
 $V \rightarrow N \text{ reads } N$

词汇化 (Lexicalized) 是指每一部分由语法允准的树都包括一个词汇成分。(30) 中的语法允准了图 11.1 中的树。²⁸

经典的短语结构语法和依存语法的一个重要的区别是短语结构规则给予节点强加了特定的语序。也就是说, (30) 中的 V 规则暗示了第一个名词性投影、动词和第二个名词性投影必须按照规则表明的语序来出现。当然, 这个语序限制可以松一些, 正如 GPSG 中所作的。这基本上允准了规则右边任意顺序的子节点。剩下的问题是说明语的整合。由于说明语也依存于中心语 (请看图 11.4 on page 340), 可以提出一条规则来允准论元外的任意多的说明语。由此, (30) 中的 V 规则应该变成 (31) 中的形式:²⁹

²⁸ 正如在第 339 页所提到的, Gaifman (1965: 305)、Hays (1964: 513)、Baumgärtner (1970: 57) 和 Heringer (1996: 37) 针对依存规则提出了一条普遍的规则, 即它有一个特殊的标记 (分别为 “*” 和 “~”) 替代了 (30) 中的词汇词。Heringer 的规则具有 (31) 中的形式:

(i) $X[Y_1, Y_2, \sim, Y_3]$

X 是中心语的范畴, Y_1 、 Y_2 和 Y_3 是中心语的从属语, 而且 “~” 是中心语所插入的位置。

²⁹ 请看第 170 页中 GPSG 的相似规则, 并且参考德语的 HPSG 理论的分析 Kasper (1994), 它提出了完全的平铺结构并且整合进了任意数量的说明语。

(31) $V \rightarrow N \text{ reads } N \text{ Adv}^*$

这种概化的短语结构等同于投影的依存语法。³⁰但是，正如我们看到的，一些研究人员对非连续的成分允许了交叉边的存在。接着，我来说明依存语法是如何在 HPSG 理论中被形式化的。

11.7.2.2 非投影依存语法与带有非连续成分的短语结构语法

等同于非投影依存图的是短语结构语法中的非连续成分。接下来，我将给出允许非连续结构的基于短语结构的理论的一个例子。因为，正如我将展示的，非连续性也可以进行模拟，短语结构语法和依存语法之间的区别归结于词的单位是否被给予了标签（比如说 NP）。

在 HPSG 理论这类框架下来模拟非连续成分的技术最早追溯到 Mike Reape 在德语中所作的工作 (1991; 1992; 1994)。Reape 使用了一个叫做 **DOMAIN** 的列表来表示按照话语表层出现顺序的符号的子节点。(32) 给出了一个例子，其中，中心语短语的 **DOM value** 从中心语的 **DOM value** 和非中心语子节点的列表中计算出来。

$$(32) \quad \textit{headed-phrase} \Rightarrow \left[\begin{array}{ll} \text{HEAD-DTR|DOM} & [1] \\ \text{NON-HEAD-DTRS} & [2] \\ \text{DOM} & [1 \circ 2] \end{array} \right]$$

符号“ \circ ”表示 *shuffle* 关系。*shuffle* 连接了 A、B 和 C 三个列表，当且仅当 C 包括了 A 和 B 中的所有成分，而且 A 中成分的语序和 B 中成分的语序保存在 C 中。(33) 分别显示了带有两个成分的两个集合的组合：

$$(33) \quad \langle a, b \rangle \circ \langle c, d \rangle = \langle a, b, c, d \rangle \vee \\ \langle a, c, b, d \rangle \vee \\ \langle a, c, d, b \rangle \vee \\ \langle c, a, b, d \rangle \vee \\ \langle c, a, d, b \rangle \vee \\ \langle c, d, a, b \rangle$$

结果是六个列表的析取。在所有这些列表中，*a* 位于 *b* 的前面，*c* 位于 *d* 的前面，因为在已经结合的 $\langle a, b \rangle$ 和 $\langle c, d \rangle$ 两个列表中也是这样的。但是，除了这种情况，*b* 可以放在 *c* 和 *d* 的前面，中间和后面。每个词都带有一个域值，它是一个包括这个词本身的列表：

³⁰ Sylvain Kahane (p.c. 2015) 指出，二元性对于依存语法来说是非常重要的，因为主语只有一条规则，宾语只有一条规则，以及其他（比如说 Kahane 2009，这是在 HPSG 理论的形式框架下的依存语法的实现）。不过，我没有想到有任何原因需要反对平铺结构。比如说，Ginzburg & Sag (2000: 364) 在 HPSG 理论中提出了主语助动词转换的平铺结构。在这类平铺规则中，限定语/主语和其他补足语在同一个目标下与动词组合在一起。这也适用于两个以上配价特征的语法功能范畴，如主语、直接宾语、间接宾语。也请看关于平铺规则脚注 28。

(34) 单个词的管辖范围, 这里是 *gibt* (给):

$$\boxed{1} \left[\begin{array}{l} \text{PHON} \langle \textit{gibt} \rangle \\ \text{SYNSEM} \dots \\ \text{DOM} \langle \boxed{1} \rangle \end{array} \right]$$

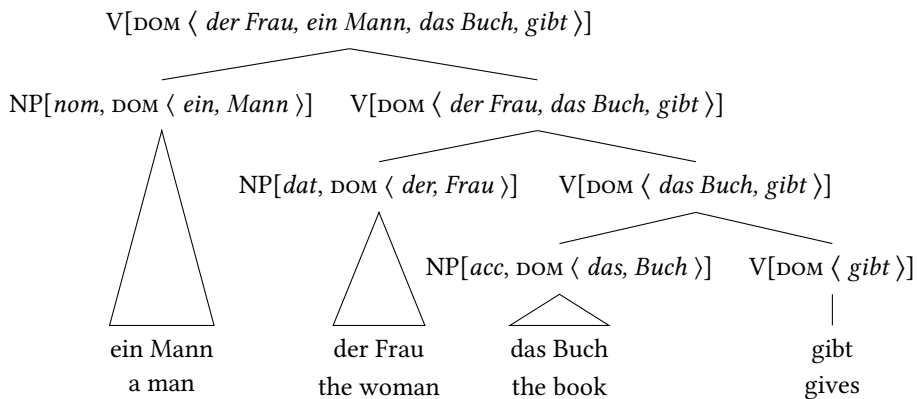
(34) 里的描述初看起来有些奇怪, 因为它是循环的, 不过它可以是被理解的, 它说明了 *gibt* 将它自己赋予到线性化域中出现的元素。

(35) 中的限制决定于短语的 PHON values :

$$(35) \quad \textit{phrase} \Rightarrow \left[\begin{array}{l} \text{PHON} \boxed{1} \oplus \dots \oplus \boxed{n} \\ \text{DOM} \left\langle \left[\begin{array}{l} \textit{sign} \\ \text{PHON} \boxed{1} \end{array} \right], \dots, \left[\begin{array}{l} \textit{sign} \\ \text{PHON} \boxed{n} \end{array} \right] \right\rangle \end{array} \right]$$

它证明了符号的 PHON value 是它的 DOMAIN 成分的 PHON values 的合一。由于 DOMAIN 成分的语序对应于他们的表层语序, 这是决定整个语言对象的 PHON value 的明显方式。

图 11.38 显示了这一机制是如何应用到带有非连续成分的二叉结构中的。由逗



二元的分叉结构被用于Kathol (1995, 2000)和Müller (1995, 1996c, 1999a, 2002a)对德语的分析, 但是正如我们在这个章节看到的, 依存语法提出了平铺的表示 (但是请看第 368 页的脚注 30)。模式 1 允准了在一个目标内实现的中心语的所有论元的结构。³¹

Schema 1 (中心语-论元模式 (平铺结构))

head-argument-phrase \Rightarrow

$$\left[\begin{array}{l} \text{SYNSEM|LOC|CAT|SUBCAT } \langle \rangle \\ \text{HEAD-DTR|SYNSEM|LOC|CAT|SUBCAT } [1] \\ \text{NON-HEAD-DTRS } [1] \end{array} \right]$$

为了简化表达, 我认为 SUBCAT list 包括了完整符号的描述。所以说, 整个列表可以等同于非中心语子节点的列表。³² DOM value 的计算可以按照下面的方式来进行限制:

$$(36) \quad \textit{headed-phrase} \Rightarrow \left[\begin{array}{l} \text{HEAD-DTR} \quad [1] \\ \text{NON-HEAD-DTRS} \quad \langle [2], \dots, [n] \rangle \\ \text{DOM} \quad \langle [1] \rangle \circ \langle [2] \rangle \circ \dots \circ \langle [n] \rangle \end{array} \right]$$

这一限制是说 DOM 的值是一个列表, 它是变换每个包括一个子节点作为成分的单一系列的结果。这种变换操作的结构上所有子节点的所有可能的排列组合。这看起来超过了 GPSG 已经取得的成果, 它通过将短语结构规则右手边的成分的序列抽象出来而得到。但是, 需要注意的是这个机制可以用来表示更加自由的语序: 通过指向子节点的 DOM values, 而不是子节点本身, 有可能将个别词插入 DOM list 中。

$$(37) \quad \textit{headed-phrase} \Rightarrow \left[\begin{array}{l} \text{HEAD-DTR|DOM} \quad [1] \\ \text{NON-HEAD-DTRS} \quad \langle [\text{DOM } [2]] \dots [\text{DOM } [n]] \rangle \\ \text{DOM} \quad \langle [1] \rangle \circ \langle [2] \rangle \circ \dots \circ \langle [n] \rangle \end{array} \right]$$

应用这一限制, 我们得到基本上按照任意排列组合的话语的所有词汇的 DOM values。我们得到的是一个不带有任何投影限制的纯粹的依存语法。根据这个语法, 我们可以分析图 11.6 on page 341 的非投影结构, 以及更多的内容。域内的术语的分析如图 ?? on page ?? 所示。关于成分/依存关系对于分析自然语言是首要的, 还是必需的问题是存在争议的: Hudson (1980) 和 Engel (1996) 认为依存关系是充分的, 这被大部分依存语法学家所认可 (根据 Engel 1996), Leiss (2003) 认为不是这样的。为了解决这个问题, 我们来看一些例子:

- (38) Dass Peter kommt, klärt nicht, ob Klaus spielt.
CONJ Peter 来 解决不 是否 Klaus 玩
'Peter 来并不能解决 Klaus 是不是想玩儿这个问题。'

³¹ 我在这里提出包括在词汇中心语, 而不是联系其中的 SUBCAT list 的所有论元成分。比如说, Borsley (1989: 339) 提出了英语的助动词转换和威尔士语中动词位于首位的模式, 它们指向主语和补足语的配价特征, 并在平铺结构实现所有的成分。

³² 除了这一假设, 我们需要一个将类型 *synsem* 的描写的列表匹配到类型 *sign* 描写的列表的关系性限制。更多细节请看 Meurers (1999c: 198)。

如果我们知道话语的意义，我们可以赋予它一个依存图。让我们假设 (38) 的意义是像 (39) 这样的：

(39) $\neg \text{resolve}'(\text{that}'(\text{come}'(\text{Peter}')), \text{whether}'(\text{play}'(\text{Klaus}'))))$

根据这个语义信息，我们当然可以给 (38) 构成一个依存图。原因是依存关系在 (39) 的语义表达式中反应为双唯一的方式。但是需要注意的是这对于普通的情况

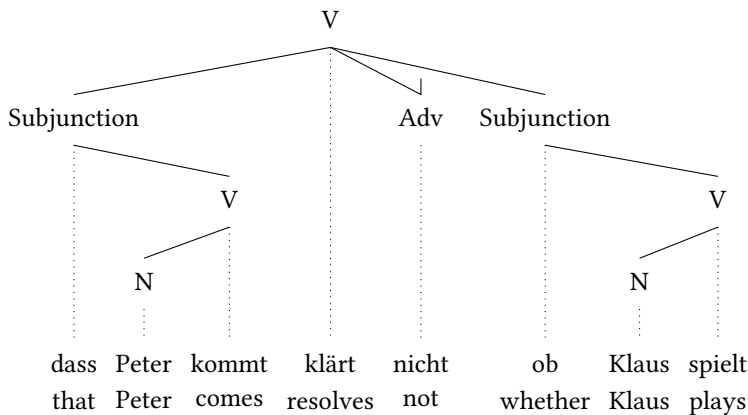


图 11.39: 能从语义表达式中推导出的 *Dass Peter kommt, klärt nicht, ob Klaus spielt.* (Peter 来并不能解决 Klaus 是不是想玩儿这个问题。) 的依存图

是不适用的。比如说 (40) 中的例子：

(40) Dass Peter kommt, klärt nicht, ob Klaus kommt.
 CONJ Peter 来 解决 不 是否 Klaus 玩儿
 ‘Peter 来并不解决 Klaus 是否来的问题。’

这里，词 *kommt* 出现了两次。没有关于邻接、线性顺序和一致性的成分或限制的任何说明，我们不能赋予一个没有歧义的依存图。比如说，图 ?? on page ?? 中的图完美地兼容了句子的语义：*dass* 统制 *kommt*，而且 *kommt* 统制 *Peter*，而 *ob* 统制 *kommt*，而且 *kommt* 统制 *Klaus*。

(41) a. Deshalb klärt, dass Peter kommt, ob Klaus spielt.
 所以 解决 CONJ Peter 来 是否 Klaus 玩
 b. *Deshalb klärt dass ob Peter Klaus kommt spielt.
 所以 解决 CONJ 是否 Peter Klaus 来 玩

(41b) 是 (41a) 的一个变体，其中两个小句的论元的成分互相按照正确的语序排列，但是所有的小句都是非连续的，每个小句的成分都按照这个方式发生了变化。依存图如图 ?? on page ?? 所示。

11.7.2.3 在中心语和投影之间没有保持一致的特征

正如Oliva (2003)指出的,依存语法和 HPSG 理论的相似之处仅限于 HEAD values。也就是,对应于 HPSG 理论中的 HEAD values 的依存图的节点标签。但是,还有像 CONT 这种表示语义的和 SLASH 这种表示非局部依存关系的额外特征。这些值通常在词汇中心语和它的短语投影之间有不同之处。为了说明这一情况,让我们看一下短语 *a book*。词汇和完整短语的语义如 (42) 所示:³³

- (42) a. $\lambda P \lambda Q \exists x (P(x) \wedge Q(x))$
 b. *book*: $\lambda y (book'(y))$
 c. *a book*: $\lambda Q \exists x (book'(x) \wedge Q(x))$

现在,依存语法标记的问题是没有能够联系到 *a book* 的语义的 NP 节点(请看第 366 页的图 11.36),树中出现的唯一事物是词汇 N 的节点: *book* 的节点。³⁴不过,这不是一个大问题:词汇属性可以表示为作为不同特征值的最高节点的部分。那么,依存图中的 N 节点会有 CONT value,它对应于完整短语的语义贡献以及对应于短语的词汇中心语的贡献的 LEX-CONT 值。所以对于 *a book* 来说,我们会得到下面的表达式:

- (43)
$$\left[\begin{array}{l} \text{CONT} \quad \lambda Q \exists x (book'(x) \wedge Q(x)) \\ \text{LEXICAL-CONT} \quad \lambda y (book'(y)) \end{array} \right]$$

根据这类表达式,我们可以保持这些分析,中心语与其从属语的语义贡献是其部分的语义共享的功能。现在,还有更多特征,其中词汇中心语与他们的投影不同。一个特征是 SLASH,它用来表示 HPSG 理论中的非局部依存关系,并能在 à la Groß & Osborne (2009)方法中用来建立提升的元素和中心语之间的关系。当然,我们可以再次应用相同的策略。这样,我们会得到特征 LEXICAL-SLASH。但是,这可以得到批准,而且词汇项的特征可以组织在一条路径下。这样,一般的架构如 (44) 所示:

- (44)
$$\left[\begin{array}{l} \text{CONT} \\ \text{SLASH} \\ \text{LEXICAL} \left[\begin{array}{l} \text{CONT} \\ \text{SLASH} \end{array} \right] \end{array} \right]$$

但是如果我们将 LEXICAL 重新命名为 HEAD-DTR,我们基本上会得到 HPSG 理论的表达式。Hellwig (2003: 602)证明了他的依存语法的特殊版本叫做依存合一语法,他认为统制中心语选择了带有所有他们的子节点的完整节点。这些节点的属性与他们的中心语不同 (p. 604)。他们实际上是成分。所以这个依存语法的非常明晰和形式化的变体与 HPSG 理论十分相似,正如 Hellwig 他自己证明的 (p. 603)。Hudson 的词语法也是明晰表示的,正如下面所示的,它与 HPSG 理论也是十分相似的。

³³ lambda 表达式的内容请看第 2.3 节。

³⁴ Hudson (2003: 391–392) 对此清楚地表述:“在依存分析中,从属词修饰中心语词的意义,所以后者带有整个短语的意义。例如,在 *long books about linguistics* 中,由于从属词的修饰效应,词 *books* 表示‘关于语言学的长书’。”对于这个观点的具体实现请看图 ?? on page ??。另一种观点是在 Meaning-Text Theory 中假定不同的表示层次 (Mel'čuk 1981)。实际上, HPSG 理论中的 CONT value 也是一个不同的表示层。但是,这个表示层跟建立起来的其他结构是同步的。

图 ?? on page ??中的表示是图 ?? on page ??表示的简缩版的详细描写。在第一个图中显示的是两个节点的组合得到一个新节点。例如, *playing* 和 *outside* 的组合得到 *playing'*, *small* 和 *children* 的组合得到 *children'*, 而且 *children'* 和 *playing'* 的组合得到 *playing''*。 *were* 和 *playing''* 的组合得到 *were'*, 以及 *children''* 和 *were'* 的组合得到 *were''*。唯一剩下需要解释的是为什么有一个节点 *children*, 它不是两个节点的组合(即 *children''*)的结果。顶点与底部的连线表示默认的承继关系。即, 上层节点继承了默认的底层节点的所有属性。默认的可以被覆盖, 即上层节点的信息会与统制节点的信息不同。这就使得按照语义的组合性的操作成为可能: 两个节点组合的结果的那个节点的语义是两个组合的节点的意义组合。再来看 *children*, *children'* 具有必须邻接到 *playing* 的属性, 但是因为这个结构是一个升位结构, 其中 *children* 提升到了 *were* 的主语位置上, 这个属性被 *children* 的一个新的实例覆盖了, 即 *children''*。这里有趣的是我们几乎得到了一个名词性短语结构树, 如果我们通过句法范畴替换图 ??中的图的词的话。这个替换的结果如图 ?? on page ??所示。这张图(由虚线标记)中唯一不同的地方是 *N'* 与 *V[ing]'* 相组合, 而且 *N'* 的母节点, 即 *N''*, 与 *V[fin]'* 相组合。正如上面解释的, 这依赖于词语法中升位的分析, 它包括提升的项目和它的中心语之间的多重依存关系。在图 ??中有两个 *N* 节点(*N'* 和 *N''*), 而且在图 ??中有 *children* 的两个实例。除了这个, 这个结构对应于 HPSG 语法可以允准的成分。Hudson 范式中的在底部用线连接到顶端的节点与使用默认承继的子节点是相关的。这对于使用默认承继关系的 HPSG 的许多版本来说也是非常相似的。比如说, Ginzburg & Sag (2000: 33)使用了广义中心语特征原则, 它默认地将中心语子节点的所有特征投影到母节点上。本节的结论是, 短语结构语法和依存语法的唯一理论区别在于中间结构是如何假设的: 是否存在没有主语的 VP 呢? 说明语附加成分是否有中间节点呢? 在没有涵盖了语义表示的全部可行的方案下回答这些问题是十分困难的。那些可行的方法, 如 Hudson 和 Hellwig 的方法, 提出了中间表达式, 它使得这些方法跟基于短语结构的方法十分相似。如果我们将全部可行的依存语法的变体的结构与短语结构语法相比较的话, 很清楚的是, 依存语法更为简单这个观点是不受欢迎的。这个观点适用于图 ??中的紧缩的范式的表示, 但是它并不适用于全部可行的分析。

11.7.2.4 非中心语构式

Hudson (1980: Section 4.E)讨论了(45)中所示的无中心语的结构:

- (45) a. the rich
b. the biggest
c. the longer the stem
d. (with) his hat over his eyes

他认为, 术语形容词(adjective)和名词(noun)应该属于术语实体(substantive), 它包括这两个术语。然后他指出“如果一条规则需要覆盖传统上指称为带或者不带中心语的名词短语, 它只称为‘名词’, 而且这可以自动允许要么是实体要么是形容词作为中心语的结构。”(p.195)但是, 这里需要提出的是, 像 *the rich* 这样的实体性短语的内部依存结构是什么样的呢。连接这些项目的唯一方式看起来是假定限定词依存于形容词。但是这可以允许像 *the rich man* 的短语的两个结构:

一个是限定词依存于形容词，另一个是它依存于名词。所以，词性的未分析不能解决这个问题。当然，所有无中心语的结构的所有问题都可以通过假定空成分来解决。³⁵在 HPSG 的关系小句的分析中是这样做的 (Pollard & Sag 1994: Chapter 5)。英语和德语的关系小句包括一个短语，它包括一个关系词和一个句子，其中关系短语是缺失的。Pollard & Sag 提出了空关系代词，它选择关系小句以及带有空位的小句 (Pollard & Sag 1994: 216–217)。相似的分析可以在依存语法中找到 (Eroms 2000: 291)。³⁶现在，另一种分析空成分的方法是短语构式。³⁷Sag (1997)研究了英语的关系小句，他提出了关系小句的短语分析，其中关系小句和它从新的短语中提取出的小句。Müller (1996c)提出了一个相似的分析，并且记录在Müller (1999a: Chapter 10)中。正如第 8.6 节所讨论的，认为关系短语中的关系代词或者其他成分作为整个关系小句的中心语是不可行的，而且认为动词作为整个小句的中心语也是不可行的 (Sag)，因为关系小句修饰 Ns，这是 (定式) 动词通常不会这样投影的。所以说，假定一个空的中心语或者短语模式看起来是唯一的选择了。第 21 章致力于讨论某些语法现象是否应该分析为包括短语结构配置或者词汇分析更为合适或者更适合于模拟某些现象。我认为所有跟定价互动的现象都应该按照词汇的来处理。但是，还有其他现象，而且依存语法被迫用来表示所有语言现象的词汇化的分析。总有一些成分是所有其他成分所依存的。Jackendoff (2008)认为，像 (46) 中的 N-P-N 构式的一个成分作为中心语的观点是没有意义的。

- (46) a. day by day, paragraph by paragraph, country by country
 b. dollar for dollar, student for student, point for point
 c. face to face, bumper to bumper
 d. term paper after term paper, picture after picture
 e. book upon book, argument upon argument

当然，有方法可以模拟所有可以在短语结构的框架下 (如 GPSG、CxG、HPSG 或者简便句法：一个空中心语) 模拟的现象。图 ?? on page ?? 显示了 *student after student* 的分析。空 N 的词汇项是非常特殊的，因为没有相似的非空的词汇名词，即没有名词选了两个光杆 Ns 和一个 P。Bragmann (2015)指出了 N-P-N 构式的一个额外的方面，这使得事情更加复杂了。这个模式没有受限于两个名词。可以是任意数量的：

- (47) Day after day after day went by, but I never found the courage to talk to her.

³⁵ 请看第 2.4.1 节关于名词短语的短语结构语法中的空中心语的假设。

³⁶ 依存语法表示通常有一个 *d-* 成分作为关系小句的中心语。但是，由于关系代词也出现在小句中，而且由于 *d-* 没有被两次发音，假定一个额外的 *d-* 中心语基本上就是假定一个空的中心语。另一个选择是假定具有多重功能的词汇：所以说，一个关系代词可以既是一个中心语，也同时是一个从属语 (Tesnière 2015: Chapter 246, §8–11, Osborne & Kahane 2015: xlvii; Kahane 2009: 129–130)。至少剑道家的分析是第 8.6 节讨论的范畴语法分析的一个例子，而且它具有相同的问题：如果关系代词是中心语，它选择了缺失关系代词的小句，不容易看到这个分析是如何扩展到 (i) 中的抽吸过程的，其中被提取的成分是一个包括关系代词，而不是代词本身的完整短语。

(i) die Frau, von deren Schwester ich ein Bild gesehen habe
 DET 女人 PREP 谁的 姐妹 我 一 图片看见 AUX
 ‘我在她的姐妹那里看到一张照片的那个女人’

³⁷ 请看第 19 章关于空成分的一般分析和第 21.10.3 小节关于关系小句的特殊分析。

所以说, Bragmann 提出了 (48) 中的模式, 而不是 N-P-N 范式, 这里 ‘+’ 表示一个序列的至少一次重复。

(48) N (P N)+

现在, 这种范式在基于选择的方法中是十分困难的, 因为我们需要假设一个空的中心语或者选择相同介词、名词或名词性短语的任意数量对儿的名词。当然, 我们可以假设 P 和 N 构成了某种成分, 但是仍然有人需要确保使用了正确的介词, 而且名词或者名词性投影具有正确的语音形式。另一种可能是要假设 N-P-N 中的第二个 N 可以是 N-P-N, 由此允许模式的循环。但是如果我们按照这个方法, 核查这种限制就非常困难了, 其中所包含的 Ns 应该具有相同或者至少类似的语音形式。解决这些问题的一种方式当然可以假定这里具有特殊的机制来指派一个新的范畴到一个或几个成分上。这在本质上可以是一个非中心语的短语结构规则, 而且这是 Tesnière 所提出的: 转用规则 (请看第 11.6.2.2 节)。但是, 这当然是存粹的依存语法向混合模型的一个扩展。请看第 21.10 节有关深入问题的讨论, 它对纯粹的基于选择的语法来说可能是有问题的。

练习题

请给出下面的三个句子的依存图:

- (49) a. Ich habe einen Mann getroffen, der blonde Haare hat.
我 AUX 一 男人 见面 谁 金色 头发 有
‘我跟一位有着金色头发的男士见面了。’
- b. Einen Mann getroffen, der blonde Haare hat, habe ich noch nie.
一 男人 见面 谁 金色 头发 有 AUX 我 yet 从未
‘我从未见到过有着金色头发的男人。’
- c. Dass er morgen kommen wird, freut uns.
CONJ 他明天 来 AUX 愉悦 我们
‘他明天会来让我们很高兴。’

你可以使用非投影的依存关系。对于关系小句的分析, 学者们通常提出一个抽象的实体, 它的功能是可修饰名词的从属语, 以及关系小句中的动词的中心语。

延伸阅读

在第 3 章的延伸阅读部分, 我推荐了 *Syntaktische Analyseperspektiven* (《句法分析的不同视角》) 这本书。该书的各个章节是由不同理论的支持者所著, 并且分析了相同的新闻语料。这本书还有 Engel (2014) 写的一章, 他提出了他的依存语法的版本, 即从属动词语法 (Dependent Verb Grammar)。Ágel, Eichinger, Eroms, Hellwig, Heringer & Lobin (2003a, 2006) 出版了依存关系和配价手册, 它讨论了依存语法涉及的所有方面。本章引用了其中的很多文献。对比依存语法和其他理论的文章在本书的语境下是尤为重要的: Lobin (2003) 比较了依存语法和范畴语法, Oliva (2003) 分析

11 依存语法

配价的表示和 HPSG 中的依存关系，而且Bangalore, Joshi & Rambow (2003)描述了配价和依存是如何覆盖在 TAG 中的。Hellwig (2006)比较了基于规则的语法与依存语法，特别关注于计算程序的剖析。Osborne & Groß (2012)比较了依存语法和构式语法，Osborne, Putnam & Groß (2011)认为某些最简方案的变体实际上是基于依存分析的重新发现。Tesnière (1959)提出的依存语法的原始工作在德语 (Tesnière 1980)中是部分可获得的，在英语中是全部可获得的 (Tesnière 2015)。

第十二章 Tree Adjoining Grammar

Tree Adjoining Grammar (TAG) was developed by Aravind Joshi at the University of Pennsylvania in the USA (Joshi, Levy & Takahashi 1975). Several important dissertations in TAG have been supervised by Aravind Joshi and Anthony Kroch at the University of Pennsylvania (e.g., Rambow 1994). Other research centers with a focus on TAG are Paris 7 (Anne Abeillé), Columbia University in the USA (Owen Rambow) and Düsseldorf, Germany (Laura Kallmeyer). Rambow (1994) and Gerdes (2002b) are more detailed studies of German.¹

TAG and its variants with relevant extensions are of interest because it is assumed that this grammatical formalism can – with regard to its expressive power – relatively accurately represent what humans do when they produce or comprehend natural language. The expressive power of Generalized Phrase Structure Grammar was deliberately constrained so that it corresponds to context-free phrase structure grammars (Type-2 languages) and it has in fact been demonstrated that this is not enough (Shieber 1985; Culy 1985).² Grammatical theories such as HPSG and CxG can generate/describe so-called Type-0 languages and are thereby far above the level of complexity presently assumed for natural languages. The assumption is that this complexity lies somewhere between context-free and context-sensitive (Type-1) languages. This class is thus referred to as *mildly context sensitive*. Certain TAG-variants are inside of this language class and it is assumed that they can produce exactly those structures that occur in natural languages. For more on complexity, see Section 12.6.3 and Chapter 17.

There are various systems for the processing of TAG grammars (Doran, Hockey, Sarkar, Srinivas & Xia 2000; Parmentier, Kallmeyer, Maier, Lichte & Dellert 2008; Kallmeyer, Lichte, Maier, Parmentier, Dellert & Evang 2008). Smaller and larger TAG fragments have been developed for the following languages:

- Arabic (Fraj, Zribi & Ahmed 2008),
- German (Rambow 1994; Gerdes 2002a; Kallmeyer & Yoon 2004; Lichte 2007),

¹ Since my knowledge of French leaves something to be desired, I just refer to the literature in French here without being able to comment on the content.

² See Pullum (1986) for a historical overview of the complexity debate and G. Müller (2011a) for argumentation for the non-context-free nature of German, which follows parallel to Culy with regard to the N-P-N construction (see Section 21.10.4).

- English (XTAG Research Group 2001; Frank 2002; Kroch & Joshi 1987),
- French (Abeillé 1988; Candito 1996, 1998, 1999; Crabbé 2005),
- Italian (Candito 1998, 1999),
- Korean (Han, Yoon, Kim & Palmer 2000; Kallmeyer & Yoon 2004),
- Vietnamese (Le, Nguyen & Roussanaly 2008)

Candito (1996) has developed a system for the representation of meta grammars which allows the uniform specification of crosslinguistic generalizations. This system was used by some of the projects mentioned above for the derivation of grammars for specific languages. For instance Kinyon, Rambow, Scheffler, Yoon & Joshi (2006) derive the verb second languages from a common meta grammar. Among those grammars for verb second languages is a grammar of Yiddish for which there was no TAG grammar until 2006.

Resnik (1992) combines TAG with a statistics component.

12.1 General remarks on representational format

12.1.1 Representation of valence information

Figure 12.1 shows so-called elementary trees. These are present in the lexicon and can be combined to create larger trees. Nodes for the insertion of arguments are specially

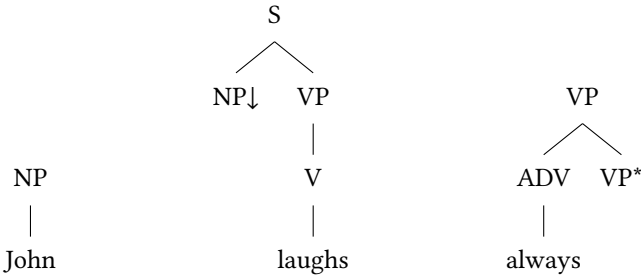


图 12.1: Elementary trees

marked (NP↓ in the tree for *laughs*). Nodes for the insertion of adjuncts into a tree are also marked (VP* in the tree for *always*). Grammars where elementary trees always contain at least one word are referred to as *Lexicalized Tree Adjoining Grammar* (LTAG).

12.1.2 Substitution

Figure 12.2 在下一页 shows the substitution of nodes. Other subtrees have to be inserted into substitution nodes such as the NP node in the tree for *laughs*. The tree for *John* is inserted there in the example derivation.

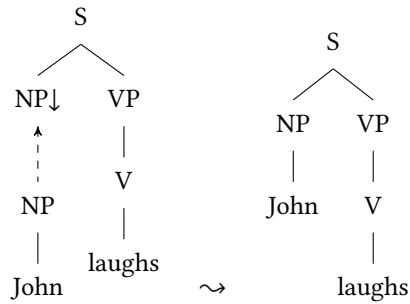


图 12.2: Substitution

12.1.3 Adjunction

Figure 12.3 shows an example of how the adjunction tree for *always* can be used.

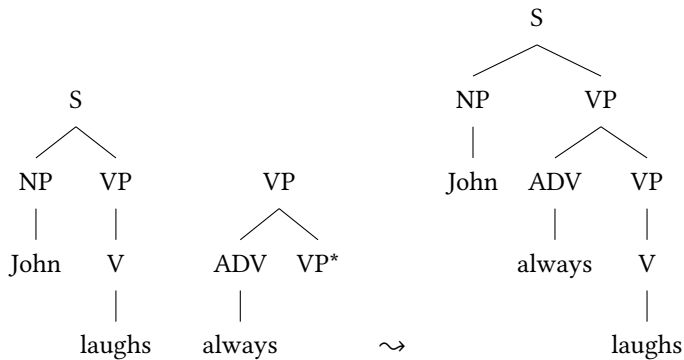


图 12.3: Adjunction

Adjunction trees can be inserted into other trees. Upon insertion, the target node (bearing the same category as the node marked with “*”) is replaced by the adjunction tree.

TAG differs considerably from the simple phrase structure grammars we encountered in Chapter 2 in that the trees extend over a larger domain: for example, there is an NP node in the tree for *laughs* that is not a sister of the verb. In a phrase structure grammar (and of course in GB and GPSG since these theories are more or less directly built on phrase structure grammars), it is only ever possible to describe subtrees with a depth of one level. For the tree for *laughs*, the relevant rules would be those in (1):

- (1) $S \rightarrow NP VP$
 $VP \rightarrow V$
 $V \rightarrow laughs$

In this context, it is common to speak of *locality domains*. The extension of the locality domain is of particular importance for the analysis of idioms (see Section 18.2).

TAG differs from other grammatical theories in that it is possible for structures to be broken up again. In this way, it is possible to use adjunction to insert any amount of material into a given tree and thereby cause originally adjacent constituents to end up being arbitrarily far away from each other in the final tree. As we will see in Section 12.5, this property is important for the analysis of long-distance dependencies without movement.

12.1.4 Semantics

There are different approaches to the syntax-semantics interface in TAG. One possibility is to assign a semantic representation to every node in the tree. The alternative is to assign each elementary tree exactly one semantic representation. The semantics construction does not make reference to syntactic structure but rather the way the structure is combined. This kind of approach has been proposed by Candito & Kahane (1998) and then by Kallmeyer & Joshi (2003), who build on it. The basic mechanisms will be briefly presented in what follows.

In the literature on TAG, a distinction is made between derived trees and derivation trees. Derived trees correspond to constituent structure (the trees for *John laughs* and *John always laughs* in Figures 12.2 and 12.3). The derivation tree contains the derivational history, that is, information about how the elementary trees were combined. The elements in a derivation tree represent predicate-argument dependencies, which is why it is possible to derive a semantic derivation tree from them. This will be shown on the basis of the sentence in (2):

(2) Max likes Anouk.

The elementary tree for (2) and the derived tree are given in Figure 12.4 在下一页. The nodes in trees are numbered from top to bottom and from left to right. The result of this numbering of nodes for *likes* is shown in Figure 12.5 在下一页. The topmost node in the tree for *likes* is S and has the position 0. Beneath S, there is an NP and a VP node. These nodes are again numbered starting at 0. NP has the position 0 and VP the position 1. The VP node has in turn two daughters: V and the object NP. V receives number 0 and the object NP 1. This makes it possible to combine these numbers and then it is possible to unambiguously access individual elements in the tree. The position for the subject NP is 00 since this is a daughter of S and occurs in first position. The object NP has the numeric sequence 011 since it is below S (0), in the VP (the second daughter of S = 1) and occurs in second position (the second daughter of VP = 1).

With these tree positions, the derivation tree for (2) can be represented as in Figure 12.6 在下一页. The derivation tree expresses the fact that the elementary tree for *likes* was combined with two arguments that were inserted into the substitution positions 00 and 011. The derivation tree also contains information about what exactly was placed into these nodes.

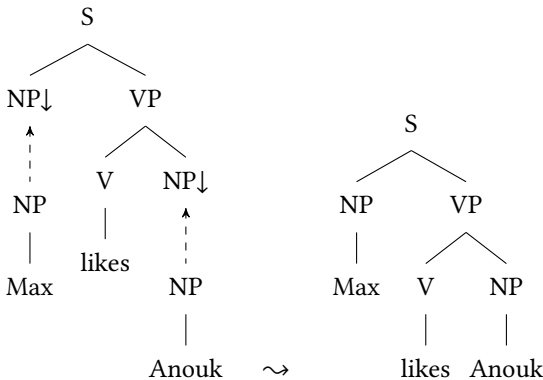


图 12.4: Elementary trees and derived tree for *Max likes Anouk*.

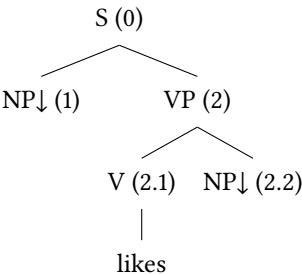


图 12.5: Node positions in the elementary tree for *likes*

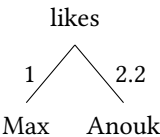


图 12.6: Derivation tree for *Max likes Anouk*.

Kallmeyer & Joshi (2003) use a variant of *Minimal Recursion Semantics* as their semantic representational formalism (Copestake, Flickinger, Pollard & Sag 2005). I will use a considerably simplified representation here, as I did in Section 9.1.6 on semantics in HPSG. For the elementary trees *Max*, *likes* and *Anouk*, we can assume the semantic representations in (3).

- (3) Semantic representations for elementary trees:

max(x)	like(x ₁ , x ₂)	anouk(y)
arg: –	arg: ⟨ x ₁ , 1 ⟩, ⟨ x ₂ , 2.2 ⟩	arg: –

In a substitution operation, a variable is assigned a value. If, for example, the elementary tree for *Max* is inserted into the subject position of the tree for *likes*, then x_1 is identified with x . In the same way, x_2 is identified with y if the tree for *Anouk* is inserted into the object position. The result of these combinations is the representation in (4):

- (4) Combination of the meaning of elementary trees:

like(x, y)
max(x)
anouk(y)
arg: –

Kallmeyer & Joshi (2003) show how an extension of TAG, Multi-Component LTAG, can handle quantifier scope and discuss complex cases with embedded verbs. Interested readers are referred to the original article.

12.2 Local reordering

In TAG, there is a family of trees for each word. In order to account for ordering variants, one can assume that there are six trees corresponding to a ditransitive verb and that each of these corresponds to a different ordering of the arguments. Trees are connected to one another via lexical rules. This lexical rule-based analysis is parallel to the one developed by Uszkoreit (1986b) in Categorical Grammar.

Alternatively, one could assume a format for TAG structures similar to what we referred to as the ID/LP format in the chapter on GPSG. Joshi (1987b) defines an elementary structure as a pair that consists of a dominance structure and linearization constraints. Unlike GPSG, linearization rules do not hold for all dominance rules but rather for a particular dominance structure. This is parallel to what we saw in Section 10.6.3 on Embodied-CxG. Figure 12.7 在下一页 shows a dominance tree with numbered nodes. If we combine this dominance structure with the linearization rules in (5), we arrive at the exact order that we would get with ordinary phrase structure rules, namely $NP_1 V NP_2$.

- (5) $LP_1^\alpha = \{ 1 < 2, 2.1 < 2.2 \}$

If one specifies the linearization restrictions as in (6), all the orders in (7) are permitted, since the empty set means that we do not state any restrictions at all.

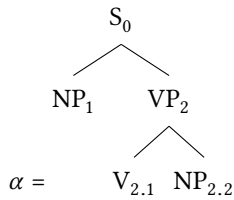


图 12.7: Dominance structure with numbered nodes

- (6) $LP_2^\alpha = \{ \}$
- (7) a. $NP_1 V NP_2$
 b. $NP_2 V NP_1$
 c. $NP_1 NP_2 V$
 d. $NP_2 NP_1 V$
 e. $V NP_1 NP_2$
 f. $V NP_2 NP_1$

This means that it is possible to derive all orders that were derived in GPSG with flat sentence rules despite the fact that there is a constituent in the tree that consists of NP and VP. Since the dominance rules include a larger locality domain, such grammars are called LD/LP grammars (local dominance/linear precedence) rather than ID/LP grammars (immediate dominance/linear precedence) (Joshi, Shanker & Weir 1990).

Simple variants of TAG such as those presented in Section 12.1 cannot deal with reordering if the arguments of different verbs are scrambled as in (8).

- (8) weil ihm das Buch jemand zu lesen versprochen hat³
 because him.DAT the.ACC book somebody.NOM to read promised has
 ‘because somebody promised him to read the book’

In (8), *das Buch* ‘the book’ is the object of *zu lesen* ‘to read’, and *ihm* ‘him’ and *jemand* ‘somebody’ are dependent on *versprochen* and *hat*, respectively. These cases can be analyzed by LD/LP-TAG developed by Joshi (1987b) and Free Order TAG (FO-TAG) (Becker, Joshi & Rambow 1991: 21) since both of these TAG variants allow for crossing edges.

Since certain restrictions cannot be expressed in FO-TAG (Rambow 1994: 48–50), so-called Multi-Component TAG was developed. Joshi, Becker & Rambow (2000) illustrate the problem that simple LTAG grammars have with sentences such as (8) using examples such as (9):⁴

³ For more on this kind of examples, see Bech (1955).

⁴ The authors use *versprochen hat* ‘has promised’ rather than *versprach* ‘promised’, which sounds better but does not correspond to the trees they use.

- (9) a. ... daß der Detektiv dem Klienten [den Verdächtigen des
 that the.NOM detective the.DAT client the.ACC suspect the.GEN
 Verbrechens zu überführen] versprach
 crime to indict promised
 ‘that the detective promised the client to indict the suspect of the crime’
- b. ... daß des Verbrechens_k der Detektiv den Verdächtigen_j
 that the.GEN crime the.NOM detective the.ACC suspect
 dem Klienten [_j –_k zu überführen] versprach
 the.DAT client to indict promised

In LTAG, the elementary trees for the relevant verbs look as shown in Figure 12.8. The

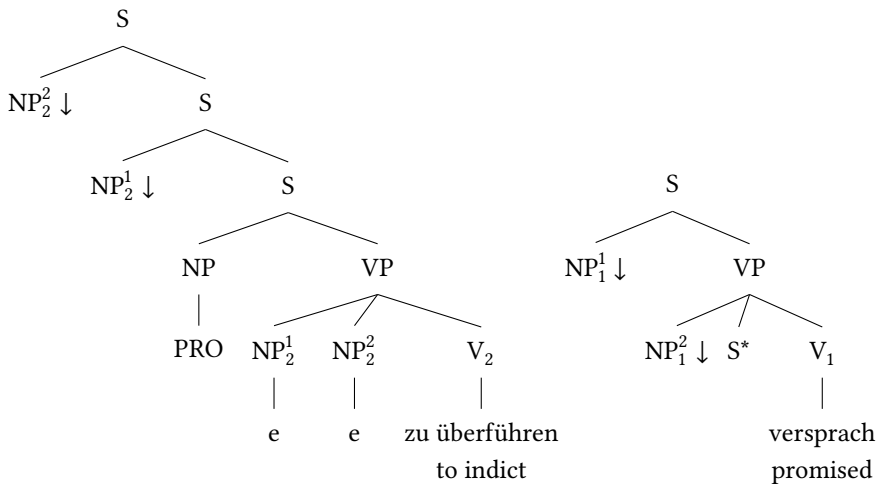


图 12.8: Elementary trees of an infinitive and a control verb

verbs are numbered according to their level of embedding. The NP arguments of a verb bear the same index as that verb and each has a superscript number that distinguishes it from the other arguments. The trees are very similar to those in GB. In particular, it is assumed that the subject occurs outside the VP. For non-finite verbs, it is assumed that the subject is realized by PRO. PRO is, like *e*, a phonologically empty pronominal category that also comes from GB. The left tree in Figure 12.8 contains traces in the normal position of the arguments and the relevant NP slots in higher trees positions. An interesting difference to other theories is that these traces only exist in the tree. They are not represented as individual entries in the lexicon as the lexicon only contains words and the corresponding trees.

The tree for *versprach* ‘promised’ can be inserted into any S node in the tree for *zu überführen* ‘to indict’ and results in trees such as those in the Figures 12.9 and 12.10. In Figure 12.9, the tree for *versprach* is inserted directly above the PRO NP and in Figure 12.10 above NP_2^1 .

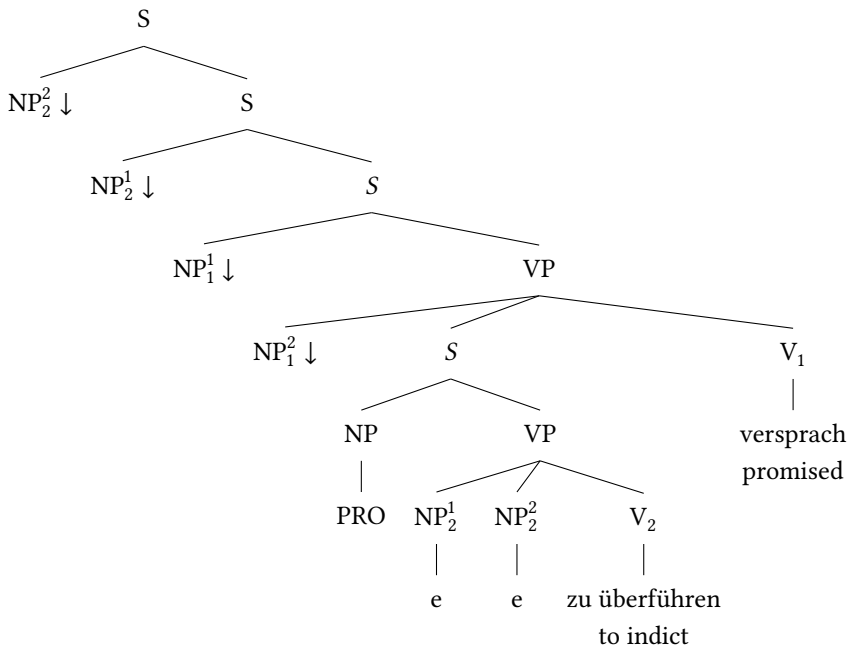


图 12.9: Analysis of the order $NP_2^2 NP_2^1 NP_1^1 NP_1^2 V_2 V_1$: adjunction to the lowest S node

It is clear that it is not possible to derive a tree in this way where an argument of *überführen* ‘to indict’ occurs between the arguments of *versprach* ‘promised’. Joshi, Becker & Rambow (2000) therefore suggest an extension of the LTAG formalism. In MC-TAG, the grammar does not consist of elementary trees but rather finite sets of elementary trees. In every derivational step, a set is selected and the elements of that set are simultaneously added to the tree. Figure 12.11 on the next page shows an elementary tree for *versprach* ‘promised’ consisting of multiple components. This tree contains a trace of NP_1^1 that was moved to the left. The bottom-left S node and the top-right S node are connected by a dashed line that indicates the dominance relation. However, immediate dominance is not required. Therefore, it is possible to insert the two subtrees into another tree separately from each other and thereby analyze the order in Figure 12.12 on page 387, for example.

Other variants of TAG that allow for other constituent orders are V-TAG (Rambow 1994) and TT-MC-TAG (Lichte 2007).

12.3 Verb position

The position of the verb can be analyzed in a parallel way to the GPSG analysis: the verb can be realized in initial or in final position in a given linearization domain.

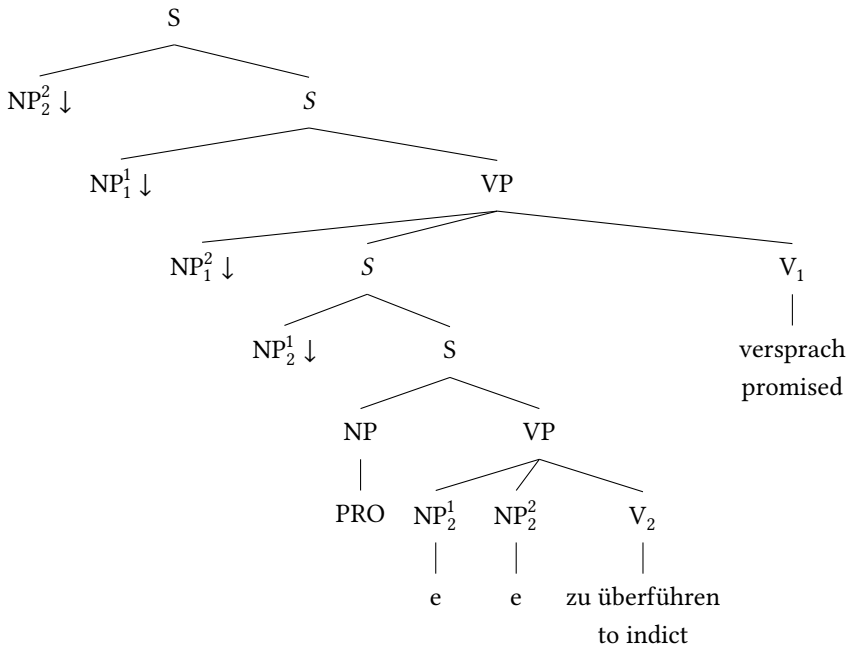


图 12.10: Analysis of the order $NP_2^2 NP_1^1 NP_1^2 NP_2^1 V_2 V_1$: adjunction to the S node between NP_2^2 and NP_2^1

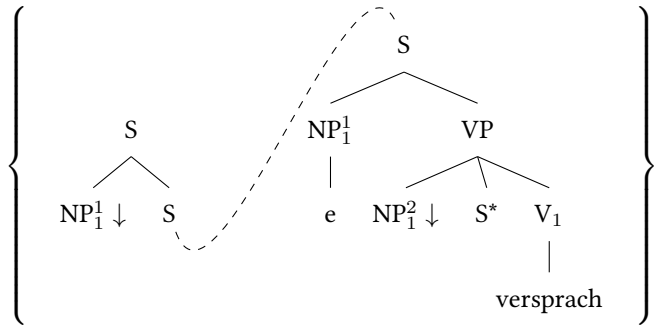


图 12.11: Elementary tree set for *versprach* consisting of multiple components

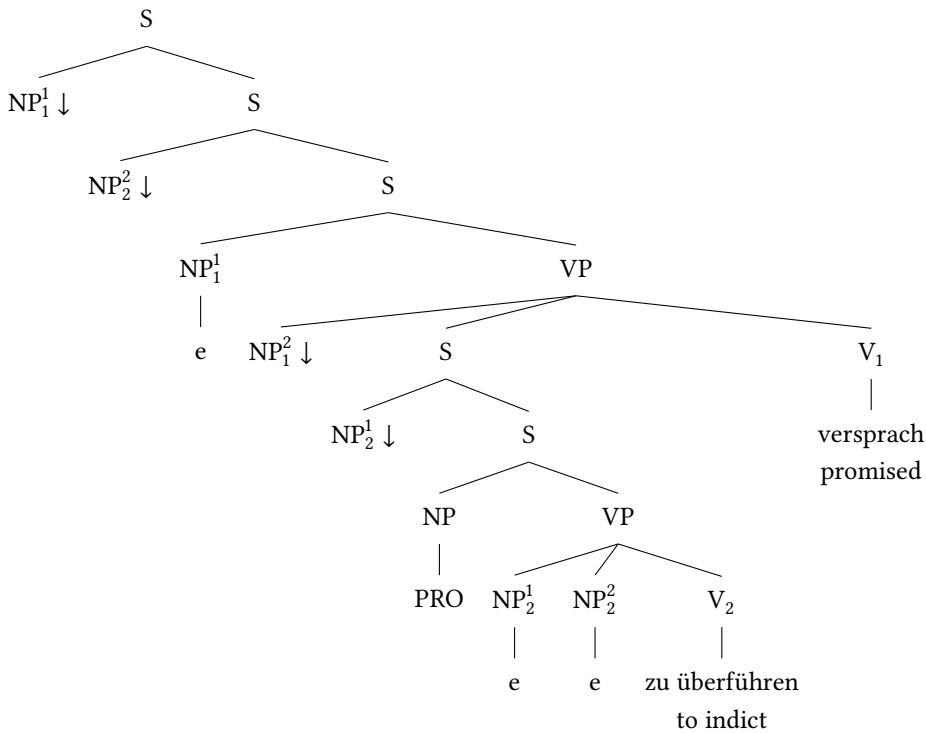


图 12.12: Analysis of the order $NP_1^1 NP_2^2 NP_1^2 NP_2^1 V_2 V_1$: adjunction to the S node between NP_2^2 and NP_1^1

Since the verb position has an effect on the clause type and hence on semantics, a lexical rule-based analysis would be also viable: a tree with the finite verb in initial position is licensed by a lexical rule that takes a tree with the verb in final position as input. This would be similar to the analyses in GB, Minimalism, and HPSG.

12.4 Passive

There is a possible analysis for the passive that is analogous to the transformations in Transformational Grammar: one assumes lexical rules that create a lexical item with a passive tree for every lexical item with an active tree (Kroch & Joshi 1985: 50–51).

Kroch & Joshi (1985: 55) propose an alternative to this transformation-like approach that more adequately handles so-called raising constructions. Their analysis assumes that arguments of verbs are represented in subcategorization lists. Verbs are entered into trees that match their subcategorization list. Kroch and Joshi formulate a lexical rule that corresponds to the HPSG lexical rule that was discussed on page 256, that is, an accusative object is explicitly mentioned in the input of the lexical rule. Kroch and Joshi

then suggest a complex analysis of the impersonal passive which uses a semantic null role for a non-realized object of intransitive verbs (p. 56). Such an analysis with abstract auxiliary entities can be avoided easily: one can instead use the HPSG analysis going back to Haider (1986a), which was presented in Section 9.2.

There are also proposals in TAG that use inheritance to deal with valence changing processes in general and the passive in particular (Candito 1996 and Kinyon, Rambow, Scheffler, Yoon & Joshi 2006 following Candito). As we saw in Section 10.2 of the Chapter on Construction Grammar, inheritance is not a suitable descriptive tool for valence changing processes. This is because these kinds of processes interact syntactically and semantically in a number of ways and can also be applied multiple times (Müller 2006, 2007c; 2007b: Section 7.5.2; 2013c; 2014a). See also Section 21.4 of this book.

12.5 Long-distance dependencies

The analysis of long-distance dependencies in TAG is handled with the standard apparatus: simple trees are inserted into the middle of other trees. Figure 12.13 在下一页 shows an example of the analysis of (10):

- (10) Who_i did John tell Sam that Bill likes _{-i}?

The tree for *WH COMP NP likes* _{-i} belongs to the tree family of *likes* and is therefore present in the lexicon. The tree for *tell* is adjoined to this tree, that is, this tree is inserted in the middle of the tree for *who that Bill likes* _{-i}. Such an insertion operation can be applied multiple times so that sentences such as (11) where *who* is moved across multiple sentence boundaries can be analyzed:

- (11) Who_i did John tell Sam that Mary said that Bill likes _{-i}?

There is another important detail: although the tree for (12) has the category S, (12) is not a grammatical sentence of English.

- (12) * who that Bill likes

This has to be captured somehow. In TAG, the marking OA ensures that a tree counts as incomplete. If a tree contains a node with marking OA, then an obligatory adjunction operation must take place at the relevant position.

12.6 New developments and theoretical variants

In Section 12.2, we introduced Multi-Component-TAG. There are a large number of TAG variants with different formal properties. Rambow (1994) gives an overview of the variants existing in 1994. In the following, I will discuss two interesting variants of TAG: Feature Structure-Based TAG (FTAG, Vijay-Shanker & Joshi 1988) and Vector-TAG (V-TAG, Rambow 1994).

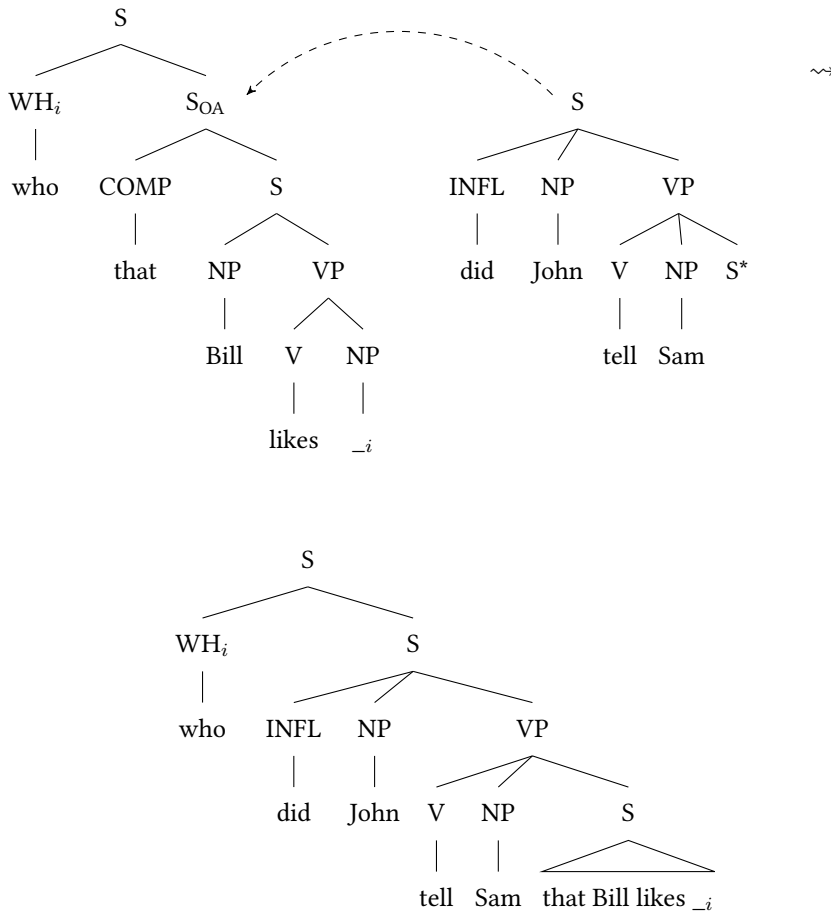


图 12.13: Analysis of long-distance dependencies in TAG

12.6.1 FTAG

In FTAG, nodes are not atomic (N, NP, VP or S), but instead consist of feature descriptions. With the exception of substitution nodes, each node has a top structure and a bottom structure. The top structure says something about what kind of properties a given tree has inside a larger structure, and the bottom structure says something about the properties of the structure below the node. Substitution nodes only have a top structure. Figure 12.14 on the following page shows an example tree for *laughs*. A noun phrase can be combined with the tree for *laughs* in Figure 12.14. Its top structure is identified with the NP node in the tree for *laughs*. The result of this combination is shown in Figure 12.15 on the next page.

In a complete tree, all top structures are identified with the corresponding bottom

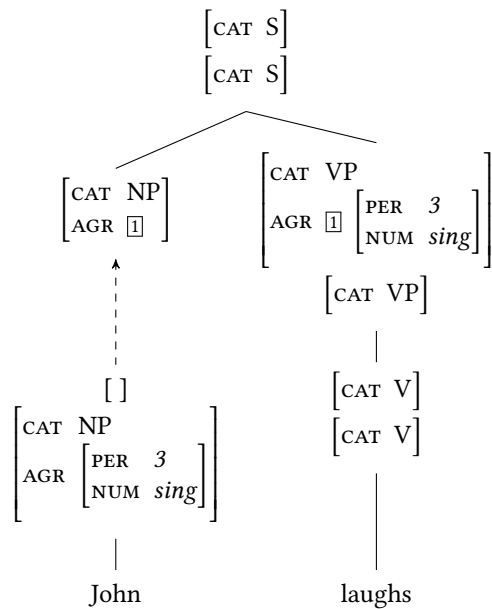


图 12.14: Elementary trees for *John* and *laughs* in FTAG

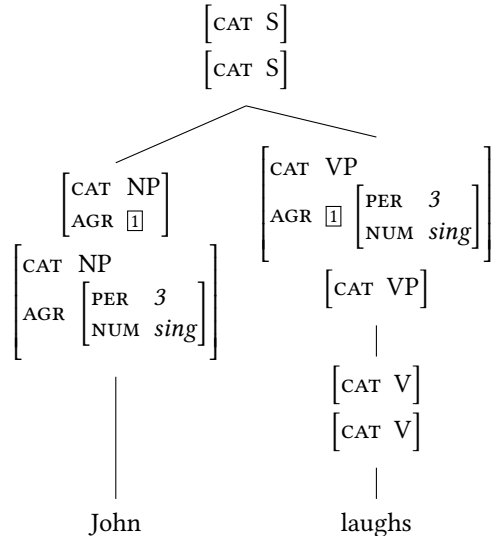


图 12.15: Combination of the trees for *John* and *laughs* in FTAG

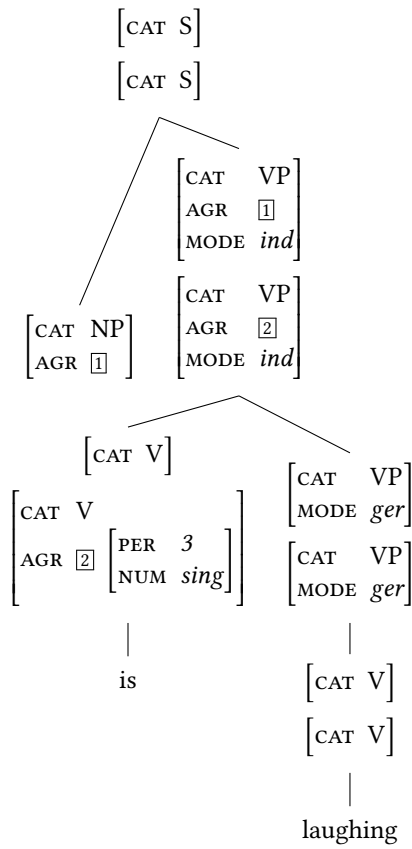


图 12.17: Result of obligatory adjunction in FTAG

one in the tree in Figure 12.17. As such, only NPs that have the same AGR value as the auxiliary can be inserted into the NP slot.

This example shows that, instead of the marking for obligatory adjunction that we saw in the section on long-distance dependencies, the same effect can be achieved by using incompatible feature specifications on the top and bottom structures. If there are incompatible top and bottom structures in a tree, then it cannot be a final derivation tree and therefore this means that at least one adjunction operation must still take place in order to yield a well-formed tree.

12.6.2 V-TAG

V-TAG is a variant of TAG proposed by Owen Rambow (1994) that also assumes feature structures on nodes. In addition, like MC-TAG, it assumes that elementary trees

consist of multiple components. Figure 12.18 shows the elementary lexical set for the ditransitive verb *geben* ‘give’. The lexicon set consists of a tree for the verb, an empty

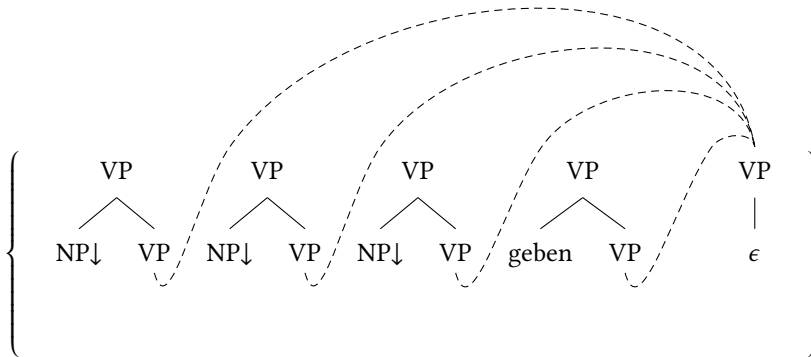


图 12.18: Lexicon set for *geben* ‘to give’ in V-TAG according to Rambow (1994: 6)

element of the category VP and three trees where a VP has been combined with an NP. As in MC-TAG, dominance relations are also indicated. The dominance constraints in Figure 12.18 ensure that all lower VP nodes dominate the highest VP node of the tree further to the right. The order of the arguments of the verb as well as the position of the verb is not given. The only thing required is that lower VP in the NP trees and lower VP in the *geben* tree dominate the empty VP node. With this lexicon set, it is possible to derive all permutations of the arguments. Rambow also shows how such lexical entries can be used to analyze sentences with verbal complexes. Figure 12.19 on the next page shows a verbal complex formed from *zu reparieren* ‘to repair’ and *versprochen* ‘promised’ and the relevant dominance constraints. Both of the first NP trees have to dominate *versprochen* and the third and fourth NP tree have to dominate *zu reparieren*. The order of the NP trees is not restricted and thus all permutations of NPs can be derived.

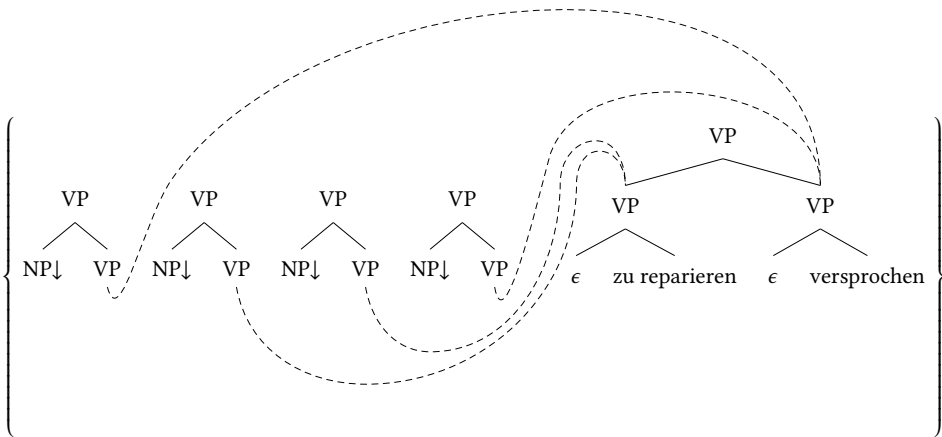


图 12.19: Analysis of the verbal complex *zu reparieren versprochen* in V-TAG

The interesting thing here is that this approach is similar to the one proposed by Berman (1996: Section 2.1.3) in LFG (see Section 7.4): in Berman's analysis, the verb projects directly to form a VP and the arguments are then adjoined.

A difference to other analyses discussed in this book is that there is always an empty element in the derived trees regardless of verb position.

12.6.3 The competence-performance distinction and the generative capacity of tree-local MC-LTAG

In many of the theories discussed in this book, a distinction is made between competence and performance (Chomsky 1965: Section I.1). Competence theories are supposed to describe linguistic knowledge, whereas a performance theory should explain how linguistic knowledge is used and why we make mistakes during speech production and comprehension, etc. See Chapter 15 for further discussion.

Joshi, Becker & Rambow (2000) discuss examples of center self embedding of relative clauses as those in (13b), and follow Chomsky & Miller (1963: 286) in the assumption that the fact that this kind of embedding is only possible up to three levels should not be described by grammar, but is rather due to processing problems with the hearer independent of their principle abilities with regard to grammar.

- (13) a. dass der Hund bellt, der die Katze jagt, die die Maus gefangen hat
 that the dog barks that the cat chases that the mouse caught has
 'that the dog that chases the cat that caught the mouse barks'

- b. dass der Hund, [₁ der die Katze, [₂ die die Maus gefangen hat, ₂] jagt ₁]
 that the dog that the cat that the mouse caught has chases
 bellt
 barks

What is interesting in this context is that it is possible to construct examples of center embedding so that they are easier to process for the hearer. In this way, it is possible to increase the number of center embeddings possible to process by one and therefore to show that all grammars that formulate a restriction that there may be at most two center-embedded relative clauses are incorrect. The following example from Hans Uszkoreit is easier to process since all embedded relative clauses are isolated and the verbs are separated by material from the higher clause.

- (14) Die Bänke, [₁ auf denen damals die Alten des Dorfes, [₂ die allen
 the benches on which back.then the old.people of.the village that all
 Kindern, [₃ die vorbeikamen ₃], freundliche Blicke zuwarfen ₂], lange Stunden
 children that came.by friendly glances gave long hours
 schweigend nebeneinander saßen ₁], mussten im letzten Jahr einem
 silent next.to.each.other sat must in.the last year a
 Parkplatz weichen.
 car.park give.way.to

‘The benches on which the older residents of the village, who used to give friendly glances to all the children who came by, used to sit silently next to one another had to give way to a car park last year.’

For other factors that play a role in processing, see Gibson (1998).

Joshi et al. (2000) discuss verbal complexes with reordered arguments. The general pattern that they discuss has the form shown in (15):

- (15) $\sigma(\text{NP}_1 \text{NP}_2 \dots \text{NP}_n) \text{V}_n \text{V}_{n-1} \dots \text{V}_1$

Here, σ stands for any permutation of noun phrases and V_1 is the finite verb. The authors investigate the properties of Lexicalized Tree Adjoining Grammar (LTAG) with regard to this pattern and notice that LTAG cannot analyze the order in (16) if the semantics is supposed to come out correctly.

- (16) $\text{NP}_2 \text{NP}_3 \text{NP}_1 \text{V}_3 \text{V}_2 \text{V}_1$

Since (17) is possible in German, LTAG is not sufficient to analyze all languages.

- (17) dass ihm₂ das Buch₃ niemand₁ zu lesen₃ versprechen₂ darf₁
 that him the book nobody to read promise be.allowed.to
 ‘that nobody is allowed to promise him to read the book’

Therefore, they propose the extension of TAG discussed in Section 12.2; so-called *tree-local multi-component LTAG* (Tree-local MC-LTAG or TL-MCTAG). They show that TL-MCTAG can analyze (17) but not (18) with the correct semantics. They claim that these

orders are not possible in German and argue that in this case, unlike the relative clause examples, one has both options, that is, the unavailability of such patterns can be explained as a performance phenomenon or as a competence phenomenon.

(18) $NP_2 NP_4 NP_3 NP_1 V_4 V_3 V_2 V_1$

If we treat this as a performance phenomenon, then we are making reference to the complexity of the construction and the resulting processing problems for the hearer. The fact that these orders do not occur in corpora can be explained with reference to the principle of cooperativeness. Speakers normally want to be understood and therefore formulate their sentences in such a way that the hearer can understand them. Verbal complexes in German with more than four verbs are hardly ever found since it is possible to simplify very complex sentences with multiple verbs in the right sentence bracket by extraposing material and therefore avoiding ambiguity (see Netter 1991: 5 and Müller 2007b: 262).

The alternative to a performance explanation would involve using a grammatical formalism which is just powerful enough to allow embedding of two verbs and reordering of their arguments, but rules out embedding of three verbs and reordering of the arguments. Joshi et al. (2000) opt for this solution and therefore attribute the impossibility of the order of arguments in (18) to competence.

In HPSG (and also in Categorical Grammar and in some GB analyses), verbal complexes are analyzed by means of argument composition (Hinrichs & Nakazawa 1989b, 1994). Under this approach, a verbal complex behaves exactly like a simplex verb and the arguments of the verbs involved can be placed in any order. The grammar does not contain any restriction on the number of verbs that can be combined, nor any constraints that ban embedding below a certain level. In the following, I will show that many reorderings are ruled out by communication rules that apply even with cases of simple two-place verbs. The conclusion is that the impossibility of embedding four or more verbs should in fact be explained as a performance issue.

Before I present arguments against a competence-based exclusion of (18), I will make a more general comment: corpora cannot help us here since one does not find any instances of verbs with four or more embeddings. Bech (1955) provides an extensive collection of material, but had to construct the examples with four embedded verbs. Meurers (1999b: 94–95) gives constructed examples with five verbs that contain multiple auxiliaries or modal verbs. These examples are barely processable and are not relevant for the discussion here since the verbs in (18) have to select their own arguments. There are therefore not that many verbs left when constructing examples. It is possible to only use subject control verbs with an additional object (e.g., *versprechen* ‘to promise’), object control verbs (e.g., *zwingen* ‘to force’) or Acl verbs (e.g., *sehen* ‘to see’ or *lassen* ‘to let’) to construct examples. When constructing examples, it is important make sure that all the nouns involved differ as much as possible with regard to their case and their selectional restrictions (e.g., animate/inanimate) since these are features that a hearer/reader could use to possibly assign reordered arguments to their heads. If we want to have patterns such as (18) with four NPs each with a different case, then we have to choose a verb that

governs the genitive. There are only a very small number of such verbs in German. Although the example constructed by Joshi et al. (2000) in (9b) fulfills these requirements, it is still very marked. It therefore becomes clear that the possibility of finding a corresponding example in a newspaper article is extremely small. This is due to the fact that there are very few situations in which such an utterance would be imaginable. Additionally, all control verbs (with the exception of *helfen* ‘to help’) require an infinitive with *zu* ‘to’ and can also be realized incoherently, that is, with an extraposed infinitival complement without verbal complex formation. As mentioned above, a cooperative speaker/author would use a less complex construction and this reduces the probability that these kinds of sentences arise even further.

Notice that tree-local MC-LTAG does not constrain the number of verbs in a sentence. The formalism allows for an arbitrary number of verbs. It is therefore necessary to assume, as in other grammatical theories, that performance constraints are responsible for the fact that we never find examples of verbal complexes with five or more verbs. Tree-local MC-LAG makes predictions about the possibility of arguments to be reordered. I consider it wrong to make constraints regarding mobility of arguments dependent on the power of the grammatical formalism since the restrictions that one finds are independent of verbal complexes and can be found with simplex verbs taking just two arguments. The problem with reordering is that it still has to be possible to assign the noun phrases to the verbs they belong to. If this assignment leads to ambiguity that cannot be resolved by case, selectional restrictions, contextual knowledge or intonation, then the unmarked constituent order is chosen. Hoberg (1981: 68) shows this very nicely with examples similar to the following:⁵

- (19) a. Hanna hat immer schon gewußt, daß das Kind sie verlassen will.
 Hanna has always already known that the child she leave wants
 ‘Hanna has always known that the child wants to leave her.’
 b. # Hanna hat immer schon gewußt, daß sie das Kind verlassen will.
 Hanna has always already known that she the child leave wants
 Preferred reading: ‘Hanna has always known that she wants to leave the child.’
 c. Hanna hat immer schon gewußt, daß sie der Mann verlassen
 Hanna has always already known that she the.NOM man leave
 will.
 wants.to
 ‘Hanna has always known that the man wants to leave her.’

⁵ Instead of *das* ‘the’, Hoberg uses the possessive pronoun *ihr* ‘her’. This makes the sentences more semantically plausible, but one then gets interference from the linearization requirements for bound pronouns. I have therefore replaced the pronouns with the definite article.

It is not possible to reorder (19a) to (19b) without creating a strong preference for another reading. This is due to the fact that neither *sie* ‘she’ nor *das Kind* ‘the child’ are unambiguously marked as nominative or accusative. (19b) therefore has to be interpreted as Hanna being the one that wants something, namely to leave the child. This reordering is possible, however, if at least one of the arguments is unambiguously marked for case as in (19c).

For noun phrases with feminine count nouns, the forms for nominative and accusative as well as genitive and dative are the same. For mass nouns, it is even worse. If they are used without an article, all cases are the same for feminine nouns (e.g., *Milch* ‘milk’) and also for masculines and neuters with exception of the genitive. In the following example from Wegener (1985: 45) it is hardly possible to switch the dative and accusative object, whereas this is possible if the nouns are used with articles as in (20c,d):

- (20)
- a. Sie mischt Wein Wasser bei.
she mixes wine water into
‘She mixes water into the wine.’
 - b. Sie mischt Wasser Wein bei.
she mixes water wine into
‘She mixes wine into the water.’
 - c. Sie mischt dem Wein das Wasser bei.
she mixes the.DAT wine the.ACC water into
‘She mixes the water into the wine.’
 - d. Sie mischt das Wasser dem Wein bei.
she mixes the.ACC water the.DAT wine into
‘She mixes the water into the wine.’

The two nouns can only be switched if the meaning of the sentence is clear from the context (e.g., through explicit negation of the opposite) and if the sentence carries a certain intonation.

The problem with verbal complexes is now that with four noun phrases, two of them almost always have the same case if one does not wish to resort to the few verbs governing the genitive. A not particularly nice-sounding example of morphologically unambiguously marked case is (21):

- (21) weil er den Mann dem Jungen des Freundes gedenken
because he.NOM the.ACC man the.DAT boy of.the.GEN friend remember
helfen lassen will
help let wants
‘because he wants to let the man help the boy remember his friend’

Another strategy is to choose verbs that select animate and inanimate objects so that animacy of the arguments can aid interpretation. I have constructed such an example where the most deeply embedded predicate is not a verb but rather an adjective. The

predicate *leer fischen* ‘to fish empty’ is a resultative construction that should be analyzed parallel to verbal complexes (Müller 2002a: Chapter 5).

- (22) weil niemand₁ [den Mann]₂ [der Frau]₃ [diesen Teich]₄ leer₄
 because nobody.NOM the.ACC man the.DAT woman this.ACC pond empty
 fischen₃ helfen₂ sah₁
 fish help saw
 ‘because nobody saw the man help the woman fish the pond empty’

If one reads the sentences with the relevant pauses, it is comprehensible. Case is unambiguously marked on the animate noun phrases and our word knowledge helps us to interpret *diesen Teich* ‘this pond’ as the argument of *leer* ‘empty’.

The sentence in (22) would correctly be analyzed by an appropriately written tree-local MC-LTAG and also by argument composition analyses for verbal complexes and resultative constructions. The sentence in (23) is a variant of (22) that corresponds exactly to the pattern of (18):

- (23) weil [der Frau]₂ [diesen Teich]₄ [den Mann]₃ niemand₁ leer₄
 because the.DAT woman this.ACC pond the.ACC man nobody.NOM empty
 fischen₃ helfen₂ sah₁
 fish help saw
 ‘because nobody saw the man help the woman fish the pond empty’

(23) is more marked than (22), but this is always the case with local reordering (Gisbert Fanselow, p. c. 2006). This sentence should not be ruled out by the grammar. Its markedness is more due to the same factors that were responsible for the markedness of reordering of arguments of simplex verbs. Tree-local MC-LTAG can not correctly analyze sentences such as (23), which shows that this TAG variant is not sufficient for analyzing natural language.

There are varying opinions among TAG researchers as to what should be counted as competence and what should be counted as performance. For instance, Rambow (1994: 15) argues that one should not exclude reorderings that cannot be processed by means of competence grammar or the grammatical formalism. In Chapter 6, he presents a theory of performance that can explain why the reordering of arguments of various verbs in the middle field is harder to process. One should therefore opt for TAG variants such as V-TAG or TT-MC-TAG (Lichte 2007) that are powerful enough to analyze the diverse reorderings and then also use a performance model that makes it possible to explain the gradual differences in acceptability.

An alternative to looking for a grammatical formalism with minimal expressive power is to not restrict the grammatical formalism at all with regard to its expressive power and instead develop as restrictive linguistic theories as possible. For further discussion of this point, see Chapter 17.

12.7 Summary and classification

In sum, we have seen the following: LTAG is lexicalized, that is, there is at least one lexical element in every tree. There are not any trees that correspond to the rule $S \rightarrow NP VP$ since no words are mentioned in this rule. Instead, there are always complex trees that contain the subject NP and the VP. Inside the VP, there can be as much structure as is necessary to ensure that the verb is contained in the tree. As well as the head, elementary trees in LTAG always contain the arguments of the head. For transitive verbs, this means that both the subject and the object have to be components of the elementary tree. This is also true of the trees used to analyze long-distance dependencies. As shown in Figure 12.13, the object must be part of the tree. The fact that the object can be separated from the verb by multiple sentence boundaries is not represented in the elementary tree, that is, recursive parts of grammar are not contained in elementary trees. The relevant effects are achieved by adjunction, that is, by insertion of material into elementary trees. The elementary tree for extraction in Figure 12.13 differs from the elementary tree for *likes* given in Figure 12.4 for the use in normal SVO clauses. Every minimal construction, in which *likes* can occur (subject extraction, topicalization, subject relative clauses, object relative clauses, passive, ...) needs its own elementary tree (Kallmeyer & Joshi 2003: 10). The different elementary trees can be connected using lexical rules. These lexical rules map a particular tree treated as underlying to other trees. In this way, it is possible to derive a passive tree from an active tree. These lexical rules are parallel to transformations in Transformational Grammar, however, one should bear in mind that there is always a lexical element in the tree, which makes the entire grammar more restrictive than grammars with free transformations.

An interesting difference to GB and variants of LFG, CG, and HPSG that assume empty elements is that the variants of TAG presented here⁶ do not contain empty elements in the lexicon. They can be used in trees but trees are listed as a whole in the lexicon.

Elementary trees can be of any size, which makes TAG interesting for the analysis of idioms (see Section 18.2). Since recursion is factored out, trees can contain elements that appear very far away from each other in the derived tree (extended domains of locality).

Kasper, Kiefer, Netter & Shanker (1995) show that it is possible to transfer HPSG grammars that fulfill certain requirements into TAG grammars. This is interesting as in this way one arrives at a grammar whose complexity behavior is known. Whereas HPSG grammars are generally in the Type-0 area, TAG grammars can, depending on the variant, fall into the realm of Type-2 languages (context-free) or even in the larger set of the mildly context-sensitive grammars (Joshi 1985). Yoshinaga, Miyao, Torisawa & Tsujii (2001) have developed a procedure for translating FB-LTAG grammars into HPSG grammars.

⁶ See Rambow (1994) and Kallmeyer (2005: 194), however, for TAG analyses with an empty element in the lexicon.

Comprehension questions

1. How are long-distance dependencies analyzed in TAG? Does one need empty elements for this?
2. Is it possible to analyze the reordering of arguments of multiple verbs using standard TAG processes?

Exercises

1. Analyze the following string in LTAG:
 (24) der dem König treue Diener
 the the.DAT king loyal servant
 ‘the servant loyal to the king’

Further reading

Some important articles are Joshi, Levy & Takahashi (1975), Joshi (1987a), and Joshi & Schabes (1997). Many works discuss formal properties of TAG and are therefore not particularly accessible for linguistically interested readers. Kroch & Joshi (1985) give a good overview of linguistic analyses. An overview of linguistic and computational linguistic works in TAG can be found in the volume edited by Abeillé and Rambow from 2000. Rambow (1994) compares his TAG variant (V-TAG) to Karttunen’s *Radical Lexicalism* approach, Uszkoreit’s GPSG, Combinatorial Categorical Grammar, HPSG and Dependency Grammar.

Shieber & Johnson (1993) discuss psycholinguistically plausible processing models and show that it is possible to do incremental parsing with TAG. They also present a further variant of TAG: synchronous TAG. In this TAG variant, there is a syntactic tree and a semantic tree connected to it. When building syntactic structure, the semantic structure is always built in parallel. This structure built in parallel corresponds to the level of Logical Form derived from S-Structure using transformations in GB.

Rambow (1994: Chapter 6) presents an automaton-based performance theory. He applies it to German and shows that the processing difficulties that arise when reordering arguments of multiple verbs can be explained.

Kallmeyer & Romero (2008) show how it is possible to derive MRS representations directly via a derivation tree using FTAG. In each top node, there is a reference to the semantic content of the entire structure and each bottom node makes reference to the semantic content below the node. In this way, it becomes possible to insert an adjective (e.g., *mutmaßlichen* ‘suspected’) into an NP tree *alle Mörder* ‘all murderers’ so that the adjective has scope over the nominal part of the NP (*Mörder* ‘murderers’): for adjunction of the adjective to the N node, the adjective can access the semantic content of the noun. The top node of *mutmaßlichen* is then the top node of the combination

12 Tree Adjoining Grammar

mutmaßlichen Mörder ‘suspected murderers’ and this ensures that the meaning of *mutmaßlichen Mörder* is correctly embedded under the universal quantifier.

第二部分 讨论

第十三章 语言知识的天赋性

如果我们试着将本书提到的诸多理论进行比较的话，我们会发现这些理论之间有很多相似之处。¹ 在所有理论框架中都有利用属性-值对来描述语言对象的理论变体。这些理论提出的句法结构有时候也相似。但是，有很多差异会导致不同流派之间的激烈争论。这些差异在于他们是否会提出转换、空成分、基于短语或基于词的分析，两分支或扁平结构这样的假设。

每一种理论不仅要描述自然语言，还要解释自然语言。为某一给定语言设定允准其结构的无限数量的文法是可行的（见 73 页的练习 1）。这些文法在观察上是充分的（*observationally adequate*）。如果一种语法能够符合观察以及母语者语感的话，该语法就实现了描述上的充分性（*descriptive adequacy*）。² 如果一种语言学理论能够用于为每一种自然语言建立一个描述充分的语法，那么这种语言学理论就是描述充分的。但是，能够实现描述充分的语法不一定能够实现解释充分（*explanatory adequacy*）。能够实现解释充分的语法需要与语言习得的数据相符，也就是说那些貌似能被人类说话者习得的语法（Chomsky 1965: 24–25）。

Chomsky (1965: 25) 假设儿童原则上已经具有语法在特定领域中是怎样的知识，然后从语言输入中提取出一个给定语法事实上应该是怎样的信息。主流的生成语法（MGG）中最著名的习得理论变体就是原则 & 参数理论，它认为参数化的原则限定了可能的语法结构，并且儿童只需要在语言习得中设置参数（见 3.1.2 节）。

这些年来，天赋性假说，也称作先天论历经几经修改。特别是，关于先天语言知识组成部分的假设，所谓的普遍语法（UG），一直在变化。

先天论经常遭到构式语法、认知语法的支持者以及其他理论学派研究人员的反对。对于那些一般用于支持句法范畴、句法结构或者句法结构中语言对象关系

¹ 我们需要区分理论（*theory*）和框架（*framework*）这两个术语。框架是指构建理论时使用的一些共同的假设和工具。在本书中，我讨论了德语的理论。这些理论是在一些不同框架（GB、GPSG、HPSG、LFG...）中形成的，另外当然有其他语言的其他理论也遵循同样的基本假设。这些理论不同于这里讨论的德语理论，但他们是在同一框架内形成的。Haspelmath (2010b) 倡导一种无框架的语法理论。如果语法理论使用了不相容的工具，那么就难以展开语言的对比。所以针对英语的非局部依存关系所提出的转换假设跟为德语提出的 SLASH 机制之间就无法进行比较了。我同意 Haspelmath 所说的形式工具可能会导致偏见的观点，但是不管怎样语言事实终究需要描写。如果理论之间没有共性，我们得到的将是依据个人的框架得出的孤立的理论。如果一个框架包含共同的术语以及建立无框架语法理论的标准，那么这一框架就是无框架语法理论。进一步论述见 Müller (2015a) 以及本书的第 22 章。

² 因为有主观因素的作用，所以这一术语有时不是特别有用。对于世界上的语言的每一个可以观察到的语序都可以从共同的限定语-中心语-补足语到构型中推导出来并且也只能通过移位移到左边（见 4.6.1 节对这些观点的论述）的这样的假设，并不是每一个人都认为这些语法理论是自然正确的。

具有天赋性的事实，人们提出了其它的解释。受到批评的另外一点是，由于很多规定被简单化地假设为普遍语法的一部分，语法分析真实存在的复杂性被模糊掉了。下面是 GB/最简方案分析中一些论证的基本程序。

1. 我已经为 S 语言中的 P 现象提供了一种分析；
2. 这种分析是优雅的/概念上简洁的/我的³；
3. 不可能去学习相关的结构或原则；
4. 所以，这一分析中的假设 A1 到 An 一定是说话者天赋知识的一部分。

通过向 UG 中随意增加假设，可以让其余的分析十分简单。

下面的章节会简要评价一些支持语言特有的天赋知识的证据。我们会发现所有证据都存在争议。在下面的章节中，我会讨论语法建构的一些基本问题，语言能力与语言使用之间的差异、如何为语言使用现象建模、语言习得理论以及其他存在争议的问题，例如，在语言表征中设置空成分是否理想以及语言是应该主要基于词的属性还是短语的模式进行解释。

在我们转向这些激烈争论的话题之前，我想先讨论一下争论最为激烈的一个话题，即天赋语言知识的问题。在文献中，有人找到了以下支持天赋知识的证据：

- 存在句法普遍性，
- 习得的速度，
- 语言习得存在一个“关键期”，
- 所有儿童都可以习得语言，但是灵长类动物不可以，
- 儿童会自发地将皮钦语规范化，
- 语言处理位于大脑的特定部分，
- 所谓的语言能力与普遍认知能力的分离：
 - 威廉氏综合症，
 - FoxP2 基因发生突变的 KE 氏家族，以及
- 刺激贫乏论。

Pinker (1994) 对这些证据作了很好的概述。Tomasello (1995) 对这本书作了批评性评论。其中个别观点会在下面进行讨论。

³ 也参见 <http://www.youtube.com/watch?v=cAYDiPizDIIs>. 2015 年 12 月 1 日.

13.1 句法普遍性

句法的普遍性被作为论据来证明语言知识的天赋性（如Chomsky 1998: 33、Pinker 1994: 237–238）。在文献中，对于什么是普遍性的、什么是具体的有不同的看法。支持普遍性的突出代表有：⁴

- 中心语导向参数
- \bar{X} 结构
- 语法功能（如主语或宾语）
- 约束原则
- 长距离依存的属性
- 时、体、态的语法要素
- 词性
- 递归或自嵌套

我们将在下面的内容中逐一讨论这些普遍性。应该强调的是，这些普遍性是没有任何形式的一致性的，而且这些观察到的属性实际上需要考察天赋的语言知识。

13.1.1 中心语导向参数

中心语导向参数在第 3.1.2 节就已经介绍过了。第 80 页（7）中的例子在这里重复表示为下面的（1），这些例子说明了日语的结构是英语结构的镜像：

- （1） a. be showing pictures of himself
 b. zibun -no syasin-o mise-te iru
 他自己 PREP 照片 显示 COP

为了捕捉到这些事实，我们提出一个参数来表示中心语相对于其所带论元的位置（如 Chomsky 1986b: 146、1988: 70）。

Radford (1990: 60–61; 1997: 19–22), Pinker (1994: 234, 238), Baker (2003: 350) 和其他作者声称，在假定了中心语导向性参数的基础上，在动词所管辖的方向和介词所管辖的方向之间具有某种外在或隐含的关系，也就是说，动词末位语序的语言具有后置词，或具有 VO 语序的语言具有前置词。这一观点可以通过英语和日语这

⁴ Frans Plank 在康斯坦斯有一个关于什么是普遍性的存档文集 (Plank & Filimonova 2000): <http://typo.uni-konstanz.de/archive/intro/>。截至 2015 年 12 月 23 日，一共有 2029 条记录。这些记录根据它们的质量来进行标注，结果发现，许多的普遍性具有统计上的普遍性，它们适用于绝大多数语言，但是有一些例外。一些普遍性被标记为绝对的，也就是说，极少有例外。有 1153 条被标记为带有问号的绝对，有 1021 条被标记为不带问号的绝对值。许多普遍性是通过蕴含的普遍性捕捉到的，即，它们具有这样的形式：如果一种语言具有属性 X，那么它也有属性 Y。在存档文集中列出的普遍性部分上是非常具体的，而且指向具体语法属性的历时变化。比如说，第四条表示：“如果呼格的典型元素是一个前缀，那么这个前缀来自于第一人称领有者或第二人称主语。”

两种语言来验证，也可以如例（1）所示：在介词短语中，*no* 出现在代词的后面，名词 *syasin-o*（“图片”）位于其所属的 PP 后面，主动词在它的宾语后面，助词 *iru* 位于主动词 *mise-te* 之后。具体的短语就是英语中各自短语的镜像。

对此，一个反例足以说明这一说法是站不住脚的。实际上，像波斯语这种语言是遵循动词末位的语序的，但它是有前置词的，如（2）所示：

- (2) man ketâb-â-ro be Sepide dâd-am.
我 书-PL-RÂ INF Sepide 给-1SG
‘我把书给了 Sepide。’

在第 3.1.4 节，我们说明了德语不能简单地按照这一参数来描写：德语是动词位于末位的语言，但是它既有前置词，也有后置词。《语言结构的世界地图》这本书囊括了 41 种带有 VO 语序和后置词的语言，还有 14 种带有 OV 语序和前置词的语言 (Dryer 2013b,a)。⁵ Dryer (1992) 更早之前在小范围样本的语言研究中也指出，按照中心语倾向性参数所预测的结果是有例外的。

不仅如此，Gibson & Wexler (1994: 422) 指出，用一个参数来表示中心语的位置是不够的，因为英语和德语及荷兰语中补足语可以出现在他们的补语前面；不过，英语是一个 VO 语序的语言，而德语和荷兰语是 OV 语序的语言。

如果我们希望通过句法范畴来决定管辖的方向 (Gibson & Wexler 1994: 422, Chomsky 2005: 15)，那么我们就必须假设句法范畴属于普遍语言的一部分（更多内容请看第 13.1.7 节）。对于这类假设来说，前置词和后置词还是有问题的，因为他们通常都被指派给同一个范畴 (P)。如果我们要对前置词和后置词引入特殊的范畴，那么像第 87 页所示的词类的四分法就不管用了。相反，我们需要一个额外的二元特征来自动预测出八种范畴，虽然只有五种（四种常规的，加上一个额外的）是实际上需要的。

我们可以看到，Pinker 构造的管辖方向之间的关系作为普遍的规则实际上是正确的，但是这只是一种倾向性，而不是一条严格的规则，也就是说，还有许多语言，其中前置词或后置词的使用与动词的位置之间具有相关性 (Dryer 1992: 83)。⁶

许多语言中，介词由动词演变而来。在汉语语法中，有一类词通常被称作副动词。这些词可以用作介词，也可以用作动词。如果我们历时地观察语言，那么我们就可以找到这些倾向性的解释，而不用参考天赋的语言学知识（请看 Evans & Levinson 2009a: 445）。

进而，我们可以解释与语言处理的倾向性相关的一些事实：具有相同管辖方向的语言（图 13.1a-b）与具有相反管辖方向的语言（图 13.1c-d）相比，动词和前/后置词之间的距离更小。从语言处理的角度来看，具有相同管辖方向的语言更易于被人们接受，因为他们允许听者更好地识别动词短语的组成部分 (Newmeyer (2004a: 219–221) 引用了 Hawkins (2004: 32) 有关语言处理的相关性的讨论，也可以参考 Dryer (1992: 131))。这一倾向性就可以这样解释为表现偏好的语法化（请看第 15 章有关语言能力和语言运用的区分），并且与天赋的语言特有的知识是没有关系的。

⁵ http://wals.info/combinations/83A_85A#2/15.0/153.0, 2015 年 12 月 23 日。

⁶ Pinker (1994: 234) 在他的观点中使用了 *usually*（通常）这个词。由此，他暗示了这里是有例外的，而且介词的顺序与动词管辖的方向之间的关系实际上是一种倾向性，而不是一个具有普遍意义的可应用的规则。但是，在随后的内容中，他认为，中心语倾向性参数构成了天赋的语言学知识的一部分。Travis (1984: 55) 讨论了汉语中与这个关系不对应的例子。之后她提出将中心语倾向性函数作为一种缺省的函数，这个函数可以被语言中的其他限制所覆盖。

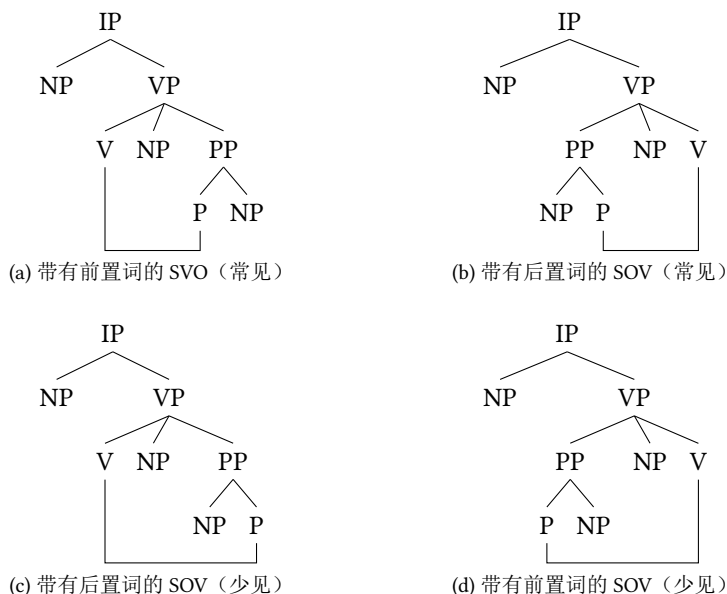


图 13.1: 基于Newmeyer (2004a: 221)的遵循不同的中心语语序的动词与介词之间的距离

13.1.2 \bar{X} 结构

通常认为, 所有语言都有对应于 \bar{X} 模式的句法结构(请看第2.5节)(Pinker 1994: 238; Meisel 1995: 11, 14; Pinker & Jackendoff 2005: 216)。但是, 也有像Dyirbal(澳大利亚)的这种语言, 其中句子是无法用层级结构来表示的。由此, Bresnan (2001: 110)认为, 塔加路语、匈牙利语、马拉雅拉姆语、瓦勒皮里语、Jiwarli 语、Wambaya 语、雅卡语和其他相应的语言没有 VP 节点, 而是带有(3)这样格式的规则:

$$(3) \quad S \rightarrow C^*$$

这里, C^* 表示任意数目的短语成分, 而且结构中没有中心语。其他不带中心语的结构例子将在第21.10节讨论。

\bar{X} 结构被用来限制可能规则的形式。假设是这些限制减少了我们可以构建的语言类型, 并且, 根据这个假设, 使得语言更容易被获得。但是正如Kornai & Pullum (1990)所展示的, \bar{X} 结构的假说不能限制可能语法的数量, 如果我们允准空的中心语的话。管约论使用了空的中心语, 而且在最简方案中, 他们还有显著数量的增加。比如说, (3) 中的规则可以被重构为:

$$(4) \quad V' \rightarrow V^0 C^*$$

这里, V^0 是一个空中心语。由于限定语是可选的, V' 可以投射到 VP 上, 这样我们就得到了一个对应于 \bar{X} 模式的结构。

除了具有自由语序的语言带来的问题以外，附加结构还有一些问题：Chomsky 在 \bar{X} 理论中有关形容词性结构的分析（Chomsky 1970: 210；还可以参考本书的第 2.5 节，尤其是图 2.8 on page 69）不能直接应用到德语，因为，与英语不同的是，德语中的形容词短语是中心语后置，并且程度修饰语必须直接位于形容词的前面：

- (5) a. der auf seinen Sohn sehr stolze Mann
 DET PREP 他的 儿子 非常骄傲 男人
 ‘对自己的儿子感到非常骄傲的男人’
 b. *der sehr auf seinen Sohn stolze Mann
 DET 非常PREP 他的 儿子 骄傲 男人
 c. *der auf seinen Sohn stolze sehr Mann
 DET PREP 他的 儿子 骄傲 非常男人

根据 \bar{X} 模式，*auf seinen Sohn* 必须与 *stolze* 相组合，而且只有这样才能得到与它的限定语相组合的 \bar{A} 投射（请看图 2.8 on page 69 中英语形容词性短语的结构）。这样就只可能推导出诸如（5b）或（5c）的语序。这两个语序在德语中都是不可能的。如果我们假设德语跟英语完全一致，而且基于某种原因，形容词的补足语必须移到左边的话，这样才有可能拯救 \bar{X} 范式。如果我们允许这种补救措施，那么当然任何一种语言都可以用 \bar{X} 范式来进行描写。结果是我们必须针对许多语言而提出巨量的移位规则，而这从心理语言学的角度来看是异常复杂和困难的。请看第 15 章有关与语言运用相适应的语法的讨论。

\bar{X} 理论的最严格形式的问题在第 2.5 节进行了说明，它被叫做 *hydra* 从句（Perlmutter & Ross 1970; Link 1984; Kiss 2005）：

- (6) a. [[der Kater] und [die Katze]], die einander lieben
 DET 公猫 和 DET 猫 CONJ 互相 爱
 ‘公猫和（母）猫相互爱慕’
 b. 约会的 [[男孩儿] 和 [女孩儿]] 是我的朋友。

因为（6）中的关系从句指称一组人，他们只能附加在并列的结果上。整个并列结构是一个 NP，但是，附加语实际上应该附加在 \bar{X} 层面上。与德语和英语中的关系小句相反的情况是波斯语中的形容词：Samvelian (2007) 提出的分析是，形容词直接与名词相组合，而且只有名词和形容词的组合才能与 PP 论元相组合。

关于德语和英语的讨论说明了限定语和附加语的导入不能在具体的投射层面进行约束，而且前面非构型语言的讨论已经显示了中间层的假设并不适用于每一种语言。

还需要指出的是，Chomsky 本人在 1970 年提出，有的语言可以脱离 \bar{X} 范式 (1970: 210)。

如果我们希望把所有的信息编码在词汇中，那么我们就需要通过能够表示共性的极为抽象的组合规则。这种组合性规则的一个例子是范畴语法的多重应用规则（请看第 8 章），以及最简方案中的合并（请看第 4 节）。上面这些规则仅仅说明了两个语言学对象被组合在一起。这种组合性当然存在于每一种语言中。但是，对于完全词汇化的语法而言，如果我们允许空中心语并且提出某些特异的假设的话，才有可能来描写语言。这将在第 21.10 节具体讨论。

13.1.3 主语和宾语的语法功能

Bresnan & Kaplan (1982: xxv)、Pinker (1994: 236–237)、Baker (2003: 349) 和其他人认为所有的语言都有主语和宾语。为了说明这个观点到底意味着什么，我们必须探究这些术语。对于大多数的欧洲语言来说，很容易说明主语和宾语是什么（请看第 1.7 节）；但是，并不是所有的语言都是这样的，或者说在有些语言中用这些术语是完全没有意义的（Croft 2001: Chapter 4; Evans & Levinson 2009a: Section 4）。

在 Pinker 研究的 LFG 这类理论中，语法功能充当了重要的角色。事实上，我们是否应该把句子看作是主语、宾语或者是在我们要指派语法功能给论元时用来显示至少有一些讨论余地的具体界定的句子论元（xCOMP）（Dalrymple & Lødrup 2000; Berman 2003b, 2007; Alsina, Mohanan & Mohanan 2005; Forst 2006），都是有争议的。由此，我们有可能在所有语言中找到指派给函项的论元的语法功能。

与 LFG 不同的是，语法功能在 GB（请看 Williams 1984; Sternefeld 1985a）和范畴语法中是毫不相关的，语法功能只能间接地通过指向树的位置来决定。由此，在第 3 章讨论的方法中，主语是位于 IP 的限定语位置上的短语。

在乔姆斯基式理论的后半期，也有一些看起来对应于语法功能的功能性节点（AgrS, AgrO, AgrIO 请看第 132 页）。但是，Chomsky (1995b: Section 4.10.1) 指出，这些功能性范畴只是出于理论内部的原因而提出的，而且应该排除在普遍语法的范畴之外。请看 Haider (1997a) 和 Sternefeld (2006: 509–510) 对德语的描写，其中无功能性的投射不能被激发。

HPSG 所持的观点较为中立：它用一个特殊的配价特征来表示主语（在德语语法中，有一个中心语特征包括了非定式动词的主语的表达）。但是，SUBJ feature 的值是从更为普遍的理论性角度推导出来的：在德语中，带有结构格的最旁格成分就是主语（Müller 2002a: 153; 2007b: 311）。

在 GB 理论（扩展的投射原则，EPP、Chomsky (1982: 10)）和 LFG（主语条件）中，都有用来确保每个句子都有一个主语的原则。通常认为这些原则是具有语言共性的。⁷

正如前面提到的，GB 中没有语法功能，但是有对应于语法功能的结构位置。对应于主语的位置是 IP 的限定语。EPP 说明了在 SpecIP 中必须有一个元素。如果我们假定这个原则具有语言共性，那么每一种语言都必须在这个位置有一个元素。正如我们看到的，这个普遍性的观点是有反例的：德语。德语有无人称被动句（7a），其中有无主语的动词（7b,c）和形容词（7d–f）。⁸

- (7) a. dass noch gearbeitet wird
 CONJ 仍然 工作 COP
 ‘人们仍在工作’
 b. Ihm graut vor der Prüfung.
 他.DAT 害怕 PREP DET 考试
 ‘他怕考试。’

⁷ 但是，Chomsky (1981a: 27) 允许语言不带主语。他假设，这可以通过一个参数来解决。Bresnan (2001: 311) 提出了主语条件，但是在脚注中指出，也许有必要将这一条件参数化，从而使得它适合于某些语言。

⁸ 更多有关德语中无主语动词的讨论，请参考 Haider (1993: Sections 6.2.1, 6.5)、Fanselow (2000b)、Nerbonne (1986b: 912) 和 Müller (2007b: Section 3.2)。

- c. Mich friert.
我.ACC 冷
‘我冷死了。’
- d. weil schulfrei ist
因为假期 COP
‘因为今天不用上学’
- e. weil ihm schlecht ist
因为他.DAT 生病 COP
‘因为他生病了’
- f. Für dich ist immer offen.⁹
对于你 COP 总是 开放
‘我们一直为你敞开大门。’

大多数不带主语的谓词也可以带形式主语。如例（8）所示：

- (8) dass es ihm vor der Prüfung graut
CONJ EXPL 他 PREP DET 考试 害怕
‘他怕考试。’

但是，也有不能带形式主语 *es* 的动词，如Reis (1982: 185)提出的例（9a）中的 *liegen*（位于）。

- (9) a. Mir liegt an diesem Plan.
我.DAT 位于PREP 这 计划
‘这个计划对我很重要。’
b. *Mir liegt es an diesem Plan.
我.DAT 位于it PREP 这 计划

尽管如此，EPP 和主语条件也通常用在德语中。Grewendorf (1993: 1311) 提出了一个空形式主语来填充到没有主语的结构的主语位置上。

Berman (1999: 11, 2003a: Chapter 4) 在 LFG 框架下提出，动词形态可以在德语中填充主语角色，这样即使在无主句中，主语位置也在 f-结构中被填充了。有一个限制证明所有不带 PRED value 的 f-结构必须是第三人称单数才能应用到没有主语的结构中。在定式动词的主谓一致信息必须匹配到没有主语的结构，这样无主语结构的动词屈折变化就被限制为第三人称单数 (Berman 1999)。Berman (1999: 11;

正如我们在第 149 页看到的，有些在最简方案下工作的研究人员甚至认为每个句子都有一个宾语 (Stabler 在 Veenstra (1998: 61, 124) 中这样引述)。单价动词的宾语被假定为一个空成分。

如果我们允准这样的机制，当然就很容易维持许多普遍性的假设：我们假设一种语言 X 具有属性 Y，然后假设结构性的成分是不可见而且没有意义的。这种分析只能用在理论内部以取得理论的一致性（请看 Culicover & Jackendoff 2005: Section 2.1.2）。¹⁰

⁹ Haider (1986a: 18).

¹⁰ 关于语言习得的内容，请看第 16 章。

13.1.4 约束原则

管辖代词约束的原则也被看作是普遍语法的一部分 (Chomsky 1998: 33、Crain, Thornton & Khlentzos 2009: 146、Rizzi 2009a: 468)。GB 理论中的约束理论包括三条原则：原则 A 规定诸如 *sich* 或 *himself* 的反身代词指称某个局部域（约束域）内的一个成分（先行词）。简言之，我们可以这样说，一个反身代词必须要指称一个共指的论元成分。

- (10) Klaus_i sagt, dass Peter_j sich_{*i/j} rasiert hat.
Klaus 说 CONJ Peter 他自己刮胡子has

原则 B 是针对人称代词的，它规定人称代词不能指称他们的约束域内的成分。

- (11) Klaus_i sagt, dass Peter_j ihn_{i/*j} rasiert hat.
Klaus 说 CONJ Peter 他 刮胡子has

原则 C 规定了有指的表达式的指称内容。根据原则 C，表达式 A₁ 不能指称表达式 A₂，如果 A₂c-统制 A₁ 的话。c-统制被界定为话语的结构。对于 c-统制有许多不同的定义；一个简单的版本是 A c-统制 B，如果在短语成分结构中有一条路径从 A 向上直到下一个节点，然后只能往下到 B。

对于 (12a) 中的例句来说，这就意味着 Max 和 *er*（他）不能指称相同的个体，因为 *er* c-统制 Max。

- (12) a. Er sagt, dass Max Brause getrunken hat.
他说 that Max 苏打水喝 AUX
'他说 Max 喝了苏打水。'
b. Max sagt, dass er Brause getrunken hat.
Max 说 that 他苏打水喝 AUX
'Max 说他喝了苏打水'
c. Als er hereinkam, trank Max Brause.
当 他进来 喝 Max 苏打水
'当他进来的时候，Max 正在喝苏打水。'

这在 (12b) 是可能的，但是，这里并没有 c-统制的关系。对于 *er*（他）来说，只有他在不指称动词 *getrunken*（喝）的另一个论元的时候是具有 c-统制关系的，而这就是 (12b) 中的情况。相似地，在 (12c) 的 *er*（他）和 Max 之间没有统制关系，因为代词位于复杂结构中。*er*（他）和 Max 就可以指称 (12b) 和 (12c) 中相同和不同的成分了。

Crain, Thornton & Khlentzos (2009: 147) 指出，(12b,c) 和对应的英语例子是有歧义的，而 (12a) 不是，这跟原则 C 有关。这意味着有一个含义是不适用的。为了获得正确的约束原则，学习者需要知道哪些意义是表达式不具备的。作者指出，儿童在三岁就已经熟练掌握了原则 C，而且他们由此得到结论说，原则 C 貌似属于天赋的语言知识。（这是一个经典的论述。关于刺激贫乏论，请看第 13.8 节，而有关否定证据的更多内容请看第 13.8.4 节）。

Evans & Levinson (2009b: 483)指出, 原则 C 具有很强的跨语言的倾向性, 但还是有一些例外的。比如说, 他们提到 Abaza 语中的相互表达式, 其中表示 *each other* (互相) 的词缀出现在主语位置上, 而不是宾语位置上, Guugu Yimidhirr 语也是这样, 其中上层句可以跟从属句中的完全 NP 共指。

进而, Fanselow (1992b: 351)提出 (13) 中的例子, 他们说明原则 C 不是一个好的句法原则。

- (13) a. Mord ist ein Verbrechen.
谋杀 是 ART 犯罪
- b. Ein gutes Gespräch hilft Probleme überwinden.
ART 好 谈话 帮助问题 克服
'一次好的交谈可以帮助解决问题。'

(13a) 是说当某人杀死某人时, 这是一项犯罪, 而 (13b) 是倾向于跟他人聊天, 而不是跟自己聊天。在这些句子中, 名词化的 *Mord* (谋杀) 和 *Gespräch* (谈话) 使用时没有带上原始动词的任何论元成分。所以说, 没有表示句法统制关系的任何论元成分。尽管如此, 名词化的动词的论元不能互指。所以说, 这样就有一条原则说明, 谓词的论元槽上必须解释为非互指的, 只要论元的指认没有通过语言学手段明确地指出来。

综上所述, 约束理论还有一些尚未解决的问题。英语中对于原则 A-C 的 HPSG 理论变体不能用于德语 (Müller 1999a: Chapter 20)。在 LFG 理论框架下, Dalrymple (1993)提出了一个约束理论的变体, 其中代词性表达的约束属性在词汇中得到约束。按照这个方法, 就可以描述具体语言的代词属性了。

13.1.5 长距离依存的属性

在前面章节中讨论的长距离依存受限于某种条件。例如, 在英语中, 不能把名词短语中的部分成分从句子中提取出来。Ross (1967: 70)把这个限制称为复杂 NP 限制 (*Complex NP Constraint*)。在后面的研究中, 还有 Ross (1967: Section 5.1.2)构建的右界限制 (*Right Roof Constraint*)。对此, 有人倾向于把这些限制归纳为一个单一的、更具有普遍性的约束, 即邻接原则 (Chomsky 1973: 271、1986a: 40、Baltin 1981, 2006)。邻接原则被认为是具有语言共性的。邻接限制说明移位操作可以跨越一个约束节点, 而到底什么是约束节点, 则取决于研究的语言 (Baltin 1981: 262、2006; Rizzi 1982: 57; Chomsky 1986a: 38–40)。¹¹

目前, 对于邻接是否应该被看作是天赋的语言知识这个问题, 在 GB 和最简方案中有不同的看法。Hauser, Chomsky & Fitch (2002)认为邻接并不属于天赋的语言能力的一部分, 至少严格来说不是, 而是广义上的跟语言相关的限制, 这些限制可以从更为普遍的认知角度推导出来 (请看第422页)。由于邻接在其他当代

¹¹ Newmeyer (2004b: 539–540) 指出了约束节点的不同语言差异的一个概念性的问题: 邻接性是天赋的语言特有的原则, 因为它太抽象了, 说话者是无法学习的。但是, 如果参数化需要说话者从语言输入中挑选一些范畴的话, 那么相应的限制就必须从输入中推导出来, 至少是有可能判断出所涵盖的范畴。这就导致一个问题, 最初有关语言习得不可能性的论断是对的。请看第 13.8 节有关刺激缺乏论的内容, 以及第 16.1 节有关基于参数的语言习得的理论。

值得注意的是, 具有某种词性的参数成分也要求在普遍语法中保有相应的词性信息。

的文献 (Newmeyer 2005: 15, 74–75、2004a: 184、Baltin 2006¹²、Baker 2009、Freidin 2009和Rizzi 2009a,b) 中还被看作是普遍语法原则的重要一环, 我们在这里会对邻接原则进行更为详细的说明。

我们可以区分两种移位: 左向移位 (通常叫做提取) 和右向移位 (通常叫做外置)。所有的移位类型都构成长距离依存关系。在下面的章节中, 我将讨论外置的一些限制。提取将在随后的第 13.1.5.2 节详细讨论。

13.1.5.1 外置

Baltin (1981)和Chomsky (1986a: 40)声称, 例 (14) 中的外置关系从句必须要根据嵌套的 NP 进行解释, 即这些句子不匹配那些出现在 *t* 位置上的关系从句, 而是对应于出现在 *t'* 位置上的例子。

- (14) a. [_{NP} Many books [_{PP} with [stories *t*]] *t'*] were sold [that I wanted to
[_{NP} 许多书 [_{PP} PREP [故事 *t*]] *t'*] AUX 卖 [CONJ] 我想 INF
read].
读]
'有许多我想读的带故事的书都被卖了。'
- b. [_{NP} Many proofs [_{PP} of [the theorem *t*]] *t'*] appeared [that I wanted
[_{NP} 许多证据 [_{PP} PREP [DET 定理 *t*]] *t'*] 出现 [CONJ] 我想
to think about].
INF 思考 PREP].
'有许多关于这个定理的证据显示是我想思考的。'

这里, 我们假定 NP、PP、VP 和 AP (至少在英语中) 都是右向移位的约束节点, 这里所指的释义就被邻接原则排除了 (Baltin 1981: 262)。

如果我们构建一个对应于 (14a) 的德语例子, 并且把嵌套的名词替换了, 这样它就被规则排除了, 或者不能作为所指, 这样我们就得到了 (15):

- (15) weil viele Schallplatten mit Geschichten verkauft wurden, die ich noch
因为许多 唱片 PREP 故事 卖 AUX CONJ 我 仍
lesen wollte
读 想
'因为很多带有我想读的故事的唱片被卖了'

这句话可以在这样的场景中使用, 其中某人在一个唱片店中看到有些特殊的唱片, 并且想起来他以前想读这些唱片上的童话故事。因为他并没有读这些唱片, 附加到上位名词是不可能的, 所以就附加到 *Geschichten* (故事) 上了。通过仔细地挑选名词, 就有可能构建出诸如 (16) 的例子了, 这些例子表示外置允许跨越多个 NP 节点:¹³

¹² 不过, 请参考Baltin (2004: 552)。

¹³ 请看Müller (1999a: 211)和 Müller (2004b; 2007e: Section 3)。相关的荷兰语的例子, 请看Koster (1978: 52)。

- (16) a. Karl hat mir [ein Bild [einer Frau $_{-i}$]] gegeben, [die schon lange tot ist]_i.
 Karl AUX 我 一 照片 一 女人 给 CONJ PRT 长 死 is
 ‘Karl 给了我一张死了一段时间的女人的照片。’
- b. Karl hat mir [eine Fälschung [des Bildes [einer Frau $_{-i}$]]] gegeben, [die schon lange tot ist]_i.
 Karl AUX 我 一 伪造 PREP.DET 照片 PREP. 一 女人
 给 CONJ PRT 长 死 is
 ‘Karl 给了我一张伪造的死了一段时间的女人的照片。’
- c. Karl hat mir [eine Kopie [einer Fälschung [des Bildes [einer Frau $_{-i}$]]]] gegeben, [die schon lange tot ist]_i.
 Karl AUX 我 一 复制 PREP. 一 伪造 PREP.DET 照片 PREP. 一 女人 给 CONJ PRT 长 死 is
 ‘Karl 给了我一张复制的伪造的死了一段时间的女人的照片。’

这种嵌套可以无限地继续下去，只要能找到语义上可行的嵌套的名词。NP 在德语中被看作是是约束节点 (Grewendorf 1988: 81、2002: 17–18、Haider 2001: 285)。这些例子说明右向的外置关系从句可以跨越任意数量的约束节点。

Koster (1978: 52–54) 讨论了例 (16) 中语言事实的一些可能的解释，其中关系从句被看作是移位到了 NP/PP 的边界，然后由此移向更远的位置上（这种移位需要所谓的逃跑机制或者逃跑路径）。他认为，这些方法也适用于那些被邻接规则排除在外的句子，也就是说，那些像 (14) 的例子。这就意味着，要么诸如 (14) 的事实可以根据邻接原则来解释，并且 (16) 中的句子是反例，要么是存在一个逃跑机制的，而且 (14) 中的例子是无关的，推导出的句子并不能用邻接原则进行解释。

在 (16) 的例句中，关系从句都被外置了。这些关系从句被看作是附加语，而且有人认为这些外置的附加语并不是移位而来的，而是在他们的位置上原位生成的，而且通过特殊的机制来表示共指关系 (Kiss 2005)。对于这类分析的支持者来说，(16) 中的例句对邻接原则来说是无效的，因为邻接原则只对移位进行限制。但是，外置可以跨越短语的界限这一点并不局限于关系从句；句子的补足语也可以被外置：

- (17) a. Ich habe [von [der Vermutung $_{-i}$]] gehört, [dass es Zahlen gibt, die die folgenden Bedingungen erfüllen]_i.
 我 AUX PREP DET 猜测 听说 CONJ EXPL 数字 给
 CONJ DET 下面的 要求 满足
 ‘我听说了一种猜测，有能够满足如下要求的数字。’
- b. Ich habe [von [einem Beweis [der Vermutung $_{-i}$]]] gehört, [dass
 我 AUX PREP ART 证据 PREP.DET 猜测 听说 CONJ

es Zahlen gibt, die die folgenden Bedingungen erfüllen]_i.

EXPL 数字 给 CONJ DET 下面的 要求 满足

‘我听说了猜测的证据，有能够满足如下要求的数字。’

- c. Ich habe [von [dem Versuch [eines Beweises [der Vermutung
我 AUX PREP DET 尝试 PREP.ART 证据 PREP.DET 猜测
_i]]]] gehört, [dass es Zahlen gibt, die die folgenden Bedingungen
听说 CONJ EXPL 数字 给 CONJ DET 下面的 要求
erfüllen]_i.
满足

‘我听说了试图证明猜测的事情，有能够满足如下要求的数字。’

由于有的名词选择带 *zu* 的不定式或介词短语，而且由于他们可以像上述句子一样被外置，就必须确保后置成分的句法范畴对应于名词所需要的范畴。这就意味着，管辖的名词和外置的成分之间必须存在某种关系。据此，(17) 中的例句必须被分析为外置的例子，并且针对上述讨论的观点提供相反的证据。

如果想要讨论循环嵌套的可能性，那么我们就必须想到例 (16) 中那些跨越几组句子的可能性，而例 (17) 是非常远的。但是，我们有可能找到深度嵌套的个例：(18) 给出了 Tiger 语料库中找到的关系从句外置和补足语外置的一些例子¹⁴ (Müller 2007e: 78–79, Meurers & Müller 2009: Section 2.1)。

- (18) a. Der 43jährige will nach eigener Darstellung damit [NP den Weg
DET 43 岁 想要 PREP 自己的 描述 那儿.PREP DET 道路
[PP für [NP eine Diskussion [PP über [NP den künftigen Kurs [NP der
PREP ART 讨论 关于 DET 未来 路线 PREP.DET
stärksten Oppositionsgruppierung]]]]] freimachen, [die aber mit 10,4
强 反对. 派 自由. 做 CONJ 但是 PREP 10.4
Prozent der Stimmen bei der Wahl im Oktober weit hinter
百分点 PREP.DET 选票 PREP DET 选举 PREP 十月 远 PREP
den Erwartungen zurückgeblieben war]. (s27639)
DET 预期 停留. 回 AUX

‘按照他自己的话来说，这位 43 岁的人想要为最强的反对派的有关未来路线的讨论扫清道路，但是，他的表现低于预期，在十月的选举中只得到了 10.4 个百分点。’

- b. [...] die Erfindung der Guillotine könnte [NP die Folge [NP
DET 发明 PREP.DET 断头台 AUX DET 结果
eines verzweifelten Versuches des gleichnamigen Doktors] gewesen
PREP.ART 绝望 尝试 DET 相同. 名字 医生 AUX
sein, [seine Patienten ein für allemal von Kopfschmerzen infolge
AUX 他的 病人 一次 PREP 永远 PREP 头疼 由于
schlechter Kissen zu befreien]. (s16977)
坏 枕头 to 解放

‘断头台的发明可能是由于与之同名的医生想要永远解除那些由于差枕

¹⁴ 有关 Tiger 语料库的更多信息请参考 Brants et al. (2004)。

头而导致头疼的病人的痛苦所做一种绝望的尝试。’

我们也可以造出违反邻接条件的英语句子。Uszkoreit (1990: 2333)举出了下面的例子:

- (19) [NP Only letters [PP from [NP those people _{-i}]]] remained unanswered [that [NP 只有 信 [PP PREP [NP 那些人 _{-i}]]] 保持 未回答的 [CONJ had received our earlier reply]_i.
AUX 收到 我们的早些 回复]_i.
‘只有那些一直没有回答的人收到了我们的回信。’

Jan Strunk (p.c. 2008) 找到了外置的例子, 既包括跨越多重短语界限的限制性关系从句, 也包括跨越多重短语界限的非限制性关系从句:

- (20) a. For example, we understand that Ariva buses have won [NP a PREP 例子, 我们理解 CONJ Ariva 巴士 AUX 赢 [NP ART number [PP of [NP contracts [PP for [NP routes in London _{-i}]]]] 数字 [PP PREP [NP 合同 [PP PREP [NP 路线 PREP 伦敦 _{-i}]]]] recently, [which will not be run by low floor accessible buses]_i.
最近, [CONJ AUX 不 AUX 运营PREP 低 层 可及的 巴士]_i.
‘例如, 我们知道 Ariva 巴士最近已经取得了一些伦敦路线的合同, 这些路线不会由低层的巴士运营了。’¹⁵
- b. I picked up [NP a copy of [NP a book _{-i}]] today, by a 我拾 起 [NP ART 复印本PREP [NP ART 书 _{-i}]] 今天, PREP ART law professor, about law, [that is not assigned or in any way 法律教授, 关于 法律, [CONJ AUX 不 指派 或者PREP 任何方式 required to read]_i.
要求 INF 阅读]_i.
‘我拾起了一本今天法律教授讲的有关法律的一本书的复印本, 不过这本书并没有被要求阅读。’¹⁶
- c. We drafted [NP a list of [NP basic demands _{-i}]] that night [that 我们起草 [NP ART 列表PREP [NP 基本 需求 _{-i}]] 那 晚上 [CONJ had to be unconditionally met or we would stop making and AUX INF AUX 无条件地 满足 或者 我们可以 停止制作 和 delivering pizza and go on strike]_i.
递送 披萨 和 走 PREP 罢工]_i.
‘我们那晚上起草了一个基本需求的清单, 这些需求必须无条件地满足, 不然我们就不再做披萨和运送披萨, 而是继续罢工。’¹⁷

¹⁵ <http://www.publications.parliament.uk/pa/cm199899/cmselect/cmenvtra/32ii/32115.htm>, 2007 年 2 月 24 日.

¹⁶ <http://greyhame.org/archives/date/2005/09/>, 2008 年 9 月 27 日.

¹⁷ <http://portland.indymedia.org/en/2005/07/321809.shtml>, 2008 年 9 月 27 日.

(20a) 也发表在Strunk & Snider (2013: 111)上。我们可以在这篇论文中找到更多的德语和英语的例子。

上述分析表明，右向的邻接限制并不适用于英语或德语，由此不能被看作是具有语言普遍性的。我们只能简单地说 NP 和 PP 在英语或德语中不是约束节点。那么，这些外置的事实对于认定邻接原则的理论来说就不存在问题了。不过，邻接限制也被看作是左向移位。我们将在下一节来详细讨论。

13.1.5.2 提取

在特定条件下，有些成分是不能左向移位的 (Ross 1967)。这些成分被称为提取的孤岛。Ross (1967: Section 4.1)构建了复杂 NP 限制 (CNPC)，该限制表明复杂名词短语是无法提取的。比如说在一个名词短语内部的关系小句的提取就是不可行的：

- (21) *Who_i did he just read [NP the report [S that was about _{-i}]]?
谁_i AUX 他刚读 [NP DET 报告 [S CONJ AUX 关于 _{-i}]]

尽管 (21) 在语义上是行得通的，这个句子仍是不合语法的。因为这里的疑问代词超越了关系小句的界限，也超越了 NP 的界限，由此它跨越了两个约束节点。有人认为，复杂 NP 限制适用于所有语言。但是，事实并非如此，因为在丹麦语 (Erteschik-Shir & Lappin 1979: 55)、挪威语、瑞典语、日语、韩语、泰米尔语和库阿语中是可以的（请看Hawkins (1999: 245, 262) 和其中的参考文献）。由于复杂 NP 限制这个条件被整合进邻接原则中，相应地邻接原则不能被普遍地适用，除非有人认为在分析的语言中，NP 不是一个约束节点。然而，确实是大部分语言不允许复杂名词短语的提取。Hawkins 基于处理难度来分析与之相关的结构（第 4.1 节）。他解释了允许这类提取的语言和不允许这类提取的语言之间的区别，这与各自语言中结构的不同处理容量是有关的，这些结构来自于与其他语法属性的提取的互动，如动词位置和其他约定的语法结构（第 4.2 节）。

与复杂名词短语的提取不同的是，跨越单句界限的提取 (22) 没有被邻接原则排除。

- (22) Who_i did she think that he saw _{-i}?
谁_i AUX 她认为 CONJ 他看见_{-i}
‘她认为他看到的是谁？’

正如在前面章节提到的，跨越多重句子边界的移位，在转换理论中被叫做循环移位：疑问代词移到限定语位置，然后移到下一个最高的限定语上。这些移位步骤的每一步都受到邻接原则的限制。邻接原则一举排除了长距离移位。

邻接原则无法解释为什么嵌套在言说动词 (23a) 或叙实动词 (23b) 下的句子的提取是异常的 (Erteschik-Shir & Lappin 1979: 68–69)。

- (23) a. ?? Who_i did she mumble that he saw _{-i}?
谁_i AUX 她咕哝 CONJ 他看见_{-i}
b. ?? Who_i did she realize that he saw _{-i}?
谁_i AUX 她意识到 CONJ 他看见_{-i}

这些句子的结构看起来跟 (22) 是一样的。在完全句法理论中, 还试图将这些差异解释为邻接违反或对 Ross 限制的违反。由此, (Stowell 1981: 401–402) 认为, (23) 中的句子与 (22) 中的句子具有不同结构。Stowell 将这些表示言说动词的方式的句子论元看作是状语。由于状语从句对提取来说是孤岛条件, 这就可以解释为什么 (23a) 是标记的了。状语的分析也跟这些句子论元是可以省略的事实是一致的。

- (24) a. She shouted that he left.
她 大喊 CONJ 他 离开
‘她大喊着他离开了。’
b. She shouted.
她 大喊
‘她大喊大叫。’

Ambridge & Goldberg (2008: 352) 指出, 将这些从句看作是状语的分析是不合理的, 因为他们只跟非常有限的动词类别有关, 即言说类动词和认为类动词。这一属性是论元的属性, 不是状语的属性。像地点修饰语这类状语可以适用于很多动词类。而且, 如果 (24b) 中的句子论元被省略, 它的含义也变了: 而 (24a) 要求交流一些信息, 而 (24b) 并不需要。也可以像 (25) 那样用 NP 来替代句子论元, 这些情况一定不希望被处理为状语。

- (25) She shouted the remark/the question/something I could not understand.
她 大喊 DET 话/DET 问题/某事 我AUX 不 理解
‘她大喊着我听不懂的话问题/某事’

将这些句子论元归类为状语不能扩展到叙实动词上, 因为他们的句子论元不是可选的 (Ambridge & Goldberg 2008: 352):

- (26) a. She realized that he left.
她 意识到 CONJ 他 离开
b. ?? She realized.
她 意识到

Kiparsky & Kiparsky (1970) 针对叙实动词提出了一个带有多词性中心语的复杂名词短语。一个可选的“事实”删除-转换去掉了诸如 (27a) 句中的中心语名词和限定语, 从而推导出诸如 (27b) 的句子 (第 159 页)。

- (27) a. She realized [_{NP} the fact [_S that he left]].
她 意识到 [_{NP} DET 事实 [_S CONJ 他 离开]]
‘她意识到他离开了的事实。’
b. She realized [_{NP} [_S that he left]].
她 意识到 [_{NP} [_S CONJ 他 离开]]
‘她意识到他离开了。’

从这类句子中提取的不可能性可以通过假定两个约束节点有交叉来解释，这种交叉被认为是不可可能的（有关结构的孤岛地位，请看Kiparsky & Kiparsky 1970: Section 4）。这个分析预测了从叙实动词的补足语从句中提取就跟在明显的NP论元中提取一样是不可行的，因为这两个结构是一样的。但是，根据Ambridge & Goldberg (2008: 353)，事实并不是如此：

- (28) a. * Who did she realize the fact that he saw _i?
 谁 AUX 她 意识到DET 事实CONJ 他 看见
 b. ?? Who did she realize that he saw _i?
 谁 AUX 她 意识到CONJ 他 看见

跟Erteschik-Shir (1981)、Erteschik-Shir & Lappin (1979)、Takami (1988) 和Van Valin (1998)一样，Goldberg (2006: Section 7.2)认为，空位一定属于可以潜在构成话语的焦点的那部分（请看Cook (2001)、De Kuthy (2002)，以及德语的请看 Fanselow (2003c)）。这就意味着这个部分一定不能被预设。¹⁸ 如果我们考虑这对邻接原则的事实意味着什么，那么我们会注意到每个例子中，提取都发生在预设的材料中：

- (29) a. 复杂 NP
 She didn't see the report that was about him. → The report was
 她 AUX. 不看见DET 报告 CONJ AUX 关于 他 → DET 报告 AUX
 about him.
 关于 他
 ‘她没看见关于他的报告。→ 报告是关于他的。’
 b. 认为动词或言说动词的补足语
 She didn't whisper that he left. → He left.
 她 AUX. 不低声说 CONJ 他 离开→ 他 离开
 ‘她没有低声说他离开了。→ 他离开了。’
 c. 叙实动词
 She didn't realize that he left. → He left.
 她 AUX. 不意识到CONJ 他 离开→ 他 离开
 ‘她没意识到他离开了。→ 他离开了。’

Goldberg 指出，属于背景信息的短语成分是孤岛 (*Backgrounded constructions are islands* (BCI))。Ambridge & Goldberg (2008)对这个语义/语用分析进行了实验测试，

¹⁸ 不管整句话是不是否定的，预设的为真仍为真。所以说，下面 (i.a) 和 (i.b) 都表明有一位法国国王。

- (i) a. The King of France is bald.
 DET 国王 PREP 法国 COP 秃的
 ‘法国国王是秃子。’
 b. The King of France is not bald.
 DET 国王 PREP 法国 COP 不 秃的
 ‘法国国王不是秃子。’

并且将之与纯粹的句法方法进行了比较。他们能够证明信息结构属性对于成分的提取能力发挥了重要的作用。跟Erteschik-Shir (1973: Section 3.H)一起, Ambridge & Goldberg (2008: 375)认为, 不同语言为了排除提取, 在有关多少成分必须属于背景信息方面是有差异的。任何情况下, 我们不应该认为所有语言都应排除状语的提取, 因为像丹麦语这些语言是可以从关系从句中进行提取的。¹⁹ Erteschik-Shir (1973: 61)举出了如下的例子:

- (30) a. Det_i er der mange [der kan lide _{-i}].
CONJ COP 那 许多 CONJ 可以喜欢
'有许多喜欢那个的人。'(lit.: '那个, 有许多人喜欢。')
- b. Det hus_i kender jeg en mand [som har købt _{-i}].
CONJ 房子认识 我 ART 男人 CONJ AUX 买
'我认识一个买了那栋房子的人。'(lit.: '这个房子, 我认识一个买了它的人。')

Rizzi 提出的有关邻接限制的参数化在许多作品中被摒弃了, 而且相关的效应被归为语法其他方面的差异 (Adams 1984; Chung & McCloskey 1983; Grimshaw 1986; Kluender 1992)。

在本小节, 我们看到有一些不属于结构的句法属性因素, 导致了左向移位可能受到了限制。除了信息结构的属性, 语言处理方面的因素也起到了重要的作用 (Grosu 1973; Ellefson & Christiansen 2000; Gibson 1998; Kluender & Kutas 1993; Hawkins 1999; Sag, Hofmeister & Snider 2007)。所包括成分的长度, 填充语和空位之间的距离、句法结构的复杂性, 以及在填充语和空位置之间的空间内相似的话语所指之间的相互效应都是话语的可接受度的重要因素。由于不同语言在他们的句法结构方面是不同的, 自然会有语言运用的不同效应, 比如那些在外置和提取方面的差异。

总而言之, 我们可以说邻接限制不能说明德语和英语中的外置现象, 而且我们可以比邻接原则更好地解释与信息结构相关的提取限制, 以及语言处理现象。由此, 假定邻接原则作为普遍的语法能力中的句法限制来解释语言事实就是不必要的了。

13.1.6 表示时态、情态和体的语法语素

Pinker (1994: 238)在对许多语言中表示时态、情态、体、格和否定的语素方面的观点是正确的。但是, 在关于一种语言中包括哪些语法属性以及他们是如何表达的方面是有着极大的差异的。

时态系统方面的差异的例子请看Dahl & Velupillai (2013b,a)。现代汉语显然是一个例子: 它几乎没有形态。在几乎每一种语言中, 相同的语素可以按照一种或

¹⁹ 在讨论基于普遍语法的方法是否可被检验的问题时, Crain, Khlentzos & Thornton (2010: 2669)声称是不可能从关系从句中进行提取的, 而且这些语言的存在需要质问普遍语法这个概念。("如果儿童习得任何语言中, 学会从关系从句中提取语言表达式, 那么这就会对普遍语法的基本原则提出严重的质疑。")由此, 他们反驳了 Evans 和 Levinson, 以及 Tomasello, 这些人认为普遍语法方法是不可被检验的。如果 Crain, Khlentzos 和 Thornton 的观点是正确的, 那么 (30) 就会证明普遍语法是假的, 这样讨论就结束了。

另一种形式的事实可以归结为某些事物需要被反复表达的事实，然后那些经常重复出现的事物就被语法化了。

13.1.7 词类

在第 4.6 节，我们提到了所谓的制图方法，其中有些人提出了超过三十种功能范畴（请看第 133 页有关 Cinque 的功能性中心语的表格 4.1），并且认为这些范畴与相应固定的句法结构一起构成了普遍语法的一部分。Cinque & Rizzi (2010: 55, 57) 甚至提出了超过 400 种被认为在所有语言的语法中起到了重要作用的功能范畴。²⁰再有，像 Infl（屈折变化）和 Comp（补足语）的特殊词类在构建所谓的普遍性的原则时也被提及了（Baltin 1981: 262、2006、Rizzi 1982、Chomsky 1986a: 38、Hornstein 2013: 397）。

Chomsky (1988: 68、1991、1995b: 131)、Pinker (1994: 284, 286)、Briscoe (2000: 270) 和 Wunderlich (2004: 621) 关于词类的天赋机制做了相对较少的假设：Chomsky 认为所有的词汇范畴（动词、名词、形容词和介词）都属于普遍语法，而语言具有处理他们的权利。Pinker、Briscoe 和 Wunderlich 认为所有的语言都有名词和动词。再者，对普遍语法的批评提出了这样的疑问，这些句法范畴能否在其他语言中以我们所熟知的语言（如德语和英语）的方式来找到。

Braine (1987: 72) 认为，诸如动词和名词的词类应该被看作是从像论元和谓词这样的基础概念推导而来的（请看 Wunderlich (2008: 257)）。这就意味着对于这些范畴的存在有一个独立的解释，他们并不是基于语言特有的天赋知识的。

Evans & Levinson (2009a: Section 2.2.4) 讨论了类型学的文献，并且举出了缺少副词和形容词的语言的例子。作者引用了 Straits Salish 作为一种在动词和名词之间没有区别的语言的例子（请看 Evans & Levinson 2009b: 481）。他们评论说，对于经常使用的排在前四位或前五位的非印欧语系的语言来说，是有必要增加额外的词类来说明状貌词、方位词、副动词和量词的。²¹如果我们认为语言可以从可能性的集合（工具包）中进行选择且不会穷尽的话，这样的情况就不是问题了（Jackendoff 2002: 263、Newmeyer 2005: 11、Fitch, Hauser & Chomsky 2005: 204、Chomsky 2007: 6–7、Cinque & Rizzi 2010: 55, 58, 65）。但是，如果我们容忍这个观点，那么必然有些武断。我们可以假设任何一个需要对至少一个语言进行分析的词类，将他们归为普遍语言，并且认为大部分（也许甚至是所有）语言并没有利用所有的词类。Villavicencio (2002: 157) 就是这样认为的，他在范畴语法的框架下提出了范畴 S、NP、N、PP 和 PRT。这种假设是不可证伪的（请看 Evans & Levinson 2009a: 436、Tomasello 2009: 471 关于相似情况的讨论以及更为普遍性的讨论）。

而 Evans 和 Levinson 认为，我们需要其他范畴，Haspelmath (2009: 458) 和 Croft (2009: 453) 进而否定了跨语言共性的词类的存在。我认为这过于极端，我认为更好的搜索策略是试图找到不同语言的共同点。²²然后，我们应该期待能够找到不能套进我们带有印欧语系偏见的语法概念。

²⁰ 问题是这些范畴到底是否属于普遍语法是不确定的。

²¹ 相反的观点，请看 Jackendoff & Pinker (2009: 465)。

²² 对比 Chomsky (1999: 2): “在缺乏相反的强有力的证据下，假定语言是统一的，带有简单可预测的话语属性的不同限制。”

13.1.8 递归与无限

在《科学》杂志的一篇文章上, Hauser, Chomsky & Fitch (2002)进一步阐明了只有具体领域的普遍性是可递归的假设,“提供一个有限成分的集合中生成无限的表达的能力”(请看第 60 页关于递归的短语结构规则的例子 (37b))。²³这个假设是具有争议的,而且既有形式化的,也有事实上的反对理由。

13.1.8.1 形式化的问题

语言能力是无限的这个观点是广为流传的,而且早在 Humboldt 的论著中就能够找到相关的说法:²⁴

Das Verfahren der Sprache ist aber nicht bloß ein solches, wodurch eine einzelne Erscheinung zustande kommt; es muss derselben zugleich die Möglichkeit eröffnen, eine unbestimmbare Menge solcher Erscheinungen und unter allen, ihr von dem Gedanken gestellten Bedingungen hervorzubringen. Denn sie steht ganz eigentlich einem unendlichen und wahrhaft grenzenlosen Gebiete, dem Inbegriff alles Denkbaren gegenüber. Sie muss daher von endlichen Mitteln einen unendlichen Gebrauch machen, und vermag dies durch die Identität der gedanken- und spracherzeugenden Kraft. (Humboldt 1988: 108)

如果我们只看语言事实,我们能看到语句的长度是有上限的。这是因为特别长的语句无法处理的事实,说话者需要睡觉或者最终会在某个时刻死去。如果我们设定一个长度为 100,000 语素的句子,然后假设一个语素为 X 的子集,我们可以构造出少于 $X^{100,000}$ 个表达式。如果我们在每个 100,000 位置上使用每个语素的话,就会得到 $X^{100,000}$ 这个数。由于不是所有这些表达式都是合乎语法的,那么实际上是少于 $X^{100,000}$ 个可能的表达式的(请看 Weydt 1972 中类似的但是更为具体的说明)。这个数已经够大的了,但是仍是有限的。对于思维来说也是一样的:我们不需要无限多的思想(如果“无限”这个概念是按照这个词的数学含义来理解的话),尽管 Humboldt 和 Chomsky (2008: 137)的观点是相反的。²⁵

²³ 在 *Cognition* 中的一篇讨论性文章中, Fitch, Hauser & Chomsky (2005)澄清了他们的观点,递归是唯一的针对具体语言和具体人类的属性,这是一个假设,而且可以根本就没有语言特有/种族特有的属性。然后,能力与属性的具体组合应该是人类特有的(第 182–201 页)。他们提出的另一个观点是,天赋的语言特有的知识具有对应于早期主流的生成语法所谓的复杂性(第 182 页)。Chomsky (2007: 7)指出递归并可以是非具体语言的操作,但是仍将之归为普遍语法。

²⁴ 语言的生成并不是简单地创造一个话语的过程;与之同时,它还允许这类表达的无限集合,而且尤其重要的是允许思想赋予表达的条件。语言面对着一个无限且真正没有边界的事实,一个人们可能想象的所有事情的缩影。由此,它必须能够利用有限的手段来表达无限的内容,而且这是可以通过负责思维和语言的生成的能力所认证的。

²⁵ Weydt (1972)讨论了 Chomsky 关于无限多的句子的可能性的观点,以及是否将 Chomsky 与 Humboldt 相提并论是合理的。Chomsky 在《当下语言学理论的问题》(Chomsky 1964a: 17)的引述中遗漏了“Denn sie steht ganz eigentlich einem unendlichen und wahrhaft grenzenlosen Gebiete, dem Inbegriff alles Denkbaren gegenüber.”这句话。Weydt (1972: 266)认为, Humboldt、Bühler 和 Martinet 声称有无限的思想可以被表达。Weydt 认为这并不意味着句子必须要特别长。相反,他认为文本的长度没有上限。这个论述是很有意思的,但我猜想文本只不过是更大的单位,而 Weydt 针对语言对句子长度没有上限的说法也同样适用于文本。文本可以通过 (i) 中相当简化的规则生成,它将一个句子 U 和一个文本 T 组合成更大的文本 T:

在文献中，我们有时可以发现我们可以创造出无限长的句子的说法（请看Nowak, Komarova & Niyogi (2001: 117)和Kim & Sells (2008: 3)，以及O'Neill & Wood (2012)中的Dan Everett）。当然不是这样的。我们在第2章提到的重写语法也不是这样的，这些语法允许无限句子的生成，因为规则右手边的符号集合被界定为有限的。但是，我们有可能生成出无限数量的句子，句子本身不能是无限的，因为总是一个符号被无限地许多其他符号所替代，由此会得到非有限符号的序列。

Chomsky (1965: Section I.1)在de Saussure (1916b)影响下，区分语言能力和语言运用：语言能力是有关哪种语言结构是合乎语法知识，而语言运用是对这个知识的应用（请看第12.6.3节和第15章）。我们有限的大脑容量以及其他限制导致我们无法处理任意数量的嵌套，而且我们无法造出超过100,000个语素长度的句子。语言能力与语言运用之间的区分是有意义的，并且允许我们构造出对诸如(31)的句子的分析的规则：

- (31) a. Richard is sleeping.
Richard AUX 睡觉
'Richard 在睡觉。'
- b. Karl suspects that Richard is sleeping.
Karl 怀疑 CONJ Richard AUX 睡觉
'Karl 怀疑 Richard 在睡觉。'
- c. Otto claims that Karl suspects that Richard is sleeping.
Otto 声称 CONJ Karl 怀疑 CONJ Richard AUX 睡觉
'Otto 声称 Karl 怀疑 Richard 在睡觉。'
- d. Julius believes that Otto claims that Karl suspects that Richard is sleeping.
Julius 认为 CONJ Otto 声称 CONJ Karl 怀疑 CONJ Richard AUX 睡觉
'Julius 认为 Otto 声称 Karl 怀疑 Richard 在睡觉。'
- e. Max knows that Julius believes that Otto claims that Karl suspects
Max 知道 CONJ Julius 认为 CONJ Otto 声称 CONJ Karl 怀疑

(i) $T \rightarrow TU$

U可以是一个文本的一个句子或者一个短语。如果我们准备好承认文本的长度没有上限的话，这就意味着句子的长度也没有上限，因为我们可以通过“和”来将文本中所有的短语连接起来构成长句。这样的长句是将短句连接在一起的产物，他们跟乔姆斯基观点下所允准的非常长的句子在本质上是不同的，因为他们不包括一个任意深度（请看第15节）的自我中心的嵌套，但是无论如何从任意长度的文本造出的句子的数量是无限的。

对于任意长的文本来说，有一个有趣的问题：让我们来设想一个人造句并将这些句子加进一个现有的文本中。当这个人死了的时候，这项工作就被打断了。我们可以说另一个人可以接下这个文本，直到这个人去世，再这样继续下去。同样，问题是人们能否理解这个有几千万页数的文本的意义和结构。如果这不足以成为一个问题，人们可以自问一直加句子到文本中直到2731年的这个人还是不是在2015年开始这个文本的同一个人。如果这个问题的答案是否定的，那么这个文本就不是由一个语言L的句子构成的，而是由许多种语言构成的，这样就与争论的焦点无关了。

that Richard is sleeping.

CONJ Richard AUX 睡觉

‘Max 知道 Julius 认为 Otto 声称 Karl 怀疑 Richard 在睡觉。’

规则采用如下的形式：将一个名词短语跟某个类别的动词和一个从句相组合。通过连续地应用这条规则，就可以造出任意长度的字符串。Pullum & Scholz (2010)指出，我们必须把两件事区分开：语言是否是递归系统的问题，还是我们能够设计的一种具体语言的最佳模型碰巧是递归的。更多有关这个方面和大脑处理的内容，请看Luuk & Luuk (2011)。当我们应用上面的系统来构造字符串时，不能说明（一个特别的）语言是无限的，即使通常是这样认为的（Bierwisch 1966: 105–106、Pinker 1994: 86、Hauser, Chomsky & Fitch 2002: 1571; Müller 2007b: 1、Hornstein, Nunes & Grohmann 2005: 7、Kim & Sells 2008: 3）。

这个语言的无限性的“证据”作为一个对应于没有最大自然数的间接证据（Bierwisch 1966: 105–106、Pinker 1994: 86）。在自然数方面，是这样运作的：假设 x 是最大自然数。然后构造 $x + 1$ ，因为这在定义上是一个自然数，我们现在找到一个比 x 更大的自然数。由此，我们说明了 x 是最大的数的假设导致了一个矛盾，所以说，就不可能有最大自然数这个说法。

当我们把这个证据转化到自然语言领域，出现的问题是我们能否仍就希望有一个 1,000,000,000 词的字符串作为我们想要描述的语言的一部分。如果我们还想这样做，那么这个证据就不能用了。

如果我们把语言看作是一个生物构造，那么我们就需要接受它是有限的事实。否则，我们必须假定它是无限的，但是生物学上真实的对象的一个无限大的部分在生物学上是不真实的 (Postal 2009: 111)。Luuk & Luuk (2011)将语言看作是物理学上不可数但是是字符串的有限集合。他们指出一定要做出这样的区分，即想要无限地扩展句子的能力，还是采取一个不可数的字符串的集合然后真正地将其扩展的能力。我们具有第一种能力，不具有第二种。

对于语言的无限性提供论据的一个可能性是认为只有生成语言适合于模拟语言，而且我们需要递归规则来捕捉语言事实，这就是为什么心智表达式具有递归机制，从而生成无限数量的表达式 (Chomsky, 1956: 115、2002: 86–87)，然后这就暗示了语言包括无限多的表达式。Pullum & Scholz (2010)指出，这个论据有两个错误Pullum & Scholz (2010)：即使我们认同生成语法，仍有可能是上下文相关语法，即使带有递归规则也只能生成有限的集合。Pullum & Scholz (2010: 120–121)引述了András Kornai举出的一个有趣的例子。

更为重要的问题是，没有必要假设语言生成集合。有三个明显的形式化的方案，但是只有第三种在这里提到了，即模型理论，也是基于约束的 approaches（请看第 14 章）。Johnson & Postal 的弧对语法 (1980)、根据Kaplan (1995)形式化的 LFG、对Rogers (1997)重新形式化的 GPSG，以及根据King (1999)假设的 HPSG，Pollard (1999)和Richter (2007)是模型理论的方法。在基于约束的理论中，我们可以分析像 (31) 的例子，并且说特定的态度动词选择一个名词性 NP 和一个 *that* 从句，而且这些只能在特定的局部句法配置中出现，他们具有所含的成分之间的特殊关系。这些关系中的一个就是主谓一致。按照这个方式，我们可以表示诸如 (31) 的表达式，并且不需要说明许多句子是图和嵌套的。这就意味着基于约束的理论是跟是否有有限还是无限数量的结构的问题的两个答案是兼容的。按照这个方式来使用匹配的语法，有可能开发出语言使用的模型，它能够解释为什么有些字符串，比如说特

别长的那些，是不可接受的（请看第 15 章）。

13.1.8.2 语言事实的问题

有时认为所有的自然语言都是递归的，而且在所有语言中一个任意长度的句子是可能的（概述请看Hornstein, Nunes & Grohmann 2005: 7，更多内容请看Pullum & Scholz (2010: Section 2)）。当有人提到递归的时候，通常是指带有自我嵌套的结构，正如我们在 (31) 的分析中所看到的 (Fitch 2010)。但是，有可能有的语言不允许自我嵌套。Everett (2005) 认为 Pirahã 就是这样一种语言（但是，请看Nevins, Pesetsky & Rodrigues (2009) 和 Everett (2009)）。没有递归的语言的另一个例子是 Warlpiri 语，它通常跟Hale (1976) 一起引用。不过，Hale 的规则针对带关系从句的句子的组合是递归的（第 85 页）。在第 98 页递归被明确地表示出来。²⁶ Pullum & Scholz (2010: 131) 讨论了 Hixkaryana 语，这是属于加勒比语族的亚马逊语，它与 Pirahã 语没有关系。这个语言有嵌套，但是嵌套的材料与其主句具有不同的形式。这些嵌套不能带有非限定性的特征。在 Hixkaryana 语中，也没有并列短语和从句 (Derbyshire (1979: 45) 通过 Pullum & Scholz (2010: 131) 这样引述)，这就是为什么这种语言中无法产生嵌套的句子。其他没有自嵌套的语言有 Akkadian 语、Dyirbal 语和原始-Uralic 语。

当然，说所有语言都是递归的这个观点是有一定意义的：他们遵守这样的规则，一个具体数量的符号可以被组合以构成另一个符号。²⁷

(32) $X \rightarrow X \dots X$

据此，所有的自然语言都是递归的，而且简单符号的组合以构成更为复杂的符号是语言的一个基本属性 (Hockett 1960: 6)。关于 Pirahã 语的争论的事实是如此的尖锐，以至于可以说明这不是所谓的那种递归。同样，请看Fitch (2010)。

同样认为，范畴语法的组合性规则也具有语言的普遍性。我们可以利用这些规则来将函项跟它的论元相组合 ($X/Y * Y = X$)。这些规则跟 (32) 的规则几乎一样抽象。区别是其中一个成分必须是函项。在最简方案中也有对应的限制，比如说选择性特征（请看第 4.6.4 节）和语义角色指派的限制。但是，是否是范畴语法允许了递归结构并不取决于普遍的组合模式，而是词汇项。使用 (33) 中的词汇项，有可能分析两个句子，并且一定不会构建递归的结构。

- (33) a. the: np/n
b. woman: n
c. cat: n
d. sees: (s\ np)/np

²⁶ 但是，他在第 78 页说明关系从句通过逗号跟包含中心语名词的句子区分开。Warlpiri 语中的关系从句总是外围的，也就是说，他们出现在他们所指的名词的句子的左边或右边。相似结构可以在德语中找到：

(i) Es war einmal ein Mann. Der hatte sieben Söhne.
那 AUX 从前 ART 男人 他 有 七 儿子
‘从前有个男人。他有七个儿子。’

也有可能是我们在文本层处理句子的连接，而在句子层没有递归。

²⁷ Chomsky (2005: 11) 认为归并组合了 n 个对象。一个特殊的例子是二元归并。

13 语言知识的天赋性

如果我们扩展词汇来囊括范畴 n/n 的修饰语，范畴的连词 (X/X) $/X$ ，那么我们就得到一个递归语法。

Fitch, Hauser & Chomsky (2005: 203)指出，不允许递归结构的语言事实对于基于普遍语法的理论来说并不是一个问题，因为并不是普遍语法的所有可能性都能够通过具体的语言来应用。按照这个观点，我们实际上跟词类一样面临同样的情况（请看第 13.1.7 节），你可以提出任意数量属于普遍语法的属性，然后再根据他们是否发挥重要作用的语言基础来决定具体语言的情况。这个方法的一个极端的变体就是所有语言的语法变成了普遍语法中的一部分（也许带有诸如 NP_{Spanish} 、 NP_{German} 的不同符号）。这个针对语言的有关人类能力的基于普遍语法的变体事实上是不可证伪的（Evans & Levinson 2009a: 436, 443、Tomasello 2009: 471）。

13.1.8.3 在认知的其他领域中的递归

在语言领域之外也有可以用递归规则描写的现象：Hauser, Chomsky & Fitch (2002: 1571)提到了导航、家族关系和计数系统。²⁸ 我们也可以认为，相关的能力是后来习得的，更高水平的数学属于个人成就，跟大多数人的认知能力是无关的，但是即使是 3 岁 9 个月的小孩儿都已经可以产出递归的结构了：2008 年，有篇新闻报道说从一架飞机上拍到了一个土生土长的巴西部落。我把这张照片给我儿子看，并告诉他土著美洲人用弓和箭射飞机。然后他问我那是哪种飞机。我告诉你你看不到，因为拍照片的人坐在飞机里。然后，他就说，如果你要拍下既有这架飞机，也有土著美洲人的照片的话，你就需要另一架飞机。他对他的主意感到很开心，然后他说“然后又有一架。然后又有一架。一架接着一架。”据此，他一定可以想象出嵌套的结果。

Culicover & Jackendoff (2005: 113–114)讨论了感知和音乐作为独立于语言的递归系统。Jackendoff (2011)将这个讨论扩展到视觉感知和音乐，并加上了程式域（以做咖啡为例）和无词的连环画。Chomsky (2007: 7–8)声称，视觉感知的例子是无关的，但是之后承认构建递归结构的能力可以属于更为普遍的认知能力（第 8 页）。他还是把这个能力归结为普遍语法。他把普遍语法看作是语言能力的一个子集，也就是说，作为语言需要的非特定域能力（广义的语言官能）和特定域能力（狭义的语言官能）的一个子集。

13.1.9 小结

综上所述，我们可以说，并没有一个用来解释语言普遍性的人们需要假定特定领域的天赋知识的一致性。在 2008 年召开的“德国语言学学会”会议，Wolfgang Klein 承诺支付 100 欧元，如果有人能够指出所有语言共享的不平凡的属性（请参考 Klein 2009）。这就涉及到对“平凡”的界定。看上去很清楚的是，所有语言都共享谓词论元结构和某些意义上（Hudson 2010a、Longobardi & Roberts 2010: 2701）的依存关系，而且，所有语言都有可以根据组合关系构成的复杂表达式（由于 Manfred

²⁸ 但是，Pinker & Jackendoff (2005: 230)指出，导航与 Chomsky 描述的递归系统不同，在所有的文化中递归不属于计数系统的一部分。他们认为，所有的文明都由于他们的语言能力而发展出了无限的计数系统。Fitch, Hauser & Chomsky (2005: 203)也这样认为。后者认为在其他领域中的递归的所有形式都取决于语言。关于这点的更多内容，请参考 Chomsky (2007: 7–8)。Luuk & Luuk (2011)指出，自然数被界定为递归的，但是数学定义并不必然对人类采用的数学运算发挥重要的作用。

Krifka提出了组合性，被允诺了 20 欧)。但是，正如在很多地方所提到的，普遍性绝不意味着天赋性 (Bates 1984: 189、Newmeyer 2005: 205)：Newmeyer 举例说，表示太阳和月亮的词几乎在所有的语言中都有。这跟天体在每个人的生活中起到的重要作用的事实有关系，所以人们需要用词来指称他们。不能由此来排除说相应的概念是天赋的。相似地，用来表示两个物体之间的关系的词 (如“接”) 必须按照明显的方式与描述这两个物体的词 (“我”、“大象”) 有联系。但是，这并不必然说明语言的这个属性是天赋的。

即使我们可以找到跟所有语言共享的结构属性，仍然没有天赋的语言知识的证据，因为这些相似性可以被追溯到其他因素上。有争论说，所有语言都必须按照这个方式以使得少量的有限资源能够被儿童习得 (Hurford 2002: Section 10.7.2、Behrens 2009: 433)。按照这一点，在它发展的相关阶段下，我们的大脑就是一个限制因素。语言需要适合我们的大脑，而因为我们的的大脑是相似的，语言也在某些方面是相似的 (请参考Kluender 1992: 251)。

13.2 语言习得的速度

通常认为，儿童学习语言异常快，而这只能是因为他们已经掌握了不需要习得的关于语言的知识 (如Chomsky 1976b: 144、Hornstein 2013: 395)。为了让这个论述更为严谨，它必须证明相对来说更为复杂的其他方面的知识需要更长的习得时间 (Sampson 1989: 214–218)。并没有看到相关的证据。语言习得跨越了多年的时间，而且不可能简单地证明语言是通过“简单的暴露”而习得的。Chomsky 将语言学跟物理学相比，并指出我们要习得物理学的知识更为困难。但是，Sampson (1989: 215)指出，人们在学校或大学习得的物理学知识并不能构成比较的基础，相反，人们应该考虑我们每天身处的物理世界的知识的习得。比如说，我们需要将液体倒进容器、从逃生绳逃走的那种知识，或者物体的发射轨道的属性的知识。对于语言习得进行说明，跟这些领域的知识相比较的复杂性远不能说是平凡的。对于这方面的深入讨论，请参考Sampson (1989: 214–218)。Müller & Riemer (1998: 1)指出，六岁的儿童可以理解 23,700 个词汇，并会运用 5000 个。据此，在四年半的时间里，他们平均每天学 14 个新词。这确实是令人惊叹的，但是不能作为天赋的语言知识的论据，因为所有习得的理论都认为词是通过语言事实学习的，而不是通过基因决定的普遍句法先天决定的。任何情况下，基因编码的假设对于新造词来说都是不太可能的，如 *fax* (传真)、*iPod*、*e-mail* (电子邮件)、*Tamagotchi* (电子鸡)。

进而，跟第二语言习得相比，第一语言习得是不费劲和迅速的这个说法是一个神话，正如Klein (1986: 9)的估算所显示的：如果我们假设儿童每天听五个小时的语言会话 (保守估计)，那么在他们生命中的头五年中，他们有 9100 小时的语言训练。但是到五岁的时候，他们仍没有习得所有的复杂结构。相比较而言，第二语言学习者，假设有必要的动机，可以在为期六周每天十二个小时 (总共 500 小时) 的集训课上学会一门语言的语法。

13.3 习得的关键期

鸭子有一个关键期，期间他们受到父母行为的显著影响。通常，小鸭子跟随他们的妈妈。但是，如果在这个特殊的时期，有一个人出现，而不是他们的妈妈，这些

鸭子就会跟着这个人。过了关键期后,这种对他们行为的影响就不存在了 (Lorenz 1970)。这种关键期在其他动物和其他认知领域中也可以被指出,例如灵长类动物的视觉能力的习得。特定的能力在特定的时间框架下习得,而相关输入的存在对决定这个时间范围的开始是重要的。

Lenneberg (1964)认为,语言习得只持续到 12 岁,并且从儿童可以比大人学语言学得更好这点得出结论,这也有赖于关键期,而且语言习得必须具有跟鸭子的印随行为相似的属性,所以说,语言习得的倾向必须是天赋的 (Lenneberg 1967: Chapter 4)。

对于关键期的长度的假说有相当大的分歧。有人认为是 5 年、6 年、12 年,甚至是 15 年 (Hakuta et al. 2003: 31)。另一个跟关键期相关的假说是假定习得语言的能力随着时间持续减弱。Johnson & Newport (1989)试图判断出第二语言习得的关键期,并且认为第二语言的学习从 15 岁开始变得尤为糟糕。但是,Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett (1996)指出,Johnson 和 Newport 的数据有一个不同的曲线,更适合于个人数据。另一个曲线表明学习语言的能力没有突然的变化,而是一个稳定的下降过程,由此没有证据证明关键期带来的影响。

Hakuta, Bialystok & Wiley (2003)评估了移民到美国的 2,016,317 位操西班牙语者和 324,444 位操汉语者所做的调查问卷的数据。他们调查了年龄、移民时间点、说话者的教育水平和他们习得的英语水平。他们无法确定出从何时开始语言习得受到了严重的限制。相反,随着年龄的增长,学习能力有着稳定的下降过程。这也可以在其他领域中被观察到:比如说,越晚学开车就越难。

由此可见,没有证据证明第二语言习得存在关键期。有时,有人认为第二语言习得不是由天赋的普遍语法驱动的,而是实际上是一个在关键期就已经接触知识的学习过程 (Lenneberg 1967: 176)。由此,我们可以说第一语言习得有关键期。但是,出于伦理学的原因,我们不能直接地用实验来控制输入量。我们不能说,找 20 个孩子,然后让他们在没有语言输入的环境中长到 3 岁、4 岁、5 岁、6 岁……或 15 岁,然后比较结果。这种研究有赖于非常少见的出于被忽视的案例。比如说,Curtiss (1977)研究了一个叫 Genie 的女孩儿。那时,Genie 13 岁,并且一直被囚禁地长大。她也被叫做野孩。正如 Curtiss 所介绍的,她没能再学会一定的语言规则。为了进行客观的比较,我们需要其他没有成长在完全封闭环境和在非人的环境中的测试数据。唯一的收集相关实验数据的可能性是研究到一定年龄没有接受过任何手语输入的聋儿。Johnson & Newport (1989: 63)用美国手语做了一些相关的试验。这里也表明了在学习能力方面有一个线性的下降过程,但是语言习得并没有从某个年龄就突然下降或者完全的能力丧失。

13.4 非人类的灵长动物的习得的缺乏

非人类的灵长动物不能学会自然语言的事实被当作是我们具有的语言能力的证据。所有的科学家都认同这样的事实,人类和灵长类动物之间的差异是由基因决定的,而且这跟语言能力是相关的。Friederici (2009)对文献进行了综述,并提出,在黑猩猩和猕猴(和儿童)中,大脑各部分之间的连接不如成年人的发达。大脑的相连区域一起负责词汇-语义知识的处理,并且可以构成语言发展的重要前提(第 179 页)。

但是,问题是,我们与其他灵长类动物在特殊认识能力上的区别是否只在于语

言,或者我们习得语言的能力是否取决于认知上一般范围上的区别。不过, Fanselow (1992b: Section 2) 认为人类特有的形式能力不必然是语言。相似地, Chomsky (2007: 7-8) 考虑归并(按照他的观点,这是唯一的构建结构的操作)是否属于语言特有的天赋能力,还是一般意义上人类特有的能力(但是,请看第 13.1.8 节,尤其是脚注 28)。

我们可以确知的是非人类灵长动物不懂得特殊的指示手势。人类喜欢模仿。其他灵长类动物也模仿,不过,不是出于社会因素 (Tomasello 2006b: 9-10)。根据 Tomasello et al. (2005: 676), 只有人类具有能力和动力按照共同的目标和社会合作的行动计划来施行合作的活动。灵长类动物懂得意向性的动作,但是,只有人类在大脑中带着共同的目标进行行动(共享意识 *shared intentionality*)。只有人类使用和懂得手语 (Tomasello et al. 2005: 685, 724, 726)。语言在更高的层面进行合作: 符号被用来指称物体,并且有时也指向说话者或听话者。为了能够使用这种交流系统,我们必须能够在交谈中设身处地为他人着想,并且发展出共同的期待和目标 (Tomasello et al. 2005: 683)。这样,非人类的灵长动物缺乏语言的社会和认知的前提条件,即人类和其他灵长类动物之间的区别不能用天赋的语言学知识来解释 (Tomasello 2003: Section 8.1.2; Tomasello et al. 2005)。

13.5 克里奥尔语和手语

当说话者没有一个共同语可以互相交流的话,他们就会发展出所谓的皮钦语。这些语言是使用所涵盖语言的词汇与非常基础的语法。需要指出的是,这些说皮钦语的子孙会对这些语言进行规范化。这些说话者的下一代就创造了一种带有独立语法的语言。这些语言被叫做克里奥尔语 (*creole languages*)。有一个假说是,从克里奥尔语发展出的语言形式受到天赋的普遍语法的限制 (Bickerton 1984b)。一般认为,克里奥尔语的参数设置对应于参数的默认值 (Bickerton 1984a: 217, 1984b: 178),也就是说,在出生时参数就有值,而且这些对应于克里奥尔语具有的值。这些默认的值在学习其他语言的时候需要进行修饰。²⁹ Bickerton 认为克里奥尔语包括语言学者无法从输入中习得的成分,即从皮钦语中获得。他的论断是经典的刺激贫乏论的变体,我们将在第 13.8 节具体讨论。

Bickerton 的观点倍受争议,因为它无法证实儿童是否从大人的个体语言中得到输入 (Samarin 1984: 207, Seuren 1984: 209)。对于这种证据的缺乏,能够说明的是有一些人口统计学上的事实显示,至少对一些克里奥尔语来说是这样的 (Arends 2008)。这意味着儿童不只有从皮钦语作为输入的字符串,还有从父母及周围其他人的个人语言中的句子的输入。许多说克里奥尔语的人认为大人对新涌现的语言贡献了具体的语法形式。例如,在夏威夷克里奥尔英语中,人们观察到,它受到相关的说话者的母语的影响: 操日语者既使用 SOV 语序,也使用 SVO 语序,而操菲律宾语者既使用 VOS 语序,也使用 SVO 语序。总之,在语言中有相当多的变体可以追溯到个别说话者的不同母语上。

我们也可以不用语言特有的天赋知识的假说来解释克里奥语化的效应: 儿童规范语言的事实可以归功于独立于语言的现象。在实验中,给参与者看两个灯泡,然后被试必须预测哪个灯泡接下来会亮。如果其中一个灯泡 70% 的时间都是亮的,

²⁹ 对于默认值假说带来的问题,请参考 Meisel (1995: 17)。Bickerton (1997: 56, fn. 13) 不同意克里奥尔语具有默认的参数值的说法。

参与者也会 70% 的时间也选这个（尽管他们实际上可以有更高的成功率，如果他们总是选择有 70% 的概率会亮的那个灯泡）。这个行为被叫做概率匹配（*Probability Matching*）。如果我们再加一个灯泡，然后把这个灯泡在 70% 的情况下点亮，这样，他们就在这个最为常见的情况的方向下进行了规范（Gardner 1957; Weir 1964）。

儿童比成人更多地进行规范化（Hudson & Newport 1999; Hudson Kam & Newport 2005），这个事实跟有限的脑容量有关（“更少就是更多”—假说，Newport 1990; Elman 1993）。

跟克里奥尔语一样，我们可以在手语的习得的某些社会环境中找到类似的情境：Singleton & Newport (2004)展示了，一个学习美国手语（ASL）的孩子（Simon）比他的父母犯的错误少很多。他的父母是在 15 或 16 岁第一次学习美国手语的，并且只有在 70% 的时间运用特殊的规定的动作。Simon 则 90% 的时间都做这些动作。他把从他的父母那里得到的输入规范化了，其中形式-意义对儿的持续使用发挥了重要的作用，也就是说，他没有简单地使用概率匹配，而是有选择地学习。Singleton & Newport (2004: 401)怀疑，这种规范化在克里奥尔语和手语的涌现过程中也起到了重要的作用。但是，我们还没有可以证明这个假说的相关的统计学数据。

13.6 大脑部位的定位

通过在语言生成或处理阶段测量大脑的活动，以及探查有脑损伤的病人，我们可以辨认出大脑的某些部分（布洛卡区和维尼克区）在语言生成和处理中发挥了重要的作用（有关当代的概况，请参考Friederici (2009)）。Chomsky 说有一个语言中心，甚至把它比作一个器官（*organ*）（Chomsky 1977: 164、Chomsky 2005: 1、Chomsky 2008: 133）。这种定位被当作是我们的语言知识的内在基础的证据（也请参考Pinker 1994: 297–314）。

但是，如果这些部位被损坏了，大脑的其他区域可以接管相应的功能。如果损害发生在幼儿早期，语言也可以不在大脑中的这些部位学习（相关资源，请看Dąbrowska 2004: Section 4.1）。

除此之外，也可以观察到，大脑的某些部位在阅读时被激活。如果在大脑的某些部位处理的定位导致语言知识的天赋机制是有效的，那么在阅读时，大脑某些部位的激活就会让我们得出这样的结论，阅读的能力是天赋的（Elman et al. 1996、Bishop 2002: 57）。但是，并没有这样假设（也请参考Fitch, Hauser & Chomsky 2005: 196）。

还可以指出的是，语言处理影响大脑的几个区域，而且不只是布洛卡区和维尼克区（Fisher & Marcus 2005: 11、Friederici 2009）。另一方面，布洛卡区和维尼克区在非语言的任务中也是活跃的，如模仿、运动协调和音乐的处理（Maess et al. 2001）。有关内容的概括和更多的资源，请参考Fisher & Marcus (2005)。

Musso et al. (2003)考察了第二语言习得中的大脑活动。他们给德语为母语者意大利语和日语的语料，然后观察布洛卡区的活动。然后，他们将这个与用意大利语和日语词的人工语言进行对比，这些语言并不符合作者提出的普遍语法的原则。比如说，他们的人工语言的处理包括例（34）中显示的调换语序的问句。

- (34) a. This is a statement.
 这 COP ART 声明
 ‘这是一个声明。’
 b. Statement a is this?
 声明 a COP 这
 ‘声明 a 是这个?’

然后作者观察大脑在学习人工语言时的不同的活跃区域。这是一个有趣的结果，但是不能说明我们具有天赋的语言知识。它只说明了在我们处理我们的自然语言时活跃的区域在学习其他语言时也是活跃的，而且像调换词序这种玩词游戏也会影响大脑的其他区域。

关于大脑的特殊部位的语言的定位部分的详细讨论可以参考Dąbrowska (2004: Chapter 4)。

13.7 语言跟一般认知的区别

不认同天赋的语言知识的学者们相信语言可以通过一般的认知手段来获得。如果可以证明带有严重认知缺陷的人仍能获得正常的语言能力或者具有正常智力水平的人的语言能力十分有限的话，那么我们就可以说明语言和一般的认知是没有关系的。

13.7.1 威廉综合症

有一些智商 (IQ) 特别低，但是能说出合法句子的人。这些人当中有的人有威廉综合症（关于具有威廉综合症的人的能力的讨论，请参考Bellugi, Lichtenberger, Jones, Lai & George (2000)）。Yamada (1981)将这些作为证据来证明是有一个独立于其他智力的语法模型的。

IQ 是根据实足年龄在智商测验（心理年龄）中取得的分数决定的。被研究的青少年都有对应于四到六岁儿童的心理年龄。不过这个年龄的儿童在很多方面已经接近成年人的语言能力。Gosch, Ståding & Pankau (1994: 295)说明了具有威廉综合症的儿童确实显示了语言缺陷，并且他们的语言能力跟他们的心理年龄相称。对于患有威廉综合症的人在形态句法方面的问题，请见Karmiloff-Smith et al. (1997)。有关威廉综合症的讨论在Karmiloff-Smith (1998)中有很好的总结。

13.7.2 带有 FoxP2 基因突变的 KE 家族

这是一个有语言问题的英国家庭，即所谓的 KE 家族。这个备受语言问题困扰的家族成员具有基因缺陷。Fisher et al. (1998)和Lai et al. (2001)发现，这跟 FoxP2 基因 (FoxP2 表示 *Forkhead-Box P2*) 的变异有关。Gopnik & Cargo (1991)从形态方面的问题遗传于基因缺陷这样的事实中总结出，一定有负责某个具体的语法模型（形态）的基因。但是，Vargha-Khadem et al. (1995: 930)证明了，KE 家族不只在形态句法上有问题：受到影响的家庭成员在智力、语言、以及面部肌肉的运用上都有问题。考虑到面部肌肉受到相当多的限制动作，可以假定他们的语言困难也来自于运动的问题 (Tomasello 2003: 285)。不过，KE 家族的语言学问题不仅限于语言生成的问题，也有语言理解的问题 (Bishop 2002: 58)。尽管如此，我们不能将语言缺陷

直接联系到 FoxP2 上，因为还有一些其他能力也受到 FoxP2 变异的影响：发音受阻、形态和句法，也对非语言的智商和面部肌肉的运动问题有影响，以及非语言任务的处理也有问题 (Vargha-Khadem et al. 1995)。

进而，FoxP2 也出现在其他身体问题上：它还负责肺、心、肠和大脑的许多区域 (Marcus & Fisher 2003)。Marcus & Fisher (2003: 260–261)指出，FoxP2 也许不是直接对器官和器官的区域的发展有影响，而是规范了一串不同的基因。由此，FoxP2 不能被叫做语言基因，它只是一个以复杂方式跟其他基因相互影响的基因。只不过，跟其他相比，它对我们的语言能力十分重要，但是，同样这不能来说明叫 FoxP2 是语言基因是有意义的，就像没人会因为肌病而不能直立行走，就将遗传的肌肉功能失调叫做“行走基因” (Bishop 2002: 58)。Karmiloff-Smith (1998: 392)也有相似的观点：有一种基因缺陷导致有些人在十岁开始失去听力，并在三十岁时彻底聋了。这个基因缺陷导致耳朵内部用来听的毛发发生了变化。这样，人们也就不愿意说“听觉基因”这样的词了。

Fitch, Hauser & Chomsky (2005: 190)也认同 FoxP2 不是语言知识的根源。对这个话题的概述，请参考 Bishop (2002)和 Dąbrowska (2004: Section 6.4.2.2)，一般意义上的基因问题，请参考 Fisher & Marcus (2005)。

13.8 刺激贫乏

关于语言知识的天赋机制的一个重要证据是所谓的刺激贫乏论 (PSA) (Chomsky 1980: 34)。在文献中可以找到不同的版本，Pullum & Scholz (2002)对此进行了详细的讨论。在讨论这些变体之后，他们总结了论证的逻辑结构，如下所示 (第 18 页)：

- (35)
- a. 人类儿童学习第一语言时，要么用数据驱动的学习方法，要么用内在知识支持的方法（假说的选言前提）
 - b. 如果儿童学习第一语言是通过数据驱动的方法，那么他们就无法获得他们没有得到必要根据的任何知识（数据驱动学习的定义）
 - c. 但是，儿童确实会学习他们之前没有（实证的前提）的关键性证据。
 - d. 所以说，儿童不是通过数据驱动来学习第一语言的。（b 和 c 的拒取式 *modus tollens*）
 - e. 结论：儿童学习语言是通过天赋知识支持的学习过程。（a 和 d 的析取结论）

Pullum 和 Scholz 随后讨论了作为天赋的语言知识的短语成分证据的四种现象。他们有英语复合词开头部分的复数 (Gordon 1986)、英语助动词的顺序 (Kimball 1973)、英语无指的 *one* (Baker 1978)，以及英语助动词的位置 (Chomsky 1971: 29–33)。在第 13.8.2 节分析这些问题之前，我将讨论一个 PSA 的变体，并将之作为短语结构语法的形式属性。

13.8.1 Gold 定理

在形式语言理论中，语言被看作是包括这门语言的所有表达式的集合。这种集合可以通过各种复杂的重写语法来获得。在第 2 章，我们介绍了一种重写语法，

即上下文无关文法。在上下文无关文法中，在规则的左边总是有一个符号（所谓的非终结符），而在规则的右边有更多的符号。在右边还可以是符号（所谓的非终结符）或所描写语言的词/语素（所谓的终结符）。语法中的词也叫做词汇（V）。部分形式语法的起始符，通常是 S。在文献中，这点备受争议，因为不是所有的表达式都是句子（请看 Deppermann 2006: 44）。但是，这样假设是没有必要的。我们可以用话语作为起始符，然后界定生成 S、NP、VP 的规则或区分出其他语句的任何其他条件。³⁰

从起始符开始，我们可以在一个文法中应用短语结构规则，直到我们得到只包括词（终结符）的序列。我们能够生成的所有序列的集合属于该文法所允准的语言的表达式。这个集合是任意组合可以得到的所有词或语素的子集。包含所有可能序列的集合叫做 V^* 。

Gold (1967) 证明了，在环境 E 下，只给出有限的语言输入，且没有额外的知识，是不可能解决识别特殊的语言类型下任意一种语言的问题的。Gold 关心从给定的语言类型中对一种语言的识别。当在 t_n 时间内的某个点上，操某种语言的人可以认定语言 L 是当下的语言，而且没有改变这个看法的话，这门语言 L 就算被识别出来了。不过，这个时间点不是提前定好的，识别总要在某个时刻实现。Gold 将之叫做“受限的识别”（*identification in the limit*）。环境是任意无限的句子序列 $\langle a_1, a_2, a_3, \dots \rangle$ ，其中，语言中的每个句子都至少在这个序列中出现一次。为了说明识别问题甚至不能解决非常简单的语言类别的问题，Gold 认为包括词表 V 中所有可能的词的序列的语言类别会有一个序列：让 V 是词表，而且 x_1, x_2, x_3, \dots 是出自这个词表的词语序列。出自这个词表的所有字符串的集合是 V^* 。对于例 (36) 中的语言类型来说，它包括 V 中所有可能的成分序列和所应有的一个序列，我们可以证明人们是如何从文本中学会这些语言的。

$$(36) \quad L_1 = V^* - x_1, L_2 = V^* - x_2, L_3 = V^* - x_3, \dots$$

在每次输入后，我们可以猜想语言是 $V^* - \sigma$ ，其中 σ 表示按照字母顺序排列的第一个最短长度，但是还没有看到的序列。如果这个序列后来出现了，那么这个假说就被相应地修正了。按照这个方式，我们最终会得到正确的语言。

如果我们从必须选择的 V^* 语言的集合中进行扩展，那么我们的学习过程就不再起作用了，因为，如果 V^* 是目标语，那么猜想最终会得到错误的结果。如果有一个能够学会这种语言类型的程序，那么它就必须有一些输入后正确地识别出 V^* 。让我们假定，这个输入是 x_k 。学习的过程如何能够告诉我们在这一点上，我们要找的语言不是为了 $j \neq k$ 的 $V^* - x_j$ ？如果 x_k 导致人们猜出了错误的语法 V^* ，那么随后的每个输入会跟正确结果和错误结果都兼容。因为我们只有正向的数据，没有输入允许我们在这些假说之间进行区分，并且给出我们找到的所寻找的语言的超集信息。Gold 指出，没有一种形式语言理论（比如说，正则文法、上下文无关文法和上下文相关文法）假定的语法类型可以在有一些例句的文本的输入后按照有限的步骤识别出来。这对包括所有有限语言和至少一种无限语言的语言的所有类型来说都是正确的。如果正向证据和负向证据都被用来学习，而不是文本的话，情况会有不同。

³⁰ 在第 250 页，我讨论了短语结构语法中属于 S 符号的描写。如果我们在这个描写中省略中心语特征的区别，那么就会得到所有完整句子的描述，即 *the man*（“男人”）或 *now*（“现在”）。

从 Gold 的结果中得出的结论是，对于语言习得来说，人们需要帮助他们从最开始就避免特殊假说的知识。Pullum (2003)批评了将 Gold 的发现作为语言知识必须是天赋的证据的应用。他列出了为了得到 Gold 的结果与自然语言习得相关的一些观点。然后，他证明这些中的每一条都不是没有争议的。

1. 自然语言可以属于可学习的文本的类型，这跟上文提到的上下文无关语法相反。
2. 学习者能知道哪些词语序列是不合语法的（请看第 453–454 页 Gold 的文章中相似的观点）。正如在那之后展示的，儿童确实有直接的负向证据，而且也有非直接的负向证据（请看第 13.8.4 节）。
3. 学习者是否将他们自己限制到一个文法中是不清楚的。Feldman (1972)发展了一个学习程序，淘汰了某点上所有不正确的文法，而且是无限次正确的，但是并不总是需要选择一个正确的文法，然后坚持相应的假说。应用这个程序，我们就有可能学会所有的递归可枚举语言，也就是说，所有的语言都有一个生成语法。Pullum 指出，即使 Feldman 的学习机制也能证明是过于具有限制性的。它需要学习者花费整整一生的时间去追求正确的文法，而且他们在这个过程中还会有不正确的，但是好一些的假说。
4. 学习者可以在改进的条件下工作。如果人们允许某种程度的容忍，那么习得就会更为简单，而且它还会有可能学会递归可枚举语言的类型 (Wharton 1974)。
5. 语言习得并不必然构成关于序列的特定集合的知识的习得，也就是说，生成语法的习得可以创造出这个集合。如果文法被看作是部分描写语言结构的约束的集合，而不必要是语言结构的唯一集合的话，情况就完全不同了（关于这点的更多内容，请看第 6.7 节和第 14 章）。

进而，Pullum 指出，也有可能有限步骤内使用正向输入的 Gold 的程序来学习上下文相关语法的类型，当规则的数量有一个上限 k 的时候，其中 k 是一个任意数。有可能 k 特别大，以至于人类大脑的认知能力不能使用比其有更多规则的文法了。因为一般认为自然语言可以根据上下文相关语法来描述，所以它可以显示出 Gold 意义上的自然语言的句法可以通过文本来学习（也请看 Scholz & Pullum 2002: 195–196）。

Johnson (2004)补充说，在有关语言习得的讨论中还有一个重要方面被忽略了。Gold 的识别问题跟先天论大讨论中起到重要作用的语言习得的问题是不同的。为了让区别明晰化，Johnson 区分了（Gold 意义上的）识别能力和语言习得意义上的先天主义。对于语种类型 C 的识别意味着必须有一个函项 f ，当目标语处于有限时间内时，对于每个环境 E ， C 中的每一种语言 L 永远收敛到假说 L 。

Johnson 对学习力 (*learnability*) (第 585 页)做了如下的界定：“自然语言的类型 C 是可以学习的，当且仅当，任何一个普通的人类儿童和 C 的任意一种语言 L 的几乎任何正常的语言环境下，儿童会在 1 岁到 5 岁间把 L （或某种足够类似于 L 的语言）作为母语来习得。”Johnson 补充道，这个定义跟心理语言学中的学习能力理论没有关系，而是在习得的现实概念上的一个暗示。

Johnson 指出, 在大部分对 Gold 理论的解读中, 识别力和学习力被当作同一回事, 而这在逻辑上是不正确的: 这两个概念的主要区别在于两个量词的使用。对于属于类型 C 中的“一”种语言 L 的识别需要学习者在有限时间内在每个环境中收敛到 L 上。这个时间在不同的环境中的区别可以非常大。这还不是时间的上限。对于语言 L , 我们可以直接构造出环境的序列 E_1, E_2, \dots , 这样在环境 E_i 中的学习者不会在早于时间 t_i 前猜出 L 。与识别力不同, 学习力是指在每个正常环境中一个时间点之后, “每个”正常的孩子都可以收敛到正确的语言上。这就意味着儿童是在特定的时间段之后习得语言的。Johnson 引用了Morgan (1989: 352)的话, 他说, 儿童在大约听到 4,280,000 个句子后学会他们的母语的。如果我们假设学习力的概念对于可能时间有一个有限的上限的话, 那么很少的语言类型会在限制内被识别出。Johnson 是这样说明的: 让 C 是包括 L 和 L' 语言的类型, 其中 L 和 L' 有一些共同的成分。我们有可能构造出这样的文本, 其中第一个句子在 L 和 L' 中都有。如果学习者将 L 当作是工作假说, 那么继续这个属于 L' 的句子的文本, 如果他将 L' 作为他的假说, 则继续属于 L 的句子。每一种情况下, 学者着都在 n 个步骤后得到一个错误的假说。这就意味着识别力不是语言习得的一个貌似正确的模型。

除了识别能力不具备心理现实性的事实外, 它跟学习能力也不相容 (Johnson 2004: 586)。对于识别能力来说, 只需要找到一个学习者 (上面提到的功能 f), 但是, 学习能力对 (几乎) 所有正常的孩子都适用。如果我们把所有的因素都保持不变, 那么相比于学习力, 更容易证明语言类型的识别力。一方面, 识别能力普遍地满足所有环境的要求, 不管这些是不是看起来很奇怪或者他们能包括多少重复。另一方面, 学习能力具有除了正常环境的 (几乎) 通用的条件。所以说, 跟识别能力相比, 学习能力指更少的环境, 这样就导致有问题的文本作为输入以使得语言是不可学习的可能性更少。进而, 学习能力是这样界定的, 学习者不需要真正地学习 L , 而是学会某种相似的识别能力。所以说, 学习能力不在识别能力之后, 反之亦然。

最后, Gold 分析了没有考虑语义信息的句法知识的习得。但是, 儿童在学习语言时, 从上下文中获得了大量的信息 (Tomasello et al. 2005)。正如Klein (1986: 44)指出的, 如果把人在屋子中, 人没有学会任何东西, 而是汉语的句子在跟他们做游戏。语言是在社会和文化语境中习得的。

总之, 我们应该指出, 天赋的语言知识的存在性不能从语言学习能力在数学上的发现推导出来。

13.8.2 四个案例

Pullum & Scholz (2002)深入地探讨了刺激贫乏论的四个著名例子。我们将在下面讨论这些内容。Pullum 和 Scholz 的文章收录在一本讨论性的文集中。由Scholz & Pullum (2002)发表的反对他们文章的观点的文章也收录在这本文集中。Eisenberg (1992)反驳了Chomsky (1986b)和德语文献中有关 PoS 论元的观点。

13.8.2.1 名名组合的复数

Gordon (1986)提出, 英语中的复合词只允许不规则的复数变化, 即 *mice-eater*, 而不是 **rats-eater*。Gordon 认为, 以不规则的复数作为开头成分的复合词非常少见, 以致于儿童不能仅从语言数据中学会这样的复合词。

在第 25-26 页, Pullum 和 Scholz 讨论了英语中的数据, 并且说明规则的复数形式确实只能在复合词的第一个成分中出现 (*chemicals-maker, forms-reader, generics-maker, securities-dealer, drinks trolley, rules committee, publications catalogue*)。³¹ 这说明, 所谓的没有从语言事实中学习实际上在语言学上是不充分的理由, 所以不能用来解释习得。

13.8.2.2 助动词的位置

第二项研究分析情态动词和助动词的位置。Kimball (1973: 73-75) 讨论了例 (37) 中的数据和 (38) 中的规则, 该规则类似于 Chomsky (1957: 39) 提出的一种规则, 并且用来描述下面的语言事实:

- (37)
- a. It rains.
PRON 下雨
‘下雨了。’
 - b. It may rain.
PRON 会 下雨
‘会下雨。’
 - c. It may have rained.
PRON 会 AUX 下雨
‘可能已经下雨了。’
 - d. It may be raining.
PRON 会 AUX 下雨
‘可能正在下雨。’
 - e. It has rained.
PRON AUX 下雨
‘已经下雨了。’
 - f. It has been raining.
PRON AUX AUX 下雨
‘一直在下雨。’
 - g. It is raining.
PRON AUX 下雨
‘正在下雨。’
 - h. It may have been raining.
PRON 会 AUX AUX 下雨
‘有可能一直在下雨。’

- (38) Aux → T(M)(have+en)(be+ing)

³¹ 也可以参考 Abney (1996: 7) 摘自《华尔街日报》中的例子。

T 表示时态, M 代表情态动词, *en* 代表分词语速 (*been/seen/...* 中的 *-en* 和 *rained* 中的 *-ed*)。这里的括弧表明了表达式的可选择性。Kimball 指出, 如果 (37h) 是合乎语法的, 就只能构造出这个规则。如果不是这样, 那么我们就必须识别出规则中的材料, 这样才能覆盖 (M)(have+en)、(M)(be+ing) 和 (have+en)(be+ing) 这三种情况。Kimball 假定, 儿童掌握了复杂的规则, 因为他们知道诸如 (37h) 这样的句子是合乎语法的, 以及他们知道情态动词和助动词必须出现的位次。Kimball 认为, 儿童对于 (37h) 的语序没有正向的证据, 并由此得出结论说, 关于 (38) 的规则的知识必须是天赋的。

Pullum 和 Scholz 指出了这个刺激贫乏论的两个问题: 首先, 他们找到了上百个例子, 其中有些是来源于儿童故事的, 所以 Kimball 的观点是, 诸如 (37h) 的“非常少”的句子应该被纳入研究范围。对于 PSA 来说, 我们至少应该区分多少情况是被允许的, 如果我们仍然希望说明没有什么能够从中学到的话 (Pullum & Scholz 2002: 29)。

第二个问题是, 认为 (37h) 的规则在我们的语言知识中发挥了作用是讲不通的。语言事实方面的发现表明这个规则在描写上是不充分的。如果 (38) 中的规则在描写上不充分, 那么它就不具备描写上的充分性, 也就不能解释它是如何获得的。

除了 (38) 的规则之外, 这里所有的理论都认为助动词或情态动词嵌套了一个短语, 即没有包括所有助动词和情态动词的 Aux 节点, 但是有像下面这样的 (37h) 的结构:

- (39) It [may [have [been raining]]].
PRON [会 [AUX [AUX 下雨]]]
‘有可能一直在下雨。’

这里, 助动词和情态动词总是选择嵌套的短语。现在, 习得问题看起来完全不同了: 说话者必须学会动词投射中选择助动词或情态动词的中心语动词的形式。如果这个信息学会了, 那么嵌套的动词性投射有多复杂就是无关的了: *may* 可以跟非定式词汇动词 (37b) 或非定式助动词 (37c,d) 相组合。

13.8.2.3 *one* 的指称

Pullum 和 Scholz 调查的第三个方面是英语的代词 *one*。Baker (1978: 413–425, 327–340) 认为, 儿童学不会 *one* 可以指代比单个词大的成分, 如例 (40)。

- (40) a. I would like to tell you another funny story, but I've already told
我要 喜欢INF 告诉你 另一个 有趣的故事 但是我.AUX 已经 告诉
you the only *one* I know.
你 DET 唯一 PRON 我知道
‘我想给您讲另一个有趣的故事, 但是我已经给你讲了我唯一知道的一个故事。’
b. The old man from France was more erudite than the young *one*.
DET 老人 从 法国 AUX 更加 博学 比 DET 年轻 PRON
‘从法国来的老人比年轻人更加博学。’

Baker (第 416–417 页) 认为, *one* 决不能指代 NPs 内部的单个名词, 并且以 (41) 中的例句作为证据:

- (41) * The student of chemistry was more thoroughly prepared than the one
 DET 学生 PREP 化学 AUX 更加 全看 准备 比 DET PRON
 of physics.
 PREP 物理学
 ‘化学专业的同学比物理学专业的同学准备更为充分。’

按照 Baker 的观点, 学习者需要负向数据来获得不合乎语法知识。根据他的推论, 由于学习者从没有接触过负向证据, 他们就不可能学会相关的知识, 并且由此必须是已经掌握它了。

Pullum & Scholz (2002: 33) 指出, 有带有相同结构的可以接受的例子, 如 (41) 中的例子所示:

- (42) a. I'd rather teach linguistics to a student of mathematics than
 我.AUX 更 教 语言学 INF ART 学生 PREP 数学 比
 to one of any discipline in the humanities.
 INF PRON PREP 任何专业 PREP DET 人文学科
 ‘我更愿意给数学专业的学生教语言学, 而不是人文学科中的任何一个专业的学生。’
 b. An advocate of Linux got into a heated discussion with one
 ART 支持者 PREP Linux 卷进 ART 热烈的讨论 PREP PRON
 of Windows NT and the rest of the evening was nerd talk.
 PREP Windows NT 和 DET 剩余的 PREP DET 晚上 AUX 书呆子谈话
 ‘一位 Linux 系统的支持者被卷进了一场跟 Windows NT 操作系统的支持者的热烈的讨论之中, 晚上的剩余时间就成了书呆子间的讨论。’

这意味着, 关于 (41) 中结构的合格性, 没有什么是要学习的。而且, 要获得 *one* 可以指代更大的成分这个事实的数据并不像 Baker (第 416 页) 声称的那样毫无希望: 有例子表明, 只有在 *one* 指代一个更大的字符串时是可以解释的。Pullum 和 Scholz 从各种语料中寻找例子。他们也从 CHILDES 语料库中找到了例子, CHILDES 语料库是一个包括了儿童的交流语言的语料库 (MacWhinney 1995)。下面的例子选自一档日间电视节目:

- (43) A: “Do you think you will ever remarry again? I don’t.”
 B: “Maybe I will, someday. But he’d have to be somebody very special. Sensitive and supportive, giving. Hey, wait a minute, where do they make guys like this?”
 A: “I don’t know. I’ve never seen one up close.”

显然, 这里的 *one* 不能指代 *guys*, 因为 A 已经看到了 *guys*。相反, 它指代 *guys like this*, 即感性又愿意提供帮助的人。

再次, 有问题出现了。听者需要听多少个例子才能算是 PSA 理论的支持者眼中认可的例子呢。

13.8.2.4 极性问句中助动词的位置

Pullum 和 Scholz 提出的 PoS 的第四个问题源自 Chomsky, 它是关于英语助动词在极性问句中的位置的。正如在第 90 页所展现的, 在 GB 理论中, 极性问句是由助动词从句中 I 位置移到开头位置 C 而生成的。在转换语法的早期版本中, 确切的分析是不同的, 但是主要的观点还是最高阶的助动词被移到了从句的开头。Chomsky (1971: 29–33) 讨论了 (44) 中的句子, 并且认为, 儿童知道他们必须移动最高阶的助动词, 即使没有正向的证据。³² 比如说, 如果他们乐于提出这样的观点, 人们简单地将第一个助动词放在句子的开头的话, 那么这个假说就会针对 (44b) 而得到正确的结论 (44a), 而不是针对 (44c) 的, 因为极性问句应该是 (44d), 而不是 (44e)。

- (44) a. The dog in the corner is hungry.
DET 狗 在 DET 角落 COP 饿
‘在角落里的狗饿了。’
- b. Is the dog in the corner hungry?
COP DET 狗 在 DET 角落 饿
‘在角落里的狗饿了吗?’
- c. The dog that is in the corner is hungry.
DET 狗 CONJ COP 在 DET 角落 COP 饿
‘在角落里的那只狗饿了。’
- d. Is the dog that is in the corner hungry?
COP DET 狗 CONJ COP 在 DET 角落 饿
‘在角落里的那只狗饿了吗?’
- e. *Is the dog that in the corner is hungry?
COP DET 狗 CONJ 在 DET 角落 COP 饿

Chomsky 表示, 儿童没有任何证据说明这样的事实, 人们简单地把第一个助动词线性地提前是错的, 这就是为什么他们可以按照数据驱动的学习过程来验证这个观点。他甚至进一步提出, 操英语者只是很少, 甚至是从未生成过 (44d) 这样的例子 (Piattelli-Palmarini (1980: 114–115))。在语料库数据和貌似真实的构造出的例子的帮助下, Pullum (1996) 证明了这个观点明显是错误的。Pullum (1996) 在《华尔街日报》中找到了例子, 并且 Pullum & Scholz (2002) 更为细致地讨论了相关的例子, 并且加上 CHILDES 语料库中的例子, 一同说明成年人不能只造出相关类型的句子, 而是他们也出现在儿童的输入中。³³ CHILDES 语料库中的例子反驳了第一助动词需要前置的假说, 如 (45) 所示:³⁴

³² 助动词变换的例子也用在最近的 PoS 观点中, 如 Berwick, Pietroski, Yankama & Chomsky (2011) 和 Chomsky (2013: 39)。Bod (2009b) 的工作并没有得到讨论。更多有关 Bod 的方法, 请看第 13.8.3 节。

³³ 关于该点的更多信息, 请参考 Sampson (1989: 223)。Sampson 引用了英语学校研究的 William Blake 的诗的一部分, 以及一本儿童百科全书。这些例子当然在助动词位置的习得中没有起到重要的作用, 因为这个语序是在 3 岁 2 个月时学会的, 也就是说, 在孩子们达到上学的年龄时早就已经学会了。

³⁴ 请参考 Lewis & Elman (2001)。语言习得方面的研究者认同, 在跟孩子们的交谈中, 这类句子出现的频率实际上是非常低的。请参考 Ambridge et al. (2008: 223)。

- (45) a. Is the ball you were speaking of in the box with the bowling
COP DET 球 你 AUX 谈 PREP 在 DET 盒子 PREP DET 保龄球
pin?
瓶
‘你说的盒子里的那个球是保龄球瓶吗?’
- b. Where's this little boy who's full of smiles?
哪儿.AUX 这 小 男孩 CONJ.AUX 充满 PREP 笑容
‘这个满脸笑容的小男孩哪儿去了?’
- c. While you're sleeping, shall I make the breakfast?
当 你.AUX 睡觉 AUX 我做 DET 早饭
‘在你睡觉的时候, 我能做早饭吗?’

Pullum 和 Scholz 指出, 诸如 (46b) 的 *wh*-问句也是相关的, 如果我们认为他们是从极性问句派生而来的话 (请看本书第 90 页), 而且如果我们希望展示儿童说如何学会基于结构的假说的话。这可以通过 (46) 中的例子来解释: (46a) 被派生出的基础形式是 (46b)。如果我们要把 (46b) 中的第一个助动词前置, 那么我们会得到 (46c)。

- (46) a. Where's the application Mark promised to fill out?
哪儿.AUX DET 申请表 Mark 承诺 INF 填 PREP
‘Mark 承诺要填的表在哪儿呢?’³⁵
- b. the application Mark [AUX PAST] promised to fill out [AUX is] there
DET 申请表 Mark [AUX 过去] 承诺 INF 填 PREP [AUX AUX] 那儿
‘Mark 过去承诺要填的表在那儿’
- c. *Where did the application Mark promised to fill out is?
哪儿 AUX DET 申请表 Mark 承诺 INF 填 PREP AUX

但是, (46c) 是不正确的证据也可以在跟孩子有关的语言中找到。Pullum 和 Scholz 举出了 (47) 中的例句:³⁶

- (47) a. Where's the little blue crib that was in the house before?
哪儿.AUX DET 小 蓝色螃蟹 CONJ COP PREP DET 房子 以前
‘以前房子里的小蓝螃蟹在哪儿?’
- b. Where's the other dolly that was in here?
哪儿.AUX DET 另一个娃娃 CONJ COP PREP 这儿
‘原来在这儿的另一个娃娃在哪儿?’
- c. Where's the other doll that goes in there?
哪儿.AUX DET 另一个娃娃 CONJ 走 在那儿
‘放在那儿的另一个娃娃在哪儿?’

³⁵ 译自 CHILDES 语料库中的一档电视节目。

³⁶ 这些句子选自 NINA05。CHA in DATABASE/ENG/SUPPES/。

这些问句具有 *Where's NP?* 这样的形式，其中 NP 包含一个关系从句。

在 (45c) 中，有另一个从句位于实际的疑问句的前面，一个包括助动词的状态从句。这个句子就可以证明线性上第一位助动词必须前置的假说是错误的 (Sampson 1989: 223)。

总之，在儿童的语言输入中，有许多可验证的句子类型允许他们在两种假说中选择。再一次，问题来了，有多少证据可以被认为是足够的呢？

Lasnik & Uriagereka (2002) 和 Legate & Yang (2002) 评论了 Pullum 和 Scholz 的文章。Lasnik 和 Uriagereka 论证道，习得问题要比 Pullum 和 Scholz 提出的大得多，因为不知道他要习得语言的什么知识的学习者不能只有我们已经讨论过的 (48) 中的假说，还应该有 (49) 中的其他假说：

- (48) a. 将第一个助动词放在从句的开头。
b. 将第一个助动词放在从句的开头的 matrix-Infl 上。

- (49) a. 将任意一个助动词放在从句的开头。
b. 将任意一个定式的助动词放在从句的开头。

(49) 中的所有假说都可以通过 (50) 中的句子得到许可：

- (50) a. Is the dog in the corner hungry?
COP DET 狗 PREP DET 角落 饿
'在角落里的狗饿吗？'
b. Is the dog that is in the corner hungry?
COP DET 狗 CONJ COP PREP DET 角落 饿
'在角落里的那只狗饿吗？'

但是，他们也可以允准 (51) 中的句子：

- (51) *Is the dog that in the corner is hungry?
COP DET 狗 CONJ PREP DET 角落 COP 饿

必须要指出的是为什么所有允准 (51) 的假说应该被丢弃，这是因为学习者在他们的自然语言的输入中并没有关于 (51) 是不可能的任何信息。他们缺乏负向证据。如果 (50b) 跟正向证据一起出现，那么这就无论如何暗示了 (48b) 中的假说必然是正确的了。Lasnik 和 Uriagereka 提出了跟 (50b) 也相容的假说，如下所示：

- (52) a. 将第一个助动词放在首位（随后会有声调上的变化）。
b. 将第一个助动词放在首位（随后是第一个完整的成分）。
c. 将第一个助动词放在首位（随后是第一个剖析的语义单位）。

这些假说不能说明像 (53) 这样包括连词的句子：

- (53) Will those who are coming and those who are not coming raise their
AUX 那些 CONJ AUX 来 并且那些 CONJ AUX 不 来 举起他们的
hands?
手
'要来或不来的人请举起手'

(52) 中的假说也可以允准诸如 (54) 的句子:

- (54) * Are those who are coming and those who not coming will raise their
AUX 那些 CONJ AUX 来 并且那些 CONJ AUX 不 来 举起他们的
hands?
手

听者听到 (53) 这样的句子时会反对假说 (52), 并由此排除 (54)。但是, 仍有可能想到一个跟之前讨论的所有数据相同的一个相似的看似正确的假说。

Legate & Yang (2002) 接受了 Pullum 和 Scholz 的挑战, 并明确表示要习得一个特定的现象, 人们需要多少次事件。他们这样写道:

假设我们有两个不同的习得问题, P_1 和 P_2 , 其中每个问题都包括一个二元决策。对于 P_1 来说, 让 F_1 是可以解决是 P_1 还是其他的数据的频率, 然后 F_2 是 P_2 的频率。进一步假设, 儿童在几乎相同的发展阶段成功地习得了 P_1 和 P_2 。那么, 在任何有关语言发展的定量研究的理论中, 我们期望 F_1 和 F_2 也大体是一致的。相反, 如果 F_1 和 F_2 的结果非常不同, 那么 P_1 和 P_2 就必须表示性质不同的学习问题了。

现在, 让 P_1 是助动词变换的问题。两个选项是基于结构的假说 (3b-i) 和第一助动词假说 (3a-i)。 (Legate & Yang 2002: 155)

有关英语中助动词的位置的知识是在儿童的 3 岁 2 个月时习得的。根据 Legate 和 Yang, 我们需要另一个在 3 岁 2 个月时学会的习得现象来进行比较。作者重点讨论了主语脱落问题³⁷, 该现象是在 36 个月时学会的 (比助动词变换早两个月)。根据作者的观点, 习得问题包括一个二元决策过程: 第一步, 我们需要在 (48) 中的两个假说进行选择。第二步, 学习者需要决定一种语言是否使用显性主语。作者认为, 像 *there* 的虚位词的使用, 可以用来证明学习者所学习的语言并不是可以带可选主语的那种语言。然后, 他们在 CHILDES 语料库中计算了包括 *there*-主语的句子, 并且估算了学习中听到的 1.2% 句子的 F_2 。按照他们的观点, 由于我们这里处理的是相同难度的现象, 助动词变换要是能够可以学会的话, 诸如 (44d) 和 (47) 的句子就应该包括 1.2% 的输入。

然后, 作者检索了 Nina 和 Adam 的语料 (都属于 CHILDES), 并注意到 0.068 到 0.045% 的语料具有 (47) 的形式, 而没有语料具有 (44d) 的形式。他们总结道, 这个数字不足以作为正向证据。

在指出 Pullum 和 Scholz 从《华尔街日报》上得到的数据不必然跟语言习得相关这点上, Legate 和 Yang 是正确的。而且, 他们还指出了在数据中没有发现带有复杂主语名词短语的例子, 或者至少是可以忽略不计的。但是, 他们的论述还是有三个严重的问题: 首先, 在虚位主语和 pro-脱落语言的属性之间没有关系: 加利西亚语 (Raposo & Uriagereka 1990: Section 2.5) 是带有虚位代词的 pro-脱落语言。意大利语有一个虚位成分 *ci*,³⁸ 尽管意大利语可以算是 pro-脱落语言, Franks (1995) 把上索布语和下索布语列为 pro-脱落语言, 他们的主语位置有虚位成分。因此, 由于虚位代词跟 pro-脱落参数没有关系, 他们的出现频率就跟参数值的习得

³⁷ 这个现象也叫做 *pro-drop* (pro-脱落)。关于 pro-脱落参数的更为详细的讨论请看第 16.1 节。

³⁸ 不过, 并不是所有人都把 *ci* 看作是虚位成分。相关的概述请看 Remberger (2009)。

无关。如果在省略主语的可能性和主语虚位的出现频率之间有关系的话，那么说挪威语和丹麦语的儿童就应该比说英语的儿童更早学会他们的语言必须要有主语，因为虚位成分在丹麦语和挪威语中的出现频率更高 (Scholz & Pullum 2002: 220)。在丹麦语中，对应于英语 *there*-结构的出现频率是英语的两倍。目前，在习得比率上是否真的存在差异仍是不清楚的 (Pullum 2009: 246)。

第二，在构造他们的刺激贫乏论时，Legate 和 Yang 认为是有天赋的语言知识的 (pro-脱落参数)。这样，他们就陷入了循环论证，因为他们理应说明内在的语言知识的假说是不可缺少的。(Scholz & Pullum 2002: 220)。

Legate 和 Yang 的观点的第三个问题是，他们认为转换分析是唯一可能的分析。下面的引述清楚地说明了这个观点 (Legate & Yang 2002: 153):

当然，构成问句的正确操作是基于结构的：它包括将句子分析为结构化组织的短语，并将位于主语 NP 后的助词前置，这可以是任意长度的：

- (4) a. Is [the woman who is singing] e happy?
COP [DET 女人 CONJ AUX 唱歌] e 高兴
‘唱歌的女人高兴吗？’
- b. Has [the man that is reading a book] e eaten supper?
AUX [DET 人 CONJ AUX 读 ART 书] e 吃 晚饭
‘看书的那个人吃晚饭了吗？’

由 Chomsky 推进的这个分析是基于转换的（请看第 90 页），也就是说，学习者需要按照 Legate 和 Yang 所描述的学习：助动词必须移到主语名词短语的前面。不过，还有其他分析说不需要变换或其他相当的机制。如果我们的语言知识不包括任何有关变换的信息，那么他们关于需要学习什么的观点就是错误的。例如，我们可以假设，像在范畴语法中，助动词构成了一组具有特殊分布属性的词类。有可能替代他们是疑问句中观察到的首位，另一种是在主语后 (Villavicencio 2002: 104)。这样就需要习得主语是在中心语的左边还是右边实现的信息。除了这个基于词汇的分析，另一种方法是，我们可以采用构式语法 (Fillmore 1988: 44、1999、Kay & Fillmore 1999: 18)、认知语法 (Dąbrowska 2004: Chapter 9) 或 HPSG 理论 (Ginzburg & Sag 2000) 来分析。在这些框架中，只有简单的两个模式³⁹来分析这两个序列，他们是根据动词和主语的语序来赋予不同的意义的。习得的问题就是学习者需要在输入中识别相应的短语范式。他们需要认识到 Aux NP VP 在英语中是合乎语法的，并且有疑问的语义。构式语法导向的文献对习得的相关理论作了非常好的论述（请看第 16.3 节和 16.4）。基于构式的习得理论也被我们能看到的频率效应这样的事实所支持，即，助词变换首先由儿童用在一些助词上，而且只在发展的后期，然后扩展到所有的助词上。如果说话者学会助词构式具有 Aux NP VP 的范式，那么 Lasnik 和 Uriagereka 在 (53) 中提出的并列数据不再是问题了。这是因为，如果我们只将第一个并列成分指派到范式 Aux NP VP 的 NP 上，然后并列结构 (*and those who are not coming*) 中剩下的部分还是未分析的，而且不能被整合进整个句子中。由此，听者被迫修改他的假设，即 *will those who are coming* 对应于 Aux NP VP 中 Aux NP 的序列，然后使用整个 NP *those who are coming and those who are not coming*。

³⁹ Fillmore (1999) 提出了主语助动词变换构式的次类型，因为这类变换并不只在问句中出现。

由此,对于习得英语来说,先学一些助动词的范式是 Aux NP VP,然后再学所有的助动词这样是足够的。Lewis & Elman (2001)也证明了这个观点,他训练了一个神经网络,其中使用了助动词构式中不包括带有关系从句的 NPs 的数据。但是,关系小句在其他结构中有。训练语料的复杂度一点一点增加,就像儿童接收语言输入一样 (Elman 1993)。⁴⁰ 神经网络可以预测一个词语序列的下一个符号。对于带有疑问语序的句子来说,预测的结果是正确的。即使是 (55) 中的关系代词也被预测了,尽管序列 Aux Det N Relp 在训练语料中并没有出现。

- (55) Is the boy who is smoking crazy?
COP DET 男孩 CONJ AUX 抽烟 疯狂
'抽烟的那个男孩疯狂吗?'

进而,如果这个网络显示了不合乎语法的句子 (56),那么系统就会给出一个出错的信号:

- (56) *Is the boy who smoking is crazy?
COP DET 男孩 CONJ 抽烟 COP 疯狂

关系代词后面不应该是动名词,应该是定式动词。构建的神经网络当然不是我们在习得和言语生成的过程中对我们大脑中活动的充分模拟。⁴¹ 但是,实验表明,学习者接收到的输入包括可以用来习得语言的丰富的统计信息。Lewis 和 Elman 指出,输入中词的分布的统计信息不是说话者所有的唯一信息。除了关于分布的信息,他们也暴露在上下文的信息中,而且可以利用词在语音上的相似性。

跟 (56) 中的不合乎语法的句子有关的是,有人认为,永远不会造出这类句子的事实说明儿童已经知道语法操作是基于结构的,而且这就是为什么他们不会有这样的想法,只有线性上第一个动词移位了 (Crain & Nakayama 1987)。这个观点并不易于验证,因为儿童一般不说相对复杂的句子。由此,只有可能测试他们会犯的相应的错误的非法句子。Crain & Nakayama (1987)做了这样的实验。他们的研究被 Ambridge, Rowland & Pine (2008)批评了,因为这些作者可以证明儿童真的在对助动词前置的时候会犯错误。作者将 Crain 和 Nakayama 的第一次研究的结果跟 Crain 和 Nakayama 的研究中助动词的错误选择区分开。由于助动词 *is* 的使用,不合乎语法的例子具有从不或很少互相挨着出现的词对儿 (57a 中的 *who running*)。

- (57) a. The boy who is running fast can jump high.
DET 男孩 CONJ AUX 跑快 AUX 跳 高
'跑得快的那个男孩跳的高.'
→
* Is the boy who running fast can jump high?
AUX DET 男孩 CONJ 跑 快 AUX 跳 高

⁴⁰ 这里有文化差异。在一些文化中,成年人不跟还没掌握全部语言能力的儿童讲话 (Ochs 1982; Ochs & Schieffelin 1985) (也请参考第 13.8.4 节)。由此,儿童就必须从环境中学习,即他们听到的句子反应了语言的全部复杂性。

⁴¹ 请看 Hurford (2002: 324) 和 Jackendoff (2007: Section 6.2) 有关神经网络的特定类型带来的问题,以及 Pulvermüller (2003, 2010) 提出的不具有这些问题的另一种观点。

- b. The boy who can run fast can jump high.
DET 男孩 CONJ AUX 跑 快 AUX 跳 高

‘跑得快的男孩可以跳得高。’

→

- * Can the boy who run fast can jump high?
AUX DET 男孩 CONJ 跑 快 AUX 跳 高

如果我们使用助动词 *can*，这个问题就消失了，因为 *who* 和 *run* 一定是一起出现的。然后这就导致儿童实际上会犯他们不应该犯的错误，因为不正确的句子确实违反了理应属于天赋的语言知识的那部分。

Estigarribia (2009) 具体调查了英语的极性问句。他指出，孩子们输入的具有 Aux NP VP 形式的极性问句连一半都不到（第 74 页）。相反，父母会用简化的方式跟孩子沟通，并且使用如下的句子：

- (58) a. That your tablet?
那 你的 药片
‘那是你的药片？’
b. He talking?
他 说话
‘他在说话？’
c. That taste pretty good?
CONJ 品尝 特别 好
‘那个尝起来特别好吃？’

Estigarribia 将不同的范式分成复杂的类型，如下所示：FRAG (*fragmentary*)、SPRED (*subject predicate*) 和 AUX-IN (*auxiliary inversion*)。(59) 指出了相应的例子：

- (59) a. coming tomorrow? (FRAG)
来 明天
‘明天来吗？’
b. you coming tomorrow? (SPRED)
你 来 明天
‘你明天来吗？’
c. Are you coming tomorrow? (AUX-IN)
AUX 你 来 明天
‘你明天来吗？’

我们看到的是，复杂度一类比一类高。Estigarribia 提出了一种语言习得的系统，其中更简单的类型在更为复杂的类型之前习得，而后者从更为简单类型的外围修饰成分发展而来（第 76 页）。他认为，问句形式是从右到左学习的 (*right to left elaboration*)，也就是说，(59a) 是先学的，然后是 (59b) 中在 (59a) 的基础上包括主语的范式，再然后，在第三步，才出现附加的助动词模式 (59c) (第 82 页)。在这种学习过程中，没有包括助动词的变换。这个观点跟基于约束的分析是一致的。

13 语言知识的天赋性

的, 如Ginzburg & Sag (2000)。Freudenthal, Pine, Aguado-Orea & Gobet (2007)提出的一个类似的方法将在第 16.3节讨论。

Bod (2009b)提出了一项更为有趣的研究。他指出, 如果我们假设带有任意种类分叉的树的话, 是有可能学会助动词变换的, 即使在输入中没有复杂名词短语的助动词变换。他使用的分析策略和他得到的结果非常有趣, 我们将在第 13.8.3节来详细探讨。

总之, 我们可以说儿童在助动词位置的使用上是会犯错误的, 而这些错误也许不会犯, 如果这些相关的知识是天赋的话。关于输入中词的分布的统计信息足以学会输入中实际不带这种复杂句子的复杂句子的结构。

13.8.2.5 小结

Pullum & Scholz (2002: 19)提出, 刺激贫乏论 (APS) 应按照下面的结构来表示:

(60) APS 的具体策略:

- a. 习得特征: 详细描述应该知道的知识。
- b. 缺陷规范: 识别出学习者必须接触到的句子的集合, 这样关于习得的数据驱动的学习就会得到支持。
- c. 不可或缺论: 给出如果学习是数据驱动的理论, 然后在没有接触到缺陷的句子的時候, 习得就不会发生。
- d. 不可及证据: 支持有缺陷的句子的类例在习得过程中对学习者是接触不到的观点。
- e. 习得证据: 给出理由以相信习得实际上在童年时期就被学习者熟知了。

正如上面四个问题的研究所示, 有许多反对习得论的理由。如果没有必要获得习得, 那么就没有任何内在的语言知识的证据了。习得论必须至少是足够可描述地。这是语言学家可以回答的语言事实方面的问题。在 Pullum 和 Scholz 讨论的四个刺激贫乏论中的三个观点中, 有没有被充分描写的部分。在前面的章节中, 我们已经接触到了其他的刺激贫乏论, 其中包括语言数据无法在事实上支持的观点 (如邻接原则)。对于 (60) 中剩下的观点, 需要跨学科的研究工作: 缺陷的具体化要落实到形式语言的理论中 (句子的集合的具体化), 不可或缺论是一个学习理论领域中的数学任务, 不可及的证据是一个语言事实方面的问题, 可以通过语料库获得, 最终习得的证据是实验发展心理学的问题 (Pullum & Scholz 2002: 19–20)。

Pullum & Scholz (2002: 46)指出了关于 (60c) 的一个有趣的悖论: 没有学习的数学理论的结果, 我们无法获得 (60c)。如果我们希望提出有效的刺激贫乏理论, 我们就需要自动在学习理论中得到进步, 也就是说, 有可能比之前认为的学会更多。

13.8.3 无指导的数据驱动的剖析 (U-DOP)

Bod (2009b)提出了不需要任何有关句中的词类或词间关系信息的程式。我们唯一需要做的假设是, 有着某种结构。这个程式包括三个步骤:

1. 针对给定句子的集合计算所有可能的 (不带范畴符号的) (二叉) 树。
2. 将这些树分成子树。

3. 计算出每个句子的理想树。

这个过程可以解释 (61) 中的句子：

- (61) a. Watch the dog.
小心 DET 狗
‘小心狗。’
b. The dog barks.
DET 狗 叫
‘狗叫。’

指派给这些语句的树只使用了范畴符号 X，因为相关短语的范畴还未知。为了让例子具有可读性，这些词本身不会被给予范畴 X，尽管我们当然可以这样做。图 13.2 说明了 (61) 的树。下一步，这些树被分成了子树。图 13.2 中的树具有图 13.3

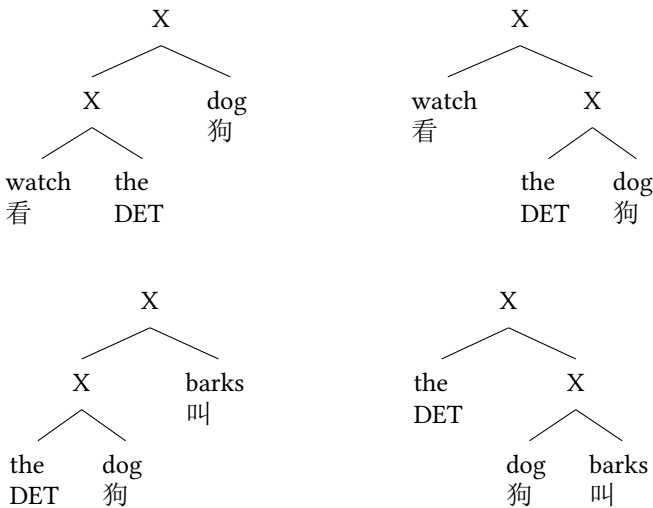


图 13.2: *Watch the dog* 和 *The dog barks* 的可能的二叉树结构.

on the next page中可见的子树。第三步，我们现在需要计算每个语句的最优树。对于 *The dog barks*. 来说，子树的集合中有两棵树完全对应于这个语句。但是，也可以从子树中构造结构。由此，对 *The dog barks*. 来说，就有多重推导的可能性，他们都使用了图 13.3 中的树：一方面是使用了整棵树的许多小的推导过程，另一方面是从小的子树构建树的推导过程。图 13.4 展现了这些子树是如何构造的。如果我们现在想判断 (62) 中的哪个分析是最优的，我们就必须要计算每棵树的概率。

- (62) a. [[the dog] barks]
[[DET 狗] 叫]
‘狗叫’

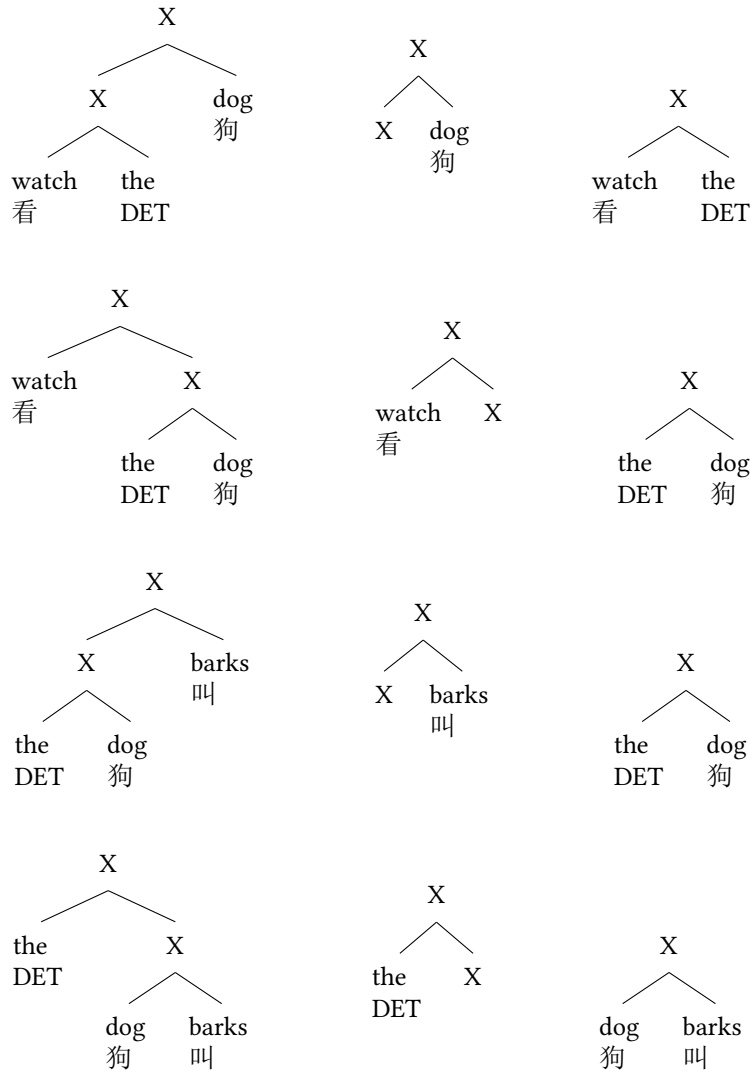
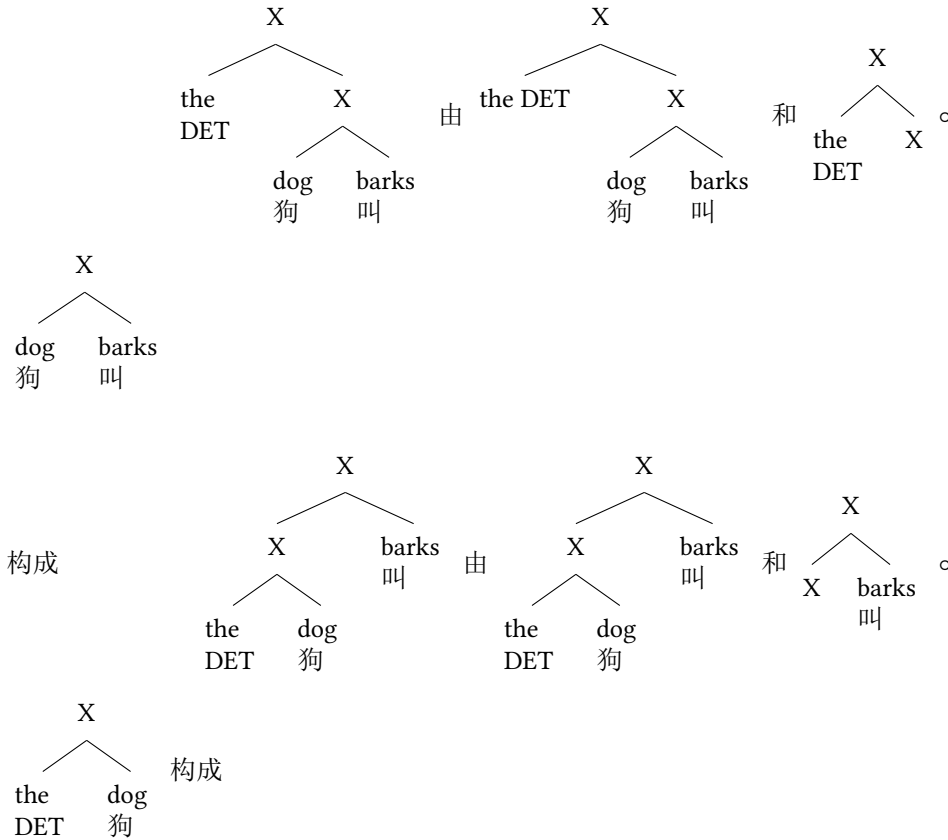


图 13.3: 图 13.2 中的树的子树

图 13.4: 应用图 13.3 中的子树的 *The dog barks* 的分析

- b. [the [dog barks]]
 [DET [狗 叫]]
 ‘狗叫’

一棵树的概率是它的所有分析的可能性的总和。我们可以在图 13.4 中找到 (62b) 的两种分析。(62b) 的第一种分析的可能性对应于从所有子树的集合中选择 [the [dog barks]] 的完整树的概率。因为有十二棵子树，选择这个的概率是 $1/12$ 。第二个分析的概率是组合的子树的概率的结果，由此是 $1/12 \times 1/12 = 1/144$ 。(62b) 的分析的概率就是 $1/12 + (1/12 \times 1/12) = 13/144$ 。我们可以就此按照相同方式计算 (62a) 中的树的概率。这里唯一的区别是 [the dog] 的树在子树集合中出现了两次。由此，它的概率是 $2/12$ 。[[the dog] barks] 这棵树的概率就是： $1/12 + (1/12 \times 2/12) = 14/144$ 。这样，我们就从语料中提取了貌似正确的结构知识。这个知识也可以在人们听到一个没有完整树的新的语句时使用。然后，就有可能使用已经知道的子树来计算

这个新语句的可能分析的概率了。Bod 的模型也考虑了权重：那些说话者很长时间以前听到的句子会得到较低的权重。据此，我们也可以解释这样的事实，儿童并不是使用他们同时可听到的所有句子。这个扩展使得 UDOP 模型对于语言习得来说更为可信了。

在上面的例子中，我们没有给词指派范畴信息。如果我们这么做了，就会得到图 13.5 中的树作为子树。subtree。如果我们想获得在给定树的不同子树间出现成

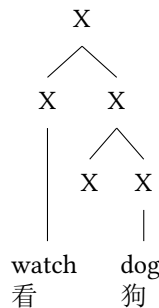


图 13.5: 非连续的部分树

分的依存关系，这些非连续的子树是很重要的。比如说下面的句子：

- (63) a. BA carried *more people than* cargo in 2005.
波音运载 更多人 比 货 在 2005
‘在 2005 年，波音公司运载的人比货多。’
- b. *What’s this scratch doing on the table?*
什么.AUX 这 划痕 做 在 DET 桌子
‘桌子上怎么有个划痕？’
- c. Most software *companies* in Vietnam *are* small sized.
大多数软件 公司 在越南 COP 小 型号
‘大多数越南的软件公司都是小型的。’

然后，也有可能用这些非连续的树来学习英语的助动词变换。为了能得到正确的句子 (65a)，而不是错误的句子 (65b)，我们所需要的是 (64) 中的两个句子的树结构。

- (64) a. The man who is eating is hungry.
DET 人 CONJ AUX 吃 COP 饿
‘正吃东西的那个人饿了。’
- b. Is the boy hungry?
COP DET 男孩 饿
‘男孩饿了吗？’

- (65) a. Is the man who is eating hungry?
COP DET 人 CONJ AUX 吃 饿
‘吃东西的那个人饿了吗?’
- b. *Is the man who eating is hungry?
COP DET 人 CONJ 吃 AUX 饿

U-DOP 可以从 (66) 中的句子学会图 13.6 on the next page 中 (64) 的结构。

- (66) a. The man who is eating mumbled.
DET 人 CONJ AUX 吃 嘟囔
‘正吃东西的那个人嘟囔了。’
- b. The man is hungry.
DET 人 COP 饿
‘那个人饿了。’
- c. The man mumbled.
DET 人 嘟囔
‘那个人嘟囔了。’
- d. The boy is eating.
DET 男孩AUX 吃
‘男孩正在吃。’

请注意这些句子不包括 (65a) 中的任何结构的实例。根据这里学会的结构，有可能证明助词的位置的最短推导过程也有可能是正确的：正确的语序是 *Is the man who is eating hungry?* 只需要图 13.7 on page 455 中的部分被组合起来，而 **Is the man who eating is hungry?* 的结构需要将图 13.6 中的至少四棵子树互相组合起来。这如图 13.8 on page 455 所示。

我们总是选择那些包括最少子树的动机是因为我们对已知的材料进行最大化的类比。

包括一个助词的 (67) 的树也可以从只有两棵子树（带有 $[_X \text{ is}_X X]$ 的树和 *The man who is eating is hungry* 的整棵树）的图 13.6 中得到。

- (67) *Is the man who is eating is hungry?
COP DET 人 CONJ AUX 吃 COP 饿

有趣的是，儿童确实会造出这类错误的句子 (Crain & Nakayama 1987: 530、Ambridge, Rowland & Pine 2008)。但是，如果我们考虑到子树加上所组合的部分的数量的概率，我们就会得到正确的结果，即 (65a)，而不是 (67)。这是因为 *the man who is eating* 在语料库中出现了两次，一个是在 (65a) 中，一个是在 (66a) 中。所以，*the man who is eating* 的概率跟 *the man who is eating is hungry* 的概率一样高，这样就得到图 13.7 中的推导式，而不是 (67) 中的推导式。这适用于这里构建的例子，但是，我们可以想象在一个真实的语料库中，具有 *the man who is eating* 形式的序列比具有更多词的序列更为常用，因为 *the man who is eating* 也可以在其他语境中出现。Bod 将这个过程用到了成人语言的语料库（英语、德语和汉语）上，也用到

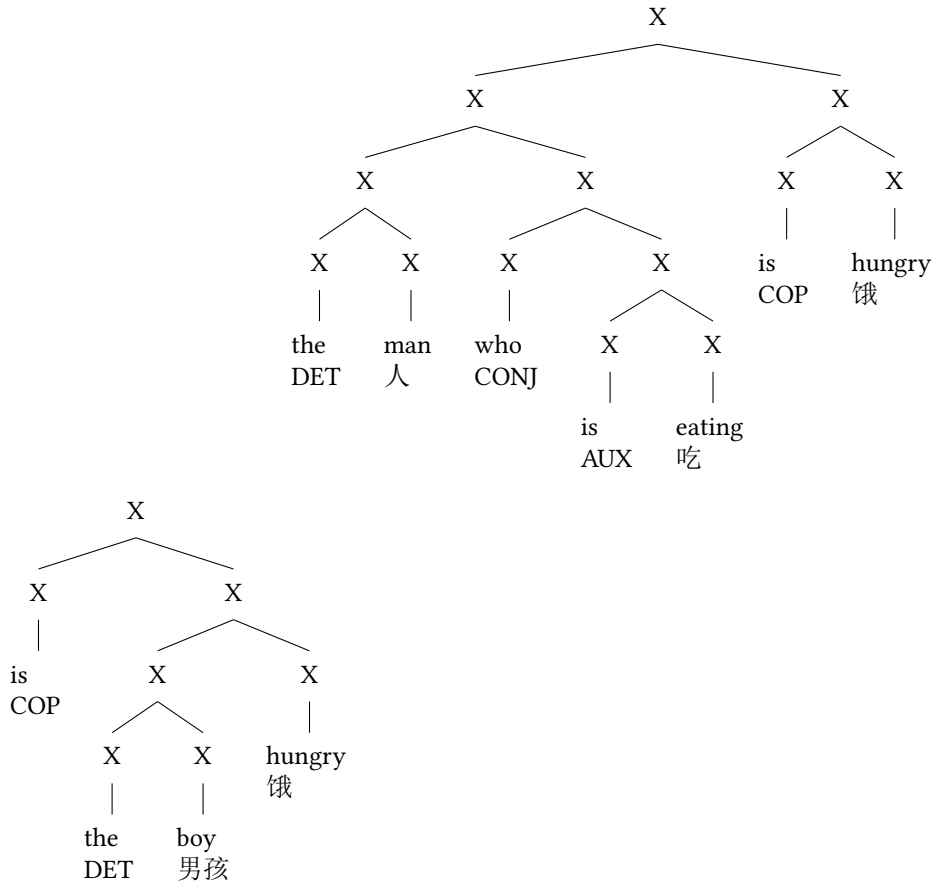


图 13.6: 从 (64) 和 (66) 中的例子学会的 U-DOP 的结构

了 CHILDES 数据库的 Eve 语料上，他这样做是为了考察类比构造是否构成了人类语言习得的模型。他可以证明，我们上面展示的例子也适用于大规模的自然语言的语料库：尽管在 Eve 语料中没有跨越复杂 NP 的助词移位的例子，有可能通过类比来学会复杂 NP 内的助词是不能前置的。

这样就有可能从没有关于词类或语言的抽象属性的先验知识的语料库中学会句法结构。Bod 所作的唯一一个假设是，（二叉）结构是存在的。二叉性的假设并不是必要的。但是如果我们把平铺结构囊括进计算中，树的集合就会变得相当大。这样，Bod 在他的实验中只使用了二叉的结构。在他的树中，X 包括两个其他的 X's 或一个词。我们就在分析循环结构。由此，Bod 的工作提出了只需要循环的句法结构的习得理论，而循环被 Hauser, Chomsky & Fitch (2002) 看作是语言的基本属性。

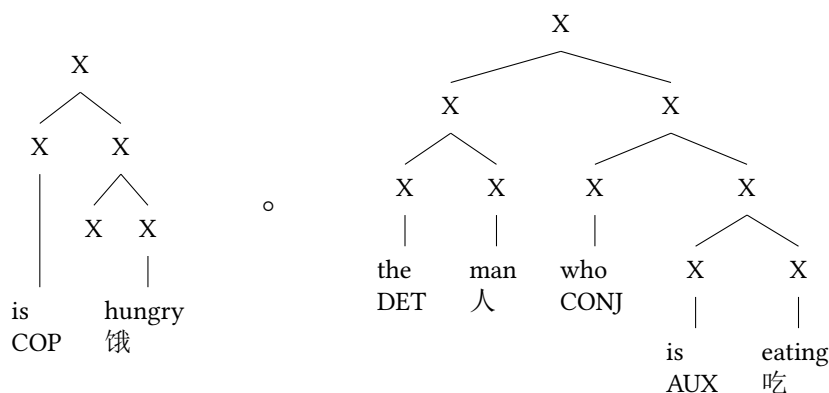


图 13.7: 使用图 13.6 中的两棵子树的助词组合的正确结构的推导过程

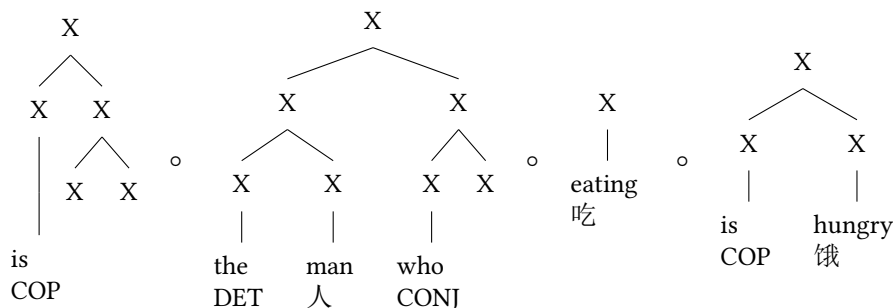


图 13.8: 使用图 13.6 中的两棵子树的助词组合的错误结构的推导过程

正如在第 13.1.8 节展示的，有证据显示循环并不限于语言，这样我们可以总结出，为了能够从现有输入中学会句法结构，我们没有必要假设内在的语言学知识。

尽管如此，这里有必要指出的是：Bod 证明的是句法结构是可以学习的。在他的结构中还没有总结的所涵盖的每个词的词性信息也可以用统计的方法推导出来 (Redington et al. 1998; Clark 2000)。⁴² 在所有的可能性中，可以学习的结构对应于表层导向的语言学理论也会假设的结构。但是，并不是所有的语言学分析都是必要的。在 Bod 的模型中，只考察了结构中词的出现情况。没有说明词之间是否具有一个具体的常规关系（例如，连接被动分词和完成时分词的词汇规则）。进而，没有说明表达式的寓意是如何表示的（他们是按照构式语法所说的整体含义还是从词汇中投射的？）。还有跟理论语言学（请看第 ??）有关的问题，并且不能直接从词的统计分布以及由其计算出的结构中推导出来（关于这点的更多内容请看第 21.8.1 节）。

⁴² 用来区分词类的计算语言学的算法是考察整个语料的。但是儿童总是处理其中的一部分。那么，相应的学习过程也一定会包括一个记忆的曲线。请参考 Braine (1987: 67)。

还需要指出的是：我们已经看到统计信息可以用来推导出复杂的语言表达式的结构。现在，问题的实质是这跟乔姆斯基早期反对的统计学方法有什么关系 (Chomsky 1957: 16)。Abney (1996: Section 4.2) 详细地讨论了这点。乔姆斯基的早期观点的问题是 Chomsky 讨论的是马尔科夫模型。这是有限状态自动机的统计版本。有限状态自动机只能描写 3 型语言，这样就不适合分析自然语言了。但是，Chomsky 的评论不能适用于普遍的统计学方法。

13.8.4 负向证据

在一些认为支持天赋知识的研究中，他们提出儿童没有接触到负向证据，即没有人告诉他们诸如 (44e) 的句子——这里重复为 (68)——是不合乎语法的 (Brown & Hanlon 1970: 42–52, Marcus 1993)。

- (68) *Is the dog that in the corner is hungry?
COP DET 狗 CONJ PREP DET 角落 COP 饿

确实是，大人们不会用不合乎语法的句子来每天叫醒他们的孩子，但是，儿童实际上能够接触到不同类型的负向证据。例如，Chouinard & Clark (2003) 指出，说英语和法语的父母会纠正孩子们不合乎语法的句子。例如，他们重复那些没有对动词正确变位的句子。儿童可以从重复的话语中和重复中修正错误的变化中推导出事实，并且 Chouinard 和 Clark 也证实了，他们确实是这样做的。作者们观察了五个儿童的数据，他们的父母都有学术背景。他们还讨论了其他文化中父母和子女之间的关系（相关概述请参考 Ochs (1982); Ochs & Schieffelin (1985) 和 Marcus (1993: 71)），并且涉及了低社会经济阶层的美国家庭的情况（第 660 页）。

负向证据的深层形式是间接的负向证据，Chomsky (1981a: 9) 提出这也在习得中起到了作用。Goldberg (1995: Section 5.2) 举出 (69a) 中的句子作为例子：⁴³

- (69) a. Look! The magician made the bird disappear.
看 DET 魔术师 使得 DET 鸟 消失
‘看！魔术师把鸟变没了。’
b. *The magician disappeared the bird.
DET 魔术师 消失 DET 鸟

儿童可以从大人更多地使用 *make* 和动词 *disappear* 构成的致使构式的事实得出结论，其他动词，如 *melt*，不能像及物动词那样使用。间接负向证据起到作用的直接例子来自形态学。有一些产生式规则无论如何无法使用，如果有词限制了这个规则的应用的话。一个例子是德语中的名词化后缀 *-er*。通过在动词词干上加上 *-er*，我们可以得到一个名词，它指（通常是习惯上）采取某种行动的人（*Raucher* “吸烟者”，*Maler* “画家”，*Sänger* “歌手”，*Tänzer* “舞者”）。但是，*Stehler*（“小偷”）是非常特殊的。*Stehler*（“小偷”）的构成被 *Dieb*（“贼”）的存在限制住了。由此，语言学习者必须从 *Stehler*（“小偷”）的缺失中推导出的名词化规则并不适用于 *stehlen*（“偷”）。

相似地，对于方式副词的位置没有任何限制的英语语法而言，具有这种知识的说话者会认为 (70) 中的所有语序都是可能的 (Scholz & Pullum 2002: 206):

⁴³ 也请参考 Tomasello (2006a: 277)。

- (70) a. call the police immediately
 叫 DET 警察 直接
 ‘直接叫警察’
 b. *call immediately the police
 叫 直接 DET 警察

学习者可以从这样的事实中间接地得出结论，诸如 (70b) 的动词短语（几乎）从不出现在不属于这个语言的输入中。这可以用相关的统计学习的算法进行模拟。

截至目前，负向证据提供的例子更多的是貌似正确的论断。Stefanowitsch (2008) 将语料库语言学的研究整合进可接受性实验的统计分布中，并且证明了所期待的频率中获得的负向证据跟说话者的可接受性的判断有关。我们在下面简短地讨论一下这个过程：Stefanowitsch 提出了下面的原则：

- (71) 根据语言特征或要素出现的个体频率来构造出他们的共现频率规律，并且根据共现的实际频率来核查这些规律。(Stefanowitsch 2008: 518)

Stefanowitsch 研究的是包括英式英语的英语国际语料库 (*International Corpus of English*, ICE-GB)。在这个语料库中，动词 *say* 出现了 3333 次，带有双及物动词的句子 (Subj Verb Obj Obj) 出现了 1824 次。数据库中动词的总数达 136551 个。如果所有的动词在所有类型的句子中以相同的频率出现的话，那么我们会希望得到这样的结果，*say* 在双及物构式中出现了 44.52 次 ($X / 1,824 = 3,333 / 136,551$ 所以 $X = 1,824 \times 3,333 / 136,551$)。但是，这个表达式出现的实际数量是 0，跟 (72b) 不同，说英语的人不适用于 (72a) 这样的句子。

- (72) a. *Dad said Sue something nice.
 爸爸说 Sue 某事 好
 b. Dad said something nice to Sue.
 爸爸说 某事 好 INF Sue
 ‘爸爸跟 Sue 说了某件好事。’

Stefanowitsch 证明了双及物句式中 *say* 没有出现是非常重要的。进而，他考察了可接受性判断是如何跟特定构式中动词不出现的频率相比较的。在第一个实验中，他能够证明特殊构式中不出现要素的频率与说话者的可接受性判断有关，而跟构式中动词的出现频率无关。

总之，我们可以说间接的负向证据可以从语言输入中推导出来，而且它在语言习得中起到了重要的作用。

13.9 总结

综上所述，没有一个支持内在的语言知识的论断是没有争议的。这当然不会排除有内在的语言知识的可能性，但是那些希望将这个假说整合进他们的理论的学者们要比之前想要证明他们假设的内在性真正属于我们的语言知识并且能够只从语言输入学习时更为小心了。

思考题

1. 哪些学说假设了天赋的语言学知识？

延伸阅读

Pinker 的 (1994)这本书是关于语言的天赋模型的最好的一本书。

Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett (1996)讨论了支持天赋语言知识的所有观点，并且证明了相关的现象可以有不同的解释。作者们采用了联结主义的观点。他们用神经网络进行研究，神经网络被认为是相对准确地模拟我们大脑工作过程的方法。这本书也包括了遗传学的基本知识和大脑结构的章节内容，并深入讨论了为什么将语言知识直接编码进我们的基因组是不可能的。

有些使用了神经网络的方法遭到了批评，这是因为他们无法捕捉到人类能力的某些方法，如循环或者话语中相同词的多次使用。Pulvermüller (2010)讨论了一个具有记忆的架构，并且使用它来分析循环结构。在他的概述性文章中，引用了一些研究来证明许多更为抽象的规则或理论语言学中认为理所当然的模式都可以在神经层面进行证明。但是，Pulvermüller 并不认为语言知识是内在的（第 173 页）。

Pullum 和 Scholz 详细地分析了刺激贫乏论 (Pullum & Scholz 2002; Scholz & Pullum 2002)。

Goldberg (2006)和Tomasello (2003)是最为著名的构式语法学家，构式语法明确地不支持天赋语言知识的假说。

第十四章 **Generative-enumerative vs. model-theoretic approaches**

Generative-enumerative approaches assume that a grammar generates a set of sequences of symbols (strings of words). This is where the term Generative Grammar comes from. Thus, it is possible to use the grammar on page 49, repeated here as (1), to derive the string *er das Buch dem Mann gibt* ‘he the book the man gives’.

(1)	$NP \rightarrow D, N$	$NP \rightarrow er$	$N \rightarrow Buch$
	$S \rightarrow NP, NP, NP, V$	$D \rightarrow das$	$N \rightarrow Mann$
		$D \rightarrow dem$	$V \rightarrow gibt$

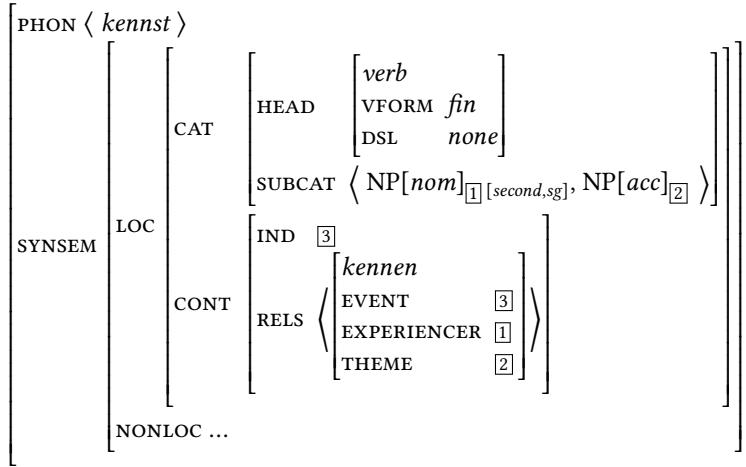
Beginning with the start symbol (S), symbols are replaced until one reaches a sequence of symbols only containing words. The set of all strings derived in this way is the language described by the grammar.

The following are classed as generative-enumerative approaches:

- all phrase structure grammars
- Transformational Grammars in almost all variants
- GPSG in the formalism of Gazdar, Klein, Pullum & Sag (1985)
- many variants of Categorical Grammar
- many variants of TAG
- Chomsky’s Minimalist Grammars

LFG was also originally designed to be a generative grammar.

The opposite of such theories of grammar are model-theoretic or constraint-based approaches (MTA). MTAs formulate well-formedness conditions on the expressions that the grammar describes. In Section 6.7, we already discussed a model-theoretic approach for theories that use feature structures to model phenomena. To illustrate this point, I will discuss another HPSG example: (2) shows the lexical item for *kennst* ‘know’. In the description of (2), it is ensured that the PHON value of the relevant linguistic sign is $\langle kennst \rangle$, that is, this value of PHON is constrained. There are parallel restrictions for

(2) Lexical item for *kennst*:

the features given in (2): the SYNSEM value is given. In SYNSEM, there are restrictions on the LOC and NONLOC value. In CAT, there are individual restrictions for HEAD and SUBCAT. The value of SUBCAT is a list with descriptions of dependent elements. The descriptions are given as abbreviations here, which actually stand for complex feature descriptions that also consist of feature-value pairs. For the first argument of *kennst*, a HEAD value of type *noun* is required, the PER value in the semantic index has to be *second* and the NUM value has to be *sg*. The structure sharings in (2) are a special kind of constraint. Values that are not specified in the descriptions of lexical entries can vary in accordance with the feature geometry given by the type system. In (2), neither the SLASH value of the nominative NP nor the one of the accusative NP is fixed. This means that SLASH can either be an empty or non-empty list.

The constraints in lexical items such as (2) interact with further constraints that hold for the signs of type *phrase*. For instance, in head-argument structures, the non-head daughter must correspond to an element from the SUBCAT list of the head daughter.

Generative-enumerative and model-theoretic approaches view the same problem from different sides: the generative side only allows what can be generated by a given set of rules, whereas the model-theoretic approach allows everything that is not ruled out by constraints.¹

Pullum & Scholz (2001: 19–20) and Pullum (2007) list the following model-theoretic approaches:²

- the non-procedural variant of Transformational Grammar of Lakoff, that formu-

¹ Compare this to an old joke: in dictatorships, everything that is not allowed is banned, in democracies, everything that is not banned is allowed and in France, everything that is banned is allowed. Generative-enumerative approaches correspond to the dictatorships, model-theoretic approaches are the democracies and France is something that has no correlate in linguistics.

² See Pullum (2007) for a historical overview of Model Theoretic Syntax (MTS) and for further references.

lates constraints on potential tree sequences,

- Johnson and Postal’s formalization of Relational Grammar (1980),
- GPSG in the variants developed by Gazdar et al. (1988), Blackburn et al. (1993) and Rogers (1997),
- LFG in the formalization of Kaplan (1995)³ and
- HPSG in the formalization of King (1999).

Categorial Grammars (Bouma & van Noord 1994), TAG (Rogers 1994) and Minimalist approaches (Veenstra 1998) can be formulated in model-theoretic terms.

Pullum & Scholz (2001) point out various differences between these points of view. In the following sections, I will focus on two of these differences.⁴ Section 14.3 deals with ten Hacken’s objection to the model-theoretic view.

14.1 Graded acceptability

Generative-enumerative approaches differ from model-theoretic approaches in how they deal with the varying degrees of acceptability of utterances. In generative-enumerative approaches, a particular string is either included in the set of well-formed expressions or it is not. This means that it is not straightforwardly possible to say something about the degree of deviance: the first sentence in (3) is judged grammatical and the following three are equally ungrammatical.

- (3)
- a. Du kennst diesen Aufsatz.
you know.2SG this.ACC essay
 - b. *Du kennen diesen Aufsatz.
you know.3PL this.ACC essay
 - c. *Du kennen dieser Aufsatz.
you know.3PL this.NOM essay
 - d. *Du kennen Aufsatz dieser.
you know.3PL essay this.NOM

At this point, critics of this view raise the objection that it is in fact possible to determine degrees of acceptability in (3b–d): in (3b), there is no agreement between the subject and the verb, in (3c), *dieser Aufsatz* ‘this essay’ has the wrong case in addition, and in (3d), *Aufsatz* ‘essay’ and *dieser* ‘this’ occur in the wrong order. Furthermore, the sentence in (4) violates grammatical rules of German, but is nevertheless still interpretable.

³ According to Pullum (2013: Section 3.2), there seems to be a problem for model-theoretic formalizations of so-called *constraining equations*.

⁴ The reader should take note here: there are differing views with regard to how generative-enumerative and MTS models are best formalized and not all of the assumptions discussed here are compatible with every formalism. The following sections mirror the important points in the general discussion.

- (4) Studenten stürmen mit Flugblättern und Megafon die Mensa und rufen
 students storm with flyers and megaphone the canteen and call
 alle auf zur Vollversammlung in der Glashalle zum kommen. Vielen bleibt
 all up to plenary.meeting in the glass.hall to.the come many.DAT stays
 das Essen im Mund stecken und kommen sofort mit.⁵
 the food in.the mouth stick and come immediately with
 ‘Students stormed into the university canteen with flyers and a megaphone call-
 ing for everyone to come to a plenary meeting in the glass hall. For many, the
 food stuck in their throats and they immediately joined them.’

Chomsky (1975: Chapter 5; 1964b) tried to use a string distance function to determine the relative acceptability of utterances. This function compares the string of an ungrammatical expression with that of a grammatical expression and assigns an ungrammaticality score of 1, 2 or 3 according to certain criteria. This treatment is not adequate, however, as there are much more fine-grained differences in acceptability and the string distance function also makes incorrect predictions. For examples of this and technical problems with calculating the function, see Pullum & Scholz (2001: 29).

In model-theoretic approaches, grammar is understood as a system of well-formedness conditions. An expression becomes worse, the more well-formedness conditions it violates (Pullum & Scholz 2001: 26–27). In (3b), the person and number requirements of the lexical item for the verb *kennst* are violated. In addition, the case requirements for the object have not been fulfilled in (3c). There is a further violation of a linearization rule for the noun phrase in (3d).

Well-formedness conditions can be weighted in such a way as to explain why certain violations lead to more severe deviations than others. Furthermore, performance factors also play a role when judging sentences (for more on the distinction between performance and competence, see Chapter 15). As we will see in Chapter 15, constraint-based approaches work very well as performance-compatible grammar models. If we combine the relevant grammatical theory with performance models, we will arrive at explanations for graded acceptability differences owing to performance factors.

14.2 Utterance fragments

Pullum & Scholz (2001: Section 3.2) point out that generative-enumerative theories do not assign structure to fragments. For instance, neither the string *and of the* nor the string *the of and* would receive a structure since none of these sequences is well-formed as an utterance and they are therefore not elements of the set of sequences generated by the grammar. However, *and of the* can occur as part of the coordination of PPs in sentences such as (5) and would therefore have some structure in these cases, for example the one given in Figure 14.1 在下一頁.

- (5) That cat is afraid of the dog and of the parrot.

⁵ Streikzeitung der Universität Bremen, 04.12.2003, p. 2. The emphasis is mine.

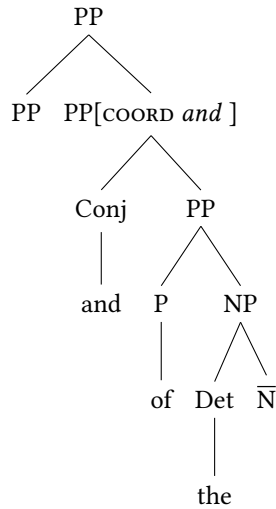


图 14.1: Structure of the fragment *and of the* following Pullum & Scholz (2001: 32)

As a result of the interaction of various constraints in a constraint-based grammar, it emerges that *the* is part of an NP and this NP is an argument of *of* and furthermore *and* is combined with the relevant *of*-PP. In symmetric coordination, the first conjunct has the same syntactic properties as the second, which is why the partial structure of *and of the* allows one to draw conclusions about the category of the conjunct despite it not being part of the string.

Ewan Klein noted that Categorical Grammar and Minimalist Grammars, which build up more complex expressions from simpler ones, can sometimes create this kind of fragments (Pullum 2013: 507). This is certainly the case for Categorical Grammars with composition rules, which allow one to combine any sequence of words to form a constituent. If one views derivations as logical proofs, as is common in some variants of Categorical Grammar, then the actual derivation is irrelevant. What matters is whether a proof can be found or not. However, if one is interested in the derived structures, then the argument brought forward by Pullum and Scholz is still valid. For some variants of Categorical Grammar that motivate the combination of constituents based on their prosodic and information-structural properties (Steedman 1991: Section 3), the problem persists since fragments have a structure independent of the structure of the entire utterance and independent of their information-structural properties within this complete structure. This structure of the fragment can be such that it is not possible to analyze it with type-raising rules and composition rules.

In any case, this argument holds for Minimalist theories since it is not possible to have a combination of *the* with a nominal constituent if this constituent was not already built up from lexical material by Merge.

14.3 A problem for model-theoretic approaches?

Ten Hacken (2007: 237–238) discusses the formal assumptions of HPSG. In HPSG, feature descriptions are used to describe feature structures. Feature structures must contain all the features belonging to a structure of a certain type. Additionally, the features have to have a maximally-specific value (see Section 6.7). Ten Hacken discusses gender properties of the English noun *cousin*. In English, gender is important in order to ensure the correct binding of pronouns (see page 251 for German):

- (6) a. The man_{*i*} sleeps. He_{*i*} snores.
- b. The woman_{*i*} sleeps. He_{**i*} snores.

While *he* in (6a) can refer to *man*, *woman* is not a possible antecedent. Ten Hacken's problem is that *cousin* is not marked with respect to gender. Thus, it is possible to use it to refer to both male and female relatives. As was explained in the discussion of the case value of *Frau* 'woman' in Section 6.7, it is possible for a value in a description to remain unspecified. Thus, in the relevant feature structures, any appropriate and maximally specific value is possible. The case of *Frau* can therefore be nominative, genitive, dative or accusative in an actual feature structure. Similarly, there are two possible genders for *cousin* corresponding to the usages in (7).

- (7) a. I have a cousin_{*i*}. He_{*i*} is very smart.
- b. I have a cousin_{*i*}. She_{*i*} is very smart.

Ten Hacken refers to examples such as (8) and claims that these are problematic:

- (8) a. Niels has two cousins.
- b. How many cousins does Niels have?

In plural usage, it is not possible to assume that *cousins* is feminine or masculine since the set of relatives can contain either women or men. It is interesting to note that (9a) is possible in English, whereas German is forced to use (9b) to express the same meaning.

- (9) a. Niels and Odette are cousins.
- b. Niels und Odette sind Cousin und Cousine.
 Niels and Odette are cousin.M and cousin.F

Ten Hacken concludes that the gender value has to remain unspecified and this shows, in his opinion, that model-theoretic analyses are unsuited to describing language.

If we consider what exactly ten Hacken noticed, then it becomes apparent how one can account for this in a model-theoretic approach: Ten Hacken claims that it does not make sense to specify a gender value for the plural form of *cousin*. In a model-theoretic approach, this can be captured in two ways. One can either assume that there are no gender features for referential indices in the plural, or that one can add a gender value that plural nouns can have.

The first approach is supported by the fact that there are no inflectional differences between the plural forms of pronouns with regard to gender. There is therefore no reason to distinguish genders in the plural.

- (10) a. Niels and Odette are cousins. They are very smart.
 b. The cousins/brothers/sisters are standing over there. They are very smart.

No distinctions are found in plural when it comes to nominal inflection (*brothers, sisters, books*). In German, this is different. There are differences with both nominal inflection and the reference of (some) noun phrases with regard to the sexus of the referent. Examples of this are the previously mentioned examples *Cousin* ‘male cousin’ and *Cousine* ‘female cousin’ as well as forms with the suffix *-in* as in *Kindergärtnerin* ‘female nursery teacher’. However, gender is normally a grammatical notion that has nothing to do with sexus. An example is the neuter noun *Mitglied* ‘member’, which can refer to both female and male persons.

The question that one has to ask when discussing Ten Hacken’s problem is the following: does gender play a role for pronominal binding in German? If this is not the case, then the gender feature is only relevant within the morphology component, and here the gender value is determined for each noun in the lexicon. For the binding of personal pronouns, there is no gender difference in German.

- (11) Die Schwestern / Brüder / Vereinsmitglieder / Geschwister stehen dort.
 the sisters.F brothers.M club.members.N siblings stand there
 Sie lächeln.
 they smile.
 ‘The sisters/brothers/club members/siblings are standing there. They are smiling.’

Nevertheless, there are adverbials in German that agree in gender with the noun to which they refer (Höhle 1983: Chapter 6):

- (12) a. Die Fenster wurden eins nach dem anderen geschlossen.
 the windows.N were one.N after the other closed
 ‘The windows were closed one after the other.’
 b. Die Türen wurden eine nach der anderen geschlossen.
 the doors.F were one.F after the other closed
 ‘The doors were closed one after the other.’
 c. Die Riegel wurden einer nach dem anderen zugeschoben.
 the bolts.M were one.M after the other closed
 ‘The bolts were closed one after the other.’

For animate nouns, it is possible to diverge from the gender of the noun in question and use a form of the adverbial that corresponds to the biological sex:

- (13) a. Die Mitglieder des Politbüros wurden eines / einer nach dem anderen
 the members.N of.the politburo were one.N one.M after the other
 aus dem Saal getragen.
 out.of the hall carried
 ‘The members of the politburo were carried out of the hall one after the other.’

- b. Die Mitglieder des Frauentanzklubs verließen eines / eine nach
 the members.N of.the women's.dance.club left one.N one.F after
 dem / der anderen im Schutze der Dunkelheit den Keller.
 the.N the.F other in.the protection of.the dark the basement
 'The members of the women's dance club left the basement one after the
 other under cover of darkness.'

This deviation from gender in favor of sexus can also be seen with binding of personal and relative pronouns with nouns such as *Weib* 'woman' (pej.) and *Mädchen* 'girl':

- (14) a. „Farbe bringt die meiste Knete!“ verriet ein 14jähriges türkisches *Mädchen*,
 color brings the most money revealed a 14-year.old Turkish girl.N
die die Mauerstückchen am Nachmittag am Checkpoint Charlie an
 who.F the wall.pieces in.the afternoon at Checkpoint Charlie at
 Japaner und US-Bürger verkauft.⁶
 Japanese and US-citizens sells
 '“Color gets the most money” said a 14-year old Turkish girl who sells pieces
 of the wall to Japanese and American citizens at Checkpoint Charlie.'
- b. Es ist ein junges *Mädchen*, *die* auf der Suche nach CDs bei Bolzes
 it is a young girl.N who.F on the search for CDs at Bolzes
 reinschaut.⁷
 stops.by
 'It is a young girl looking for CDs that stops by Bolzes.'

For examples from Goethe, Kafka and Thomas Mann, see Müller (1999a: 417–418).

For inanimate nouns such as those in (12), agreement is obligatory. For the analysis of German, one therefore does in fact require a gender feature in the plural. In English, this is not the case since there are no parallel examples with pronouns inflecting for gender. One can therefore either assume that plural indices do not have a gender feature or that the gender value is *none*. In the latter case, the feature would have a value and hence fulfill the formal requirements. (15) shows the first solution: plural indices are modeled by feature structures of type *pl-ind* and the GENDER feature is just not appropriate for such objects.

- (15) a. singular index: b. plural index:
- | | |
|---|--|
| $\begin{array}{l} sg-ind \\ PER \text{ } per \\ NUM \text{ } sg \\ GEN \text{ } gender \end{array}$ | $\begin{array}{l} pl-ind \\ PER \text{ } per \\ NUM \text{ } pl \end{array}$ |
|---|--|

The second solution requires the type hierarchy in Figure 14.2 在下一頁 for the subtypes of *gender*. With such a type hierarchy *none* is a possible value of the GEN feature and no problem will arise.

⁶ taz, 14.06.1990, p. 6.

⁷ taz, 13.03.1996, p. 11.

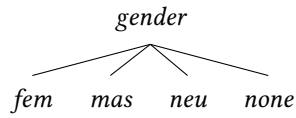


图 14.2: Type hierarchy for one of the solutions of ten Hacken's problem

In general, it is clear that cases such as the one constructed by ten Hacken will never be a problem since there are either values that make sense, or there are contexts for which there is no value that makes sense and one therefore does not require the features.

So, while ten Hacken's problem is a non-issue, there are certain problems of a more technical nature. I have pointed out one such technical problem in Müller (1999a: Section 14.4). I show that spurious ambiguities arise for a particular analysis of verbal complexes in German when one resolves the values of a binary feature (FLIP). I also show how this problem can be avoided by the complicated stipulation of a value in certain contexts.

第十五章 The competence/performance distinction

The distinction between competence and performance (Chomsky 1965: Section 1.1), which is assumed by several theories of grammar, was already discussed in Section 12.6.3 about the analysis of scrambling and verbal complexes in TAG. Theories of competence are intended to describe linguistic knowledge and performance theories are assigned the task of explaining how linguistic knowledge is used as well as why mistakes are made in speech production and comprehension. A classic example in the competence/performance discussion are cases of center self-embedding. Chomsky & Miller (1963: 286) discuss the following example with recursively embedded relative clauses:

(1) (the rat (the cat (the dog chased) killed) ate the malt)

(2b) is a corresponding example in German:

- (2) a. dass der Hund bellt, der die Katze jagt, die die Maus kennt, die
that the dog.M barks that.M the cat chases that.F the mouse knows who
im Keller lebt
in.the basement lives
‘that the dog that chases the cat that knows the mouse who is living in the
basement is barking’
- b. dass er Hund, [₁ der die Katze, [₂ die die Maus, [₃ die im Keller
that the dog that the cat that the mouse who in.the basement
lebt, ₃] kennt, ₂] jagt ₁] bellt
lives knows chases barks

The examples in (1) and (2b) are entirely incomprehensible for most people. If one rearranges the material somewhat, it is possible to process the sentences and assign a meaning to them.¹ For sentences such as (2b), it is often assumed that they fall within

¹ The sentence in (2a) can be continued following the pattern that was used to create the sentence. For instance by adding *die unter der Treppe lebte, die meine Freunde repariert haben* ‘who lived under the staircase which my friends repaired’. This shows that a restriction of the number of elements that depend on one head to seven (Leiss 2003: 322) does not restrict the set of the sentences that are generated or licensed by a grammar to be finite. There are at most two dependents of each head in (2a). The extraposition of the relative clauses allows the hearer to group material into processable and reducible chunks, which reduces the cognitive burden during processing.

This means that the restriction to seven dependents does not cause a finitization of recursion (“Ver-

our grammatical competence, that is, we possess the knowledge required to assign a structure to the sentence, although the processing of utterances such as (2b) exceeds language-independent abilities of our brain. In order to successfully process (2b), we would have to retain the first five noun phrases and corresponding hypotheses about the further progression of the sentence in our heads and could only begin to combine syntactic material when the verbs appear. Our brains become overwhelmed by this task. These problems do not arise when analyzing (2a) as it is possible to immediately begin to integrate the noun phrases into a larger unit.

Nevertheless, center self-embedding of relative clauses can also be constructed in such a way that our brains can handle them. Hans Uszkoreit (p. c. 2009) gives the following example:

- (3) Die Bänke, [₁ auf denen damals die Alten des Dorfes, [₂ die allen the benches on which back.then the old.people of.the village that all Kindern, [₃ die vorbeikamen₃], freundliche Blicke zuwarfen₂], lange Stunden children that came.by friendly glances gave long hours schweigend nebeneinander saßen₁], mussten im letzten Jahr einem silent next.to.each.other sat must in.the last year a Parkplatz weichen.
car.park give.way.to
'The benches on which the older residents of the village, who used to give friendly glances to all the children who came by, used to sit silently next to one another for hours had to give way to a car park last year.'

Therefore, one does not wish to include in the description of our grammatical knowledge that relative clauses are not allowed to be included inside each other as in (2b) as this would also rule out (3).

We can easily accept the fact that our brains are not able to process structures past a certain degree of complexity and also that corresponding utterances then become unacceptable. The contrast in the following examples is far more fascinating:²

- (4) a. # The patient [who the nurse [who the clinic had hired] admitted] met Jack.
b. * The patient who the nurse who the clinic had hired met Jack.

Although (4a) is syntactically well-formed and (4b) is not, Gibson & Thomas (1999) were able to show that (4b) is rated better by speakers than (4a). It does not occur to some people that an entire VP is missing. There are a number of explanations for this fact, all of which in some way make the claim that previously heard words are forgotten as soon

endlichung von Rekursivität") as was claimed by Leiss (2003: 322). Leiss argued that Miller could not use his insights regarding short term memory, since he worked within Transformational Grammar rather than in Dependency Grammar. The discussion shows that dependency plays an important role, but that linear order is also important for processing.

² See Gibson & Thomas (1999: 227). Frazier (1985: 178) attributes the discovery of this kind of sentences to Janet Fodor.

as new words are heard and a particular degree of complexity is exceeded (Frazier 1985: 178; Gibson & Thomas 1999).

Instead of developing grammatical theories that treat (2b) and (4a) as unacceptable and (3) and (4b) as acceptable, descriptions have been developed that equally allow (2b), (3), and (4a) (competence models) and then additionally investigate the way utterances are processed in order to find out what kinds of structures our brains can handle and what kinds of structures it cannot. The result of this research is then a performance model (see Gibson (1998), for example). This does not rule out that there are language-specific differences affecting language processing. For example, Vasissth, Suckow, Lewis & Kern (2010) have shown that the effects that arise in center self-embedding structures in German are different from those that arise in the corresponding English cases such as (4): due to the frequent occurrence of verb-final structures in German, speakers of German were able to better store predictions about the anticipated verbs into their working memory (p. 558).

Theories in the framework of Categorical Grammar, GB, LFG, GPSG and HPSG are theories about our linguistic competence.³ If we want to develop a grammatical theory that directly reflects our cognitive abilities, then there should also be a corresponding performance model to go with a particular competence model. In the following two sections, I will recount some arguments from Sag & Wasow (2011) in favor of constraint-based theories such as GPSG, LFG and HPSG.

15.1 The derivational theory of complexity

The first point discussed by Sag & Wasow (2011) is the Derivational Theory of Complexity. In the early days of Transformational Grammar, it was assumed that transformations were cognitively real, that is, it is possible to measure the consumption of resources that transformations have. A sentence that requires more transformations than the analysis of another sentence should therefore also be more difficult for humans to process. The corresponding theory was dubbed the *Derivational Theory of Complexity* (DTC) and initial experiments seemed to confirm it (Miller & McKean 1964; Savin & Perchonock 1965; Clifton & Odom 1966), so that in 1968 Chomsky still assumed that the

³ For an approach where the parser is equated with UG, see Abney & Cole (1986: Section 3.4). For a performance-oriented variant of Minimalism, see Phillips (2003).

In Construction Grammar, the question of whether a distinction between competence and performance would be justified at all is controversially discussed (see Section 10.6.8.1.1). Fanselow, Schlesewsky, Cavar & Kliegl (1999) also suggest a model – albeit for different reasons – where grammatical properties considerably affect processing properties. The aforementioned authors work in the framework of Optimality Theory and show that the OT constraints that they assume can explain parsing preferences. OT is not a grammatical theory on its own but rather a meta theory. It is assumed that there is a component GEN that creates a set of candidates. A further component EVAL then chooses the most optimal candidate from this set of candidates. GEN contains a generative grammar of the kind that we have seen in this book. Normally, a GP/MP variant or also LFG is assumed as the base grammar. If one assumes a transformational theory, then one automatically has a problem with the Derivational Theory of Complexity that we will encounter in the following section. If one wishes to develop OT parsing models, then one has to make reference to representational variants of GB as the aforementioned authors seem to.

15 The competence/performance distinction

Derivational Theory of Complexity was in fact correct (Chomsky 1976a: 249–250).⁴ Some

years later, however, most psycholinguists rejected the DTC. For discussion of several experiments that testify against the DTC, see Fodor, Bever & Garrett (1974: 320–328). One set of phenomena where the DTC makes incorrect predictions for respective analyses is that of elliptical constructions, for example (Fodor, Bever & Garrett 1974: 324): in elliptical constructions, particular parts of the utterance are left out or replaced by auxiliaries. In transformation-based approaches, it was assumed that (5b) is derived from (5a) by means of deletion of *swims* and (5c) is derived from (5b) by inserting *do*.

- (5) a. John swims faster than Bob swims.
- b. John swims faster than Bob.
- c. John swims faster than Bob does.

The DTC predicts that (5b) should require more time to process than (5a), since the analysis of (5b) first requires to build up the structure in (5a) and then delete *swims*. This prediction was not confirmed.

Similarly, no difference could be identified for the pairs in (6) and (7) even though one of the sentences, given the relevant theoretical assumptions, requires more transformations for the derivation from a base structure (Fodor, Bever & Garrett 1974: 324).

- (6) a. John phoned up the girl.
- b. John phoned the girl up.
- (7) a. The bus driver was nervous after the wreck.
- b. The bus driver was fired after the wreck.

In (6), we are dealing with local reordering of the particle and the object. (7b) contains a passive clause that should be derived from an active clause under Transformational Grammar assumptions. If we compare this sentence with an equally long sentence with an adjective, like (7a), the passive clause should be more difficult to process. This is, however, not the case.

It is necessary to add two qualifications to Sag & Wasow's claims: if one has experimental data that show that the DTC makes incorrect predictions for a particular analysis, this does not necessarily mean that the DTC has been disproved. One could also try to find a different analysis for the phenomenon in question. For example, instead of a transformation that deletes material, one could assume empty elements for the analysis

⁴ In the Transformational Grammar literature, transformations were later viewed as a metaphor (Lohnstein 2014: 170, also in Chomsky 2001: Footnote 4), that is, it was no longer assumed to have psycholinguistic reality. In *Derivation by phase* and *On phases*, Chomsky refers once again to processing aspects such as computational and memory load (Chomsky 2001: 11, 12, 15; 2007: 3, 12; 2008: 138, 145, 146, 155). See also Marantz (2005: 440) and Richards (2015).

A structure building operation that begins with words and is followed by transformations, as recently assumed by theories in the Minimalist Program, is psycholinguistically implausible for sentence parsing. See Labelle (2007) and Section 15.2 for more on incremental processing.

Chomsky (2007: 6) (written later than *On phases*) seems to adopt a constraint-based view. He writes that “a Merge-based system involves parallel operations” and compares the analysis of an utterance with a proof and explicitly mentions the competence/performance distinction.

of elliptical structures that are inserted directly into the structure without deleting any material (see page 63 for the assumption of an empty nominal head in structures with noun ellipsis in German). Data such as (5) would then be irrelevant to the discussion.⁵ However, reordering such as (6b) and the passive in (7b) are the kinds of phenomena that are typically explained using transformations.

The second qualification pertains to analyses for which there is a representational variant: it is often said that transformations are simply metaphors (Jackendoff 2000: 22–23; 2007: 5, 20): for example, we have seen that extractions with a transformational grammar yield structures that are similar to those assumed in HPSG. Figure 15.1 shows cyclic movement in GB theory compared to the corresponding HPSG analysis.

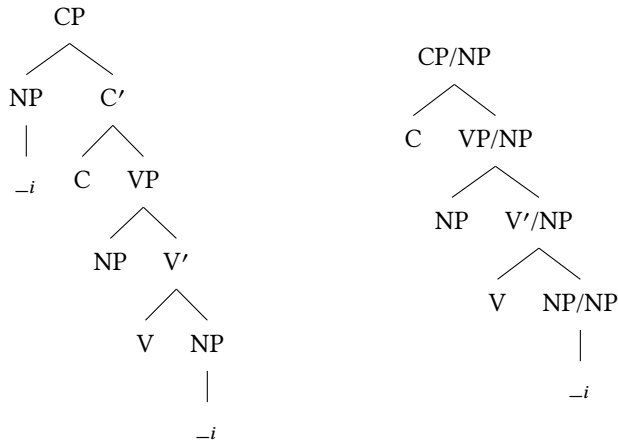


图 15.1: Cyclic movement vs. feature percolation

In GB, an element is moved to the specifier position of CP (SpecCP) and can then be moved from there to the next higher SpecCP position.

- (8) a. Chris_i , we think $[_{CP} \text{ } _i \text{ Anna claims } [_{CP} \text{ } _i \text{ that David saw } _i]]$. (GB)
 b. Chris_i , we think $[_{CP/NP} \text{ Anna claims } [_{CP/NP} \text{ that David saw } _i]]$. (HPSG)

In HPSG, the same effect is achieved by structure sharing. Information about a long-distance dependency is not located in the specifier node but rather in the mother node of the projection itself. In Section 19.2, I will discuss various ways of eliminating empty elements from grammars. If we apply these techniques to structures such as the GB structure in Figure 15.1, then we arrive at structures where information about missing elements is integrated into the mother node (CP) and the position in SpecCP is unfilled. This roughly corresponds to the HPSG structure in Figure 15.1.⁶ It follows from this that

⁵ Culicover & Jackendoff (2005: Chapters 1 and 7) argue in favor of analyzing ellipsis as a semantic or pragmatic phenomenon rather than a syntactic one anyway.

⁶ In Figure 15.1, additionally the unary branching of C' to CP was omitted in the tree on the right so that C combines directly with VP/NP to form CP/NP.

there are classes of phenomena that can be spoken about in terms of transformations without expecting empirical differences with regard to performance when compared to transformation-less approaches. However, it is important to note that we are dealing with an S-structure in the left-hand tree in Figure 15.1. As soon as one assumes that this is derived by moving constituents out of other structures, this equivalence of approaches disappears.

15.2 Incremental processing

The next important point mentioned by Sag & Wasow (2011) is the fact that both comprehension and production of language take places incrementally. As soon as we hear or read even the beginning of a word, we begin to assign meaning and to create structure. In the same way, we sometimes start talking before we have finished planning the entire utterance. This is shown by interruptions and self-correction in spontaneous speech (Clark & Wasow 1998; Clark & Fox Tree 2002). When it comes to processing spoken speech, Tanenhaus et al. (1996) have shown that we access a word as soon as we have heard a part of it (see also Marslen-Wilson 1975). The authors of the study carried out an experiment where participants were instructed to pick up particular objects on a grid and reorganize them. Using eye-tracking measurements, Tanenhaus and colleagues could then show that the participants could identify the object in question earlier if the sound sequence at the beginning of the word was unambiguous than in cases where the initial sounds occurred in multiple words. An example for this is a configuration with a candle and candy: *candy* and *candle* both begin with *can* such that speakers could not yet decide upon hearing this sequence which lexical entry should be accessed. Therefore, there was a slight delay in accessing the lexical entry when compared to words where the objects in question did not contain the same segment at the start of the word (Tanenhaus et al. 1995: 1633).

If complex noun phrases were used in the instructions (*Touch the starred yellow square*), the participants' gaze fell on the object in question 250ms after it was unambiguously identifiable. This means that if there was only a single object with stars on it, then they looked at it after they heard *starred*. In cases where there were starred yellow blocks as well as squares, they looked at the square only after they had processed the word *square* (Tanenhaus et al. 1995: 1632). The planning and execution of a gaze lasts 200ms. From this, one can conclude that hearers combine words directly and as soon as enough information is available, they create sufficient structure in order to capture the (potential) meaning of an expression and react accordingly. This finding is incompatible with models that assume that one must have heard a complete noun phrase or even a complete utterance of even more complexity before it is possible to conclude anything about the meaning of a phrase/utterance. In particular, analyses in the Minimalist Program which assume that only entire phrases or so-called phases⁷ are interpreted (see Chomsky 1999 and also Marantz 2005: 441, who explicitly contrasts the MP to Catego-

⁷ Usually, only CP and vP are assumed to be phases.

rial Grammar) must therefore be rejected as inadequate from a psycholinguistic perspective.^{8,9}

With contrastive emphasis of individual adjectives in complex noun phrases (e.g., *the BIG blue triangle*), hearers assumed that there must be a corresponding counterpart to the reference object, e.g., a small blue triangle. The eye-tracking studies carried out by Tanenhaus et al. (1996) have shown that taking this kind of information into account results in objects being identified more quickly.

Similarly, Arnold et al. (2004) have shown, also using eye-tracking studies, that hearers tend to direct their gaze to previously unmentioned objects if the interlocutor interrupts their speech with *um* or *uh*. This can be traced back to the assumption that hearers assume that describing previously unmentioned objects is more complex than referring to objects already under discussion. The speaker can create more time for himself by using *um* or *uh*.

Examples such as those above constitute evidence for approaches that assume that when processing language, information from all available channels is used and that this information is also used as soon as it is available and not only after the structure of the entire utterance or complete phrase has been constructed. The results of experimental research therefore show that the hypothesis of a strictly modular organization of linguistic knowledge must be rejected. Proponents of this hypothesis assume that the output of one module constitutes the input of another without a given module having access to the inner states of another module or the processes taking place inside it. For example, the morphology module could provide the input for syntax and then this would be processed later by the semantic module. One kind of evidence for this kind of organization of linguistic knowledge that is often cited are so-called *garden path sentences* such as (9):

- (9) a. The horse raced past the barn fell.
- b. The boat floated down the river sank.

The vast majority of English speakers struggle to process these sentences since their parser is led down a garden path as it builds up a complete structure for (10a) or (10b) only then to realize that there is another verb that cannot be integrated into this structure.

- (10) a. The horse raced past the barn.
- b. The boat floated down the river.

⁸ Sternefeld (2006: 729–730) points out that in theories in the Minimalist Program, the common assumption of uninterpretable features is entirely unjustified. Chomsky assumes that there are features that have to be deleted in the course of a derivation since they are only relevant for syntax. If they are not checked, the derivation crashes at the interface to semantics. It follows from this that NPs should not be interpretable under the assumptions of these theories since they contain a number of features that are irrelevant for the semantics and have to therefore be deleted (see Section 4.1.2 of this book and Richards 2015). As we have seen, these kinds of theories are incompatible with the facts.

⁹ It is sometimes claimed that current Minimalist theories are better suited to explain production (generation) than perception (parsing). But these models are as implausible for generation as they are for parsing. The reason is that it is assumed that there is a syntax component that generates structures that are then shipped to the interfaces. This is not what happens in generation though. Usually speakers know what they want to say (at least partly), that is, they start with semantics.

However, the actual structure of (9) contains a reduced relative clause (*raced past the barn* or *floated down the river*). That is the sentences in (9a) are semantically equivalent to the sentences in (11):

- (11) a. The horse that was raced past the barn fell.
b. The boat that was floated down the river sank.

The failure of the parser in these cases was explained by assuming that syntactic processing such as constructing a sentence from NP and VP take place independently of the processing of other constraints. As Crain & Steedman (1985) and others have shown, yet there are data that make this explanation seem less plausible: if (9a) is uttered in a relevant context, the parser is not misled. In (12), there are multiple horses under discussion and each NP is clearly identified by a relative clause. The hearer is therefore prepared for a relative clause and can process the reduced relative clause without being led down the garden path, so to speak.

- (12) The horse that they raced around the track held up fine. The horse that was raced down the road faltered a bit. And the horse raced past the barn fell.

By exchanging lexical material, it is also possible to modify (9a) in such way as to ensure that processing is unproblematic without having to add additional context. It is necessary to choose the material so that the interpretation of the noun as the subject of verb in the reduced relative clause is ruled out. Accordingly, *evidence* in (13) refers to an inanimate noun. It is therefore not a possible agent of *examined*. A hypothesis with *evidence* as the agent of *examined* is therefore never created when processing this sentence (Sag & Wasow 2011).

- (13) The evidence examined by the judge turned out to be unreliable.

Since processing proceeds incrementally, it is sometimes assumed that realistic grammars should be obliged to immediately assign a constituent structure to previously heard material (Ades & Steedman 1982; Hausser 1992). Proponents of this view would assume a structure for the following sentence where every word forms a constituent with the preceding material:

- (14) [[[[[[[[[[[Das britische] Finanzministerium] stellt] dem] angeschlagenen]
the British treasury provides the crippled
Bankensystem] des] Landes] mindestens] 200] Milliarden] Pfund] zur]
banking.system of.the country at.least 200 billion pounds to
Verfügung].
use
'The British Treasury is making at least 200 billion pounds available to the crippled banking system.'

Pulman (1985), Stabler (1991) and Shieber & Johnson (1993: 301–308) have shown, however, that it is possible to build semantic structures incrementally, using the kind of

phrase structure grammars we encountered in Chapter 2. This means that a partial semantic representation for the string *das britische* ‘the British’ can be computed without having to assume that the two words form a constituent in (14). Therefore, one does not necessarily need a grammar that licenses the immediate combination of words directly. Furthermore, Shieber & Johnson (1993) point out that from a purely technical point of view, synchronous processing is more costly than asynchronous processing since synchronous processing requires additional mechanisms for synchronization whereas asynchronous processing processes information as soon as it becomes available (p. 297–298). Shieber and Johnson do not clarify whether this also applies to synchronous/asynchronous processing of syntactic and semantic information. See Shieber & Johnson (1993) for incremental processing and for a comparison of Steedman’s Categorical Grammar and TAG.

What kind of conclusions can we draw from the data we have previously discussed? Are there further data that can help to determine the kinds of properties a theory of grammar should have in order to count as psycholinguistically plausible? Sag, Wasow & Bender (2003) and Sag & Wasow (2011, 2015) list the following properties that a performance-compatible competence grammar should have:¹⁰

- surface-oriented
- model-theoretic and therefore constraint-based
- sign-based organization
- strictly lexicalist
- representational underspecification of semantic information

Approaches such as CG, GPSG, LFG, HPSG, CxG and TAG are surface-oriented since they do not assume a base structure from which other structures are derived via transformations. Transformational approaches, however, require additional assumptions.¹¹ This will be briefly illustrated in what follows. In Section 3.1.5, we encountered the following analysis of English interrogatives:

¹⁰ Also, see Jackendoff (2007) for reflections on a performance model for a constraint-based, surface-oriented linguistic theory.

¹¹ An exception among transformational approaches is Phillips (2003). Phillips assumes that structures relevant for phenomena such as ellipsis, coordination and fronting are built up incrementally. These constituents are then reordered in later steps by transformations. For example, in the analysis of (i), the string *Wallace saw Gromit in* forms a constituent where *in* is dominated by a node with the label P(P). This node is then turned into a PP in a subsequent step (p. 43–44).

(i) Wallace saw Gromit in the kitchen.

While this approach is a transformation-based approach, the kind of transformation here is very idiosyncratic and incompatible with other variants of the theory. In particular, the modification of constituents contradicts the assumption of Structure Preservation when applying transformations as well as the *No Tampering Condition* of Chomsky (2008). Furthermore, the conditions under which an incomplete string such as *Wallace saw Gromit in* forms a constituent are not entirely clear.

(15) [_{CP} What_i [_{C'} will_k [_{IP} Ann [_{I'} _{-k} [_{VP} read _{-i}]]]]].

This structure is derived from (16a) by two transformations (two applications of Move- α):

- (16) a. Ann will read what?
b. * Will Ann read what

The first transformation creates the order in (16b) from (16a), and the second creates (15) from (16b).

When a hearer processes the sentence in (15), he begins to build structure as soon as he hears the first word. Transformations can, however, only be carried out when the entire utterance has been heard. One can, of course, assume that hearers process surface structures. However, since – as we have seen – they begin to access semantic knowledge early into an utterance, this begs the question of what we need a deep structure for at all.

In analyses such as those of (15), deep structure is superfluous since the relevant information can be reconstructed from the traces. Corresponding variants of GB have been proposed in the literature (see page 111). They are compatible with the requirement of being surface-oriented. Chomsky (1981a: 181; 1986a: 49) and Lasnik & Saito (1992: 59–60) propose analyses where traces can be deleted. In these analyses, the deep structure cannot be directly reconstructed from the surface structure and one requires transformations in order to relate the two. If we assume that transformations are applied ‘online’ during the analysis of utterances, then this would mean that the hearer would have to keep a structure derived from previously heard material as well as a list of possible transformations during processing in his working memory. In constraint-based grammars, entertaining hypotheses about potential upcoming transformation steps is not necessary since there is only a single surface structure that is processed directly. At present, it is still unclear whether it is actually possible to distinguish between these models empirically. But for Minimalist models with a large number of movements (see Figure 4.20 on page 135, for example), it should be clear that they are unrealistic since storage space is required to manage the hypotheses regarding such movements and we know that such short-term memory is very limited in humans.

Frazier & Clifton (1996: 27) assume that a transformation-based competence grammar yields a grammar with pre-compiled rules or rather templates that is then used for parsing. Therefore, theorems derived from UG are used for parsing and not axioms of UG directly. Johnson (1989) also suggests a parsing system that applies constraints from different sub-theories of GB as early as possible. This means that while he does assume the levels of representation D-Structure, S-Structure, LF and PF, he specifies the relevant constraints (\bar{X} theory, Theta-Theory, Case Theory, ...) as logical conditions that can be reorganized, then be evaluated in a different but logically equivalent order and be used for structure building.¹² Chomsky (2007: 6) also compares human parsing to working through a proof, where each step of the proof can be carried out in different orders.

¹² Stabler (1992: Section 15.7) also considers a constraint-based view, but arrives at the conclusion that parsing and other linguistic tasks should use the structural levels of the competence theory. This would again pose problems for the DTC.

This view does not assume the psychological reality of levels of grammatical representation when processing language, but simply assumes that principles and structures play a role when it comes to language acquisition. As we have seen, the question of whether we need UG to explain language acquisition was not yet decided in favor of UG-based approaches. Instead, all available evidence seems to point in the opposite direction. However, even if innate linguistic knowledge does exist, the question arises as to why one would want to represent this as several structures linked via transformations when it is clear that these do not play a role for humans (especially language learners) when processing language. Approaches that can represent this knowledge using fewer technical means, e.g., without transformations, are therefore preferable. For more on this point, see Kuhn (2007: 615).

The requirement for constraint-based grammars is supported by incremental processing and also by the ability to deduce what will follow from previously heard material. Stabler (1991) has pointed out that Steedman's argumentation with regard to incrementally processable grammars is incorrect, and instead argues for maintaining a modular view of grammar. Stabler has developed a constraint-based grammar where syntactic and semantic knowledge can be accessed at any time. He formulates both syntactic structures and the semantic representations attached to them as conjoined constraints and then presents a processing system that processes structures based on the availability of parts of syntactic and semantic knowledge. Stabler rejects models of performance that assume that one must first apply all syntactic constraints before the semantic ones can be applied. If one abandons this strict view of modularity, then we arrive at something like (17):

$$(17) \quad (\text{Syn}_1 \wedge \text{Syn}_2 \wedge \dots \wedge \text{Syn}_n) \wedge (\text{Sem}_1 \wedge \text{Sem}_2 \wedge \dots \wedge \text{Sem}_n)$$

Syn_1 – Syn_n stand for syntactic rules or constraints and Sem_1 – Sem_n stand for semantic rules or constraints. If one so desires, the expressions in brackets can be referred to as modules. Since it is possible to randomly reorder conjoined expressions, one can imagine performance models that first apply some rules from the syntax module and then, when enough information is present, respective rules from the semantic module. The order of processing could therefore be as in (18), for example:

$$(18) \quad \text{Syn}_2 \wedge \text{Sem}_1 \wedge \text{Syn}_1 \wedge \dots \wedge \text{Syn}_n \wedge \text{Sem}_2 \wedge \dots \wedge \text{Sem}_n$$

If one subscribes to this view of modularity, then theories such as HPSG or CxG also have a modular structure. In the representation assumed in the HPSG variant of Pollard & Sag (1987) and Sign-Based CxG (see Section 10.6.2), the value of SYN would correspond to the syntax module, the value of SEM to the semantic module and the value of PHON to the phonology module. If one were to remove the respective other parts of the lexical entries/dominance schemata, then one would be left with the part of the theory corresponding exactly to the level of representation in question.¹³ Jackendoff (2000) argues

¹³ In current theories in the Minimalist Program, an increasing amount of morphological, syntactic, semantic and information-structural information is being included in analyses (see Section 4.6.1). While there are

for this form of modularity with the relevant interfaces between the modules for phonology, syntax, semantics and further modules from other areas of cognition. Exactly what there is to be gained from assuming these modules and how these could be proved empirically remains somewhat unclear to me. For skepticism with regard to the very concept of modules, see Jackendoff (2000: 22,27). For more on interfaces and modularization in theories such as LFG and HPSG, see Kuhn (2007).

Furthermore, Sag & Wasow (2015: 53–54) argue that listeners often leave semantic interpretation underspecified until enough information is present either in the utterance itself or the context. They do not commit to a certain reading early and run into garden paths or backtrack to other readings. This is modeled appropriately by theories that use a variant of underspecified semantics. For a concrete example of underspecification in semantics see Section 19.3.

In conclusion, we can say that surface-oriented, model-theoretic and strongly lexicalist grammatical theories such as CG, LFG, GPSG, HPSG, CxG and the corresponding GB/MP variants (paired with appropriate semantic representations) can plausibly be combined with processing models, while this is not the case for the overwhelming majority of GB/MP theories.

suggestions for using feature-value pairs (Sauerland & Elbourne 2002: 290–291), a strict structuring of information as in GPSG, LFG, HPSG, CxG and variants of CG and TAG is not present. This means that there are the levels for syntax, Phonological Form and Logical Form, but the information relevant for these levels is an unstructured part of syntax, smeared all over syntactic trees.

第十六章 Language acquisition

Linguists and philosophers are fascinated by the human ability to acquire language. Assuming the relevant input during childhood, language acquisition normally takes place completely effortlessly. Chomsky (1965: 24–25) put forward the requirement that a grammatical theory must provide a plausible model of language acquisition. Only then could it actually explain anything and would otherwise remain descriptive at best. In this section, we will discuss theories of acquisition from a number of theoretical standpoints.

16.1 Principles & Parameters

A very influential explanation of language acquisition is Chomsky's Principles & Parameters model (1981a). Chomsky assumes that there is an innate Universal Grammar that contains knowledge that is equally relevant for all languages. Languages can then vary in particular ways. For every difference between languages in the area of core grammar, there is a feature with a specific value. Normally, the value of a parameter is binary, that is, the value is either '+' or '-'. Depending on the setting of a parameter, a language will have certain properties, that is, setting a parameter determines whether a language belongs to a particular class of languages. Parameters are assumed to influence multiple properties of a grammar simultaneously (Chomsky 1981a: 6). For example, Rizzi (1986) claims that the pro-drop parameter affects whether referential subjects can be omitted, the absence of expletives, subject extraction from clauses with complementizers (*that*-t contexts) and interrogatives and finally the possibility of realizing the subject postverbally in VO-languages (see Chomsky 1981a: Section 4.3; Meisel 1995: 12). It has been noted that there are counter-examples to all the correlations assumed.¹ Another example of a parameter is the Head Directionality Parameter discussed in Section 13.1.1. As was shown, there are languages where heads govern in different directions. In his overview article, Haider (2001) still mentions the parametrized Subjacency Principle but

¹ See Haider (1994) and Haider (2001: Section 2.2) for an overview. Haider assumes that there is at least a correlation between the absence of expletive subjects and pro-drop. However, Galician is a pro-drop language with expletive subject pronouns (Raposo & Uriagereka 1990: Section 2.5). Franks (1995: 314) cites Upper and Lower Sorbian as pro-drop languages with expletive subjects. Scholz & Pullum (2002: 218) point out that there is an expletive pronoun *ci* in modern Italian although Italian is classed as a pro-drop language.

notes that subadjacency is no longer assumed as a principle in newer versions of the theory (see Section 13.1.5.2 for more on subadjacency).

Snyder (2001) discovered a correlation of various phenomena with productive root compounding as it is manifested for instance in compounding of two nouns. He argues that the acquisition of complex predicate formation is connected to the acquisition of compound structures and that there is a parameter that is responsible for this type of compounding and simultaneously for the following set of phenomena:

- | | | |
|-----|--|---------------------------|
| (1) | a. John painted the house red. | (resultative) |
| | b. Mary picked the book up/picked up the book. | (verb-particle) |
| | c. Fred made Jeff leave. | (<i>make</i> -causative) |
| | d. Fred saw Jeff leave. | (perceptual report) |
| | e. Bob put the book on the table. | (<i>put</i> -locative) |
| | f. Alice sent the letter to Sue. | (<i>to</i> -dative) |
| | g. Alice sent Sue the letter. | (double-object dative) |

Snyder examined languages from various language groups: Afroasiatic, Austroasiatic, Austronesian, Finno-Ugric, Indo-European (Germanic, Romance, Slavic), Japanese-Korean, Niger-Kordofanian (Bantu), and Sino-Tibetan, as well as American Sign Language and the language isolate Basque. The languages that were examined either had all of these phenomena or none. This was tested with native speakers of the respective languages. In addition the claim that these phenomena are acquired once noun-noun compounds are used productively was tested for English using CHILDES data. The result was positive with the exception of the double object construction, for which an explanation was provided. The correlation of the phenomena in (1) is interesting and was interpreted as proof of the existence of a parameter that correlates several phenomena in a language. However, Son (2007) and Son & Svenonius (2008) showed that Snyder's claims for Japanese were wrong and that there are further languages like Korean, Hebrew, Czech, Malayalam, Javanese in which some of the phenomena show no correlations.

Gibson & Wexler (1994) discuss the acquisition of constituent order and assume three parameters that concern the position of the verb relative to the subject (SV vs. VS) and relative to the object (VO vs. OV) as well as the V2-property. There is no consensus in the literature about which parameters determine the make-up of languages (see Newmeyer 2005: Section 3.2 and Haspelmath 2008 for an overview and critical discussion). Fodor (1998a: 346–347) assumes that there are 20 to 30 parameters, Gibson & Wexler (1994: 408) mention the number 40, Baker (2003: 349) talks of 10 to 20 and Roberts & Holmberg (2005: 541) of 50 to 100. There is no consensus in the literature as to which parameters one should assume, how they interact and what they predict. However, it is nevertheless possible to contemplate how a grammar of an individual language could be derived from a UG with parameters that need to be set. Chomsky's original idea (1986b: Section 3.5.1) was that the child sets the value of a parameter based on the language input as soon as the relevant evidence is present from the input (see also Gibson & Wexler 1994; Nowak, Komarova & Niyogi 2001). At a given point in time, the learner has a grammar with certain parameter settings that correspond to the input seen so far.

In order to fully acquire a grammar, all parameters must be assigned a value. In theory, thirty utterances should be enough to acquire a grammar with thirty parameters if these utterances provide unambiguous evidence for a particular parameter value.

This approach has often been criticized. If setting a parameter leads to a learner using a different grammar, one would expect sudden changes in linguistic behavior. This is, however, not the case (Bloom 1993: 731). Fodor (1998a: 343–344) also notes the following three problems: 1) Parameters can affect things that are not visible from the perceptible constituent order. 2) Many sentences are ambiguous with regard to the setting of a particular parameter, that is, there are sometimes multiple combinations of parameters compatible with one utterance. Therefore, the respective utterances cannot be used to set any parameters (Berwick & Niyogi 1996; Fodor 1998b). 3) There is a problem with the interaction of parameters. Normally multiple parameters play a role in an utterance such that it can be difficult to determine which parameter contributes what and thus how the values should be determined.

Points 1) and 2) can be explained using the constituent order parameters of Gibson & Wexler: imagine a child hears sentences such as the English and the German examples in (2):

- (2) a. Daddy drinks juice.
 b. Papa trinkt Saft.
 daddy drinks juice

These sentences look exactly the same, even though radically different structures are assumed for each. According to the theories under discussion, the English sentence has the structure shown in Figure 3.8 on page 91 given in abbreviated form in (3a). The German sentence, on the other hand, has the structure in Figure 3.13 on page 98 corresponding to (3b):

- (3) a. [_{IP} [Daddy [_{I'} _{-k} [_{VP} drinks_k juice]]]].
 b. [_{CP} Papa_i [_{C'} trinkt_k [_{IP} _{-i} [_{I'} [_{VP} Saft _{-k}] _{-k}]]]].

English has the basic constituent order SVO. The verb forms a constituent with the object (VP) and this is combined with the subject. The parameter setting must therefore be SV, VO and –V2. German, on the other hand, is analyzed as a verb-final and verb-second language and the parameter values would therefore have to be SV, OV and +V2. If we consider the sentences in (2), we see that both sentences do not differ from one another with regard to the order of the verb and its arguments.

Fodor (1998a,b) concludes from this that one first has to build a structure in order to see what grammatical class the grammar licensing the structure belongs to since one first needs the structure in (3b) in order to be able to see that the verb in the partial constituent occurs after its argument in the VP (Saft _{-k}). The question is now how one achieves this structure. A UG with 30 parameters corresponds to $2^{30} = 1,073,741,824$ fully instantiated grammars. It is an unrealistic assumption that children try out these grammars successively or simultaneously.

Gibson & Wexler (1994) discuss a number of solutions for this problem: parameters have a default value and the learner can only change a parameter value if a sentence that could previously not be analyzed can then be analyzed with the new parameter

setting (*Greediness Constraint*). In this kind of procedure, only one parameter can be changed at a time (*Single Value Constraint*), which aims at ruling out great leaps leading to extremely different grammars (see Berwick & Niyogi 1996: 612–613, however). This reduces the processing demands, however with 40 parameters, the worst case could still be that one has to test 40 parameter values separately, that is, try to parse a sentence with 40 different grammars. This processing feat is still unrealistic, which is why Gibson & Wexler (1994: 442) additionally assume that one hypothesis is tested per input sentence. A further modification of the model is the assumption that certain parameters only begin to play a role during the maturation of the child. At a given point in time, there could be only a few accessible parameters that also need to be set. After setting these parameters, new parameters could become available.

In their article, Gibson & Wexler show that the interaction between input and parameter setting is in no way trivial. In their example scenario with three parameters, a situation can arise in which a learner sets a parameter in order to analyze a new sentence, however setting this parameter leads to the fact that the target grammar cannot be acquired because only one value can be changed at a time and changes can only be made if more sentences can be analyzed than before. The learner reaches a so-called local maximum in these problematic cases.² Gibson & Wexler then suggest assigning a default value to particular parameters, whereby the default value is the one that will cause the learner to avoid problematic situations. For the V2 parameter, they assume ‘–’ as the default value.

Berwick & Niyogi (1996) show that Gibson & Wexler calculated the problematic conditions incorrectly and that, if one shares their assumptions, it is even more frequently possible to arrive at parameter combinations from which it is not possible to reach the target grammar by changing individual parameter values. They show that one of the problematic cases not addressed by Gibson & Wexler is –V2 (p. 609) and that the assumption of a default value for a parameter does not solve the problem as both ‘+’ and ‘–’ can lead to problematic combinations of parameters.³ In their article, Berwick and Niyogi show that learners in the example scenario above (with three parameters) learn the target grammar faster if one abandons the Greediness or else the Single Value Constraint. They suggest a process that simply randomly changes one parameter if a sentence cannot be analyzed (*Random Step*, p. 615–616). The authors note that this approach does not share the problems with the local maxima that Gibson & Wexler had in their example and that it also reaches its goal faster than theirs. However, the fact that *Random Step* converges more quickly has to do with the quality of the parameter space (p. 618). Since there is no consensus about parameters in the literature, it is not possible to assess how the entire system works.

Yang (2004: 453) has criticized the classic Principles & Parameters model since abrupt switching between grammars after setting a parameter cannot be observed. In-

² If one imagines the acquisition process as climbing a hill, then the Greediness Constraint ensures that one can only go uphill. It could be the case, however, that one begins to climb the wrong hill and can no longer get back down.

³ Kohl (1999, 2000) has investigated this acquisition model in a case with twelve parameters. Of the 4096 possible grammars, 2336 (57%) are unlearnable if one assumes the best initial values for the parameters.

stead, he proposes the following learning mechanism:

- (4) For an input sentence, s , the child: (i) with probability P_i selects a grammar G_i , (ii) analyzes s with G_i , (iii) if successful, reward G_i by increasing P_i , otherwise punish G_i by decreasing P_i .

Yang discusses the example of the pro-drop and topic drop parameters. In pro-drop languages (e.g., Italian), it is possible to omit the subject and in topic drop languages (e.g., Mandarin Chinese), it is possible to omit both the subject and the object if it is a topic. Yang compares English-speaking and Chinese-speaking children noting that English children omit both subjects and objects in an early linguistic stage. He claims that the reason for this is that English-speaking children start off using the Chinese grammar.

The pro-drop parameter is one of the most widely discussed parameters in the context of Principles & Parameters theory and it will therefore be discussed in more detail here. It is assumed that speakers of English have to learn that all sentences in English require a subject, whereas speakers of Italian learn that subjects can be omitted. One can observe that children learning both English and Italian omit subjects (German children too in fact). Objects are also omitted notably more often than subjects. There are two possible explanations for this: a competence-based one and a performance-based one. In competence-based approaches, it is assumed that children use a grammar that allows them to omit subjects and then only later acquire the correct grammar (by setting parameters or increasing the rule apparatus). In performance-based approaches, by contrast, the omission of subjects is traced back to the fact that children are not yet capable of planning and producing long utterances due to their limited brain capacity. Since the cognitive demands are greatest at the beginning of an utterance, this leads to subjects being increasingly left out. Valian (1991) investigated these various hypotheses and showed that the frequency with which children learning English and Italian respectively omit subjects is not the same. Subjects are omitted more often than objects. She therefore concludes that competence-based explanations are not empirically adequate. The omission of subjects should then be viewed more as a performance phenomenon (see also Bloom 1993). Another argument for the influence of performance factors is the fact that articles of subjects are left out more often than articles of objects (31% vs. 18%, see Gerken 1991: 440). As Bloom notes, no subject article-drop parameter has been proposed so far. If we explain this phenomenon as a performance phenomenon, then it is also plausible to assume that the omission of complete subjects is due to performance issues.

Gerken (1991) shows that the metrical properties of utterances also play a role: in experiments where children had to repeat sentences, they omitted the subject/article of the subject more often than the object/article of the object. Here, it made a difference whether the intonation pattern was iambic (weak-strong) or trochaic (strong-weak). It can even be observed with individual words that children leave out weak syllables at the beginning of words more often than at the end of the word. Thus, it is more probable that “giRAFFE” is reduced to “RAFFE” than “MONkey” to “MON”. Gerken assumes the following for the metrical structure of utterances:

1. A metrical foot contains one and only one strong syllable.
2. Create maximally binary left-to-right feet.
3. Metrical structure is independent of syntactic structure.

Subject pronouns in English are sentence-initial and form a iambic foot with the following strongly emphasized verb as in (5a). Object pronouns, however, can form the weak syllable of a trochaic foot as in (5b).

- (5) a. she KISSED + the DOG
 b. the DOG + KISSED her
 c. PETE + KISSED the + DOG

Furthermore, articles in iambic feet as in the object of (5a) and the subject of (5b) are omitted more often than in trochaic feet such as with the object of (5c).

It follows from this that there are multiple factors that influence the omission of elements and that one cannot simply take the behavior of children as evidence for switching between two grammars.

Apart from what has been discussed so far, the pro-drop parameter is of interest for another reason: there is a problem when it comes to setting parameters. The standard explanation is that learners identify that a subject must occur in all English sentences, which is suggested by the appearance of expletive pronouns in the input.

As discussed on page 483, there is no relation between the pro-drop property and the presence of expletives in a language. Since the pro-drop property does not correlate with any of the other putative properties either, only the existence of subject-less sentences in the input constitutes decisive evidence for setting a parameter. The problem is that there are grammatical utterances where there is no visible subject. Examples of this are imperatives such as (6), declaratives with a dropped subject as in (7a) and even declarative sentences without an expletive such as the example in (7b) found by Valian (1991: 32) in the New York Times.

- (6) a. Give me the teddy bear!
 b. Show me your toy!
- (7) a. She'll be a big hit. Sings like a dream.
 b. Seems like she always has something twin-related perking.

The following title of a Nirvana song also comes from the same year as Valian's article:

- (8) Smells like Teen Spirit.

Teen Spirit refers to a deodorant and *smell* is a verb that, both in German and English, requires a referential subject but can also be used with an expletive *it* as subject. The usage that Kurt Cobain had in mind cannot be reconstructed⁴, independent of the intended

⁴ See http://de.wikipedia.org/wiki/Smells_Like_Teen_Spirit. 06.03.2016.

meaning, however, the subject in (8) is missing. Imperatives do occur in the input children have and are therefore relevant for acquisition. Valian (1991: 33) says the following about them:

What is acceptable in the adult community forms part of the child's input, and is also part of what children must master. The utterances that I have termed "acceptable" are not grammatical in English (since English does not have pro subjects, and also cannot be characterized as a simple VP). They lack subjects and therefore violate the extended projection principle (Chomsky 1981a), which we are assuming.

Children are exposed to fully grammatical utterances without subjects, in the form of imperatives. They are also exposed to acceptable utterances which are not fully grammatical, such as [(7a)], as well as forms like, "Want lunch now?" The American child must grow into an adult who not only knows that overt subjects are grammatically required, but also knows when subjects can acceptably be omitted. The child must not only acquire the correct grammar, but also master the discourse conditions that allow relaxation of the grammar. (Valian 1991: 33)

This passage turns the relations on their head: we cannot conclude from the fact that a particular grammatical theory is not compatible with certain data, that these data should not be described by this theory, instead we should modify the incompatible grammar or, if this is not possible, we should reject it. Since utterances with imperatives are entirely regular, there is no reason to categorize them as utterances that do not follow grammatical rules. The quotation above represents a situation where a learner has to acquire two grammars: one that corresponds to the innate grammar and a second that partially suppresses the rules of innate grammar and also adds some additional rules.

The question we can pose at this point is: how does a child distinguish which of the data it hears are relevant for which of the two grammars?

Fodor (1998a: 347) pursues a different analysis that does not suffer from many of the aforementioned problems. Rather than assuming that learners try to find a correct grammar among a billion others, she instead assumes that children work with a single grammar that contains all possibilities. She suggests using parts of trees (*treelets*) rather than parameters. These treelets can also be underspecified and in extreme cases, a treelet can consist of a single feature (Fodor 1998b: 6). A language learner can deduce whether a language has a given property from the usage of a particular treelet. As an example, she provides a VP treelet consisting of a verb and a prepositional phrase. This treelet must be used for the analysis of the VP occurring in *Look at the frog*. Similarly, the analysis of an interrogative clause with a fronted *who* would make use of a treelet with a *wh*-NP in the specifier of a complementizer phrase (see Figure 3.7 on page 91). In Fodor's version of Principles and Parameters Theory, this treelet would be the parameter that licenses *wh*-movement in (overt) syntax. Fodor assumes that there are defaults that allow a learner to parse a sentence even when no or very few parameters have been set. This allows one to learn from utterances that one would have not otherwise been able to use since there would have been multiple possible analyses for them. Assuming a default can lead to misanalyses, however: due to a default value, a second parameter could be set

because an utterance was analyzed with a treelet t_1 and t_3 , for example, but t_1 was not suited to the particular language in question and the utterance should have instead been analyzed with the non-default treelet t_2 and the treelet t_{17} . In this acquisition model, there must therefore be the possibility to correct wrong decisions in the parameter setting process. Fodor therefore assumes that there is a frequency-based degree of activation for parameters (p. 365): treelets that are often used in analyses have a high degree of activation, whereas those used less often have a lower degree of activation. In this way, it is not necessary to assume a particular parameter value while excluding others.

Furthermore, Fodor proposes that parameters should be structured hierarchically, that is, only if a parameter has a particular value does it then make sense to think about specific other parameter values.

Fodor's analysis is – as she herself notes (Fodor 2001: 385) – compatible with theories such as HPSG and TAG. Pollard & Sag (1987: 147) characterize UG as the conjunction of all universally applicable principles:

$$(9) \quad \text{UG} = P_1 \wedge P_2 \wedge \dots \wedge P_n$$

As well as principles that hold universally, there are other principles that are specific to a particular language or a class of languages. Pollard & Sag give the example of the constituent ordering principle that only holds for English. English can be characterized as follows if one assumes that P_{n+1} – P_m are language-specific principles, L_1 – L_p a complete list of lexical entries and R_1 – R_q a list of dominance schemata relevant for English.

$$(10) \quad \text{English} = P_1 \wedge P_2 \wedge \dots \wedge P_m \wedge (L_1 \vee \dots \vee L_p \vee R_1 \vee \dots \vee R_q)$$

In Pollard & Sag's conception, only those properties of language that equally hold for all languages are part of UG. Pollard & Sag do not count the dominance schemata as part of this. However, one can indeed also describe UG as follows:

$$(11) \quad \text{UG} = P_1 \wedge P_2 \wedge \dots \wedge P_n \wedge (R_{\text{en-1}} \vee \dots \vee R_{\text{en-q}} \vee R_{\text{de-1}} \vee \dots \vee R_{\text{de-r}} \vee \dots)$$

P_1 – P_n are, as before, universally applicable principles and $R_{\text{en-1}}$ – $R_{\text{en-q}}$ are the (core) dominance schemata of English and $R_{\text{de-1}}$ – $R_{\text{de-r}}$ are the dominance schemata in German. The dominance schemata in (11) are combined by means of disjunctions, that is, not every disjunct needs to have a realization in a specific language. Principles can make reference to particular properties of lexical entries and rule out certain phrasal configurations. If a language only contains heads that are marked for final-position in the lexicon, then grammatical rules that require a head in initial position as their daughter can never be combined with these heads or their projections. Furthermore, theories with a type system are compatible with Fodor's approach to language acquisition because constraints can easily be underspecified. As such, constraints in UG do not have to make reference to all properties of grammatical rules: principles can refer to feature values, the language-specific values themselves do not have to already be contained in UG. Similarly, a supertype describing multiple dominance schemata that have similar but language-specific instantiations can also be part of UG, however the language-specific details remain open and are then deduced by the learner upon parsing (see Ackerman & Webelhuth 1998: Section 9.2). The differences in activation assumed by Fodor can be captured by

weighting the constraints: the dominance schemata R_{en-1} – R_{en-q} etc. are sets of feature-value pairs as well as path equations. As explained in Chapter 15, weights can be added to such constraints and also to sets of constraints. In Fodor’s acquisition model, given a German input, the weights for the rules of English would be reduced and those for the German rules would be increased. Note that in Pollard & Sag’s acquisition scenario, there are no triggers for parameter setting unlike in Fodor’s model. Furthermore, properties that were previously disjunctively specified as part of UG will now be acquired directly. Using the treelet t_{17} (or rather a possibly underspecified dominance schema), it is not the case that the value ‘+’ is set for a parameter P_5 but rather the activation potential of t_{17} is increased such that t_{17} will be prioritized for future analyses.

16.2 Principles and the lexicon

A variant of the UG-driven theory of language acquisition would be to assume that principles are so general that they hold for all languages and individual languages simply differ with regard to their lexicon. Principles then refer to properties of combined entities. Parameters therefore migrate from principles into the lexicon (Chomsky 1999: 2). See Mensching & Remberger (2011) for a study of Romance languages in this model and Son & Svenonius (2008: 395) for an analysis of Snyder’s examples that were discussed in the previous subsection.

At this point, one can observe an interesting convergence in these approaches: most of the theories discussed here assume a very general structure for the combination of heads with their arguments. For example, in Categorical Grammar and the Minimalist Program, these are always binary functor-argument combinations. The way in which constituents can be ordered in a particular language depends on the lexical properties of the combined elements.

The question that is being discussed controversially at present is whether the spectrum of lexical properties is determined by UG (Chomsky 2007: 6–7) and whether all areas of the language can be described with the same general combinatorial possibilities (see Section 21.10 on phrasal constructions).

In Section 16.1, I have shown what theories of acquisition assuming innate language specific knowledge can look like and also that variants of such acquisition theories are compatible with all the theories of grammar we have discussed. During this discussion, one should bear in mind the question of whether it makes sense at all to assume that English children use parts of a Chinese grammar during some stages of their acquisition process (as suggested by Yang 2004: 453), or whether the relevant phenomena can be explained in different ways. In the following, I will present some alternative approaches that do not presuppose innate language specific knowledge, but instead assume that language can simply be acquired from the input. The following section will deal with pattern-based approaches and Section 16.4 will discuss the lexically-oriented variant of input-based language acquisition.

16.3 Pattern-based approaches

Chomsky (1981a: 7–8) proposed that languages can be divided into a core area and a periphery. The core contains all regular aspects of language. The core grammar of a language is seen as an instantiation of UG. Idioms and other irregular parts of language are then part of the periphery. Critics of the Principles & Parameters model have pointed out that idiomatic and irregular constructions constitute a relatively large part of our language and that the distinction, both fluid and somewhat arbitrary, is only motivated theory-internally (Jackendoff 1997: Chapter 7; Culicover 1999; Ginzburg & Sag 2000: 5; Newmeyer 2005: 48; Kuhn 2007: 619). For example, it is possible to note that there are interactions between various idioms and syntax (Nunberg, Sag & Wasow 1994). Most idioms in German with a verbal component allow the verb to be moved to initial position (12b), some allow that parts of idioms can be fronted (12c) and some can undergo passivization (12d).

- (12) a. dass er ihm den Garaus macht
that he him the GARAUS makes
‘that he finishes him off (kills him)’
- b. Er macht ihm den Garaus.
he makes him the GARAUS
‘He finishes him off.’
- c. In Amerika sagte man der Kamera nach, die größte Kleinbildkamera der Welt zu sein. Sie war laut Schleiffer am Ende der Sargnagel der Mühlheimer Kameraproduktion.
Den Garaus machte ihr die Diskussion um die Standardisierung des the GARAUS made her the discussion around the standardization of the 16-Millimeter-Filmformats, an dessen Ende die DIN-Norm 19022 (Patrone 16-millimeter-film.format at whose end the DIN-norm 19022 cartridge mit Spule für 16-Millimeter-Film) stand, die im März 1963 zur Norm with coil for 16-millimeter-film stood that in March 1963 to.the norm wurde.⁵
became
‘In America, one says that this camera was the biggest compact camera in the world. According to Schleiffer, it was the last nail in the coffin for camera production in Mühlheim. What finished it off was the discussion about standardizing the 16 millimeter format, which resulted in the DIN-Norm 19022 (cartridge with coil for 16 millimeter film) that became the norm in March 1963.’
- d. in Heidelberg wird „parasitären Elementen“ unter den Professoren *der* in Heidelberg are parasitic elements among the professors the

⁵ Frankfurter Rundschau, 28.06.1997, p. 2.

*Garaus gemacht*⁶

GARAUS made

‘In Heidelberg, “parasitic elements” among professors are being killed off.’

It is assumed that the periphery and lexicon are not components of UG (Chomsky 1986b: 150–151; Fodor 1998a: 343) but rather are acquired using other learning methods – namely inductively directly from the input. The question posed by critics is now why these methods should not work for regular aspects of the language as well (Abney 1996: 20; Goldberg 2003a: 222; Newmeyer 2005: 100; Tomasello 2006c: 36; 2006b: 20): the areas of the so-called ‘core’ are by definition more regular than components of the periphery, which is why they should be easier to learn.

Tomasello (2000, 2003) has pointed out that a Principles & Parameters model of language acquisition is not compatible with the observable facts. The Principles and Parameters Theory predicts that children should no longer make mistakes in a particular area of grammar once they have set a particular parameter correctly (see Chomsky 1986b: 146, Radford 1990: 21–22 and Lightfoot 1997: 175). Furthermore, it is assumed that a parameter is responsible for very different areas of grammar (see the discussion of the pro-drop parameter in Section 16.1). When a parameter value is set, then there should be sudden developments with regard to a number of phenomena (Lightfoot 1997: 174). This is, however, not the case. Instead, children acquire language from utterances in their input and begin to generalize from a certain age. Depending on the input, they can reorder certain auxiliaries and not others, although movement of auxiliaries is obligatory in English.⁷ One argument put forward against these kinds of input-based theories is that children produce utterances that cannot be observed to a significant frequency in the input. One much discussed phenomenon of this kind are so called *root infinitives* (RI) or *optional infinitives* (OI) (Wexler 1998). These are infinitive forms that can be used in non-embedded clauses (*root sentences*) instead of a finite verb. Optional infinitives are those where children use both a finite (13a) and non-finite (13b) form (Wexler 1998: 59):

- (13) a. Mary likes ice cream.
 b. Mary like ice cream.

Wijnen, Kempen & Gillis (2001: 656) showed that Dutch children use the order object infinitive 90 % of the time during the two-word phase although these orders occur in less than 10 % of their mother’s utterances that contained a verb. Compound verb forms, e.g., with a modal in initial position as in (14) that contain another instance of this pattern only occurred in 30 % of the input containing a verb (Wijnen, Kempen & Gillis 2001: 647).

- (14) Willst du Brei essen?
 want you porridge eat
 ‘Do you want to eat porridge?’

⁶ Mannheimer Morgen, 28.06.1999, Sport; Schrauben allein genügen nicht.

⁷ Here, Yang’s suggestion to combine grammars with a particular probability does not help since one would have to assume that the child uses different grammars for different auxiliaries, which is highly unlikely.

At first glance, there seems to be a discrepancy between the input and the child's utterances. However, this deviation could also be explained by an utterance-final bias in learning (Wijnen et al. 2001; Freudenthal, Pine & Gobet 2006). A number of factors can be made responsible for the salience of verbs at the end of an utterance: 1) restrictions of the infant brain. It has been shown that humans (both children and adults) forget words during the course of an utterance, that is, the activation potential decreases. Since the cognitive capabilities of small children are restricted, it is clear why elements at the end of an utterance have an important status. 2) Easier segmentation at the end of an utterance. At the end of an utterance, part of the segmentation problem for hearers disappears: the hearer first has to divide a sequence of phonemes into individual words before he can understand them and combine them to create larger syntactic entities. This segmentation is easier at the end of an utterance since the word boundary is already given by the end of the utterance. Furthermore according to Wijnen, Kempen & Gillis (2001: 637), utterance-final words have an above average length and do bear a pitch accent. This effect occurs more often in language directed at children.

Freudenthal, Pine, Aguado-Orea & Gobet (2007) have modeled language acquisition for English, German, Dutch, and Spanish. The computer model could reproduce differences between these languages based on input. At first glance, it is surprising that there are even differences between German and Dutch and between English and Spanish with regard to the use of infinitives as German and Dutch have a very similar syntax (SOV+V2). Similarly, English and Spanish are both languages with SVO order. Nevertheless, children learning English make OI mistakes, whereas this is hardly ever the case for children learning Spanish.

Freudenthal, Pine, Aguado-Orea & Gobet (2007) trace the differences in error frequencies back to the distributional differences in each language: the authors note that 75 % of verb-final utterances⁸ in English consist of compound verbs (finite verb + dependent verb, e.g., *Can he go?*), whereas this is only the case 30 % of the time in Dutch.

German also differs from Dutch with regard to the number of utterance-final infinitives. Dutch has a progressive form that does not exist in Standard German:

- (15) Wat ben je aan het doen?
 what are you on it do.INF
 'What are you doing?'

Furthermore, verbs such as *zitten* 'to sit', *lopen* 'to run' and *staan* 'to stand' can be used in conjunction with the infinitive to describe events happening in that moment:

- (16) Zit je te spelen?
 sit you to play
 'Are you sitting and playing?'

Furthermore, there is a future form in Dutch that is formed with *ga* 'go'. These factors contribute to the fact that Dutch has 20 % more utterance-final infinitives than German.

⁸ For English, the authors only count utterances with a subject in third person singular since it is only in these cases that a morphological difference between the finite and infinitive form becomes clear.

Spanish differs from English in that it has object clitics:

- (17) (Yo) Lo quiero.
 I it want
 ‘I want it.’

Short pronouns such as *lo* in (17) are realized in front of the finite verb so that the verb appears in final position. In English, the object follows the verb, however. Furthermore, there are a greater number of compound verb forms in the English input (70 %) than in Spanish (25 %). This is due to the higher frequency of the progressive in English and the presence of *do*-support in question formation.

The relevant differences in the distribution of infinitives are captured correctly by the proposed acquisition model, whereas alternative approaches that assume that children possess an adult grammar but use infinitives instead of the finite forms cannot explain the gradual nature of this phenomenon.

Freudenthal, Pine & Gobet (2009) could even show that input-based learning is superior to other explanations for the distribution of NPs and infinitives. They can explain why this order is often used with a modal meaning (e.g., *to want*) in German and Dutch (Ingram & Thompson 1996). In these languages, infinitives occur with modal verbs in the corresponding interrogative clauses. Alternative approaches that assume that the linguistic structures in question correspond to those of adults and only differ from them in that a modal verb is not pronounced cannot explain why not all utterances of object and verb done by children learning German and Dutch do have a modal meaning. Furthermore, the main difference to English cannot be accounted for: in English, the number of modal meanings is considerably less. Input-based models predict this exactly since English can use the dummy verb *do* to form questions:

- (18) a. Did he help you?
 b. Can he help you?

If larger entities are acquired from the end of an utterance, then there would be both a modal and non-modal context for *he help you*. Since German and Dutch normally do not use the auxiliary *tun* ‘do’, the relevant endings of utterances are always associated with modals contexts. One can thereby explain why infinitival expressions have a modal meaning significantly more often in German and Dutch than in English.

Following this discussion of the arguments against input-based theories of acquisition, I will turn to Tomasello’s pattern-based approach. According to Tomasello (2003: Section 4.2.1), a child hears a sentence such as (19) and realizes that particular slots can be filled freely (see also Dąbrowska (2001) for analogous suggestions in the framework of Cognitive Grammar).

- (19) a. Do you want more juice/milk?
 b. Mommy is gone.

From these utterances, it is possible to derive so-called pivot schemata such as those in (20) into which words can then be inserted:

- (20) a. more ____ → more juice/milk
 b. ____ gone → mommy/juice gone

In this stage of development (22 months), children do not generalize using these schemata, these schemata are instead construction islands and do not yet have any syntax (Tomasello et al. 1997). The ability to use previously unknown verbs with a subject and an object in an SVO order is acquired slowly between the age of three and four (Tomasello 2003: 128–129). More abstract syntactic and semantic relations only emerge with time: when confronted with multiple instantiations of the transitive construction, the child is then able to generalize:

- (21) a. [_S [_{NP} The man/the woman] sees [_{NP} the dog/the rabbit/it]].
 b. [_S [_{NP} The man/the woman] likes [_{NP} the dog/the rabbit/it]].
 c. [_S [_{NP} The man/the woman] kicks [_{NP} the dog/the rabbit/it]].

According to Tomasello (2003: 107), this abstraction takes the form [Sbj TrVerb Obj]. Tomasello's approach is immediately plausible since one can recognize how abstraction works: it is a generalization about reoccurring patterns. Each pattern is then assigned a semantic contribution. These generalizations can be captured in inheritance hierarchies (see page ??) (Croft 2001: 26). The problem with this kind of approach, however, is that it cannot explain the interaction between different areas of phenomena in the language: it is possible to represent simple patterns such as the use of transitive verbs in (21), but transitive verbs interact with other areas of the grammar such as negation. If one wishes to connect the construction one assumes for the negation of transitive verbs with the transitive construction, then one arrives at a problem since this is not possible in inheritance hierarchies.

- (22) The woman did not kick the dog.

The problem is that the transitive construction has a particular semantic contribution but that negated transitive construction has the opposite meaning. The values of SEM features would therefore be contradictory. There are technical tricks to avoid this problem, however, since there are a vast number of these kinds of interactions between syntax and semantics, this kind of technical solution will result in something highly implausible from a cognitive perspective (Müller 2006, 2007c,b, 2010b; Müller & Wechsler 2014a). For discussion of Croft's analysis, see Section 21.4.1.

At this point, proponents of pattern-based analyses might try and argue that these kinds of problems are only the result of a poor/inadequate formalization and would rather do without a formalization (Goldberg 2009: Section 5). However, this does not help here as the problem is not the formalization itself, rather the formalization allows one to see the problem more clearly.

An alternative to an approach built entirely on inheritance is a TAG-like approach that allows one to insert syntactic material into phrasal constructions. Such a proposal was discussed in Section 10.6.3. Bergen & Chang (2005: 170) working in Embodied Construction Grammar suggest an Active-Ditransitive Construction with the form

[RefExpr Verb RefExpr RefExpr], where RefExpr stands for a referential expression and the first RefExpr and the verb may be non-adjacent. In this way, it is possible to analyze (23a,b), while ruling out (23c):

- (23) a. Mary tossed me a drink.
 b. Mary happily tossed me a drink.
 c. * Mary tossed happily me a drink.

While the compulsory adjacency of the verb and the object correctly predicts that (23c) is ruled out, the respective constraint also rules out coordinate structures such as (24):

- (24) Mary tossed me a juice and Peter a water.

Part of the meaning of this sentence corresponds to what the ditransitive construction contributes to *Mary tossed Peter a water*. There is, however, a gap between *tossed* and *Peter*. Similarly, one can create examples where there is a gap between both objects of a ditransitive construction:

- (25) He showed me and bought for Mary the book that was recommended in the Guardian last week.

In (25), *me* is not adjacent to *the book* It is not my aim here to request a coordination analysis. Coordination is a very complex phenomenon for which most theories do not have a straightforward analysis (see Section 21.6.2). Instead, I would simply like to point out that the fact that constructions can be realized discontinuously poses a problem for approaches that claim that language acquisition is exclusively pattern-based. The point is the following: in order to understand coordination data in a language, a speaker must learn that a verb which has its arguments somewhere in the sentence has a particular meaning together with these arguments. The actual pattern [Sbj V Obj1 Obj2] can, however, be interrupted in all positions. In addition to the coordination examples, there is also the possibility of moving elements out of the pattern either to the left or the right. In sum, we can say that language learners have to learn that there is a relation between functors and their arguments. This is all that is left of pattern-based approaches but this insight is also covered by the selection-based approaches that we will discuss in the following section.

A defender of pattern-based approaches could perhaps object that there is a relevant construction for (25) that combines all material. This means that one would have a construction with the form [Sbj V Obj1 Conj V PP Obj2]. It would then have to be determined experimentally or with corpus studies whether this actually makes sense. The generalization that linguists have found is that categories with the same syntactic properties can be coordinated (N, \bar{N} , NP, V, \bar{V} , VP, ...). For the coordination of verbs or verbal projections, it must hold that the coordinated phrases require the same arguments:

- (26) a. Er [arbeitet] und [liest viele Bücher].
 he works and reads many books

- b. Er [kennt und liebt] diese Schallplatte.
he knows and loves this record
- c. Er [zeigt dem Jungen] und [gibt der Frau] die Punk-Rock-CD.
he shows the boy and gives the woman the punk rock CD
- d. Er [liebt diese Schallplatte] und [schenkt ihr ein Buch].
he loves this record and gives her a book

In an approach containing only patterns, one would have to assume an incredibly large number of constructions and so far we are only considering coordinations that consist of exactly two conjuncts. However, the phenomenon discussed above is not only restricted to coordination of two elements. If we do not wish to abandon the distinction between competence and performance (see Chapter 15), then the number of conjuncts is not constrained at all (by the competence grammar):

- (27) Er [kennt, liebt und verborgt] diese Schallplatte.
he knows loves and lends.out this record

It is therefore extremely unlikely that learners have patterns for all possible cases in their input. It is much more likely that they draw the same kind of generalizations as linguists from the data occurring in their input: words and phrases with the same syntactic properties can be coordinated. If this turns out to be true, then all that is left for pattern-based approaches is the assumption of discontinuously realized constructions and thus a dependency between parts of constructions that states that they do not have to be immediately adjacent to one another. The acquisition problem is then the same as for selection-based approaches that will be the topic of the following section: what ultimately has to be learned are dependencies between elements or valences (see Behrens (2009: 439), the author reaches the same conclusion following different considerations).

16.4 Selection-based approaches

I will call the alternative to pattern-based approaches *selection-based*. A selection-based approach has been proposed by Green (2011).

The generalizations about the pattern in (21) pertain to the valence class of the verb. In Categorical Grammar, the pattern [Sbj TrVerb Obj] corresponds to the lexical entry (s\vp)/np (for the derivation of a sentence with this kind of lexical entry, see Figure 8.3 on page 221). A TAG tree for *likes* was given on page 381. Here, one can see quite clearly that lexical entries determine the structure of sentences in these models. Unlike pattern-based approaches, these analyses allow enough room for semantic embedding: the lexical entries in Categorical Grammar can be combined with adjuncts, and elementary trees in TAG also allow for adjunction to the relevant nodes.

Now, we face the question of how the jump from a pivot schema to a lexical entry with an argument structure takes place. In Tomasello's approach, there is no break between them. Pivot schemata are phrasal patterns and [Sbj TrVerb Obj] is also a phrasal pattern. Both schemata have open slots into which certain elements can be inserted. In

selection-based approaches, the situation is similar: the elements that are fixed in the pivot schema are functors in the selection-based approach. Green (2011) proposes a theory of acquisition in HPSG that can do without UG. For the two-word phase, she assumes that *where's* is the head of an utterance such as (28) and that *where's* selects *Robin* as its argument.

(28) Where's Robin?

This means that, rather than assuming that there is a phrasal pattern *Where's X?* with an empty slot X for a person or thing, she assumes that there is a lexical entry *where's*, which contains the information that it needs to be combined with another constituent. What needs to be acquired is the same in each case: there is particular material that has to be combined with other material in order to yield a complete utterance.

In her article, Green shows how long-distance dependencies and the position of English auxiliaries can be acquired in later stages of development. The acquisition of grammar proceeds in a monotone fashion, that is, knowledge is added – for example, knowledge about the fact that material can be realized outside of the local context – and previous knowledge does not have to be revised. In her model, mistakes in the acquisition process are in fact mistakes in the assignment of lexical entries to valence classes. These mistakes have to be correctable.

In sum, one can say that all of Tomasello's insights can be applied directly to selection-based approaches and the problems with pattern-based approaches do not surface with selection-based approaches. It is important to point out explicitly once again here that the selection-based approach discussed here also is a construction-based approach. Constructions are just lexical and not phrasal. The important point is that, in both approaches, words and also more complex phrases are pairs of form and meaning and can be acquired as such.

In Chapter 21, we will discuss pattern-based approaches further and we will also explore areas of the grammar where phrasal patterns should be assumed.

16.5 Summary

We should take from the preceding discussion that models of language acquisition that assume that a grammar is chosen from a large set of grammars by setting binary parameters are in fact inadequate. All theories that make reference to parameters have in common that they are purely hypothetical since there is no non-trivial set of parameters that all proponents of the model equally agree on. In fact there is not even a trivial one.

In a number of experiments, Tomasello and his colleagues have shown that, in its original form, the Principles & Parameters model makes incorrect predictions and that language acquisition is much more pattern-based than assumed by proponents of P&P analyses. Syntactic competence develops starting from verb islands. Depending on the frequency of the input, certain verbal constructions can be mastered even though the same construction has not yet been acquired with less frequent verbs.

The interaction with other areas of grammar still remains problematic for pattern-based approaches: in a number of publications, it has been shown that the interaction of phenomena that one can observe in complex utterances can in fact not be explained with phrasal patterns since embedding cannot be captured in an inheritance hierarchy. This problem is not shared by selection-based approaches. All experimental results and insights of Tomasello can, however, be successfully extended to selection-based approaches.

Further reading

Meisel (1995) gives a very good overview of theories of acquisition in the Principles & Parameters model.

Adele Goldberg and Michael Tomasello are the most prominent proponents of Construction Grammar, a theory that explicitly tries to do without the assumption of innate linguistic knowledge. They published many papers and books about topics related to Construction Grammar and acquisition. The most important books probably are Goldberg (2006) and Tomasello (2003).

An overview of different theories of acquisition in German can be found in Klann-Delius (2008) an English overview is Ambridge & Lieven (2011).

第十七章 生成能力与语言形式描述

前面有几章曾提到过形式语言的复杂性等级。最简单的语言是所谓的正则语言（3 型），比正则语言更复杂的是上下文无关语法（2 型），比上下文无关语法更复杂的是上下文有关语法（1 型），最复杂的是无限制语法（0 型）。无限制语法可以产生递归可枚举语言（recursively enumerable languages）。在提出理论时，大家都有意识地努力采用能与自然语言真实现象相符的形式手段。所以，大家放弃了无限制转换语法（unrestricted Transformational Grammars），因为该语法的生成能力相当于 0 型语法（参看第 79 页）。GPSG 设计得只能分析上下文无关语言，不能分析生成能力更强的语言。80 年代中期，研究发现自然语言的复杂性高于上下文无关语言（Shieber 1985; Culy 1985）。现在都认为弱上下文有关语法（mildly context sensitive grammars）足以分析自然语言。在树邻接语法（TAG）框架中工作的学者正致力于提出这种类型的树邻接语法（TAG）方案。与此相似，研究发现 Stabler 最简语法（Minimalist Grammars）（参见第 4.6.4 节和 Stabler 2001, 2011b）的不同变体都有弱上下文有关能力（Michaelis 2001）。Peter Hellwig 的依存合一语法也是弱上下文有关的（Hellwig 2003: 595）。现在的问题是：理想目标是不是寻找一种描述语言其生成能力与其描述对象完全一致。Carl Pollard (1996a) 曾经说过，如果仅仅因为某些物理学理论采用了过于有力的数学工具，而断言这些物理学理论不够完备，这种断言则是非常奇怪的。¹ 不应该是描述语言限制理论而应该是理论包含必须适用于研究对象的限制。这是 Chomsky (1981b: 277, 280) 所持的观点。也可以参看 Berwick (1982: Section 4), Kaplan & Bresnan (1982: Section 8) 对于 LFG 的论述以及 Johnson (1988: Section 3.5) 对于 LFG 中离线句法分析能力限制（Off-Line Parsability Constraint）的论述和对于属性-值语法的总体论述。当然，在条件允许的情况下选择复杂性最低的语法也有技术方面的原因：我们知道对于计算机来说处理复杂性较低的语法更加容易。要获知一个任务的复杂性，就需要确定相关计算的所谓“最坏情况”，即确定一个程序在最差情况下使用某种语法处理一定长度输入时需要多长时间。但是，我们不禁提出这样一个问题：最差情况对确定任务复杂性是否真的有用。例如，一些允许非连续成分的语法在最差情况下比那些只允许连续

¹ 如果物理学家要求形式描述限制理论：

编辑：爱因斯坦教授，我恐怕不能接收你在广义相对论方面的论文。

爱因斯坦：为什么呢？难道方程式是错的？

编辑：不，但是我们注意到你的微分方程式是用集合论中的一阶语言表达的。这是一个完全不受限制的形式描述！为什么这样做呢？你本可以写出任意集合微分方程式的！（Pollard 1996a）

字符串组合的常规短语结构语法表现更差 (Reape 1991: Section 8)。正如我在Müller (2004c)所指出的那样：从单词出发构建更大单位的句法分析器（一个自底向上的句法分析器）在处理假设动词移位分析的语法时比处理假设非连续成分的语法更加低效。这与动词语迹没有语音形式有关，由于动词语迹没有语音形式所以句法分析器需要借助另外的手段来定位它们。因此，需要假设动词语迹可以出现在字符串的任意位置，并且在大多数情况下这些语迹对于完整输入的分析不起任何作用。因为动词语迹没有指定其价 (valence) 信息，所以能与句子中任意成分进行组合，这造成了很大的计算负担。相反，如果允许不连续成分，那么就可以放弃使用动词语迹，计算负担因此降低了。但是，使用非连续成分的分析最终因为语言学原因被放弃了 (Müller 2005b,c, 2007b, 2015b)。尽管如此，研究两种语法的句法分析行为仍然具有价值。因为这种研究显示最差情况对确定任务复杂性并非总是有效。下面我将讨论另外一个例子来说明特定语言的限制会限制语法的复杂性：Gärtner & Michaelis (2007: Section 3.2)认为 Stabler 提出的增加了后附接 (late adjunction) 和外置的最简语法 (参看第 4.6.4节) 比弱上下文有关语法的生成力更强。如果一种语法禁止从附接语位置提取成分 (Frey & Gärtner 2002: 46) 并且认同最短移位限制 (参看第 149页脚注 30)，那么该语法就是弱上下文有关语法 (Gärtner & Michaelis 2007: 178)。包含最短移位限制并限制从指定语位置提取成分的语法也是弱上下文有关语法。是否能从指定语位置提取成分取决于所讨论的语法的组织。在一些语法中，所有论元都充当指定语 (Kratzer 1996: 120–123，也可以参看第 511页的图 18.4)。在这些语法中，禁止从指定语位置提取成分意味着提取论元是不可能的。当然，总体来说这是不对的。通常情况下，主语被当作指定语 (Frey & Gärtner 2002: 44也这样认为)。主语经常被认为是阻碍提取的岛 (参看 Grewendorf 1989: 35, 41; G. Müller 1996b: 220; 1998: 32, 163; Sabel 1999: 98; Fanselow 2001: 422)。但是，一些研究者发现在德语中可以从主语位置提取成分 (参看 Dürscheid 1989: 25; Haider 1993: 173; Pafel 1993; Fortmann 1996: 27; Suchsland 1997: 320; Vogel & Steinbach 1998: 87; Ballweg 1997: 2066; Müller 1999a: 100–101; De Kuthy 2002: 7)。下面是经过验证的例子：

- (1) a. [Von den übrigbleibenden Elementen]_i scheinen [die Determinantien _{-i}]
 of the left.over elements seem the determinants
 die wenigsten Klassifizierungsprobleme aufzuwerfen.²
 the fewest classification.problems to.throw.up
 ‘Of the remaining elements, the determinants seem to pose the fewest problems for classification.’
 b. [Von den Gefangenen]_i hatte eigentlich [keine _{-i}] die Nacht der Bomben
 of the prisoners had actually none the night of.the bombs
 überleben sollen.³
 survive should
 ‘None of the prisoners should have actually survived the night of the bombings.’
 c. [Von der HVA]_i hielten sich [etwa 120 Leute _{-i}] dort in ihren Gebäuden
 of the HVA held REFL around 120 people there in their buildings

² In the main text of Engel (1970: 102).

³ Bernhard Schlink, *Der Vorleser*, Diogenes Taschenbuch 22953, Zürich: Diogenes Verlag, 1997, p. 102.

auf.⁴

PRT

‘Around 120 people from the HVA stayed there inside their buildings.’

- d. [Aus dem „Englischen Theater“]_i stehen [zwei Modelle __i] in den Vitrinen.⁵
from the English theater stand two models in the cabinets

‘Two models from the ‘English Theater’ are in the cabinets.’

- e. [Aus der Fraktion]_i stimmten ihm [viele __i] zu darin, dass die
from the faction agreed him many PRT there.in that the
Kaufkraft der Bürger gepäppelt werden müsse, nicht die gute Laune
buying.power.of.the citizens boosted become must not the good mood
der Wirtschaft.⁶
of.the economy

‘Many of the fraction agreed with him that it is the buying power of citizens that needed to be increased, not the good spirits of the economy.’

- f. [Vom Erzbischof Carl Theodor Freiherr von Dalberg]_i gibt es
from archbishop Carl Theodor Freiherr from Dalberg gives it
beispielsweise [ein Bild __i] im Stadtarchiv.⁷
for.example a picture in.the city.archives

‘For example, there is a picture of archbishop Carl Theodor Freiherr of Dalberg in the city archives.’

- g. [Gegen die wegen Ehebruchs zum Tod durch Steinigen verurteilte
against the because.of adultery to.the death by stoning sentenced
Amina Lawal]_i hat gestern in Nigeria [der zweite Berufungsprozess __i]
Amina Lawal has yesterday in Nigeria the second appeal.process
begonnen.⁸
begun

‘The second appeal process began yesterday against Amina Lawal, who was sentenced to death by stoning for adultery.’

- h. [Gegen diese Kahlschlagspolitik]_i finden derzeit bundesweit
against this clear.cutting.politics happen at.the.moment statewide
[Proteste und Streiks __i] statt.⁹
protests and strikes PRT

‘At the moment, there are state-wide protests and strikes against this destructive politics.’

⁴ Spiegel, 3/1999, p. 42.

⁵ Frankfurter Rundschau, quoted from De Kuthy (2001: 52).

⁶ taz, 16.10.2003, p. 5.

⁷ Frankfurter Rundschau, quoted from De Kuthy (2002: 7).

⁸ taz, 28.08.2003, p. 2.

⁹ Streikaufruf, Universität Bremen, 03.12.2003, p. 1.

- i. [Von den beiden, die hinzugestoßen sind], hat [einer $_{-i}$] eine Hacke, der
of the both that joined are has one a pickaxe the
andere einen Handkarren.¹⁰
other a handcart
'Of the two that joined, one had a pickaxe and the other a handcart.'
- j. ein Plan, [gegen den]_i sich nun [ein Proteststurm $_{-i}$] erhebt¹¹
a plan against which REFL now a storm.of.protests rises
'a plan against which a storm of protests has now risen'
- k. Dagegen_i jedoch regt sich jetzt [Widerstand $_{-i}$]: [...]¹²
against however rises REFL now resistance
'Resistance to this is now rising, however.'
- l. [Aus der Radprofiszene]_i kennt ihn [keiner $_{-i}$] mehr.¹³
from the cycling.professional.scene knows him nobody anymore
'Nobody from the professional cycling scene acts like they know him any-
more.'
- m. [Über das chinesische Programm der Deutschen Welle] tobt dieser
about the Chinese program of.the Deutsche Welle rages these
Tage [ein heftiger Streit $_{-i}$].¹⁴
days a hefty controversy
'Recently, there has been considerable controversy about the Chinese pro-
gram by the Deutsche Welle.'

这意味着禁止从指定语位置提取成分对于德语来说并不适用。因此，对于所有语言来说禁止从指定语位置提取成分也不对。

我们的语法与允许不连续成分的语法相似：因为不能将禁止提取整合到语法形式描述中，所以语法还是比描述自然语言所需语法的生成能力强。但是，真实语法中的限制——在当前研究中，是相关语言中对从指定语位置提取成分的限制——能确保相应具体语言的语法的生成能力与弱上下文有关语法相同。

¹⁰ Haruki Murakami, *Hard-boiled Wonderland und das Ende der Welt*, suhrkamp taschenbuch, 3197, 2000, Translation by Annelie Ortmanns and Jürgen Stalph, p. 414.

¹¹ taz, 30.12.2004, p. 6.

¹² taz, 02.09.2005, p. 18.

¹³ taz, 04.07.2005, p. 5.

¹⁴ taz, 21.10.2008, p. 12.

第十八章 Binary branching, locality, and recursion

This chapter discusses three points: section 18.1 deals with the question of whether all linguistic structures should be binary branching or not. Section 18.2 discusses the question what information should be available for selection, that is, whether governing heads can access the internal structure of selected elements or whether everything should be restricted to local selection. Finally, Section 18.3 discusses recursion and how/whether it is captured in the different grammar theories that are discussed in this book.

18.1 Binary branching

We have seen that the question of the kind of branching structures assumed has received differing treatments in various theories. Classical \bar{X} theory assumes that a verb is combined with all its complements. In later variants of GB, all structures are strictly binary branching. Other frameworks do not treat the question of branching in a uniform way: there are proposals that assume binary branching structures and others that opt for flat structures.

Haegeman (1994: Section 2.5) uses learnability arguments (rate of acquisition, see Section 13.2 on this point). She discusses the example in (1) and claims that language learners have to choose one of eight structures if flat-branching structures can occur in natural language. If, on the other hand, there are only binary-branching structures, then the sentence in (1) cannot have the structures in Figure 18.1 to start with, and therefore a learner would not have to rule out the corresponding hypotheses.

- (1) Mummy must leave now.

图 18.1: Structures with partial flat-branching

However, Haegeman (1994: 88) provides evidence for the fact that (1) has the structure in (2):

- (2) [Mummy [must [leave now]]]

The relevant tests showing this include elliptical constructions, that is, the fact that it is possible to refer to the constituents in (2) with pronouns. This means that there is actually evidence for the structure of (1) that is assumed by linguists and we therefore do not have to assume that it is just hard-wired in our brains that only binary-branching structures are allowed. Haegeman (1994: 143) mentions a consequence of the binary branching hypothesis: if all structures are binary-branching, then it is not possible to straightforwardly account for sentences with ditransitive verbs in \bar{X} theory. In \bar{X} theory, it is assumed that a head is combined with all its complements at once (see Section 2.5). So in order to account for ditransitive verbs in \bar{X} theory, an empty element (little v) has to be assumed (see Section 4.1.4).

It should have become clear in the discussion of the arguments for the Poverty of the Stimulus in Section 13.8 that the assumption that only binary-branching structures are possible is part of our innate linguistic knowledge is nothing more than pure speculation. Haegeman offers no kind of evidence for this assumption. As shown in the discussions of the various theories we have seen, it is possible to capture the data with flat structures. For example, it is possible to assume that, in English, the verb is combined with its complements in a flat structure (Pollard & Sag 1994: 39). There are sometimes theory-internal reasons for deciding for one kind of branching or another, but these are not always applicable to other theories. For example, Binding Theory in GB theory is formulated with reference to dominance relations in trees (Chomsky 1981a: 188). If one assumes that syntactic structure plays a crucial role for the binding of pronouns (see page 83), then it is possible to make assumptions about syntactic structure based on the observable binding relations (so also Section 4.1.4). Binding data have, however, received a very different treatment in various theories. In LFG, constraints on f-structure are used for Binding Theory (Dalrymple 1993), whereas Binding Theory in HPSG operates on argument structure lists (valence information that are ordered in a particular way, see Section 9.6.1).

The opposite of Haegeman's position is the argumentation for flat structures put forward by Croft (2001: Section 1.6.2). In his *Radical Construction Grammar FAQ*, Croft observes that a phrasal construction such as the one in (3a) can be translated into a Categorical Grammar lexical entry like (3b).

- (3) a. [_{VP} V NP]
 b. VP/NP

He claims that a disadvantage of Categorical Grammar is that it only allows for binary-branching structures and yet there exist constructions with more than two parts (p. 49). The exact reason why this is a problem is not explained, however. He even acknowledges himself that it is possible to represent constructions with more than two arguments in Categorical Grammar. For a ditransitive verb, the entry in Categorical Grammar of English would take the form of (4):

- (4) ((s\NP)/NP)/NP

If we consider the elementary trees for TAG in Figure 18.2 在下一页, it becomes clear that it is equally possible to incorporate semantic information into a flat tree and a binary-

branching tree. The binary-branching tree corresponds to a Categorical Grammar deriva-

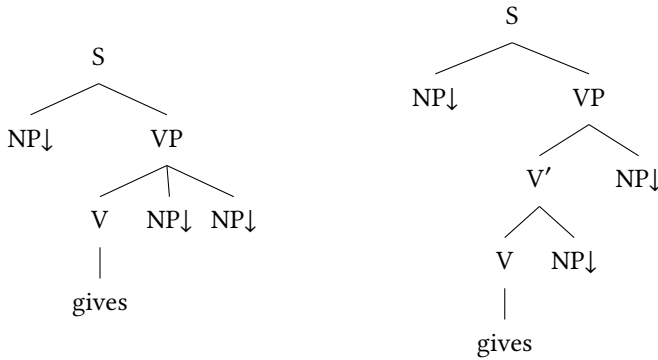


图 18.2: Flat and binary-branching elementary trees

tion. In both analyses in Figure 18.2, a meaning is assigned to a head that occurs with a certain number of arguments. Ultimately, the exact structure required depends on the kinds of restrictions on structures that one wishes to formulate. In this book, such restrictions are not discussed, but as explained above some theories model binding relations with reference to tree structures. Reflexive pronouns must be bound within a particular local domain inside the tree. In theories such as LFG and HPSG, these binding restrictions are formulated without any reference to trees. This means that evidence from binding data for one of the structures in Figure 18.2 (or for other tree structures) constitutes nothing more than theory-internal evidence.

Another reason to assume trees with more structure is the possibility to insert adjuncts on any node. In Chapter 9, an HPSG analysis for German that assumes binary-branching structures was proposed. With this analysis, it is possible to attach an adjunct to any node and thereby explain the free ordering of adjuncts in the middle field:

- (5) a. [weil] der Mann der Frau das Buch *gestern* gab
 because the man the woman the book yesterday gave
 ‘because the man gave the woman the book 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

This analysis is not the only one possible, however. One could also assume an entirely flat structure where arguments and adjuncts are dominated by one node. Kasper (1994) suggests this kind of analysis in HPSG (see also Section 5.1.5 for GPSG analyses that make

use of metarules for the introduction of adjuncts). Kasper requires complex relational constraints that create syntactic relations between elements in the tree and also compute the semantic contribution of the entire constituent using the meaning of both the verb and the adjuncts. The analysis with binary-branching structures is simpler than those with complex relational constraints and – in the absence of theory-external evidence for flat structures – should be preferred to the analysis with flat structures. At this point, one could object that adjuncts in English cannot occur in all positions between arguments and therefore the binary-branching Categorical Grammar analysis and the TAG analysis in Figure 18.2 are wrong. This is not correct, however, as it is the specification of adjuncts with regard to the adjunction site that is crucial in Categorical Grammar. An adverb has the category $(s\backslash np)\backslash(s\backslash np)$ or $(s\backslash np)/(s\backslash np)$ and can therefore only be combined with constituents that correspond to the VP node in Figure 18.2. In the same way, an elementary tree for an adverb in TAG can only attach to the VP node (see Figure 12.3 on page 379). For the treatment of adjuncts in English, binary-branching structures therefore do not make any incorrect predictions.

18.2 Locality

The question of local accessibility of information has been treated in various ways by the theories discussed in this book. In the majority of theories, one tries to make information about the inner workings of phrases inaccessible for adjacent or higher heads, that is, *glaubt* ‘believe’ in (6) selects a sentential argument but it cannot “look inside” this sentential argument.

- (6) a. Karl glaubt, dass morgen seine Schwester kommt.
 Karl believes that tomorrow his sister comes
 ‘Karl believes that his sister is coming tomorrow.’
 b. Karl glaubt, dass seine Schwester morgen kommt.
 Karl believes that his sister tomorrow comes

Thus for example, *glauben* cannot enforce that the subject of the verb has to begin with a consonant or that the complementizer has to be combined with a verbal projection starting with an adjunct. In Section 1.5, we saw that it is a good idea to classify constituents in terms of their distribution and independent of their internal structure. If we are talking about an NP box, then it is not important what this NP box actually contains. It is only of importance that a given head wants to be combined with an NP with a particular case marking. This is called *locality of selection*.

Various linguistic theories have tried to implement locality of selection. The simplest form of this implementation is shown by phrase structure grammars of the kind discussed in Chapter 2. The rule in (17) on page 55, repeated here as (7), states that a ditransitive verb can occur with three noun phrases, each with the relevant case:

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

Since the symbols for NPs do not have any further internal structure, the verb cannot require that there has to be a relative clause in an NP, for example. The internal properties of the NP are not visible to the outside. We have already seen in the discussion in Chapter 2, however, that certain properties of phrases have to be outwardly visible. This was the information that was written on the boxes themselves. For noun phrases, at least information about person, number and case are required in order to correctly capture their relation to a head. The gender value is important in German as well, since adverbial phrases such as *einer nach dem anderen* ‘one after the other’ have to agree in gender with the noun they refer to (see example (12) on page 465). Apart from that, information about the length of the noun phrases is required, in order to determine their order in a clause. Heavy constituents are normally ordered after lighter ones, and are also often extraposed (cf. Behaghel’s *Gesetz der wachsenden Glieder* ‘Law of increasing constituents’ (1909: 139; 1930: 86)).

Theories that strive to be as restrictive as possible with respect to locality therefore have to develop mechanisms that allow one to only access information that is required to explain the distribution of constituents. This is often achieved by projecting certain properties to the mother node of a phrase. In \bar{X} theory, the part of speech a head belongs to is passed up to the maximal projection: if the head is an N, for example, then the maximal projection is an NP. In GPSG, HPSG and variants of CxG, there are Head Feature Principles responsible for the projection of features. Head Feature Principles ensure that an entire group of features, so-called head features, are present on the maximal projection of a head. Furthermore, every theory has to be capable of representing the fact that a constituent can lack one of its parts and this part is then realized via a long-distance dependency in another position in the clause. As previously discussed on page 272, there are languages in which complementizers inflect depending on whether their complement is missing a constituent or not. This means that this property must be somehow accessible. In GPSG, HPSG and variants of CxG, there are additional groups of features that are present at every node between a filler and a gap in a long-distance dependency. In LFG, there is f-structure instead. Using Functional Uncertainty, one can look for the position in the f-structure where a particular constituent is missing. In GB theory, movement proceeds cyclically, that is, an element is moved into the specifier of CP and can be moved from there into the next highest CP. It is assumed in GB theory that heads can look inside their arguments, at least they can see the elements in the specifier position. If complementizers can access the relevant specifier positions, then they can determine whether something is missing from an embedded phrase or not. In GB theory, there was also an analysis of case assignment in infinitive constructions in which the case-assigning verb governs into the embedded phrase and assigns case to the element in SpecIP. Figure 18.3 shows the relevant structure taken from Haegeman (1994: 170). Since the Case Principle is formulated in such a way that only finite I can assign

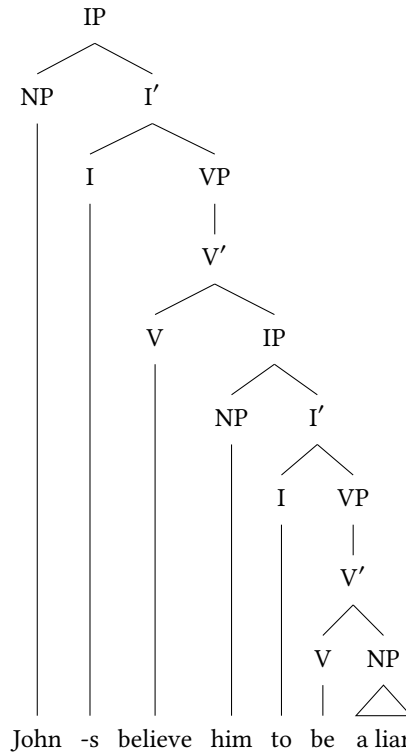


图 18.3: Analysis of the AcI construction with *Exceptional Case Marking*

case to the subject (cf. page 101), *him* does not receive case from *I*. Instead, it is assumed that the verb *believe* assigns case to the subject of the embedded infinitive.

Verbs that can assign case across phrase boundaries are referred to as ECM verbs, where ECM stands for *Exceptional Case Marking*. As the name suggests, this instance of case assignment into a phrase was viewed as an exception. In newer versions of the theory (e.g., Kratzer 1996: 120–123), all case assignment is to specifier positions. For example, the Voice head in Figure 18.4 在下一页 assigns accusative to the DP in the specifier of VP. Since the Voice head governs into the VP, case assignment to a run-of-the-mill object in this theory is an instance of exceptional case assignment as well. The same is true in Adger’s version of Minimalism, which was discussed in Chapter 4: Adger (2010) argues that his theory is more restrictive than LFG or HPSG since it is only one feature that can be selected by a head, whereas in LFG and HPSG complex feature bundles are selected. However, the strength of this kind of locality constraint is weakened by the operation Agree, which allows for nonlocal feature checking. As in Kratzer’s proposal, case is assigned nonlocally by little *v* to the object inside the VP (see Section 4.1.5.2).

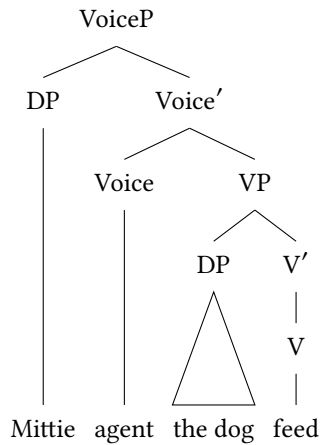


图 18.4: Analysis of structures with a transitive verb following Kratzer

Adger discusses PP arguments of verbs like *depend* and notes that these verbs need specific PPs, that is, the form of the preposition in the PP has to be selectable. While this is trivial in Dependency Grammar, where the preposition is selected right away, the respective information is projected in theories like HPSG and is then selectable at the PP node. However, this requires that the governing verb can determine at least two properties of the selected element: its part of speech and the form of the preposition. This is not possible in Adger's system and he left this for further research. Of course it would be possible to assume an onP (a phrasal projection of *on* that has the category 'on'). Similar solutions have been proposed in Minimalist theories (see Section 4.6.1 on functional projections), but such a solution would obviously miss the generalization that all prepositional phrases have something in common, which would not be covered in a system with atomic categories that are word specific.

In theories such as LFG and HPSG, case assignment takes place locally in constructions such as those in (8):

- (8)
- a. John believes him to be a liar.
 - b. Ich halte ihn für einen Lügner.
I hold him for a.ACC liar
'I take him to be a liar.'
 - c. Er scheint ein Lügner zu sein.
he seems a.NOM liar to be
'He seems to be a liar.'
 - d. Er fischt den Teich leer.
he fishes the.ACC pond empty
'He fishes (in) the pond (until it is) empty.'

Although *him*, *ihn* ‘him’, *er* ‘he’ and *den Teich* ‘the pond’ are not semantic arguments of the finite verbs, they are syntactic arguments (they are raised) and can therefore be assigned case locally. See Bresnan (1982a: 348–349 and Section 8.2) and Pollard & Sag (1994: Section 3.5) for an analysis of raising in LFG and HPSG respectively. See Meurers (1999c), Przepiórkowski (1999b), and Müller (2007b: Section 17.4) for case assignment in HPSG and for its interaction with raising.

There are various phenomena that are incompatible with strict locality and require the projection of at least some information. For example, there are question tags in English that must match the subject of the clause with which they are combined:

- (9) a. She is very smart, isn’t she / * he?
b. They are very smart, aren’t they?

Bender & Flickinger (1999), Flickinger & Bender (2003) therefore propose making information about agreement or the referential index of the subject available on the sentence node.¹ In Sag (2007), all information about phonology, syntax and semantics of the subject is represented as the value of a feature *XARG* (EXTERNAL ARGUMENT). Here, *external argument* does not stand for what it does in GB theory, but should be understood in a more general sense. For example, it makes the possessive pronoun accessible on the node of the entire NP. Sag (2007) argues that this is needed to force coreference in English idioms:

- (10) a. He_i lost [his_i / *her_j marbles].
b. They_i kept/lost [their_i / *our_j cool].

The use of the *XARG* feature looks like an exact parallel to accessing the specifier position as we saw in the discussion of GB. However, Sag proposes that complements of prepositions in Polish are also made accessible by *XARG* since there are data suggesting that higher heads can access elements inside PPs (Przepiórkowski 1999a: Section 5.4.1.2).

In Section 10.6.2 about Sign-based Construction Grammar, we already saw that a theory that only makes the reference to one argument available on the highest node of a projection cannot provide an analysis for idioms of the kind given in (11). This is because the subject is made available with verbal heads, however, it is the object that needs to be accessed in sentences such as (11). This means that one has to be able to formulate constraints affecting larger portions of syntactic structure.

- (11) a. Ich glaube, mich / # dich tritt ein Pferd.²
I believes me you kicks a horse
‘I am utterly surprised.’
b. Jonas glaubt, ihn tritt ein Pferd.³
Jonas believes him kicks a horse
‘Jonas is utterly surprised.’

¹ See also Sag & Pollard (1991: 89).

² Richter & Sailer (2009: 311).

³ <http://www.machandel-verlag.de/der-katzenschatz.html>, 06.07.2015.

- c. # Jonas glaubt, dich tritt ein Pferd.
 Jonas believes you kicks a horse
 'Jonas believes that a horse kicks you.'

Theories of grammar with extended locality domains do not have any problems with this kind of data.⁴ An example for this kind of theory is TAG. In TAG, one can specify trees of exactly the right size (Abeillé 1988; Abeillé & Schabes 1989). All the material that is fixed in an idiom is simply determined in the elementary tree. Figure 18.5 shows the tree for *kick the bucket* as it is used in (12a).

- (12) a. The cowboys kicked the bucket.
 b. Cowboys often kick the bucket.
 c. He kicked the proverbial bucket.

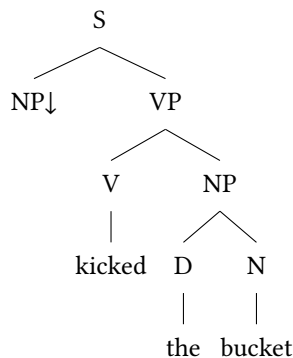


图 18.5: Elementary tree for *kick the bucket*

Since TAG trees can be split up by adjunction, it is possible to insert elements between the parts of an idiom as in (12b,c) and thus explain the flexibility of idioms with regard to adjunction and embedding.⁵ Depending on whether the lexical rules for the passive and long-distance dependencies can be applied, the idiom can occur in the relevant variants.

⁴ Or more carefully put: they do not have any serious problems since the treatment of idioms in all their many aspects is by no means trivial (Sailer 2000).

⁵ Interestingly, variants of Embodied CxG are strikingly similar to TAG. The Ditransitive Construction that was discussed on page 311 allows for additional material to occur between the subject and the verb.

The problems that arise for the semantics construction are also similar. Abeillé & Schabes (1989: 9) assume that the semantics of *John kicked the proverbial bucket* is computed from the parts *John'*, *kick-the-bucket'* and *proverbial'*, that is, the added modifiers always have scope over the entire idiom. This is not adequate for all idioms (Fischer & Keil 1996):

- (i) Er band ihr einen großen Bären auf.
 he tied her a big bear on
 'He pulled (a lot of) wool over her eyes.'

In the idiom in (i), *Bär* 'bear' actually means 'lie' and the adjective has to be interpreted accordingly. The

In cases where the entire idiom or parts of the idiom are fixed, it is possible to rule out adjunction to the nodes of the idiom tree. Figure 18.6 shows a pertinent example from Abeillé & Schabes (1989: 7). The ban on adjunction is marked by a subscript NA.

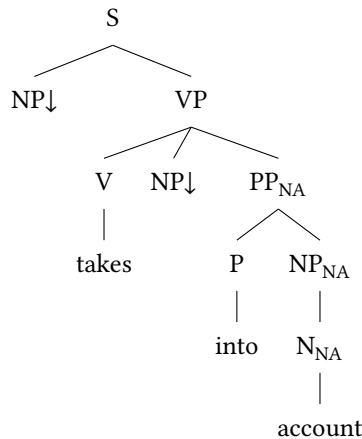


图 18.6: Elementary tree for *take into account*

The question that also arises for other theories is whether the efforts that have been made to enforce locality should be abandoned altogether. In our box model in Section 1.5, this would mean that all boxes were transparent. Since plastic boxes do not allow all of the light through, objects contained in multiple boxes cannot be seen as clearly as those in the topmost box (the path of Functional Uncertainty is longer). This is parallel to a suggestion made by Kay & Fillmore (1999) in CxG. Kay and Fillmore explicitly represent all the information about the internal structure of a phrase on the mother node and therefore have no locality restrictions at all in their theory. In principle, one can motivate this kind of theory in parallel to the argumentation in Chapter 17. The argument there made reference to the complexity of the grammatical formalism: the kind of complexity that the language of description has is unimportant, it is only important what one does with it. In the same way, one can say that regardless of what kind of information is in principle accessible, it is not accessed if this is not permitted. This was the approach taken by Pollard & Sag (1987: 143–145).

It is also possible to assume a world in which all the boxes contain transparent areas where it is possible to see parts of their contents. This is more or less the LFG world: the information about all levels of embedding contained in the f-structure is visible to both the inside and the outside. We have already discussed Nordlinger's (1998) LFG analysis

relevant tree should therefore contain nodes that contribute semantic information and also say something about the composition of these features.

In the same way, when computing the semantics of noun phrases in TAG and Embodied Construction Grammar, one should bear in mind that the adjective that is combined with a discontinuous NP Construction (see page 309) or an NP tree can have narrow scope over the noun (*all alleged murderers*).

of Wambaya on page 278. In Wambaya, words that form part of a noun phrase can be distributed throughout the clause. For example, an adjective that refers to a noun can occur in a separate position from it. Nordlinger models this by assuming that an adjective can make reference to an argument in the *f*-structure and then agrees with it in terms of case, number and gender. Bender (2008c) has shown that this analysis can be transferred to HPSG: instead of no longer representing an argument on the mother node after it has been combined with a head, simply marking the argument as realized allows us to keep it in the representation (Meurers 1999c; Przepiórkowski 1999b; Müller 2007b: Section 17.4). Detmar Meurers compares both of these HPSG approaches to different ways of working through a shopping list: in the standard approach taken by Pollard & Sag (1994), one tears away parts of the shopping list once the relevant item has been found. In the other case, the relevant item on the list is crossed out. At the end of the shopping trip, one ends up with a list of what has been bought as well as the items themselves.

I have proposed the crossing-out analysis for depictive predicates in German and English (Müller 2004a, 2008a). Depictive predicates say something about the state of a person or object during the event expressed by a verb:

- (13) a. Er sah sie nackt.⁶
 he saw her naked
 b. He saw her naked.

In (13), the depictive adjective can either refer to the subject or the object. However, there is a strong preference for readings where the antecedent noun precedes the depictive predicate (Lötscher 1985: 208). Figure 18.7 on the following page shows analyses for the sentences in (14):

- (14) a. dass er_i die Äpfel_j ungewaschen_{i/j} isst
 that he the apples unwashed eats
 ‘that he eats the apples unwashed’
 b. dass er_i ungewaschen_{i/*j} die Äpfel_j isst
 that he unwashed the apples eats
 ‘that he eats the apples (while he is) unwashed’

Arguments that have been realized are still represented on the upper nodes, however, they are crossed-out and thereby marked as “realized”. In German, this preference for the antecedent noun can be captured by assuming a restriction that states that the antecedent noun must not yet have been realized.

It is commonly assumed for English that adjuncts are combined with a VP.

- (15) a. John [[_{VP} ate the apples_i] unwashed_i].
 b. You can’t [[_{VP} give them_i injections] unconscious_i].⁷

⁶ Haider (1985a: 94).

⁷ Simpson (2005a: 17).

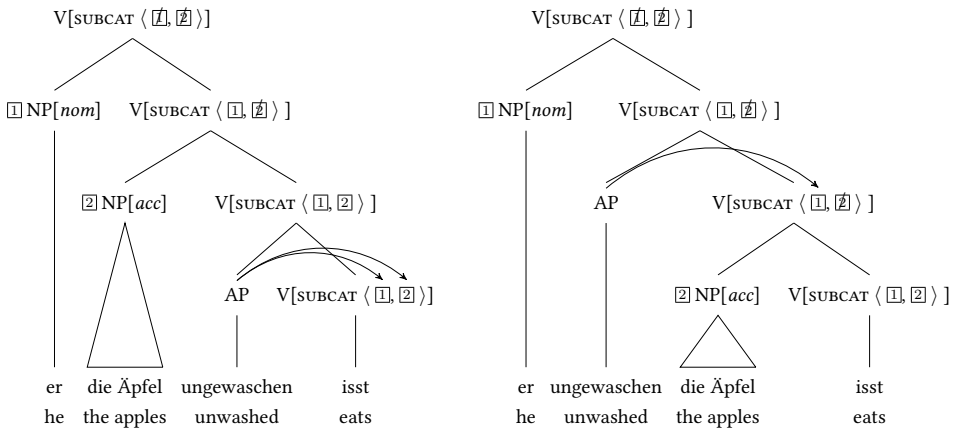


图 18.7: Analysis of *dass er die Äpfel ungewaschen isst* ‘that he the apples unwashed eats’ and *dass er ungewaschen die Äpfel isst* ‘that he unwashed the apples eat’

In approaches where the arguments of the verb are accessible at the VP node, it is possible to establish a relation between the depictive predicate and an argument although the antecedent noun is inside the VP. English differs from German in that depictives can refer to both realized (*them* in (15b)) and unrealized (*you* in (15b)) arguments.

Higginbotham (1985: 560) and Winkler (1997) have proposed corresponding non-cancellation approaches in GB theory. There are also parallel suggestions in Minimalist theories: checked features are not deleted, but instead marked as already checked (Stabler 2011b: 14). However, these features are still viewed as inaccessible.

Depending on how detailed the projected information is, it can be possible to see adjuncts and argument in embedded structures as well as their phonological, syntactic and semantic properties. In the CxG variant proposed by Kay and Fillmore, all information is available. In LFG, information about grammatical function, case and similar properties is accessible. However, the part of speech is not contained in the f-structure. If the part of speech does not stand in a one-to-one relation to grammatical function, it cannot be restricted using selection via f-structure. Nor is phonological information represented completely in the f-structure. If the analysis of idioms requires nonlocal access to phonological information or part of speech, then this has to be explicitly encoded in the f-structure (see Bresnan (1982c: 46–50) for more on idioms).

In the HPSG variant that I adopt, only information about arguments is projected. Since arguments are always represented by descriptions of type *synsem*, no information about their phonological realization is present. However, there are daughters in the structure so that it is still possible to formulate restrictions for idioms as in TAG or Construction Grammar (see Richter & Sailer (2009) for an analysis of the ‘horse’ example in (11a)). This may seem somewhat like overkill: although we already have the tree structure, we are still projecting information about arguments that have already been

realized (unfortunately these also contain information about their arguments and so on). At this point, one could be inclined to prefer TAG or LFG since these theories only make use of one extension of locality: TAG uses trees of arbitrary or rather exactly the necessary size and LFG makes reference to a complete f-structure. However, things are not quite that simple: if one wants to create a relation to an argument when adjoining a depictive predicate in TAG, then one requires a list of possible antecedents. Syntactic factors (e.g., reference to dative vs. accusative noun phrases, to argument vs. adjuncts, coordination of verbs vs. nouns) play a role in determining the referent noun, this cannot be reduced to semantic relations. Similarly, there are considerably different restrictions for different kinds of idioms and these cannot all be formulated in terms of restrictions on f-structure since f-structure does not contain information about parts of speech.

One should bear in mind that some phenomena require reference to larger portions of structure. The majority of phenomena can be treated in terms of head domains and extended head domains, however, there are idioms that go beyond the sentence level. Every theory has to account for this somehow.

18.3 Recursion

Every theory in this book can deal with self-embedding in language as it was discussed on page 4. The example (2) is repeated here as (16):

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

Most theories capture this directly with recursive phrase structure rules or dominance schemata. TAG is special with regard to recursion since recursion is factored out of the trees. The corresponding effects are created by an adjunction operation that allows any amount of material to be inserted into trees. It is sometimes claimed that Construction Grammar cannot capture the existence of recursive structure in natural language (e.g., Leiss 2009: 269). This impression is understandable since many analyses are extremely surface-oriented. For example, one often talks of a [Sbj TrVerb Obj] construction. However, the grammars in question also become recursive as soon as they contain a sentence embedding or relative clause construction. A sentence embedding construction could have the form [Sbj that-Verb that-S], where a that-Verb is one that can take a sentential complement and that-S stands for the respective complement. A *that*-clause can then be inserted into the that-S slot. Since this *that*-clause can also be the result of the application of this construction, the grammar is able to produce recursive structures such as those in (17):

- (17) Otto claims [_{that-S} that Karl suspects [_{that-S} that Richard sleeps]].

In (17), both *Karl suspects that Richard sleeps* and the entire clause are instances of the [Sbj that-Verb that-S] construction. The entire clause therefore contains an embedded subpart that is licensed by the same construction as the clause itself. (17) also contains a

constituent of the category *that*-S that is embedded inside of *that*-S. For more on recursion and self-embedding in Construction Grammar, see Verhagen (2010).

Similarly, every Construction Grammar that allows a noun to combine with a genitive noun phrase also allows for recursive structures. The construction in question could have the form [Det N NP[gen]] or [N NP[gen]]. The [Det N NP[gen]] construction licenses structures such as (18):

- (18) [NP des Kragens [NP des Mantels [NP der Vorsitzenden]]]
 the collar of.the coat of.the chairwoman
 ‘the collar of the coat of the chairwoman’

Jurafsky (1996) and Bannard, Lieven & Tomasello (2009) use probabilistic context-free grammars (PCFG) for a Construction Grammar parser with a focus on psycholinguistic plausibility and modeling of acquisition. Context-free grammars have no problems with self-embedding structures like those in (18) and thus this kind of Construction Grammar itself does not encounter any problems with self-embedding.

Goldberg (1995: 192) assumes that the resultative construction for English has the following form:

- (19) [SUBJ [V OBJ OBL]]

This corresponds to a complex structure as assumed for elementary trees in TAG. LTAG differs from Goldberg’s approach in that every structure requires a lexical anchor, that is, for example (19), the verb would have to be fixed in LTAG. But in Goldberg’s analysis, verbs can be inserted into independently existing constructions (see Section 21.1). In TAG publications, it is often emphasized that elementary trees do not contain any recursion. The entire grammar is recursive however, since additional elements can be added to the tree using adjunction and – as (17) and (18) show – insertion into substitution nodes can also create recursive structures.

第十九章 Empty elements

This chapter deals with empty elements, I first discuss the general attitude of various research traditions towards empty elements and then show how they can be eliminated from grammars (Section 19.2). Section 19.3 discusses empty elements that have been suggested in order to facilitate semantic interpretation. Section 19.4 discusses possible motivation for empty elements with a special focus on cross-linguistic comparison and the final Section 19.5 shows that certain accounts with transformations, lexical rules, and empty elements can be translated into each other.

19.1 Views on empty elements

One point that is particularly controversial among proponents of the theories discussed in this book is the question of whether one should assume empty elements or not. The discussion of empty elements is quite old: there was already some investigation in 1961 with reference to phrase structure grammars (Bar-Hillel, Perles & Shamir 1961). The discussion of the status of empty elements has carried on ever since (see Löbner 1986; Wunderlich 1987, 1989; von Stechow 1989; Haider 1997a; Sag 2000; Bouma, Malouf & Sag 2001a; Levine & Hukari 2006; Müller 2014c; Arnold & Spencer 2015, for example). There are sometimes empirical differences between analyses that assume empty elements and those that do not (Arnold & Spencer 2015), but often this is not the case. Since empty elements often feature prominently in the argumentation for or against particular theories, I will discuss how they have been used in somewhat more detail here.

In GB theory, empty elements were assumed for traces of movement (verb movement and fronting of phrases) as well as for deleted elements in elliptical constructions. Starting with the analysis of Larson (1988), more and more empty heads have been introduced to ensure uniformity of structures and certain semantic interpretations (binding and scope, see Section 4.1.4 on little *v*). Other examples of an empty element that was introduced in order to maintain particular generalizations are the empty expletives of Coopmans (1989: 734) and Postal (2004: Chapter 1). These fill the subject position in inversion structures in English, where the position preceding the verb is occupied by a PP and not by an overt subject NP. Similarly, Grewendorf (1993: 1311) assumes that the subject position in impersonal passives and passives without subject movement is in fact occupied by an empty expletive. Also, see Newmeyer (2005: 91) and Lohnstein

(2014: 180) for this assumption with regard to the passive in German. Sternefeld (2006: Section II.3.3.3) assumes that there is an empty expletive subject in impersonal passives and subjectless sentences such as (1).

- (1) a. Mir graut.
 me.DAT scares
 ‘I am scared.’
 b. Mich düstet.
 me.ACC is.thirsty
 ‘I am thirsty.’

On page 150, we discussed Stabler’s proposal for the analysis of sentences with intransitive verbs. Since, following Chomsky (2008: 146), the element that first merges with a head is the complement, intransitive verbs pose a problem for the theory. This problem is solved by Stabler by assuming that intransitive verbs are combined with an empty object (Veenstra 1998: 61, 124). Since these silent elements do not contribute to the meaning of an expression, we are also dealing with empty expletive pronouns.

In other theories, there are researchers that reject empty elements as well as those who assume them. In Categorical Grammar, Steedman suggests an analysis of nonlocal dependencies that does without empty elements (see Section 8.5), but as Pollard (1988) has shown, Steedman’s analysis requires various kinds of type raising for NPs or a correspondingly high number of complex lexical items for relative pronouns (see Section 8.5.3). On the other hand, König (1999) uses traces. In GPSG, there is the traceless analysis of extraction by Uszkoreit (1987: 76–77) that we discussed in Section 5.4, but there is also the analysis of Gazdar, Klein, Pullum & Sag (1985: 143) that uses traces. In LFG, there are both analyses with traces (Bresnan 2001: 67) and those without (see Section 7.3 and Section 7.5). Many of the phrasal analyses in HPSG are born out of the wish to avoid empty elements (see Section 21.10). An example for this is the relative clause analysis by Sag (1997) that replaces the empty relativizer in Pollard & Sag (1994) with a corresponding phrasal rule. On the other hand we have Bender (2000) and Sag, Wasow & Bender (2003: 464), who assume a silent copula. Another attempt to eliminate empty elements from HPSG was to handle long-distance dependencies not by traces but rather in the lexicon (Bouma, Malouf & Sag 2001a). As Levine & Hukari (2006) could show, however, theories of extraction that introduce long-distance dependencies lexically have problems with the semantic interpretation of coordinate structures. For a suggestion of how to solve these problems, see Chaves (2009). There are many TAG analyses without silent elements in the lexicon (see Section 12.5 and Kroch (1987), for example), however there are variants of TAG such as that of Kallmeyer (2005: 194), where a trace is assumed for the reordering of constituents in sentences with a verbal complex. Rambow (1994: 10–11) assumes an empty head in every verb phrase (see Section 12.6.2 on V-TAG).¹ In Dependency Grammar, Mel’čuk (1988: 303; 2003: 219), Starosta (1988: 253),

¹ Note that empty elements in TAG are slightly different from empty elements in other theories. In TAG the empty elements are usually part of elementary trees, that is, they are not lexical items that are combined with other material.

Eroms (2000: 471–472), Hudson (2007: Section 3.7; 2010b: 166) and Engel (2014) assume empty elements for determiners, nouns, ellipsis, imperatives, controlled infinitives, and for coordinate structures, but Groß & Osborne (2009: 73) reject empty elements (with the exception of ellipsis, Osborne 2016).

No empty elements are assumed in Construction Grammar (Michaelis & Ruppenhofer 2001: 49–50; Goldberg 2003a: 219; Goldberg 2006: 10), the related Simpler Syntax (Culicover & Jackendoff 2005) as well as in Cognitive Grammar.² The argumentation against empty elements runs along the following lines:

1. There is no evidence for invisible objects.
2. There is no innate linguistic knowledge.
3. Therefore, knowledge about empty elements cannot be learned, which is why they cannot be assumed as part of our grammar.

This begs the question of whether all the premises on which the conclusion is based actually hold. If we consider an elliptical construction such as (2), then it is clear that a noun has been omitted:

- (2) Ich nehme den roten Ball und du den blauen.
 I take the.ACC red.ACC ball and you the.ACC blue.ACC
 ‘I’ll take the red ball and you take the blue one.’

Despite there being no noun in *den blauen* ‘the blue’, this group of words behaves both syntactically and semantically just like a noun phrase. (2) is of course not necessarily evidence for there being empty elements, because one could simply say that *den blauen* is a noun phrase consisting only of an article and an adjective (Wunderlich 1987).

Similar to the fact that it is understood that a noun is missing in (2), speakers of English know that something is missing after *like*:

- (3) Bagels, I like.

Every theory of grammar has to somehow account for these facts. It must be represented in some way that *like* in (3) behaves just like a verb phrase that is missing something. One possibility is to use traces. Bar-Hillel, Perles & Shamir (1961: 153, Lemma 4.1) have shown that it is possible to turn phrase structure grammars with empty elements into those without any. In many cases, the same techniques can be applied to the theories presented here and we will therefore discuss the point in more detail in the following section.

19.2 Eliminating empty elements from grammars

It is possible to turn a grammar with empty elements (also called *epsilon*) into a grammar without these by removing all categories that can be rewritten by an epsilon

² However, Fillmore (1988: 51) did not rule them out.

in every rule that uses such categories and then add the respective rules without the empty elements to the grammar. The following example has an epsilon rule for np. One therefore has to replace all rules containing the symbol np with new rules without this np symbol. (5) shows the result of this conversion of the grammar in (4):

- (4) $\bar{v} \rightarrow \text{np}, v$
 $\bar{v} \rightarrow \text{np}, \text{pp}, v$
 $\text{np} \rightarrow \epsilon$
- (5) $\bar{v} \rightarrow \text{np}, v$
 $\bar{v} \rightarrow v$
 $\bar{v} \rightarrow \text{np}, \text{pp}, v$
 $\bar{v} \rightarrow \text{pp}, v$

This can also lead to cases where all elements on the right-hand side of a rule are removed. Thus, what one has done is actually create a new empty category and then one has to apply the respective replacement processes again. We will see an example of this in a moment. Looking at the pair of grammars in (4)–(5), it is clear that the number of rules has increased in (5) compared to (4) despite the grammars licensing the same sequences of symbols. The fact that an NP argument can be omitted is not expressed directly in (5) but instead is implicitly contained in two rules.

If one applies this procedure to the HPSG grammar in Chapter 9, then the trace does not have a specific category such as NP. The trace simply has to be compatible with a non-head daughter. As the examples in (6) show, adjuncts, arguments and parts of verbal complexes can be extracted.

- (6) a. Er_i liest t_i die Berichte.
 he reads the reports
- b. Oft_i liest er die Berichte t_i nicht.
 often reads he the reports not
 ‘Often, he does not read the reports.’
- c. Lesen_i wird er die Berichte t_i müssen.
 read will he the reports must
 ‘He will have to read the reports.’

The relevant elements are combined with their head in a specific schema (Head-Argument Schema, Head-Adjunct Schema, Predicate Complex Schema). See Chapter 9 for the first two schemata; the Predicate Complex Schema is motivated in detail in Müller (2002a: Chapter 2; 2007b: Chapter 15). If one wishes to do without traces, then one needs further additional schemata for the fronting of adjuncts, of arguments and of parts of predicate complexes. The combination of a head with a trace is given in Figure 19.1 在下一頁. The trace-less analysis is shown in Figure 19.2 在下一頁. In Figure 19.1, the element in the SUBCAT list of *kennen* is identified with the SYNSEM value of the trace [4]. The lexical entry of the trace prescribes that the LOCAL value of the trace should be identical to the element in the INHER|SLASH list.

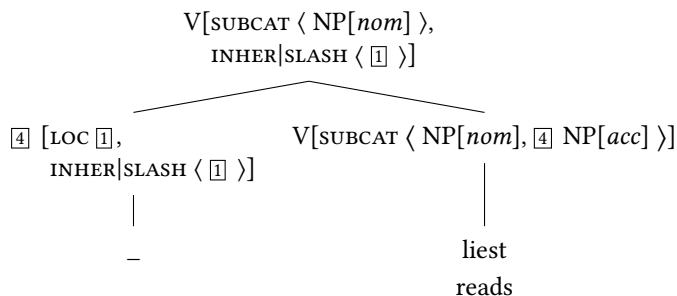


图 19.1: Introduction of information about long-distance dependencies with a trace

The Non-Local Feature Principle (page 270) ensures that the SLASH information is present on the mother node. Since an argument position gets saturated in Head-Argument structures, the accusative object is no longer contained in the SUBCAT list of the mother node.

Figure 19.2 shows the parallel trace-less structure. The effect that one gets by com-

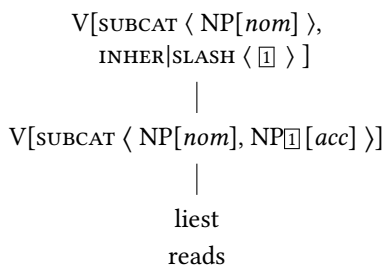


图 19.2: Introduction of information about long-distance dependencies using a unary projection

binning a trace in argument position in Head-Argument structures is represented directly on the mother node in Figure 19.2: the LOCAL value of the accusative object was identified with the element in INHER|SLASH on the mother node and the accusative object does not occur in the valence list any more.

The grammar presented in Chapter 9 contains another empty element: a verb trace. This would then also have to be eliminated.

- (7) a. Er_i liest_j t_i die Berichte t_j .
 he reads the reports
 b. Oft_i liest_j er die Berichte t_i nicht t_j .
 often reads he the reports not
 ‘Often, he does not read the reports.’

- c. Lesen_i wird_j er die Berichte t_i müssen t_j.
 read will he the reports must
 'He will have to read the reports.'

Figure 19.3 shows the combination of a verb trace with an accusative object. The verb

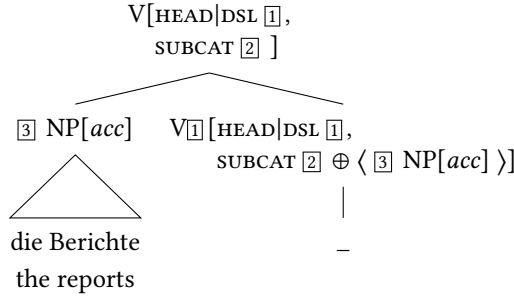


图 19.3: Analysis of verb position with verb trace

trace is specified such that the DSL value is identical to the LOCAL value of the trace (see p. 265). Since DSL is a head feature, the corresponding value is also present on the mother node. Figure 19.4 shows the structures that we get by omitting the empty node. This structure may look odd at first sight since a noun phrase is projected to a verb

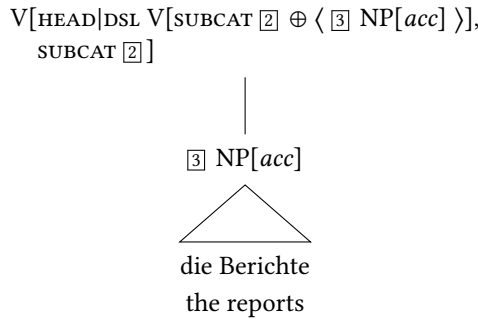


图 19.4: Analysis of verb position using a unary projection

(see page 210 for similar verb-less structures in LFG). The information about the fact that a verb is missing in the structure is equally contained in this structure as in the structure with the verb trace. It is the DSL value that is decisive for the contexts in which the structure in Figure 19.4 can appear. This is identical to the value in Figure 19.3 and contains the information that a verb that requires an accusative object is missing in the structure in question. Until now, we have seen that extraction traces can be removed from the grammar by stipulating three additional rules. Similarly, three new rules are

needed for the verb trace. Unfortunately, it does not stop here as the traces for extraction and head movement can also interact. For example, the NP in the tree in Figure 19.4 could be an extraction trace. Therefore, the combination of traces can result in more empty elements that then also have to be eliminated. Since we have three schemata, we will have three new empty elements if we combine the non-head daughter with an extraction trace and the head daughter with a verb trace. (8) shows these cases:

- (8) a. Er_i [schläft_j t_i t_j]. (Extraction trace (argument) + verb trace)
 he sleeps
 ‘He is sleeping.’
 b. Jetzt_i [schlaf_j t_i t_j]! (Extraction trace (adjunct) + verb trace)
 now sleep
 ‘Go to sleep now!’
 c. Geschlafen_i [wird_j t_i t_j]! (Extraction trace (complex) + verb trace)
 slept is
 ‘Now is time to sleep!’

These three new traces can occur as non-head daughters in the Head-Argument Schema and thus one would require three new schemata for Head-Argument structures. Using these schemata, it then becomes possible to analyze the sentences in (8).

Six further schemata are required for the examples in (9) and (10) since the three new traces can each occur as heads in Head-Argument structures (9) and Head-Adjunct structures (10):

- (9) a. Den Aufsatz_i liest_j [er t_i t_j].
 the essay reads he
 ‘He is reading the essay.’
 b. Oft_i liest_j er [ihn t_i t_j].
 often reads he it
 ‘He often reads it.’
 c. Lesen_i wird_j er [ihn t_i t_j].
 read will he it
 ‘He will read it.’
 (10) a. Den Aufsatz_i liest_j er [nicht t_i t_j].
 the essay reads he not
 ‘He isn’t reading the essay.’
 b. Oft_i liest_j er ihn [nicht t_i t_j].
 often reads he it not
 ‘He often doesn’t read it’
 c. Lesen_i wird_j er ihn [nicht t_i t_j].
 reads will he it not
 ‘He won’t read it.’

Eliminating two empty elements therefore comes at the price of twelve new rules. These rules are not particularly transparent and it is not immediately obvious why the mother node describes a linguistic object that follows general grammatical laws. For example, there are no heads in the structures following the pattern in Figure 19.4. Since there is no empirical difference between the theoretical variant with twelve additional schemata and the variant with two empty elements, one should prefer the theory that makes fewer assumptions (Occam's Razor) and that is the theory with two empty elements.

One might think that the problem discussed here is just a problem specific to HPSG not shared by trace-less analyses such as the LFG approach that was discussed in Section 7.5. If we take a closer look at the rule proposed by Dalrymple (2006: Section 2.2), we see that the situation in LFG grammars is entirely parallel. The brackets around the category symbols mark their optionality. The asterisk following the PP means that any number of PPs (zero or more) can occur in this position.

$$(11) \quad V' \rightarrow (V) (NP) PP^*$$

This means that (11) is a shorthand for rules such as those in (12):

- $$(12) \quad \begin{array}{ll} \text{a. } V' \rightarrow V \\ \text{b. } V' \rightarrow V NP \\ \text{c. } V' \rightarrow V NP PP \\ \text{d. } V' \rightarrow V NP PP PP \\ \text{e. } \dots \\ \text{f. } V' \rightarrow NP \\ \text{g. } V' \rightarrow NP PP \\ \text{h. } V' \rightarrow NP PP PP \\ \text{i. } \dots \end{array}$$

Since all the elements on the right-hand side of the rule are optional, the rule in (11) also stands for (13):

$$(13) \quad V' \rightarrow \epsilon$$

Thus, one does in fact have an empty element in the grammar although the empty element is not explicitly listed in the lexicon. This follows from the optionality of all elements on the right-hand side of a rule. The rule in (12f) corresponds to the schema licensed by the structure in Figure 19.4. In the licensed LFG structure, there is also no head present. Furthermore, one has a large number of rules that correspond to exactly the schemata that we get when we eliminate empty elements from an HPSG grammar. This fact is, however, hidden in the representational format of the LFG rules. The rule schemata of LFG allow for handy abbreviations of sometimes huge sets of rules (even infinite sets when using ‘*’).

Pollard (1988) has shown that Steedman's trace-less analysis of long-distance dependencies is not without its problems. As discussed in Section 8.5.3, a vast number of recategorization rules or lexical entries for relative pronouns are required.

19.3 Empty elements and semantic interpretation

In this section, I discuss an analysis that assumes empty elements in order to allow for different readings of particular sentences. I then show how one can use so-called underspecification approaches to do without empty elements.

Sentences such as (14) are interesting since they have multiple readings (see Dowty 1979: Section 5.6) and it is not obvious how these can be derived.

- (14) dass Max alle Fenster wieder öffnete
 that Max all windows again opened
 ‘that Max opened all the windows again’

There is a difference between a repetitive and a restitutive reading: for the repetitive reading of (14), Max has to have opened every window at least once before, whereas the restitutive reading only requires that all windows were open at some point, that is, they could have been opened by someone else.

These different readings are explained by decomposing the predicate *open'* into at least two sub-predicates. Egg (1999) suggests the decomposition into CAUSE and *open'*:

- (15) CAUSE(x , *open'*(y))

This means that there is a CAUSE operator that has scope over the relation *open'*. Using this kind of decomposition, it is possible to capture the varying scope of *wieder* ‘again’: in one of the readings, *wieder* scopes over CAUSE and it scopes over *open'* but below CAUSE in the other. If we assume that *öffnen* has the meaning in (15), then we still have to explain how the adverb can modify elements of a word’s meaning, that is, how *wieder* ‘again’ can refer to *open'*. Von Stechow (1996: 93) developed the analysis in Figure 19.5 on the next page. AgrS and AgrO are functional heads proposed for subject and object agreement in languages like Basque and have been adopted for German (see Section 4.6). Noun phrases have to be moved from the VoiceP into the specifier position of the AgrS and AgrO heads in order to receive case. T stands for Tense and corresponds to Infl in the GB theory (see Section 3.1.5 and Section 4.1.5). What is important is that there is the Voice head and the separate representation of *offen* ‘open’ as the head of its own phrase. In the figure, everything below Voice' corresponds to the verb *öffnen*. By assuming a separate Voice head that contributes causative meaning, it becomes possible to derive both readings in syntax: in the reading with narrow scope of *wieder* ‘again’, the adverb is adjoined to the XP and has scope over *open*(x). In the reading with wide scope, the adverb attaches to VoiceP or some higher phrase and therefore has scope over CAUSE(BECOME(*open*(x))).

Jäger & Blutner (2003) point out that this analysis predicts that sentences such as (16) only have the repetitive reading, that is, the reading where *wieder* ‘again’ has scope over CAUSE.

- (16) dass Max wieder alle Fenster öffnete
 that Max again all windows opened

This is because *wieder* precedes *alle Fenster* and therefore all heads that are inside VoiceP. Thus, *wieder* can only be combined with AgrOP or higher phrases and therefore has (too)

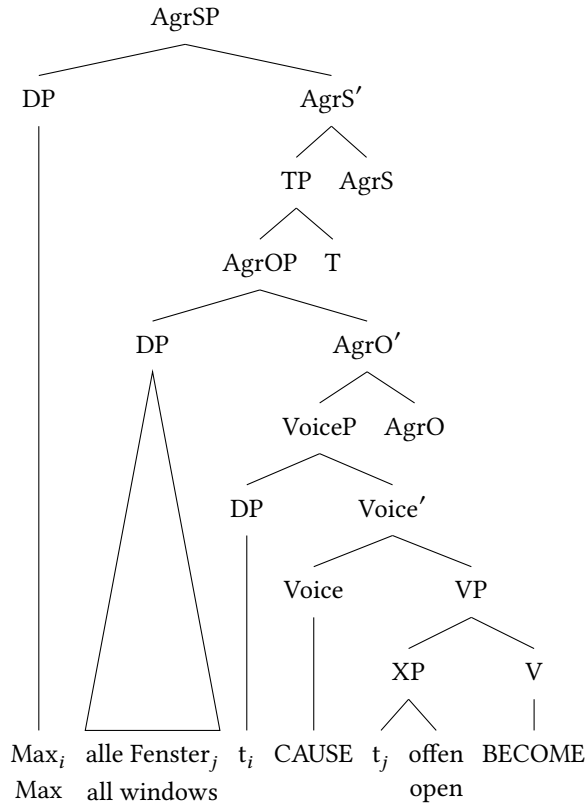


图 19.5: Decomposition in syntactic structures

wide scope. (16) does permit a restitutive reading, however: all windows were open at an earlier point in time and Max reestablishes this state.

Egg (1999) develops an analysis for these *wieder* cases using Constraint Language for Lambda-Structures (CLLS). CLLS is an underspecification formalism, that is, no logical formulae are given but instead expressions that describe logical formulae. Using this kind of expressions, it is possible to leave scope relations underspecified. I have already mentioned Minimal Recursion Semantics (MRS) (Copestake, Flickinger, Pollard & Sag 2005) in several chapters of this book. As well as CLLS, MRS together with Underspecified Discourse Representation Theory (Reyle 1993; Frank & Reyle 1995) and Hole Semantics (Bos 1996; Blackburn & Bos 2005) all belong to the class of underspecification formalisms. See Baldridge & Kruijff (2002) for an underspecification analysis in Categorical Grammar and Nerbonne (1993) for an early underspecification analysis in HPSG. In the following, I will reproduce Egg's analysis in an MRS-like notation.

Before we turn to (14) and (16), let us consider the simpler sentence in (17):

- (17) dass Max alle Fenster öffnete
 that Max all windows opened
 ‘that Max opened all the windows’

This sentence can mean that in a particular situation, it is true of all windows that Max opened them. A less readily accessible reading is the one in which Max causes all of the windows to be open. It is possible to force this reading if one rules out the first reading through contextual information (Egg 1999):

- (18) Erst war nur die Hälfte der Fenster im Bus auf, aber dann öffnete Max
 first was only the half of the windows in the bus open but then opened Max
 alle Fenster.
 all windows
 ‘At first, only half of the windows in the bus were open, but then Max opened
 all of the windows.’

Both readings under discussion here differ with regard to the scope of the universal quantifier. The reading where Max opens all the windows himself corresponds to wide scope in (19a). The reading where some windows could have already been open corresponds to (19b):

- (19) a. $\forall x \text{ window}'(x) \rightarrow \text{CAUSE}(\text{max}', \text{open}'(x))$
 b. $\text{CAUSE}(\text{max}', \forall x \text{ window}'(x) \rightarrow \text{open}'(x))$

Using underspecification, both of these readings can be represented in one dominance graph such as the one given in Figure 19.6. Each relation in Figure 19.6 has a name that

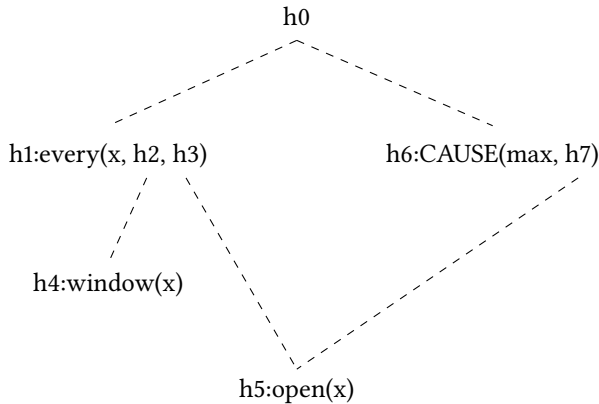


图 19.6: Dominance graph for *Max alle Fenster öffnete*

one can use to refer to the relation or “grasp” it. These names are referred to as *handle*. The dominance graph states that h_0 dominates both h_1 and h_6 and that h_2 dominates h_4 , h_3 dominates h_5 , and h_7 dominates h_5 . The exact scopal relations are underspecified:

the universal quantifier can have scope over CAUSE or CAUSE can have scope over the universal quantifier. Figures 19.7 and 19.8 show the variants of the graph with resolved scope. The underspecified graph in Figure 19.6 does not say anything about the relation

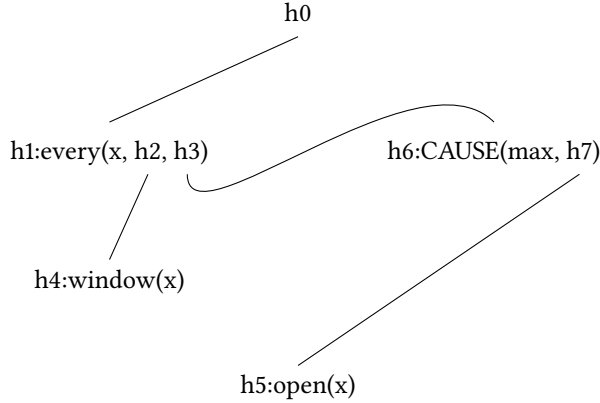


图 19.7: Dominance graph for the reading $\forall x \text{ window}(x) \rightarrow \text{CAUSE}(\text{max}, \text{open}(x))$.

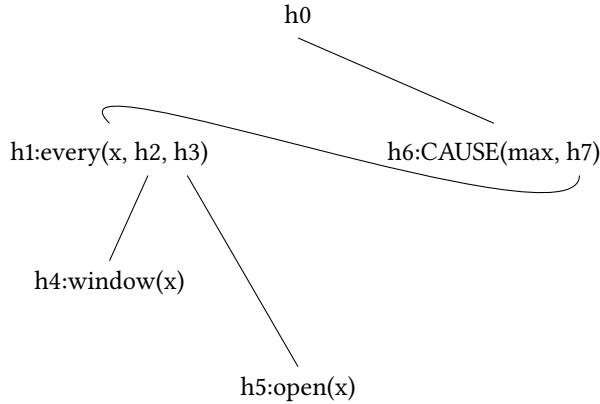


图 19.8: Graph for the reading $\text{CAUSE}(\text{max}, \forall x \text{ window}(x) \rightarrow \text{open}(x))$.

between $h3$ and $h6$. The only thing it says is that $h3$ somehow has to dominate $h5$. In Figure 19.7 $every$ ($h3$) dominates $CAUSE$ ($h6$) and $CAUSE$ dominates $open$ ($h5$). So, *every*' dominates *open*' indirectly. In Figure 19.8, $CAUSE$ dominates *every*' and *every*' dominates *open*'. Again the constraints of Figure 19.6 are fulfilled, but $h7$ dominates $h5$ only indirectly.

The fact that the quantifier dominates $h4$ is determined by the lexical entry of the quantifier. The fact that the quantifier dominates $h5$ does not have to be made explicit in the analysis since the quantifier binds a variable in the relation belonging to $h5$, namely

x. The dominance relation between $h7$ and $h5$ is always determined in the lexicon since CAUSE and *open'* both belong to the semantic contribution of a single lexical entry.

The exact syntactic theory that one adopts for this analysis is, in the end, not of great importance. I have chosen HPSG here. As Figure 19.9 shows, the analysis of *alle Fenster öffnet* contains a simple structure with a verb and an object. This structure does not differ

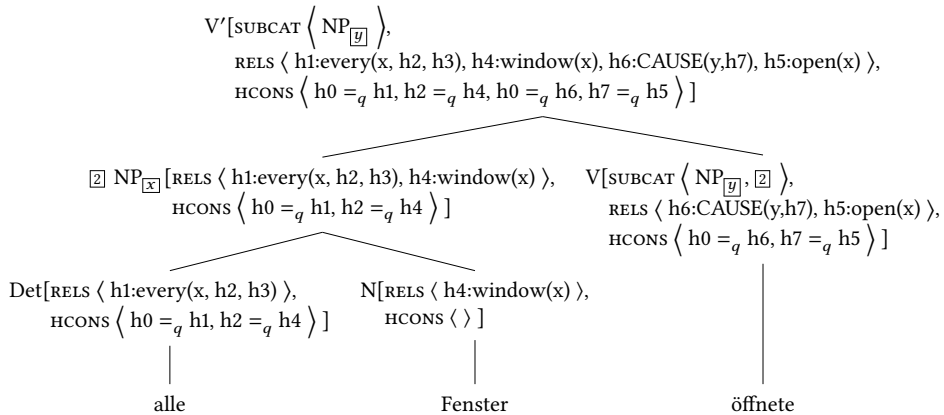


图 19.9: MRS analysis of *alle Fenster öffnet*

from the one that would be assumed for *alle Kinder kennt* ‘all children know’, involving the semantically simplex verb *kennen* ‘to know’. The only difference comes from the meaning of the individual words involved. As shown in Section 9.1.6, relations between individual words are passed on upwards. The same happens with scopal restrictions. These are also represented in lists. HCONS stands for *handle constraints*. $=_q$ in $h0 =_q h6$ stand for the equality *modulo* quantifier scope.

Egg lists the following readings for the sentence in (16) – repeated here as (20):

- (20) dass Max wieder alle Fenster öffnet
 that Max again all windows opened
 ‘that Max opened all the windows again’

1. Max opened every window and he had already done that at least once for each window (*again'*(\forall (CAUSE(open))); repetitive)
2. Max caused every window to be open and he had done that at least once before (*again'*(CAUSE(\forall (open))); repetitive)
3. At some earlier point in time, all windows were simultaneously open and Max re-established this state (CAUSE(*again'*(\forall (open))); restitutive)

These readings correspond to the dominance graph in Figure 19.10 on the following page. Figure 19.11 on the next page shows the graph for (14) – repeated here as (21):

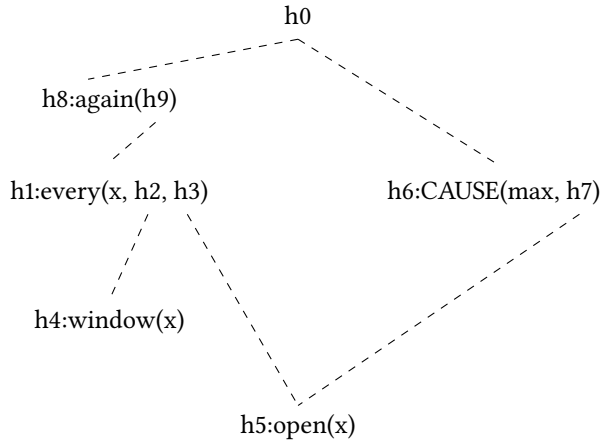


图 19.10: Dominance graph for *Max wieder alle Fenster öffnete* ‘that Max opened all the windows again’

- (21) dass Max alle Fenster wieder öffnete
that Max all windows again opened

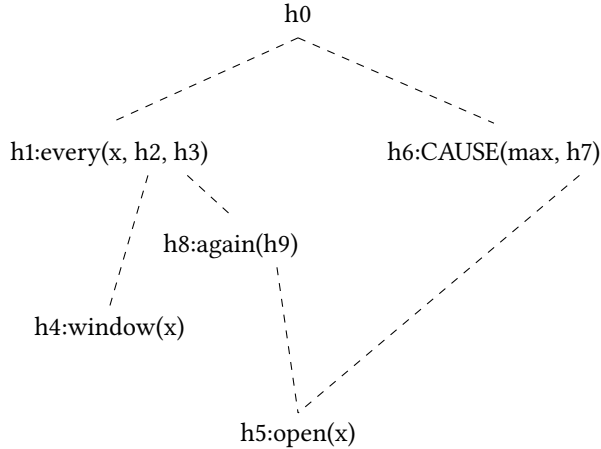


图 19.11: Dominance graph for *Max alle Fenster wieder öffnete* ‘that Max opened all the windows again’

To derive these dominance graphs from the ones without *wieder* ‘again’, all one has to do is add the expression $h8:again(h9)$ and the dominance requirements that demand that $h9$ dominates quantifiers occurring to the right of *wieder* and that it is dominated by quantifiers to the left of *wieder*.

It is therefore unproblematic to derive the relevant readings for modification by *wieder* without empty elements for CAUSE and BECOME. The meaning of the word *öffnen* is decomposed in a similar way but the decomposed meaning is assigned to a single element, the verb. By underspecification of the scopal relations in the lexicon, the relevant readings can then be derived.

19.4 Evidence for empty elements

As previously discussed, grammarians agree that both linguists and speakers notice when there is a constituent missing from a string of words. For cases where it can be shown that analyses with or without traces are indistinguishable empirically, then one can assume empty elements. Nevertheless, the learnability argument put forward by Construction Grammarians has some validity: if one assumes that there is no or little innate linguistic knowledge, then it is not possible to motivate empty elements with data from other languages. This means that just because Basque shows object agreement, this does not mean that one can assume an empty head for object agreement (AgrO) in a grammar of German as for instance von Stechow (1996) and Meinunger (2000) do. Since there is no object agreement in German, there would be no way for the child to learn the fact that there is an AgrO head. Knowledge about AgrO must therefore be innate. Since the assumption of innate linguistic knowledge is controversial (see Chapter 13), any theory that uses cross-linguistic data to motivate the use of empty elements is on shaky ground.

Cross-linguistic considerations can only be drawn upon if there are no empirical differences between multiple alternative analyses compatible and motivated by the language under consideration. In this case, one should follow Occam's Razor and choose the analysis which is compatible with analyses of other languages (see Müller 2015a and Chapter 22.2).

19.5 Transformations, lexical rules, and empty elements

In the discussion of the passive in the framework of TAG, it became clear that lexical rules correspond to particular transformations, namely those which have some relation to a lexical item (lexically governed transformations, Dowty 1978; for the discussion of transformations and lexical rules, see Bresnan (1978) and Bresnan & Kaplan (1982)). In the respective variants of TAG, lexical rules establish a relation between a lexical item for an active tree with a lexical item of a passive tree. Both the active and passive tree can be extended by adjunction.

In theories such as Categorical Grammar, the situation is similar: since the direction in which a functor expects to find its argument is fixed for languages such as English, the lexical item stands for an entire tree. Only the attachment of adjuncts is not yet specified in lexical items. The positions in the tree where the adjuncts can occur depend on the properties of the adjuncts. In Section 8.4, we have seen suggestions for treatments of languages with free constituent order. If the direction of combination is not fixed in the

lexicon, then the lexical item can occur in a number of trees. If we compare lexical rules that can be applied to this kind of lexical items with transformations, we see that lexical rules create relations between different sets of trees.

In HPSG analyses, this works in a similar way: lexical rules relate lexical items with differing valence properties to each other. In HPSG grammars of English, there is normally a schema that licenses a VP containing the verb and all its complements as well as a schema that connects the subject to the VP (Pollard & Sag 1994: 39). In the lexical items for finite verbs, it is already determined what the tree will look like in the end. As in Categorical Grammar, adjuncts in HPSG can be combined with various intermediate projections. Depending on the dominance schemata used in a particular grammar, the lexical item will determine the constituent structure in which it can occur or allow for multiple structures. In the grammar of German proposed in Chapter 9, it is possible to analyze six different sequences with a lexical item for a ditransitive verb, that is, the lexical item can – putting adjuncts aside – occur in six different structures with verb-final order. Two sequences can be analyzed with the passive lexical item, which only has two arguments. As in Categorical Grammar, sets of licensed structures are related to other sets of licensed structures. In HPSG theorizing and also in Construction Grammar, there have been attempts to replace lexical rules with other mechanisms since their “status is dubious and their interaction with other analyses is controversial” (Bouma, Malouf & Sag 2001a: 19). Bouma et al. (2001a) propose an analysis for extraction that, rather than connecting lexical items with differing valence lists, establishes a relation between a subset of a particular list in a lexical item and another list in the same lexical item. The results of the two alternative analyses are shown in (22) and (23), respectively:

- (22) a.
$$\left[\begin{array}{l} \text{SUBCAT} \langle \text{NP}[\textit{nom}], \text{NP}[\textit{acc}] \rangle \\ \text{SLASH} \langle \rangle \end{array} \right]$$
- b.
$$\left[\begin{array}{l} \text{SUBCAT} \langle \text{NP}[\textit{nom}] \rangle \\ \text{SLASH} \langle \text{NP}[\textit{acc}] \rangle \end{array} \right]$$

In (22), (22a) is the basic entry and (22b) is related to (22a) via a lexical rule. The alternative analysis would only involve specifying the appropriate value of the ARG-ST feature³ and the SUBCAT and SLASH value is then derived from the ARG-ST value using the relevant constraints. (23) shows two of the licensed lexical items.

- (23) a.
$$\left[\begin{array}{l} \text{ARG-ST} \langle \text{NP}[\textit{nom}], \text{NP}[\textit{acc}] \rangle \\ \text{SUBCAT} \langle \text{NP}[\textit{nom}], \text{NP}[\textit{acc}] \rangle \\ \text{SLASH} \langle \rangle \end{array} \right]$$
- b.
$$\left[\begin{array}{l} \text{ARG-ST} \langle \text{NP}[\textit{nom}], \text{NP}[\textit{acc}] \rangle \\ \text{SUBCAT} \langle \text{NP}[\textit{nom}] \rangle \\ \text{SLASH} \langle \text{NP}[\textit{acc}] \rangle \end{array} \right]$$

³ ARG-ST stands for *Argument Structure*. The value of ARG-ST is a list containing all the arguments of a head. For more on ARG-ST, see Section 9.6.1.

If we want to eliminate lexical rules entirely in this way, then we would require an additional feature for each change.⁴ Since there are many interacting valence-changing processes, things only work out with the stipulation of a large number of auxiliary features. The consequences of assuming such analyses have been discussed in detail in Müller (2007b: Section 7.5.2.2). The problems that arise are parallel for inheritance-based approaches for argument structure-changing processes: they also require auxiliary features since it is not possible to model embedding and multiple changes of valence information with inheritance. See Section 10.2.

Furthermore, the claim that the status of lexical rules is dubious must be rejected: there are worked-out formalizations of lexical rules (Meurers 2001; Copestake & Briscoe 1992; Lascarides & Copestake 1999) and their interaction with other analyses is not controversial. Most HPSG implementations make use of lexical rules and the interaction of a number of rules and constraints can be easily verified by experiments with implemented fragments.

Jackendoff (1975) presents two possible conceptions of lexical rules: in one variant, the lexicon contains all words in a given language and there are just redundancy rules saying something about how certain properties of lexical entries behave with regard to properties of other lexical entries. For example, *les-* ‘read-’ and *lesbar* ‘readable’ would both have equal status in the lexicon. In the other way of thinking of lexical rules, there are a few basic lexical entries and the others are derived from these using lexical rules. The stem *les-* ‘read-’ would be the basic entry and *lesbar* would be derived from it. In HPSG, the second of the two variants is more often assumed. This is equivalent to the assumption of unary rules. In Figure 9.8 on page 261, this has been shown accordingly: the verb *kennt* ‘knows’ is mapped by a lexical rule to a verb that selects the projection of an empty verbal head. With this conception of lexical rules, it is possible to remove lexical rules from the grammar by assuming binary-branching structures with an empty head rather than unary rules. For example, in HPSG analyses of resultative constructions such as (24), lexical rules have been proposed (Verspoor 1997; Wechsler 1997; Wechsler & Noh 2001; Müller 2002a: Chapter 5).

- (24) [dass] Peter den Teich leer fischt
 that Peter the pond empty fishes
 ‘that Peter fishes the pond empty’

In my own analysis, a lexical rule connects a verb used intransitively to a verb that selects an accusative object and a predicate. Figure 19.12 on the following page shows the corresponding tree. If we consider what (24) means, then we notice that the fishing act causes the pond to become empty. This causation is not contained in any of the basic lexical items for the words in (24). In order for this information to be present in the semantic representation of the entire expression, it has to be added by means of a lexical

⁴ Alternatively, one could assume a very complex relation that connects ARG-ST and SUBCAT. But this would then have to deliver the result of an interaction of a number of phenomena and the interaction of these phenomena would not be captured in a transparent way.

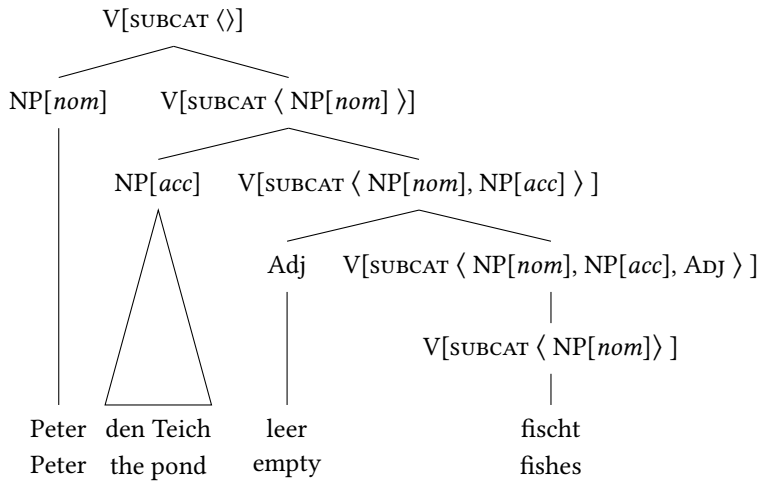


图 19.12: Analysis of the resultative construction with a lexical rule

rule. The lexical rule says: if a verb is used with an additional predicate and accusative object, then the entire construction has a causative meaning.

Figure 19.13 在下一页 shows how a lexical rule can be replaced by an empty head. The empty head requires the intransitive verb and additionally an adjective, an accusative object and a subject. The subject of *fischt* ‘fishes’ must of course be identical to the subject that is selected by the combination of *fischt* and the empty head. This is not shown in the figure. It is possible, however, to establish this identity (see Hinrichs & Nakazawa 1994). The causative semantics is contributed by the empty head in this analysis. The trick that is being implemented here is exactly what was done in Section 19.2, just in the opposite direction: in the previous section, binary-branching structures with an empty daughter were replaced by unary-branching structures. In this section, we have replaced unary-branching structures with binary-branching structures with an empty daughter.⁵

We have therefore seen that certain transformations can be replaced by lexical rules and also that lexical rules can be replaced by empty heads. The following chapter deals with the question of whether phenomena like extraction, scrambling, and passive should be described with the same tool as in GB/Minimalism or with different tools as in LFG

⁵ Here, we are discussing lexical rules, but this transformation trick can also be applied to other unary rules. Semanticists often use such rules for type shifting. For example, a rule that turns a referential NP such as *a trickster* in (i.a) into a predicative one (i.b) (Partee 1987).

- (i) a. A trickster laughs.
b. He is a trickster.

These changes can be achieved by a unary rule that is applied to an NP or with a special empty head that takes an NP as its argument. In current Minimalist approaches, empty heads are used (Ramchand 2005: 370), in Categorical Grammar and HPSG unary-branching rules are more common (Flickinger 2008: 91–92; Müller 2009c, 2012).

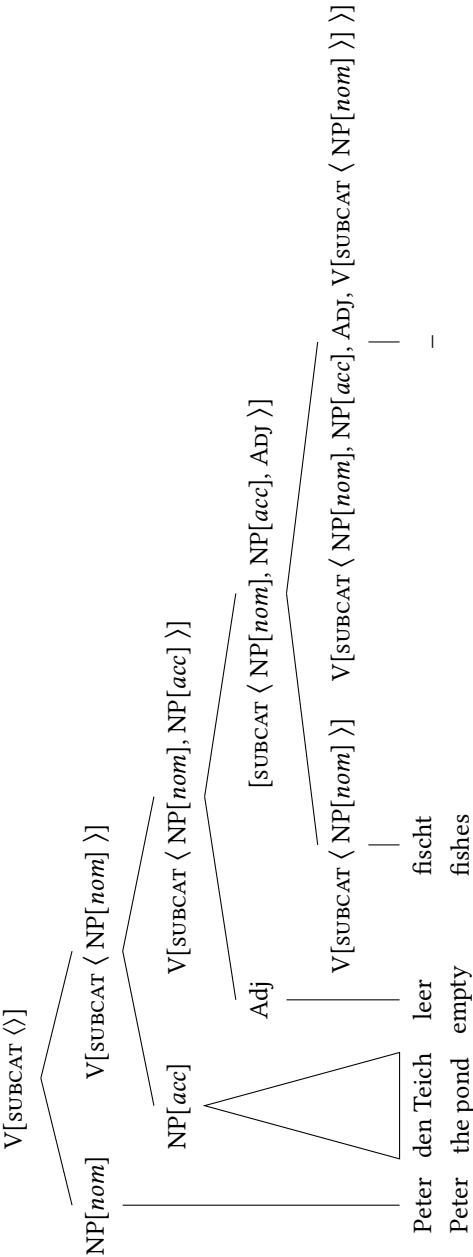


图 19.13: Analysis of the resultative construction with an empty head

19 Empty elements

and HPSG.

第二十章 **Extraction, scrambling, and passive: one or several descriptive devices?**

An anonymous reviewer suggested discussing one issue in which transformational theories differ from theories like LFG and HPSG. The reviewer claimed that Transformational Grammars use just one tool for the description of active/passive alternations, scrambling, and extraction, while theories like LFG and HPSG use different techniques for all three phenomena. If this claim were correct and if the analyses made correct predictions, the respective GB/Minimalism theories would be better than their competitors, since the general aim in science is to develop theories that need a minimal set of assumptions. I already commented on the analysis of passive in GB in Section 3.4, but I want to extend this discussion here and include a Minimalist analysis and one from Dependency Grammar.

The task of any passive analysis is to explain the difference in argument realization in examples like (1):

- (1) a. She beats him.
- b. He was beaten.

In these examples about chess, the accusative object of *beat* is realized as the nominative in (1b). In addition, it can be observed that the position of the elements is different: while *him* is realized postverbally in object position in (1a), it is realized preverbally in (1b). In GB this is explained by a movement analysis. It is assumed that the object does not get case in passive constructions and hence has to move into the subject position where case is assigned by the finite verb. This analysis is also assumed in Minimalist work as in David Adger's textbook (2003), for instance. Figure 20.1 on the following page shows his analysis of (2):

- (2) Jason was killed.

TP stands for Tense Phrase and corresponds to the IP that was discussed in Chapter 3. PassP is a functional head for passives. *v*P is a special category for the analysis of verb phrases that was originally introduced for the analysis of ditransitives (Larson 1988) and VP is the normal VP that consists of verb and object. In Adger's analysis, the verb *kill* moves from the verb position in VP to the head position of *v*, the passive auxiliary *be* moves from the head position of PassP to the head position of the Tense Phrase. Features

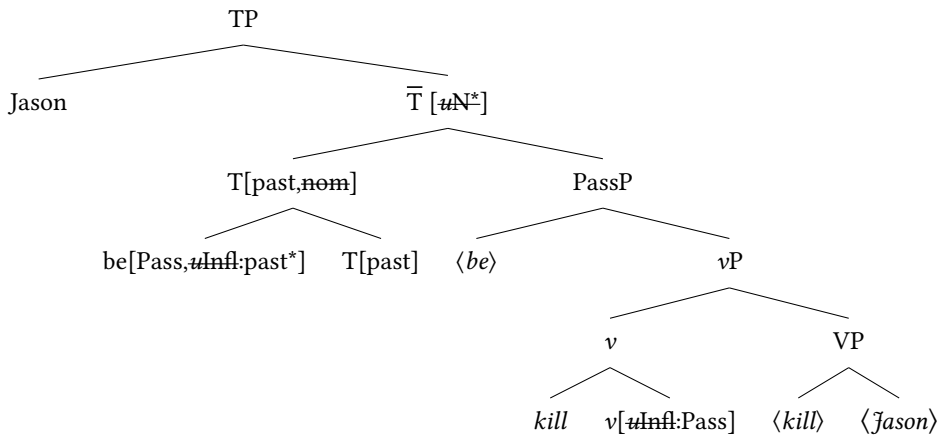


图 20.1: Adger's Minimalist movement-based analysis of the passive (p. 231)

like Infl are 'checked' in combination with such movements. The exact implementation of these checking and valuing operations does not matter here. What is important is that *Jason* moves from the object position to a position that was formerly known as the specifier position of T (see Footnote 26 on page 145 on the notion of specifier). All these analyses assume that the participle cannot assign accusative to its object and that the object has to move to another position to get case or check features. How exactly one can formally represent the fact that the participle cannot assign case is hardly ever made explicit in the GB literature. The following is a list of statements that can be found in the literature:

- (3) a. We shall assume that a passivized verb loses the ability to assign structural ACCUSATIVE case to its complement. (Haegeman 1994: 183)
- b. das Objekt des Aktivsatzes wird zum Subjekt des Passivsatzes, weil die passivische Verbform keinen Akkusativ-Kasus regieren kann (Akk-Kasus-Absorption). (Lohnstein 2014: 172)

In addition, it is sometimes said that the external theta-role is absorbed by the verb morphology (Jaeggli 1986; Haegeman 1994: 183). Now, what would it entail if we made this explicit? There is some lexical item for verbs like *beat*. The active form has the ability to assign accusative to its object, but the passive form does not. Since this is a property that is shared by all transitive verbs (by definition of the term transitive verb), this is some regularity that has to be captured. One way to capture this is the assumption of a special passive morpheme that suppresses the agent and changes something in the case specification of the stem it attaches too. How this works in detail was never made explicit. Let us compare this morpheme-based analysis with lexical rule-based analyses: as was explained in Section 19.5, empty heads can be used instead of lexical rules in those cases in which the phonological form of the input and the output do not differ. So for ex-

ample, lexical rules that license additional arguments as in resultative constructions, for instance, can be replaced by an empty head. However, as was explained in Section 9.2, lexical rules are also used to model morphology. This is also true for Construction Grammar (see Gert Booij's work on Construction Morphology (2010), which is in many ways similar to Riehemann's work in HPSG (1993; 1998)). In the case of the passive lexical rule, the participle morphology is combined with the stem and the subject is suppressed in the corresponding valence list. This is exactly what is described in the GB/MP literature. The respective lexical rule for the analysis of *ge-lieb-t* 'loved' is depicted in Figure 20.2 to the left. The morpheme-based analysis is shown to the right. To keep things simple, I

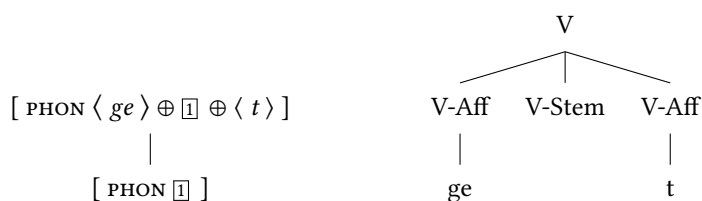


图 20.2: Lexical rule-based/constructionist vs. morpheme-based analysis

assume a flat analysis, but those who insist on binary branching structures would have to come up with a way of deciding whether the *ge-* or the *-t* is combined first with the stem and in which way selection and percolation of features takes place. Independent of how morphology is done, the fact that the inflected form (the top node in both figures) has different properties than the verb stem has to be represented somehow. In the morpheme-based world, the morpheme is responsible for suppressing the agent and changing the case assignment properties, in the lexical rule/construction world this is done by the respective lexical rule. There is no difference in terms of needed tools and necessary stipulations.

The situation in Minimalist theories is a little bit different. For instance, (Adger 2003: 229, 231) writes the following:

Passives are akin to unaccusatives, in that they do not assign accusative case to their object, and they do not appear to have a thematic subject. [...] Moreover, the idea that the function of this auxiliary is to select an unaccusative little *vP* simultaneously explains the lack of accusative case and the lack of a thematic subject. (Adger 2003: 229, 231)

So this is an explicit statement. The relation between a stem and a passive participle form that was assumed in GB analyses is now a verb stem that is combined with two different versions of little *v*. Which *v* is chosen is determined by the governing head, a functional Perf head or a Pass head. This can be depicted as in Figure 20.3 on the following page. When *kill* is used in the perfect or the passive, it is spelled out as *killed*. If it is used in the active with a 3rd person singular subject it is spelled out as *kills*. This can be compared with a lexical analysis, for instance the one assumed in HPSG. The analysis is shown in

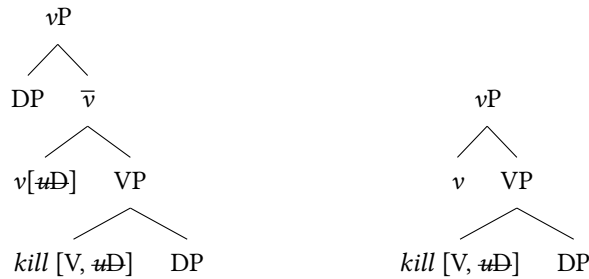


图 20.3: Analysis of the passive and the perfect and the passive in a Minimalist theory involving two different versions of little *v*

Figure 20.4. The left figure shows a lexical item that is licensed by a lexical rule that is

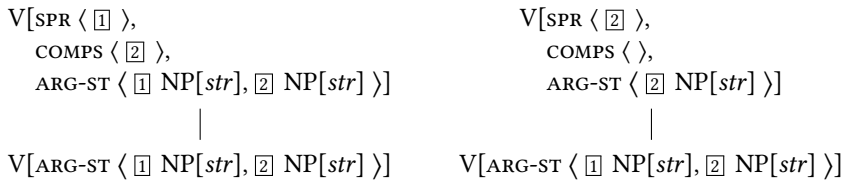


Figure 20.4: Lexical rule-based analysis of the perfect and the passive in HPSG

applied to the stem *kill-*. The stem has two elements in its argument structure list and for the active forms the complete argument structure list is shared between the licensed lexical item and the stem. The first element of the ARG-ST list is mapped to SPR and the other elements to COMPS (in English). Passive is depicted in the right figure: the first element of the ARG-ST with structural case is suppressed and since the element that was the second element in the ARG-ST list of the stem ($\boxed{2}$) is now the first element, this item is mapped to SPR. See Section 9.2 for passive in HPSG and Section 9.6.1 for comments on ARG-ST and the differences between German and English.

The discussion of Figures 20.3 and 20.4 are a further illustration of a point made in Section 19.5: lexical rules can be replaced by empty heads and vice versa. While HPSG says there are stems that are related to inflected forms and corresponding to the inflection the arguments are realized in a certain way, Minimalist theories assume two variants of little *v* that differ in their selection of arguments. Now, the question is: are there empirical differences between the two approaches? I think there are differences if one considers the question of language acquisition. What children can acquire from data is that there are various inflected forms and that they are related somehow. What remains questionable is whether they really would be able to detect empty little *vs*. One could claim of course that children operate with chunks of structures such as the ones in Figure 20.3. But then a verb would be just a chunk consisting of little *v* and *V* and

having some open slots. This would be indistinguishable from what the HPSG analysis assumes.

As far as the “lexical rules as additional tool” aspect is concerned, the discussion is closed, but note that the standard GB/Minimalism analyses differ in another way from LFG and HPSG analyses, since they assume that passive has something to do with movement, that is, they assume that the same mechanisms are used that are used for nonlocal dependencies.¹ This works for languages like English in which the object has to be realized in postverbal position in the active and in preverbal position in the passive, but it fails for languages like German in which the order of constituents is more free. Lenerz (1977: Section 4.4.3) discussed the examples in (44) on page 101 – which are repeated here as (4) for convenience:

- (4)
- a. weil das Mädchen dem Jungen den Ball schenkt
 because the girl the.DAT boy the.ACC Ball gives
 ‘because the girl gives the ball to the boy’
 - b. weil dem Jungen der Ball geschenkt wurde
 because the.DAT boy the.NOM ball given was
 - c. weil der Ball dem Jungen geschenkt wurde
 because the.NOM ball the.DAT boy given was
 ‘because the ball was given to the boy’

While both orders in (4b) and (4c) are possible, the one with dative–nominative order in (4b) is the unmarked one. There is a strong linearization preference in German demanding that animate NPs be serialized before inanimate ones (Hoberg 1981: 46). This linearization rule is unaffected by passivization. Theories that assume that passive is movement either have to assume that the passive of (4a) is (4c) and (4b) is derived from (4c) by a further reordering operation (which would be implausible since usually one assumes that more marked constructions require more transformations), or they would have to come up with other explanations for the fact that the subject of the passive sentence has the same position as the object in active sentences. As was already explained in Section 3.4, one such explanation is to assume an empty expletive subject that is placed in the position where nominative is assigned and to somehow connect this expletive element to the subject in object position. While this somehow works, it should be clear that the price for rescuing a movement-based analysis of passive is rather high: one has to assume an empty expletive element, that is, something that neither has a form nor a meaning. The existence of such an object could not be inferred from the input unless it is assumed that the structures in which it is assumed are given. Thus, a rather rich UG would have to be assumed.

¹ There is another option in Minimalist theories. Since Agree can check features nonlocally, T can assign nominative to an embedded element. So, in principle the object may get nominative in the VP without moving to T. However, Adger (2003: 368) assumes that German has a strong EPP feature on T, so that the underlying object has to move to the specifier of T. This is basically the old GB analysis of passive in German with all its conceptual problems and disadvantages.

The question one needs to ask here is: why does the movement-based analysis have these problems and why does the valence-based analysis not have them? The cause of the problem is that the analysis of the passive mixes two things: the fact that SVO languages like English encode subjecthood positionally, and the fact that the subject is suppressed in passives. If these two things are separated the problem disappears. The fact that the object of the active sentence in (1a) is realized as the subject in (1b) is explained by the assumption that the first NP on the argument structure list with structural case is realized as subject and mapped to the respective valence feature: *SPR* in English. Such mappings can be language specific (see Section 9.6.1 and Müller (2016a) where I discuss Icelandic, which is an SVO language with subjects with lexical case).

In what follows, I discuss another set of examples that are sometimes seen as evidence for a movement-based analysis. The examples in (5) are instances of the so-called remote passive (Höhle 1978: 175–176).²

- (5) a. daß er auch von mir zu überreden versucht wurde³
 that he.NOM also from me to persuade tried got
 ‘that an attempt to persuade him was also made by me’
 b. weil der Wagen oft zu reparieren versucht wurde
 because the car.NOM often to repair tried was
 ‘because many attempts were made to repair the car’

What is interesting about these examples is that the subject is the underlying object of a deeply embedded verb. This seems to suggest that the object is extracted out of the verb phrase. So the analysis of (5b) would be (6):

- (6) weil [_{IP} der Wagen_i [_{VP} oft [_{VP} [_{VP} [_{VP} _{-i} zu reparieren] versucht] wurde]
 because the car.NOM often to repair tried was

While this is a straight-forward explanation of the fact that (5b) is grammatical, another explanation is possible as well. In the HPSG analysis of German (and Dutch) it is assumed that verbs like those in (5b) form a verbal complex, that is, *zu reparieren versucht wurde* ‘to repair tried was’ forms one unit. When two or more verbs form a complex, the highest verb attracts the arguments from the verb it embeds (Hinrichs & Nakazawa 1989c, 1994; Bouma & van Noord 1998). A verb like *versuchen* ‘to try’ selects a subject, an infinitive with *zu* ‘to’ and all complements that are selected by this infinitive. In the analysis of (7), *versuchen* ‘to try’ selects for its subject, the object of *reparieren* ‘to repair’ and for the verb *zu reparieren* ‘to repair’.

- (7) weil er den Wagen zu reparieren versuchen will
 because he.NOM the.ACC car to repair try wants
 ‘because he wants to try to repair the car’

² See Müller (2002a: Section 3.1.4.1) and Wurmbrand (2003b) for corpus examples.

³ Oppenrieder (1991: 212).

Now if the passive lexical rule applies to *versuch-*, it suppresses the first argument of *versuch-* with structural case, which is the subject of *versuch-*. The next argument of *versuch-* is the object of *zu reparieren*. Since this element is the first NP with structural case, it gets nominative as in (5b). So, this shows that there is an analysis of the remote passive that does not rely on movement. Since movement-based analyses were shown to be problematic and since there are no data that cannot be explained without movement, analyses without movement have to be preferred.

This leaves us with movement-based accounts of local reordering (scrambling). The reviewer suggested that scrambling, passive, and nonlocal extraction may be analyzed with the same mechanism. It was long thought that scope facts made the assumption of movement-based analyses of scrambling necessary, but it was pointed out by Kiss (2001: 146) and Fanselow (2001: Section 2.6) that the reverse is true: movement-based accounts of scrambling make wrong predictions with regard to available quantifier scopings. I discussed the respective examples in Section 3.5 already and will not repeat the discussion here. The conclusion that has to be drawn from this is that passive, scrambling, and long distance extraction are three different phenomena that should be treated differently. The solution for the analysis of the passive that is adopted in HPSG is based on an analysis by Haider (1986a), who worked within the GB framework. The “scrambling-as-base generation” approach to local reordering that was used in HPSG right from the beginning (Gunji 1986) is also adopted by some practitioners of GB/Minimalism, e.g., Fanselow (2001).

Having discussed the analyses in GB/Minimalism, I now turn to Dependency Grammar. Groß & Osborne (2009) suggest that *w*-fronting, topicalization, scrambling, extraposition, splitting, and also the remote passive should be analyzed by what they call *rising*. The concept was already explained in Section 11.5. The Figures 20.5 and 20.6 show examples for the fronting and the scrambling of an object. Groß and Osborne assume

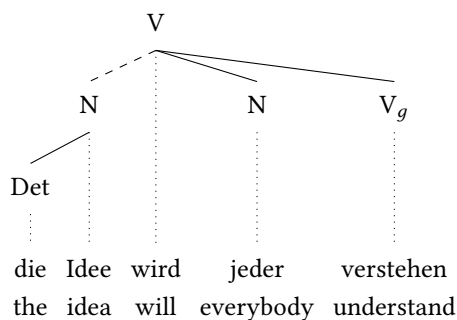


图 20.5: Analysis of *Die Idee wird jeder verstehen*. ‘Everybody will understand the idea.’ involving rising

that the object depends on the main verb in sentences with auxiliary verbs, while the subject depends on the auxiliary. Therefore, the object *die Idee* ‘the idea’ and the object *sich* ‘himself’ have to rise to the next higher verb in order to keep the structures projec-

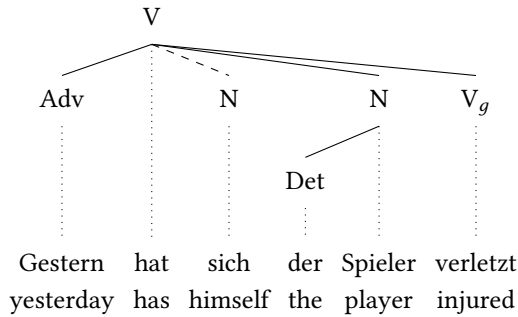


图 20.6: Analysis of *Gestern hat sich der Spieler verletzt*. ‘Yesterday, the player injured himself.’ involving rising of the object of the main verb *verletzt* ‘injured’

tive. Figure 20.7 shows the analysis of the remote passive. The object of *zu reparieren* ‘to

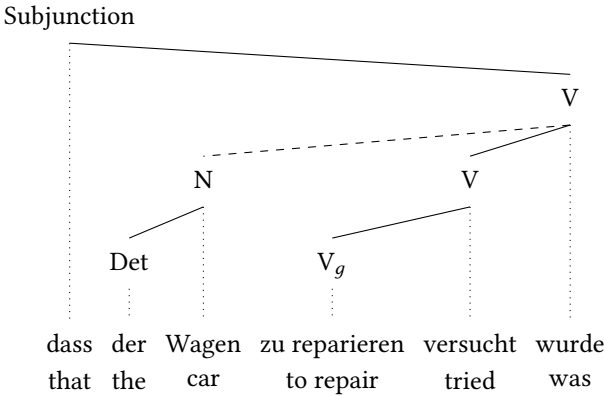


图 20.7: Analysis of the remote passive *dass der Wagen zu reparieren versucht wurde* ‘that it was tried to repair the car’ involving rising

repair’ rises to the auxiliary *wurde* ‘was’.

Groß and Osborne use the same mechanism for all these phenomena, but it should be clear that there have to be differences in the exact implementation. Groß and Osborne say that English does not have scrambling, while German does. If this is to be captured, there must be a way to distinguish the two phenomena, since if this were not possible, one would predict that English has scrambling as well, since both German and English allow long distance fronting. Groß & Osborne (2009: 58) assume that object nouns that rise must take the nominative. But if the kind of rising that they assume for remote passives is identical to the one that they assume for scrambling, they would predict that *den Wagen* gets nominative in (8) as well:

- (8) dass den Wagen niemand repariert hat
 that the.ACC car nobody.NOM repaired has
 ‘that nobody repaired the car’

Since *den Wagen* ‘the car’ and *repariert* ‘repaired’ are not adjacent, *den Wagen* has to rise to the next higher head in order to allow for a projective realization of elements. So in order to assign case properly, one has to take into account the arguments that are governed by the head to which a certain element rises. Since the auxiliary *hat* ‘has’ already governs a nominative, the NP *den Wagen* has to be realized in the accusative. An analysis that assumes that both the accusative and nominative depend on *hat* ‘has’ in (8) is basically the verbal complex analysis assumed in HPSG and some GB variants.

Note, however, that this does not extend to nonlocal dependencies. Case is assigned locally by verbs or verbal complexes, but not to elements that come from far away. The long distance extraction of NPs is more common in southern variants of German and there are only a few verbs that do not take a nominative argument themselves. The examples below involve *dünken* ‘to think’, which governs an accusative and a sentential object and *scheinen* ‘to seem’, which governs a dative and a sentential object. If (9a) is analyzed with *den Wagen* rising to *dünkt*, one might expect that *den Wagen* ‘the car’ gets nominative since there is no other element in the nominative. However, (8b) is entirely out.

- (9) a. Den Wagen dünkt mich, dass er repariert.
 the.ACC car thinks me.ACC that he.NOM repairs
 ‘I think that he repairs the car’
 b. *Der Wagen dünkt mich, dass er repariert.
 the.NOM car thinks me.ACC that he.NOM repairs

Similarly there is no agreement between the fronted element and the verb to which it attaches:

- (10) a. Mir scheint, dass die Wagen ihm gefallen.
 me.DAT.1PL seems.3SG that the cars.3PL him please.3PL
 ‘He seems to me to like the cars.’
 b. Die Wagen scheint mir, dass ihm gefallen.
 the cars.3PL seem.3SG me.DAT that him please.3PL
 ‘The cars, he seems to me to like.’
 c. *Die Wagen scheinen mir, dass ihm gefällt.
 the cars.3PL seem.3PL me.DAT that him pleases.3SG
 d. *Die Wagen scheinen mir, dass ihm gefallen.
 the cars.3PL seem.3PL me.DAT that him please.3PL

This shows that scrambling/remote passive and extraction should not be dealt with by the same mechanism or if they are dealt with by the same mechanism one has to make sure that there are specialized variants of the mechanism that take the differences into

account. I think what Groß and Osborne did is simply recode the attachment relations of phrase structure grammars. *die Idee* ‘the idea’ has some relation to *wird jeder verstehen* ‘will everybody understand’ in Figure 20.5, as it does in GB, LFG, GPSG, HPSG, and other similar frameworks. In HPSG, *die Idee* ‘the idea’ is the filler in a filler-head configuration. The remote passive and local reorderings of arguments of auxiliaries, modal verbs, and other verbs that behave similarly are explained by verbal complex formation where all non-verbal arguments depend on the highest verb (Hinrichs & Nakazawa 1994).

Concluding this chapter, it can be said that local reorderings and long-distance dependencies are two different things that should be described with different tools (or there should be further constraints that differ for the respective phenomena when the same tool is used). Similarly, movement-based analyses of the passive are problematic since passive does not necessarily imply reordering.

第二十一章 Phrasal vs. lexical analyses

coauthored with Stephen Wechsler

This section deals with a rather crucial aspect when it comes to the comparison of the theories described in this book: valence and the question whether sentence structure, or rather syntactic structure in general, is determined by lexical information or whether syntactic structures have an independent existence (and meaning) and lexical items are just inserted into them. Roughly speaking, frameworks like GB/Minimalism, LFG, CG, HPSG, and DG are lexical, while GPSG and Construction Grammar (Goldberg 1995, 2003a; Tomasello 2003, 2006b; Croft 2001) are phrasal approaches. This categorization reflects tendencies, but there are non-lexical approaches in Minimalism (Borer's exoskeletal approach, 2003) and LFG (Alsina 1996; Asudeh et al. 2008, 2013) and there are lexical approaches in Construction Grammar (Sign-Based Construction Grammar, see Section 10.6.2). The phrasal approach is wide-spread also in frameworks like Cognitive Grammar (Dąbrowska 2001; Langacker 2009: 169) and Simpler Syntax (Culicover & Jackendoff 2005; Jackendoff 2008) that could not be discussed in this book.

The question is whether the meaning of an utterance like (1a) is contributed by the verb *give* and the structure needed for the NPs occurring together with the verb does not contribute any meaning or whether there is a phrasal pattern [X Verb Y Z] that contributes some “ditransitive meaning” whatever this may be.¹

- (1) a. Peter gives Mary the book.
b. Peter fishes the pond empty.

¹ Note that the prototypical meaning is a transfer of possession in which Y receives Z from X, but the reverse holds in (i.b):

- (i) a. Er gibt ihr den Ball.
he.NOM gives her.DAT the.ACC ball
b. Er stiehlt ihr den Ball.
he.NOM steals her.DAT the.ACC ball
'He steals the ball from her.'

Similarly, there is the question of how the constituents in (1b) are licensed. This sentence is interesting since it has a resultative meaning that is not part of the meaning of the verb *fish*: Peter’s fishing causes the pond to become empty. Nor is this additional meaning part of the meaning of any other item in the sentence. On the lexical account, there is a lexical rule that licenses a lexical item that selects for *Peter*, *the pond*, and *empty*. This lexical item also contributes the resultative meaning. On the phrasal approach, it is assumed that there is a pattern [Subj V Obj Obl]. This pattern contributes the resultative meaning, while the verb that is inserted into this pattern just contributes its prototypical meaning, e.g., the meaning that *fish* would have in an intransitive construction. I call such phrasal approaches *plugging approaches*, since lexical items are plugged into ready-made structures that do most of the work.

In what follows I will examine these proposals in more detail and argue that the lexical approaches to valence are the correct ones. The discussion will be based on earlier work of mine (Müller 2006, 2007c, 2010b) and work that I did together with Steve Wechsler (Müller & Wechsler 2014a,b). Some of the sections in Müller & Wechsler (2014a) started out as translations of Müller (2013b), but the material was reorganized and refocused due to intensive discussion with Steve Wechsler. So rather than using a translation of Section 11.11 of Müller (2013b), I use parts of Müller & Wechsler (2014a) here and add some subsections that had to be left out of the article due to space restrictions (Subsections 21.3.6 and 21.7.3). Because there have been misunderstandings in the past (e.g., Boas (2014), see Müller & Wechsler (2014b)), a disclaimer is necessary here: this section is not an argument against Construction Grammar. As was mentioned above Sign-Based Construction Grammar is a lexical variant of Construction Grammar and hence compatible with what I believe to be correct. This section is also not against phrasal constructions in general, since there are phenomena that seem to be best captured with phrasal constructions. These are discussed in detail in Subsection 21.10. What I will argue against in the following subsections is a special kind of phrasal construction, namely phrasal argument structure constructions (phrasal ASCs). I believe that all phenomena that have to do with valence and valence alternations should be treated lexically.

21.1 Some putative advantages of phrasal models

In this section we examine certain claims to purported advantages of phrasal versions of Construction Grammar over lexical rules. Then in the following section, we will turn to positive arguments for lexical rules.

21.1.1 Usage-based theories

For many practitioners of Construction Grammar, their approach to syntax is deeply rooted in the ontological strictures of *usage-based* theories of language (Langacker 1987; Goldberg 1995; Croft 2001; Tomasello 2003). Usage-based theorists oppose the notion of “linguistic rules conceived of as algebraic procedures for combining symbols that do not themselves contribute to meaning” (Tomasello 2003: 99). All linguistic entities are

symbolic of things in the realm of denotations; “all have communicative significance because they all derive directly from language use” (*ibid*). Although the formatives of language may be rather abstract, they can never be divorced from their functional origin as a tool of communication. The usage-based view of constructions is summed up well in the following quote:

The most important point is that constructions are nothing more or less than patterns of usage, which may therefore become relatively abstract if these patterns include many different kinds of specific linguistic symbols. But never are they empty rules devoid of semantic content or communicative function. (Tomasello 2003: 100)

Thus constructions are said to differ from grammatical rules in two ways: they must carry meaning; and they reflect the actual “patterns of usage” fairly directly.

Consider first the constraint that every element of the grammar must carry meaning, which we call the *semiotic dictum*. Do lexical or phrasal theories hew the most closely to this dictum? Categorical Grammar, the paradigm of a lexical theory (see Chapter 8), is a strong contender: it consists of meaningful words, with only a few very general combinatorial rules such as $X/Y * Y = X$. Given the rule-to-rule assumption, those combinatorial rules specify the meaning of the whole as a function of the parts. Whether such a rule counts as meaningful in itself in Tomasello’s sense is not clear.

What does seem clear is that the combinatorial rules of Construction Grammar, such as Goldberg’s Correspondence Principle for combining a verb with a construction (1995: 50), have the same status as those combinatorial rules:

- (2) The Correspondence Principle: each participant that is lexically profiled and expressed must be fused with a profiled argument role of the construction. If a verb has three profiled participant roles, then one of them may be fused with a non-profiled argument role of a construction. (Goldberg 1995: 50)

Both verbs and constructions are specified for participant roles, some of which are *profiled*. Argument profiling for verbs is “lexically determined and highly conventionalized” (Goldberg 1995: 46). Profiled argument roles of a construction are mapped to direct grammatical functions, i. e., SUBJ, OBJ, or OBJ2. By the Correspondence Principle the lexically profiled argument roles must be direct, unless there are three of them, in which case one may be indirect.² With respect to the semiotic dictum, the Correspondence Principle has the same status as the Categorical Grammar combinatorial rules: a meaningless algebraic rule that specifies the way to combine meaningful items.

Turning now to the lexicalist syntax we favor, some elements abide by the semiotic dictum while others do not. Phrase structure rules for intransitive and transitive VPs (or the respective HPSG ID schema) do not. Lexical valence structures clearly carry meaning since they are associated with particular verbs. In an English ditransitive, the first object expresses the role of “intended recipient” of the referent of the second object.

² We assume that the second sentence of (2) provides for exceptions to the first sentence.

Hence *He carved her a toy* entails that he carved a toy with the intention that she receive it. So the lexical rule that adds a benefactive recipient argument to a verb adds meaning. Alternatively, a phrasal ditransitive construction might contribute that “recipient” meaning.³ Which structures have meaning is an empirical question for us.

In Construction Grammar, however, meaning is assumed for all constructions *a priori*. But while the ditransitive construction plausibly contributes meaning, no truth-conditional meaning has yet been discovered for either the intransitive or bivalent transitive constructions. Clearly the constructionist’s evidence for the meaningfulness of *certain* constructions such as the ditransitive does not constitute evidence that *all* phrasal constructions have meaning. So the lexical and phrasal approaches seem to come out the same, as far as the semiotic dictum is concerned.

Now consider the second usage-based dictum, that the elements of the grammar directly reflect patterns of usage, which we call *the transparency dictum*. The Construction Grammar literature often presents their constructions informally in ways that suggest that they represent surface constituent order patterns: the transitive construction is “[X VERB Y]” (Tomasello) or “[Subj V Obj]” (Goldberg 1995, 2006)⁴; the passive construction is “X was VERBed by Y” (Tomasello 2003: 100) or “Subj aux Vpp (PPby)” (Goldberg 2006: 5). But a theory in which constructions consist of surface patterns was considered in detail and rejected by Müller (2006: Section 2), and does not accurately reflect Goldberg’s actual theory.⁵ The more detailed discussions present *argument structure constructions*, which are more abstract and rather like the lexicalists’ grammatical elements (or perhaps an LFG f-structure): the transitive construction resembles a transitive valence structure (minus the verb itself); the passive construction resembles the passive lexical rule.

With respect to fulfilling the desiderata of usage-based theorists, we do not find any significant difference between the non-lexical and lexical approaches.

21.1.2 Coercion

Researchers working with plugging proposals usually take coercion as an indication of the usefulness of phrasal constructions. For instance, Anatol Stefanowitsch (Lecture in the lecture series *Algorithmen und Muster – Strukturen in der Sprache*, 2009) discussed the example in (3):

- (3) Das Tor zur Welt Hrnglb öffnete sich ohne Vorwarnung und verschlang [sie] ...
die Welt Hrnglb wird von Magiern erschaffen, die Träume zu Realität formen

³ In Section 21.2.1 we argue that the recipient should be added in the lexical argument structure, not through a phrasal construction. See Wechsler (1991: 111–113; 1995: 88–89) for an analysis of English ditransitives with elements of both constructional and lexical approaches. It is based on Kiparsky’s notion of a *thematically restricted positional linker* (1987; 1988).

⁴ Goldberg et al. (2004: 300) report about a language acquisition experiment that involves an SOV pattern. The SOV order is mentioned explicitly and seen as part of the construction.

⁵ This applies to argument structure constructions only. In some of her papers Goldberg assumes that very specific phrase structural configurations are part of the constructions. For instance in her paper on complex predicates in Persian (Goldberg 2003b) she assigns V^0 and \bar{V} categories. See Müller (2010b: Section 4.9) for a critique of that analysis.

können, aber nicht in der Lage sind zu träumen. Haltet aus, Freunde. Und ihr da draußen, bitte träumt ihnen ein Tor.⁶

The crucial part is *bitte träumt ihnen ein Tor* ‘Dream a gate for them’. In this fantasy context the word *träumen*, which is intransitive, is forced into the ditransitive construction and therefore gets a certain meaning. This forcing of a verb corresponds to overwriting or rather extending properties of the verb by the phrasal construction.

In cases in which the plugging proposals assume that information is overwritten or extended, lexical approaches assume mediating lexical rules. Briscoe & Copestake (1999: Section 4) have worked out a lexical approach in detail.⁷ They discuss the ditransitive sentences in (4), which either correspond to the prototypical ditransitive construction (4a) or deviate from it in various ways.

- (4)
- a. Mary gave Joe a present.
 - b. Joe painted Sally a picture.
 - c. Mary promised Joe a new car.
 - d. He tipped Bill two pounds.
 - e. The medicine brought him relief.
 - f. The music lent the party a festive air.
 - g. Jo gave Bob a punch.
 - h. He blew his wife a kiss.
 - i. She smiled herself an upgrade.

For the non-canonical examples they assume lexical rules that relate transitive (*paint*) and intransitive (*smile*) verbs to ditransitive ones and contribute the respective semantic information or the respective metaphorical extension. The example in (4i) is rather similar to the *träumen* example discussed above and is also analyzed with a lexical rule (page 509). Briscoe and Copestake note that this lexical rule is much more restricted in its productivity than the other lexical rules they suggest. They take this as motivation for developing a representational format in which lexical items (including those that are derived by lexical rules) are associated with probabilities, so that differences in productivity of various patterns can be captured.

Looking narrowly at such cases, it is hard to see any rational grounds for choosing between the phrasal analysis and the lexical rule. But if we broaden our view, the lexical rule approach can be seen to have a much wider application. Coercion is a very general pragmatic process, occurring in many contexts where no construction seems to be responsible (Nunberg 1995). Nunberg cites many cases such as the restaurant waiter asking *Who is the ham sandwich?* (Nunberg 1995: 115). Copestake & Briscoe (1992: 116)

⁶ http://www.elbenwaldforum.de/showflat.php?Cat=&Board=Tolkiens_Werke&Number=1457418&page=3&view=collapsed&sb=5&o=&fpart=16, 27.02.2010.

‘The gate to the world Hrnglb opened without warning and swallowed them. The world Hrnglb is created by magicians that can form reality from dreams but cannot dream themselves. Hold out, friends! And you out there, please, dream a gate for them.’

⁷ Kay (2005), working in the framework of CxG, also suggests unary constructions.

discuss the conversion of terms for animals to mass nouns (see also Copestake & Briscoe (1995: 36–43)). Example (5) is about a substance, not about a cute bunny.

- (5) After several lorries had run over the body, there was rabbit splattered all over the road.

The authors suggest a lexical rule that maps a count noun onto a mass noun. This analysis is also assumed by Fillmore (1999: 114–115). Such coercion can occur without any syntactic context: one can answer the question *What's that stuff on the road?* or *What are you eating?* with the one-word utterance *Rabbit*. Some coercion happens to affect the complement structure of a verb, but this is simply a special case of a more general phenomenon that has been analyzed by rules of systematic polysemy.

21.1.3 Aspect as a clause level phenomenon

Alsina (1996), working in the framework of LFG, argues for a phrasal analysis of resultative constructions based on the aspectual properties of sentences, since aspect is normally viewed as a property that is determined by the sentence syntax. Intransitive verbs such as *bark* refer to activities, a resultative construction with the same verb, however, stands for an accomplishment (an extended change of state). Alsina supports this with the following data:

- (6) a. (*) The dog barked in five minutes.
b. The dog barked the neighbors awake in five minutes.

The latter sentence means that the *barking* event was completed after five minutes. A reading referring to the time span of the event is not available for (6a). If (6a) is grammatical at all, then a claim is being made about the time frame in which the event begun.

If we now consider examples such as (7c), however, we see that Alsina's argumentation is not cogent since the resultative meaning is already present at the word-level in nominalizations. As the examples in (7) show, this contrast can be observed in nominal constructions and is therefore independent of the sentence syntax:

- (7) a. weil sie die Nordsee in fünf Jahren leer fischten
because they the North.Sea in five years empty fished
'because they fished the North Sea (until it was) empty in five years'
b. # weil sie in fünf Jahren fischten
because they in five years fished
c. das Leerfischen der Nordsee in fünf Jahren
the empty.fishing of.the North.Sea in five years
d. # das Fischen in fünf Jahren
the fishing in five years

In a lexical approach there is a verb stem selecting for two NPs and a resultative predicate. This stem has the appropriate meaning and can be inflected or undergo derivation

und successive inflection. In both cases we get words that contain the resultative semantics and hence are compatible with respective adverbials.

21.1.4 Simplicity and polysemy

Much of the intuitive appeal of the plugging approach stems from its apparent simplicity relative to the use of lexical rules. But the claim to greater simplicity for Construction Grammar is based on misunderstandings of both lexical rules and Construction Grammar (specifically of Goldberg's (1995; 2006) version). It draws the distinction in the wrong place and misses the real differences between these approaches. This argument from simplicity is often repeated and so it is important to understand why it is incorrect.

Tomasello (2003) presents the argument as follows. Discussing first the lexical rules approach, Tomasello (2003: 160) writes that

One implication of this view is that a verb must have listed in the lexicon a different meaning for virtually every different construction in which it participates [...]. For example, while the prototypical meaning of *cough* involves only one participant, the cougher, we may say such things as *He coughed her his cold*, in which there are three core participants. In the lexical rules approach, in order to produce this utterance the child's lexicon must have as an entry a ditransitive meaning for the verb *cough*. (Tomasello 2003: 160)

Tomasello (2003: 160) then contrasts a Construction Grammar approach, citing Fillmore et al. (1988), Goldberg (1995), and Croft (2001). He concludes as follows:

The main point is that if we grant that constructions may have meaning of their own, in relative independence of the lexical items involved, then we do not need to populate the lexicon with all kinds of implausible meanings for each of the verbs we use in everyday life. The construction grammar approach in which constructions have meanings is therefore both much simpler and much more plausible than the lexical rules approach. (Tomasello 2003: 161)

This reflects a misunderstanding of lexical rules, as they are normally understood. There is no implausible sense populating the lexicon. The lexical rule approach to *He coughed her his cold* states that when the word *coughed* appears with two objects, the whole complex has a certain meaning (see Müller 2006: 876). Furthermore we explicitly distinguish between listed elements (lexical entries) and derived ones. The general term subsuming both is *lexical item*.

The simplicity argument also relies on a misunderstanding of a theory Tomasello advocates, namely the theory due to Goldberg (1995, 2006). For his argument to go through, Tomasello must tacitly assume that verbs can combine freely with constructions, that is, that the grammar does not place extrinsic constraints on such combinations. If it is necessary to also stipulate which verbs can appear in which constructions, then the claim to greater simplicity collapses: each variant lexical item with its "implausible meaning"

under the lexical rule approach corresponds to a verb-plus-construction combination under the phrasal approach.

Passages such as the following may suggest that verbs and constructions are assumed to combine freely:⁸

Constructions are combined freely to form actual expressions as long as they can be construed as not being in conflict (invoking the notion of construal is intended to allow for processes of accommodation or coercion). [...] Allowing constructions to combine freely as long as there are no conflicts, allows for the infinitely creative potential of language. [...] That is, a speaker is free to creatively combine constructions as long as constructions exist in the language that can be combined suitably to categorize the target message, given that there is no conflict among the constructions. (Goldberg 2006: 22)

But in fact Goldberg does not assume free combination, but rather that a verb is “conventionally associated with a construction” (Goldberg 1995: 50): verbs specify their participant roles and which of those are obligatory direct arguments (*profiled*, in Goldberg’s terminology). In fact, Goldberg herself (2006: 211) argues against Borer’s putative assumption of free combination (2003) on the grounds that Borer is unable to account for the difference between *dine* (intransitive), *eat* (optionally transitive), and *devour* (obligatorily transitive).⁹ Despite Tomasello’s comment above, Construction Grammar is no simpler than the lexical rules.

The resultative construction is often used to illustrate the simplicity argument. For example, Goldberg (1995: Chapter 7) assumes that the same lexical item for the verb *sneeze* is used in (8a) and (8b). It is simply inserted into different constructions:

- (8) a. He sneezed.
b. He sneezed the napkin off the table.

The meaning of (8a) corresponds more or less to the verb meaning, since the verb is used in the Intransitive Construction. But the Caused-Motion Construction in (8b) contributes additional semantic information concerning the causation and movement: his sneezing caused the napkin to move off the table. *sneeze* is plugged into the Caused-Motion Construction, which licenses the subject of *sneeze* and additionally provides two slots: one for the theme (*napkin*) and one for the goal (*off the table*). The lexical approach is essentially parallel, except that the lexical rule can feed further lexical processes like passivization (*The napkin was sneezed off the table*), and conversion to nouns or adjectives (see Sections 21.2.2 and 21.6).

In a nuanced comparison of the two approaches, Goldberg (1995: 139–140) considers again the added recipient argument in *Mary kicked Joe the ball*, where *kick* is lexically

⁸ The context of these quotes makes clear that the verb and the argument structure construction are considered constructions. See Goldberg (2006: 21, ex. (2)).

⁹ Goldberg’s critique cites a 2001 presentation by Borer with the same title as Borer (2003). See Section 21.3.4 for more discussion of this issue. As far as we know, the *dine* / *eat* / *devour* minimal triplet originally came from Dowty (1989: 89–90).

a 2-place verb. She notes that on the constructional view, “the composite fused structure involving both verb and construction is stored in memory”. The verb itself retains its original meaning as a 2-place verb, so that “we avoid implausible verb senses such as ‘to cause to receive by kicking’”. The idea seems to be that the lexical approach, in contrast, must countenance such implausible verb senses since a lexical rule adds a third argument.

But the lexical and constructional approaches are actually indistinguishable on this point. The lexical rule does not produce a verb with the “implausible sense” in (9a). Instead it produces the sense in (9b):

- (9) a. cause-to-receive-by-kicking(*x*, *y*, *z*)
 b. cause(kick(*x*, *y*),receive(*z*,*y*))

The same sort of “composite fused structure” is assumed under either view. With respect to the semantic structure, the number and plausibility of senses, and the polyadicity of the semantic relations, the two theories are identical. They mainly differ in the way this representation fits into the larger theory of syntax. They also differ in another respect: on the lexical view, the derived three-argument valence structure is associated with the phonological string *kicked*. Next, we present evidence for this claim.

21.2 Evidence for lexical approaches

21.2.1 Valence and coordination

On the lexical account, the verb *paint* in (4b), for example, is lexically a 2-argument verb, while the unary branching node immediately dominating it is effectively a 3-argument verb. On the constructional view there is no such predicate seeking three arguments that dominates only the verb. Coordination provides evidence for the lexical account.

A generalization about coordination is that two constituents which have compatible syntactic properties can be coordinated and that the result of the coordination is an object that has the syntactic properties of each of the conjuncts. This is reflected by the Categorical Grammar analysis which assumes the category $(X \setminus X)/X$ for the conjunction: the conjunction takes an *X* to the right, an *X* to the left and the result is an *X*.

For example, in (10a) we have a case of the coordination of two lexical verbs. The coordination *know and like* behaves like the coordinated simplex verbs: it takes a subject and an object. Similarly, two sentences with a missing object are coordinated in (10b) and the result is a sentence with a missing object.

- (10) a. I know and like this record.
 b. Bagels, I like and Ellison hates.

The German examples in (11) show that the case requirement of the involved verbs has to be respected. In (11b,c) the coordinated verbs require accusative and dative respectively and since the case requirements are incompatible with unambiguously case marked nouns both of these examples are out.

- (11) a. Ich kenne und unterstütze diesen Mann.
I know and support this.ACC man
b. *Ich kenne und helfe diesen Mann.
I know and help this.ACC man
c. *Ich kenne und helfe diesem Mann.
I know and help this.DAT man

Interestingly, it is possible to coordinate basic ditransitive verbs with verbs that have additional arguments licensed by the lexical rule. (12) provides examples in English and German ((12b) is quoted from Müller (2013b: 420)):

- (12) a. She then offered and made me a wonderful espresso – nice.¹⁰
b. ich hab ihr jetzt diese Ladung Muffins mit den Herzchen drauf
I have her now this load Muffins with the little.heart there.on
gebacken und gegeben.¹¹
baked and given
'I have now baked and given her this load of muffins with the little heart on top.'

These sentences show that both verbs are 3-argument verbs at the V^0 level, since they involve V^0 coordination:

- (13) [V^0 offered and made] [_{NP} me] [_{NP} a wonderful espresso]

This is expected under the lexical rule analysis but not the non-lexical constructional one.¹²

Summarizing the coordination argument: coordinated verbs generally must have compatible syntactic properties like valence properties. This means that in (12b), for example, *gebacken* 'baked' and *gegeben* 'given' have the same valence properties. On the lexical approach the creation verb *gebacken*, together with a lexical rule, licenses a ditransitive verb. It can therefore be coordinated with *gegeben*. On the phrasal approach however, the verb *gebacken* has two argument roles and is not compatible with the verb *gegeben*, which has three argument roles. In the phrasal model, *gebacken* can only realize three arguments when it enters the ditransitive phrasal construction or argument

¹⁰ <http://www.thespinroom.com.au/?p=102> 07.07.2012

¹¹ <http://www.musiker-board.de/diverses-ot/35977-die-liebe-637-print.html>. 08.06.2012

¹² One might wonder whether these sentences could be instances of Right Node Raising (RNR) out of coordinated VPs (Bresnan 1974; Abbott 1976):

- (i) She [offered ____] and [made me ____] a wonderful espresso.

But this cannot be correct. Under such an analysis the first verb has been used without a benefactive or recipient object. But *me* is interpreted as the recipient of both the offering and making. Secondly, the second object can be an unstressed pronoun (*She offered and made me it*), which is not possible in RNR. Note that *offered and made* cannot be a pseudo-coordination meaning 'offered to make'. This is possible only with stem forms of certain verbs such as *try*.

structure construction. But in sentences like (12) it is not *gebacken* alone that enters the phrasal syntax, but rather the combination of *gebacken* and *gegeben*. On this view, the verbs are incompatible as far as the semantic roles are concerned.

To fix this under the phrasal approach, one could posit a mechanism such that the semantic roles that are required for the coordinate phrase *baked and given* are shared by each of its conjunct verbs and that they are therefore compatible. But this would amount to saying that there are several verb senses for *baked*, something that the anti-lexicalists claim to avoid, as discussed in the next section.

A reviewer of Theoretical Linguistics correctly observes that a version of the (phrasal) ASC approach could work in the exactly same way as our lexical analysis. Our ditransitive lexical rule would simply be rechristened as a “ditransitive ASC”. This construction would combine with *baked*, thus adding the third argument, prior to its coordination with *gave*. As long as the ASC approach is a non-distinct notational variant of the lexical rule approach then of course it works in exactly the same way. But the literature on the ASC approach represents it as a radical alternative to lexical rules, in which constructions are combined through inheritance hierarchies, instead of allowing lexical rules to alter the argument structure of a verb prior to its syntactic combination with the other words and phrases.

The reviewer also remarked that examples like (14) show that the benefactive argument has to be introduced on the phrasal level.

(14) I designed and built him a house.

Both *designed* and *built* are bivalent verbs and *him* is the benefactive that extends both *designed* and *built*. However, we assume that sentences like (14) can be analyzed as coordination of two verbal items that are licensed by the lexical rule that introduces the benefactive argument. That is, the benefactive is introduced before the coordination.

The coordination facts illustrate a more general point. The output of a lexical rule such as the one that would apply in the analysis of *gebacken* in (12b) is just a word (an X^0), so it has the same syntactic distribution as an underived word with the same category and valence feature. This important generalization follows from the lexical account while on the phrasal view, it is mysterious at best. The point can be shown with any of the lexical rules that the anti-lexicalists are so keen to eliminate in favor of phrasal constructions. For example, active and passive verbs can be coordinated, as long as they have the same valence properties, as in this Swedish example:

(15) Golfklubben begärde och beviljade-s marklov för banbygget
 golf.club.DEF requested and granted-PASS ground.permit for track.build.DEF
 efter en hel del förhandlingar och kompromisser med Länsstyrelsen
 after a whole part negotiations and compromises with county.board.DEF
 och Naturvårdsverket.¹³
 and nature.protection.agency.DEF

‘The golf club requested and was granted a ground permit for fairlane construc-

¹³ <http://www.lyckselegolf.se/index.asp?Sida=82>

tion after a lot of negotiations and compromises with the County Board and the Environmental Protection Agency.’

(English works the same way, as shown by the grammatical translation line.) The passive of the ditransitive verb *bevilja* ‘grant’ retains one object, so it is effectively transitive and can be coordinated with the active transitive *begära* ‘request’.

Moreover, the English passive verb form, being a participle, can feed a second lexical rule deriving adjectives from verbs. All categories of English participles can be converted to adjectives (Bresnan, 1982c, 2001: Chapter 3):

- (16) a. active present participles (cf. The leaf is falling): *the falling leaf*
 b. active past participles (cf. The leaf has fallen): *the fallen leaf*
 c. passive participles (cf. The toy is being broken (by the child).): *the broken toy*

That the derived forms are adjectives, not verbs, is shown by a host of properties, including negative *un-* prefixation: *unbroken* means ‘not broken’, just as *unkind* means ‘not kind’, while the *un-* appearing on verbs indicates, not negation, but action reversal, as in *untie* (Bresnan, 1982c: 21, 2001: Chapter 3). Predicate adjectives preserve the subject of predication of the verb and for prenominal adjectives the rule is simply that the role that would be assigned to the subject goes to the modified noun instead (*The toy remained (un-)broken.*; *the broken toy*). Being an A^0 , such a form can be coordinated with another A^0 , as in the following:

- (17) a. The suspect should be considered [armed and dangerous].
 b. any [old, rotting, or broken] toys

In (17b), three adjectives are coordinated, one underived (*old*), one derived from a present participle (*rotting*), and one from a passive participle (*broken*). Such coordination is completely mundane on a lexical theory. Each A^0 conjunct has a valence feature (in HPSG it would be the *SPR* feature for predicates or the *MOD* feature for the prenominal modifiers), which is shared with the mother node of the coordinate structure. But the point of the phrasal (or ASC) theory is to deny that words have such valence features.

The claim that lexical derivation of valence structure is distinct from phrasal combination is further supported with evidence from deverbal nominalization (Wechsler 2008a). To derive nouns from verbs, *-ing* suffixation productively applies to all inflectable verbs (*the shooting of the prisoner*), while morphological productivity is severely limited for various other suffixes such as *-(a)tion* (**the shootation of the prisoner*). So forms such as *destruction* and *distribution* must be retrieved from memory while *-ing* nouns such as *looting* or *growing* could be (and in the case of rare verbs or neologisms, must be) derived from the verb or the root through the application of a rule (Zucchi 1993). This difference explains why *-ing* nominals always retain the argument structure of the cognate verb, while other forms show some variation. A famous example is the lack of the agent argument for the noun *growth* versus its retention by the noun *growing*: **John’s growth of tomatoes* versus *John’s growing of tomatoes* (Chomsky 1970).¹⁴

¹⁴ See Section 21.3.3 for further discussion.

But what sort of rule derives the *-ing* nouns, a lexical rule or a phrasal one? In Marantz's (1997) phrasal analysis, a phrasal construction (notated as νP) is responsible for assigning the agent role of *-ing* nouns such as *growing*. For him, none of the words directly selects an agent via its argument structure. The *-ing* forms are permitted to appear in the νP construction, which licenses the possessive agent. Non-*-ing* nouns such as *destruction* and *growth* do not appear in νP . Whether they allow expression of the agent depends on semantic and pragmatic properties of the word: *destruction* involves external causation so it does allow an agent, while *growth* involves internal causation so it does not allow an agent.

However, a problem for Marantz is that these two types of nouns can coordinate and share dependents (example (18a) is from Wechsler (2008a: Section 7)):

- (18) a. With nothing left after the soldier's [destruction and looting] of their home, they reboarded their coach and set out for the port of Calais.¹⁵
 b. The [cultivation, growing or distribution] of medical marijuana within the County shall at all times occur within a secure, locked, and fully enclosed structure, including a ceiling, roof or top, and shall meet the following requirements.¹⁶

On the phrasal analysis, the nouns *looting* and *growing* occur in one type of syntactic environment (namely νP), while forms *destruction*, *cultivation*, and *distribution* occur in a different syntactic environment. This places contradictory demands on the structure of coordinations like those in (18). As far as we know, neither this problem nor the others raised by Wechsler (2008a) have even been addressed by advocates of the phrasal theory of argument structure.

Consider one last example. In an influential phrasal analysis, Hale and Keyser (1993a) derived denominal verbs like *to saddle* through noun incorporation out of a structure akin to [PUT a saddle ON x]. Again, verbs with this putative derivation routinely coordinate and share dependents with verbs of other types:

- (19) Realizing the dire results of such a capture and that he was the only one to prevent it, he quickly [saddled and mounted] his trusted horse and with a grim determination began a journey that would become legendary.¹⁷

As in all of these X^0 coordination cases, under the phrasal analysis the two verbs place contradictory demands on a single phrase structure.

A lexical valence structure is an abstraction or generalization over various occurrences of the verb in syntactic contexts. To be sure, one key use of that valence structure is simply to indicate what sort of phrases the verb must (or can) combine with, and the result of semantic composition; if that were the whole story then the phrasal theory would be viable. But it is not. As it turns out, this lexical valence structure, once abstracted,

¹⁵ <http://www.amazon.com/review/R3IG4M3Q6YYNFT>, 21.07.2012

¹⁶ <http://www.scribd.com/doc/64013640/Tulare-County-medical-cannabis-cultivation-ordinance>, 05.03.2016

¹⁷ http://www.jouethouse.org/index.php?option=com_content&view=article&id=56&Itemid=63, 21.07.2012

can alternatively be used in other ways: among other possibilities, the verb (crucially including its valence structure) can be coordinated with other verbs that have similar valence structures; or it can serve as the input to lexical rules specifying a new word bearing a systematic relation to the input word. The coordination and lexical derivation facts follow from the lexical view, while the phrasal theory at best leaves these facts as mysterious and at worst leads to irreconcilable contradictions for the phrase structure.

21.2.2 Valence and derivational morphology

Goldberg & Jackendoff (2004), Alsina (1996), and Asudeh, Dalrymple & Toivonen (2008, 2013) suggest analyzing resultative constructions and/or caused-motion constructions as phrasal constructions.¹⁸ As was argued in Müller (2006) this is incompatible with the assumption of lexical integrity. Lexical integrity means that word formation happens before syntax and that the morphological structure is inaccessible to syntactic processes (Bresnan & Mchombo 1995).¹⁹ Let us consider a concrete example, such as (20):

- (20)
- a. Er tanzt die Schuhe blutig / in Stücke.
he dances the shoes bloody into pieces
 - b. die in Stücke / blutig getanzten Schuhe
the into pieces bloody danced shoes
 - c. * die getanzten Schuhe
the danced shoes

The shoes are not a semantic argument of *tanzt*. Nevertheless the referent of the NP that is realized as accusative NP in (20a) is the element the adjectival participle in (20b) predicates over. Adjectival participles like the one in (20b) are derived from a passive participle of a verb that governs an accusative object. If the accusative object is licensed phrasally by configurations like the one in (20a), then it is not possible to explain why the participle *getanzten* can be formed despite the absence of an accusative object in the valence specification of the verb. See Müller (2006: Section 5) for further examples of the interaction of resultatives and morphology. The conclusion drawn by Dowty (1978: 412) and Bresnan (1982c: 21) in the late 70s and early 80s is that phenomena which feed morphology should be treated lexically. The natural analysis in frameworks like HPSG, CG, CxG, and LFG is therefore one that assumes a lexical rule for the licensing of resultative

¹⁸ Asudeh & Toivonen (2014: Section 2.3) argue that their account is not constructional. If a construction is a form-meaning pair, their account is constructional, since a certain c-structure is paired with a semantic contribution. Asudeh & Toivonen (2014: Section 2.2) compare their approach with approaches in Constructional HPSG (Sag 1997) and Sign-Based Construction Grammar (see Section 10.6.2), which they term constructional. The only difference between these approaches and the approach by Asudeh, Dalrymple & Toivonen is that the constructions in the HPSG-based theories are modeled using types and hence have a name.

¹⁹ Asudeh et al. (2013: 14) claim that the Swedish Directed Motion Construction does not interact with derivational morphology. However, the parallel German construction does interact with derivational morphology. The absence of this interaction in Swedish can be explained by other factors of Swedish grammar and given this I believe it to be more appropriate to assume an analysis that captures both the German and the Swedish data in the same way.

constructions. See Verspoor (1997), Wechsler (1997), Wechsler & Noh (2001), Wunderlich (1992: 45; 1997: 120–126), Kaufmann & Wunderlich (1998), Müller (2002a: Chapter 5), Kay (2005), and Simpson (1983) for lexical proposals in some of these frameworks.

This argument is similar to the one that was discussed in connection with the GPSG representation of valence in Section 5.5: morphological processes have to be able to see the valence of the element they apply to. This is not the case if arguments are introduced by phrasal configurations after the level of morphology.

Asudeh, Dalrymple & Toivonen's papers are about the concept of lexical integrity and about constructions. Asudeh & Toivonen (2014) replied to our target article and pointed out (again) that their template approach makes it possible to specify the functional structure of words and phrases alike. In the original paper they discussed the Swedish word *vägen*, which is the definite form of *väg* 'way'. They showed that the f-structure is parallel to the f-structure for the English phrase *the way*. In our reply (2014b), we gave in too early, I believe. Since the point is not about being able to provide the f-structure of words, the point is about morphology, that is – in LFG terms – about deriving the f-structure by a morphological analysis. More generally speaking, one wants to derive all properties of the involved words, that is, their valence, their meaning, and the linking of this meaning to their dependents. What we used in our argument based on the sentences in (20) was parallel to what Bresnan (1982c: 21; 2001: 31) used in her classical argument for a lexical treatment of the passive. So either Bresnan's argument (and ours) is invalid or both arguments are valid and there is a problem for Asudeh, Dalrymple & Toivonen's approach and for phrasal approaches in general. I want to give another example that was already discussed in Müller (2006: 869) but was omitted in Müller & Wechsler (2014a) due to space limitations. I will first point out why this example is problematic for phrasal approaches and then explain why it is not sufficient to be able to assign certain f-structures to words: in (21a), we are dealing with a resultative construction. According to the plugging approach, the resultative meaning is contributed by a phrasal construction into which the verb *fischt* is inserted. There is no lexical item that requires a resultative predicate as its argument. If no such lexical item exists, then it is unclear how the relation between (21a) and (21b) can be established:

- (21) a. [dass] jemand die Nordsee leer fischt
 that somebody the North.Sea empty fishes
 'that somebody fishes the North Sea empty'
 b. wegen der *Leerfischung* der Nordsee²⁰
 because of.the empty.fishing of.the North.Sea
 'because of the fishing that resulted in the North Sea being empty'

As Figure 21.1 on the next page shows, both the arguments selected by the heads and the structures are completely different. In (21b), the element that is the subject of the related construction in (21a) is not realized. As is normally the case in nominalizations, it is possible to realize it in a PP with the preposition *durch* 'by':

²⁰ taz, 20.06.1996, p. 6.

- (22) wegen der Leerfischung der Nordsee durch die Anrainerstaaten
 because of the empty.fishing of the North.Sea by the neighboring.states
 ‘because of the fishing by the neighboring states that resulted in the North Sea
 being empty’

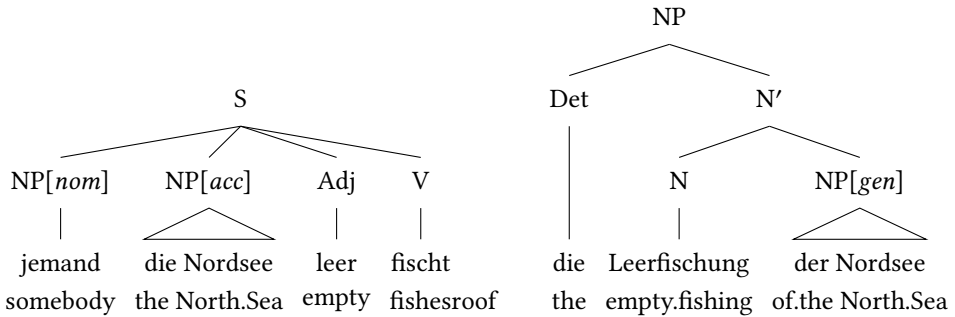


图 21.1: Resultative construction and nominalization

If one assumes that the resultative meaning comes from a particular configuration in which a verb is realized, there would be no explanation for (21b) since no verb is involved in the analysis of this example. One could of course assume that a verb stem is inserted into a construction both in (21a) and (21b). The inflectional morpheme *-t* and the derivational morpheme *-ung* as well as an empty nominal inflectional morpheme would then be independent syntactic components of the analysis. However, since Goldberg (2003b: 119) and Asudeh et al. (2013) assume lexical integrity, only entire words can be inserted into syntactic constructions and hence the analysis of the nominalization of resultative constructions sketched here is not an option for them.

One might be tempted to try and account for the similarities between the phrases in (21) using inheritance. One would specify a general resultative construction standing in an inheritance relation to the resultative construction with a verbal head and the nominalization construction. I have discussed this proposal in more detail in Müller (2006: Section 5.3). It does not work as one needs embedding for derivational morphology and this cannot be modeled in inheritance hierarchies (Krieger & Nerbonne (1993), see also Müller (2006) for a detailed discussion).

It would also be possible to assume that both constructions in (23), for which structures such as those in Figure 21.1 would have to be assumed, are connected via metarules.^{21, 22}

²¹ Goldberg (p. c. 2007, 2009) suggests connecting certain constructions using GPSG-like metarules. Deppermann (2006: 51), who has a more Croftian view of CxG, rules this out. He argues for active/passive alternations that the passive construction has other information structural properties. Note also that GPSG metarules relate phrase structure rules, that is, local trees. The structure in Figure 21.2, however, is highly complex.

²² The structure in (23b) violates a strict interpretation of lexical integrity as is commonly assumed in LFG. Booi (2005, 2009), working in Construction Grammar, subscribes to a somewhat weaker version, however.

- (23) a. [Sbj Obj Obl V]
 b. [Det [[Adj V -ung]] NP[*gen*]]

The construction in (23b) corresponds to Figure 21.2.²³ The genitive NP is an argument

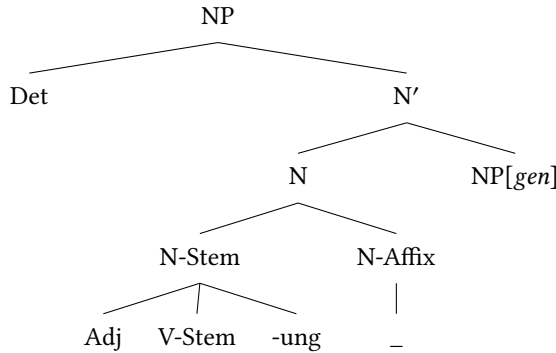


图 21.2: Resultative construction and nominalization

of the adjective. It has to be linked semantically to the subject slot of the adjective. Alternatively, one could assume that the construction only has the form [Adj V -ung], that is, that it does not include the genitive NP. But then one could also assume that the verbal variant of the resultative construction has the form [OBL V] and that Sbj and Obj are only represented in the valence lists. This would almost be a lexical analysis, however.

Turning to lexical integrity again, I want to point out that all that Asudeh & Toivonen can do is assign some f-structure to the N in Figure 21.2. What is needed, however, is a principled account of how this f-structure comes about and how it is related to the resultative construction on the sentence level.

Before I turn to approaches with radical underspecification of argument structure in the next section, I want to comment on a more recent paper by Asudeh, Giorgolo & Toivonen (2014). The authors discuss the phrasal introduction of cognate objects and benefactives. (24a) is an example of the latter construction.

- (24) a. The performer sang the children a song.
 b. The children were sung a song.

According to the authors, the noun phrase *the children* is not an argument of *sing* but contributed by the c-structure rule that optionally licenses a benefactive.

- (25) V' → V DP DP
 ↑ = ↓ (↑ OBJ) = ↓ (↑ OBJ_θ) = ↓
 (@BENEFACTIVE)

²³ I do not assume zero affixes for inflection. The respective affix in Figure 21.2 is there to show that there is structure. Alternatively one could assume a unary branching rule/construction as is common in HPSG/Construction Morphology.

Whenever this rule is evoked, the template *BENEFACTIVE* can add a benefactive role and the respective semantics of this is compatible with the verb that is inserted into the structure. The authors show how the mappings for the passive example in (24b) work, but they do not provide the c-structure that licenses such examples. In order to analyze these examples one would need a c-structure rule for passive VPs and this rule has to license a benefactive as well. So it would be:

$$(26) \quad V' \rightarrow \begin{array}{cc} V[\text{pass}] & \text{DP} \\ \uparrow = \downarrow & (\uparrow \text{OBJ}_{\theta}) = \downarrow \\ (@\text{BENEFACTIVE}) \end{array}$$

Note that a benefactive cannot be added to any verb: adding a benefactive to an intransitive verb as in (27a) is out and the passive that would correspond to (27a) is ungrammatical as well, as (27b) shows:

- (27) a. * He laughed the children.
 b. * The children were laughed.

So one could not just claim that all c-structure rules optionally introduce a benefactive argument. Therefore there is something special about the two rules in (25) and (26). The problem is that there is no relation between these rules. They are independent statements saying that there can be a benefactive in the active and that there can be one in the passive. This is what Chomsky (1957: 43) criticized in 1957 and this was the reason for the introduction of transformations (see Section 3.1.1 of this book). Bresnan-style LFG captured the generalizations by lexical rules and later by Lexical Mapping Theory. But if elements are added outside the lexical representations, the representations where these elements are added have to be related too. One could say that our knowledge about formal tools has changed since 1957. We now can use inheritance hierarchies to capture generalizations. So one can assume a type (or a template) that is the supertype of all those c-structure rules that introduce a benefactive. But since not all rules allow for the introduction of a benefactive element, this basically amounts to saying: c-structure rule A, B, and C allow for the introduction of a benefactive. In comparison, lexical rule-based approaches have one statement introducing the benefactive. The lexical rule states what verbs are appropriate for adding a benefactive and syntactic rules are not affected.

In Müller & Wechsler (2014a) we argued that the approach to Swedish Caused-Motion Constructions in Asudeh et al. (2008, 2013) would not carry over to German since the German construction interacts with derivational morphology. Asudeh & Toivonen (2014) argued that Swedish is different from German and hence there would not be a problem. However, the situation is different with the benefactive constructions. Although English and German do differ in many respects, both languages have similar dative constructions:

- (28) a. He baked her a cake.
 b. Er buk ihr einen Kuchen.
 he baked her.DAT a.ACC cake

Now, the analysis of the free constituent order was explained by assuming binary branching structures in which a VP node is combined with one of its arguments or adjuncts (see Section 7.4). The c-structure rule is repeated in (29):

- (29) $VP \rightarrow \begin{array}{cc} NP & VP \\ (\uparrow \text{SUBJ} | \text{OBJ} | \text{OBJ}_\theta) = \downarrow & \uparrow = \downarrow \end{array}$

The dependent elements contribute to the f-structure of the verb and coherence/completeness ensure that all arguments of the verb are present. One could add the introduction of the benefactive argument to the VP node of the right-hand side of the rule. However, since the verb-final variant of (28b) would have the structure in (30), one would get spurious ambiguities, since the benefactive could be introduced at every node:

- (30) weil $[_{VP} \text{er } [_{VP} \text{ihr } [_{VP} \text{einen Kuchen } [_{VP} [_V \text{buk}]]]]]$
 because he her a cake baked

So the only option seems to be to introduce the benefactive at the rule that got the recursion going, namely the rule that projected the lexical verb to the VP level. The rule (39) from page 212 is repeated as (31) for convenience.

- (31) $VP \rightarrow (V)$
 $\uparrow = \downarrow$

Note also that benefactive datives appear in adjectival environments as in (32):

- (32) a. der seiner Frau einen Kuchen backende Mann
 the his.DAT wife a.ACC cake baking man
 ‘the man who is baking a cake for her’
 b. der einen Kuchen seiner Frau backende Mann
 the a.ACC cake his.DAT wife baking man
 ‘the man who is baking a cake for her’

In order to account for these datives one would have to assume that the adjective-to-AP rule that would be parallel to (31) introduces the dative. The semantics of the benefactive template would have to somehow make sure that the benefactive argument is not added to intransitive verbs like *lachen* ‘to laugh’ or participles like *lachende* ‘laughing’. While this may be possible, I find the overall approach unattractive. First it does not have anything to do with the original constructional proposal but just states that the benefactive may be introduced at several places in syntax, second the unary branching syntactic rule is applying to a lexical item and hence is very similar to a lexical rule and third the analysis does not capture cross-linguistic commonalities of the construction. In a lexical rule-based approach as the one that was suggested by Briscoe & Copestake (1999: Section 5), a benefactive argument is added to certain verbs and the lexical rule is parallel in all languages that have this phenomenon. The respective languages differ simply in the way the arguments are realized with respect to their heads. In languages that have adjectival participles, these are derived from the respective verbal stems. The morphological rule is the same independent of benefactive arguments and the syntactic rules for adjectival phrases do not have to mention benefactive arguments.

21.3 Radical underspecification: the end of argument structure?

21.3.1 Neo-Davidsonianism

In the last section we examined proposals that assume that verbs come with certain argument roles and are inserted into prespecified structures that may contribute additional arguments. While we showed that this is not without problems, there are even more radical proposals that the construction adds all agent arguments, or even all arguments. The notion that the agent argument should be severed from its verbs is put forth by Marantz (1984, 1997), Kratzer (1996), Embick (2004) and others. Others suggest that no arguments are selected by the verb. Borer (2003) calls such proposals *exoskeletal* since the structure of the clause is not determined by the predicate, that is, the verb does not project an inner “skeleton” of the clause. Counter to such proposals are *endoskeletal* approaches, in which the structure of the clause is determined by the predicate, that is, lexical proposals. The radical exoskeletal approaches are mainly proposed in Mainstream Generative Grammar (Borer 1994, 2003, 2005; Schein 1993; Hale & Keyser 1997; Lohndal 2012) but can also be found in HPSG (Haugereid 2009). We will not discuss these proposals in detail here, but we review the main issues insofar as they relate to the question of lexical argument structure.²⁴ We conclude that the available empirical evidence favors the lexical argument structure approach over such alternatives.

Exoskeletal approaches usually assume some version of Neo-Davidsonianism. Davidson (1967) argued for an event variable in the logical form of action sentences (33a). Dowty (1989) coined the term *neo-Davidsonian* for the variant in (33b), in which the verb translates to a property of events, and the subject and complement dependents are translated as arguments of secondary predicates such as *agent* and *theme*.²⁵ Kratzer (1996) further noted the possibility of mixed accounts such as (33c), in which the agent (subject) argument is severed from the *kill*' relation, but the theme (object) remains an argument of the *kill*' relation.²⁶

- | | | |
|------|---|-------------------|
| (33) | a. <i>kill</i> : $\lambda y \lambda x \exists e[kill(e, x, y)]$ | (Davidsonian) |
| | b. <i>kill</i> : $\lambda y \lambda x \exists e[kill(e) \wedge agent(e, x) \wedge theme(e, y)]$ | (neo-Davidsonian) |
| | c. <i>kill</i> : $\lambda y \lambda x \exists e[kill(e, y) \wedge agent(e, x)]$ | (mixed) |

Kratzer (1996) observed that a distinction between Davidsonian, neo-Davidsonian and mixed can be made either “in the syntax” or “in the conceptual structure” (Kratzer 1996: 110–111). For example, on a lexical approach of the sort we advocate here, any of the three alternatives in (33) could be posited as the semantic content of the verb *kill*. A lexical entry for *kill* in the mixed model is given in (34).

²⁴ See Müller (2010a: Section 11.11.3) for a detailed discussion of Haugereid's approach.

²⁵ Dowty (1989) called the system in (33a) an *ordered argument system*.

²⁶ The event variable is shown as existentially bound, as in Davidson's original account. As discussed below, in Kratzer's version it must be bound by a lambda operator instead.

But Kratzer (1996), among others, has gone further and argued for an account that is neo-Davidsonian (or rather, mixed) “in the syntax”. Kratzer’s claim is that the verb specifies only the internal argument(s), as in (35a) or (35b), while the agent (external argument) role is assigned by the phrasal structure. On the “neo-Davidsonian in the syntax” view, the lexical representation of the verb has no arguments at all, except the event variable, as shown in (35c).

- (35) a. *kill*: $\lambda y\lambda e[\textit{kill}(e, y)]$ (agent is severed)
 b. *kill*: $\lambda y\lambda e[\textit{kill}(e) \wedge \textit{theme}(e, y)]$ (agent is severed)
 c. *kill*: $\lambda e[\textit{kill}(e))]$ (all arguments severed)

In the following subsections we address arguments that have been put forth in favor of the little *v* hypothesis, from idiom asymmetries (Section 21.3.2) and deverbal nominals (Section 21.3.3). We argue that the evidence actually favors the lexical view. Then we turn to problems for exoskeletal approaches, from idiosyncratic syntactic selection (Section 21.3.4) and expletives (Section 21.3.5). We conclude with a look at the treatment of idiosyncratic syntactic selection under Borer's exoskeletal theory (Section 21.3.7), and a summary (Section 21.3.8).

Marantz (1984) and Kratzer (1996) argued for severing the agent from the argument structure as in (35a), on the basis of putative idiom asymmetries. Marantz (1984) observed that while English has many idioms and specialized meanings for verbs in which the internal argument is the fixed part of the idiom and the external argument is free, the reverse situation is considerably rarer. To put it differently, the nature of the role played

by the subject argument often depends on the filler of the object position, but not vice versa. To take Kratzer's examples (Kratzer 1996: 114):

- (36)
- a. kill a cockroach
 - b. kill a conversation
 - c. kill an evening watching TV
 - d. kill a bottle (i.e. empty it)
 - e. kill an audience (i.e., wow them)

On the other hand, one does not often find special meanings of a verb associated with the choice of subject, leaving the object position open (examples from Marantz (1984: 26)):

- (37)
- a. Harry killed NP.
 - b. Everyone is always killing NP.
 - c. The drunk refused to kill NP.
 - d. Silence certainly can kill NP.

Kratzer observes that a mixed representation of *kill* as in (38a) allows us to specify varying meanings that depend upon its sole NP argument.

- (38)
- a. *kill*: $\lambda y\lambda e[\textit{kill}(e, y)]$
 - b. If *a* is a time interval, then $\textit{kill}(e, a) = \textit{truth}$ if *e* is an event of wasting *a*
 If *a* is animate, then $\textit{kill}(e, a) = \textit{truth}$ if *e* is an event in which *a* dies
 ... etc.

On the polyadic (Davidsonian) theory, the meaning could similarly be made to depend upon the filler of the agent role. On the polyadic view, "there is no technical obstacle" (Kratzer 1996: 116) to conditions like those in (38b), except reversed, so that it is the filler of the agent role instead of the theme role that affects the meaning. But, she writes, this could not be done if the agent is not an argument of the verb. According to Kratzer, the agent-severed representation (such as (38a)) disallows similar constraints on the meaning that depend upon the agent, thereby capturing the idiom asymmetry.

But as noted by Wechsler (2005), "there is no technical obstacle" to specifying agent-dependent meanings even if the Agent has been severed from the verb as Kratzer proposes. It is true that there is no variable for the agent in (38a). But there is an event variable *e*, and the language user must be able to identify the agent of *e* in order to interpret the sentence. So one could replace the variable *a* with "the agent of *e*" in the expressions in (38b), and thereby create verbs that violate the idiom asymmetry.

While this may seem to be a narrow technical or even pedantic point, it is nonetheless crucial. Suppose we try to repair Kratzer's argument with an additional assumption: that modulations in the meaning of a polysemous verb can only depend upon arguments of the *relation* denoted by that verb, and not on other participants in the event. Under that additional assumption, it makes no difference whether the agent is severed from

the lexical entry or not. For example, consider the following (mixed) neo-Davidsonian representation of the semantic content in the lexical entry of *kill*:

- (39) *kill*: $\lambda y \lambda x \lambda e [kill(e, y) \wedge agent(e, x)]$

Assuming that sense modulations can only be affected by arguments of the *kill*(*e*, *y*) relation, we derive the idiom asymmetry, even if (39) is the lexical entry for *kill*. So suppose that we try to fix Kratzer's argument with a different assumption: that modulations in the meaning of a polysemous verb can only depend upon an argument of the lexically denoted function. Kratzer's "neo-Davidsonian in the syntax" lexical entry in (35a) lacks the agent argument, while the lexical entry in (39) clearly has one. But Kratzer's entry still fails to predict the asymmetry because, as noted above, it has the *e* argument and so the sense modulation can be conditioned on the "agent of *e*". As noted above, that event argument cannot be eliminated (for example through existential quantification) because it is needed in order to undergo event identification with the event argument of the silent light verb that introduces the agent (Kratzer 1996: 22).

Moreover, recasting Kratzer's account in lexicalist terms allows for verbs to vary. This is an important advantage, because the putative asymmetry is only a tendency. The following are examples in which the subject is a fixed part of the idiom and there are open slots for non-subjects:²⁷

- (40) a. A little bird told X that S.
 'X heard the rumor that S.'
 b. The cat's got X's tongue.
 'X cannot speak.'
 c. What's eating X?
 'Why is X so galled?'

Further data and discussion of subject idioms in English and German can be found in Müller (2007b: Section 3.2.1).

The tendency towards a subject-object asymmetry plausibly has an independent explanation. Nunberg, Sag & Wasow (1994) argue that the subject-object asymmetry is a side-effect of an animacy asymmetry. The open positions of idioms tend to be animate while the fixed positions tend to be inanimate. Nunberg et al. (1994) derive these animacy generalizations from the figurative and proverbial nature of the metaphorical transfers that give rise to idioms. If there is an independent explanation for this tendency, then a lexicalist grammar successfully encodes those patterns, perhaps with a mixed neo-Davidsonian lexical decomposition, as explained above (see Wechsler (2005) for such a lexical account of the verbs *buy* and *sell*). But the little *v* hypothesis rigidly predicts this asymmetry for all agentive verbs, and that prediction is not borne out.

²⁷ (40a) is from Nunberg, Sag & Wasow (1994: 526), (40b) from Bresnan (1982a: 349–350), and (40c) from Bresnan (1982a: 349–350).

21.3.3 Deverbal nominals

An influential argument against lexical argument structure involves English deverbal nominals and the causative alternation. It originates from a mention in Chomsky (1970), and is developed in detail by Marantz (1997); see also Pesetsky (1996) and Harley & Noyer (2000). The argument is often repeated, but it turns out that the empirical basis of the argument is incorrect, and the actual facts point in the opposite direction, in favor of lexical argument structure (Wechsler 2008b,a).

Certain English causative alternation verbs allow optional omission of the agent argument (41), while the cognate nominal disallows expression of the agent (42):

- (41) a. that John grows tomatoes
b. that tomatoes grow
- (42) a. *John's growth of tomatoes
b. the tomatoes' growth, the growth of the tomatoes

In contrast, nominals derived from obligatorily transitive verbs such as *destroy* allow expression of the agent, as shown in (44a):

- (43) a. that the army destroyed the city
b. *that the city destroyed
- (44) a. the army's destruction of the city
b. the city's destruction

Following a suggestion by Chomsky (1970), Marantz (1997) argued on the basis of these data that the agent role is lacking from lexical entries. In verbal projections like (41) and (43) the agent role is assigned in the syntax by little *v*. Nominal projections like (42) and (44) lack little *v*. Instead, pragmatics takes over to determine which agents can be expressed by the possessive phrase: the possessive can express “the sort of agent implied by an event with an external rather than an internal cause” because only the former can “easily be reconstructed” (quoted from Marantz (1997: 218)). The destruction of a city has a cause external to the city, while the growth of tomatoes is internally caused by the tomatoes themselves (Smith 1970). Marantz points out that this explanation is unavailable if the noun is derived from a verb with an argument structure specifying its agent, since the deverbal nominal would inherit the agent of a causative alternation verb.

The empirical basis for this argument is the putative mismatch between the allowability of agent arguments, across some verb-noun cognate pairs: e.g., *grow* allows the agent but *growth* does not. But it turns out that the *grow/growth* pattern is rare. Most deverbal nominals precisely parallel the cognate verb: if the verb has an agent, so does the noun. Moreover, there is a ready explanation for the exceptional cases that exhibit the *grow/growth* pattern (Wechsler 2008a). First consider non-alternating theme-only intransitives (unaccusatives), as in (45) and non-alternating transitives as in (46). The pattern is clear: if the verb is agentless, then so is the noun:

- (45) *arriv(al), disappear(ance), fall* etc.:
- a. A letter arrived.
 - b. the arrival of the letter
 - c. *The mailman arrived a letter.
 - d. *the mailman's arrival of the letter
- (46) *destroy/destruction, construct(ion), creat(ion), assign(ment)* etc.:
- a. The army is destroying the city.
 - b. the army's destruction of the city

This favors the view that the noun inherits the lexical argument structure of the verb. For the anti-lexicalist, the badness of (45c) and (45d), respectively, would have to receive independent explanations. For example, on Harley and Noyer's 2000 proposal, (45c) is disallowed because a feature of the root ARRIVE prevents it from appearing in the context of *v*, but (45d) is instead ruled out because the cause of an event of arrival cannot be easily reconstructed from world knowledge. This exact duplication in two separate components of the linguistic system would have to be replicated across all non-alternating intransitive and transitive verbs, a situation that is highly implausible.

Turning to causative alternation verbs, Marantz's argument is based on the implicit generalization that noun cognates of causative alternation verbs (typically) lack the agent argument. But apart from the one example of *grow/growth*, there do not seem to be any clear cases of this pattern. Besides *grow(th)*, Chomsky (1970: examples (7c) and (8c)) cited two experienter predicates, *amuse* and *interest*: *John amused (interested) the children with his stories* versus **John's amusement (interest) of the children with his stories*. But this was later shown by Rappaport (1983) and Dowty (1989) to have an independent aspectual explanation. Deverbal experienter nouns like *amusement* and *interest* typically denote a mental state, where the corresponding verb denotes an event in which such a mental state comes about or is caused. These result nominals lack not only the agent but all the eventive arguments of the verb, because they do not refer to events. Exactly to the extent that such nouns can be construed as representing events, expression of the agent becomes acceptable.

In a response to Chomsky (1970), Carlota Smith (1972) surveyed Webster's dictionary and found no support for Chomsky's claim that deverbal nominals do not inherit agent arguments from causative alternation verbs. She listed many counterexamples, including "*explode, divide, accelerate, expand, repeat, neutralize, conclude, unify*, and so on at length." (Smith 1972: 137). Harley and Noyer (2000) also noted many so-called "exceptions": *explode, accumulate, separate, unify, disperse, transform, dissolve/dissolution, detach(ment), disengage(-ment)*, and so on. The simple fact is that these are not exceptions because there is no generalization to which they can be exceptions. These long lists of verbs represent the norm, especially for suffix-derived nominals (in *-tion*, *-ment*, etc.). Many zero-derived nominals from alternating verbs also allow the agent, such as *change, release*, and *use*: *my constant change of mentors from 1992–1997; the frequent release of the*

prisoners by the governor; the frequent use of sharp tools by underage children (examples from Borer (2003: fn. 13)).²⁸

Like the experiencer nouns mentioned above, many zero-derived nominals lack event readings. Some reject all the arguments of the corresponding eventive verb, not just the agent: **the freeze of the water*, **the break of the window*, and so on. According to Stephen Wechsler, *his drop of the ball* is slightly odd, but *the drop of the ball* has exactly the same degree of oddness. The locution *a drop in temperature* matches the verbal one *The temperature dropped*, and both verbal and nominal forms disallow the agent: **The storm dropped the temperature*. **the storm's drop of the temperature*. In short, the facts seem to point in exactly the opposite direction from what has been assumed in this oft-repeated argument against lexical valence. Apart from the one isolated case of *grow/growth*, event-denoting deverbal nominals match their cognate verbs in their argument patterns.

Turning to *grow/growth* itself, we find a simple explanation for its unusual behavior (Wechsler 2008a). When the noun *growth* entered the English language, causative (transitive) *grow* did not exist. The OED provides these dates of the earliest attestations of *grow* and *growth*:

- | | | | |
|------|-------------------------------|------|--|
| (47) | a. intransitive <i>grow</i> : | c725 | 'be verdant' ... 'increase' (intransitive) |
| | b. the noun <i>growth</i> : | 1587 | 'increase' (intransitive) |
| | c. transitive <i>grow</i> : | 1774 | 'cultivate (crops)' |

Thus *growth* entered the language at a time when transitive *grow* did not exist. The argument structure and meaning were inherited by the noun from its source verb, and then preserved into present-day English. This makes perfect sense if, as we claim, words have predicate argument structures. Nominalization by *-th* suffixation is not productive in English, so *growth* is listed in the lexicon. To explain why *growth* lacks the agent we need only assume that a lexical entry's predicate argument structure dictates whether it takes an agent argument or not. So even this one word provides evidence for lexical argument structure.

21.3.4 Idiosyncratic syntactic selections

The notion of lexical valence structure immediately explains why the argument realization patterns are strongly correlated with the particular lexical heads selecting those arguments. It is not sufficient to have general lexical items without valence information and let the syntax and world knowledge decide about argument realizations, because not all realizational patterns are determined by the meaning. The form of the preposition of a prepositional object is sometimes loosely semantically motivated but in other cases arbitrary. For example, the valence structure of the English verb *depend* captures the fact that it selects an *on*-PP to express one of its semantic arguments:

²⁸ Pesetsky (1996: 79, ex. (231)) assigns a star to *the thief's return of the money*, but it is acceptable to many speakers. The *Oxford English Dictionary* lists a transitive sense for the noun *return* (definition 11a), and corpus examples like *her return of the spoils* are not hard to find.

(48) a. John depends on Mary. (*counts, relies, etc.*)

b. John trusts (*on) Mary.

c.
$$\left[\begin{array}{ll} \text{PHON} & \langle \textit{depend} \rangle \\ \text{ARG-ST} & \langle \text{NP}_x, \text{PP}[\textit{on}]_y \rangle \\ \text{CONTENT} & \textit{depend}(x,y) \end{array} \right]$$

Such idiosyncratic lexical selection is utterly pervasive in human language. The verb or other predicator often determines the choice between direct and oblique morphology, and for obliques, it determines the choice of adposition or oblique case. In some languages such as Icelandic even the subject case can be selected by the verb (Zaenen, Maling & Thráinsson 1985).

Selection is language-specific. English *wait* selects *for* (German *für*) while German *warten* selects *auf* ‘on’ with an accusative object:

(49) a. I am waiting for my man.

b. Ich warte auf meinen Mann.

I wait on my man.ACC

It is often impossible to find semantic motivation for case. In German there is a tendency to replace genitive (50a) with dative (50b) with no apparent semantic motivation:

(50) a. dass der Opfer gedacht werde

that the victims.GEN remembered is

‘that the victims would be remembered’

b. daß auch hier den Opfern des Faschismus gedacht werde [...] ²⁹

that also here the victims.DAT of.the fascism remembered is

‘that the victims of fascism would be remembered here too’

The synonyms *treffen* and *begegnen* ‘to meet’ govern different cases (example from Polard & Sag (1987: 126)).

(51) a. Er traf den Mann.

he.NOM met the.ACC man

b. Er begegnete dem Mann.

he.NOM met the.DAT man

One has to specify the case that the respective verbs require in the lexical items of the verbs.³⁰

²⁹ Frankfurter Rundschau, 07.11.1997, p. 6.

³⁰ Or at least mark the fact that *treffen* takes an object with the default case for objects and *begegnen* takes a dative object in German. See Haider (1985a), Heinz & Matiassek (1994), and Müller (2001) on structural and lexical case.

A radical variant of the plugging approach is suggested by Haugereid (2009).³¹ Haugereid (pages 12–13) assumes that the syntax combines a verb with an arbitrary combination of a subset of five different argument roles. Which arguments can be combined with a verb is not restricted by the lexical item of the verb.³² A problem for such views is that the meaning of an ambiguous verb sometimes depends on which of its arguments are expressed. The German verb *borgen* has the two translations ‘borrow’ and ‘lend’, which basically are two different perspectives on the same event (see Kunze (1991, 1993) for an extensive discussion of verbs of exchange of possession). Interestingly, the dative object is obligatory only with the ‘lend’ reading (Müller 2010a: 403):

- (52) a. Ich borge ihm das Eichhörnchen.
 I lend him the squirrel
 ‘I lend the squirrel to him.’
 b. Ich borge (mir) das Eichhörnchen.
 I borrow me the squirrel
 ‘I borrow the squirrel.’

If we omit it, we get only the ‘borrow’ reading. So the grammar must specify for specific verbs that certain arguments are necessary for a certain verb meaning or a certain perspective on an event.

Synonyms with differing valence specifications include the minimal triplet mentioned earlier: *dine* is obligatorily intransitive (or takes an *on*-PP), *devour* is transitive, and *eat* can be used either intransitively or transitively (Dowty 1989: 89–90). Many other examples are given in Levin (1993) and Levin & Rappaport Hovav (2005).

In a phrasal constructionist approach one would have to assume phrasal patterns with the preposition or case, into which the verb is inserted. For (49b), the pattern includes a prepositional object with *auf* and an accusative NP, plus an entry for *warten* specifying that it can be inserted into such a structure (see Kroch & Joshi (1985: Section 5.2) for such a proposal in the framework of TAG). Since there are generalizations regarding verbs with such valence representations, one would be forced to have two inheritance hierarchies: one for lexical entries with their valence properties and another one for specific phrasal patterns that are needed for the specific constructions in which these lexical items can be used.

More often, proponents of neo-constructionist approaches either make proposals that are difficult to distinguish from lexical valence structures (see Section 21.3.7 below) or simply decline to address the problem. For instance, Lohndal (2012) writes:

An unanswered question on this story is how we ensure that the functional heads occur together with the relevant lexical items or roots. This is a general problem for the view that Case is assigned by functional heads, and I do not have anything to say about this issue here. (Lohndal 2012)

³¹ Technical aspects of Haugereid’s approach are discussed in Section 21.3.6.

³² Haugereid has the possibility to impose valence restrictions on verbs, but he claims that he uses this possibility just in order to get a more efficient processing of his computer implementation (p. 13).

We think that getting case assignment right in simple sentences, without vast overgeneration of ill-formed word sequences, is a minimal requirement for a linguistic theory.

21.3.5 Expletives

A final example for the irreducibility of valence to semantics are verbs that select for expletives and reflexive arguments of inherently reflexive verbs in German:

- (53)
- a. weil es regnet
because it rains
 - b. weil (es) mir (vor der Prüfung) graut
because EXPL me.DAT before the exam dreads
'because I am dreading the exam'
 - c. weil er es bis zum Professor bringt
because he EXPL until to.the professor brings
'because he made it to professor'
 - d. weil es sich um den Montag handelt
because EXPL REFL around the Monday trades
'because it is about the Monday'
 - e. weil ich mich (jetzt) erhole
because I myself now recreate
'because I am relaxing now'

The lexical heads in (53) need to contain information about the expletive subjects/objects and/or reflexive pronouns that do not fill semantic roles. Note that German allows for subjectless predicates and hence the presence of expletive subjects cannot be claimed to follow from general principles. (53c) is an example with an expletive object. Explanations referring to the obligatory presence of a subject would fail on such examples in any case. Furthermore it has to be ensured that *erholen* is not realized in the [Sbj IntrVerb] construction for intransitive verbs or respective functional categories in a Minimalist setting although the relation *erholen*' (*relax*') is a one-place predicate and hence *erholen* is semantically compatible with the construction.

21.3.6 An exoskeletal approach

In what follows I discuss Haugereid's proposal (2007) in more detail. His analysis has all the high-level problems that were mentioned in the previous subsections, but since it is worked out in detail it is interesting to see its predictions.

Haugereid (2007), working in the framework of HPSG, suggests an analysis along the lines of Borer (2005) where the meaning of an expression is defined as depending on the arguments that are present. He assumes that there are five argument slots that are assigned to semantic roles as follows:

- Arg1: agent or source
- Arg2: patient
- Arg3: benefactive or recipient
- Arg4: goal
- Arg5: antecedent

Here, antecedent is a more general role that stands for instrument, comitative, manner and source. The roles Arg1–Arg3 correspond to subject and objects. Arg4 is a resultative predicate of the end of a path. Arg4 can be realized by a PP, an AP or an NP. (54) gives examples for the realization of Arg4:

- (54) a. John smashed the ball *out of the room*.
 b. John hammered the metal *flat*.
 c. He painted the car *a brilliant red*.

Whereas Arg4 follows the other participants in the causal chain of events, the antecedent precedes the patient in the order of events. It is realized as a PP. (55) is an example of the realization of Arg5:

- (55) John punctured the balloon *with a needle*.

Haugereid now assumes that argument frames consist of these roles. He provides the examples in (56):

- | | | |
|------|---|----------------|
| (56) | a. John smiles. | (arg1-frame) |
| | b. John smashed the ball. | (arg12-frame) |
| | c. The boat arrived. | (arg2-frame) |
| | d. John gave Mary a book. | (arg123-frame) |
| | e. John gave a book to Mary. | (arg124-frame) |
| | f. John punctured the ball with a needle. | (arg125-frame) |

Haugereid points out that multiple verbs can occur in multiple argument frames. He provides the variants in (57) for the verb *drip*:

- | | | |
|------|---|----------------|
| (57) | a. The roof drips. | (arg1-frame) |
| | b. The doctor drips into the eyes. | (arg14-frame) |
| | c. The doctor drips with water. | (arg15-frame) |
| | d. The doctor drips into the eyes with water. | (arg145-frame) |
| | e. The roof drips water. | (arg12-frame) |
| | f. The roof drips water into the bucket. | (arg124-frame) |
| | g. The doctor dripped the eyes with water. | (arg125-frame) |

21.3 Radical underspecification: the end of argument structure?

- h. The doctor dripped into the eyes with water. (arg145-frame)
- i. John dripped himself two drops of water. (arg123-frame)
- j. John dripped himself two drops of water into his eyes. (arg1234-frame)
- k. John dripped himself two drops of water into his eyes with a drop counter. (arg12345-frame)
- l. Water dripped. (arg2-frame)
- m. It drips. (arg0-frame)

He proposes the inheritance hierarchy in Figure 21.3 in order to represent all possible argument combinations. The Arg5 role is omitted due to space considerations.

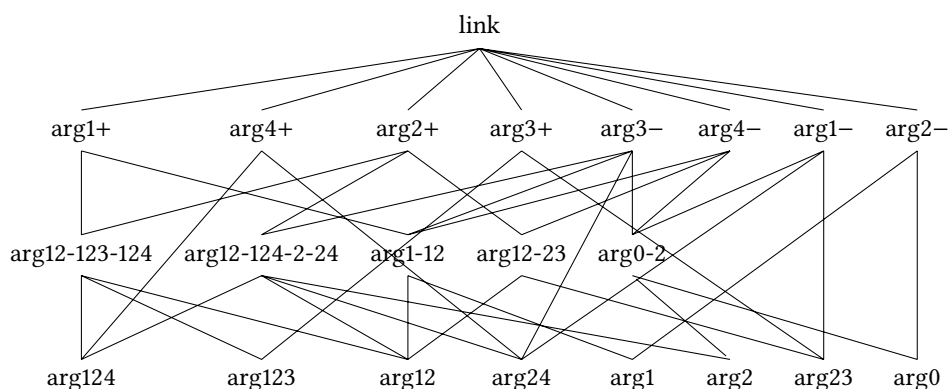


图 21.3: Hierarchy of argument frames following Haugereid (2007)

Haugereid assumes binary-branching structures where arguments can be combined with a head in any order. There is a dominance schema for each argument role. The schema realizing the argument role 3 provides a link value *arg3+*. If the argument role 2 is provided by another schema, we arrive at the frame *arg23*. For unergative intransitive verbs, it is possible to determine that it has an argument frame of *arg1*. This frame is only compatible with the types *arg1+*, *arg2-*, *arg3-* and *arg4-*. Verbs that have an optional object are assigned to *arg1-12* according to Haugereid. This type allows for the following combinations: *arg1+*, *arg2-*, *arg3-* and *arg4-* such as *arg1+*, *arg2+*, *arg3-* and *arg4-*.

This approach comes very close to an idea by Goldberg: verbs are underspecified with regard to the sentence structures in which they occur and it is only the actual realization of arguments in the sentence that decides which combinations of arguments are realized. One should bear in mind that the hierarchy in Figure 21.3 corresponds to a considerable disjunction: it lists all possible realizations of arguments. If we say that *essen* ‘to eat’ has the type *arg1-12*, then this corresponds to the disjunction *arg1* \vee *arg12*. In addition to the information in the hierarchy above, one also requires information about the syntactic properties of the arguments (case, the form of prepositions, verb forms in verbal complements). Since this information is in part specific to each verb (see

Section 21.1), it cannot be present in the dominance schemata and must instead be listed in each individual lexical entry. For the lexical entry for *warten auf* ‘wait for’, there must be information about the fact that the subject has to be an NP and that the prepositional object is an *auf*-PP with accusative. The use of a type hierarchy then allows one to elegantly encode the fact that the prepositional object is optional. The difference to a disjunctively specified SUBCAT list with the form of (58) is just a matter of formalization.

(58) SUBCAT $\langle \text{NP}[\textit{str}] \rangle \vee \langle \text{NP}[\textit{str}], \text{PP}[\textit{auf}, \textit{acc}] \rangle$

Since Haugereid’s structures are binary-branching, it is possible to derive all permutations of arguments (59a–b), and adjuncts can be attached to every branching node (59c–d).

- (59)
- a. dass $[\textit{arg}_1 \text{ keiner } [\textit{arg}_2 \text{ Pizza isst}]]$
 that nobody pizza eats
 ‘that nobody eats pizza’
 - b. dass $[\textit{arg}_2 \text{ Pizza } [\textit{arg}_1 \text{ keiner isst}]]$
 that pizza nobody eats
 - c. dass $[\textit{arg}_1 \text{ keiner } [\textit{gerne } [\textit{arg}_2 \text{ Pizza isst}]]]$
 that nobody gladly pizza eats
 ‘that nobody eats pizza gladly’
 - d. dass $[[\textit{hier } [\textit{arg}_1 \text{ keiner } [\textit{arg}_2 \text{ Pizza isst}]]]]$
 that here nobody pizza eats
 ‘that nobody eats pizza here’

Haugereid has therefore found solutions for some of the problems in Goldberg’s analysis that were pointed out in Müller (2006). Nevertheless, there are a number of other problems, which I will discuss in what follows. In Haugereid’s approach, nothing is said about the composition of meaning. He follows the so-called Neo-Davidsonian approach. In this kind of semantic representation, arguments of the verb are not directly represented on the verb. Instead, the verb normally has an event argument and the argument roles belonging to the event in question are determined in a separate predication. (60) shows two alternative representations, where *e* stands for the event variable.

- (60)
- a. Der Mann isst eine Pizza.
 the man eats a pizza
 ‘The man is eating a pizza’
 - b. $\textit{eat}'(e, x, y) \wedge \textit{man}'(x) \wedge \textit{pizza}'(y)$
 - c. $\textit{eat}'(e) \wedge \textit{agent}(e, x) \wedge \textit{theme}(e, y) \wedge \textit{man}'(x) \wedge \textit{pizza}'(y)$

Haugereid adopts Minimal Recursion Semantics (MRS) as his semantic formalism (see also Section 9.1.6 and 19.3). The fact that arguments belong to a particular predicate is represented by the fact that the relevant predicates have the same handle. The representation in (60c) corresponds to (61):

(61) $h1:essen'(e), h1:arg1(x), h1:arg2(y), h2:mann'(x), h3:pizza'(y)$

This analysis captures Goldberg's main idea: meaning arises from particular constituents being realized together with a head.

For the sentence in (62a), Haugereid (2007, p. c.) assumes the semantic representation in (62b):³³

- (62) a. der Mann den Teich leer fischt
 the man the pond empty fishes
 b. $h1:mann'(x), h2:teich'(y), h3:leer'(e),$
 $h4:fischen'(e2), h4:arg1(x), h4:arg2(y), h4:arg4(h3)$

In (62b), the *arg1*, *arg2* and *arg4* relations have the same handle as *fischen'*. Following Haugereid's definitions, this means that *arg2* is the patient of the event. In the case of (62a), this makes incorrect predictions since the accusative element is not a semantic argument of the main verb. It is a semantic argument of the secondary predicate *leer* 'empty' and has been raised to the object of the resultative construction. Depending on the exact analysis assumed, the accusative object is either a syntactic argument of the verb or of the adjective, however, it is never a semantic argument of the verb. In addition to this problem, the representation in (62b) does not capture the fact that *leer* 'empty' predicates over the object. Haugereid (2007, p. c.) suggests that this is implicit in the representation and follows from the fact that all *arg4s* predicate over all *arg2s*. Unlike Haugereid's analysis, analyses using lexical rules that relate a lexical item of a verb to another verbal item with a resultative meaning allow for a precise specification of the semantic representation that then captures the semantic relation between the predicates involved. In addition, the lexical rule-based analysis makes it possible to license lexical items that do not establish a semantic relation between the accusative object and the verb (Wechsler 1997; Wechsler & Noh 2001; Müller 2002a: Chapter 5).

Haugereid sketches an analysis of the syntax of the German clause and tackles active/passive alternations. However, certain aspects of the grammar are not elaborated on. In particular, it remains unclear how complex clauses containing AcI verbs such as *sehen* 'to see' and *lassen* 'to let' should be analyzed. Arguments of embedded and embedding verbs can be permuted in these constructions. Haugereid (2007, p. c.) assumes special rules that allow to saturate arguments of more deeply embedded verbs, for example, a special rule that combines an *arg2* argument of an argument with a verb. In order to combine *das Nilpferd* and *nicht füttern helfen lässt* in sentences such as (63), he is forced to assume a special grammatical rule that combines an argument of a doubly embedded verb with another verb:

³³ See Haugereid (2009: 165) for an analysis of the Norwegian examples in (i).

(i) Jon maler veggen rød.
 Jon paints wall.DEF red
 'Jon paints the wall red.'

- (63) weil Hans Cecilia John das Nilpferd nicht füttern helfen lässt
 because Hans Cecilia John the hippo not feed help let
 ‘because Hans is not letting Cecilia help John feed the hippo.’

In Müller (2004c: 220), I have argued that embedding under complex-forming predicates is only constrained by performance factors (see also Section 12.6.3). In German, verbal complexes with more than four verbs are barely acceptable. Evers (1975: 58–59) has pointed out, however, that the situation in Dutch is different since Dutch verbal complexes have a different branching: in Dutch, verbal complexes with up to five verbs are possible. Evers attributes this difference to a greater processing load for German verbal complexes (see also Gibson 1998: Section 3.7). Haugereid would have to assume that there are more rules for Dutch than for German. In this way, he would give up the distinction between competence and performance and incorporate performance restrictions directly into the grammar. If he wanted to maintain a distinction between the two, then Haugereid would be forced to assume an infinite number of schemata or a schema with functional uncertainty since depth of embedding is only constrained by performance factors. Existing HPSG approaches to the analysis of verbal complexes do without functional uncertainty (Hinrichs & Nakazawa 1994). Since such raising analyses are required for object raising anyway (as discussed above), they should be given preference.

Summing up, it must be said that Haugereid’s exoskeletal approach does account for different orderings of arguments, but it neither derives the correct semantic representations nor does it offer a solution for the problem of idiosyncratic selection of arguments and the selection of expletives.

21.3.7 Is there an alternative to lexical valence structure?

The question for theories denying the existence of valence structure is what replaces it to explain idiosyncratic lexical selection. In her exoskeletal approach, Borer (2005) explicitly rejects lexical valence structures. But she posits post-syntactic interpretive rules that are difficult to distinguish from them. To explain the correlation of *depend* with an *on*-PP, she posits the following interpretive rule (Borer 2005: Vol. II, p. 29):

- (64) MEANING $\Leftrightarrow \pi_9 + [\langle e^{on} \rangle]$

Borer refers to all such cases of idiosyncratic selection as idioms. In a rule such as (64), “MEANING is whatever the relevant idiom means” (Borer 2005: Vol. II, p. 27). In (64), π_9 is the “phonological index” of the verb *depend* and e^{on} “corresponds to an open value that must be assigned range by the f-morph *on*” (Borer 2005: Vol. II, p. 29), where f-morphs are function words or morphemes. Hence this rule brings together much the same information as the lexical valence structure in (48c). Discussing such “idiom” rules, Borer writes

Although by assumption a listeme cannot be associated with any grammatical properties, one device used in this work has allowed us to get around the formidable

restrictions placed on the grammar by such a constraint – the formation of idioms. [...] Such idiomatic specification could be utilized, potentially, not just for *arrive* and *depend on*, but also for obligatorily transitive verbs [...], for verbs such as *put*, with their obligatory locative, and for verbs which require a sentential complement.

The reader may object that subcategorization, of sorts, is introduced here through the back door, with the introduction, in lieu of lexical syntactic annotation, of an articulated listed structure, called an *idiom*, which accomplishes, de facto, the same task. The objection of course has some validity, and at the present state of the art, the introduction of idioms may represent somewhat of a concession.

(Borer 2005: Vol. II, p. 354–355)

Borer goes on to pose various questions for future research, related to constraining the class of possible idioms. With regard to that research program it should be noted that a major focus of lexicalist research has been narrowing the class of subcategorization and extricating derivable properties from idiosyncratic subcategorization. Those are the functions of HPSG lexical hierarchies, for example.

21.3.8 Summary

In Sections 21.3.2–21.3.5 we showed that the question of which arguments must be realized in a sentence cannot be reduced to semantics and world knowledge or to general facts about subjects. The consequence is that valence information has to be connected to lexical items. One therefore must either assume a connection between a lexical item and a certain phrasal configuration as in Croft's approach (2003) and in LTAG or assume our lexical variant. In a Minimalist setting the right set of features must be specified lexically to ensure the presence of the right case assigning functional heads. This is basically similar to the lexical valence structures we are proposing here, except that it needlessly introduces various problems discussed above, such as the problem of coordination raised in Section 21.2.1.

21.4 Relations between constructions

On the lexical rules approach, word forms are related by lexical rules: a verb stem can be related to a verb with finite inflection and to a passive verb form; verbs can be converted to adjectives or nouns; and so on. The lexical argument structure accompanies the word and can be manipulated by the lexical rule. In this section we consider what can replace such rules within a phrasal or ASC approach.

21.4.1 Inheritance hierarchies for constructions

For each valence structure that the lexicalist associates with a root lexeme (transitive, ditransitive, etc.), the phrasal approach requires multiple phrasal constructions, one to replace each lexical rule or combination of lexical rules that can apply to the word.

Taking ditransitives, for example, the phrasal approach requires an active-ditransitive construction, a passive-ditransitive construction, and so on, to replace the output of every lexical rule or combination of lexical rules applied to a ditransitive verb. (Thus Bergen & Chang (2005: 169–170) assume an active-ditransitive and a passive-ditransitive construction and Kallmeyer & Osswald (2012: 171–172) assume active and passive variants of the transitive construction.) On that view some of the active voice constructions for German would be:

- (65) a. Nom V
- b. Nom Acc V
- c. Nom Dat V
- d. Nom Dat Acc V

The passive voice constructions corresponding to (65) would be:

- (66) a. V V-Aux
- b. Nom V V-Aux
- c. Dat V V-Aux
- d. Dat Nom V V-Aux

Merely listing all these constructions is not only uneconomical but fails to capture the obvious systematic relation between active and passive constructions. Since phrasalists reject both lexical rules and transformations, they need an alternative way to relate phrasal configurations and thereby explain the regular relation between active and passive. The only proposals to date involve the use of inheritance hierarchies, so let us examine them.

Researchers working in various frameworks, both with lexical and phrasal orientation, have tried to develop inheritance-based analyses that could capture the relation between valence patterns such as those in (65) and (66) (see for instance Kay & Fillmore 1999: 12; Michaelis & Ruppenhofer 2001: Chapter 4; Candito 1996; Clément & Kinyon 2003: 188; Kallmeyer & Osswald 2012: 171–172; Koenig 1999: Chapter 3; Davis & Koenig 2000; Kordoni 2001 for proposals in CxG, TAG, and HPSG). The idea is that a single representation (lexical or phrasal, depending on the theory) can inherit properties from multiple constructions. In a phrasal approach the description of the pattern in (65b) inherits from the transitive and the active construction and the description of (66b) inherits from both the transitive and the passive constructions. Figure 21.4 在下一頁 illustrates the inheritance-based lexical approach: a lexical entry for a verb such as *read* or *eat* is combined with either an active or passive representation. The respective representations for the active and passive are responsible for the expression of the arguments.

As was already discussed in Section 10.2, inheritance-based analyses cannot account for multiple changes in valence as for instance the combination of passive and impersonal construction that can be observed in languages like Lithuanian (Timberlake 1982: Section 5), Irish (Noonan 1994), and Turkish (Özkaragöz 1986). Özkaragöz's Turkish examples are repeated here with the original glossing as (67) for convenience:

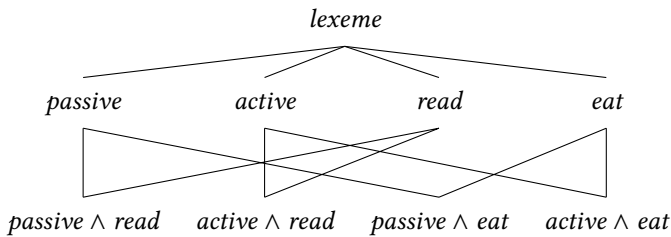


图 21.4: Inheritance Hierarchy for active and passive

- (67) a. Bu şato-da boğ-ul-un-ur. (Turkish)
 this château-LOC strangle-PASS-PASS-AOR
 ‘One is strangled (by one) in this château.’
 b. Bu oda-da döv-ül-ün-ür.
 this room-LOC hit-PASS-PASS-AOR
 ‘One is beaten (by one) in this room.’
 c. Harp-te vur-ul-un-ur.
 war-LOC shoot-PASS-PASS-AOR
 ‘One is shot (by one) in war.’

Another example from Section 10.2 that cannot be handled with inheritance is multiple causativization in Turkish. Turkish allows double and even triple causativization (Lewis 1967: 146):

- (68) Öl-dür-t-tür-t- (Turkish)
 ‘to cause somebody to cause somebody to kill somebody’

An inheritance-based analysis would not work, since inheriting the same information several times does not add anything new. Krieger & Nerbonne (1993) make the same point with respect to derivational morphology in cases like *preprepreversion*: inheriting information about the prefix *pre-* twice or more often, does not add anything.

So assuming phrasal models, the only way to capture the generalization with regard to (65) and (66) seems to be to assume GPSG-like metarules that relate the constructions in (65) to the ones in (66). If the constructions are lexically linked as in LTAG, the respective mapping rules would be lexical rules. For approaches that combine LTAG with the Goldbergian plugging idea such as the one by Kallmeyer & Osswald (2012) one would have to have extended families of trees that reflect the possibility of having additional arguments and would have to make sure that the right morphological form is inserted into the respective trees. The morphological rules would be independent of the syntactic structures in which the derived verbal lexemes could be used. One would have to assume two independent types of rules: GPSG-like metarules that operate on trees and morphological rules that operate on stems and words. We believe that this is an unnecessary

complication and apart from being complicated the morphological rules would not be acceptable as form-meaning pairs in the CxG sense since one aspect of the form namely that additional arguments are required is not captured in these morphological rules. If such morphological rules were accepted as proper constructions then there would not be any reason left to require that the arguments have to be present in a construction in order for it to be recognizable, and hence, the lexical approach would be accepted.³⁴

Inheritance hierarchies are the main explanatory device in Croft's Radical Construction Grammar (Croft 2001). He also assumes phrasal constructions and suggests representing these in a taxonomic network (an inheritance hierarchy). He assumes that every idiosyncrasy of a linguistic expression is represented on its own node in this kind of network. Figure 21.5 shows part of the hierarchy he assumes for sentences. There are

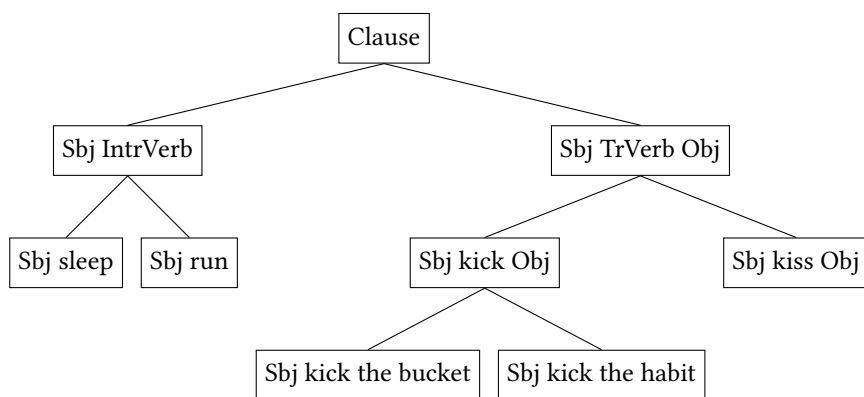


图 21.5: Classification of phrasal patterns in Croft (2001: 26)

sentences with intransitive verbs and sentences with transitive verbs. Sentences with the form [Sbj kiss Obj] are special instances of the construction [Sbj TrVerb Obj]. The [Sbj kick Obj] construction also has further sub-constructions, namely the constructions [Sbj kick the bucket] and [Subj kick the habit]. Since constructions are always pairs of form and meaning, this gives rise to a problem: in a normal sentence with *kick*, there is a kicking relation between the subject and the object of *kick*. This is not the case for the idiomatic use of *kick* in (69):

(69) He kicked the bucket.

This means that there cannot be a normal inheritance relation between the [Sbj kick Obj] and the [Sbj kick the bucket] construction. Instead, only parts of the information may be inherited from the [Sbj kick Obj] construction. The other parts are redefined by the sub-construction. This kind of inheritance is referred to as *default inheritance*.

kick the bucket is a rather fixed expression, that is, it is not possible to passivize it or front parts of it without losing the idiomatic reading (Nunberg, Sag & Wasow 1994:

³⁴ Compare the discussion of *Totschießen* 'shoot dead' in example (81) below.

508). However, this is not true for all idioms. As Nunberg, Sag & Wasow (1994: 510) have shown, there are idioms that can be passivized (70a) as well as realizations of idioms where parts of idioms occur outside of the clause (70b).

- (70) a. The beans were spilled by Pat.
b. The strings [that Pat pulled] got Chris the job.

The problem is now that one would have to assume two nodes in the inheritance hierarchy for idioms that can undergo passivization since the realization of the constituents is different in active and passive variants but the meaning is nevertheless idiosyncratic. The relation between the active and passive form would not be captured. Kay (2002) has proposed an algorithm for computing objects (Construction-like objects = CLOs) from hierarchies that then license active and passive variants. As I have shown in Müller (2006: Section 3), this algorithm does not deliver the desired results and it is far from straightforward to improve it to the point that it actually works. Even if one were to adopt the changes I proposed, there are still phenomena that cannot be described using inheritance hierarchies (see Section 10.2 in this book).

A further interesting point is that the verbs have to be explicitly listed in the constructions. This begs the question of how constructions should be represented where the verbs are used differently. If a new node in the taxonomic network is assumed for cases like (71), then Goldberg's criticism of lexical analyses that assume several lexical entries for a verb that can appear in various constructions³⁵ will be applicable here: one would have to assume constructions for every verb and every possible usage of that verb.

- (71) He kicked the bucket into the corner.

For sentences with negation, Croft assumes the hierarchy with multiple inheritance given in Figure 21.6. The problem with this kind of representation is that it remains

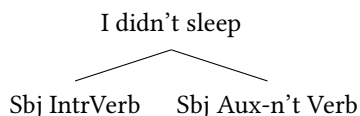


图 21.6: Interaction of phrasal patterns following Croft (2001: 26)

unclear as to how the semantic embedding of the verb meaning under negation can be represented. If all constructions are pairs of form and meaning, then there would have to be a semantic representation for [Sbj IntrVerb] (CONT value or SEM value). Similarly, there would have to be a meaning for [Sbj Aux-n't Verb]. The problem now arises that the meaning of [Sbj IntrVerb] has to be embedded under the meaning of the negation and

³⁵ Note the terminology: I used the word *lexical entry* rather than *lexical item*. The HPSG analysis uses lexical rules that correspond to Goldberg's templates. What Goldberg criticizes is lexical rules that relate lexical entries, not lexical rules that licence new lexical items, which may be stored or not. HPSG takes the latter approach to lexical rules. See Section 9.2.

this cannot be achieved directly using inheritance since X and $\text{not}(X)$ are incompatible. There is a technical solution to this problem using auxiliary features. Since there are a number of interactions in grammars of natural languages, this kind of analysis is highly implausible if one claims that features are a direct reflection of observable properties of linguistic objects. For a more detailed discussion of approaches with classifications of phrasal patterns, see Müller (2010b) as well as Müller (2007b: Section 18.3.2.2) and for the use of auxiliary features in inheritance-based analyses of the lexicon, see Müller (2007b: Section 7.5.2.2).

21.4.2 Mappings between different levels of representations

Culicover & Jackendoff (2005: Chapter 6.3) suggest that passive should be analyzed as one of several possible mappings from the Grammatical Function tier to the surface realization of arguments. Surface realizations of referential arguments can be NPs in a certain case, with certain agreement properties, or in a certain position. While such analyses that work by mapping elements with different properties onto different representations are common in theories like LFG and HPSG (Koenig 1999; Bouma, Malouf & Sag 2001a), a general property of these analyses is that one needs one level of representation per interaction of phenomena (ARG-ST, SEM-ARG, ADD-ARG in Koenig's proposal, ARG-ST, DEPS, SPR, COMPS in Bouma, Malouf and Sag's proposal). This was discussed extensively in Müller (2007b: Section 7.5.2.2) with respect to extensions that would be needed for Koenig's analysis.

Since Culicover and Jackendoff argue for a phrasal model, we will discuss their proposal here. Culicover and Jackendoff assume a multilayered model in which semantic representations are linked to grammatical functions, which are linked to tree positions. Figure 21.7 shows an example for an active sentence. GF stands for Grammatical Func-

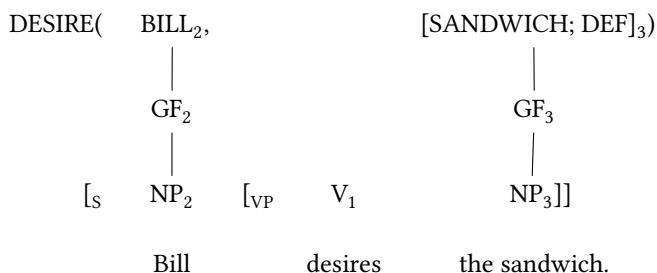


图 21.7: Linking grammatical functions to tree positions: active

tion. Culicover & Jackendoff (2005: 204) explicitly avoid names like Subject and Object since this is crucial for their analysis of the passive to work. They assume that the first GF following a bracket is the subject of the clause the bracket corresponds to (p. 195–196) and hence has to be mapped to an appropriate tree position in English. Note that this view of grammatical functions and obliqueness does not account for subjectless sentences that

are possible in some languages, for instance in German.³⁶

Regarding the passive, the authors write:

we wish to formulate the passive not as an operation that deletes or alters part of the argument structure, but rather as a piece of structure in its own right that can be unified with the other independent pieces of the sentence. The result of the unification is an alternative licensing relation between syntax and semantics. (Culicover & Jackendoff 2005: 203)

They suggest the following representation of the passive:

$$(72) \quad [GF_i > [GF \dots]]_k \Leftrightarrow [\dots V_k + \text{pass} \dots (\text{by NP}_i) \dots]_k$$

The italicized parts are the normal structure of the sentence and the non-italicized parts are an overlay on the normal structure, that is, additional constraints that have to hold in passive sentences. Figure 21.8 shows the mapping of the example discussed above that corresponds to the passive.

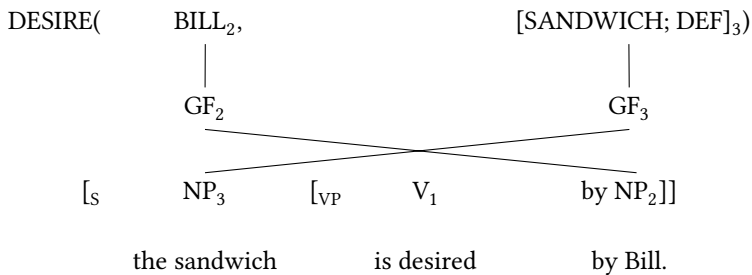


图 21.8: Linking grammatical functions to tree positions: passive

Although Culicover and Jackendoff emphasize the similarity between their approach and Relational Grammar (Perlmutter 1983), there is an important difference: in Relational Grammar additional levels (strata) can be stipulated if additional remappings are needed. In Culicover and Jackendoff's proposal there is no additional level. This causes problems for the analysis of languages which allow for multiple argument alternations. Examples from Turkish were provided in (67). Approaches that assume that the personal passive is the unification of a general structure with a passive-specific structure will not be able to capture this, since they committed to a certain structure too early. The problem for approaches that state syntactic structure for the passive is that such a structure, once stated, cannot be modified. Culicover and Jackendoff's proposal works in this respect since there are no strong constraints in the right-hand side of their constraint in (72). But there is a different problem: when passivization is applied the second time, it has to apply to the innermost bracket, that is, the result of applying (72) should be:

³⁶ Of course one could assume empty expletive subjects, as was suggested by Grewendorf (1993: 1311), but empty elements and especially those without meaning are generally avoided in the constructionist literature. See Müller (2010a: Section 3.4, Section 11.1.1.3) for further discussion.

$$(73) \quad [GF_i > [GF_j \dots]]_k \Leftrightarrow [\dots V_k + \text{pass } \dots (\text{by } NP_i) \dots (\text{by } NP_j) \dots]_k$$

This cannot be done with unification, since unification checks for compatibility and since the first application of passive was possible it would be possible for the second time as well. Dots in representations are always dangerous and in the example at hand one would have to make sure that NP_i and NP_j are distinct, since the statement in (72) just says there has to be a *by*-PP somewhere. What is needed instead of unification would be something that takes a GF representation and searches for the outermost bracket and then places a bracket to the left of the next GF. But this is basically a rule that maps one representation onto another one, just like lexical rules do.

If Culicover and Jackendoff want to stick to a mapping analysis, the only option to analyze the data seems to be to assume an additional level for impersonal passives from which the mapping to phrase structure is done. In the case of Turkish sentences like (74), which is a personal passive, the mapping to this level would be the identity function.

- (74) Arkadaş-ım bu oda-da döv-ül-dü.
 friend-my this room-LOC hit-PASS-AOR
 ‘My friend is beaten (by one) in this room.’

In the case of passivization + impersonal construction, the correct mappings would be implemented by two mappings between the three levels that finally result in a mapping as the one that is seen in (67b), repeated here as (75) for convenience.

- (75) Bu oda-da döv-ül-ün-ür.
 this room-LOC hit-PASS-PASS-AOR
 ‘One is beaten (by one) in this room.’

Note that passivization + impersonal construction is also problematic for purely inheritance based approaches. What all these approaches can suggest though is that they just stipulate four different relations between argument structure and phrase structure: active, passive, impersonal construction, passive + impersonal construction. But this misses the fact that (75) is an impersonal variant of the passive in (74).

In contrast, the lexical rule-based approach suggested by Müller (2003b) does not have any problems with such multiple alternations: the application of the passivization lexical rule suppresses the least oblique argument and provides a lexical item with the argument structure of a personal passive. Then the impersonal lexical rule applies and suppresses the now least oblique argument (the object of the active clause). The result is impersonal constructions without any arguments as the one in (75).

21.4.3 Is there an alternative to lexical rules?

In this section we have reviewed the attempts to replace lexical rules with methods of relating constructions. These attempts have not been successful, in our assessment. We believe that the essential problem with them is that they fail to capture the derivational character of the relationship between certain word forms. Alternations signaled

by passive voice and causative morphology are relatively simple and regular when formulated as operations on lexical valence structures that have been abstracted from their phrasal context. But non-transformational rules or systems formulated on the phrasal structures encounter serious problems that have not yet been solved.

21.5 Further problems for phrasal approaches

Müller (2006) discussed the problems shared by proposals that assume phrasal constructions to be a fixed configuration of adjacent material as for instance the one by Goldberg & Jackendoff (2004). I showed that many argument structure constructions allow great flexibility as far as the order of their parts is concerned. Back then I discussed resultative constructions in their interaction with free datives, passive and other valence changing phenomena and showed that for all these constructions licensed by such interactions the construction parts can be scrambled, the verb can appear in different positions, arguments can be extracted and so on. The following subsection discusses particle verbs, which pose similar problems for theories that assume a phrasal construction with fixed order of verb and particle.

21.5.1 Particle verbs and commitment to phrase structure configurations

A general problem of approaches that assume phrase structure configurations paired with meaning is that the construction may appear in different contexts: the construction parts may be involved in derivational morphology (as discussed in the previous subsection) or the construction parts may be involved in dislocations. A clear example of the latter type is the phrasal analysis of particle verbs that was suggested by Booij (2002: Section 2; 2012) and Blom (2005), working in the frameworks of Construction Grammar and LFG, respectively. The authors working on Dutch and German assume that particle verbs are licensed by phrasal constructions (pieces of phrase structure) in which the first slot is occupied by the particle.

(76) $[X []_V]_{V'}$ where $X = P, Adv, A, \text{ or } N$

Examples for specific Dutch constructions are:

- (77) a. $[af []_V]_{V'}$
 b. $[door []_V]_{V'}$
 c. $[op []_V]_{V'}$

This suggestion comes with the claim that particles cannot be fronted. This claim is made frequently in the literature, but it is based on introspection and wrong for languages like Dutch and German. On Dutch see Hoeksema (1991: 19), on German, Müller (2002a,c, 2003d, 2007e).³⁷ A German example is given in (78); several pages of attested examples

³⁷ Some more fundamental remarks on introspection and corpus data with relation to particle verbs can also be found in Müller (2007e) and Meurers & Müller (2009).

can be found in the cited references and some more complex examples will also be discussed in Section 21.7.3 on page 604.

- (78) *Los damit geht es schon am 15. April.*³⁸
 PART there.with goes it already at.the 15 April
 ‘It already starts on April the 15th.’

Particle verbs are mini-idioms. So the conclusion is that idiomatic expressions that allow for a certain flexibility in order should not be represented as phrasal configurations describing adjacent elements. For some idioms, a lexical analysis along the lines of Sag (2007) seems to be required.³⁹ The issue of particle verbs will be taken up in Section 21.7.3 again, where we discuss evidence for/against phrasal analyses from neuroscience.

21.6 Arguments from language acquisition

The question whether language acquisition is pattern-based and hence can be seen as evidence for the phrasal approach has already been touched upon in the Sections 16.3 and 16.4. It was argued that constructions can be realized discontinuously in coordinations and hence it is the notion of dependency that has to be acquired; acquiring simple continuous patterns is not sufficient.

Since the present discussion about phrasal and lexical approaches deals with specific proposals, I would like to add two more special subsections: Section 21.6.1 deals with the recognizability of constructions and Section 21.6.2 discusses specific approaches to coordination in order to demonstrate how frameworks deal with the discontinuous realization of constructions.

21.6.1 Recognizability of constructions

I think that a purely pattern-based approach is weakened by the existence of examples like (79):

- (79) a. John tried to sleep.
 b. John tried to be loved.

Although no argument of *sleep* is present in the phrase *to sleep* and neither a subject nor an object is realized in the phrase *to be loved*, both phrases are recognized as phrases containing an intransitive and a transitive verb, respectively.⁴⁰

³⁸ taz, 01.03.2002, p. 8, see also Müller (2005c: 313).

³⁹ Note also that the German example is best described as a clause with a complex internally structured constituent in front of the finite verb and it is doubtful whether linearization-based proposals like the ones in Kathol (1995: 244–248) or Wetta (2011) can capture this. See also the discussion of multiple frontings in connection to Dependency Grammar in Section 11.7.1.

⁴⁰ Constructionist theories do not assume empty elements. Of course, in the GB framework the subject would be realized by an empty element. So it would be in the structure, although inaudible.

The same applies to arguments that are supposed to be introduced/licensed by a phrasal construction: in (80) the resultative construction is passivized and then embedded under a control verb, resulting in a situation in which only the result predicate (*tot* ‘dead’) and the matrix verb (*geschossen* ‘shot’) are realized overtly within the local clause, bracketed here:

- (80) Der kranke Mann wünschte sich, [totgeschossen zu werden].⁴¹
 the sick man wished SELF dead.shot to be
 ‘The sick man wanted to be shot dead.’

Of course passivization and control are responsible for these occurrences, but the important point here is that arguments can remain unexpressed or implicit and nevertheless a meaning usually connected to some overt realization of arguments is present (Müller 2007c: Section 4). So, what has to be acquired by the language learner is that when a result predicate and a main verb are realized together, they contribute the resultative meaning. To take another example, NP arguments that are usually realized in active resultative constructions may remain implicit in nominalizations like the ones in (81):

- (81) a. dann scheint uns das Totschießen mindestens ebensoviel Spaß zu machen⁴²
 then seems us the dead-shooting at.least as.much fun to make
 ‘then the shooting dead seems to us to be as least as much fun’
 b. Wir lassen heut das Totgeschieße,
 we let today the annoying.repeated.shooting.dead
 Weil man sowas heut nicht tut.
 since one such.thing today not does
 Und wer einen Tag sich ausruht,
 and who a day SELF rests
 Der schießt morgen doppelt gut.⁴³
 this shoots tomorrow twice good
 ‘We do not shoot anybody today, since one does not do this today, and those who rest a day shoot twice as well tomorrow.’

The argument corresponding to the patient of the verb (the one who is shot) can remain unrealized, because of the syntax of nominalizations. The resultative meaning is still understood, which shows that it does not depend upon the presence of a resultative construction involving Subj V Obj and Obl.

⁴¹ Müller (2007c: 387).

⁴² <https://www.elitepartner.de/forum/wie-gehen-die-maenner-mit-den-veraenderten-anspruechen-der-frauen-um-26421-6.html>. 26.03.0212.

⁴³ Gedicht für den Frieden, Oliver Kalkofe, <http://www.golyr.de/oliver-kalkofe/songtext-gedicht-fuer-den-frieden-417329.html>. 04.03.2016.

21.6.2 Coordination and discontinuousness

The following subsection deals with analyses of coordination in some of the frameworks that were introduced in this book. The purpose of the section is to show that simple phrasal patterns have to be broken up in coordination structures. This was already mentioned in Section 16.3, but I think it is illuminative to have a look at concrete proposals.

In Categorical Grammar, there is a very elegant treatment of coordination (see Steedman 1991). A generalization with regard to so-called symmetric coordination is that two objects with the same syntactic properties are combined to an object with those properties. We have already encountered the relevant data in the discussion of the motivation for feature geometry in HPSG on page 244. Their English versions are repeated below as (82):

- (82) a. the man and the woman
 b. He knows and loves this record.
 c. He is dumb and arrogant.

Steedman (1991) analyzes examples such as those in (82) with a single rule:

- (83) $X \text{ conj } X \Rightarrow X$

This rule combines two categories of the same kind with a conjunction in between to form a category that has the same category as the conjuncts.⁴⁴ Figure 21.9 shows the analysis of (82a) and Figure 21.10 在下一页 gives an analysis of the corresponding English example of (82b).

$$\begin{array}{ccccccc}
 \textit{the} & \textit{man} & \textit{and} & \textit{the} & \textit{woman} \\
 \hline
 \textit{np/n} & \textit{n} & \textit{conj} & \textit{np/n} & \textit{n} \\
 \hline
 \textit{np} & & & \textit{np} & \\
 & & & & \textit{&} \\
 & & \textit{np} & &
 \end{array}$$

图 21.9: Coordination of two NPs in Categorical Grammar

If we compare this analysis to the one that would have to be assumed in traditional phrase structure grammars, it becomes apparent what the advantages are: one rule was required for the analysis of NP coordination where two NPs are coordinated to form an NP and another was required for the analysis of V coordination. This is not only undesirable from a technical point of view, neither does it capture the basic property of

⁴⁴ Alternatively, one could analyze all three examples using a single lexical entry for the conjunction *and*: *and* is a functor that takes a word or phrase of any category to its right and after this combination then needs to be combined with an element of the same category to its left in order to form the relevant category after combining with this second element. This means that the category for *and* would have the form $(X \backslash X) / X$. This analysis does not require any coordination rules. If one wants to assume, as is common in GB/MP, that every structure has a head, then a headless analysis that assumes a special rule for coordination like the one in (83) would be ruled out.

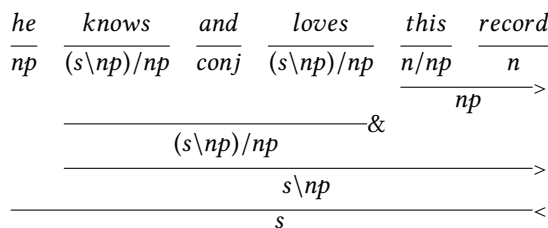


图 21.10: Coordination of two transitive verbs in Categorical Grammar

symmetric coordination: two symbols with the same syntactic category are combined with each other.

It is interesting to note that it is possible to analyze phrases such as (84) in this way:

(84) give George a book and Martha a record

In Section 1.3.2.4, we have seen that this kind of sentences is problematic for constituent tests. However, in Categorical Grammar, it is possible to analyze them without any problems if one adopts rules for type raising and composition as Dowty (1988) and Steedman (1991) do. In Section 8.5, we have already seen forward type raising as well as forward and backward composition. In order to analyze (84), one would require backward type raising repeated in (85) and backward composition repeated in (86):

(85) Backward type raising ($< T$)

$$X \Rightarrow T(T/X)$$

(86) Backward composition ($< B$)

$$Y \setminus Z * X \setminus Y = X \setminus Z$$

Dowty's analysis of (84) is given in Figure 21.11. vp stands for $s \setminus np$.

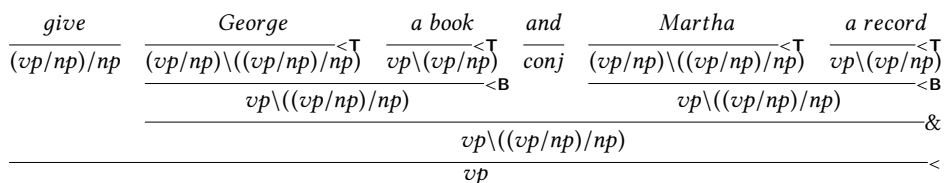


图 21.11: Gapping in Categorical Grammar

This kind of type-raising analysis was often criticized because raising categories leads to many different analytical possibilities for simple sentences. For example, one could first combine a type-raised subject with the verb and then combine the resulting constituent with the object. This would mean that we would have a $[[S\ V]\ O]$ in addition to the standard $[S\ [V\ O]]$ analysis. Steedman (1991) argues that both analyses differ in

terms of information structure and it is therefore valid to assume different structures for the sentences in question.

I will not go into these points further here. However, I would like to compare Steedman's lexical approach to phrasal analyses: all approaches that assume that the ditransitive construction represents a continuous pattern encounter a serious problem with the examples discussed above. This can be best understood by considering the TAG analysis of coordination proposed by Sarkar & Joshi (1996). If one assumes that [Sbj TransVerb Obj] or [S [V O]] constitutes a fixed unit, then the trees in Figure 21.12 form the starting point for the analysis of coordination.

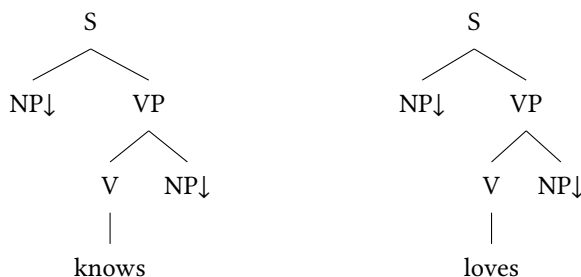


图 21.12: Elementary trees for *knows* and *loves*

If one wants to use these trees/constructions for the analysis of (87), there are in principle two possibilities: one assumes that two complete sentences are coordinated or alternatively, one assumes that some nodes are shared in a coordinated structure.

(87) He knows and loves this record.

Abeillé (2006) has shown that it is not possible to capture all the data if one assumes that cases of coordination such as those in (87) always involve the coordination of two complete clauses. It is also necessary to allow for lexical coordination of the kind we saw in Steedman's analysis (see also Section 4.6.3). Sarkar & Joshi (1996) develop a TAG analysis in which nodes are shared in coordinate structures. The analysis of (87) can be seen in Figure 21.13 在下一页. The subject and object nodes are only present once in this figure. The S nodes of both elementary trees both dominate the *he* NP. In the same way, the object NP node belongs to both VPs. The conjunction connects two verbs indicated by the thick lines. Sarkar and Joshi provide an algorithm that determines which nodes are to be shared. The structure may look strange at first, but for TAG purposes, it is not the derived tree but rather the derivation tree that is important, since this is the one that is used to compute the semantic interpretation. The authors show that the derivation trees for the example under discussion and even more complex examples can be constructed correctly.

In theories such as HPSG and LFG where structure building is, as in Categorical Grammar, driven by valence, the above sentence is unproblematic: both verbs are conjoined and then the combination behaves like a simple verb. The analysis of this is

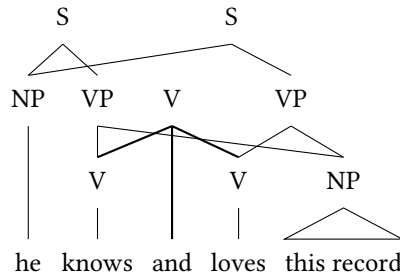


图 21.13: TAG analysis of *He knows and loves this record.*

given in Figure 21.14. This analysis is similar to the Categorical Grammar analysis in Figure 21.10.⁴⁵ With Goldberg's plugging analysis one could also adopt this approach to

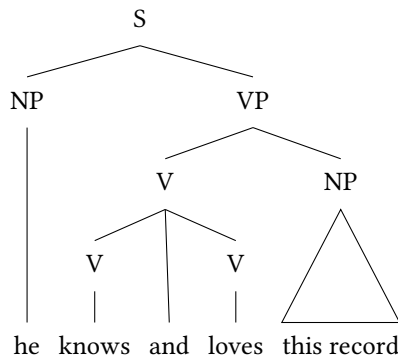


图 21.14: Selection-based analysis of *He knows and loves this record.* in tree notation

coordination: here, *knows* and *loves* would first be plugged into a coordination construction and the result would then be plugged into the transitive construction. Exactly how the semantics of *knows and loves* is combined with that of the transitive construction is unclear since the meaning of this phrase is something like $\text{and}'(\text{know}'(x, y), \text{love}'(x, y))$, that is, a complex event with at least two open argument slots x and y (and possibly additionally an event and a world variable depending on the semantic theory that is used). Goldberg would probably have to adopt an analysis such as the one in Figure 21.13 in order to maintain the plugging analysis.

⁴⁵ A parallel analysis in Dependency Grammar is possible as well. Tesnière's original analysis was different though. See Section 11.6.2.1 for discussion.

Croft would definitely have to adopt the TAG analysis since the verb is already present in his constructions. For the example in (84), both Goldberg and Croft would have to draw from the TAG analysis in Figure 21.15.

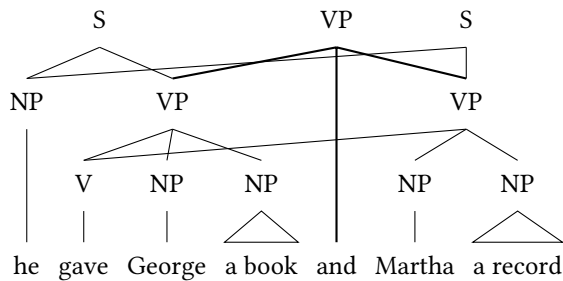


图 21.15: TAG analysis of *He gave George a book and Martha a record.*

The consequence of this is that one requires discontinuous constituents. Since coordination allows a considerable number of variants, there can be gaps between all arguments of constructions. An example with a ditransitive verb is given in (88):

(88) He gave George and sent Martha a record.

See Crysmann (2008) and Beavers & Sag (2004) for HPSG analyses that assume discontinuous constituents for particular coordination structures.

The result of these considerations is that the argument that particular elements occur next to each other and that this occurrence is associated with a particular meaning is considerably weakened. What competent speakers do acquire is the knowledge that heads must occur with their arguments somewhere in the utterance and that all the requirements of the heads involved have to somehow be satisfied (θ -Criterion, coherence/completeness, empty SUBCAT list). The heads themselves need not necessarily occur directly adjacent to their arguments. See the discussion in Section 16.3 about pattern-based models of language acquisition.

The computation of the semantic contribution of complex structures such as those in Figure 21.15 is by no means trivial. In TAG, there is the derivation tree in addition to the derived tree that can then be used to compute the semantic contribution of a linguistic object. Construction Grammar does not have this separate level of representation. The question of how the meaning of the sentences discussed here is derived from their component parts still remains open for phrasal approaches.

Concluding the section on language acquisition, we assume that a valence representation is the result of language acquisition, since this is necessary for establishing the dependency relations in various possible configurations in an utterance. See also Behrens (2009: 439) for a similar conclusion.

21.7 Arguments from psycho- and neurolinguistics

This section has three parts: in the first part we compare approaches which assume that valence alternations are modeled by lexical rules, underspecification, or disjunctions with phrasal approaches. In Subsection 21.7.2 part we discuss approaches to light verb constructions and Subsection 21.7.3 is devoted to neurolinguistic findings.

21.7.1 Lexical rules vs. phrasal constructions

Goldberg (1995: Section 1.4.5) uses evidence from psycholinguistic experiments to argue against lexical approaches that use lexical rules to account for argument structure alternations: Carlson & Tanenhaus (1988) showed that sentences with true lexical ambiguity like those in (89) and sentences with two verbs with the same core meaning have different processing times.

- (89) a. Bill set the alarm clock onto the shelf.
 b. Bill set the alarm clock for six.
- (90) a. Bill loaded the truck onto the ship.
 b. Bill loaded the truck with bricks.

Errors due to lexical ambiguity cause a bigger increase in processing time than errors in the use of the same verb. Experiments showed that there was a bigger difference in processing times for the sentences in (89) than for the sentences in (90). The difference in processing times between (90a) and (90b) would be explained by different preferences for phrasal constructions. In a lexicon-based approach one could explain the difference by assuming that one lexical item is more basic, that is, stored in the mental dictionary and the other is derived from the stored one. The application of lexical rules would be time consuming, but since the lexical items are related, the overall time consumption is smaller than the time needed to process two unrelated items (Müller 2002a: 405).

Alternatively one could assume that the lexical items for both valence patterns are the result of lexical rule applications. As with the phrasal constructions, the lexical rules would have different preferences. This shows that the lexical approach can explain the experimental results as well, so that they do not force us to prefer phrasal approaches.

Goldberg (1995: 18) claims that lexical approaches have to assume two variants of *load* with different meaning and that this would predict that *load* alternations would behave like two verbs that really have absolutely different meanings. The experiments discussed above show that such predictions are wrong and hence lexical analyses would be falsified. However, as was shown in Müller (2010a: Section 11.11.8.2), the argumentation contains two flaws: let's assume that the construction meaning of the construction that licenses (90a) is C_1 and the construction meaning of the construction that licenses (90b) is C_2 . Under such assumptions the semantic contribution of the two lexical items in the lexical analysis would be (91). *load*(...) is the contribution of the verb that would be assumed in phrasal analyses.

- (91) a. load (onto): $C_1 \wedge \text{load}(\dots)$
 b. load (with): $C_2 \wedge \text{load}(\dots)$

(91) shows that the lexical items partly share their semantic contribution. We hence predict that the processing of the dispreferred argument realization of *load* is simpler than the dispreferred meaning of *set*: in the latter case a completely new verb has to be activated while in the first case parts of the meaning are activated already.⁴⁶

Goldberg (1995: 107) argues against lexical rule-based approaches for locative alternations like (92), since according to her such approaches have to assume that one of the verb forms has to be the more basic form.

- (92) a. He loaded hay onto the wagon.
 b. He loaded the wagon with hay.

She remarks that this is problematic since we do not have clear intuitions about what the basic and what the derived forms are. She argues that the advantage of phrasal approaches is that various constructions can be related to each other without requiring the assumption that one of the constructions is more basic than the other. There are two phrasal patterns and the verb is used in one of the two patterns. This criticism can be addressed in two ways: first one could introduce two lexical types (for instance *onto-verb* and *with-verb*) into a type hierarchy. The two types correspond to two valence frames that are needed for the analysis of (92a) and (92b). These types can have a common supertype (*onto-with-verb*) which is relevant for all *spray/load* verbs. One of the subtypes or the respective lexical item of the verb is the preferred one. This corresponds to a disjunction in the lexicon, while the phrasal approach assumes a disjunction in the set of phrasal constructions.

A variant of this approach is to assume that the lexical description of *load* just contains the supertype describing all *spray/load* verbs. Since model theoretic approaches assume that all structures that are models of utterances contain only maximally specific types (see for instance King (1999) and Pollard & Sag (1994: 21)), it is sufficient to say about verbs like *load* that they are of type *onto-with-verb*. As this type has exactly two subtypes, *load* has to be either *onto-verb* or *with-verb* in an actual model.⁴⁷

A second option is to stick with lexical rules and to assume a single representation for the root of a verb that is listed in the lexicon. In addition, one assumes two lexical rules that map this basic lexical item onto other items that can be used in syntax after being inflected. The two lexical rules can be described by types that are part of a type hierarchy and that have a common supertype. This would capture commonalities between the lexical rules. We therefore have the same situation as with phrasal constructions (two lexical rules vs. two phrasal constructions). The only difference is that the action is one level deeper in the lexical approach, namely in the lexicon (Müller 2002a: 405–406).

⁴⁶ See also Croft (2003: 64–65) for a brief rejection of Goldberg's interpretation of the experiment that corresponds to what is said here.

⁴⁷ This analysis does not allow the specification of verb specific preferences for one of the realization patterns since the lexicon contains the general type only.

The argumentation with regard to the processing of resultative constructions like (93c) is parallel:

- (93) a. He drinks.
 b. He drinks the milk.
 c. He drinks the pub empty.

When humans parse a sentence they build up structure incrementally. If one hears a word that is incompatible with the current hypothesis, the parsing process breaks down or the current hypothesis is revised. In (93c) *the pub* does not correspond to the normal transitive use of *drink*, so the respective hypothesis has to be revised. In the phrasal approach the resultative construction would have to be used instead of the transitive construction. In the lexical analysis the lexical item that is licensed by the resultative lexical rule would have to be used rather than the bivalent one. Building syntactic structure and lexicon access in general place different demands on our processing capacities. However, when (93c) is parsed, the lexical items for *drink* are active already, we only have to use a different one. It is currently unclear to us whether psycholinguistic experiments can differentiate between the two approaches, but it seems to be unlikely.

21.7.2 Light verbs

Wittenberg, Jackendoff, Kuperberg, Paczynski, Snedeker & Wiese (2014) report on a number of experiments that test predictions made by various approaches to light verb constructions. (94a) shows a typical light verb construction: *take* is a light verb that is combined with the nominal that provides the main predication.

- (94) a. take a walk to the park
 b. walk to the park

Wittenberg & Piñango (2011) examined two psychologically plausible theories of light verb constructions. The phrasal approach assumes that light verb constructions are stored objects associated with semantics (Goldberg 2003b). The alternative compositional view assumes that the semantics is computed as a fusion of the semantics of the event noun and the semantics of the light verb (Grimshaw 1997; Butt 2003; Jackendoff 2002; Culicover & Jackendoff 2005; Müller 2010b; Beavers et al. 2008). Since light verb constructions are extremely frequent (Piñango, Mack & Jackendoff 2006; Wittenberg & Piñango 2011: 399), the phrasal approaches assuming that light verb constructions are stored items with the object and verb fixed predict that light verb constructions should be retrievable faster than non-light verb constructions like (95) (Wittenberg & Piñango 2011: 396).

- (95) take a frisbee to the park

This is not the case. As Wittenberg and Piñango found, there is no difference in processing at the licensing condition (the noun in VO languages like English and the verb in OV languages like German).

However, Wittenberg & Piñango (2011) found an increased processing load 300ms *after* the light verb construction is processed. The authors explain this by assuming that semantic integration of the noun with the verbal meaning takes place after the syntactic combination. While the syntactic combination is rather fast, the semantic computation takes additional resources and this is measurable at 300ms. The verb contributes aspectual information and integrates the meaning of the nominal element. The semantic roles are fused. The resource consumption effect would not be expected if the complete light verb construction were a stored item that is retrieved together with the complete meaning (p. 404). We can conclude that Wittenberg and Piñango's results are compatible with the lexical proposal, but are incompatible with the phrasal view.

21.7.3 Arguments from neurolinguistics

Pulvermüller, Cappelle & Shtyrov (2013) discuss neurolinguistic facts and relate them to the CxG view of grammar theory. One important finding is that deviant words (lexical items) cause brain responses that differ in polarity from brain responses on incorrect strings of words, that is, syntactic combinations. This suggests that there is indeed an empirical basis for deciding the issue.

Concerning the standard example of the Caused-Motion Construction in (96) the authors write the following:

(96) She sneezed the foam off the cappuccino.⁴⁸

this constellation of brain activities may initially lead to the co-activation of the verb *sneeze* with the DCNAs for *blow* and thus to the sentence mentioned. Ultimately, such co-activation of a one-place verb and DCNAs associated with other verbs may result in the former one-place verb being subsumed into a three-place verb category and DCNA set, a process which arguably has been accomplished for the verb *laugh* as used in the sequence *laugh NP off the stage*. (Pulvermüller, Cappelle & Shtyrov 2013)

A DCNA is a discrete combinatorial neuronal assembly. Regarding the specifics of DCNAs the authors write that

Apart from linking categories together, typical DCNAs establish a temporal order between the category members they bind to. DCNAs that do not impose temporal order (thus acting, in principle, as AND units for two constituents) are thought to join together constituents whose sequential order is free or allow for scrambling. (Pulvermüller, Cappelle & Shtyrov 2013: 404)

I believe that this view is entirely compatible with the lexical view outlined above: the lexical item or DCNA requires certain arguments to be present. A lexical rule that relates an intransitive verb to one that can be used in the Caused-Motion Construction is an explicit representation of what it means to activate the valence frame of *blow*.

⁴⁸ Goldberg (2006: 42).

The authors cite earlier work (Cappelle, Shtyrov & Pulvermüller 2010) and argue that particle verbs are lexical objects, admitting for a discontinuous realization despite their lexical status (p. 21). They restrict their claim to frequently occurring particle verbs. This claim is of course compatible with our assumptions here, but the differences in brain behavior are interesting when it comes to fully productive uses of particle verbs. For instance any semantically appropriate monovalent verb in German can be combined with the aspectual particle *los*: *lostanzen* ‘start to dance’, *loslachen* ‘start to laugh’, *lossingen* ‘start to sing’, Similarly, the combination of monovalent verbs with the particle *an* with the reading *directed-towards* is also productive: *anfahren* ‘drive towards’, *anlachen* ‘laugh in the direction of’, *ansegeln* ‘sail towards’, ... (see Stiebels (1996) on various productive patterns). The interesting question is how particle verbs behave that follow these patterns but occur with low frequency. This is still an open question as far as the experimental evidence is concerned, but as I argue below lexical proposals to particle verbs as the one suggested by Müller (2003d) are compatible with both possible outcomes.

Summarizing the discussion so far, lexical approaches are compatible with the accumulated neurobiological evidence and as far as particle verbs are concerned they seem to be better suited than the phrasal proposals by Booij (2002: Section 2) and Blom (2005) (See Section 21.5.1 for discussion). However, in general, it remains an open question what it means to be a discontinuous lexical item. The idea of discontinuous words is pretty old (Wells 1947), but there have not been many formal accounts of this idea. Nunberg, Sag & Wasow (1994) suggest a representation in a linearization-based framework of the kind that was proposed by Reape (1994) and Kathol (1995: 244–248) and Crysmann (2002) worked out such analyses in detail. Kathol’s lexical item for *aufwachen* ‘to wake up’ is given in (97):

(97) *aufwachen* (following Kathol 1995: 246):

$$\left[\begin{array}{l} \dots | \text{HEAD } [1] \text{ verb} \\ \dots | \text{VCOMP } \langle \rangle \\ \text{DOM } \left\langle \left[\begin{array}{l} \langle \text{wachen} \rangle \\ \dots | \text{HEAD } [1] \\ \dots | \text{VCOMP } \langle [2] \rangle \end{array} \right] \right\rangle \circ \left\langle \left[\begin{array}{l} \text{vc} \\ \langle \text{auf} \rangle \\ \text{SYNSEM } [2] \left[\begin{array}{l} \dots | \text{HEAD } \left[\begin{array}{l} \text{sepref} \\ \text{FLIP } - \end{array} \right] \end{array} \right] \end{array} \right] \right\rangle \end{array} \right]$$

The lexical representation contains the list-valued feature *DOM* that contains a description of the main verb and the particle (see Section 11.7.2.2 for details). The *DOM* list is a list that contains the dependents of a head. The dependents can be ordered in any order provided no linearization rule is violated (Reape 1994). The dependency between the particle and the main verb was characterized by the value of the *VCOMP* feature, which is a valence feature for the selection of arguments that form a complex predicate with their head. The shuffle operator \circ concatenates two lists without specifying an order between the elements of the two lists, that is, both *wachen*, *auf* and *auf*, *wachen* are possible. The little marking *vc* is an assignment to a topological field in the clause.

I criticized such linearization-based proposals since it is unclear how analyses that

claim that the particle is just linearized in the domain of its verb can account for sentences like (98), in which complex syntactic structures are involved (Müller 2007c). German is a V2 language and the fronting of a constituent into the position before the finite verb is usually described as some sort of nonlocal dependency; that is, even authors who favor linearization-based analyses do not assume that the initial position is filled by simple reordering of material (Kathol 2000; Müller 1999a, 2002a; Bjerre 2006).⁴⁹

- (98) a. [_{vf} [_{lmf} Den Atem] [_{vc} an]] hielt die ganze Judenheit.⁵⁰
 the breath PART held the whole Jewish.community
 ‘The whole Jewish community held their breath.’
- b. [_{vf} [_{lmf} Wieder] [_{vc} an]] treten auch die beiden Sozialdemokraten.⁵¹
 again PART kick also the two Social.Democrats
 ‘The two Social Democrats are also running for office again.’
- c. [_{vf} [_{vc} Los] [_{nf} damit]] geht es schon am 15. April.⁵²
 PART there.with went it already at.the 15 April
 ‘It already starts on April the 15th.’

The conclusion that has to be drawn from examples like (98) is that particles interact in complex ways with the syntax of sentences. This is captured by the lexical treatment that was suggested in Müller (2002a: Chapter 6) and Müller (2003d): the main verb selects for the verbal particle. By assuming that *wachen* selects for *auf*, the tight connection between verb and particle is represented.⁵³ Such a lexical analysis provides an easy way to account for fully nontransparent particle verbs like *an-fangen* ‘to begin’. However, I also argued for a lexical treatment of transparent particle verbs like *losfahren* ‘to start to drive’ and *jemanden/etwas anfahren* ‘drive directed towards somebody/something’. The analysis involves a lexical rule that licenses a verbal item selecting for an adjunct particle. The particles *an* and *los* can modify verbs and contribute arguments (in the case of *an*) and the particle semantics. This analysis can be shown to be compatible with the neuro-mechanical findings: if it is the case that even transparent particle verb combinations with low frequency are stored, then the rather general lexical rule that I

⁴⁹ Kathol (1995: Section 6.3) working in HPSG suggested such an analysis for simple sentences, but later changed his view. Wetta (2011) also working in HPSG assumes a purely linearization-based approach. Similarly Groß & Osborne (2009) working in Dependency Grammar assume that there is a simple dependency structure in simple sentences while there are special mechanisms to account for extraction out of embedded clauses. I argue against such proposals in Müller (2015b) referring to the scope of adjuncts, coordination of simple with complex sentences and Across the Board Extraction and apparent multiple frontings. See also Section 11.7.1.

⁵⁰ Lion Feuchtwanger, *Jud Süß*, p. 276, quoted from Grubačić (1965: 56).

⁵¹ taz, bremen, 24.05.2004, p. 21.

⁵² taz, 01.03.2002, p. 8.

⁵³ Cappelle et al. (2010: 197) write: “the results provide neurophysiological evidence that phrasal verbs are lexical items. Indeed, the increased activation that we found for existing phrasal verbs, as compared to infelicitous combinations, suggests that a verb and its particle together form one single lexical representation, i. e. a single lexeme, and that a unified cortical memory circuit exists for it, similar to that encoding a single word”. I believe that my analysis is compatible with this statement.

suggested in the works cited above is the generalization of the relation between a large amount of lexical particle verb items and their respective main verb. The individual particle verbs would be special instantiations that have the form of the particle specified as it is also the case for non-transparent particle verbs like *anfangen*. If it should turn out that productive combinations with particle verbs of low frequency cause syntactic reflexes in the brain, this could be explained as well: the lexical rule licenses an item that selects for an adverbial element. This selection would then be seen as parallel to the relation between the determiner and the noun in the NP *der Mut* ‘the courage’, which Cappelle et al. (2010: 191) discuss as an example of a syntactic combination. Note that this analysis is also compatible with another observation made by Shtyrov, Pihko & Pulvermüller (2005): morphological affixes also cause the lexical reflexes. In my analysis the stem of the main verb is related to another stem that selects for a particle. This stem can be combined with (derivational and inflectional) morphological affixes causing the lexical activation pattern in the brain. After this combination the verb is combined with the particle and the dependency can be either a lexical or a syntactic one, depending on the results of the experiments to be carried out. The analysis is compatible with both results.

Note that my analysis allows the principle of lexical integrity to be maintained. I therefore do not follow Cappelle, Shtyrov & Pulvermüller (2010: 198), who claim that they “provide proof that potentially separable multi-word items can nonetheless be word-like themselves, and thus against the validity of a once well-established linguistic principle, the Lexical Integrity Principle”. I agree that non-transparent particle verbs are multi-word lexemes, but the existence of multi-word lexemes does not show that syntax has access to the word-internal morphological structure. The parallel between particle verbs and clearly phrasal idioms was discussed in Müller (2002a,c) and it was concluded that idiom-status is irrelevant for the question of wordhood. Since the interaction of clearly phrasal idioms with derivational morphology as evidenced by examples like (99) did not force grammarians to give up on lexical integrity, it can be argued that particle verbs are not convincing evidence for giving up the Lexical Integrity Principle either.⁵⁴

- (99) a. Er hat ins Gras gebissen.
 he has in.the grass bit
 ‘He bit the dust.’
 b. „Heath Ledger“ kann ich nicht einmal schreiben, ohne dass mir sein
 Heath Ledger can I not even write without that me his
 ins Gras-Gebeiß wieder so wahnsinnig leid tut⁵⁵
 in.the grass.biting again so crazy sorrow does
 ‘I cannot even write “Heath Ledger” without being sad again about his biting
 the dust.’

The example in (99b) involves the discontinuous derivation with the circumfix *Ge- -e* (Lüdeling 2001: Section 3.4.3; Müller 2002a: 324–327, 372–377; Müller 2003d: Section 2.2.1, ■

⁵⁴ However, see Booij (2009) for some challenges to lexical integrity.

⁵⁵ <http://www.coffee2watch.at/egala>. 23.03.2012

Section 5.2.1). Still the parts of the idiom *ins Gras beiß*- ‘bite the dust’ are present and with them the idiomatic reading. See Sag (2007) for a lexical analysis of idioms that can explain examples like (99).

So, while I think that it is impossible to distinguish phrasal and lexical approaches for phenomena where heads are used with different valence patterns (Section 21.7.1), there seem to be ways to test whether patterns with high frequency and strong collocations should be analyzed as one fixed chunk of material with a fixed form and a fixed meaning or whether they should be analyzed compositionally.

21.8 Arguments from statistical distribution

In this section, we want to look at arguments from statistics that have been claimed to support a phrasal view. We first look at data-oriented parsing, a technique that was successfully used by Bod (2009b) to model language acquisition and then we turn to the collostructional analysis by Stefanowitsch & Gries (2009). Lastly, we argue that these distributional analyses cannot decide the question whether argument structure constructions are phrasal or lexical.

21.8.1 Unsupervised Data-Oriented Parsing

In Section 13.8.3, we saw Bod’s approach to the structuring of natural language utterances (2009b). If one assumes that language is acquired from the input without innate knowledge, the structures that Bod extracts from the distribution of words would have to be the ones that children also learn (parts of speech, meaning, and context would also have to be included). These structures would then also have to be the ones assumed in linguistic theories. Since Bod does not have enough data, he carried out experiments under the assumption of binary-branching trees and, for this reason, it is not possible to draw any conclusions from his work about whether rules license flat or binary-branching structures. There will almost certainly be interesting answers to this question in the future. What can certainly not be determined in a distribution-based analysis is the exact node in the tree where meaning is introduced. Bod (2009a: 132) claims that his approach constitutes “a testable realization of CxG” in the Goldbergerian sense, but the trees that he can construct do not help us to decide between phrasal or lexical analyses or analyses with empty heads. These alternative analyses are represented in Figure 21.16 在下一頁.⁵⁶ The first figure stands for a complex construction that contributes the meaning as a whole. The second figure corresponds to the analysis with a lexical rule and the third corresponds to the analysis with an empty head. A distributional analysis cannot decide between these theoretical proposals. Distribution is computed with reference to words; what the words actually mean is not taken into account. As such, it is only possible to say that the word *fischt* ‘fishes’ occurs in a particular utterance, however it is not possible to see if this word contains resultative semantics or not. Similarly, a distributional analysis

⁵⁶ The discussion is perhaps easier to follow if one assumes flat structures rather than binary-branching ones.

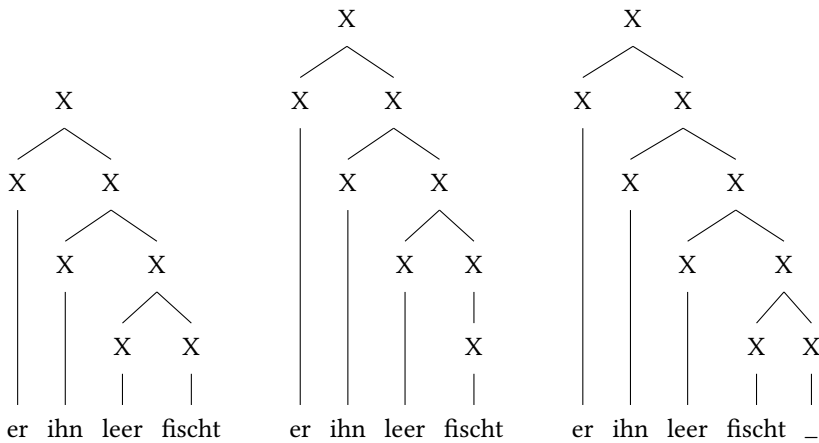
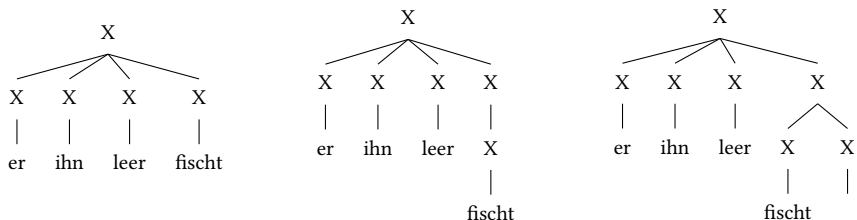


图 21.16: Three possible analyses for resultative construction: holistic construction, lexical rule, empty head

does not help to distinguish between theoretical analyses with or without a lexical head. The empty head is not perceptible in the signal. It is a theoretical construct and, as we have seen in Section 19.5, it is possible to translate an analysis using an empty head into one with a lexical rule. For the present example, any argumentation for a particular analysis will be purely theory-internal.

Although Unsupervised Data-Oriented Parsing (U-DOP) cannot help us to decide between analyses, there are areas of grammar for which these structures are of interest: under the assumption of binary-branching structures, there are different branching possibilities depending on whether one assumes an analysis with verb movement or not. This means that although one does not see an empty element in the input, there is a reflex in statistically-derived trees. The left tree in Figure 21.17 on the next page shows a structure that one would expect from an analysis following Steedman (2000: 159), see



The first figure corresponds to the Goldbergian view of phrasal constructions where the verb is inserted into the construction and the meaning is present at the topmost node. In the second figure, there is a lexical rule that provides the resultative semantics and the corresponding valence information. In the third analysis, there is an empty head that combines with the verb and has ultimately the same effect as the lexical rule.

Section 8.3. The tree on the right shows a structure that would be expected from a GB-type verb movement analysis (see Section 3.2). But at present, there is no clear finding

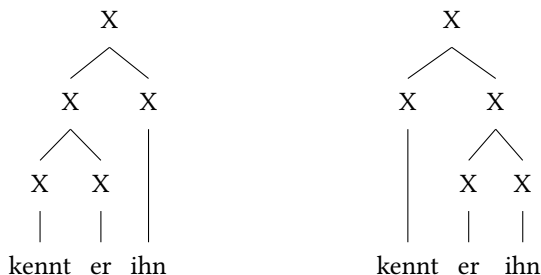


图 21.17: Structures corresponding to analyses with or without verb movement

in this regard (Bod, p. c. 2009). There is a great deal of variance in the U-DOP trees. The structure assigned to an utterance depends on the verb (Bod, referring to the Wall Street Journal). Here, it would be interesting to see if this changes with a larger data sample. In any case, it would be interesting to look at how all verbs as well as particular verb classes behave. The U-DOP procedure applies to trees containing at least one word each. If one makes use of parts of speech in addition, this results in structures that correspond to the ones we have seen in the preceding chapters. Sub-trees would then not have two Xs as their daughters but rather NP and V, for example. It is also possible to do statistic work with this kind of subtrees and use the part of speech symbols of words (the preterminal symbols) rather than the words themselves in the computation. For example, one would get trees for the symbol V instead of many trees for specific verbs. So instead of having three different trees for *küssen* ‘kiss’, *kennen* ‘know’ and *sehen* ‘see’, one would have three identical trees for the part of speech “verb” that corresponds to the trees that are needed for transitive verbs. The probability of the V tree is therefore higher than the probabilities of the trees for the individual verbs. Hence one would have a better set of data to compute structures for utterances such as those in Figure 21.17. I believe that there are further results in this area to be found in the years to come.

Concluding this subsection, we contend that Bod’s paper is a milestone in the Poverty of the Stimulus debate, but it does not and cannot show that a particular version of constructionist theories, namely the phrasal one, is correct.

21.8.2 Collostructions

Stefanowitsch & Gries (2009: Section 5) assume a plugging analysis: “words occur in (slots provided by) a given construction if their meaning matches that of the construction”. The authors claim that their *collostructional analysis has confirmed [the plugging analysis] from various perspectives*. Stefanowitsch and Gries are able to show that certain verbs occur more often than not in particular constructions, while other verbs never occur in the respective constructions. For instance, *give*, *tell*, *send*, *offer* and *show*

are attracted by the Ditransitive Construction, while *make* and *do* are repelled by this construction, that is they occur significantly less often in this construction than what would be expected given the overall frequency of verbs in the corpus. Regarding this distribution the authors write:

These results are typical for collexeme analysis in that they show two things. First, there are indeed significant associations between lexical items and grammatical structures. Second, these associations provide clear evidence for semantic coherence: the strongly attracted collexemes all involve a notion of ‘transfer’, either literally or metaphorically, which is the meaning typically posited for the ditransitive. This kind of result is typical enough to warrant a general claim that collostructional analysis can in fact be used to identify the meaning of a grammatical construction in the first place. (Stefanowitsch & Gries 2009: 943)

We hope that the preceding discussion has made clear that the distribution of words in a corpus cannot be seen as evidence for a phrasal analysis. The corpus study shows that *give* usually is used with three arguments in a certain pattern that is typical for English (Subject Verb Object1 Object2) and that this verb forms a cluster with other verbs that have a transfer component in their meaning. The corpus data do not show whether this meaning is contributed by a phrasal pattern or by lexical entries that are used in a certain configuration.

21.9 Conclusion

The essence of the lexical view is that a verb is stored with a valence structure indicating how it combines semantically and syntactically with its dependents. Crucially, that structure is abstracted from the actual syntactic context of particular tokens of the verb. Once abstracted, that valence structure can meet other fates besides licensing the phrasal structure that it most directly encodes: it can undergo lexical rules that manipulate that structure in systematic ways; it can be composed with the valence structure of another predicate; it can be coordinated with similar verbs; and so on. Such an abstraction allows for simple explanations of a wide range of robust, complex linguistic phenomena. We have surveyed the arguments against the lexical valence approach and in favor of a phrasal representation instead. We find the case for a phrasal representation of argument structure to be unconvincing: there are no compelling arguments in favor of such approaches, and they introduce a number of problems:

- They offer no account for the interaction of valence changing processes and derivational morphology.
- They offer no account for the interaction of valence changing processes and coordination of words.
- They offer no account for the iteration of valence changing processes.

- They overgenerate, unless a link between lexical items and phrasal constructions is assumed.
- They offer no account for distribution of arguments in partial fronting examples.

Assuming a lexical valence structure allows us to solve all the problems that arise with phrasal approaches.

21.10 Why (phrasal) constructions?

In previous sections, I have argued against assuming too much phrasality in grammatical descriptions. If one wishes to avoid transformations in order to derive alternative patterns from a single base structure, while still maintaining lexical integrity, then phrasal analyses become untenable for analyzing all those phenomena where changes in valence and derivational morphology interact. There are, however, some areas in which these two do not interact. In these cases, there is mostly a choice between analyses with silent heads and those with phrasal constructions. In this section, I will discuss some of these cases.

21.10.1 Verbless directives

Jacobs (2008) showed that there are linguistic phenomena where it does not make sense to assume that there is a head in a particular group of words. These configurations are best described as phrasal constructions, in which the adjacency of particular constituents leads to a complete meaning that goes beyond the sum of its parts. Examples of the phenomena that are discussed by Jacobs are phrasal templates such as those in (100) and verbless directives as in (105):

- (100) Pro_{+w,caus/purp} NP
- Wozu Konstruktionen?
why constructions
'Why constructions?'
 - Warum ich?
why I
'Why me?'
- (101) NP_{acc} Y_{PP/A/Adv}
- Den Hut in der Hand (kam er ins Zimmer).
the hat in the hand came he into.the room
'(He came into the room) hat in hand.'

In (100), we are dealing with abbreviated questions:

- (102) a. Wozu braucht man Konstruktionen? / Wozu sollte man Konstruktionen
 to.what needs one constructions to.what should one constructions
 annehmen?
 assume
 ‘Why do we need constructions?’ / ‘Why should we assume constructions?’
- b. Warum soll ich das machen? / Warum wurde ich ausgewählt? / Warum
 why should I that do why was I chosen why
 passiert mir sowas?
 happens me something.like.that
 ‘Why should I do that?’ / ‘Why was I chosen?’ / ‘Why do things like that
 happen to me?’

In (101), a participle has been omitted:

- (103) Den Hut in der Hand haltend kam er ins Zimmer.
 the hat.ACC in the hand holding came he in.the room
 ‘He came into the room hat in hand.’

Cases such as (101) can be analyzed with an empty head that corresponds to *haltend* ‘holding’. For (100), on the other hand, one would require either a syntactic structure with multiple empty elements, or an empty head that selects both parts of the construction and contributes the components of meaning that are present in (102). If one adopts the first approach with multiple silent elements, then one would have to explain why these elements cannot occur in other constructions. For example, it would be necessary to assume an empty element corresponding to *man* ‘one’/‘you’. But such an empty element could never occur in embedded clauses since subjects cannot simply be omitted there:

- (104) * weil dieses Buch gerne liest
 because this book gladly reads
 Intended: ‘because he/she/it likes to read this book’

If one were to follow the second approach, one would be forced to assume an empty head with particularly odd semantics.

The directives in (105) and (106) are similarly problematic (see also Jackendoff & Pinker (2005: 220) for parallel examples in English):

- (105) a. Her mit dem Geld / dem gestohlenen Geld!
 here with the money the stolen money
 ‘Hand over the (stolen) money!’
- b. Weg mit dem Krempel / dem alten Krempel!
 away with the junk the old junk
 ‘Get rid of this (old) junk!’

- c. Nieder mit den Studiengebühren / den sozialfeindlichen Studiengebühren!
 down with the tuition.fees the antisocial tuition.fees

‘Down with (antisocial) tuition fees!’

- (106) a. In den Müll mit diesen Klamotten!
 in the trash with these clothes
 ‘Throw these clothes into the trash!’
 b. Zur Hölle mit dieser Regierung!
 to.the hell with this government
 ‘To hell with this government!’

Here, it is also not possible to simply identify an elided verb. It is, of course, possible to assume an empty head that selects an adverb or a *mit*-PP, but this would be *ad hoc*. Alternatively, it would be possible to assume that adverbs in (105) select the *mit*-PP. Here, one would have to disregard the fact that adverbs do not normally take any arguments. The same is true of Jacobs’s examples in (106). For these, one would have to assume that *in* and *zur* ‘to the’ are the respective heads. Each of the prepositions would then have to select a noun phrase and a *mit*-PP. While this is technically possible, it is as unattractive as the multiple lexical entries that Categorical Grammar has to assume for pied-piping constructions (see Section 8.6).

A considerably more complicated analysis has been proposed by G. Müller (2009a). Müller treats verbless directives as antipassive constructions. Antipassive constructions involve either the complete suppression of the direct object or its realization as an oblique element (PP). There can also be morphological marking on the verb. The subject is normally not affected by the antipassive but can, however, receive a different case in ergative case systems due to changes in the realization of the object. According to G. Müller, there is a relation between (107a) and (107b) that is similar to active-passive pairs:

- (107) a. [dass] jemand diese Klamotten in den Müll schmeißt
 that somebody these clothes in the trash throws
 ‘that somebody throws these clothes into the trash’
 b. In den Müll mit diesen Klamotten!
 in the trash with these clothes
 ‘Throw these clothes into the trash!’

An empty passive morpheme absorbs the capability of the verb to assign accusative (see also Section 3.4 on the analysis of the passive in GB theory). The object therefore has to be realized as a PP or not at all. It follows from Burzio’s Generalization that as the accusative object has been suppressed, there cannot be an external argument. G. Müller assumes, like proponents of Distributed Morphology (e.g., Marantz 1997), that lexical entries are inserted into complete trees post-syntactically. The antipassive morpheme creates a feature bundle in the relevant tree node that is not compatible with German verbs such as *schmeißen* ‘throw’ and this is why only a null verb with the corresponding specifications

can be inserted. Movement of the directional PP is triggered by mechanisms that cannot be discussed further here. The antipassive morpheme forces an obligatory reordering of the verb in initial position (to C, see Section 3.2 and Section 4.2). By stipulation, filling the prefield is only possible in sentences where the C position is filled by a visible verb and this is why G. Müller's analysis does only derive V1 clauses. These are interpreted as imperatives or polar questions. Figure 21.18 gives the analysis of (107b). Budde (2010)

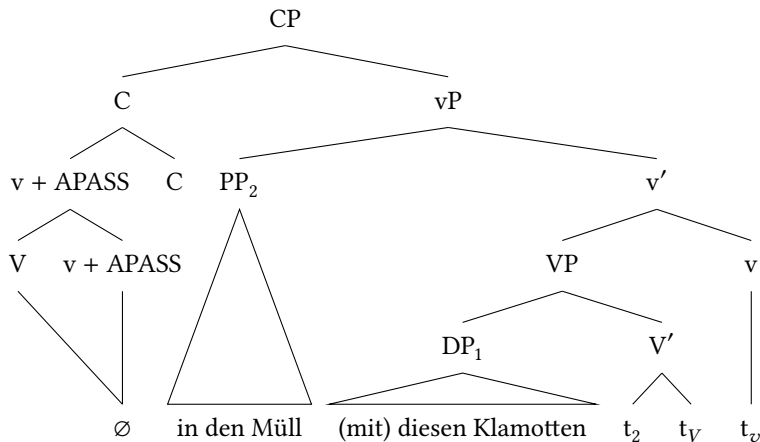


图 21.18: *In den Müll mit diesen Klamotten* ‘in the trash with these clothes’ as an antipassive following G. Müller (2009a)

and Maché (2010) note that the discussion of the data has neglected the fact that there are also interrogative variants of the construction:

- (108) a. Wohin mit den Klamotten?
 where.to with the clothes
 ‘Where should the clothes go?’
 b. Wohin mit dem ganzen Geld?
 where.to with the entire money
 ‘Where should all this money go?’

Since these questions correspond to V2 sentences, one does not require the constraint that the prefield can only be filled if the C position is filled.

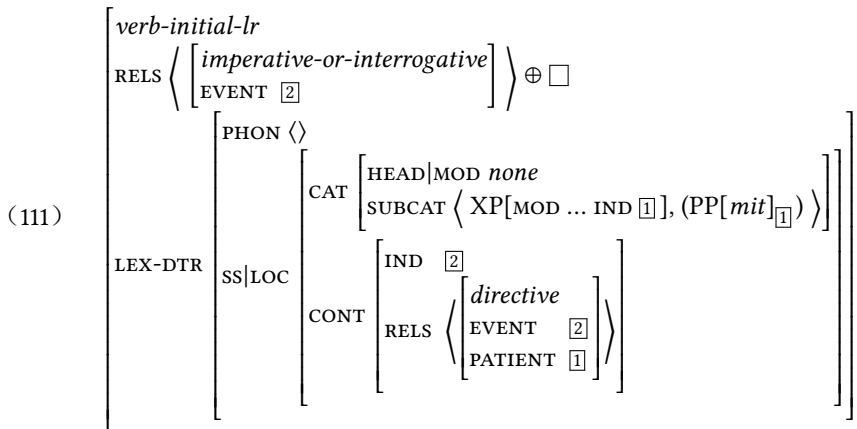
One major advantage of this analysis is that it derives the different sentence types that are possible with this kind of construction: the V1-variants correspond to polar questions and imperatives, and the V2-variants with a question word correspond to *wh*-questions. A further consequence of the approach pointed out by G. Müller is that no further explanation is required for other interactions with the grammar. For example, the way in which the constructions interact with adverbs follows from the analysis:

- (109) a. Schmeiß den Krempel weg!
 throw the junk away
 b. Schmeiß den Krempel schnell weg!
 throw the junk quickly away
 c. ?* Schmeiß den Krempel sorgfältig weg!
 throw the junk carefully away
- (110) a. Weg mit dem Krempel!
 away with the junk
 b. Schnell weg mit dem Krempel!
 quickly away with the junk
 c. ?* Sorgfältig weg mit dem Krempel!
 carefully away with the junk

Nevertheless one should still bear the price of this analysis in mind: it assumes an empty antipassive morpheme that is otherwise not needed in German. It would only be used in constructions of the kind discussed here. This morpheme is not compatible with any verb and it also triggers obligatory verb movement, which is something that is not known from any other morpheme used to form verb diatheses.

The costs of this analysis are, of course, less severe if one assumes that humans already have this antipassive morpheme anyway, that is, this morpheme is part of our innate Universal Grammar. But if one follows the argumentation from the earlier sections of this chapter, then one should only assume innate linguistic knowledge if there is no alternative explanation.

G. Müller's analysis can be translated into HPSG. The result is given in (111):



(111) contains a lexical entry for an empty verb in verb-initial position. *directive'* is a placeholder for a more general relation that should be viewed as a supertype of all possible meanings of this construction. These subsume both *schmeißen* 'to throw' and cases such as (112) that were pointed out to me by Monika Budde:

- (112) Und mit dem Klavier ganz langsam durch die Tür!
 and with the piano very slowly through the door
 ‘Carry the piano very slowly through the door!’

Since only verb-initial and verb-second orders are possible in this construction, the application of the lexical rule for verb-initial position (see page 264) is obligatory. This can be achieved by writing the result of the application of this lexical rule into the lexicon, without having the object to which the rule should have applied actually being present in the lexicon itself. Koenig (1999: Section 3.4.2, 5.3) proposed something similar for English *rumored* ‘it is rumored that ...’ and *aggressive*. There is no active variant of the verb *rumored*, a fact that can be captured by the assumption that only the result of applying a passive lexical rule is present in the lexicon. The actual verb or verb stem from which the participle form has been derived exists only as the daughter of a lexical rule but not as an independent linguistic object. Similarly, the verb **aggress* only exists as the daughter of a (non-productive) adjective rule that licenses *aggressive* and a nominalization rule licensing *aggression*.

The optionality of the *mit*-PP is signaled by the brackets in (111). If one adds the information inherited from the type *verb-initial-lr* under *SYNSEM*, then the result is (113).

$$\begin{array}{c}
 \text{(113)} \quad \left[\begin{array}{c} \text{verb-initial-lr} \\ \text{SYNSEM|LOC} \left[\begin{array}{c} \text{HEAD} \left[\begin{array}{c} \text{verb} \\ \text{VFORM } \textit{fin} \\ \text{INITIAL } + \\ \text{DSL } \textit{none} \end{array} \right] \\ \text{SUBCAT} \left\langle \left[\text{LOC|CAT} \left[\begin{array}{c} \text{HEAD} \left[\begin{array}{c} \text{verb} \\ \text{DSL } \boxed{3} \end{array} \right] \right] \right] \right\rangle \end{array} \right] \\ \text{RELS} \left\langle \left[\begin{array}{c} \textit{imperative-or-interrogative} \\ \text{EVENT } \boxed{2} \end{array} \right] \right\rangle \oplus \boxed{4} \end{array} \right] \\ \left[\begin{array}{c} \text{LEX-DTR} \\ \text{SS|LOC } \boxed{3} \left[\begin{array}{c} \text{PHON } \langle \rangle \\ \text{CAT} \left[\begin{array}{c} \text{HEAD|MOD } \textit{none} \\ \text{SUBCAT } \left\langle \text{XP[MOD ... IND } \boxed{1} \text{], (PP[mit] } \boxed{1} \text{)} \right\rangle \right] \\ \text{CONT} \left[\begin{array}{c} \text{IND } \boxed{2} \\ \text{RELS } \boxed{4} \left\langle \left[\begin{array}{c} \textit{directive} \\ \text{EVENT } \boxed{2} \\ \text{PATIENT } \boxed{1} \end{array} \right] \right\rangle \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right]
 \end{array}$$

The valence properties of the empty verb in (113) are to a large extent determined by the lexical rule for verb-initial order: the V1-LR licenses a verbal head that requires a VP to its right that is missing a verb with the local properties of the LEX-DTR ($\boxed{3}$).

Semantic information dependent on sentence type (assertion, imperative or question) is determined inside the V1-LR depending on the morphological make-up of the verb and the SLASH value of the selected VP (see Müller 2007b: Section 10.3; 2016b; 2015b). Setting the semantics to *imperative-or-interrogative* rules out *assertion* as it occurs in V2-clauses. Whether this type is resolved in the direction of *imperative* or *interrogative* is ultimately decided by further properties of the utterance such as intonation or the use of interrogative pronouns.

The valence of the lexical daughters in (113) as well as the connection to the semantic role (the linking to the patient role) are simply stipulated. Every approach has to stipulate that an argument of the verb has to be expressed as a *mit*-PP. Since there is no antipassive in German, the effect that could be otherwise achieved by an antipassive lexical rule in (113) is simply written into the LEX-DTR of the verb movement rule.

The SUBCAT list of LEX-DTR contains a modifier (adverb, directional PP) and the *mit*-PP. This *mit*-PP is co-indexed with the patient of *directive'* and the modifier refers to the referent of the *mit*-PP. The agent of *directive'* is unspecified since it depends on the context (speaker, hearer, third person).

This analysis is shown in Figure 21.19. Here, V[LOC 2] corresponds to the LEX-DTR

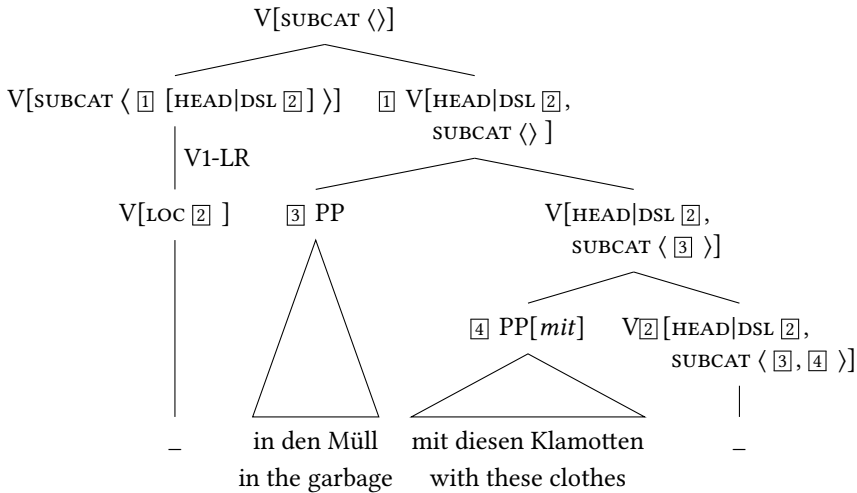


图 21.19: HPSG variant of the analysis of *In den Müll mit diesen Klamotten!/?*

in (113). The V1-LR licenses an element that requires a maximal verb projection with that exact DSL value 2. Since DSL is a head feature, the information is present along the head path. The DSL value is identified with the LOCAL value (2) in Figure 21.19) in the verb movement trace (see page 265). This ensures that the empty element at the end of sentence has exactly the same local properties that the LEX-DTR in (113) has. Thus, both the correct syntactic and semantic information is present on the verb trace and structure building involving the verb trace follows the usual principles. The structures correspond to the

structures that were assumed for German sentences in Chapter 9. Therefore, there are the usual possibilities for integrating adjuncts. The correct derivation of the semantics, in particular embedding under imperative or interrogative semantics, follows automatically (for the semantics of adjuncts in conjunction with verb position, see Müller (2007b: Section 9.4)). Also, the ordering variants with the *mit*-PP preceding the direction (112) and direction preceding the *mit*-PP (107b) follow from the usual mechanisms.

If one rejects the analyses discussed up to this point, then one is only really left with phrasal constructions or dominance schemata that connect parts of the construction and contribute the relevant semantics. Exactly how one can integrate adjuncts into the phrasal construction in a non-stipulative way remains an open question; however, there are already some initial results by Jakob Maché (2010) suggesting that directives can still be sensibly integrated into the entire grammar provided an appropriate phrasal schema is assumed.

21.10.2 Serial verbs

There are languages with so-called serial verbs. For example, it is possible to form sentences in Mandarin Chinese where there is only one subject and several verb phrases (Li & Thompson 1981: Chapter 21). There are multiple readings depending on the distribution of aspect marking inside the VP:⁵⁷ if the first VP contains a perfect marker, then we have the meaning ‘VP1 in order to do/achieve VP2’ (114a). If the second VP contains a perfect marker, then the entire construction means ‘VP2 because VP1’ (114b), and if the first VP contains a durative marker and the verb *hold* or *use*, then the entire construction means ‘VP2 using VP1’ (114c).

- (114) a. Ta1 qu3 le qian2 qu4 guang1jie1.
 he withdraw PRF money go shop
 ‘He withdrew money to go shopping.’
 b. Ta1 zhu4 Zhong1guo2 xue2 le Han4yu3.
 he live China learn PRF Chinese
 ‘He learned Chinese because he lived in China.’
 c. Ta1 na2 zhe kuai4zi chi1 fan4.
 he take DUR chopsticks eat food
 ‘He eats with chopsticks.’

If we consider the sentences, we only see two adjacent VPs. The meanings of the entire sentences, however, contain parts of meaning that go beyond the meaning of the verb phrases. Depending on the kind of aspect marking, we arrive at different interpretations with regard to the semantic combination of verb phrases. As can be seen in the translations, English sometimes uses conjunctions in order to express relations between clauses or verb phrases.

There are three possible ways to capture these data:

⁵⁷ See Müller & Lipenkova (2009) for a detailed discussion and further references.

1. One could claim that speakers of Chinese simply deduce the relation between the VPs from the context,
2. one could assume that there are empty heads in Chinese corresponding to *because* or *to*, or
3. one could assume a phrasal construction for serial verbs that contributes the correct semantics for the complete meaning depending on the aspect marking inside the VPs.

The first approach is unsatisfactory because the meaning does not vary arbitrarily. There are grammaticalized conventions that should be captured by a theory. The second solution has a stipulative character and thus, if one wishes to avoid empty elements, only the third solution remains. Müller & Lipenkova (2009) have presented a corresponding analysis.

21.10.3 Relative and interrogative clauses

Sag (1997) develops a phrasal analysis of English relative clauses as have Ginzburg & Sag (2000) for interrogative clauses. Relative and interrogative clauses consist of a fronted phrase and a clause or a verb phrase missing the fronted phrase. The fronted phrase contains a relative or interrogative pronoun.

- (115) a. the man [who] sleeps
 b. the man [who] we know
 c. the man [whose mother] visited Kim
 d. a house [in which] to live
- (116) a. I wonder [who] you know.
 b. I want to know [why] you did this.

The GB analysis of relative clauses is given in Figure 21.20. In this analysis, an empty head is in the C position and an element from the IP is moved to the specifier position.

The alternative analysis shown in Figure 21.21 在下一页 involves combining the subparts directly in order to form a relative clause. Borsley (2006) has shown that one would require six empty heads in order to capture the various types of relative clauses possible in English if one wanted to analyze them lexically. These heads can be avoided and replaced by corresponding schemata (see Chapter 19 on empty elements). A parallel argument can also be found in Webelhuth (2011) for German: grammars of German would also have to assume six empty heads for the relevant types of relative clause.

Unlike the resultative constructions that were already discussed, there is no variability among interrogative and relative clauses with regard to the order of their parts. There are no changes in valence and no interaction with derivational morphology. Thus, nothing speaks against a phrasal analysis. If one wishes to avoid the assumption of empty heads, then one should opt for the analysis of relative clauses by Sag, or the variant in

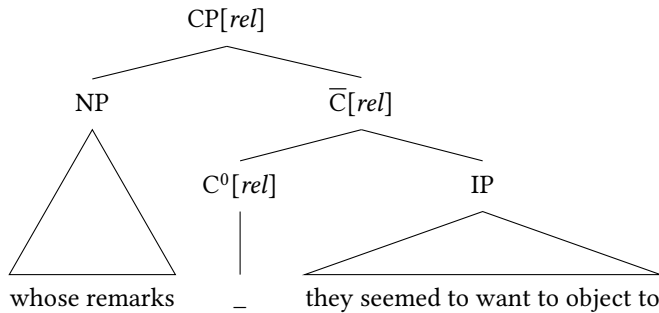


图 21.20: Analysis of relative clauses in GB theory

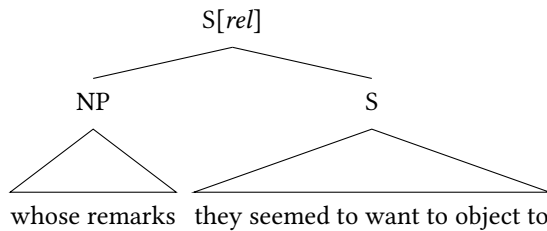


图 21.21: Analysis of relative clauses in HPSG following Sag (1997)

Müller (1999a: Chapter 10; 2007b: Chapter 11). The latter analysis does without a special schema for noun-relative clause combinations since the semantic content of the relative clause is provided by the relative clause schema.

Sag (2010) discusses long-distance dependencies in English that are subsumed under the term *wh*-movement in GB theory and the Minimalist Program. He shows that this is by no means a uniform phenomenon. He investigates *wh*-questions (117), *wh*-exclamatives (118), topicalization (119), *wh*-relative clauses (120) and *the*-clauses (121):

- (117) a. How foolish is he?
 b. I wonder *how foolish he is*.
- (118) a. What a fool he is!
 b. It's amazing *how odd it is*.
- (119) The bagels, I like.
- (120) a. I met the person *who they nominated*.
 b. I'm looking for a bank *in which to place my trust*.
- (121) a. The more people I met, *the happier I became*.

- b. *The more people I met*, the happier I became.

These individual constructions vary in many respects. Sag lists the following questions that have to be answered for each construction:

- Is there a special *wh*-element in the filler daughter and, if so, what kind of element is it?
- Which syntactic categories can the filler daughters have?
- Can the head-daughter be inverted or finite? Is this obligatory?
- What is the semantic and/or syntactic category of the mother node?
- What is the semantic and/or syntactic category of the head-daughter?
- Is the sentence an island? Does it have to be an independent clause?

The variation that exists in this domain has to be captured somehow by a theory of grammar. Sag develops an analysis with multiple schemata that ensure that the category and semantic contribution of the mother node correspond to the properties of both daughters. The constraints for both classes of constructions and specific constructions are represented in an inheritance hierarchy so that the similarities between the constructions can be accounted for. The analysis can of course also be formulated in a GB-style using empty heads. One would then have to find some way of capturing the generalizations pertaining to the construction. This is possible if one represents the constraints on empty heads in an inheritance hierarchy. Then, the approaches would simply be notational variants of one another. If one wishes to avoid empty elements in the grammar, then the phrasal approach would be preferable.

21.10.4 The N-P-N construction

Jackendoff (2008) discusses the English N-P-N construction. Examples of this construction are given in (122):

- (122)
- a. day by day, paragraph by paragraph, country by country
 - b. dollar for dollar, student for student, point for point
 - c. face to face, bumper to bumper
 - d. term paper after term paper, picture after picture
 - e. book upon book, argument upon argument

This construction is relatively restricted: articles and plural nouns are not allowed. The phonological content of the first noun has to correspond to that of the second. There are also similar constructions in German:

- (123)
- a. Er hat Buch um Buch verschlungen.
 he has book around book swallowed
 ‘He binge-read book after book.’

- b. Zeile für Zeile⁵⁸
 line for line
 ‘line by line’

Determining the meaning contribution of this kind of N-P-N construction is by no means trivial. Jackendoff suggests the meaning *many Xs in succession* as an approximation.

Jackendoff points out that this construction is problematic from a syntactic perspective since it is not possible to determine a head in a straightforward way. It is also not clear what the structure of the remaining material is if one is working under assumptions of \bar{X} theory. If the preposition *um* were the head in (123a), then one would expect that it is combined with an NP, however this is not possible:

- (124) a. *Er hat dieses Buch um jenes Buch verschlungen.
 he has this book around this book swallowed
 b. *Er hat ein Buch um ein Buch verschlungen.
 he has a book around a book swallowed

For this kind of structures, it would be necessary to assume that a preposition selects a noun to its right and, if it find this, it then requires a second noun of this exact form to its left. For N-*um*-N and N-*für*-N, it is not entirely clear what the entire construction has to do with the individual prepositions. One could also try to develop a lexical analysis for this phenomenon, but the facts are different to those for resultative constructions: in resultative constructions, the semantics of simplex verbs clearly plays a role. Furthermore, unlike with the resultative construction, the order of the component parts of the construction is fixed in the N-P-N construction. It is not possible to extract a noun or place the preposition in front of both nouns. Syntactically, the N-P-N combination with some prepositions behaves like an NP (Jackendoff 2008: 9):

- (125) Student after/upon/*by student flunked.

This is also strange if one wishes to view the preposition as the head of the construction.

Instead of a lexical analysis, Jackendoff proposes the following phrasal construction for N-*after*-N combinations:

- (126) Meaning: MANY X_i s IN SUCCESSION [or however it is encoded]
 Syntax: [_{NP} N_i P_j N_i]
 Phonology: Wd_i after_j Wd_i

The entire meaning as well as the fact that the N-P-N has the syntactic properties of an NP would be captured on the construction level.

I already discussed examples by Bragmann (2015) in Section 11.7.2.4 that show that N-P-N constructions may be extended by further P-N combinations:

- (127) Day after day after day went by, but I never found the courage to talk to her.

⁵⁸ *Zwölf Städte*. Einstürzende Neubauten. Fünf auf der nach oben offenen Richterskala, 1987.

So rather than an N-P-N pattern Bragmann suggests the pattern in (128), where ‘+’ stands for at least one repetition of a sequence.

(128) N (P N)+

As was pointed out on page 375 this pattern is not easy to cover in selection-based approaches. One could assume that an N takes arbitrarily many P-N combinations, which would be very unusual for heads. Alternatively, one could assume recursion, so N would be combined with a P and with an N-P-N to yield N-P-N-P-N. But such an analysis would make it really difficult to enforce the restrictions regarding the identity of the nouns in the complete construction. In order to enforce such an identity the N that is combined with N-P-N would have to impose constraints regarding deeply embedded nouns inside the embedded N-P-N object (see also Section 18.2).

G. Müller (2011a) proposes a lexical analysis of the N-P-N construction. He assumes that prepositions can have a feature REDUP. In the analysis of *Buch um Buch* ‘book after book’, the preposition is combined with the right noun *um Buch*. In the phonological component, reduplication of *Buch* is triggered by the REDUP feature, thereby yielding *Buch um Buch*. This analysis also suffers from the problems pointed out by Jackendoff: in order to derive the semantics of the construction, the semantics would have to be present in the lexical entry of the reduplicating preposition (or in a relevant subsequent component that interprets the syntax). Furthermore it is unclear how a reduplication analysis would deal with the Bragmann data.

第二十二章 Universal Grammar and doing comparative linguistics without an a priori assumption of a (strong) UG

The following two sections deal with the tools that I believe to be necessary to capture generalizations and the way one can derive such generalizations.

22.1 Formal tools for capturing generalizations

In Chapter 13, it was shown that all the evidence that has previously been brought forward in favor of innate linguistic knowledge is in fact controversial. In some cases, the facts are irrelevant to the discussion and in other cases, they could be explained in other ways. Sometimes, the chains of argumentation are not logically sound or the premises are not supported. In other cases, the argumentation is circular. As a result, the question of whether there is innate linguistic knowledge still remains unanswered. All theories that presuppose the existence of this kind of knowledge are making very strong assumptions. If one assumes, as Kayne (1994) for example, that all languages have the underlying structure [specifier [head complement]] and that movement is exclusively to the left, then these two basic assumptions must be part of innate linguistic knowledge since there is no evidence for the assumption that utterances in all natural languages have the structure that Kayne suggests. As an example, the reader may check Laenzlinger's proposal for German (2004: 224), which is depicted in Figure 4.20 on page 135. According to Laenzlinger, (1a) is derived from the underlying structure in (1b):

- (1) a. weil der Mann wahrscheinlich diese Sonate nicht oft gut gespielt
because the man probably this sonata not often well played
hat
has
'because the man probably had not played this sonata well often'
- b. *weil der Mann wahrscheinlich nicht oft gut hat gespielt diese
because the man probably not often well has played this
Sonate
sonata

(1b) is entirely unacceptable, so the respective structure cannot be acquired from input and hence the principles and rules that license it would have to be innate.

As we have seen, there are a number of alternative theories that are much more surface-oriented than most variants of Transformational Grammar. These alternative theories often differ with regard to particular assumptions that have been discussed in the preceding sections. For example, there are differences in the treatment of long-distance dependencies that have led to a proliferation of lexical items in Categorical Grammar (see Section 8.6). As has been shown by Jacobs (2008), Jackendoff (2008) and others, approaches such as Categorical Grammar that assume that every phrase must have a functor/head cannot explain certain constructions in a plausible way. Inheritance-based phrasal analyses that only list heads with a core meaning in the lexicon and have the constructions in which the heads occur determine the meaning of a complex expression turn out to have difficulties with derivational morphology and with accounting for alternative ways of argument realization (see Section 21.2.2, 21.4.1, and 21.4.2). We therefore need a theory that handles argument structure changing processes in the lexicon and still has some kind of phrase structure or relevant schemata. Some variants of GB/MP as well as LFG, HPSG, TAG and variants of CxG are examples of this kind of theory. Of these theories, only HPSG and some variants of CxG make use of the same descriptive tools ((typed) feature descriptions) for roots, stems, words, lexical rules and phrases. By using a uniform description for all these objects, it is possible to formulate generalizations over the relevant objects. It is therefore possible to capture what particular words have in common with lexical rules or phrases. For example, the *-bar* ‘able’ derivation corresponds to a complex passive construction with a modal verb. (2) illustrates.

- (2) a. Das Rätsel ist lösbar.
 the puzzle is solvable
 b. Das Rätsel kann gelöst werden.
 the puzzle can solved be
 ‘The puzzle can be solved.’

By using the same descriptive inventory for syntax and morphology, it is possible to capture cross-linguistic generalizations: something that is inflection/derivation in one language can be syntax in another.

It is possible to formulate principles that hold for both words and phrases and furthermore, it is possible to capture cross-linguistic generalizations or generalizations that hold for certain groups of languages. For example, languages can be divided into those with fixed constituent order and those with more flexible or completely free constituent order. The corresponding types can be represented with their constraints in a type hierarchy. Different languages can use a particular part of the hierarchy and also formulate different constraints for each of the types (see Ackerman & Webelhuth 1998: Section 9.2). HPSG differs from theories such as LFG and TAG in that phrases are not ontologically different from words. This means that there are no special c-structures or tree structures. Descriptions of complex phrases simply have additional features that say something about their daughters. In this way, it is possible to formulate cross-linguistic generalizations about dominance schemata. In LFG, the c-structure rules are normally specified separately for each language. Another advantage of consistent description is that one

can capture similarities between words and lexical rules, as well as between words and phrases. For example, a complementizer such as *dass* ‘that’ shares a number of properties with a simple verb or with coordinated verbs in initial position:

- (3) a. [dass] Maria die Platte kennt und liebt
 that Maria the record knows and loves
 ‘that Maria knows and loves the record’
 b. [Kennt und liebt] Maria die Platte?
 knows and loves Maria the record
 ‘Does Mary know and love the record?’

The difference between the two linguistic objects mainly lies in the kind of phrase they select: the complementizer requires a sentence with a visible finite verb, whereas the verb in initial position requires a sentence without a visible finite verb.

In Section 9.1.5, a small part of an inheritance hierarchy was presented. This part contains types that probably play a role in the grammars of all natural languages: there are head-argument combinations in every language. Without this kind of combinatorial operation, it would not be possible to establish a relation between two concepts. The ability to create relations, however, is one of the basic properties of language.

In addition to more general types, the type hierarchy of a particular language contains language-specific types or those specific to a particular class of languages. All languages presumably have one and two-place predicates and for most languages (if not all), it makes sense to talk about verbs. It is then possible to talk about one and two-place verbs. Depending on the language, these can then be subdivided into intransitive and transitive. Constraints are formulated for the various types that can either hold generally or be language-specific. In English, verbs have to occur before their complements and therefore have the INITIAL value +, whereas verbs in German have the INITIAL value – and it is the lexical rule for initial position that licenses a verb with an INITIAL value +.

The differing settings of the INITIAL value for German and English is reminiscent of parameters from GB-Theory. There is one crucial difference, however: it is not assumed that a language learner sets the INITIAL value for all heads once and for all. The use of an INITIAL value is compatible with models of acquisition that assume that learners learn individual words with their positional properties. It is certainly possible for the respective words to exhibit different values for a particular feature. Generalizations about the position of entire word classes are only learned at a later point in the acquisition process.

A hierarchy analogous to the one proposed by Croft (see Section 21.4.1) is given in Figure 22.1 on the next page. For inflected words, the relevant roots are in the lexicon. Examples of this are *schlaf*- ‘sleep’, *lieb*- ‘love’ and *geb*- ‘give’. In Figure 22.1, there are different subtypes of *root*, the general type for roots: e.g., *intrans-verb* for intransitive verbs and *trans-verb* for transitive verbs. Transitive verbs can be further subdivided into strictly transitive verbs (those with nominative and accusative arguments) and ditransitive verbs (those with nominative and both accusative and dative arguments). The

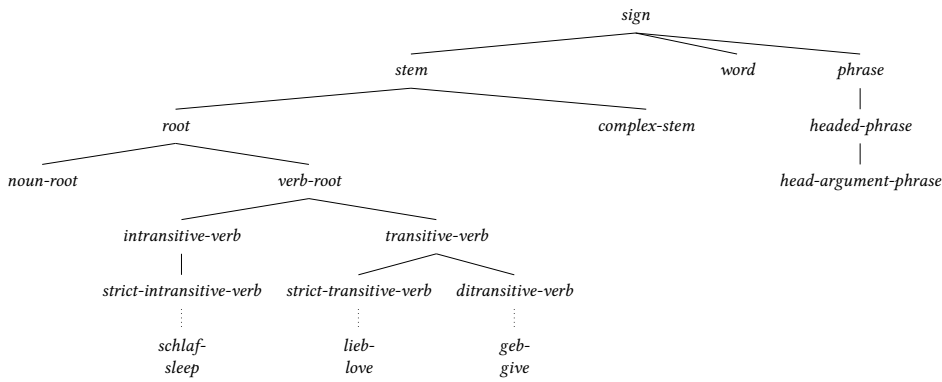


图 22.1: Section of an inheritance hierarchy with lexical entries and dominance schemata

hierarchy above would of course have to be refined considerably as there are even further sub-classes for both transitive and intransitive verbs. For example, one can divide intransitive verbs into unaccusative and unergative verbs and even strictly transitive verbs would have to be divided into further sub-classes (see Welke 2009: Section 2).

In addition to a type for roots, the above figure contains types for stems and words. Complex stems are complex objects that are derived from simple roots but still have to be inflected (*lesbar*- ‘readable’, *besing*- ‘to sing about’). Words are objects that do not inflect. Examples of these are the pronouns *er* ‘he’, *sie* ‘she’ etc. as well as prepositions. An inflected form can be formed from a verbal stem (*geliebt* ‘loved’, *besingt* ‘sings about’). Relations between inflected words and (complex) stems can be formed again using derivation rules. In this way, *geliebt* ‘loved’ can be recategorized as an adjective stem that must then be combined with adjectival endings (*geliebt-e*). The relevant descriptions of complex stems/words are subtypes of *complex-stem* or *word*. These subtypes describe the form that complex words such as *geliebte* must have. For a technical implementation of this, see Müller (2002a: Section 3.2.7). Using dominance schemata, all words can be combined to form phrases. The hierarchy given here is of course by no means complete. There are a number of additional valence classes and one could also assume more general types that simply describe one, two and three-place predicates. Such types are probably plausible for the description of other languages. Here, we are only dealing with a small part of the type hierarchy in order to have a comparison to the Croftian hierarchy: in Figure 22.1, there are no types for sentence patterns with the form [Sbj IntrVerb], but rather types for lexical objects with a particular valence (V[SUBCAT (NP[*str*])]). Lexical rules can then be applied to the relevant lexical objects that license objects with another valence or introduce information about inflection. Complete words can be combined in the syntax with relatively general rules, for example in head-argument structures. The problems from which purely phrasal approaches suffer are thereby avoided. Nevertheless generalizations about lexeme classes and the utterances that can be formed can be captured in the hierarchy.

There are also principles in addition to inheritance hierarchies: the Semantics Principle presented in Section 9.1.6 holds for all languages. The Case Principle that we also saw is a constraint that only applies to a particular class of languages, namely nominative-accusative languages. Other languages have an ergative-absolutive system.

The assumption of innate linguistic knowledge is not necessary for the theory of language sketched here. As the discussion in Section 13 has shown, the question of whether this kind of knowledge exists has still not been answered conclusively. Should it turn out that this knowledge actually exists, the question arises of what exactly is innate. It would be a plausible assumption that the part of the inheritance hierarchy that is relevant for all languages is innate together with the relevant principles (e.g., the constraints on Head-Argument structures and the Semantics Principle). It could, however, also be the case that only a part of the more generally valid types and principles is innate since something being innate does not follow from the fact that it is present in all languages (see also Section 13.1.9).

In sum, one can say that theories that describe linguistic objects using a consistent descriptive inventory and make use of inheritance hierarchies to capture generalizations are the ones best suited to represent similarities between languages. Furthermore, this kind of theory is compatible with both a positive and a negative answer to the question of whether there is innate linguistic knowledge.

22.2 How to develop linguistic theories that capture cross-linguistic generalizations

In the previous section I argued for a uniform representation of linguistic knowledge at all descriptive levels and for type hierarchies as a good tool for representing generalizations. This section explores a way to develop grammars that are motivated by facts from several languages.

If one looks at the current practice in various linguistic schools one finds two extreme ways of approaching language. On the one hand, we have the Mainstream Generative Grammar (MGG) camp and, on the other hand, we have the Construction Grammar/Cognitive Grammar camp. I hasten to say that what I state here does not hold for all members of these groups, but for the extreme cases. The caricature of the MGG scientist is that he is looking for underlying structures. Since these have to be the same for all languages (poverty of the stimulus), it is sufficient to look at one language, say English. The result of this research strategy is that one ends up with models that were suggested by the most influential linguist for English and that others then try to find ways to accommodate other languages. Since English has an NP VP structure, all languages have to have it. Since English reorders constituents in passive sentences, passive is movement and all languages have to work this way. I discussed the respective analyses of German in more detail in Section 3.4.2 and in Chapter 20 and showed that the assumption that passive is movement makes unwanted predictions for German, since the subject of passives stays in the object position in German. Furthermore, this analysis requires the assumption of invisible expletives, that is, entities that cannot be seen and do not have any meaning.

On the other extreme of the spectrum we find people working in Construction Grammar or without any framework at all (see footnote 1 on page 1 for discussion) who claim that all languages are so different that we cannot even use the same vocabulary to analyze them. Moreover, within languages, we have so many different objects that it is impossible (or too early) to state any generalizations. Again, what I describe here are extreme positions and clichés.

In what follows, I sketch the procedure that we apply in the CoreGram project¹ (Müller 2013a, 2015a). In the CoreGram project we work on a set of typologically diverse languages in parallel:

- German (Müller 2007b, 2009c, 2012; Müller & Ørsnes 2011, 2013a; Müller 2014b, 2015b)
- Danish (Ørsnes 2009b; Müller 2009c, 2012; Müller & Ørsnes 2011, 2013a,b, 2015)
- Persian (Müller 2010b; Müller & Ghayoomi 2010)
- Maltese (Müller 2009b)
- Mandarin Chinese (Lipenkova 2009; Müller & Lipenkova 2009, 2013, 2016)

¹ <http://hpsg.fu-berlin.de/Projects/CoreGram.html>, March 20, 2017.

22.2 How to develop linguistic theories that capture cross-linguistic generalizations

- Yiddish (Müller & Ørsnes 2011)
- English (Müller 2009c, 2012; Müller & Ørsnes 2013a)
- Hindi
- Spanish (Machicao y Priemer 2015)
- French

These languages belong to diverse language families (Indo-European, Afro-Asiatic, Sino-Tibetan) and among the Indo-European languages the languages belong to different groups (Germanic, Romance, Indo-Iranian). Figure 22.2 provides an overview. We work out fully formalized, computer-processable grammar fragments in the framework of HPSG that have a semantics component. The details will not be discussed here, but the interested reader is referred to Müller (2015a).

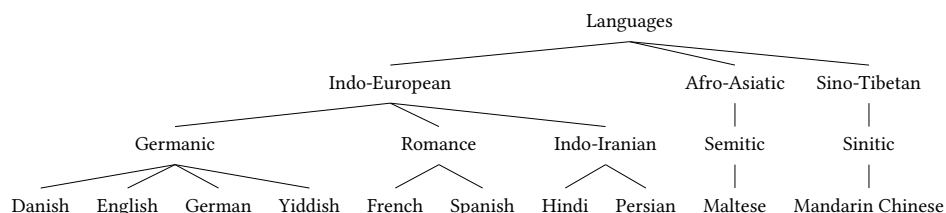


图 22.2: Language families and groups of the languages covered in the CoreGram project

As was argued in previous sections, the assumption of innate language-specific knowledge should be kept to a minimum. This is also what Chomsky suggested in his Minimalist Program. There may even be no language-specific innate knowledge at all, a view taken in Construction Grammar/Cognitive Grammar. So, instead of imposing constraints from one language onto other languages, a bottom-up approach seems to be more appropriate: grammars for individual languages should be motivated language-internally. Grammars that share certain properties can be grouped in classes. This makes it possible to capture generalizations about groups of languages and natural language as such. Let us consider a few example languages: German, Dutch, Danish, English and French. If we start developing grammars for German and Dutch, we find that they share a lot of properties: for instance, both are SOV and V2 languages and both have a verbal complex. One main difference is the order of elements in the verbal complex. The situation can be depicted as in Figure 22.3 on the following page. There are some properties that are shared between German and Dutch (Set 3). For instance, the argument structure of lexical items, a list containing descriptions of syntactic and semantic properties of arguments and the linking of these arguments to the meaning of the lexical items, is contained in Set 3. In addition to the constraints for SOV languages, the verb position and the fronting of a constituent in V2 clauses are contained in Set 3. The respective

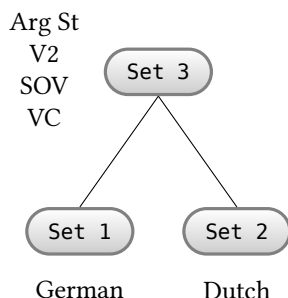


图 22.3: Shared properties of German and Dutch

constraints are shared between the two grammars. Although these sets are arranged in a hierarchy in Figure 22.3 and the following figures this has nothing to do with the type hierarchies that have been discussed in the previous subsection. These type hierarchies are part of our linguistic theories and various parts of such hierarchies can be in different sets: those parts of the type hierarchy that concern more general aspects can be in Set 3 in Figure 22.3 and those that are specific to Dutch or German are in the respective other sets. When we add another language, say Danish, we get further differences. While German and Dutch are SOV, Danish is an SVO language. Figure 22.4 在下一页 shows the resulting situation: the topmost node represents constraints that hold for all the languages considered so far (for instance the argument structure constraints, linking and V2) and the node below it (Set 4) contains constraints that hold for German and Dutch only.² For instance, Set 4 contains constraints regarding verbal complexes and SOV order. The union of Set 4 and Set 5 is Set 3 of Figure 22.3.

If we add further languages, further constraint sets will be distinguished. Figure 22.5 在下一页 shows the situation that results when we add English and French. Again, the picture is not complete since there are constraints that are shared by Danish and English but not by French, but the general idea should be clear: by systematically working this way, we should arrive at constraint sets that directly correspond to those that have been established in the typological literature.

The interesting question is what will be the topmost set if we consider enough languages. At first glance, one would expect that all languages have valence representations and linkings between these and the semantics of lexical items (argument structure lists in the HPSG framework). However, Koenig & Michelson (2012) argue for an analysis of Oneida (a Northern Iroquoian language) that does not include a representation of syntactic valence. If this analysis is correct, syntactic argument structure would not be universal. It would, of course, be characteristic of a large number of languages, but it

² In principle, there could be constraints that hold for Dutch and Danish, but not for German or for German and Danish, but not for Dutch. These constraints would be removed from Set 1 and Set 2, respectively, and inserted into another constraint set higher up in the hierarchy. These sets are not illustrated in the figure and I keep the names Set 1 and Set 2 from Figure 22.3 for the constraint sets for German and Dutch.

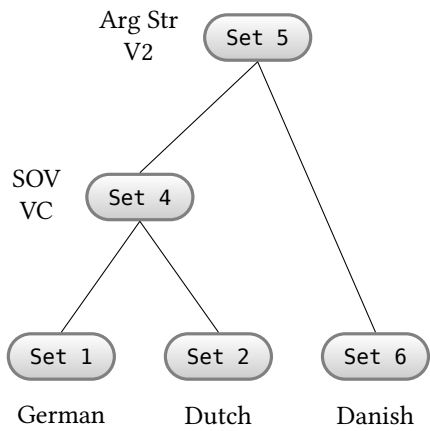


图 22.4: Shared properties of German, Dutch, and Danish

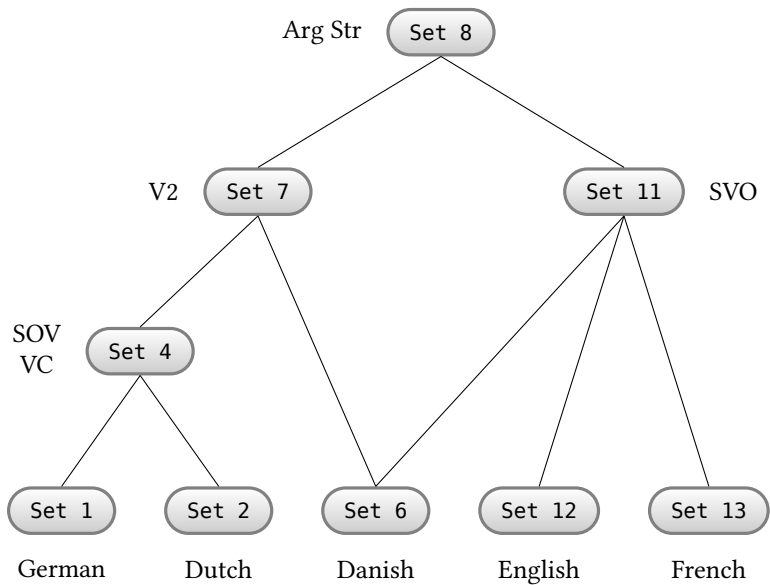


图 22.5: Languages and language classes

would not be part of the topmost set. So this leaves us with just one candidate for the topmost set from the area of syntax: the constraints that license the combination of two or more linguistic objects. This is basically Chomsky's External Merge without the binarity restriction³. In addition, the topmost set would, of course, contain the basic machinery for representing phonology and semantics.

It should be clear from what has been said so far that the goal of every scientist who works this way is to find generalizations and to describe a new language in a way that reuses theoretical constructs that have been found useful for a language that is already covered. However, as was explained above, the resulting grammars should be motivated by data of the respective languages and not by facts from other languages. In situations where more than one analysis would be compatible with a given dataset for language X, the evidence from language Y with similar constructs is most welcome and can be used as evidence in favor of one of the two analyses for language X. I call this approach the *bottom-up approach with cheating*: unless there is contradicting evidence, we can reuse analyses that have been developed for other languages.

Note that this approach is compatible with the rather agnostic view advocated by Haspelmath (2010a), Dryer (1997), Croft (2001: Section 1.4.2–1.4.3), and others, who argue that descriptive categories should be language-specific, that is, the notion of *subject* for Tagalog is different from the one for English, the category *noun* in English is different from the category *noun* in Persian and so on. Even if one follows such extreme positions, one can still derive generalizations regarding constituent structure, head-argument relations and so on. However, I believe that some categories can fruitfully be used cross-linguistically; if not universally, then at least for language classes. As Newmeyer (2010: 692) notes with regard to the notion of *subject*: calling two items *subject* in one language does not entail that they have identical properties. The same is true for two linguistic items from different languages: calling a Persian linguistic item *subject* does not entail that it has exactly the same properties as an English linguistic item that is called *subject*. The same is, of course, true for all other categories and relations, for instance, parts of speech: Persian nouns do not share all properties with English nouns.⁴ Haspelmath (2010c: 697) writes: "Generative linguists try to use as many crosslinguistic categories in the description of individual languages as possible, and this often leads to insurmountable problems." If the assumption of a category results in problems, they have to be

³ Note that binarity is more restrictive than flat structures: there is an additional constraint that there have to be exactly two daughters. As was argued in Section 21.10.4 one needs phrasal constructions with more than two constituents.

⁴ Note that using labels like *Persian Noun* and *English Noun* (see for instance Haspelmath 2010a: Section 2 for such a suggestion regarding case, e.g., Russian Dative, Korean Dative, ...) is somehow strange since it implies that both Persian nouns and English nouns are somehow nouns. Instead of using the category *Persian Noun* one could assign objects of the respective class to the class *noun* and add a feature *LANGUAGE* with the value *persian*. This simple trick allows one to assign both objects of the type *Persian Noun* and objects of the type *English Noun* to the class *noun* and still maintain the fact that there are differences. Of course, no theoretical linguist would introduce the *LANGUAGE* feature to differentiate between Persian and English nouns, but nouns in the respective languages have other features that make them differ. So the part of speech classification as noun is a generalization over nouns in various languages and the categories *Persian Noun* and *English Noun* are feature bundles that contain further, language-specific information.

solved. If this is not possible with the given set of categories/features, new ones have to be assumed. This is not a drawback of the methodology, quite the opposite is true: if we have found something that does not integrate nicely into what we already have, this is a sign that we have discovered something new and exciting. If we stick to language-particular categories and features, it is much harder to notice that a special phenomenon is involved, since all categories and features are specific to one language anyway. Note also that not all speakers of a language community have exactly the same categories. If one were to take the idea of language-particular category symbols to an extreme, one would end up with person specific category symbols like *Klaus-English-noun*.

After my talk at the MIT in 2013, members of the linguistics department objected to the approach taken in the CoreGram project and claimed that it would not make any predictions as far as possible/impossible languages are concerned. Regarding predictions two things must be said: firstly, predictions are being made on a language particular basis. As an example consider the following sentences from Netter (1991):

- (4) a. [Versucht, zu lesen], hat er das Buch nicht.
 tried to read has he.NOM the.ACC book not
 ‘He did not try to read the book.’
 b. [Versucht, einen Freund vorzustellen], hat er ihr noch nie.
 tried a.ACC friend to.introduce has he.NOM her.DAT yet never
 ‘He never before tried to introduce a friend to her.’

When I first read these sentences I had no idea about their structure. I switched on my computer and typed them in and within milliseconds I got an analysis of the sentences and by inspecting the result I realized that these sentences are combinations of partial verb phrase fronting and the so-called third construction (Müller 1999a: 439). I had previously implemented analyses of both phenomena but had never thought about the interaction of the two. The grammar predicted that examples like (4) are grammatical. Similarly the constraints of the grammar can interact to rule out certain structures. So predictions about ungrammaticality/impossible structures are in fact made as well.

Secondly, the topmost constraint set holds for all languages seen so far. It can be regarded as a hypothesis about properties that are shared by all languages. This constraint set contains constraints about the connection between syntax and information structure and such constraints allow for V2 languages but rule out languages with the verb in penultimate position (see Kayne 1994: 50 for the claim that such languages do not exist. Kayne develops a complicated syntactic system that predicts this). Of course, if a language is found that places the verb in penultimate position for the encoding of sentence types or some other communicative effect, a more general topmost set has to be defined. But this is parallel for Minimalist theories: if languages are found that are incompatible with basic assumptions, the basic assumptions have to be revised. As with the language particular constraints, the constraints in the topmost set make certain predictions about what can be and what cannot be found in languages.

Frequently discussed examples such as those languages that form questions by reversing the order of the words in a string (Musso et al. 2003) need not be ruled out by

the grammar, since they are ruled out by language external constraints: we simply lack the working memory to do such complex computations.

A variant of this argument comes from David Pesetsky and was raised in Facebook discussions of an article by Paul Ibbotson and Michael Tomasello published in *The Guardian*⁵. Pesetsky claimed that Tomasello's theory of language acquisition could not explain why we find V2 languages but no V3 languages. First, I do not know of anything that blocks V3 languages in current Minimalist theories. So per se the fact that V3 languages may not exist cannot be used to support any of the competing approaches. Of course, the question could be asked whether the V3 pattern would be useful for reaching our communicative goals and whether it can be easily acquired. Now, with V2 as a pattern it is clear that we have exactly one position that can be used for special purposes in the V2 sentence (topic or focus). For monovalent and bivalent verbs we have an argument that can be placed in initial position. The situation is different for the hypothetical V3 languages, though: If we have monovalent verbs like *sleep*, there is nothing for the second position. As Pesetsky pointed out in the answer to my comment on a blog post, languages solve such problems by using expletives. For instance some languages insert an expletive to mark subject extraction in embedded interrogative sentences, since otherwise the fact that the subject is extracted would not be recognizable by the hearer. So the expletive helps to make the structure transparent. V2 languages also use expletives to fill the initial position if speakers want to avoid something in the special, designated position:

- (5) Es kamen drei Männer zum Tor hinein.
 EXPL came three man to.the gate in
 'Three man came through the gate.'

In order to do the same in V3 languages one would have to put two expletives in front of the verb. So there seem to be many disadvantages of a V3 system that V2 systems do not have and hence one would expect that V3 systems are less likely to come into existence. If they existed, they would be expected to be subject to change in the course of time; e.g., omission of the expletive with intransitives, optional V2 with transitives and finally V2 in general. With the new modeling techniques for language acquisition and agent-based community simulation one can actually simulate such processes and I guess in the years to come, we will see exciting work in this area.

Cinque (1999: 106) suggested a cascade of functional projections to account for re-occurring orderings in the languages of the world. He assumes elaborate tree structures to play a role in the analysis of all sentences in all languages even if there is no evidence for respective morphosyntactic distinctions in a particular language (see also Cinque & Rizzi 2010: 55). In the latter case, Cinque assumes that the respective tree nodes are empty. Cinque's results could be incorporated in the model advocated here. We would define part of speech categories and morpho-syntactic features in the topmost set and state linearization constraints that enforce the order that Cinque encoded directly in

⁵ *The roots of language: What makes us different from other animals?* Published 05.11.2015. <http://www.theguardian.com/science/head-quarters/2015/nov/05/roots-language-what-makes-us-different-animals>

his tree structure. In languages in which such categories are not manifested by lexical material, the constraints would never apply. Neither empty elements nor elaborate tree structures would be needed. Thus Cinque's data could be covered in a better way in an HPSG with a rich UG but I, nevertheless, refrain from introducing 400 categories (or features) into the theories of all languages and, again, I point out that such a rich and language-specific UG is implausible from a genetic point of view. Therefore, I wait for other, probably functional, explanations of the Cinque data.

Note also that implicational universals can be derived from hierarchically organized constraint sets as the ones proposed here. For instance, one can derive from Figure 22.5 the implicational statement that all SVO languages are V2 languages, since there is no language that has constraints from Set 4 that does not also have the constraints of Set 7. Of course, this implicational statement is wrong, since there are lots and lots of SOV languages and just exceptionally few V2 languages. So, as soon as we add other languages as for instance Persian or Japanese, the picture will change.

The methodology suggested here differs from what is done in MGG, since MGG stipulates the general constraints that are supposed to hold for all languages on the basis of general speculations about language. In the best case, these general assumptions are fed by a lot of experience with different languages and grammars, in the worst case they are derived from insights gathered from one or more Indo-European languages. Quite often impressionistic data is used to motivate rather far-reaching fundamental design decisions (Fanselow 2009; Sternefeld & Richter 2012; Haider 2014). It is interesting to note that this is exactly what members of the MGG camp reproach typologists for. Evans & Levinson (2009a) pointed out that counterexamples can be found for many alleged universals. A frequent response to this is that unanalyzed data cannot refute grammatical hypotheses (see, for instance, Freidin 2009: 454). In the very same way it has to be said that unanalyzed data should not be used to build theories on (Fanselow 2009). In the CoreGram project, we aim to develop broad-coverage grammars of several languages, so those constraints that make it to the top node are motivated and not stipulated on the basis of intuitive implicit knowledge about language.

Since it is data-oriented and does not presuppose innate language-specific knowledge, this research strategy is compatible with work carried out in Construction Grammar (see Goldberg 2013b: 481 for an explicit statement to this end) and in any case it should also be compatible with the Minimalist world.

第二十三章 结论

本书讨论的分析都呈现出了一些相似性。所有框架都使用了复杂范畴来描述语言对象。这在 GPSG、LFG、HPSG、CxG 和 FTAG 中尤为明显。但是, GB/最简方案和范畴语法也讨论了第三人称单数 NPs, 以及构成复杂范畴部分的词类、人称和数的相关特征。GB 理论有二元值的特征 N 和 V(Chomsky 1970: 199), Stabler (1992: 119)用特征值偶对构成语障 (*Barriers*), Sauerland & Elbourne (2002: 290–291)主张在最简方案中使用特征值偶对。同样, 请看 Veenstra (1998)使用类型特征描写构建的基于约束的最简分析。依存语法的变体, 如 Hellwig 的依存合一语法, 也使用了特征值偶对 (Hellwig 2003: 612)。

再者, 所有当代的理论框架 (除了构式语法和依存语法) 在分析德语句子结构时都体现出了一致性: 德语是一个 SOV 和 V2 型语言。在结构上, 动词首位的从句语序跟动词末位的从句语序是一致的。定式动词要么移位 (GB), 要么跟动词末位位置上的元素有关系 (HPSG)。动词二位的从句包括一个成分都没有提取的动词首位的从句。这跟被动的分析也趋于一致: 最初由 Haider (1984, 1985a, 1986a) 在 GB 框架下构建的思想被借鉴到 HPSG 中。构式语法的一些变体也利用了特殊标记的“指定论元” (Michaelis & Ruppenhofer 2001: 55–57)。

如果我们考虑单个框架的新发展, 很清楚的是所提出的分析有时在本质上有极大的区别。而 CG、LFG、HPSG 和 CxG 都是表层导向的, 有时在最简方案和有些情况下提出非常抽象的结构, 有人试图将所有语言都归结于一个基本的结构 (普遍性本质假说)¹ 只有我们假定适用于所有语言基本结构和生成表层结构那些必要操作的天赋知识是存在的, 这种方法才是有意义的。正如在第 13 章所展示的, 所有关于天赋语言知识的假说要么是站不住脚的, 要么至少是有争议的。语言能力的习得可以在很大程度上得到基于输入的解释 (第 13.8.3 节、第 16.3 节和第 16.4 节)。不是所有关于习得的问题都被一次性解决, 但是, 对那些对天赋的语言知识持谨慎态度的人来说, 基于输入的方法至少是貌似正确的。

诸如 LFG、CG、HPSG、CxG 和 TAG 的模型都兼容于语言运用的数据, 这些数据在特定的基于转换的方法中是不考虑语言运用的, 这种基于转换的方法被看作是语言能力的理论, 而不是语言运用的理论。MGG 认为还有其他机制作用于语言知识, 如结合“组块” (语言材料的片段) 的机制。如果我们想做出这些假设, 那么就需要解释组块和组块的过程是如何获得的, 而不是转换的复杂系统和转换对比的限制是如何获得的。这就意味着语言习得的问题会是非常不同的。如果我们认同基于组块的方法, 那么关于普遍的转换基础的天赋知识就只能从表层导向的

¹ 需要指出的是, 现在有很多子变体和最简方案团体的个别观点, 所以只能像 CxG 那样说倾向性。

语法生成。那么，这就引出了在语言能力的语法中转换的证据到底是什么的问题，而且是不是应该简单地认为语言能力的语法属于 LFG、CG、HPSG、CxG 或 TAG 所假设的类型。由此，我们可以总结出这样的结论，基于约束的分析与允许基于约束的重新建构的转换方法的类型是跟现在的事实唯一兼容的方法，而其他所有的分析都需要额外的假设。

最简方案的一些研究区别于其他框架下的研究，因为他们认为结构只能从其他语言的事实所激发。这就可以简化生成不同结构的整个机制，但是这个方法的所有代价并没有减少：有些代价只是被转化为 UG 的组成部分。他们得到的抽象语法是无法从输入学会的。

我们可以从这个讨论获知，只有基于约束的、表层导向的模型是合适的且具有解释力的：他们也跟心理语言学的事实兼容，并且从习得的角度来看貌似是合理的。

如果我们现在比较这些方法，我们会发现一些分析可以互相转换。LFG（和 CxG 与 DG 的一些变体）与其他所有理论不同，因为诸如主语和宾语的语法功能是基本类型。如果我们不想要这样的观点，那么就可以将这些标签替换为 Argument1、Argument2 等。论元的数量对应于他们相对的旁格性。LFG 就需要向 HPSG 靠拢了。另外，我们可以在 HPSG 和 CxG 中根据语法功能来标记论元。在被动分析中就是这样做的（DESIGNATED ARGUMENT）。

LFG、HPSG、CxG 和范畴语法的变体 (Moens et al. 1989; Briscoe 2000; Villavicencio 2002) 有对知识进行层级式组织的方式，这对普遍规律的捕捉是非常重要的。当然，可以按照这个方式来扩展其他任何的框架，但是除了计算机的应用之外，这从未被明确地表示出来，而且承继层级体系在其他框架的理论方面没有起到有效的作用。

在 HPSG 和 CxG 中，词根、词干、词、形态规则和句法规则都是可以按照相同方式描述的对象。这就允许我们总结出影响不同对象的普遍规律（见第 22 章）。在 LFG 中，c-结构被看作是本质上不同的，这就是为什么这种普遍化是不可能的。在跨语言的研究中，有一种在 f-结构中捕捉相似性的尝试，c-结构是不太重要的，甚至在一些研究都没有提及。再者，针对不同语言的计算实现有较大的差异。基于此，我个人倾向于使用相同方式描写所有语言对象的理论框架，即 HPSG 和 CxG。形式上，没有什么阻挡了使用特征值偶对的 LFG 语法的 c-结构描写，所以在可预见的未来理论之间会有更多的地方趋于一致。例如，有关 HPSG 和 LFG 的混合形式请见 Ackerman & Webelhuth (1998) 和 Hellan & Haugereid (2003)。

如果我们比较 CxG 和 HPSG，CxG 研究中形式化的程度明显是相对较低的，而且有一些问题没有得到解答。CxG（除了流体构式语法）中较为形式化的方法是 HPSG 的变体。构式语法中相对来说有较少的精细化构造的分析，而且没有能跟本书中介绍的其他方法相比的针对德语的描写。公平来说，必须要指出的是，构式语法是这里讨论的理论中最为年轻的。它对语言学理论的重要贡献被整合进 HPSG 和 LFG 这类理论框架中。

未来的理论将是表层导向、基于约束和理论模型方法的整合，TAG 和 GB/最简方案 的方法将会变革为基于约束的方法。最简方案的（变体）和构式语法的（变体）是目前最为广泛采用的方法。实际上，我对其中的事实保持怀疑。未来的语言学将是数据驱动的。内省式的方法作为唯一的数据搜集的方法被证明是不可靠的 (Müller 2007e; Meurers & Müller 2009)，而且更多地需要辅以实验的应用和基于语料库的分析。

统计信息和统计程序在机器翻译中发挥了重要的作用，而且在狭义的语言学方面变得更重要了 (Abney 1996)。我们已经看到，统计信息在习得过程中是非常重要的，而且 Abney 讨论了语言其他方面的情况，如语言变化、句法分析倾向性和合格性判断递差。随着统计程式越来越受到重视，现在计算语言学面临着向混合形式的转向，² 因为我们注意到不可能只用统计方法来超过一定的质量水平 (Steedman 2011; Church 2011; Kay 2011)。跟前述一样的是：真相位于两者之间，即整合的系统。为了整合，相关的语言学理论首先需要进一步发展。正如 Manfred Pinkal 所说的：“我们不可能在没有了解语言的情况下构造出理解语言的系统。”

² 有关 HPSG 语法的话语识别器的混合请参考 Kaufmann & Pfister (2007) 和 Kaufmann (2009)。

第二十四章 练习题答案

24.1 导言与术语

- (1) a. $\underbrace{\text{Karl}}_{\text{VF}} \underbrace{\text{isst}}_{\text{LS}} .$
- b. $\underbrace{\text{Der Mann}}_{\text{VF}} \underbrace{\text{liebt}}_{\text{LS}} \underbrace{\text{eine Frau,}}_{\text{MF}} \underbrace{\underbrace{\text{den}}_{\text{VF}} \underbrace{\text{Peter}}_{\text{MF}} \underbrace{\text{kennt.}}_{\text{RS}}}_{\text{NF}}$
- c. $\underbrace{\text{Der Mann}}_{\text{VF}} \underbrace{\text{liebt}}_{\text{LS}} \underbrace{\underbrace{\text{eine Frau,}}_{\text{MF}} \underbrace{\underbrace{\text{die}}_{\text{VF}} \underbrace{\text{Peter}}_{\text{MF}} \underbrace{\text{kennt.}}_{\text{RS}}}_{\text{NF}}}_{\text{MF}}$
- d. $\underbrace{\text{Die Studenten}}_{\text{VF}} \underbrace{\text{haben}}_{\text{LS}} \underbrace{\text{behauptet,}}_{\text{RS}} \underbrace{\underbrace{\text{nur wegen der Hitze einzuschlafen.}}_{\text{MF}}}_{\text{NF}}$
- e. $\underbrace{\underbrace{\underbrace{\text{Dass}}_{\text{LS}} \underbrace{\text{Peter nicht kommt,}}_{\text{MF}} \underbrace{\text{ärgert}}_{\text{RS}}}_{\text{VF}} \underbrace{\underbrace{\text{Klaus.}}_{\text{LS}} \underbrace{\text{}}_{\text{MF}}}_{\text{MF}}}_{\text{VF}}$
- f. $\underbrace{\underbrace{\underbrace{\text{Einen Mann küssen,}}_{\text{MF}} \underbrace{\text{der}}_{\text{VF}} \underbrace{\text{ihr nicht gefällt,}}_{\text{MF}} \underbrace{\text{würde sie nie.}}_{\text{RS}}}_{\text{NF}}}_{\text{VF}}$

对于 (1c) 而言：理论上，这也可以是关系小句外置到后场的一个例子。因为 *eine Frau, die Peter kennt* (“Peter 认识的一个女人”) 是一个组成成分，不过，一般认为这个关系小句并没有重新排序。相反，我们有一个更简单的结构，它把 *eine Frau, die Peter kennt* 作为一个完整的 NP 放在中场的位置上。

24.2 短语结构语法

1. 任何一种语法都有可能提出生成不必要结构的符号和规则，或者只是简单地从未使用过的符号和规则，因为规则的右边没有能够使用的词或短语。比如说，如果我们在文法中加入下面的规则，我们需要一个仍能分析这个语言的相同片段的更为复杂的文法。

(2) Tralala \rightarrow Trulla Trololo

2. 一般认为，有着最少规则的文法是最好的。这样，我们可以拒绝那些像 (2) 一样包括不必要规则的文法。

我们需要牢记，文法理论的目的是什么。如果我们的目标是为了描写人类的语言能力，那么有着更多规则的文法就比那些具有较少规则的文法好。这是因为心理语言学的研究显示高频的单位易于在我们的大脑中存储，而且并不是每次都由他们的个别部分构成，尽管我们当然有能力这么做。

3. 这里的问题是有可能生成一个完整的空名词短语（见图 24.1）。这个名词短语

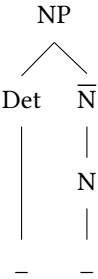


图 24.1: 不带限定词和名词的名词短语

可以被插进填充的 NP 所在的所有位置上。那么，我们可以分析例 (3) 这样的词的序列，其中 *schläft* (“睡觉”) 的主语被实现为一个空 NP:

(3) *Ich glaube, dass schläft.
我 认为 CONJ 睡觉

这个问题可以使用一个特征来解决，这个特征决定了 \bar{N} 的左边界是否是空的。至少带一个形容词的可见的 N_s 和 \bar{N} 具有 “-” 值，其他的是 “+”。空限定语就只能跟具有 “-” 值的 \bar{N}_s 相、组合。见 Netter (1994)。

4. 如果 *Bücher* (书) 在词库中是一个名词，那么为了使得诸如 (4) 的短语是可分析的，*interessant* (有趣的) 这类形容词就需要修饰 NPs。

(4) interessante Bücher
有趣的 书

不过，如果形容词跟相 NPs 组合，仍需要解释 (5) 为什么是不合乎语法的。

- (5) *interessante die Bücher
有趣的 DET 书

有关这个话题的详细讨论，见Müller (2007b: Section 6.6.2)。

5. 这类规则不能分析 (6) 中的那些名词短语：

- (6) a. interessante [Aufsätze und Bücher]
有趣的 文章 和 书
b. interessante [Aufsätze und Bücher aus Stuttgart]
有趣的 文章 和 书 PREP 斯图加特

由于形容词只能跟名词直接组合，这些短语就是无法分析的。*Bücher* (“书”) 或 *Bücher aus Stuttgart* (“来自斯图加特的书”) 就是完整 NPs 了。因为一般认为并列成分要求具有相同的句法范畴，所以 *Aufsätze* (“文章”) 就应该是一个 NP。*Aufsätze und Bücher* (“文章和书”) 和 *Aufsätze und Bücher aus Stuttgart* (“来自斯图加特的文章和书”) 也应该是 NPs，这样还是没有解释为什么形容词跟这个名词组合。因为 (5) 的关系，我们必须排除那些认为完整 NPs 跟形容词组合的分析。

关于空成分的概括性论述见第 19 章。

6. 如果某个冠词或者任何一个冠词都要跟一个形容词组合以构成一个完整的 NP，就没有地儿给后置修饰语了，如修饰属格、PPs 和关系从句的修饰语。对于 PPs 和关系从句来说，分析认为这些后置修饰语附加在完整 NPs 上 (Kiss 2005)，但是修饰性属格通常附加在更小的单位上。但是即使我们承认后置修饰语附加在完整 NPs 上，我们无法解释形容词的重叠，以及依存于省略名词的论元。

所以，最简单的处理德语的方法就是提出空名词的假设。或者，我们可以假设形容词直接投射到 \bar{N} 上。这个 \bar{N} 就被形容词和后置形容词修饰。跟冠词组合构成一个完整的 \bar{N} 。对于包括省略的关系名词短语来说，我们必须假设 *vom Gleimtunnel* (“属于格莱姆隧道的”) 这样的论元投射到 \bar{N} 上。 \bar{N} 就可以进一步被修饰或者直接跟冠词组合。

7. 我们无法分析 (7) 中的那些形容词短语，因为程度修饰语位于补足语和形容词词之间：

- (7) der auf seinen Sohn sehr stolze Vater
DET PREP 他的 儿子 非常骄傲 父亲
‘对儿子感到非常骄傲的父亲’

我们要么允准限定语跟他们的中心语组合位于补足语前，要么允准树的交叉。另一个假设是德语像英语一样，不过形容词性补足语就需要强制地位于他们的限定语前面。有关这类重新排序的描写，请参考第 3 章。第 13.1.2 节是关于 X-理论的讨论。

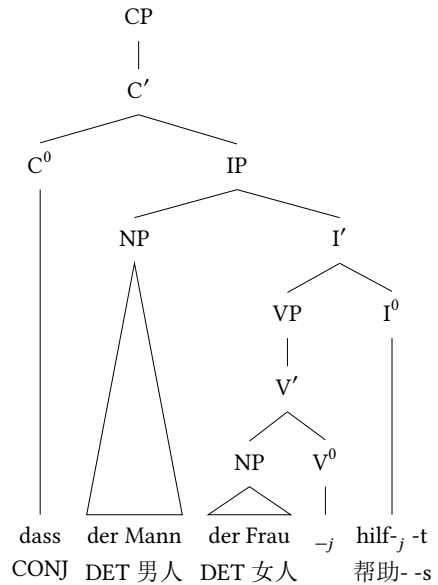
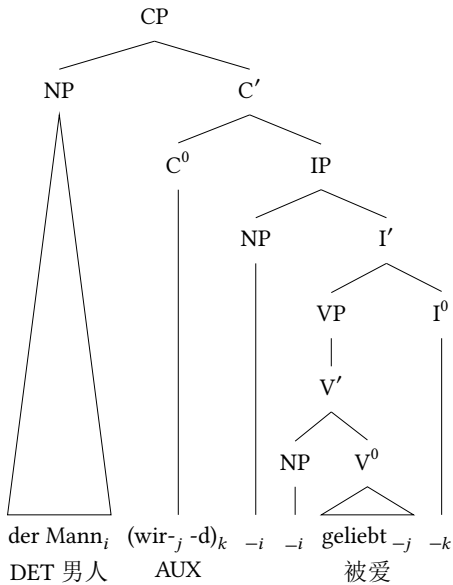
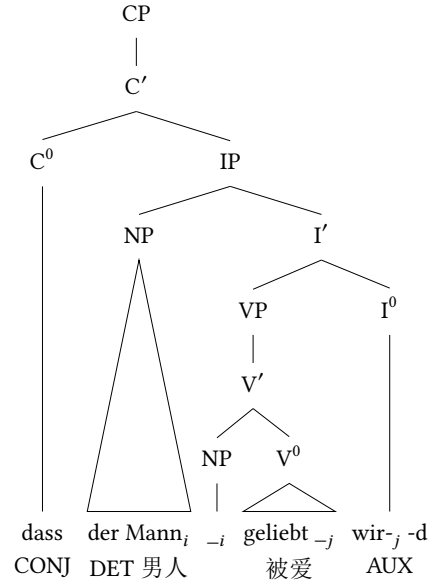
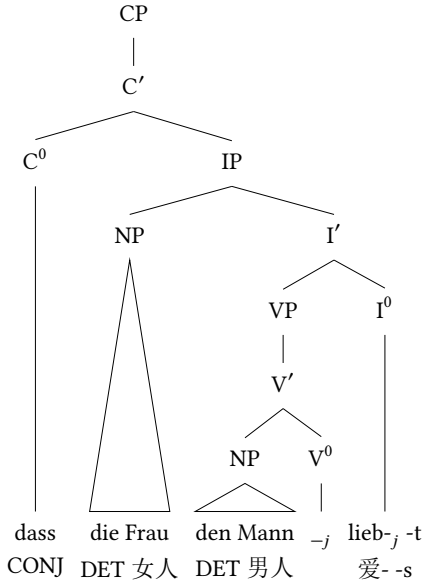
8. 请写一段能够分析例（8）中句子的短语结构文法，但是不要允准（9）那样的词串。

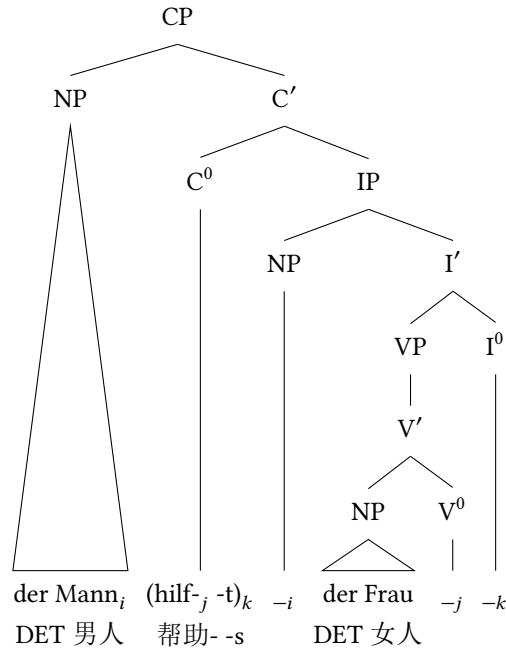
- (8) a. Der Mann hilft der Frau.
DET.NOM 男人 帮助DET.DAT 女人
‘这个男人帮助这个女人。’
- b. Er gibt ihr das Buch.
他.NOM 给 她.DAT DET.ACC 书
‘他给她这本书。’
- c. Er wartet auf ein Wunder.
他.NOM 等 PREP ART 奇迹.ACC
‘他在等一个奇迹。’
- (9) a. *Der Mann hilft er.
DET.NOM 男人 帮助他.NOM
- b. *Er gibt ihr den Buch.
他.NOM 给 她.DAT DET.ACC 书

为了排除后面两个句子，文法必须要包括格的信息。下面的文法就可以解决这个问题：

- (10) a. $s \rightarrow np(nom) v(nom_dat), np(dat)$
b. $s \rightarrow np(nom), v(nom_dat_acc), np(dat), np(acc)$
c. $s \rightarrow np(nom), v(nom_pp_auf), pp(auf, acc)$
d. $pp(Pform, Case) \rightarrow p(Pform, Case), np(Case)$
e. $np(Case) \rightarrow d(Case), n(Case)$
f. $v(nom_dat) \rightarrow hilft$
g. $v(nom_dat_acc) \rightarrow gibt$
h. $v(nom_pp_auf) \rightarrow wartet$
i. $np(nom) \rightarrow er$
j. $np(dat) \rightarrow ihr$
k. $d(nom) \rightarrow der$
l. $d(dat) \rightarrow der$
m. $d(acc) \rightarrow das$
n. $d(acc) \rightarrow ein$
o. $n(nom) \rightarrow Mann$
p. $n(dat) \rightarrow Frau$
q. $n(acc) \rightarrow Buch$
r. $n(acc) \rightarrow Wunder$
s. $p(auf, acc) \rightarrow auf$

24.3 转换语法—管辖与约束理论





24.4 广义短语结构语法

为了分析例 (11) 中的句子，我们需要一条及物动词规则和一条成分提取的元规则。进而，还需要名词短语内成分的组合规则。

- (11) a. [dass] der Mann ihn liest
CONJ DET 男人 它 读
‘这个男人读它’
b. [dass] ihn der Mann liest
CONJ 它 DET 男人 读
‘这个男人读它’
c. Der Mann liest ihn.
DET 男人 读 它
‘这个男人读它。’

可以用 (12) 的规则和 (13) 中的词汇项分析 (11a,b) 中的句子。

- (12) a. $V3 \rightarrow H[6], N2[CASE\ nom], N2[CASE\ acc]$
b. $N2 \rightarrow Det[CASE\ CAS], H1[CASE\ CAS]$
c. $N1 \rightarrow H[27]$
(13) a. $Det[CASE\ nom] \rightarrow der$

- b. $N[27] \rightarrow \text{Mann}$
- c. $V[6, +\text{FIN}] \rightarrow \text{liest}$
- d. $N2[\text{CASE acc}] \rightarrow \text{ihn}$

规则 (12b, c) 对应于我们在第 2.4.1 节讲到的 \bar{X} -规则。他们跟这些规则的不同之处在于没有在规则的右边给出中心语的词类。词类是通过中心语特征原则决定的。中心语的词类跟规则左边成分的词类保持一致，即它必须是 (12b, c) 中的 N。它也从中心语特征原则得出，整个 NP 具有跟中心语一样的格，这样就不必在上面的规则中额外提及了。27 是一个次范畴。这个数字是任意的。

为了让动词出现在正确的位置上，我们需要线性化规则：

- (14) $V[+\text{MC}] < X$
 $X < V[-\text{MC}]$

冠词前置于名词的事实通过下面的 LP-规则来实现：

- (15) $\text{Det} < X$

为了分析 (11c)，(16) 的提取元规则是必须的：

- (16) $V3 \rightarrow W, X \mapsto$
 $V3/X \rightarrow W$

其中，这个元规则允准了针对 (12a) 的规则 (17)：

- (17) $V3/N2[\text{CASE nom}] \rightarrow H[6], N2[\text{CASE acc}]$

规则 (18) 被用来描写长距离依存。

- (18) $V3[+\text{FIN}] \rightarrow X[+\text{TOP}], V3[+\text{MC}]/X$

下面的线性规则确保 +TOP-成分前置于它缺失的句子：

- (19) $[+\text{TOP}] < X$

图 24.2 on the following page 展示了该语法允准的结构。总之，我们可以说允准 (11) 句子的文法应该（至少）包括下面这几个部分：

1. ID 规则：

- (20) a. $V3 \rightarrow H[6], N2[\text{CASE nom}], N2[\text{CASE acc}]$
 b. $N2 \rightarrow \text{Det}[\text{CASE CAS}], H1[\text{CASE CAS}]$
 c. $N1 \rightarrow H[27]$

2. LP 规则：

- (21) $V[+\text{MC}] < X$
 $X < V[-\text{MC}]$
 $\text{Det} < X$
 $[+\text{TOP}] < X$

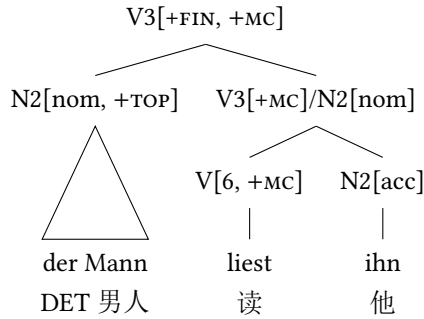


图 24.2: *Der Mann liest ihn.* (“这个男人读它。”) 的分析

3. 元规则:

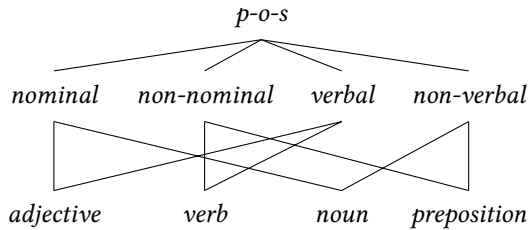
(22) $V3 \rightarrow W, X \mapsto$
 $V3/X \rightarrow W$

4. 词汇项

- (23) a. $\text{Det}[\text{CASE nom}] \rightarrow \text{der}$
 b. $\text{N}[27] \rightarrow \text{Mann}$
 c. $\text{V}[6, +\text{FIN}] \rightarrow \text{liest}$
 d. $\text{N2}[\text{CASE acc}] \rightarrow \text{ihn}$

24.5 特征描写

1. 对于 [+V] 类来说, 类型 *verbal* 被当作是跟 *adjective* 和 *verb* 一起使用的。对于 [-V] 类来说, 有类型 *non-verbal* 和它的子类型 *noun* 和 *preposition*。这跟 N 的值是相似的。相应的层级体系如下图所示:



2. 列表可以通过包括开头和其余部分列表的递归结构来描写。其余成分可以是一个非空列表 (*ne_list*) 或空列表 (*e_list*)。列表 $\langle a, b, c \rangle$ 可以按照下面这样表示:

$$(24) \left[\begin{array}{l} ne_list \\ FIRST \ a \\ \\ REST \ \left[\begin{array}{l} ne_list \\ FIRST \ b \\ \\ REST \ \left[\begin{array}{l} ne_list \\ FIRST \ c \\ REST \ e_list \end{array} \right] \end{array} \right] \end{array} \right]$$

3. 如果我们通过两个额外特征来扩展 (24) 的数据结构, 就有可能不用 *append*。关键词是 *difference list* (“差异表”)。差异表包括一个列表和指向这个列表末端的指针。

$$(25) \left[\begin{array}{l} diff-list \\ \\ LIST \ \left[\begin{array}{l} ne_list \\ FIRST \ a \\ \\ REST \ \left[\begin{array}{l} ne_list \\ FIRST \ b \\ REST \ \boxed{1} \ list \end{array} \right] \end{array} \right] \\ \\ LAST \ \boxed{1} \end{array} \right]$$

与 (24) 列表的表示不同, 这个列表末尾的 *REST* 值不是 *e_list*, 而是 *list*。这就有可能通过在它结束的指针处加上另一个列表来对它进行扩展。(25) 和 (26a) 的联结是 (26b))。

$$(26) \quad \begin{array}{l} \text{a.} \left[\begin{array}{l} diff-list \\ \\ LIST \ \left[\begin{array}{l} ne_list \\ FIRST \ c \\ REST \ \boxed{2} \ list \end{array} \right] \\ \\ LAST \ \boxed{2} \end{array} \right] \\ \\ \text{b.} \left[\begin{array}{l} diff-list \\ \\ LIST \ \left[\begin{array}{l} ne_list \\ FIRST \ a \\ \\ REST \ \left[\begin{array}{l} ne_list \\ FIRST \ b \\ \\ REST \ \left[\begin{array}{l} ne_list \\ FIRST \ c \\ REST \ \boxed{2} \ list \end{array} \right] \end{array} \right] \end{array} \right] \\ \\ LAST \ \boxed{2} \end{array} \right] \end{array}$$

为了整合列表，第二个列表的 **LIST** 值必须跟第一个列表的 **LAST** 值相同。所得列表的 **LAST** 值就对应于第二个列表（例子中的②）的 **LAST** 值。

有关不同列表的编码信息可以通过搜索关键词 *list*、*append* 和 *feature structure* 得到。在搜索结果中，我们可以找到解释不同列表的正在开发文法的网页。

24.6 词汇功能语法

1. *kannte*（“认识”）是一个及物动词：

(27) *kannte* V (↑ PRED) = 'KENNEN(SUBJ, OBJ)'
 (↑ SUBJ AGR CAS = NOM)
 (↑ OBJ AGR CAS = ACC)
 (↑ TENSE) = PAST

2. 在例句 (28) 中，*verschlingen*（“吞食”）的宾语位于前场。

(28) Den Apfel verschlingt David.
 DET.ACC 苹果 吞食 David.NOM
 ‘David 在狼吞虎咽地吃苹果。’

该分析是图 7.2 on page 210 中的分析和第 7.5 节所示的长距离依存分析的组合。宾语不在 VP 内部实现，而是在前场实现。

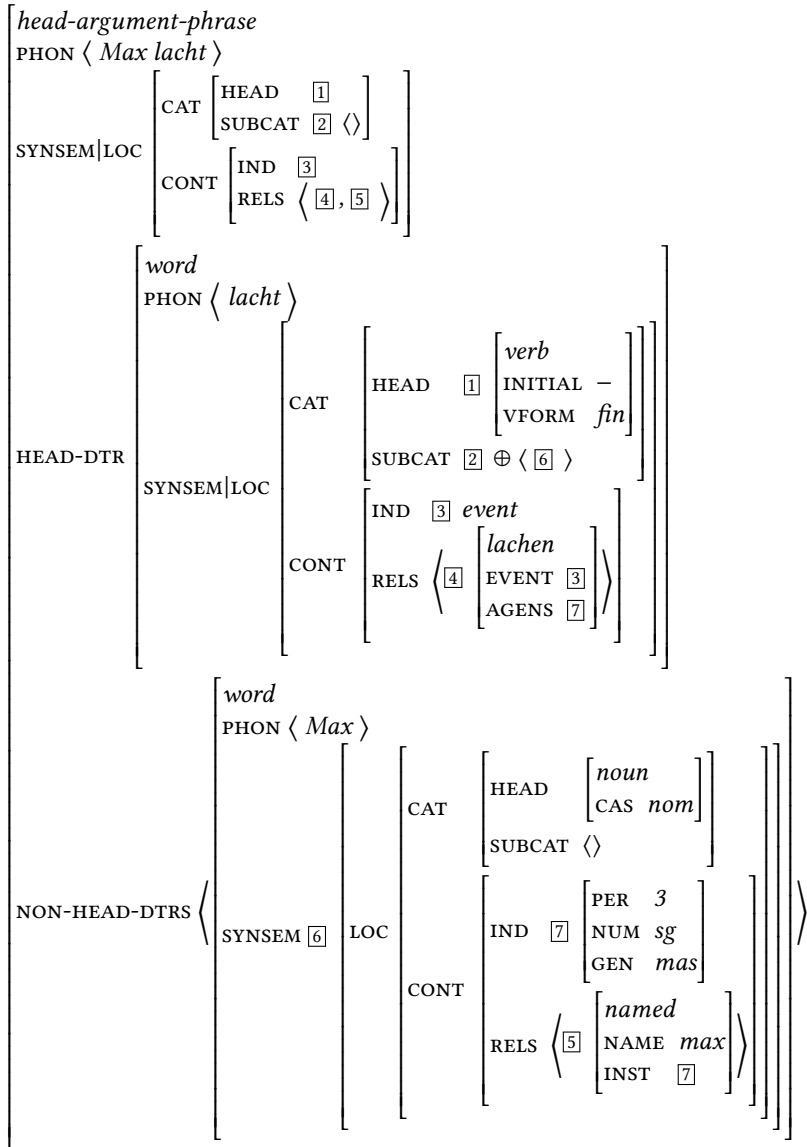
必要的 c-结构规则如 (29) 所示：

(29) a. VP → NP VP
 (↑SUBJ | OBJ) = ↓ ↑ = ↓
 b. VP → (V)
 ↑ = ↓
 c. C' → C VP
 ↑ = ↓ ↑ = ↓
 d. CP → XP C'
 (↑DF) = ↓ ↑ = ↓
 (↑DF) = (↑COMP* GF)

这些规则允准了这里所述例子的两个 f-结构：其中一个 NP *den Apfel*（“苹果”）是话题，另一个 NP 是焦点。图 24.3 显示了前场中话题化成分的分析。

24.8 中心语驱动的短语结构语法

1. 答案是:



2. 对于例(30)中的最小差比对的分析需要捕捉到这样的事实, 形容词的格需要与名词的格保持一致。在(30a)中, 我们使用了 *interessant* (有趣的) 的属格形式, 而(30b)包含的形式是与属格的单数不相容的形式。

- (30) a. eines interessanten Mannes
 一.GEN 有趣的.GEN 男人.GEN
 b. *eines interessanter Mannes
 一.GEN 有趣的.NOM 男人.GEN

(31) 说明了 *interessanten* 的 CAT value 。

(31) 带有格信息的 *interessanten* (有趣的) 的 CAT 值:

$$\left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \text{adj} \\ \text{MOD } \bar{N} [\text{CASE } \boxed{1}] \\ \text{CASE } \boxed{1} \text{ gen} \end{array} \right] \\ \text{SUBCAT } \langle \rangle \end{array} \right]$$

在 MOD 下形容词的格值与 \bar{N} 的格值的结构共享决定了名词和形容词的格值。由此, *interessanten* 可以跟 *Mannes* 组合, 但是不能跟 *Mann* 组合。类似地, *interessanter* 只能跟主格的 *Mann* 组合, 而不能跟属格的 *Mannes* 组合。

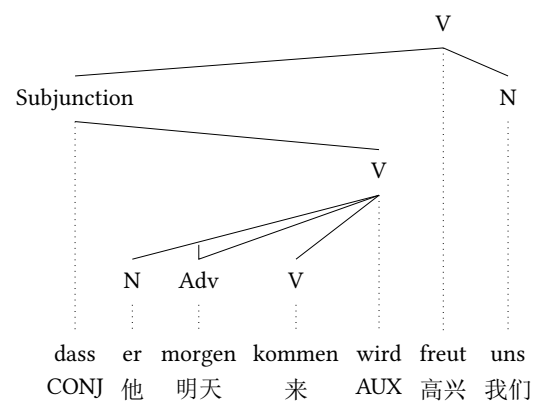
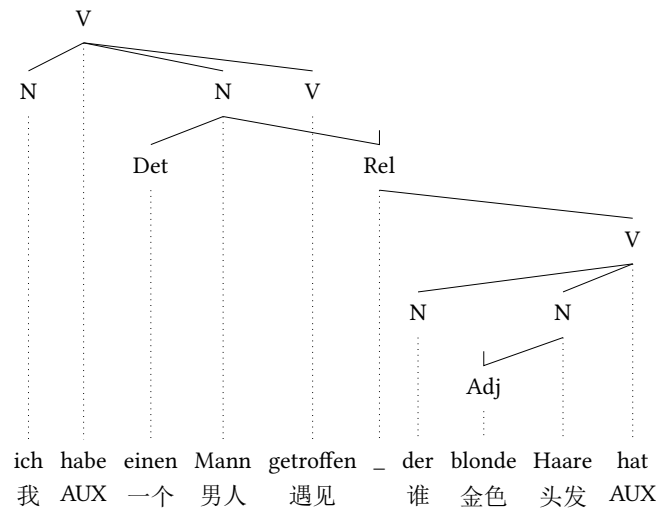
对于名词短语内部一致关系的改进, 请参考Müller (2007b: Abschnitt 13.2)。

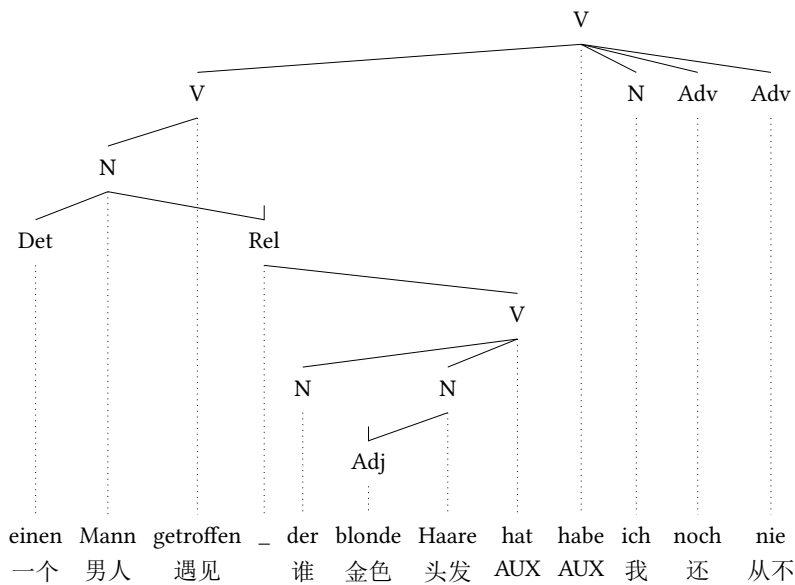
24.9 构式语法

熟语可以通过仔细阅读报纸来获得。稍微枯燥的方法是查找熟语词典, 如“熟语与短语免费辞典”¹。

¹ <http://idioms.thefreedictionary.com/>, 2015 年 3 月 4 日。

24.10 依存语法





24.11 树邻接语法

(32) 的分析需要图 24.6 中的基本树。

- (32) der dem König treue Diener
 DET.NOM DET.DAT 国王 忠诚 仆人
 ‘对国王忠诚的仆人’

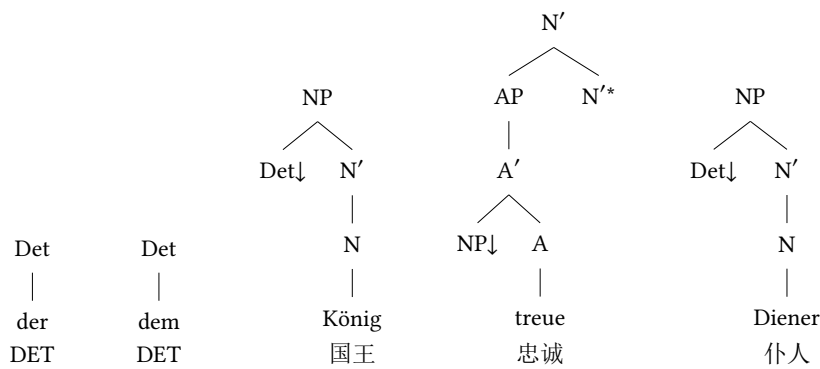


图 24.6: *der dem König treue Diener* 的基本树

通过在替换节点 *König* (“国王”) 替换 *dem* 的树，我们会得到一个完整的 NP。然

后，这就可以插进 *treue* (“忠诚的”) 的替换节点中。相似地，*der* 的树可以跟 *Diener* 的树组合。然后，我们就得到图 24.7 中的两棵树了。

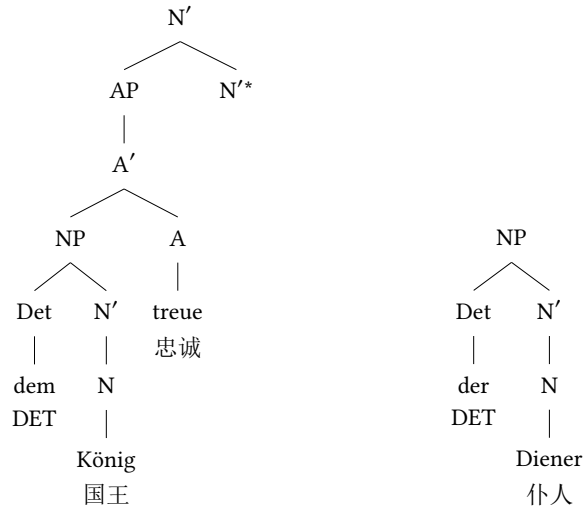


图 24.7: 替换后的 *der dem König treue* 和 *der Diener* 树

形容词树可以连接到 *der Diener* 的 N'-n 节点上，并得到图 24.8 在下一页中的结构。

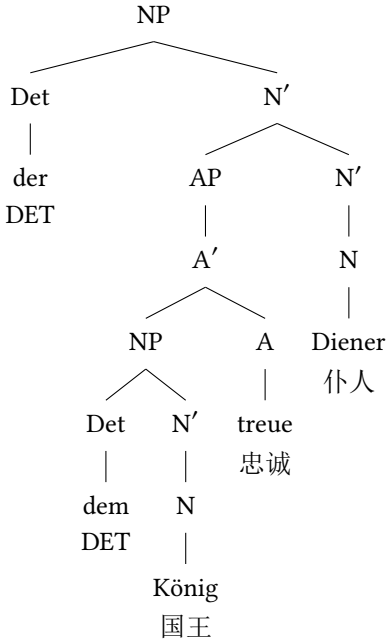


图 24.8: 连接到 N'-节点的结果

参考文献

- Abbott, Barbara. 1976. Right node raising as a test for constituenthood. *Linguistic Inquiry* 7(4). 639–642.
- Abeillé, Anne. 1988. Parsing French with Tree Adjoining Grammar: Some linguistic accounts. In Vargha (1988) 7–12. <http://www.aclweb.org/anthology/C/C88/C88-1002.pdf>.
- Abeillé, Anne. 2006. In defense of lexical coordination. In Olivier Bonami & Patricia Cabredo Hofherr (eds.), *Empirical issues in formal syntax and semantics*, vol. 6, 7–36. Paris: CNRS. <http://www.cssp.cnrs.fr/eiss6/>.
- Abeillé, Anne & Owen Rambow (eds.). 2000a. *Tree Adjoining Grammars: Formalisms, linguistic analysis and processing* (CSLI Lecture Notes 156). Stanford, CA: CSLI Publications.
- Abeillé, Anne & Owen Rambow (eds.). 2000b. *Tree Adjoining Grammars: Formalisms, linguistic analysis and processing* (CSLI Lecture Notes 156). Stanford, CA: CSLI Publications.
- Abeillé, Anne & Yves Schabes. 1989. Parsing idioms in Lexicalized TAG. In Somers & Wood (1989) 1–9.
- Abney, Steven P. 1987. *The English noun phrase in its sentential aspect*. Cambridge, MA: MIT dissertation. <http://www.vinartus.net/spa/87a.pdf>.
- Abney, Steven P. 1996. Statistical methods and linguistics. In Judith L. Klavans & Philip Resnik (eds.), *The balancing act: Combining symbolic and statistical approaches to language* (Language, Speech, and Communication), 1–26. London, England/Cambridge, MA: MIT Press.
- Abney, Steven P. & Jennifer Cole. 1986. A Government-Binding parser. In S. Berman, J-W. Choe & J. McDonough (eds.), *Proceedings of NELS 16*, 1–17. University of Massachusetts, Amherst: GLSA.
- Abney, Steven P. & Erhard W. Hinrichs (eds.). 1995. *Proceedings of the Seventh Conference of the European Chapter of the Association for Computational Linguistics*. Dublin: Association for Computational Linguistics.
- Abraham, Werner (ed.). 1982. *Satzglieder im Deutschen – Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung* (Studien zur deutschen Grammatik 15). Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.
- Abraham, Werner (ed.). 1985. *Erklärende Syntax des Deutschen* (Studien zur deutschen Grammatik 25). Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.

- Abraham, Werner. 1995. *Deutsche Syntax im Sprachenvergleich: Grundlegung einer typologischen Syntax des Deutschen* (Studien zur deutschen Grammatik 41). Tübingen: Stauffenburg Verlag.
- Abraham, Werner. 2003. The syntactic link between thema and rhema: The syntax-discourse interface. *Folia Linguistica* 37(1–2). 13–34.
- Abzianidze, Lasha. 2011. *An HPSG-based formal grammar of a core fragment of Georgian implemented in TRALE*. Charles University in Prague MA thesis.
- Ackerman, Farrell & Gert Webelhuth. 1998. *A theory of predicates* (CSLI Lecture Notes 76). Stanford, CA: CSLI Publications.
- Adams, Marianne. 1984. Multiple interrogation in Italian. *The Linguistic Review* 4(1). 1–27.
- Ades, Anthony E. & Mark J. Steedman. 1982. On the order of words. *Linguistics and Philosophy* 4(4). 517–558.
- Adger, David. 2003. *Core syntax: A Minimalist approach* (Oxford Core Linguistics 1). Oxford: Oxford University Press Oxford.
- Adger, David. 2010. A Minimalist theory of feature structure. In Anna Kibort & Greville G. Corbett (eds.), *Features: Perspectives on a key notion in linguistics* (Oxford Linguistics), 185–218. Oxford: Oxford University Press.
- Adger, David. 2013. Constructions and grammatical explanation: Comments on Goldberg. *Mind and Language* 28(4). 466–478.
- Ágel, Vilmos. 2000. *Valenztheorie* (Narr Studienbücher). Tübingen: Gunter Narr Verlag.
- Ágel, Vilmos, Ludwig M. Eichinger, Hans Werner Eroms, Peter Hellwig, Hans Jürgen Heringer & Henning Lobin (eds.). 2003a. *Dependenz und Valenz / Dependency and valency: Ein internationales Handbuch der zeitgenössischen Forschung / An international handbook of contemporary research*, vol. 25.1 (Handbücher zur Sprach- und Kommunikationswissenschaft). Berlin: Walter de Gruyter.
- Ágel, Vilmos, Ludwig M. Eichinger, Hans Werner Eroms, Peter Hellwig, Hans Jürgen Heringer & Henning Lobin (eds.). 2003b. *Dependenz und Valenz / Dependency and valency: Ein internationales Handbuch der zeitgenössischen Forschung / An international handbook of contemporary research*, vol. 25.1 (Handbücher zur Sprach- und Kommunikationswissenschaft). Berlin: Walter de Gruyter.
- Ágel, Vilmos, Ludwig M. Eichinger, Hans Werner Eroms, Peter Hellwig, Hans Jürgen Heringer & Henning Lobin (eds.). 2006. *Dependenz und Valenz / Dependency and valency: Ein internationales Handbuch der zeitgenössischen Forschung / An international handbook of contemporary research*, vol. 25.2 (Handbücher zur Sprach- und Kommunikationswissenschaft). Berlin: Walter de Gruyter.
- Ágel, Vilmos & Klaus Fischer. 2010. 50 Jahre Valenztheorie und Dependenzgrammatik. *Zeitschrift für Germanistische Linguistik* 38(2). 249–290.
- Ajdukiewicz, Kasimir. 1935. Die syntaktische Konnexität. *Studia Philosophica* 1. 1–27.
- de Alencar, Leonel. 2004. Complementos verbais oracionais – uma análise léxicofuncional. *Lingua(gem)* 1(1). 173–218.
- de Alencar, Leonel. 2013. BrGram: uma gramática computacional de um fragmento do português brasileiro no formalismo da LFG. In *Proceedings of the 9th Brazil-*

- ian Symposium in Information and Human Language Technology. Fortaleza, Ceará, Brazil, October 20–24, 183–188. Fortaleza, Ceará: Sociedade Brasileira de Computação. <http://www.aclweb.org/anthology/W13-4823>.
- Alsina, Alex. 1996. Resultatives: A joint operation of semantic and syntactic structures. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '96 conference, Rank Xerox, Grenoble*, Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/LFG/1/>.
- Alsina, Alex, KP Mohanan & Tara Mohanan. 2005. How to get rid of the COMP. In Butt & King (2005). <http://csli-publications.stanford.edu/LFG/10/lfg05amm.pdf>.
- Ambridge, Ben & Adele E. Goldberg. 2008. The island status of clausal complements: Evidence in favor of an information structure explanation. *Cognitive Linguistics* 19. 349–381.
- Ambridge, Ben & Elena V. M. Lieven. 2011. *Child language acquisition: Contrasting theoretical approaches*. Cambridge, UK: Cambridge University Press.
- Ambridge, Ben, Caroline F. Rowland & Julian M. Pine. 2008. Is structure dependence an innate constraint? New experimental evidence from children's complex-question production. *Cognitive Science: A Multidisciplinary Journal* 32(1). 222–255.
- Anderson, John M. 1971. *The grammar of case: Towards a localistic theory*, vol. 4 (Cambridge Studies in Linguistics). Cambridge, UK: Cambridge University Press.
- Anderson, Stephen R. 1992. *A-morphous morphology* (Cambridge Studies in Linguistics 62). Cambridge: Cambridge University Press.
- Aoun, Joseph & David W. Lightfoot. 1984. Government and contraction. *Linguistic Inquiry* 15(3). 465–473.
- Aoun, Joseph & Dominique Sportiche. 1983. On the formal theory of government. *The Linguistic Review* 2(3). 211–236.
- Arad Greshler, Tali, Livnat Herzig Sheinflux, Nurit Melnik & Shuly Wintner. 2015. Development of maximally reusable grammars: Parallel development of Hebrew and Arabic grammars. In Müller (2015d) 27–40. <http://csli-publications.stanford.edu/HPSG/2015/ahmw.pdf>.
- Arends, Jacques. 2008. A demographic perspective on Creole formation. In Silvia Kouwenberg & John Victor Singler (eds.), *The handbook of pidgin and creole studies*, 309–331. Oxford/Cambridge: Blackwell Publishing Ltd.
- Arka, I Wayan, Avery Andrews, Mary Dalrymple, Meladel Mistica & Jane Simpson. 2009. A linguistic and computational morphosyntactic analysis for the applicative *-i* in Indonesian. In Butt & King (2009) 85–105. <http://csli-publications.stanford.edu/LFG/14/>.
- Arnold, Doug & Andrew Spencer. 2015. A constructional analysis for the skeptical. In Müller (2015d) 41–60. <http://csli-publications.stanford.edu/HPSG/2015/arnold-spencer.pdf>.
- Arnold, Jennifer E., Michael K. Tanenhaus, Rebecca J. Altmann & Maria Fagnano. 2004. The old and thee, uh, new. *Psychological Science* 15(9). 578–582.
- Askedal, John Ole. 1986. Zur vergleichenden Stellungsfelderanalyse von Verbalsätzen und nichtverbalen Satzgliedern. *Deutsch als Fremdsprache* 23. 269–273 and 342–348.

- Asudeh, Ash. 2004. *Resumption as resource management*: Stanford University dissertation.
- Asudeh, Ash, Mary Dalrymple & Ida Toivonen. 2008. Constructions with lexical integrity: Templates as the lexicon-syntax interface. In Butt & King (2008). <http://csli-publications.stanford.edu/LFG/13/>.
- Asudeh, Ash, Mary Dalrymple & Ida Toivonen. 2013. Constructions with lexical integrity. *Journal of Language Modelling* 1(1). 1–54.
- Asudeh, Ash, Gianluca Giorgolo & Ida Toivonen. 2014. Meaning and valency. In Butt & King (2014) 68–88.
- Asudeh, Ash & Ida Toivonen. 2014. With lexical integrity. *Theoretical Linguistics* 40(1–2). 175–186.
- Attardi, Giuseppe. 2006. Experiments with a multilanguage non-projective dependency parser. In *Proceedings of the 10th Conference on Computational Natural Language Learning (CoNLL-X)*, 166–170. Association for Computational Linguistics.
- Attia, Mohammed A. 2008. *Handling Arabic morphological and syntactic ambiguity within the LFG framework with a view to machine translation*: School of Languages, Linguistics and Cultures, University of Manchester dissertation.
- Avgustinova, Tania & Yi Zhang. 2009. Exploiting the Russian national corpus in the development of a Russian Resource Grammar. In Núria Bel, Erhard Hinrichs, Kiril Simov & Petya Osenova (eds.), *Adaptation of language resources and technology to new domains at the RANLP 2009 Conference, Borovets, Bulgaria*, 1–11. Shoumen, Bulgaria: INCOMA Ltd.
- Bach, Emmon. 1962. The order of elements in a Transformational Grammar of German. *Language* 8(3). 263–269.
- Bach, Emmon. 1976. An extension of classical Transformation Grammar. In *Problems in linguistic metatheory, Proceedings of the 1976 Conference at the Michigan State University*, 183–224.
- Bahrani, Mohammad, Hossein Sameti & Mehdi Hafezi Manshadi. 2011. A computational grammar for Persian based on GPSG. *Language Resources and Evaluation* 45(4). 387–408.
- Baker, Carl Lee. 1978. *Introduction to Generative-Transformational Syntax*. Englewood Cliffs, NJ: Prentice-Hall.
- Baker, Mark C. 2003. Linguistic differences and language design. *Trends in Cognitive Sciences* 7(8). 349–353.
- Baker, Mark C. 2009. Language universals: Abstract but not mythological. *The Behavioral and Brain Sciences* 32(5). 448–449.
- Baldrige, Jason. 2002. *Lexically specified derivational control in Combinatory Categorical Grammar*: University of Edinburgh dissertation.
- Baldrige, Jason, Sudipta Chatterjee, Alexis Palmer & Ben Wing. 2007. DotCCG and VisCCG: Wiki and programming paradigms for improved grammar engineering with OpenCCG. In King & Bender (2007) 5–25. <http://csli-publications.stanford.edu/GEAF/2007/>.

- Baldrige, Jason & Geert-Jan M. Kruijff. 2002. Coupling CCG and Hybrid Logic Dependency Semantics. In Isabelle (2002) 319–326. <http://aclanthology.info/events/acl-2002>.
- Ballweg, Joachim. 1997. Stellungenregularitäten in der Nominalphrase. In Hans-Werner Eroms, Gerhard Stickel & Gisela Zifonun (eds.), *Grammatik der deutschen Sprache*, vol. 7.3 (Schriften des Instituts für deutsche Sprache), 2062–2072. Berlin: Walter de Gruyter.
- Baltin, Mark. 1981. Strict bounding. In Carl Lee Baker & John J. McCarthy (eds.), *The logical problem of language acquisition*, 257–295. Cambridge, MA/London, England: MIT Press.
- Baltin, Mark. 2004. Remarks on the relation between language typology and Universal Grammar: Commentary on Newmeyer. *Studies in Language* 28(3). 549–553.
- Baltin, Mark. 2006. Extraposition. In Martin Everaert, Henk van Riemsdijk, Rob Goedemans & Bart Hollebrandse (eds.), *The Blackwell companion to syntax* (Blackwell Handbooks in Linguistics), 237–271. Oxford: Blackwell Publishing Ltd.
- Baltin, Mark R. & Anthony S. Kroch (eds.). 1989. *Alternative conceptions of phrase structure*. Chicago/London: The University of Chicago Press.
- Bangalore, Srinivas, Aravind K. Joshi & Owen Rambow. 2003. Dependency and valency in other theories: Tree Adjoining Grammar. In Ágel et al. (2003b) 669–678.
- Bannard, Colin, Elena Lieven & Michael Tomasello. 2009. Modeling children’s early grammatical knowledge. *Proceedings of the National Academy of Sciences* 106(41). 17284–17289.
- Bar-Hillel, Yehoshua, Micha A. Perles & Eliahu Shamir. 1961. On formal properties of simple phrase-structure grammars. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 14(2). 143–172.
- Bartsch, Renate & Theo Vennemann. 1972. *Semantic structures: A study in the relation between semantics and syntax* (Athenäum-Skripten Linguistik 9). Frankfurt/Main: Athenäum.
- Barwise, Jon & John Perry. 1983. *Situations and attitudes*. Cambridge, MA/London, England: MIT Press.
- Barwise, Jon & John Perry. 1987. *Situationen und Einstellungen – Grundlagen der Situationssemantik*. Berlin, New York: de Gruyter.
- Baschung, K., G. G. Bes, A. Corluy & T. Guillotin. 1987. Auxiliaries and clitics in French UCG grammar. In Bente Maegaard (ed.), *Proceedings of the Third Conference of the European Chapter of the Association for Computational Linguistics*, 173–178. Copenhagen, Denmark: Association for Computational Linguistics.
- Bates, Elizabeth A. 1984. Bioprograms and the innateness hypothesis. *The Behavioral and Brain Sciences* 7(2). 188–190.
- Baumgärtner, Klaus. 1965. Spracherklärung mit den Mitteln der Abhängigkeitsstruktur. *Beiträge zur Sprachkunde und Informationsverarbeitung* 5. 31–53.
- Baumgärtner, Klaus. 1970. Konstituenz und Dependenz: Zur Integration beider grammatischer Prinzipien. In Hugo Steger (ed.), *Vorschläge für eine strukturelle Grammatik des Deutschen* (Wege der Forschung 144), 52–77. Darmstadt: Wissenschaftliche Buchgesellschaft.

参考文献

- Bausewein, Karin. 1990. Haben kopflose Relativsätze tatsächlich keine Köpfe? In Fanselow & Felix (1990) 144–158.
- Bayer, Josef & Jaklin Kornfilt. 1989. Restructuring effects in German. DYANA Report University of Edinburgh.
- Beavers, John. 2003. A CCG implementation for the LKB. LinGO Working Paper 2002-08 CSLI Stanford Stanford, CA. <http://lingo.stanford.edu/pubs/WP-2002-08.ps.gz>.
- Beavers, John. 2004. Type-inheritance Combinatory Categorical Grammar. In *Proceedings of COLING 2004*, 57–63. Geneva, Switzerland: Association for Computational Linguistics.
- Beavers, John, Elias Ponvert & Stephen Mark Wechsler. 2008. Possession of a controlled substantive. In T. Friedman & S. Ito (eds.), *Proceedings of Semantics and Linguistic Theory (SALT) XVIII*, 108–125. Ithaca, NY: Cornell University.
- Beavers, John & Ivan A. Sag. 2004. Coordinate ellipsis and apparent non-constituent coordination. In Müller (2004e) 48–69. <http://csli-publications.stanford.edu/HPSG/2004/>.
- Bech, Gunnar. 1955. *Studien über das deutsche Verbum infinitum* (Linguistische Arbeiten 139). Tübingen: Max Niemeyer Verlag. 2nd unchanged edition 1983.
- Becker, Tilman, Aravind K. Joshi & Owen Rambow. 1991. Long-distance scrambling and Tree Adjoining Grammars. In *Fifth Conference of the European Chapter of the Association for Computational Linguistics. Proceedings of the conference*, 21–26. Berlin: Association for Computational Linguistics. <http://www.aclweb.org/anthology/E91-1005.pdf>.
- Beermann, Dorothee & Lars Hellan. 2004. A treatment of directionals in two implemented HPSG grammars. In Müller (2004e) 357–377. <http://csli-publications.stanford.edu/HPSG/2004/>.
- Beghelli, Filippo & Timothy Stowell. 1997. Distributivity and negation: The syntax of *each* and *every*. In Anna Szabolcsi (ed.), *Ways of scope taking*, 71–107. Dordrecht: Kluwer Academic Publishers.
- Behaghel, Otto. 1909. Beziehung zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25. 110–142.
- Behaghel, Otto. 1930. Von deutscher Wortstellung. *Zeitschrift für Deutschkunde* 44. 81–89.
- Behrens, Heike. 2009. Konstruktionen im Spracherwerb. *Zeitschrift für Germanistische Linguistik* 37(3). 427–444.
- Bellugi, Ursula, Liz Lichtenberger, Wendy Jones, Zona Lai & Marie St. George. 2000. The neurocognitive profile of Williams Syndrome: A complex pattern of strengths and weaknesses. *Journal of Cognitive Neuroscience* 12. 7–29.
- Bender, Emily & Daniel P. Flickinger. 1999. Peripheral constructions and core phenomena: Agreement in tag questions. In Webelhuth et al. (1999) 199–214.
- Bender, Emily M. 2000. *Syntactic variation and linguistic competence: The case of AAVE copula absence*. Stanford University dissertation. <http://faculty.washington.edu/ebender/dissertation/>.

- Bender, Emily M. 2008a. Evaluating a crosslinguistic grammar resource: A case study of Wambaya. In Moore et al. (2008) 977–985. <http://aclweb.org/anthology-new/P/P08/P08-1111.pdf>.
- Bender, Emily M. 2008b. Grammar engineering for linguistic hypothesis testing. In Nicholas Gaylord, Alexis Palmer & Elias Ponvert (eds.), *Proceedings of the Texas Linguistics Society X Conference: Computational linguistics for less-studied languages*, 16–36. Stanford CA: CSLI Publications ONLINE.
- Bender, Emily M. 2008c. Radical non-configurationality without shuffle operators: An analysis of Wambaya. In Müller (2008b) 6–24. <http://csli-publications.stanford.edu/HPSG/2008/>.
- Bender, Emily M. 2010. Reweaving a grammar for Wambaya: A case study in grammar engineering for linguistic hypothesis testing. *Linguistic Issues in Language Technology – LiLT* 3(3). 1–34. <http://journals.linguisticsociety.org/elanguage/lilt/article/view/662/523.html>.
- Bender, Emily M., Daniel P. Flickinger, Frederik Fouvry & Melanie Siegel (eds.). 2003. *Proceedings of the ESSLLI 2003 Workshop “Ideas and Strategies for Multilingual Grammar Development”*. Vienna, Austria.
- Bender, Emily M., Daniel P. Flickinger & Stephan Oepen. 2002. The Grammar Matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. In John Carroll, Nelleke Oostdijk & Richard Sutcliffe (eds.), *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*, 8–14. Taipei, Taiwan.
- Bender, Emily M., Lori Levin, Stefan Müller, Yannick Parmentier & Aarne Ranta (eds.). 2015. *Proceedings of the Grammar Engineering Across Frameworks (GEAF) Workshop*. The Association for Computational Linguistics.
- Bender, Emily M. & Melanie Siegel. 2005. Implementing the syntax of Japanese numeral classifiers. In Su et al. (2005) 626–635.
- Bergen, Benjamin K. & Nancy Chang. 2005. Embodied Construction Grammar in simulation-based language understanding. In Jan-Ola Östman & Mirjam Fried (eds.), *Construction Grammars: Cognitive grounding and theoretical extensions*, 147–190. Amsterdam: John Benjamins Publishing Co.
- Berman, Judith. 1996. Eine LFG-Grammatik des Deutschen. In Berman & Frank (1996) 11–96.
- Berman, Judith. 1999. Does German satisfy the Subject Condition? In Butt & King (1999). <http://csli-publications.stanford.edu/LFG/4/>.
- Berman, Judith. 2003a. *Clausal syntax of German* (Studies in Constraint-Based Lexicalism). Stanford, CA: CSLI Publications.
- Berman, Judith. 2003b. Zum Einfluss der strukturellen Position auf die syntaktische Funktion der Komplementsätze. *Deutsche Sprache* 3. 263–286.
- Berman, Judith. 2007. Functional identification of complement clauses in German and the specification of COMP. In Annie Zaenen, Jane Simpson, Tracy Holloway King, Jane Grimshaw, Joan Maling & Chris Manning (eds.), *Architectures, rules, and preferences: Variations on themes by Joan W. Bresnan*, 69–83. Stanford, CA: CSLI Publications.

- Berman, Judith & Anette Frank. 1996. *Deutsche und französische Syntax im Formalismus der LFG* (Linguistische Arbeiten 344). Tübingen: Max Niemeyer Verlag.
- Berwick, Robert C. 1982. Computational complexity and Lexical-Functional Grammar. *American Journal of Computational Linguistics* 8(3–4). 97–109.
- Berwick, Robert C. & Samuel David Epstein. 1995. On the convergence of the ‘Minimalist’ Syntax and Categorical Grammar. In Anton Nijholt, Giuseppe Scollo & Rene Steetskamp (eds.), *Algebraic methods in language processing*, 143–148. Enschede: University of Twente. <http://eprints.eemcs.utwente.nl/9555/01/twlt10.pdf>.
- Berwick, Robert C. & Partha Niyogi. 1996. Learning from triggers. *Linguistic Inquiry* 27. 605–622.
- Berwick, Robert C., Paul Pietroski, Beracah Yankama & Noam Chomsky. 2011. Poverty of the Stimulus revisited. *Cognitive Science* 35(7). 1207–1242.
- Bick, Eckhard. 2001. En Constraint Grammar parser for dansk. In Peter Widell & Mette Kunøe (eds.), *8. Møde om Udforskningen af Dansk Sprog, 12.–13. October 2000*, vol. 8, 40–50. Århus: Århus University.
- Bick, Eckhard. 2003. A Constraint Grammar-based question answering system for Portuguese. In Fernando Moura Pires & Salvador Abreu (eds.), *Progress in artificial intelligence: 11th Portuguese Conference on Artificial Intelligence, EPIA 2003, Beja, Portugal, December 4–7, 2003, proceedings* (Lecture Notes in Computer Science 2902), 414–418. Berlin: Springer Verlag.
- Bick, Eckhard. 2006. A Constraint Grammar parser for Spanish. In *Proceedings of TIL 2006 – 4th Workshop on Information and Human Language Technology (Ribeirão Preto, October 27–28, 2006)*, 3–10. <http://www.nilc.icmc.usp.br/til/til2006/>.
- Bick, Eckhard. 2009. A Dependency Constraint Grammar for Esperanto. In Bick et al. (2009) 8–12.
- Bick, Eckhard. 2010. FrAG: A hybrid Constraint Grammar parser for French. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner & Daniel Tapias (eds.), *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC’10)*, 794–798. Valletta, Malta: European Language Resources Association (ELRA).
- Bick, Eckhard, Kristin Hagen, Kaili Müürisep & Trond Trosterud (eds.). 2009. *Constraint Grammar and robust parsing: Proceedings of the NODALIDA 2009 workshop* (NEALT Proceedings Series 8) Northern European Association for Language Technology Tartu: Tartu University Library.
- Bick, Eckhard & Lars Nygaard. 2007. Using Danish as a CG interlingua: A wide-coverage Norwegian-English machine translation system. In Joakim Nivre, Heiki-Jaan Kaalep, Kadri Muischnek & Mare Koit (eds.), *Proceedings of the 16th Nordic Conference of Computational Linguistics*, 21–28. Forlag uden navn.
- Bickerton, Derek. 1984a. Creol is still king. *The Behavioral and Brain Sciences* 7(2). 212–218.
- Bickerton, Derek. 1984b. The Language Bioprogram Hypothesis. *The Behavioral and Brain Sciences* 7(2). 173–188.

- Bickerton, Derek. 1997. How to acquire language without positive evidence: What acquisitionists can learn from Creoles. In Michel DeGraff (ed.), *Language creation and language change: Creolization, diachrony, and development* (Learning, Development, and Conceptual Change), 49–74. Cambridge, MA: MIT Press.
- Bierwisch, Manfred. 1963. *Grammatik des deutschen Verbs* (studia grammatica 2). Berlin: Akademie Verlag.
- Bierwisch, Manfred. 1966. Strukturalismus: Geschichte, Probleme und Methoden. *Kursbuch* 5. 77–152.
- Bierwisch, Manfred. 1992. Grammatikforschung in der DDR: Auch ein Rückblick. *Linguistische Berichte* 139. 169–181.
- Bildhauer, Felix. 2008. *Representing information structure in an HPSG grammar of Spanish*: Universität Bremen Dissertation.
- Bildhauer, Felix. 2011. Mehrfache Vorfeldbesetzung und Informationsstruktur: Eine Bestandsaufnahme. *Deutsche Sprache* 39(4). 362–379.
- Bildhauer, Felix. 2014. Head-Driven Phrase Structure Grammar. In Carnie et al. (2014) 526–555.
- Bildhauer, Felix & Philippa Helen Cook. 2010. German multiple fronting and expected topic-hood. In Stefan Müller (ed.), *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar, Université Paris Diderot*, 68–79. Stanford, CA: CSLI Publications.
- Bird, Steven & Ewan Klein. 1994. Phonological analysis in typed feature systems. *Computational Linguistics* 20(3). 455–491.
- Bishop, Dorothy V. M. 2002. Putting language genes in perspective. *TRENDS in Genetics* 18(2). 57–59.
- Bjerre, Tavs. 2006. Object positions in a topological sentence model for Danish: A linearization-based HPSG approach. Presentation at Ph.D.-Course at Sandbjerg, Denmark. <http://www.hum.au.dk/engelsk/engsv/objectpositions/workshop/Bjerre.pdf>.
- Blackburn, Patrick & Johan Bos. 2005. *Representation and inference for natural language: A first course in computational semantics*. Stanford, CA: CSLI Publications.
- Blackburn, Patrick, Claire Gardent & Wilfried Meyer-Viol. 1993. Talking about trees. In Krauwer et al. (1993) 21–29.
- Błaszczak, Joanna & Hans-Martin Gärtner. 2005. Intonational phrasing, discontinuity, and the scope of negation. *Syntax* 8(1). 1–22.
- Blevins, James P. 2003. Passives and impersonals. *Journal of Linguistics* 39(3). 473–520.
- Block, Hans-Ulrich & Rudolf Hunze. 1986. Incremental construction of c- and f-structure in a LFG-parser. In Nagao (1986) 490–493.
- Blom, Corrien. 2005. *Complex predicates in Dutch: Synchrony and diachrony* (LOT Dissertation Series 111). Utrecht: Utrecht University.
- Bloom, Paul. 1993. Grammatical continuity in language development: The case of subjectless sentences. *Linguistic Inquiry* 24(4). 721–734.
- Boas, Hans C. 2003. *A Constructional approach to resultatives* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.

- Boas, Hans C. 2014. Lexical approaches to argument structure: Two sides of the same coin. *Theoretical Linguistics* 40(1–2). 89–112.
- Boas, Hans C. & Ivan A. Sag (eds.). 2012. *Sign-based Construction Grammar* (CSLI Lecture Notes 193). Stanford, CA: CSLI Publications.
- Bobaljik, Jonathan. 1999. Adverbs: The hierarchy paradox. *Glott International* 4(9/10). 27–28.
- Bod, Rens. 2009a. Constructions at work or at rest? *Cognitive Linguistics* 20(1). 129–134.
- Bod, Rens. 2009b. From exemplar to grammar: Integrating analogy and probability in language learning. *Cognitive Science* 33(4). 752–793.
- Bögel, Tina, Miriam Butt & Sebastian Sulger. 2008. Urdu ezafe and the morphology-syntax interface. In Butt & King (2008) 129–149. <http://csli-publications.stanford.edu/LFG/13/>.
- Bohnet, Bernd. 2010. Very high accuracy and fast Dependency Parsing is not a contradiction. In Chu-Ren Huang & Dan Jurafsky (eds.), *Proceedings of the 23rd International Conference on Computational Linguistics*, 89–97. Stroudsburg, PA, USA: Association for Computational Linguistics.
- Bolc, Leonard, Krzysztof Czuba, Anna Kupść, Małgorzata Marciniak, Agnieszka Mykowiecka & Adam Przepiórkowski. 1996. A survey of systems for implementing HPSG grammars. Tech. Rep. 814 Institute of Computer Science, Polish Academy of Sciences Warsaw, Poland. <http://www.cs.cmu.edu/~kczuba/systems-wide.ps.gz>.
- Booij, Geert E. 2002. Separable complex verbs in Dutch: A case of periphrastic word formation. In Dehé et al. (2002) 21–41.
- Booij, Geert E. 2005. Construction-Dependent Morphology. *Lingue e linguaggio* 4. 31–46.
- Booij, Geert E. 2009. Lexical integrity as a formal universal: A Constructionist view. In Sergio Scalise, Elisabetta Magni & Antonietta Bisetto (eds.), *Universals of language today* (Studies in Natural Language and Linguistic Theory 76), 83–100. Berlin: Springer Verlag.
- Booij, Geert E. 2010. Construction morphology. *Language and Linguistics Compass* 4(7). 543–555. DOI:10.1111/j.1749-818X.2010.00213.x.
- Booij, Geert E. 2012. Construction morphology. Ms. Leiden University.
- Borer, Hagit. 1994. The projection of arguments. In E. Benedicto & J. Runner (eds.), *Functional projections* (UMass Occasional Papers in Linguistics (UMOP) 17), 19–47. Massachusetts: University of Massachusetts Graduate Linguistic Student Association.
- Borer, Hagit. 2003. Exo-skeletal vs. endo-skeletal explanations: Syntactic projections and the lexicon. In John Moore & Maria Polinsky (eds.), *The nature of explanation in linguistic theory*, 31–67. Stanford, CA: CSLI Publications.
- Borer, Hagit. 2005. *Structuring sense: In name only*, vol. 1. Oxford: Oxford University Press.
- Borsley, Robert D. 1987. Subjects and complements in HPSG. Report No. CSLI-87-107 Center for the Study of Language and Information Stanford, CA.
- Borsley, Robert D. 1989. Phrase-Structure Grammar and the Barriers conception of clause structure. *Linguistics* 27(5). 843–863.
- Borsley, Robert D. 1991. *Syntactic theory: A unified approach*. London: Edward Arnold.

- Borsley, Robert D. 1999. *Syntactic theory: A unified approach*. London: Edward Arnold 2nd edn.
- Borsley, Robert D. 2005. Against ConjP. *Lingua* 115(4). 461–482.
- Borsley, Robert D. 2006. Syntactic and lexical approaches to unbounded dependencies. Essex Research Reports in Linguistics 49 University of Essex. <http://core.ac.uk/download/pdf/4187949.pdf#page=35>.
- Borsley, Robert D. 2007. Hang on again! Are we ‘on the right track’? In Andrew Radford (ed.), *Martin Atkinson – the Minimalist muse* (Essex Research Reports in Linguistics 53), 43–69. Essex: Department of Language and Linguistics, University of Essex.
- Borsley, Robert D. 2012. Don’t move! *Iberia: An International Journal of Theoretical Linguistics* 4(1). 110–139.
- Borsley, Robert D. & Kersti Börjars (eds.). 2011. *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*. Oxford, UK/Cambridge, MA: Blackwell Publishing Ltd.
- Bos, Johan. 1996. Predicate logic unplugged. In Paul J. E. Dekker & M. Stokhof (eds.), *Proceedings of the Tenth Amsterdam Colloquium*, 133–143. Amsterdam: ILLC/Department of Philosophy, University of Amsterdam.
- Boukedi, Sirine & Kais Haddar. 2014. HPSG grammar treating of different forms of Arabic coordination. *Research in Computing Science* 86: Advances in Computational Linguistics and Intelligent Decision Making, 25–41.
- Boullier, Pierre & Benoît Sagot. 2005a. Analyse syntaxique profonde à grande échelle: SxLFG. *Traitement Automatique des Langues (T.A.L.)* 46(2). 65–89.
- Boullier, Pierre & Benoît Sagot. 2005b. Efficient and robust LFG parsing: SxLFG. In *Proceedings of IWPT 2005*, 1–10. Vancouver, Canada: Association for Computational Linguistics.
- Boullier, Pierre, Benoît Sagot & Lionel Clément. 2005. Un analyseur LFG efficace pour le français: SxLfg. In *Actes de TALN 05*, 403–408. Dourdan, France.
- Bouma, Gosse. 1996. Extraposition as a nonlocal dependency. In Geert-Jan Kruijff, Glynn V. Morrill & Dick Oehrle (eds.), *Proceedings of Formal Grammar 96*, 1–14. Prag. <http://www.let.rug.nl/gosse/papers.html>.
- Bouma, Gosse, Robert Malouf & Ivan A. Sag. 2001a. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19(1). 1–65.
- Bouma, Gosse & Gertjan van Noord. 1994. Constraint-based Categorical Grammar. In James Pustejovsky (ed.), *32th Annual Meeting of the Association for Computational Linguistics. Proceedings of the conference*, 147–154. Las Cruces: Association for Computational Linguistics.
- Bouma, Gosse & Gertjan van Noord. 1998. Word order constraints on verb clusters in German and Dutch. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 43–72. San Diego: Academic Press. <http://www.let.rug.nl/~vannoord/papers/>.
- Bouma, Gosse, Gertjan van Noord & Robert Malouf. 2001b. Alpino: Wide-coverage computational analysis of Dutch. In Walter Daelemans, Khalil Sima’an, Jorn Veenstra & Jakub Zavrel (eds.), *Computational linguistics in the Netherlands 2000: Selected papers*

- from the Eleventh CLIN Meeting (Language and Computers 37), Amsterdam/New York, NY: Rodopi.
- Bragmann, Sascha. 2015. Syntactically flexible VP-idioms and the N-after-N Construction. Poster presentation at the 5th General Meeting of PARSEME, Iasi, 23–24 September 2015.
- Braine, Martin D. S. 1987. What is learned in acquiring word classes—A step toward an acquisition theory. In Brian MacWhinny (ed.), *Mechanisms of language acquisition*, 65–87. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Branco, António & Francisco Costa. 2008a. A computational grammar for deep linguistic processing of Portuguese: LXGram, version A.4.1. Tech. Rep. TR-2008-17 Universidade de Lisboa, Faculdade de Ciências, Departamento de Informática.
- Branco, António & Francisco Costa. 2008b. LXGram in the shared task ‘comparing semantic representations’ of STEP 2008. In Johan Bos & Rodolfo Delmonte (eds.), *Semantics in text processing: STEP 2008 conference proceedings*, vol. 1 (Research in Computational Semantics), 299–314. London: College Publications. <http://www.aclweb.org/anthology/W08-2224>.
- Brants, Sabine, Stefanie Dipper, Peter Eisenberg, Silvia Hansen-Schirra, Esther König, Wolfgang Lezius, Christian Rohrer, George Smith & Hans Uszkoreit. 2004. TIGER: Linguistic interpretation of a German corpus. *Research on Language and Computation* 2(4). 597–620.
- Bresnan, Joan. 1974. The position of certain clause-particles in phrase structure. *Linguistic Inquiry* 5(4). 614–619.
- Bresnan, Joan. 1978. A realistic Transformational Grammar. In M. Halle, J. Bresnan & G. A. Miller (eds.), *Linguistic theory and psychological reality*, 1–59. Cambridge, MA: MIT Press.
- Bresnan, Joan. 1982a. Control and complementation. *Linguistic Inquiry* 13(3). 343–434.
- Bresnan, Joan (ed.). 1982b. *The mental representation of grammatical relations* (MIT Press Series on Cognitive Theory and Mental Representation). Cambridge, MA/London: MIT Press.
- Bresnan, Joan. 1982c. The passive in lexical theory. In Bresnan (1982b) 3–86.
- Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford, UK/Cambridge, USA: Blackwell.
- Bresnan, Joan & Jane Grimshaw. 1978. The syntax of free relatives in English. *Linguistic Inquiry* 9. 331–392.
- Bresnan, Joan & Jonni M. Kanerva. 1989. Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry* 20(1). 1–50.
- Bresnan, Joan & Ronald M. Kaplan. 1982. Introduction: Grammars as mental representations of language. In Bresnan (1982b) xvii–lii.
- Bresnan, Joan & Sam A. Mchombo. 1995. The Lexical Integrity Principle: Evidence from Bantu. *Natural Language and Linguistic Theory* 13. 181–254.
- Bresnan, Joan & Annie Zaenen. 1990. Deep unaccusativity in LFG. In Katarzyna Dziwirek, Patrick Farrell & Errapel Mejías-Bikandi (eds.), *Grammatical relations: A cross-theoretical perspective*, 45–57. Stanford, CA: CSLI Publications.
- Brew, Chris. 1995. Stochastic HPSG. In Abney & Hinrichs (1995) 83–89.

- Briscoe, Ted J. 1997. Review of Edward P. Stabler, Jr., *The logical approach to syntax: Foundations, specifications, and implementations of theories of Government and Binding*. *Journal of Linguistics* 33(1). 223–225.
- Briscoe, Ted J. 2000. Grammatical acquisition: Inductive bias and coevolution of language and the language acquisition device. *Language* 76(2). 245–296.
- Briscoe, Ted J. & Ann Copestake. 1999. Lexical rules in constraint-based grammar. *Computational Linguistics* 25(4). 487–526.
- Bröker, Norbert. 2003. Formal foundations of Dependency Grammar. In Ágel et al. (2003b) 294–310.
- Brosziewski, Ulf. 2003. *Syntactic derivations: A nontransformational view* (Linguistische Arbeiten 470). Tübingen: Max Niemeyer Verlag.
- Brown, Keith (ed.). 2006. *The encyclopedia of language and linguistics*. Oxford: Elsevier Science Publisher B.V. (North-Holland) 2nd edn.
- Brown, Roger & Camille Hanlon. 1970. Derivational complexity and order of acquisition in child speech. In John R. Hayes (ed.), *Cognition and the development of language*, 11–53. New York: John Wiley & Sons, Inc.
- Bruening, Benjamin. 2009. Selectional asymmetries between CP and DP suggest that the DP hypothesis is wrong. In Laurel MacKenzie (ed.), *Proceedings of the 32th Annual Penn Linguistics Colloquium* (Penn Working Papers in Linguistics 15.1), 26–35. Philadelphia.
- Bryant, John. 2003. *Constructional analysis*. University of California at Berkeley MA thesis. <http://www.icsi.berkeley.edu/~jbryant/old-analyzer.html>.
- Budde, Monika. 2010. Konstruktionen integrativ: Topik-Konstruktionen als rein-syntaktisches Pendant zu den lexikalisch verankerten Komplement-Konstruktionen. Vortrag auf der Tagung Konstruktionsgrammatik: Neue Perspektiven zur Untersuchung des Deutschen und Englischen. Internationale Fachtagung an der Christian-Albrechts-Universität zu Kiel vom 18. bis 20. Februar 2010.
- Bungeroth, Jan. 2002. *A formal description of Sign Language using HPSG*. Karlsruhe Department of Computer Science, University of Stellenbosch, Lehrstuhl Informatik für Ingenieure und Naturwissenschaftler, Universität Karlsruhe (TH) Diploma thesis. <http://www-i6.informatik.rwth-aachen.de/~bungeroth/diplarb.pdf>.
- Burzio, Luigi. 1981. *Intransitive verbs and Italian auxiliaries*: MIT dissertation.
- Burzio, Luigi. 1986. *Italian syntax: A Government-Binding approach* (Studies in Natural Language and Linguistic Theory 1). Dordrecht: D. Reidel Publishing Company.
- Busemann, Stephan. 1992. *Generierung natürlicher Sprache mit generalisierten Phrasenstrukturgrammatiken*, vol. 313 (Informatik-Fachberichte). Berlin: Springer Verlag.
- Bußmann, Hadumod (ed.). 1983. *Lexikon der Sprachwissenschaft* (Kröners Taschenausgabe 452). Stuttgart: Alfred Kröner Verlag.
- Bußmann, Hadumod (ed.). 1990. *Lexikon der Sprachwissenschaft* (Kröners Taschenausgabe 452). Stuttgart: Alfred Kröner Verlag 2nd edn.
- Butt, Miriam. 2003. The light verb jungle. In C. Quinn, C. Bowerman & G. Aygen (eds.), *Papers from the Harvard/Dudley House light verb workshop* (Harvard Working Papers in Linguistics 9), 1–49. Harvard University, Department of Linguistics.

- Butt, Miriam, Stefanie Dipper, Anette Frank & Tracy Holloway King. 1999a. Writing large-scale parallel grammars for English, French and German. In Butt & King (1999). <http://csli-publications.stanford.edu/LFG/4/>.
- Butt, Miriam, Helge Dyvik, Tracy Holloway King, Hiroshi Masuichi & Christian Rohrer. 2002. The Parallel Grammar Project. In *Proceedings of COLING-2002 Workshop on Grammar Engineering and Evaluation*, 1–7.
- Butt, Miriam & Tracy Holloway King (eds.). 1999. *Proceedings of the LFG '99 conference, University of Manchester*. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/LFG/4/>.
- Butt, Miriam & Tracy Holloway King (eds.). 2005. *Proceedings of the LFG 2005 conference*. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/LFG/10/>.
- Butt, Miriam & Tracy Holloway King (eds.). 2008. *Proceedings of the LFG 2008 conference*. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/LFG/13/>.
- Butt, Miriam & Tracy Holloway King (eds.). 2009. *Proceedings of the LFG 2009 conference*. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/LFG/14/>.
- Butt, Miriam & Tracy Holloway King (eds.). 2014. *Proceedings of the LFG 2014 conference*. Stanford, CA: CSLI Publications.
- Butt, Miriam, Tracy Holloway King, María-Eugenia Niño & Frédérique Segond. 1999b. *A grammar writer's cookbook* (CSLI Lecture Notes 95). Stanford, CA: CSLI Publications.
- Butt, Miriam, Tracy Holloway King & Sebastian Roth. 2007. Urdu correlatives: Theoretical and implementational issues. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2007 conference*, 107–127. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/LFG/12/>.
- Cahill, Aoife, Michael Burke, Martin Forst, Ruth O'Donovan, Christian Rohrer, Josef van Genabith & Andy Way. 2005. Treebank-based acquisition of multilingual unification grammar resources. *Research on Language and Computation* 3(2). 247–279.
- Cahill, Aoife, Michael Burke, Ruth O'Donovan, Stefan Riezler, Josef van Genabith & Andy Way. 2008. Wide-coverage deep statistical parsing using automatic dependency structure annotation. *Computational Linguistics* 34(1). 81–124.
- Calder, Jonathan, Ewan Klein & Henk Zeevat. 1988. Unification Categorical Grammar: A concise, extendable grammar for natural language processing. In Vargha (1988) 83–86.
- Callmeier, Ulrich. 2000. PET—A platform for experimentation with efficient HPSG processing techniques. *Journal of Natural Language Engineering* 1(6). 99–108. (Special Issue on Efficient Processing with HPSG: Methods, Systems, Evaluation).
- Calzolari, Nicoletta, Claire Cardie & Pierre Isabelle (eds.). 2006. *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*. Sydney, Australia: Association for Computational Linguistics.
- Candito, Marie-Hélène. 1996. A principle-based hierarchical representation of LTAGs. In Tsuji (1996) 194–199.
- Candito, Marie-Hélène. 1998. Building parallel LTAG for French and Italian. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and*

- 17th International Conference on Computational Linguistics, 211–217. Montreal, Quebec, Canada: Association for Computational Linguistics.
- Candito, Marie-Hélène. 1999. *Organisation modulaire et paramétrable de grammaires électroniques lexicalisées. Application au français et à l’italien*: Université Paris 7 dissertation.
- Candito, Marie-Hélène & Sylvain Kahane. 1998. Can the TAG derivation tree represent a semantic graph? An answer in the light of Meaning-Text Theory. In *TAG+4*, 25–28.
- Cappelle, Bert. 2006. Particle placement and the case for “allostructions”. *Constructions online* 1(7). 1–28.
- Cappelle, Bert, Yuri Shtyrov & Friedemann Pulvermüller. 2010. *Heating up or cooling up the brain?* MEG evidence that phrasal verbs are lexical units. *Brain and Language* 115. 189–201.
- Carlson, Gregory N. & Michael K. Tanenhaus. 1988. Thematic roles and language comprehension. In Wendy Wilkins (ed.), *Thematic relations* (Syntax and Semantics 21), 263–289. San Diego: Academic Press.
- Carnie, Andrew, Yosuke Sato & Dan Siddiqi (eds.). 2014. *The Routledge handbook of syntax*. Oxford: Routledge.
- Carpenter, Bob. 1992. *The logic of typed feature structures* (Tracts in Theoretical Computer Science). Cambridge: Cambridge University Press.
- Carpenter, Bob. 1994. A natural deduction theorem prover for type-theoretic Categorical Grammars. Tech. rep. Carnegie Mellon Laboratory for Computational Linguistics. http://www.essex.ac.uk/linguistics/external/clmt/papers/cg/carp_cgparser_doc.ps.
- Carpenter, Bob. 1998. *Type-logical semantics*. Cambridge, MA/London, England: MIT Press.
- Carpenter, Bob & Gerald Penn. 1996. Efficient parsing of compiled typed attribute value logic grammars. In Harry Bunt & Masaru Tomita (eds.), *Recent advances in parsing technology* (Text, Speech and Language Technology 1), 145–168. Dordrecht: Kluwer Academic Publishers.
- Çetinoğlu, Özlem & Kemal Oflazer. 2006. Morphology-syntax interface for Turkish LFG. In Calzolari et al. (2006) 153–160.
- Chang, Nancy Chih-Lin. 2008. Constructing grammar: A computational model of the emergence of early constructions. Technical Report UCB/EECS-2009-24 Electrical Engineering and Computer Sciences, University of California at Berkeley.
- Chaves, Rui P. 2009. Construction-based cumulation and adjunct extraction. In Müller (2009d) 47–67.
- Choi, Hye-Won. 1999. *Optimizing structure in scrambling: Scrambling and information structure* (Dissertations in Linguistics). Stanford, CA: CSLI Publications.
- Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions on Information Theory* 2. 113–124.
- Chomsky, Noam. 1957. *Syntactic structures* (Janua Linguarum / Series Minor 4). The Hague/Paris: Mouton.
- Chomsky, Noam. 1959. On certain formal properties of grammars. *Information and Control* 2(2). 137–167.

参考文献

- Chomsky, Noam. 1964a. *Current issues in linguistic theory* (Janua Linguarum / Series Minor 38). The Hague/Paris: Mouton.
- Chomsky, Noam. 1964b. Degrees of grammaticalness. In Fodor & Katz (1964) 384–389.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1968. Language and the mind. *Psychology Today* 1(9). 48–68. Reprint as: Chomsky (1976a).
- Chomsky, Noam. 1970. Remarks on nominalization. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in English Transformational Grammar*, chap. 12, 184–221. Waltham, MA/Toronto/London: Ginn and Company.
- Chomsky, Noam. 1971. *Problems of knowledge and freedom*. London: Fontana.
- Chomsky, Noam. 1973. Conditions on transformations. In Stephen R. Anderson & Paul Kiparski (eds.), *A festschrift for Morris Halle*, 232–286. New York: Holt, Rinehart & Winston.
- Chomsky, Noam. 1975. *The logical structure of linguistic theory*. New York: Plenum Press.
- Chomsky, Noam. 1976a. Language and the mind. In Diane D. Borstein (ed.), *Readings in the theory of grammar: From the 17th to the 20th century*, 241–251. Cambridge, MA: Winthrop. Reprint from: Chomsky (1968).
- Chomsky, Noam. 1976b. *Reflections on language*. New York: Pantheon Books.
- Chomsky, Noam. 1977. *Essays on form and interpretation*. New York: North Holland.
- Chomsky, Noam. 1980. *Rules and representations*. Oxford: Basil Blackwell.
- Chomsky, Noam. 1981a. *Lectures on government and binding*. Dordrecht: Foris Publications.
- Chomsky, Noam. 1981b. Reply to comments of Thompson. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 295(1077). 277–281.
- Chomsky, Noam. 1982. *Some concepts and consequences of the theory of Government and Binding* (Linguistic Inquiry Monographs 5). Cambridge, MA/London, England: MIT Press.
- Chomsky, Noam. 1986a. *Barriers* (Linguistic Inquiry Monographs 13). Cambridge, MA/London, England: MIT Press.
- Chomsky, Noam. 1986b. *Knowledge of language: Its nature, origin, and use* (Convergence). New York/Westport, Connecticut/London: Praeger.
- Chomsky, Noam. 1988. *Language and problems of knowledge: The Managua lectures* (Current Studies in Linguistics 16). Cambridge, MA: MIT Press.
- Chomsky, Noam. 1989. Some notes on economy of derivation and representation. In I. Laka & Anoop Mahajan (eds.), *Functional heads and clause structure* (MIT Working Papers in Linguistics 10), 43–74. Cambridge, MA: Department of Linguistics and Philosophy.
- Chomsky, Noam. 1990. On formalization and formal linguistics. *Natural Language and Linguistic Theory* 8(1). 143–147.
- Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In Robert Freidin (ed.), *Principles and parameters in Generative Grammar*, 417–454. Cambridge, MA: MIT Press. Reprint as: Chomsky (1995b: 129–166).

- Chomsky, Noam. 1993. A Minimalist Program for linguistic theory. In Hale & Keyser (1993b) 1–52.
- Chomsky, Noam. 1995a. Bare phrase structure. In Hector Campos & Paula Kempchinsky (eds.), *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero*, 51–109. Washington, DC: Georgetown U Press.
- Chomsky, Noam. 1995b. *The Minimalist Program* (Current Studies in Linguistics 28). Cambridge, MA/London, England: MIT Press.
- Chomsky, Noam. 1998. Noam Chomsky's Minimalist Program and the philosophy of mind: An interview [with] Camilo J. Cela-Conde and Gisde Marty. *Syntax* 1(1). 19–36.
- Chomsky, Noam. 1999. Derivation by phase. MIT Occasional Papers in Linguistics 18 MIT. Reprint in: Michael Kenstowicz, ed. 2001. Ken Hale. A Life in Language. Cambridge, MA: MIT Press, 1–52.
- Chomsky, Noam. 2000. *New horizons in the study of language and mind*. Cambridge, UK: Cambridge University Press.
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2002. *On nature and language*. Cambridge, UK: Cambridge University Press.
- Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22.
- Chomsky, Noam. 2007. Approaching UG from below. In Sauerland & Gärtner (2007) 1–29.
- Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2010. Restricting stipulations: Consequences and challenges. Talk given in Stuttgart.
- Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49.
- Chomsky, Noam & George A. Miller. 1963. Introduction to the formal analysis of natural languages. In R. Duncan Luce, Robert R. Bush & Eugene Galanter (eds.), *Handbook of mathematical psychology*, vol. 2, 269–321. New York: John Wiley & Sons, Inc.
- Chouinard, Michelle M. & Eve V. Clark. 2003. Adult reformulations of child errors as negative evidence. *Journal of Child Language* 30. 637–669.
- Chrupala, Grzegorz & Josef van Genabith. 2006. Using machine-learning to assign function labels to parser output for Spanish. In Calzolari et al. (2006) 136–143.
- Chung, Sandra & James McCloskey. 1983. On the interpretation of certain island facts in GPSG. *Linguistic Inquiry* 14. 704–713.
- Church, Kenneth. 2011. A pendulum swung too far. *Linguistic Issues in Language Technology* 6(5). 1–27. <http://journals.linguisticsociety.org/elanguage/lilt/article/view/2581.html>. Special Issue on Interaction of Linguistics and Computational Linguistics.
- Cinque, Guglielmo. 1994. On the evidence for partial N movement in the Romance DP. In Guglielmo Cinque, Jan Koster, Jean-Yves Pollock, Luigi Rizzi & Raffaella Zanuttini (eds.), *Paths towards Universal Grammar: Studies in honor of Richard S. Kayne*, 85–110. Washington, D.C.: Georgetown University Press.

参考文献

- Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. New York, Oxford: Oxford University Press.
- Cinque, Guglielmo & Luigi Rizzi. 2010. The cartography of syntactic structures. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, 51–65. Oxford: Oxford University Press.
- Citko, Barbara. 2008. Missing labels. *Lingua* 118(7). 907–944.
- Clark, Alexander. 2000. Inducing syntactic categories by context distribution clustering. In *Proceedings CoNLL 2000*, 91–94. Stroudsburg, PA: Association for Computational Linguistics.
- Clark, Herbert H. & Jean E. Fox Tree. 2002. Using *uh* and *um* in spontaneous speaking. *Cognition* 84(1). 73–111.
- Clark, Herbert H. & Thomas Wasow. 1998. Repeating words in spontaneous speech. *Cognitive Psychology* 37(3). 201–242.
- Clark, Stephen & James Curran. 2007. Wide-coverage efficient statistical parsing with CCG and log-linear models. *Computational Linguistics* 33(4). 493–552.
- Clark, Stephen, Julia Hockenmaier & Mark J. Steedman. 2002. Building deep dependency structures with a wide-coverage CCG parser. In Isabelle (2002) 327–334. <http://aclanthology.info/events/acl-2002>.
- Clément, Lionel. 2009. XLFG5 documentation. Translated from French by Olivier Bonami. <http://www.xlfg.org/>.
- Clément, Lionel & Alexandra Kinyon. 2001. XLFG—An LFG parsing scheme for French. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2001 conference*, Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/LFG/6/>.
- Clément, Lionel & Alexandra Kinyon. 2003. Generating parallel multilingual LFG-TAG grammars from a MetaGrammar. In Erhard Hinrichs & Dan Roth (eds.), *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*, 184–191. Sapporo, Japan: Association for Computational Linguistics.
- Clifton, Charles Jr. & Penelope Odom. 1966. Similarity relations among certain English sentence constructions. *Psychological Monographs: General and Applied* 80(5). 1–35.
- Coch, Jose. 1996. Overview of AlethGen. In *Demonstrations and posters of the Eighth International Natural Language Generation Workshop (INLG'96)*, 25–28.
- Cook, Philippa Helen. 2001. *Coherence in German: An information structure approach*: Departments of Linguistics and German, University of Manchester dissertation.
- Cook, Philippa Helen. 2006. The datives that aren't born equal: Beneficiaries and the dative passive. In Daniel Hole, André Meinunger & Werner Abraham (eds.), *Datives and similar cases: Between argument structure and event structure*, 141–184. Amsterdam: John Benjamins Publishing Co.
- Cooper, Robin, Kuniaki Mukai & John Perry (eds.). 1990. *Situation Theory and its applications*, vol. 1 (CSLI Lecture Notes 22). Stanford, CA: CSLI Publications.
- Coopmans, Peter. 1989. Where stylistic and syntactic processes meet: Locative inversion in English. *Language* 65(4). 728–751.
- Copetake, Ann. 2002. *Implementing typed feature structure grammars* (CSLI Lecture Notes 110). Stanford, CA: CSLI Publications.

- Copestake, Ann. 2007. Applying robust semantics. In *Proceedings of the 10th Conference of the Pacific Association for Computational Linguistics (PACLING)*, 1–12.
- Copestake, Ann & Ted Briscoe. 1995. Semi-productive polysemy and sense extension. *Journal of Semantics* 12(1). 15–67.
- Copestake, Ann & Ted J. Briscoe. 1992. Lexical operations in a unification based framework. In James Pustejovsky & Sabine Bergler (eds.), *Lexical semantics and knowledge representation* (Lecture Notes in Artificial Intelligence 627), 101–119. Berlin: Springer Verlag.
- Copestake, Ann & Daniel P. Flickinger. 2000. An open-source grammar development environment and broad-coverage English grammar using HPSG. In *Proceedings of the second Linguistic Resources and Evaluation Conference*, 591–600. Athens, Greece.
- Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 4(3). 281–332.
- Correa, Nelson. 1987. An Attribute-Grammar implementation of Government-Binding Theory. In Sidner (1987) 45–51.
- Costa, Francisco & António Branco. 2010. LXGram: A deep linguistic processing grammar for Portuguese. In Thiago A.S. Pardo (ed.), *Computational processing of the Portuguese language: 9th International Conference, PROPOR 2010, Porto Alegre, RS, Brazil, April 27–30, 2010. Proceedings* (Lecture Notes in Artificial Intelligence 6001), 86–89. Berlin: Springer Verlag.
- Covington, Michael A. 1990. Parsing discontinuous constituents in Dependency Grammar. *Computational Linguistics* 16(4). 234–236.
- Crabbé, Benoit. 2005. *Représentation informatique de grammaires d’arbres fortement lexicalisées: le cas de la grammaire d’arbres adjoints*: Université Nancy 2 dissertation.
- Crain, Stephen, Drew Khlentzos & Rosalind Thornton. 2010. Universal Grammar versus language diversity. *Lingua* 120(12). 2668–2672.
- Crain, Stephen & Mineharu Nakayama. 1987. Structure dependence in grammar formation. *Language* 63(3). 522–543.
- Crain, Stephen & Mark J. Steedman. 1985. On not being led up the garden path: The use of context by the psychological syntax processor. In Dowty et al. (1985) 320–358.
- Crain, Stephen, Rosalind Thornton & Drew Khlentzos. 2009. The case of the missing generalizations. *Cognitive Linguistics* 20(1). 145–155.
- Cramer, Bart & Yi Zhang. 2009. Construction of a German HPSG grammar from a detailed treebank. In King & Santaholma (2009) 37–45. <http://www.aclweb.org/anthology/W/W09/#2600>.
- Crocker, Matthew Walter & Ian Lewin. 1992. Parsing as deduction: Rules versus principles. In Bernd Neumann (ed.), *ECAI 92. 10th European Conference on Artificial Intelligence*, 508–512. John Wiley & Sons, Inc.
- Croft, William. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Croft, William. 2003. Lexical rules vs. constructions: A false dichotomy. In Hubert Cuyckens, Thomas Berg, René Dirven & Klaus-Uwe Panther (eds.), *Motivation in language*:

- Studies in honour of Günter Radden*, 49–68. Amsterdam: John Benjamins Publishing Co.
- Croft, William. 2009. Syntax is more diverse, and evolutionary linguistics is already here. *The Behavioral and Brain Sciences* 32(5). 457–458.
- Crysmann, Berthold. 2001. Clitics and coordination in linear structure. In Birgit Gerlach & Janet Grijzenhout (eds.), *Clitics in phonology, morphology and syntax* (Linguistik Aktuell/Linguistics Today 36), 121–159. Amsterdam: John Benjamins Publishing Co.
- Crysmann, Berthold. 2002. *Constraint-based co-analysis: Portuguese cliticisation and morphology-syntax interaction in HPSG* (Saarbrücken Dissertations in Computational Linguistics and Language Technology 15). Saarbrücken: Deutsches Forschungszentrum für Künstliche Intelligenz und Universität des Saarlandes.
- Crysmann, Berthold. 2003. On the efficient implementation of German verb placement in HPSG. In *Proceedings of RANLP 2003*, 112–116. Borovets, Bulgaria.
- Crysmann, Berthold. 2004. Underspecification of intersective modifier attachment: Some arguments from German. In Müller (2004e) 378–392. <http://csli-publications.stanford.edu/HPSG/2004/>.
- Crysmann, Berthold. 2005a. An inflectional approach to Hausa final vowel shortening. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2004*, 73–112. Dordrecht: Kluwer Academic Publishers.
- Crysmann, Berthold. 2005b. Relative clause extraposition in German: An efficient and portable implementation. *Research on Language and Computation* 1(3). 61–82.
- Crysmann, Berthold. 2005c. Syncretism in German: A unified approach to underspecification, indeterminacy, and likeness of case. In Stefan Müller (ed.), *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon*, 91–107. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/HPSG/2005/>.
- Crysmann, Berthold. 2008. An asymmetric theory of peripheral sharing in HPSG: Conjunction reduction and coordination of unlikes. In Jäger et al. (2008) 47–62.
- Crysmann, Berthold. 2009. Autosegmental representations in an HPSG of Hausa. In King & Santaholma (2009) 28–36. <http://www.aclweb.org/anthology/W/W09/#2600>.
- Crysmann, Berthold. 2011. A unified account of Hausa genitive constructions. In Philippe de Groote, Markus Egg & Laura Kallmeyer (eds.), *Formal Grammar: 14th International Conference, FG 2009, Bordeaux, France, July 25–26, 2009, revised selected papers* (Lecture Notes in Artificial Intelligence 5591), 102–117. Berlin: Springer Verlag.
- Crysmann, Berthold. 2012. HaG: A computational grammar of Hausa. In Michael R. Marlo, Nikki B. Adams, Christopher R. Green, Michelle Morrison & Tristan M. Purvis (eds.), *Selected proceedings of the 42nd Annual Conference on African Linguistics (ACAL 42)*, 321–337. Somerville, MA: Cascadilla Press. <http://www.lingref.com/cpp/acal/42/paper2780.pdf>.
- Crysmann, Berthold. 2016. Representing morphological tone in a computational grammar of Hausa. *Journal of Language Modelling* 3(2). 463–512.
- Culicover, Peter W. 1999. *Syntactic nuts: Hard cases, syntactic theory, and language acquisition*, vol. 1 (Foundations of Syntax). Oxford: Oxford University Press.

- Culicover, Peter W. & Ray S. Jackendoff. 2005. *Simpler Syntax*. Oxford: Oxford University Press.
- Culy, Christopher. 1985. The complexity of the vocabulary of Bambara. *Linguistics and Philosophy* 8. 345–351.
- Curtiss, Susan. 1977. *Genie: A psycholinguistic study of a modern-day “wild child”*. New York: Academic Press.
- Dąbrowska, Ewa. 2001. From formula to schema: The acquisition of English questions. *Cognitive Linguistics* 11(1–2). 83–102.
- Dąbrowska, Ewa. 2004. *Language, mind and brain: Some psychological and neurological constraints on theories of grammar*. Washington, D.C.: Georgetown University Press.
- Dahl, Östen & Viveka Velupillai. 2013a. The past tense. In Dryer & Haspelmath (2013). <http://wals.info/chapter/66>.
- Dahl, Östen & Viveka Velupillai. 2013b. Perfective/imperfective aspect. In Dryer & Haspelmath (2013). <http://wals.info/chapter/65>.
- Dahllöf, Mats. 2002. Token dependency semantics and the paratactic analysis of intensional constructions. *Journal of Semantics* 19(4). 333–368.
- Dahllöf, Mats. 2003. Two reports on computational syntax and semantics. Reports from Uppsala University (RUUL) 36 Department of Linguistics. <http://stp.ling.uu.se/~matsd/pub/ruul36.pdf>.
- Dalrymple, Mary. 1993. *The syntax of anaphoric binding* (CSLI Lecture Notes 36). Stanford, CA: CSLI Publications.
- Dalrymple, Mary (ed.). 1999. *Semantics and syntax in Lexical Functional Grammar: The Resource Logic approach*. Cambridge, MA: MIT Press.
- Dalrymple, Mary. 2001. *Lexical Functional Grammar* (Syntax and Semantics 34). New York: Academic Press.
- Dalrymple, Mary. 2006. Lexical Functional Grammar. In Brown (2006) 82–94.
- Dalrymple, Mary, Ronald M. Kaplan & Tracy Holloway King. 2004. Linguistic generalizations over descriptions. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2004 conference*, 199–208. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/LFG/9/>.
- Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell III & Annie Zaenen (eds.). 1995. *Formal issues in Lexical-Functional Grammar* (CSLI Lecture Notes 47). Stanford, CA: CSLI Publications.
- Dalrymple, Mary, John Lamping & Vijay Saraswat. 1993. LFG semantics via constraints. In Krauwer et al. (1993) 97–105. DOI:10.3115/976744.976757.
- Dalrymple, Mary, Maria Liakata & Lisa Mackie. 2006. Tokenization and morphological analysis for Malagasy. *Computational Linguistics and Chinese Language Processing* 11(4). 315–332.
- Dalrymple, Mary & Helge Lødrup. 2000. The grammatical functions of complement clauses. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2000 conference*, Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/LFG/5/lfg00dalrympl-lodrup.pdf>.

- Davidson, Donald. 1967. The logical form of action sentences. In Nicholas Rescher (ed.), *The logic of decision and action*, 81–95. Pittsburg: Pittsburg University Press.
- Davis, Anthony R. 1996. *Lexical semantics and linking in the hierarchical lexicon*: Stanford University dissertation.
- Davis, Anthony R. & Jean-Pierre Koenig. 2000. Linking as constraints on word classes in a hierarchical lexicon. *Language* 76(1). 56–91.
- De Kuthy, Kordula. 2000. *Discontinuous NPs in German — A case study of the interaction of syntax, semantics and pragmatics*. Saarbrücken: Universität des Saarlandes dissertation.
- De Kuthy, Kordula. 2001. Splitting PPs from NPs. In Meurers & Kiss (2001) 31–76.
- De Kuthy, Kordula. 2002. *Discontinuous NPs in German* (Studies in Constraint-Based Lexicalism 14). Stanford, CA: CSLI Publications.
- De Kuthy, Kordula, Vanessa Metcalf & Walt Detmar Meurers. 2004. Documentation of the implementation of the Milca English Resource Grammar in the Trale system. Ohio State University, ms.
- De Kuthy, Kordula & Walt Detmar Meurers. 2001. On partial constituent fronting in German. *Journal of Comparative Germanic Linguistics* 3(3). 143–205.
- De Kuthy, Kordula & Walt Detmar Meurers. 2003a. Dealing with optional complements in HPSG-based grammar implementations. In Müller (2003c) 88–96. <http://csli-publications.stanford.edu/HPSG/2003/>.
- De Kuthy, Kordula & Walt Detmar Meurers. 2003b. The secret life of focus exponents, and what it tells us about fronted verbal projections. In Müller (2003c) 97–110. <http://csli-publications.stanford.edu/HPSG/2003/>.
- de Saussure, Ferdinand. 1916a. *Cours de linguistique générale* (Bibliothèque Scientifique Payot). Paris: Payot. Edited by Charles Bally and Albert Sechehaye.
- de Saussure, Ferdinand. 1916b. *Grundfragen der allgemeinen Sprachwissenschaft*. Berlin: Walter de Gruyter & Co. 2nd edition 1967.
- Dehé, Nicole, Ray S. Jackendoff, Andrew McIntyre & Silke Urban (eds.). 2002. *Verb-particle explorations* (Interface Explorations 1). Berlin: Mouton de Gruyter.
- Delmonte, Rodolfo. 1990. Semantic parsing with an LFG-based lexicon and conceptual representations. *Computers and the Humanities* 24(5–6). 461–488.
- Demberg, Vera & Frank Keller. 2008. A psycholinguistically motivated version of TAG. In *Proceedings of the 9th International Workshop on Tree Adjoining Grammars and Related Formalisms TAG+9*, 25–32. Tübingen.
- Demske, Ulrike. 2001. *Merkmale und Relationen: Diachrone Studien zur Nominalphrase des Deutschen* (Studia Linguistica Germanica 56). Berlin: Walter de Gruyter Verlag.
- den Besten, Hans. 1983. On the interaction of root transformations and lexical deletive rules. In Werner Abraham (ed.), *On the formal syntax of the Westgermania: Papers from the 3rd Groningen Grammar Talks, Groningen, January 1981* (Linguistik Aktuell/Linguistics Today 3), 47–131. Amsterdam: John Benjamins Publishing Co.
- den Besten, Hans. 1985. Some remarks on the Ergative Hypothesis. In Abraham (1985) 53–74.

- Deppermann, Arnulf. 2006. Construction Grammar – eine Grammatik für die Interaktion? In Arnulf Deppermann, Reinhard Fiehler & Thomas Spranz-Fogasy (eds.), *Grammatik und Interaktion*, 43–65. Radolfzell: Verlag für Gesprächsforschung.
- Derbyshire, Desmond C. 1979. *Hixkaryana* (Lingua Descriptive Series 1). Amsterdam: North Holland.
- Devlin, Keith. 1992. *Logic and information*. Cambridge: Cambridge University Press.
- Dhonnchadha, E. Uí & Josef van Genabith. 2006. A part-of-speech tagger for Irish using finite-state morphology and Constraint Grammar disambiguation. In *Proceedings of Irec'06*, 2241–2244.
- Diesing, Molly. 1992. *Indefinites*. Cambridge, MA/London, England: MIT Press.
- Dione, Cheikh Mouhamadou Bamba. 2013. Handling Wolof Clitics in LFG. In Christine Meklenborg Salvesen & Hans Petter Helland (eds.), *Challenging clitics* (Linguistik Aktuell/Linguistics Today 206), 87–118. Amsterdam: John Benjamins Publishing Co.
- Dione, Cheikh Mouhamadou Bamba. 2014. An LFG approach to Wolof cleft constructions. In Butt & King (2014) 157–176.
- Dipper, Stefanie. 2003. *Implementing and documenting large-scale grammars – German LFG*: IMS, University of Stuttgart dissertation. Arbeitspapiere des Instituts für Maschinelle Sprachverarbeitung (AIMS), Volume 9, Number 1.
- Donati, C. 2006. On *wh*-head-movement. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), *Wh-movement: Moving on* (Current Studies in Linguistics 42), 21–46. Cambridge, MA: MIT Press.
- Donohue, Cathryn & Ivan A. Sag. 1999. Domains in Warlpiri. In *Sixth International Conference on HPSG–Abstracts. 04–06 August 1999*, 101–106. Edinburgh.
- Doran, Christine, Beth Ann Hockey, Anoop Sarkar, Bangalore Srinivas & Fei Xia. 2000. Evolution of the XTAG system. In Abeillé & Rambow (2000b) 371–403.
- Dörre, Jochen & Michael Dorna. 1993. CUF: A formalism for linguistic knowledge representation. DYANA 2 deliverable R.1.2A IMS Stuttgart, Germany.
- Dörre, Jochen & Roland Seiffert. 1991. A formalism for natural language – STUF. In Herzog & Rollinger (1991) 29–38.
- Dowty, David. 1997. Non-constituent coordination, wrapping, and Multimodal Categorical Grammars: Syntactic form as logical form. In Maria Luisa Dalla Chiara, Kees Doets, Daniele Mundici & Johan Van Benthem (eds.), *Structures and norms in science* (Synthese Library 260), 347–368. Springer.
- Dowty, David R. 1978. Governed transformations as lexical rules in a Montague Grammar. *Linguistic Inquiry* 9(3). 393–426.
- Dowty, David R. 1979. *Word meaning and Montague Grammar* (Synthese Language Library 7). Dordrecht: D. Reidel Publishing Company.
- Dowty, David R. 1988. Type raising, functional composition, and nonconstituent coordination. In Oehrle et al. (1988) 153–198.
- Dowty, David R. 1989. On the semantic content of the notion ‘thematic role’. In Gennaro Chierchia, Barbara H. Partee & Raymond Turner (eds.), *Properties, types and meaning*, vol. 2 (Studies in Linguistics and Philosophy), 69–130. Dordrecht: Kluwer Academic Publishers.

参考文献

- Dowty, David R. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619.
- Dowty, David R. 2003. The dual analysis of adjuncts and complements in Categorical Grammar. In Lang et al. (2003) 33–66.
- Dowty, David R., Lauri Karttunen & Arnold M. Zwicky (eds.). 1985. *Natural language processing*. Cambridge, UK: Cambridge University Press.
- Dras, Mark, François Lareau, Benjamin Börschinger, Robert Dale, Yasaman Motazedi, Owen Rambow, Myfany Turpin & Morgan Ulinski. 2012. Complex predicates in Arernte. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2012 conference*, 177–197. Stanford, CA: CSLI Publications.
- Drellishak, Scott. 2009. *Widespread but not universal: Improving the typological coverage of the Grammar Matrix*. University of Washington Doctoral dissertation.
- Drosdowski, Günther. 1984. *Duden: Grammatik der deutschen Gegenwartssprache*, vol. 4. Mannheim, Wien, Zürich: Dudenverlag 4th edn.
- Drosdowski, Günther. 1995. *Duden: Die Grammatik*, vol. 4. Mannheim, Leipzig, Wien, Zürich: Dudenverlag 5th edn.
- Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68(1). 81–138.
- Dryer, Matthew S. 1997. Are grammatical relations universal? In Joan Bybee, John Haiman & Sandra Thompson (eds.), *Essays on language function and language type: Dedicated to T. Givon*, 115–143. Amsterdam: John Benjamins Publishing Co.
- Dryer, Matthew S. 2013a. Order of adposition and noun phrase. In Dryer & Haspelmath (2013). <http://wals.info/chapter/85>.
- Dryer, Matthew S. 2013b. Order of object and verb. In Dryer & Haspelmath (2013). <http://wals.info/chapter/83>.
- Dryer, Matthew S. 2013c. Order of subject, object and verb. In Dryer & Haspelmath (2013). <http://wals.info/chapter/81>.
- Dryer, Matthew S. & Martin Haspelmath (eds.). 2013. *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/>.
- Dürscheid, Christa. 1989. *Zur Vorfeldbesetzung in deutschen Verbzweit-Strukturen* (FOKUS 1). Trier: Wissenschaftlicher Verlag.
- Dyvik, Helge, Paul Meurer & Victoria Rosén. 2005. LFG, Minimal Recursion Semantics and translation. Paper presented at the LFG conference 2005.
- Egg, Markus. 1999. Derivation and resolution of ambiguities in *wieder*-sentences. In Paul J. E. Dekker (ed.), *Proceedings of the 12th Amsterdam Colloquium*, 109–114.
- Eisele, Andreas & Jochen Dorre. 1986. A Lexical Functional Grammar system in Prolog. In Nagao (1986) 551–553.
- Eisenberg, Peter. 1992. Platos Problem und die Lernbarkeit der Syntax. In Suchsland (1992) 371–378.
- Eisenberg, Peter. 1994a. German. In Ekkehard König & Johan van der Auwera (eds.), *The Germanic languages* (Routledge Language Family Descriptions), 349–387. London / New York: Routledge.

- Eisenberg, Peter. 1994b. *Grundriß der deutschen Grammatik*. Stuttgart, Weimar: Verlag J. B. Metzler 3rd edn.
- Eisenberg, Peter. 2004. *Grundriß der deutschen Grammatik*, vol. 2. Der Satz. Stuttgart, Weimar: Verlag J. B. Metzler 2nd edn.
- Eisenberg, Peter, Jörg Peters, Peter Gallmann, Cathrine Fabricius-Hansen, Damaris Nübling, Irmhild Barz, Thomas A. Fritz & Reinhard Fiehler. 2005. *Duden: Die Grammatik*, vol. 4. Mannheim, Leipzig, Wien, Zürich: Dudenverlag 7th edn.
- Ellefson, Michelle R. & Morten Christiansen. 2000. Subjacency constraints without Universal Grammar: Evidence from artificial language learning and connectionist modeling. In *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*, 645–650. Mahwah, NJ: Lawrence Erlbaum Associates.
- Elman, Jeffrey L. 1993. Learning and development in neural networks: The importance of starting small. *Cognition* 48(1). 71–99.
- Elman, Jeffrey L., Elizabeth A. Bates, Mark H. Johnson, Annette Karmiloff-Smith, Domenico Parisi & Kim Plunkett. 1996. *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: Bradford Books/MIT Press.
- Embick, David. 2004. On the structure of resultative participles in English. *Linguistic Inquiry* 35(3). 355–392.
- Emirkanian, Louisette, Lyne Da Sylva & Lorne H. Bouchard. 1996. The implementation of a computational grammar of French using the Grammar Development Environment. In Tsuji (1996) 1024–1027.
- Engdahl, Elisabet & Enric Vallduví. 1996. Information packaging in HPSG. In Claire Grover & Enric Vallduví (eds.), *Edinburgh Working Papers in Cognitive Science*, vol. 12: *Studies in HPSG*, chap. 1, 1–32. Scotland: Centre for Cognitive Science, University of Edinburgh. <ftp://ftp.cogsci.ed.ac.uk/pub/CCS-WPs/wp-12.ps.gz>.
- Engel, Ulrich. 1970. Regeln zur Wortstellung. Forschungsberichte des Instituts für deutsche Sprache 5 Institut für deutsche Sprache Mannheim.
- Engel, Ulrich. 1977. *Syntax der deutschen Gegenwartssprache*, vol. 22 (Grundlagen der Germanistik). Berlin: Erich Schmidt Verlag.
- Engel, Ulrich. 1996. Tesnière mißverstanden. In Gertrud Gréciano & Helmut Schumacher (eds.), *Lucien Tesnière – Syntaxe Structurale et Opérations Mentales. Akten des deutsch-französischen Kolloquiums anlässlich der 100. Wiederkehr seines Geburtstages. Strasbourg 1993* (Linguistische Arbeiten 348), 53–61. Tübingen: Max Niemeyer Verlag.
- Engel, Ulrich. 2014. Die dependenzielle Verbgrammatik (DVG). In Hagemann & Staffeldt (2014) 43–62.
- Engelberg, Stefan, Anke Holler & Kristel Proost (eds.). 2011. *Sprachliches Wissen zwischen Lexikon und Grammatik* (Institut für Deutsche Sprache, Jahrbuch 2010). Berlin: de Gruyter.
- Erbach, Gregor. 1995. ProFIT: Prolog with features, inheritance and templates. In Abney & Hinrichs (1995) 180–187.
- Ernst, Thomas. 1992. The phrase structure of English negation. *The Linguistic Review* 9(2). 109–144.

参考文献

- Eroms, Hans-Werner. 1985. Eine reine Dependenzgrammatik für das Deutsche. *Deutsche Sprache* 13. 306–326.
- Eroms, Hans-Werner. 1987. Passiv und Passivfunktionen im Rahmen einer Dependenzgrammatik. In Centre de Recherche en Linguistique Germanique (Nice) (ed.), *Das Passiv im Deutschen* (Linguistische Arbeiten 183), 73–95. Tübingen: Max Niemeyer Verlag.
- Eroms, Hans-Werner. 2000. *Syntax der deutschen Sprache* (de Gruyter Studienbuch). Berlin: Walter de Gruyter Verlag.
- Eroms, Hans-Werner & Hans Jürgen Heringer. 2003. Dependenz und lineare Ordnung. In Ágel et al. (2003b) 247–263.
- Eroms, Hans-Werner, Gerhard Stickel & Gisela Zifonun (eds.). 1997. *Grammatik der deutschen Sprache*, vol. 7 (Schriften des Instituts für deutsche Sprache). Berlin: Walter de Gruyter.
- Erteschik-Shir, Nomi. 1973. *On the nature of island constraints*. Cambridge, MA: MIT dissertation.
- Erteschik-Shir, Nomi. 1981. More on extractability from quasi-NPs. *Linguistic Inquiry* 12(4). 665–670.
- Erteschik-Shir, Nomi & Shalom Lappin. 1979. Dominance and the functional explanation of island phenomena. *Theoretical Linguistics* 6(1–3). 41–86.
- Estigarribia, Bruno. 2009. Facilitation by variation: Right-to-left learning of English yes/no questions. *Cognitive Science* 34(1). 68–93.
- Evans, Nicholas & Stephen C. Levinson. 2009a. The myth of language universals: Language diversity and its importance for cognitive science. *The Behavioral and Brain Sciences* 32(5). 429–448.
- Evans, Nicholas & Stephen C. Levinson. 2009b. With diversity in mind: Freeing the language sciences from Universal Grammar. *The Behavioral and Brain Sciences* 32(5). 472–492.
- Evans, Roger. 1985. ProGram: A development tool for GPSG grammars. *Linguistics* 23(2). 213–244.
- Everett, Daniel L. 2005. Cultural constraints on grammar and cognition in Pirahã. *Current Anthropology* 46(4). 621–646.
- Everett, Daniel L. 2009. Pirahã culture and grammar: A response to some criticisms. *Language* 85(2). 405–442.
- Evers, Arnold. 1975. *The transformational cycle in Dutch and German*: University of Utrecht dissertation.
- Faaß, Gertrud. 2010. *A morphosyntactic description of Northern Sotho as a basis for an automated translation from Northern Sotho into English*. Pretoria, South Africa: University of Pretoria dissertation. <http://hdl.handle.net/2263/28569>.
- Fabregas, Antonio, Tom Stroik & Michael Putnam. 2016. Is simplest merge too simple? Ms. Penn State University.
- Falk, Yehuda N. 1984. The English auxiliary system: A Lexical-Functional analysis. *Language* 60(3). 483–509.

- Fan, Zhenzhen, Sanghoun Song & Francis Bond. 2015. An HPSG-based shared-grammar for the Chinese languages: ZHONG []. In Bender et al. (2015) 17–24.
- Fang, Ji & Tracy Holloway King. 2007. An LFG Chinese grammar for machine use. In King & Bender (2007) 144–160. <http://csli-publications.stanford.edu/GEAF/2007/>.
- Fanselow, Gisbert. 1981. *Zur Syntax und Semantik der Nominalkomposition* (Linguistische Arbeiten 107). Tübingen: Max Niemeyer Verlag.
- Fanselow, Gisbert. 1987. *Konfiguralität* (Studien zur deutschen Grammatik 29). Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.
- Fanselow, Gisbert. 1988. Aufspaltung von NPn und das Problem der ‚freien‘ Wortstellung. *Linguistische Berichte* 114. 91–113.
- Fanselow, Gisbert. 1990. Scrambling as NP-movement. In Grewendorf & Sternefeld (1990) 113–140.
- Fanselow, Gisbert. 1992a. „Ergative“ Verben und die Struktur des deutschen Mittelfelds. In Ludger Hoffmann (ed.), *Deutsche Syntax: Ansichten und Aussichten* (Institut für deutsche Sprache, Jahrbuch 1991), 276–303. Berlin: de Gruyter.
- Fanselow, Gisbert. 1992b. Zur biologischen Autonomie der Grammatik. In Suchsland (1992) 335–356.
- Fanselow, Gisbert. 1993. Die Rückkehr der Basisgenerierer. *Groninger Arbeiten zur Germanistischen Linguistik* 36. 1–74.
- Fanselow, Gisbert. 2000a. Does constituent length predict German word order in the Middle Field? In Josef Bayer & Christine Römer (eds.), *Von der Philologie zur Grammatiktheorie: Peter Suchsland zum 65. Geburtstag*, 63–77. Tübingen: Max Niemeyer Verlag.
- Fanselow, Gisbert. 2000b. Optimal exceptions. In Barbara Stiebels & Dieter Wunderlich (eds.), *The lexicon in focus* (studia grammatica 45), 173–209. Berlin: Akademie Verlag.
- Fanselow, Gisbert. 2001. Features, θ -roles, and free constituent order. *Linguistic Inquiry* 32(3). 405–437.
- Fanselow, Gisbert. 2002. Against remnant VP-movement. In Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers & Hans-Martin Gärtner (eds.), *Dimensions of movement: From features to remnants* (Linguistik Aktuell/Linguistics Today 48), 91–127. Amsterdam: John Benjamins Publishing Co.
- Fanselow, Gisbert. 2003a. Free constituent order: A Minimalist interface account. *Folia Linguistica* 37(1–2). 191–231.
- Fanselow, Gisbert. 2003b. Münchhausen-style head movement and the analysis of verb second. In Anoop Mahajan (ed.), *Proceedings of the workshop on head movement* (UCLA Working Papers in Linguistics 10), Los Angeles: UCLA, Linguistics Department.
- Fanselow, Gisbert. 2003c. Zur Generierung der Abfolge der Satzglieder im Deutschen. *Neue Beiträge zur Germanistik* 112. 3–47.
- Fanselow, Gisbert. 2004a. Cyclic phonology-syntax-interaction: PPT Movement in German (and other languages). In Shinichiro Ishihara, Michaela Schmitz & Anne Schwarz (eds.), *Interdisciplinary studies on information structure* (Working Papers of the SFB 632 1), 1–42. Potsdam: Universitätsverlag.

- Fanselow, Gisbert. 2004b. Münchhausen-style head movement and the analysis of verb second. In Ralf Vogel (ed.), *Three papers on German verb movement* (Linguistics in Potsdam 22), 9–49. Universität Potsdam.
- Fanselow, Gisbert. 2006. On pure syntax (uncontaminated by information structure). In Patrick Brandt & Eric Fuss (eds.), *Form, structure and grammar: A festschrift presented to Günther Grewendorf on occasion of his 60th birthday* (Studia grammatica 63), 137–157. Berlin: Akademie Verlag.
- Fanselow, Gisbert. 2009. Die (generative) Syntax in den Zeiten der Empiriediskussion. *Zeitschrift für Sprachwissenschaft* 28(1). 133–139.
- Fanselow, Gisbert & Sascha W. Felix. 1987. *Sprachtheorie 2. Die Rektions- und Bindungstheorie* (UTB für Wissenschaft: Uni-Taschenbücher 1442). Tübingen: A. Francke Verlag GmbH.
- Fanselow, Gisbert & Sascha W. Felix (eds.). 1990. *Strukturen und Merkmale syntaktischer Kategorien* (Studien zur deutschen Grammatik 39). Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.
- Fanselow, Gisbert, Matthias Schlesewsky, Damir Cavar & Reinhold Kliegl. 1999. Optimal parsing, syntactic parsing preferences, and Optimality Theory. Rutgers Optimality Archive (ROA) 367 Universität Potsdam. <http://roa.rutgers.edu/view.php3?roa=367>.
- Feldhaus, Anke. 1997. Eine HPSG-Analyse ausgewählter Phänomene des deutschen *w*-Fragesatzes. Working Papers of the Institute for Logic and Linguistics 27 Institute for Logic and Linguistics IBM Scientific Center Heidelberg.
- Feldman, Jerome. 1972. Some decidability results on grammatical inference and complexity. *Information and Control* 20(3). 244–262.
- Fillmore, Charles J. 1968. The case for case. In Emmon Bach & Robert T. Harms (eds.), *Universals of linguistic theory*, 1–88. New York: Holt, Rinehart, and Winston.
- Fillmore, Charles J. 1971. Plädoyer für Kasus. In Werner Abraham (ed.), *Kasustheorie* (Schwerpunkte Linguistik und Kommunikationswissenschaft 2), 1–118. Frankfurt/Main: Athenäum.
- Fillmore, Charles J. 1988. The mechanisms of “Construction Grammar”. In Shelley Axmaker, Annie Jaissner & Helen Singmaster (eds.), *Proceedings of the 14th Annual Meeting of the Berkeley Linguistics Society*, 35–55. Berkeley, CA: Berkeley Linguistics Society.
- Fillmore, Charles J. 1999. Inversion and Constructional inheritance. In Webelhuth et al. (1999) 113–128.
- Fillmore, Charles J., Paul Kay & Mary Catherine O’Connor. 1988. Regularity and idiomatcity in grammatical constructions: The case of *let alone*. *Language* 64(3). 501–538.
- Fillmore, Charles J., Russell R. Lee-Goldmann & Russell Rhomieux. 2012. The FrameNet constructicon. In Boas & Sag (2012) 309–372.
- Fischer, Ingrid & Martina Keil. 1996. Parsing decomposable idioms. In Tsuji (1996) 388–393.
- Fischer, Kerstin & Anatol Stefanowitsch (eds.). 2006. *Konstruktionsgrammatik: Von der Anwendung zur Theorie* (Stauffenburg Linguistik 40). Tübingen: Stauffenburg Verlag.

- Fisher, Simon E. & Gary F. Marcus. 2005. The eloquent ape: Genes, brains and the evolution of language. *Nature Reviews Genetics* 7(1). 9–20.
- Fisher, Simon E., Faraneh Vargha-Khadem, Kate E. Watkins, Anthony P. Monaco & Marcus E. Pembrey. 1998. Localisation of a gene implicated in a severe speech and language disorder. *Nature Genetics* 18(2). 168–170.
- Fitch, W. Tecumseh. 2010. Three meanings of “recursion”: Key distinctions for biolinguistics. In Richard K. Larson, Viviane Déprez & Hiroko Yamakido (eds.), *The evolution of human language: Biolinguistic perspectives* (Approaches to the Evolution of Language 2), 73–90. Cambridge, UK: Cambridge University Press.
- Fitch, W. Tecumseh, Marc D. Hauser & Noam Chomsky. 2005. The evolution of the language faculty: Clarifications and implications. *Cognition* 97(2). 179–210.
- Flickinger, Daniel P. 1983. Lexical heads and phrasal gaps. In *Proceedings of the West Coast Conference on Formal Linguistics*, vol. 2, Stanford University Linguistics Dept.
- Flickinger, Daniel P. 1987. *Lexical rules in the hierarchical lexicon*: Stanford University dissertation.
- Flickinger, Daniel P. 2000. On building a more efficient grammar by exploiting types. *Natural Language Engineering* 6(1). 15–28.
- Flickinger, Daniel P. 2008. Transparent heads. In Müller (2008b) 87–94. <http://csli-publications.stanford.edu/HPSG/2008/abstr-flickinger.shtml>.
- Flickinger, Daniel P. & Emily M. Bender. 2003. Compositional semantics in a multilingual grammar resource. In Bender et al. (2003) 33–42.
- Flickinger, Daniel P., Ann Copestake & Ivan A. Sag. 2000. HPSG analysis of English. In Wahlster (2000b) 254–263.
- Flickinger, Daniel P., Carl J. Pollard & Thomas Wasow. 1985. Structure-sharing in lexical representation. In William C. Mann (ed.), *Proceedings of the Twenty-Third Annual Meeting of the Association for Computational Linguistics*, 262–267. Chicago, IL: Association for Computational Linguistics.
- Fodor, Janet Dean. 1998a. Parsing to learn. *Journal of Psycholinguistic Research* 27(3). 339–374.
- Fodor, Janet Dean. 1998b. Unambiguous triggers. *Linguistic Inquiry* 29(1). 1–36.
- Fodor, Janet Dean. 2001. Parameters and the periphery: Reflections on syntactic nuts. *Journal of Linguistics* 37. 367–392.
- Fodor, Jerry A., Thomas G. Bever & Merrill F. Garrett. 1974. *The psychology of language: An introduction to psycholinguistics and Generative Grammar*. New York: McGraw-Hill Book Co.
- Fodor, Jerry A. & Jerrold J. Katz (eds.). 1964. *The structure of language*. Englewood Cliffs, NJ: Prentice-Hall.
- Fokkens, Antske. 2011. Metagrammar engineering: Towards systematic exploration of implemented grammars. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 1066–1076. Portland, Oregon, USA: Association for Computational Linguistics. <http://www.aclweb.org/anthology/P11-1107>.

参考文献

- Fokkens, Antske, Laurie Poulson & Emily M. Bender. 2009. Inflectional morphology in Turkish VP coordination. In Müller (2009d) 110–130.
- Fong, Sandiway. 1991. *Computational properties of principle-based grammatical theories*: MIT Artificial Intelligence Lab dissertation.
- Fong, Sandiway. 2014. Unification and efficient computation in the Minimalist Program. In L. Francis & L. Laurent (eds.), *Language and recursion*, 129–138. Berlin: Springer Verlag.
- Fong, Sandiway & Jason Ginsburg. 2012. Computation with doubling constituents: Pronouns and antecedents in Phase Theory. In Anna Maria Di Sciullo (ed.), *Towards a Biolinguistic understanding of grammar: Essays on interfaces* (Linguistik Aktuell/Linguistics Today 194), 303–338. Amsterdam: John Benjamins Publishing Co.
- Fordham, Andrew & Matthew Walter Crocker. 1994. Parsing with principles and probabilities. In Judith L. Klavans Philip Resnik (ed.), *The balancing act: Combining symbolic and statistical approaches to language*, Las Cruces, New Mexico, USA: Association for Computational Linguistics.
- Forst, Martin. 2006. COMP in (parallel) grammar writing. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2006 conference*, Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/LFG/11/lfg06forst.pdf>.
- Forst, Martin & Christian Rohrer. 2009. Problems of German VP coordination. In Butt & King (2009) 297–316. <http://csli-publications.stanford.edu/LFG/14/>.
- Fortmann, Christian. 1996. *Konstituentenbewegung in der DP-Struktur: Zur funktionalen Analyse der Nominalphrase im Deutschen* (Linguistische Arbeiten 347). Tübingen: Max Niemeyer Verlag.
- Fourquet, Jean. 1957. Review of: Heinz Anstock: Deutsche Syntax – Lehr- und Übungsbuch. *Wirkendes Wort* 8. 120–122.
- Fourquet, Jean. 1970. *Prolegomena zu einer deutschen Grammatik* (Sprache der Gegenwart – Schriften des Instituts für deutsche Sprache in Mannheim 7). Düsseldorf: Pädagogischer Verlag Schwann.
- Fouvry, Frederik. 2003. Lexicon acquisition with a large-coverage unification-based grammar. In *Proceedings of EACL 03, 10th Conference of the European Chapter of the Association for Computational Linguistics, research notes and demos, April 12–17, 2003, Budapest, Hungary*, 87–90.
- Fraj, Fériel Ben, Chiraz Zribi & Mohamed Ben Ahmed. 2008. ArabTAG: A Tree Adjoining Grammar for Arabic syntactic structures. In *Proceedings of the International Arab Conference on Information Technology*, Sfax, Tunisia.
- Francis, Elaine J. & Laura A. Michaelis (eds.). 2003. *Mismatch: Form-function incongruity and the architecture of grammar* (CSLI Lecture Notes 163). Stanford, CA: CSLI Publications.
- Frank, Anette. 1994. Verb second by lexical rule or by underspecification. Arbeitspapiere des SFB 340 No. 43 IBM Deutschland GmbH Heidelberg. <ftp://ftp.ims.uni-stuttgart.de/pub/papers/anette/v2-usp.ps.gz>.
- Frank, Anette. 1996. Eine LFG-Grammatik des Französischen. In Berman & Frank (1996) 97–244.

- Frank, Anette. 2006. (Discourse-) functional analysis of asymmetric coordination. In Miriam Butt, Mary Dalrymple & Tracy Holloway King (eds.), *Intelligent linguistic architectures: Variations on themes by Ronald M. Kaplan*, 259–285. Stanford, CA: CSLI Publications.
- Frank, Anette & Uwe Reyle. 1995. Principle based semantics for HPSG. In Abney & Hinrichs (1995) 9–16.
- Frank, Anette & Annie Zaenen. 2002. Tense in LFG: Syntax and morphology. In Hans Kamp & Uwe Reyle (eds.), *How we say when it happens: Contributions to the theory of temporal reference in natural language*, 17–52. Tübingen: Max Niemeyer Verlag. Reprint as: Frank & Zaenen (2004).
- Frank, Anette & Annie Zaenen. 2004. Tense in LFG: Syntax and morphology. In Louisa Sadler & Andrew Spencer (eds.), *Projecting morphology*, 23–66. Stanford, CA: CSLI Publications.
- Frank, Robert. 2002. *Phrase structure composition and syntactic dependencies* (Current Studies in Linguistics 38). Cambridge, MA/London: MIT Press.
- Franks, Steven. 1995. *Parameters in Slavic morphosyntax*. New York, Oxford: Oxford University Press.
- Frazier, Lyn. 1985. Syntactic complexity. In Dowty et al. (1985) 129–189.
- Frazier, Lyn & Charles Clifton. 1996. *Construal*. Cambridge, MA: MIT Press.
- Freidin, Robert. 1975. The analysis of passives. *Language* 51(2). 384–405.
- Freidin, Robert. 1997. Review article: The Minimalist Program. *Language* 73(3). 571–582.
- Freidin, Robert. 2009. A note on methodology in linguistics. *The Behavioral and Brain Sciences* 32(5). 454–455.
- Freudenthal, Daniel, Julian M. Pine, Javier Aguado-Orea & Fernand Gobet. 2007. Modeling the developmental patterning of finiteness marking in English, Dutch, German, and Spanish using MOSAIC. *Cognitive Science* 31(2). 311–341.
- Freudenthal, Daniel, Julian M. Pine & Fernand Gobet. 2006. Modeling the development of children’s use of optional infinitives in Dutch and English using MOSAIC. *Cognitive Science* 30(2). 277–310.
- Freudenthal, Daniel, Julian M. Pine & Fernand Gobet. 2009. Simulating the referential properties of Dutch, German, and English root infinitives in MOSAIC. *Language Learning and Development* 5(1). 1–29.
- Frey, Werner. 1993. *Syntaktische Bedingungen für die semantische Interpretation: Über Bindung, implizite Argumente und Skopus* (studia grammatica 35). Berlin: Akademie Verlag.
- Frey, Werner. 2000. Über die syntaktische Position der Satztopiks im Deutschen. In Ewald Lang, Marzena Rochon, Kerstin Schwabe & Oliver Teuber (eds.), *Issues on topics* (ZAS Papers in Linguistics 20), 137–172. Berlin: ZAS, Humboldt-Universität zu Berlin.
- Frey, Werner. 2001. About the whereabouts of indefinites. *Theoretical Linguistics* 27(2/3). 137–161. DOI:10.1515/thli.2001.27.2-3.137. Special Issue: NP Interpretation and Information Structure, Edited by Klaus von Heusinger and Kerstin Schwabe.

参考文献

- Frey, Werner. 2004a. The grammar-pragmatics interface and the German prefield. *Forschungsprogramm Sprache und Pragmatik* 52 Germanistisches Institut der Universität Lund.
- Frey, Werner. 2004b. A medial topic position for German. *Linguistische Berichte* 198. 153–190.
- Frey, Werner. 2005. Pragmatic properties of certain German and English left peripheral constructions. *Linguistics* 43(1). 89–129.
- Frey, Werner & Hans-Martin Gärtner. 2002. On the treatment of scrambling and adjunction in Minimalist Grammars. In Jäger et al. (2002) 41–52.
- Frey, Werner & Uwe Reyle. 1983a. Lexical Functional Grammar und Diskursrepräsentationstheorie als Grundlagen eines sprachverarbeitenden Systems. *Linguistische Berichte* 88. 79–100.
- Frey, Werner & Uwe Reyle. 1983b. A Prolog implementation of Lexical Functional Grammar as a base for a natural language processing system. In Antonio Zampolli (ed.), *First Conference of the European Chapter of the Association for Computational Linguistics: Proceedings of the conference*, 52–57. Pisa, Italy: Association for Computational Linguistics. <http://aclweb.org/anthology/E/E83/>.
- Fried, Mirjam. 2013. Principles of constructional change. In Hoffmann & Trousdale (2013) 419–437.
- Fried, Mirjam. 2015. Construction Grammar. In Kiss & Alexiadou (2015) 974–1003.
- Friederici, Angela D. 2009. Pathways to language: Fiber tracts in the human brain. *Trends in Cognitive Sciences* 13(4). 175–181.
- Friedman, Joyce. 1969. Applications of a computer system for Transformational Grammar. In Research Group for Quantitative Linguistics (ed.), *Proceedings of COLING 69*, 1–27.
- Friedman, Joyce, Thomas H. Bredt, Robert W. Doran, Bary W. Pollack & Theodore S. Martner. 1971. *A computer model of Transformational Grammar* (Mathematical Linguistics and Automatic Language Processing 9). New York: Elsevier.
- Fries, Norbert. 1988. Über das Null-Topik im Deutschen. *Forschungsprogramm Sprache und Pragmatik* 3 Germanistisches Institut der Universität Lund.
- Fukui, Naoki & Margaret Speas. 1986. Specifiers and projection. In N. Fukui, T. R. Rapoport & E. Sagey (eds.), *Papers in theoretical linguistics* (MIT Working Papers 8), 128–172. Cambridge, MA: MIT.
- Gaifman, Haim. 1965. Dependency systems and phrase-structure systems. *Information and Control* 8. 304–397.
- Gardner, R. Allen. 1957. Probability-learning with two and three choices. *The American Journal of Psychology* 70(2). 174–185.
- Gärtner, Hans-Martin & Jens Michaelis. 2007. Some remarks on locality conditions and Minimalist Grammars. In Sauerland & Gärtner (2007) 161–195.
- Gärtner, Hans-Martin & Markus Steinbach. 1997. Anmerkungen zur Vorfeldphobie pronominaler Elemente. In Franz-Josef d’Avis & Uli Lutz (eds.), *Zur Satzstruktur im Deutschen* (Arbeitspapiere des SFB 340 No. 90), 1–30. Tübingen: Eberhard-Karls-Universität Tübingen.

- Gazdar, Gerald. 1981a. On syntactic categories. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 295(1077). 267–283.
- Gazdar, Gerald. 1981b. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12. 155–184.
- Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag. 1985. *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.
- Gazdar, Gerald, Geoffrey K. Pullum, Bob Carpenter, Ewan Klein, Thomas E. Hukari & Robert D. Levine. 1988. Category structures. *Computational Linguistics* 14(1). 1–19.
- Geach, Peter Thomas. 1970. A program for syntax. *Synthese* 22. 3–17.
- Geißler, Stefan & Tibor Kiss. 1994. Erläuterungen zur Umsetzung einer HPSG im Basis-formalismus STUF III. Tech. Rep. 19 IBM Informationssysteme GmbH – Institut für Logik und Linguistik (Verbundvorhaben Verbmobil) Heidelberg.
- Gerdes, Kim. 2002a. DTAG? In *Proceedings of the Sixth International Workshop on Tree Adjoining Grammar and Related Frameworks (TAG+6)*, 242–251. Università di Venezia.
- Gerdes, Kim. 2002b. *Topologie et grammaires formelles de l'allemand*: Ecole doctorale Science du langage, UFR de linguistique, Université Paris 7 dissertation.
- Gerdes, Kim & Sylvain Kahane. 2001. Word order in German: A formal Dependency Grammar using a topological hierarchy. In *Proceedings of the 39th Annual Meeting on Association for Computational Linguistics*, 220–227. Stroudsburg, PA, USA: Association for Computational Linguistics. DOI:10.3115/1073012.1073041.
- Gerken, LouAnn. 1991. The metrical basis for children's subjectless sentences. *Journal of Memory and Language* 30. 431–451.
- Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68(1). 1–76.
- Gibson, Edward & James Thomas. 1999. Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes* 14(3). 225–248.
- Gibson, Edward & Kenneth Wexler. 1994. Triggers. *Linguistic Inquiry* 25(3). 407–454.
- Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives* (CSLI Lecture Notes 123). Stanford, CA: CSLI Publications.
- Gold, Mark E. 1967. Language identification in the limit. *Information and Control* 10(5). 447–474.
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar approach to argument structure* (Cognitive Theory of Language and Culture). Chicago/London: The University of Chicago Press.
- Goldberg, Adele E. 2003a. Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7(5). 219–224.
- Goldberg, Adele E. 2003b. Words by default: The Persian Complex Predicate Construction. In Francis & Michaelis (2003) 117–146.
- Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language* (Oxford Linguistics). Oxford, New York: Oxford University Press.

- Goldberg, Adele E. 2009. Constructions work. [response]. *Cognitive Linguistics* 20(1). 201–224.
- Goldberg, Adele E. 2013a. Argument structure Constructions vs. lexical rules or derivational verb templates. *Mind and Language* 28(4). 435–465.
- Goldberg, Adele E. 2013b. Explanation and Constructions: Response to Adger. *Mind and Language* 28(4). 479–491.
- Goldberg, Adele E. 2014. Fitting a slim dime between the verb template and argument structure construction approaches. *Theoretical Linguistics* 40(1–2). 113–135.
- Goldberg, Adele E., Devin Casenhiser & Nitya Sethuraman. 2004. Learning argument structure generalizations. *Cognitive Linguistics* 15(3). 289–316.
- Goldberg, Adele E. & Ray S. Jackendoff. 2004. The English resultative as a family of Constructions. *Language* 80(3). 532–568.
- Gopnik, Myrna & Martha B. Cargo. 1991. Familial aggregation of a developmental language disorder. *Cognition* 39(1). 1–50.
- Gordon, Peter. 1986. Level ordering in lexical development. *Cognition* 21(2). 73–93.
- Gosch, Angela, Gabriele Ståding & Rainer Pankau. 1994. Linguistic abilities in children with Williams-Beuren Syndrome. *American Journal of Medical Genetics* 52(3). 291–296.
- Götz, Thilo, Walt Detmar Meurers & Dale Gerdemann. 1997. The ConTroll manual: (ConTroll v.1.0 beta, XTroll v.5.0 beta). User's manual Seminar für Sprachwissenschaft Universität Tübingen. <http://www.sfs.uni-tuebingen.de/control/code.html>.
- Grebe, Paul & Helmut Gipper. 1966. *Duden: Grammatik der deutschen Gegenwartssprache*, vol. 4. Mannheim, Wien, Zürich: Dudenverlag 2nd edn.
- Green, Georgia M. 2011. Modelling grammar growth: Universal Grammar without innate principles or parameters. In Borsley & Börjars (2011) 378–403.
- Grewendorf, Günther. 1983. Reflexivierungen in deutschen A.c.I.-Konstruktionen – kein transformationsgrammatisches Dilemma mehr. *Groninger Arbeiten zur Germanistischen Linguistik* 23. 120–196.
- Grewendorf, Günther. 1985. Anaphern bei Objekt-Koreferenz im Deutschen: Ein Problem für die Rektions-Bindungs-Theorie. In Abraham (1985) 137–171.
- Grewendorf, Günther. 1987. Kohärenz und Restrukturierung: Zu verbalen Komplexen im Deutschen. In Brigitte Asbach-Schnitker & Johannes Roggenhofer (eds.), *Neuere Forschungen zur Wortbildung und Histographie: Festgabe für Herbert E. Brekle zum 50. Geburtstag* (Tübinger Beiträge zur Linguistik 284), 123–144. Tübingen: Gunter Narr Verlag.
- Grewendorf, Günther. 1988. *Aspekte der deutschen Syntax: Eine Rektions-Bindungs-Analyse* (Studien zur deutschen Grammatik 33). Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.
- Grewendorf, Günther. 1989. *Ergativity in German* (Studies in Generative Grammar 35). Dordrecht: Foris Publications.
- Grewendorf, Günther. 1993. German: A grammatical sketch. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax – Ein internationales Handbuch zeitgenössischer Forschung*, vol. 9.2 (Handbücher zur Sprach- und Kommunikationswissenschaft), 1288–1319. Berlin: Walter de Gruyter Verlag.

- Grewendorf, Günther. 2002. *Minimalistische Syntax* (UTB für Wissenschaft: Uni-Taschenbücher 2313). Tübingen, Basel: A. Francke Verlag GmbH.
- Grewendorf, Günther & Wolfgang Sternefeld (eds.). 1990. *Scrambling and barriers* (Linguistik Aktuell/Linguistics Today 5). Amsterdam: John Benjamins Publishing Co.
- Grewendorf, Günther. 2009. The left clausal periphery: Clitic left dislocation in Italian and left dislocation in German. In Benjamin Shear, Philippa Helen Cook, Werner Frey & Claudia Maienborn (eds.), *Dislocated elements in discourse: Syntactic, semantic, and pragmatic perspectives* (Routledge Studies in Germanic Linguistics), 49–94. New York: Routledge.
- Grimshaw, Jane. 1986. Subadjacency and the S/S' Parameter. *Linguistic Inquiry* 17(2). 364–369.
- Grimshaw, Jane. 1997. Projections, heads, and optimality. *Linguistic Inquiry* 28. 373–422.
- Grinberg, Dennis, John D. Lafferty & Daniel Dominic Sleator. 1995. A robust parsing algorithm for Link Grammars. In *Proceedings of the Fourth International Workshop on Parsing Technologies*, <http://arxiv.org/abs/cmp-lg/9508003>. Also as Carnegie Mellon University Computer Science Technical Report CMU-CS-95-125.
- Groos, Anneke & Henk van Riemsdijk. 1981. Matching effects in free relatives: A parameter of core grammar. In A. Belletti, L. Brandi & L. Rizzi (eds.), *Theory of markedness in Generative Grammar*, 171–216. Pisa: Scuola Normale Superiore.
- Groß, Thomas M. & Timothy Osborne. 2009. Toward a practical Dependency Grammar theory of discontinuities. *SKY Journal of Linguistics* 22. 43–90.
- Groß, Thomas Michael. 2003. Dependency Grammar's limits – and ways of extending them. In Ägel et al. (2003b) 331–351.
- Grosu, Alexander. 1973. On the status of the so-called Right Roof Constraint. *Language* 49(2). 294–311.
- Grover, Claire, John Carroll & Ted J. Briscoe. 1993. The Alvey Natural Language Tools grammar (4th release). Technical Report 284 Computer Laboratory, Cambridge University, UK.
- Grubačić, Emilija. 1965. *Untersuchungen zur Frage der Wortstellung in der deutschen Prosadichtung der letzten Jahrzehnte*. Zagreb: Philosophische Fakultät dissertation.
- Gruber, Jeffrey. 1965. *Studies in lexical relations*: MIT dissertation.
- Gunji, Takao. 1986. Subcategorization and word order. In William J. Poser (ed.), *Papers from the Second International Workshop on Japanese Syntax*, 1–21. Stanford, CA: CSLI Publications.
- Günther, Carsten, Claudia Maienborn & Andrea Schopp. 1999. The processing of information structure. In Peter Bosch & Rob van der Sandt (eds.), *Focus: Linguistic, cognitive, and computational perspectives* (Studies in Natural Language Processing), 18–42. Cambridge, UK: Cambridge University Press. Rev. papers orig. presented at a conference held 1994, Schloss Wolfsbrunnen, Germany.
- Guo, Yuqing, Haifeng Wang & Josef van Genabith. 2007. Recovering non-local dependencies for Chinese. In *Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Natural Language Learning, (EMNLP-CoNLL 2007)*, 257–266. Prague, Czech Republic: Association for Computational Linguistics.

- Guzmán Naranjo, Matías. 2015. Unifying everything: Integrating quantitative effects into formal models of grammar. In *Proceedings of the 6th Conference on Quantitative Investigations in Theoretical Linguistics*, DOI:10.15496/publikation-8636.
- Haddar, Kais, Sirine Boukedi & Ines Zalila. 2010. Construction of an HPSG grammar for the Arabic language and its specification in TDL. *International Journal on Information and Communication Technologies* 3(3). 52–64.
- Haegeman, Liliane. 1994. *Introduction to Government and Binding Theory* (Blackwell Textbooks in Linguistics 1). Oxford, UK/Cambridge, USA: Blackwell Publishing Ltd 2nd edn.
- Haegeman, Liliane. 1995. *The syntax of negation*. Cambridge, UK: Cambridge University Press.
- Haftka, Brigitta. 1995. Syntactic positions for topic and contrastive focus in the German middlefield. In Inga Kohlhof, Susanne Winkler & Hans-Bernhard Drubig (eds.), *Proceedings of the Göttingen Focus Workshop, 17 DGfS, March 1–3* (Arbeitspapiere des SFB 340 No. 69), 137–157. Eberhard-Karls-Universität Tübingen.
- Haftka, Brigitta. 1996. Deutsch ist eine V/2-Sprache mit Verbendstellung und freier Wortfolge. In Ewald Lang & Gisela Zifonun (eds.), *Deutsch – typologisch* (Institut für deutsche Sprache, Jahrbuch 1995), 121–141. Berlin: Walter de Gruyter.
- Hagemann, Jörg & Sven Staffeldt (eds.). 2014. *Syntaxtheorien: Analysen im Vergleich* (Stauffenburg Einführungen 28). Tübingen: Stauffenburg Verlag.
- Hagen, Kristin, Janne Bondi Johannessen & Anders Nøklestad. 2000. A constraint-based tagger for Norwegian. In C.-E. Lindberg & S. N. Lund (eds.), *17th Scandinavian Conference of Linguistic, Odense*, vol. I (Odense Working Papers in Language and Communication 19), 1–15.
- Hahn, Michael. 2011. Null conjuncts and bound pronouns in Arabic. In Müller (2011b) 60–80. <http://csli-publications.stanford.edu/HPSG/2011/>.
- Haider, Hubert. 1982. Dependenz und Konfigurationen: Zur deutschen V-Projektion. *Groninger Arbeiten zur Germanistischen Linguistik* 21. 1–60.
- Haider, Hubert. 1984. Was zu haben ist und was zu sein hat – Bemerkungen zum Infinitiv. *Papiere zur Linguistik* 30(1). 23–36.
- Haider, Hubert. 1985a. The case of German. In Toman (1985) 23–64.
- Haider, Hubert. 1985b. Über *sein* oder nicht *sein*: Zur Grammatik des Pronomens *sich*. In Abraham (1985) 223–254.
- Haider, Hubert. 1986a. Fehlende Argumente: Vom Passiv zu kohärenten Infinitiven. *Linguistische Berichte* 101. 3–33.
- Haider, Hubert. 1986b. Nicht-sententiale Infinitive. *Groninger Arbeiten zur Germanistischen Linguistik* 28. 73–114.
- Haider, Hubert. 1990a. Pro-bleme? In Fanselow & Felix (1990) 121–143.
- Haider, Hubert. 1990b. Topicalization and other puzzles of German syntax. In Grewendorf & Sternefeld (1990) 93–112.
- Haider, Hubert. 1991. Fakultativ kohärente Infinitivkonstruktionen im Deutschen. Arbeitspapiere des SFB 340 No. 17 IBM Deutschland GmbH Heidelberg.

- Haider, Hubert. 1993. *Deutsche Syntax – generativ: Vorstudien zur Theorie einer projektiven Grammatik* (Tübinger Beiträge zur Linguistik 325). Tübingen: Gunter Narr Verlag.
- Haider, Hubert. 1994. (Un-)heimliche Subjekte: Anmerkungen zur Pro-drop Causa, im Anschluß an die Lektüre von Osvaldo Jaeggli & Kenneth J. Safir, eds., *The Null Subject Parameter*. *Linguistische Berichte* 153. 372–385.
- Haider, Hubert. 1995. Studies on phrase structure and economy. Arbeitspapiere des SFB 340 No. 70 Universität Stuttgart Stuttgart.
- Haider, Hubert. 1997a. Projective economy: On the minimal functional structure of the German clause. In Werner Abraham & Elly van Gelderen (eds.), *German: Syntactic problems—Problematic syntax* (Linguistische Arbeiten 374), 83–103. Tübingen: Max Niemeyer Verlag.
- Haider, Hubert. 1997b. Typological implications of a directionality constraint on projections. In Artemis Alexiadou & T. Alan Hall (eds.), *Studies on Universal Grammar and typological variation* (Linguistik Aktuell/Linguistics Today 13), 17–33. Amsterdam: John Benjamins Publishing Co.
- Haider, Hubert. 1999. The license to license: Structural case plus economy yields Burzio's Generalization. In Eric Reuland (ed.), *Arguments and case: Explaining Burzio's Generalization* (Linguistik Aktuell/Linguistics Today 34), 31–55. Amsterdam: John Benjamins Publishing Co.
- Haider, Hubert. 2000. OV is more basic than VO. In Peter Svenonius (ed.), *The derivation of VO and OV*, 45–67. Amsterdam: John Benjamins Publishing Co.
- Haider, Hubert. 2001. Parametrisierung in der Generativen Grammatik. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Sprachtypologie und sprachliche Universalien – Language typology and language universals: Ein internationales Handbuch – An international handbook*, 283–294. Berlin: Mouton de Gruyter.
- Haider, Hubert. 2014. Scientific ideology in grammar theory. Ms. Universität Salzburg, Dept. of Linguistics and Centre for Cognitive Neuroscience.
- Haider, Hubert & Klaus Netter (eds.). 1991. *Representation and derivation in the theory of grammar* (Studies in Natural Language and Linguistic Theory 22). Dordrecht: Kluwer Academic Publishers.
- Hajičová, Eva & Petr Sgall. 2003. Dependency Syntax in Functional Generative Description. In Ágel et al. (2003b) 570–592.
- Hakuta, Kenji, Ellen Bialystok & Edward Wiley. 2003. Critical evidence: A test of the Critical-Period Hypothesis for second-language acquisition. *Psychological Science* 14(1). 31–38.
- Hale, Kenneth. 1976. The adjoined relative clause in Australia. In R.M.W. Dixon (ed.), *Grammatical categories of Australian languages* (Linguistic Series 22), 78–105. New Jersey: Humanities Press.
- Hale, Kenneth & Samuel Jay Keyser. 1993a. On argument structure and the lexical expression of syntactic relations. In Hale & Keyser (1993b) 53–109.

- Hale, Kenneth & Samuel Jay Keyser (eds.). 1993b. *The view from building 20: essays in linguistics in honor of Sylvain Bromberger* (Current Studies in Linguistics 24). Cambridge, MA/London: MIT Press.
- Hale, Kenneth & Samuel Jay Keyser. 1997. On the complex nature of simple predicators. In Alex Alsina, Joan Bresnan & Peter Sells (eds.), *Complex predicates* (CSLI Lecture Notes 64), 29–65. Stanford, CA: CSLI Publications.
- Han, Chung-hye, Juntae Yoon, Nari Kim & Martha Palmer. 2000. A feature-based Lexicalized Tree Adjoining Grammar for Korean. Technical Report IRCS-00-04 University of Pennsylvania Institute for Research in Cognitive Science. http://repository.upenn.edu/ircs_reports/35/.
- Harbour, Daniel. 2011. Mythomania? Methods and morals from ‘The myth of language universals’. *Lingua* 121(12). 1820–1830.
- Harley, Heidi & Rolf Noyer. 2000. Formal versus encyclopedic properties of vocabulary: Evidence from nominalizations. In Peeters (2000) 349–374.
- Harman, Gilbert. 1963. Generative grammars without transformation rules: A defence of phrase structure. *Language* 39. 597–616.
- Harris, Zellig S. 1957. Co-occurrence and transformation in linguistic structure. *Language* 33(3). 283–340.
- Haspelmath, Martin. 2008. Parametric versus functional explanations of syntactic universals. In T. Biberauer (ed.), *The limits of syntactic variation*, 75–107. Amsterdam: John Benjamins Publishing Co.
- Haspelmath, Martin. 2009. The best-supported language universals refer to scalar patterns deriving from processing costs. *The Behavioral and Brain Sciences* 32(5). 457–458.
- Haspelmath, Martin. 2010a. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687.
- Haspelmath, Martin. 2010b. Framework-free grammatical theory. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of grammatical analysis*, 341–365. Oxford: Oxford University Press.
- Haspelmath, Martin. 2010c. The interplay between comparative concepts and descriptive categories (reply to Newmeyer). *Language* 86(3). 696–699.
- Haugereid, Petter. 2007. Decomposed phrasal constructions. In Müller (2007d) 120–129. <http://csli-publications.stanford.edu/HPSG/2007/>.
- Haugereid, Petter. 2009. *Phrasal subconstructions: A Constructionalist grammar design, exemplified with Norwegian and English*: Norwegian University of Science and Technology dissertation.
- Haugereid, Petter, Nurit Melnik & Shuly Wintner. 2013. Nonverbal predicates in Modern Hebrew. In Stefan Müller (ed.), *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin*, 69–89. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/HPSG/2013/hmw.pdf>.
- Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298. 1569–1579. DOI:10.1126/science.298.5598.1569.

- Hausser, Roland. 1992. Complexity in left-associative grammar. *Theoretical Computer Science* 106(2). 283–308.
- Hawkins, John A. 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75(2). 244–285.
- Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
- Hays, David G. 1964. Dependency Theory: A formalism and some observations. *Language* 40(4). 511–525.
- Hays, David G. & T. W. Ziehe. 1960. Studies in machine translation: 10–Russian sentence-structure determination. Tech. rep. Rand Corporation.
- Hedeland, Hanna, Thomas Schmidt & Kai Wörner (eds.). 2011. *Multilingual resources and multilingual applications: Proceedings of the Conference of the German Society for Computational Linguistics and Language Technology (GSCL) 2011* (Arbeiten zur Mehrsprachigkeit/Working Papers in Multilingualism, Folge B/Series B 96). Hamburg: Universität Hamburg.
- Heinecke, Johannes, Jürgen Kunze, Wolfgang Menzel & Ingo Schröder. 1998. Eliminative parsing with graded constraints. In Isabelle (1998) 526–530. DOI:10.3115/980845.980953.
- Heinz, Wolfgang & Johannes Matiassek. 1994. Argument structure and case assignment in German. In Nerbonne et al. (1994b) 199–236.
- Helbig, Gerhard & Joachim Buscha. 1969. *Deutsche Grammatik: Ein Handbuch für den Ausländerunterricht*. Leipzig: VEB Verlag Enzyklopädie.
- Helbig, Gerhard & Joachim Buscha. 1998. *Deutsche Grammatik: Ein Handbuch für den Ausländerunterricht*. Leipzig Berlin München: Langenscheidt Verlag Enzyklopädie 18th edn.
- Helbig, Gerhard & Wolfgang Schenkel. 1969. *Wörterbuch zur Valenz und Distribution deutscher Verben*. Leipzig: VEB Bibliographisches Institut Leipzig.
- Hellan, Lars. 1986. The headedness of NPs in Norwegian. In Peter Muysken & Henk van Riemsdijk (eds.), *Features and projections*, 89–122. Dordrecht/Cinnaminson, U.S.A.: Foris Publications.
- Hellan, Lars. 2007. On ‘deep evaluation’ for individual computational grammars and for cross-framework comparison. In King & Bender (2007) 161–181. <http://csli-publications.stanford.edu/GEAF/2007/>.
- Hellan, Lars & Dorothee Beermann. 2006. The ‘specifier’ in an HPSG grammar implementation of Norwegian. In S. Werner (ed.), *Proceedings of the 15th NODALIDA Conference, Joensuu 2005* (Ling@JoY: University of Joensuu electronic publications in linguistics and language technology 1), 57–64. Joensuu: University of Joensuu.
- Hellan, Lars & Petter Haugereid. 2003. Norsource – An exercise in the Matrix Grammar building design. In Bender et al. (2003).
- Hellwig, Peter. 1978. PLAIN – Ein Programmsystem zur Sprachbeschreibung und maschinellen Sprachbearbeitung. *Sprache und Datenverarbeitung* 1(2). 16–31.
- Hellwig, Peter. 1986. Dependency Unification Grammar. In Nagao (1986) 195–198.
- Hellwig, Peter. 2003. Dependency Unification Grammar. In Ägel et al. (2003b) 593–635.

- Hellwig, Peter. 2006. Parsing with Dependency Grammars. In Vilmos Ágel, Ludwig M. Eichinger, Hans Werner Eroms, Peter Hellwig, Hans Jürgen Heringer & Henning Lobin (eds.), *Dependenz und Valenz / Dependency and valency: Ein internationales Handbuch der zeitgenössischen Forschung / An international handbook of contemporary research*, vol. 25.2 (Handbücher zur Sprach- und Kommunikationswissenschaft), 1081–1109. Berlin: Walter de Gruyter.
- Her, One-Soon, Dan Higinbotham & Joseph Pentheroudakis. 1991. An LFG-based machine translation system. *Computer Processing of Chinese and Oriental Languages* 5(3–4). 285–297.
- Heringer, Hans-Jürgen. 1996. *Deutsche Syntax dependentiell* (Stauffenburg Linguistik). Tübingen: Stauffenburg Verlag.
- Herzog, Otthein & Claus-Rainer Rollinger (eds.). 1991. *Text understanding in LILOG* (Lecture Notes in Artificial Intelligence 546). Berlin: Springer Verlag.
- Higinbotham, James. 1985. On semantics. *Linguistic Inquiry* 16(4). 547–593.
- Hinrichs, Erhard W., Walt Detmar Meurers, Frank Richter, Manfred Sailer & Heike Winhart (eds.). 1997. *Ein HPSG-Fragment des Deutschen. Teil 1: Theorie*, vol. No. 95 (Arbeitspapiere des SFB 340). Tübingen: Eberhard-Karls-Universität Tübingen.
- Hinrichs, Erhard W. & Tsuneko Nakazawa. 1989a. *Aspects of German VP structure* (SfS-Report-01-93). Tübingen: Eberhard-Karls-Universität Tübingen.
- Hinrichs, Erhard W. & Tsuneko Nakazawa. 1989b. Flipped out: AUX in German. In Hinrichs & Nakazawa (1989a).
- Hinrichs, Erhard W. & Tsuneko Nakazawa. 1989c. Subcategorization and VP structure in German. In Hinrichs & Nakazawa (1989a).
- Hinrichs, Erhard W. & Tsuneko Nakazawa. 1994. Linearizing AUXs in German verbal complexes. In Nerbonne et al. (1994b) 11–38.
- Hinterhölzel, Roland. 2004. Language change versus grammar change: What diachronic data reveal about the interaction between core grammar and periphery. In Carola Trips & Eric Fuß (eds.), *Diachronic clues to synchronic grammar*, 131–160. Amsterdam: John Benjamins Publishing Co.
- Hoberg, Ursula. 1981. *Die Wortstellung in der geschriebenen deutschen Gegenwartssprache* (Heutiges Deutsch. Linguistische Grundlagen. Forschungen des Instituts für deutsche Sprache 10). München: Max Hueber Verlag.
- Hockett, Charles F. 1960. The origin of speech. *Scientific American* 203. 88–96.
- Hoeksema, Jack. 1991. Theoretische Aspekten van Partikelvooropplaatsing. *TABU Bulletin voor Taalwetenschap* 21(1). 18–26.
- Hoffman, Beryl Ann. 1995. *The computational analysis of the syntax and interpretation of “free” word order in Turkish*: University of Pennsylvania dissertation.
- Hoffmann, Thomas & Graeme Trousdale (eds.). 2013. *The Oxford handbook of Construction Grammar* (Oxford Handbooks). Oxford: Oxford University Press.
- Höhle, Tilman N. 1978. *Lexikalische Syntax: Die Aktiv-Passiv-Relation und andere Infinitkonstruktionen im Deutschen* (Linguistische Arbeiten 67). Tübingen: Max Niemeyer Verlag.

- Höhle, Tilman N. 1982. Explikationen für „normale Betonung“ und „normale Wortstellung“. In Abraham (1982) 75–153. Republished as Höhle (2016d).
- Höhle, Tilman N. 1983. Topologische Felder. Köln, ms, Published as Höhle (2016h).
- Höhle, Tilman N. 1986. Der Begriff „Mittelfeld“, Anmerkungen über die Theorie der topologischen Felder. In Walter Weiss, Herbert Ernst Wiegand & Marga Reis (eds.), *Akten des VII. Kongresses der Internationalen Vereinigung für germanische Sprach- und Literaturwissenschaft. Göttingen 1985. Band 3. Textlinguistik contra Stilistik? – Wortschatz und Wörterbuch – Grammatische oder pragmatische Organisation von Rede?* (Kontroversen, alte und neue 4), 329–340. Tübingen: Max Niemeyer Verlag. Republished as Höhle (2016c).
- Höhle, Tilman N. 1988. Verum-Fokus. Netzwerk Sprache und Pragmatik 5 Universität Lund, Germananistisches Institut Lund. Republished as Höhle (2016i).
- Höhle, Tilman N. 1991a. On reconstruction and coordination. In Haider & Netter (1991) 139–197. Republished as Höhle (2016e).
- Höhle, Tilman N. 1991b. Projektionsstufen bei V-Projektionen: Bemerkungen zu F/T. Ms. Published as Höhle (2016f).
- Höhle, Tilman N. 1994. Spuren in HPSG. Vortrag auf der GGS-Tagung in Tübingen am 14. Mai 1994, published as Höhle (2016g).
- Höhle, Tilman N. 1997. Vorangestellte Verben und Komplementierer sind eine natürliche Klasse. In Christa Dürscheid, Karl Heinz Ramers & Monika Schwarz (eds.), *Sprache im Fokus: Festschrift für Heinz Vater zum 65. Geburtstag*, 107–120. Tübingen: Max Niemeyer Verlag. Republished as Höhle (2016j).
- Höhle, Tilman N. 1999. An architecture for phonology. In Robert D. Borsley & Adam Przepiórkowski (eds.), *Slavic in Head-Driven Phrase Structure Grammar*, 61–90. Stanford, CA: CSLI Publications. Republished as Höhle (2016a).
- Höhle, Tilman N. 2016a. An architecture for phonology. In Höhle (2016b). Originally published as Höhle (1999).
- Höhle, Tilman N. 2016b. *Beiträge zur Grammatik des Deutschen* (Classics in Linguistics). Berlin: Language Science Press. In Preparation.
- Höhle, Tilman N. 2016c. Der Begriff „Mittelfeld“, Anmerkungen über die Theorie der topologischen Felder. In Höhle (2016b). First published as Höhle (1986).
- Höhle, Tilman N. 2016d. Explikationen für „normale Betonung“ und „normale Wortstellung“. In Höhle (2016b). In Preparation.
- Höhle, Tilman N. 2016e. On reconstruction and coordination. In Höhle (2016b). In Preparation.
- Höhle, Tilman N. 2016f. Projektionsstufen bei V-Projektionen: Bemerkungen zu F/T. In Höhle (2016b). First circulated in 1991.
- Höhle, Tilman N. 2016g. Spuren in HPSG. In Höhle (2016b). Vortrag auf der GGS-Tagung in Tübingen am 14. Mai 1994.
- Höhle, Tilman N. 2016h. Topologische Felder. In Höhle (2016b). First circulated as draft in 1983.
- Höhle, Tilman N. 2016i. Verum-Fokus. In Höhle (2016b). Originally published as Höhle (1988).

- Höhle, Tilman N. 2016j. Vorangestellte Verben und Komplementierer sind eine natürliche Klasse. In Höhle (2016b). First published as Höhle (1997).
- Holler, Anke. 2005. *Weiterführende Relativsätze: Empirische und theoretische Aspekte* (studia grammatica 60). Berlin: Akademie Verlag.
- Hornstein, Norbert. 2013. Three grades of grammatical involvement: Syntax from a Minimalist perspective. *Mind and Language* 28(4). 392–420.
- Hornstein, Norbert, Jairo Nunes & Kleantes K. Grohmann. 2005. *Understanding Minimalism* (Cambridge Textbooks in Linguistics). Cambridge, UK: Cambridge University Press.
- Huck, Geoffrey J. & Almerindo E. Ojeda (eds.). 1987. *Discontinuous constituency* (Syntax and Semantics 20). New York: Academic Press.
- Hudson, Carla L. & Elissa L. Newport. 1999. Creolization: Could adults really have done it all? In Annabel Greenhill, Heather Littlefield & Cheryl Tano (eds.), *Proceedings of the Boston University Conference on Language Development*, vol. 23, 265–276. Somerville, MA: Cascadilla Press.
- Hudson, Richard. 1980. Constituency and dependency. *Linguistics* 18. 179–198.
- Hudson, Richard. 1984. *Word grammar*. Oxford: Basil Blackwell.
- Hudson, Richard. 1989. Towards a computer-testable Word Grammar of English. *UCL Working Papers in Linguistics* 1. 321–339.
- Hudson, Richard. 1990. *English Word Grammar*. Oxford: Basil Blackwell.
- Hudson, Richard. 1991. *English Word Grammar*. Oxford: Basil Blackwell.
- Hudson, Richard. 2003. Mismatches in default inheritance. In Francis & Michaelis (2003) 355–402.
- Hudson, Richard. 2004. Are determiners heads? *Functions of Language* 11(1). 7–42.
- Hudson, Richard. 2010a. Reaction to: “The myth of language universals and cognitive science”: On the choice between phrase structure and dependency structure. *Lingua* 120(12). 2676–2679.
- Hudson, Richard A. 1988. Coordination and grammatical relations. *Journal of Linguistics* 24(2). 303–342.
- Hudson, Richard A. 1997. German partial VP fronting. Ms. University College London. <http://dickhudson.com/papers/>.
- Hudson, Richard A. 2000. Discontinuity. *Dependency Grammars, TAL* 41(1). 15–56.
- Hudson, Richard A. 2007. *Language networks: The new Word Grammar*. Oxford: Oxford University Press.
- Hudson, Richard A. 2010b. *An introduction to Word Grammar* (Cambridge Textbooks in Linguistics). Cambridge, UK: Cambridge University Press.
- Hudson Kam, Carla L. & Elissa L. Newport. 2005. Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development* 1. 151–195.
- Humboldt, Wilhelm von. 1988. *Gesammelte Werke*. Berlin, New York: Walter de Gruyter.
- Hurford, James R. 2002. Expression/induction models of language evolution: Dimensions and issues. In Ted J. Briscoe (ed.), *Linguistic evolution through language acquisition*, 301–344. Cambridge, UK: Cambridge University Press.

- Hurskainen, Arvi. 2006. Constraint Grammar in unconventional use: Handling complex Swahili idioms and proverbs. *Suominen, Mickael et.al.: A Man of Measure: Festschrift in Honour of Fred Karlsson on his 60th Birthday. Special Supplement to SKY Journal of Linguistics* 19. 397–406.
- Ingram, David & William Thompson. 1996. Early syntactic acquisition in German: Evidence for the modal hypothesis. *Language* 72(1). 97–120.
- Iordanskaja, L., M. Kim, R. Kittredge, B. Lavoie & A. Polguère. 1992. Generation of extended bilingual statistical reports. In Zampolli (1992) 1019–1023.
- Isabelle, Pierre (ed.). 1998. *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics*. Montreal, Quebec, Canada: Association for Computational Linguistics. DOI:10.3115/980845.980953.
- Isabelle, Pierre (ed.). 2002. *40th Annual Meeting of the Association for Computational Linguistics: Proceedings of the conference*. University of Pennsylvania, Philadelphia: Association for Computational Linguistics. <http://aclanthology.info/events/acl-2002>.
- Islam, Md. Asfaul, K. M. Azharul Hasan & Md. Mizanur Rahman. 2012. Basic HPSG structure for Bangla grammar. In *15th International Conference on Computer and Information Technology (ICCIT)*, 185–189. Curran Associates. DOI:10.1109/ICCITech.2012.6509749.
- Jackendoff, Ray S. 1972. *Semantic interpretation in Generative Grammar*. Cambridge, MA/London, England: MIT Press.
- Jackendoff, Ray S. 1975. Morphological and semantic regularities in the lexikon. *Language* 51(3). 639–671.
- Jackendoff, Ray S. 1977. *\bar{X} syntax: A study of phrase structure*. Cambridge, MA/London, England: MIT Press.
- Jackendoff, Ray S. 1997. *The architecture of the language faculty* (Linguistic Inquiry Monographs 28). Cambridge, MA/London: MIT Press.
- Jackendoff, Ray S. 2000. Fodorian modularity and representational modularity. In Yosef Grodzinsky, Lewis P. Shapiro & David Swinney (eds.), *Language and the brain: Representation and processing* (Foundations of Neuropsychology), 3–30. San Diego: Academic Press.
- Jackendoff, Ray S. 2002. *Foundations of language*. Oxford: Oxford University Press.
- Jackendoff, Ray S. 2007. A parallel architecture perspective on language processing. *Brain Research* 1146. 2–22.
- Jackendoff, Ray S. 2008. Construction after Construction and its theoretical challenges. *Language* 84(1). 8–28.
- Jackendoff, Ray S. 2011. What is the human language faculty? Two views. *Language* 87(3). 586–624.
- Jackendoff, Ray S. & Steven Pinker. 2005. The nature of the language faculty and its implications for evolution of language (reply to Fitch, Hauser, and Chomsky). *Cognition* 97(2). 211–225.
- Jackendoff, Ray S. & Steven Pinker. 2009. The reality of a universal language faculty. *The Behavioral and Brain Sciences* 32(5). 465–466.

- Jacobs, Joachim. 1986. The syntax of focus and adverbials in German. In Werner Abraham & S. de Meij (eds.), *Topic, focus, and configurationality: Papers from the 6th Groningen Grammar Talks, Groningen, 1984* (Linguistik Aktuell/Linguistics Today 4), 103–127. Amsterdam: John Benjamins Publishing Co.
- Jacobs, Joachim. 1991. Bewegung als Valenztransfer. SFB 282: Theorie des Lexikons 1 Heinrich Heine Uni/BUGH Düsseldorf/Wuppertal.
- Jacobs, Joachim. 2008. Wozu Konstruktionen? *Linguistische Berichte* 213. 3–44.
- Jacobson, Pauline. 1987a. Phrase structure, grammatical relations, and discontinuous constituents. In Huck & Ojeda (1987) 27–69.
- Jacobson, Pauline. 1987b. Review of generalized phrase structure grammar. *Linguistics and Philosophy* 10(3). 389–426.
- Jaeggli, Osvaldo A. 1986. Passive. *Linguistic Inquiry* 17(4). 587–622.
- Jäger, Gerhard & Reinhard Blutner. 2003. Competition and interpretation: The German adverb *wieder* ‘again’. In Lang et al. (2003) 393–416.
- Jäger, Gerhard, Paola Monachesi, Gerald Penn & Shuly Wintner (eds.). 2002. *Proceedings of Formal Grammar 2002*. Trento.
- Jäger, Gerhard, Paola Monachesi, Gerald Penn & Shuly Wintner (eds.). 2008. *Proceedings of Formal Grammar 2003, Vienna, Austria*. Stanford, CA: CSLI Publications.
- Jäppinen, H., A. Lehtola & K. Valkonen. 1986. Functional structures for parsing dependency constraints. In Nagao (1986) 461–463. DOI:10.3115/991365.991501.
- Johnson, David E. & Shalom Lappin. 1997. A critique of the Minimalist Programm. *Linguistics and Philosophy* 20(3). 273–333.
- Johnson, David E. & Shalom Lappin. 1999. *Local constraints vs. economy* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.
- Johnson, David E. & Paul M. Postal. 1980. *Arc Pair Grammar*. Princeton, NJ: Princeton University Press.
- Johnson, Jacqueline S. & Elissa L. Newport. 1989. Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21(1). 60–99.
- Johnson, Kent. 2004. Gold’s theorem and cognitive science. *Philosophy of Science* 71(4). 571–592.
- Johnson, Mark. 1986. A GPSG account of VP structure in German. *Linguistics* 24(5). 871–882.
- Johnson, Mark. 1988. *Attribute-value logic and the theory of grammar* (CSLI Lecture Notes 14). Stanford, CA: CSLI Publications.
- Johnson, Mark. 1989. Parsing as deduction: The use of knowledge of language. *Journal of Psycholinguistic Research* 18(1). 105–128.
- Johnson, Mark, Stuart Geman, Stephen Canon, Zhiyi Chi & Stefan Riezler. 1999. Estimators for stochastic “unification-based” grammars. In Robert Dale & Ken Church (eds.), *Proceedings of the Thirty-Seventh Annual Meeting of the ACL*, 535–541. Association for Computational Linguistics.
- Joshi, Aravind K. 1985. Tree Adjoining Grammars: How much context-sensitivity is required to provide reasonable structural descriptions? In David Dowty, Lauri Kart-

- tunen & Arnold Zwicky (eds.), *Natural language parsing*, 206–250. Cambridge University Press.
- Joshi, Aravind K. 1987a. Introduction to Tree Adjoining Grammar. In Alexis Manaster-Ramer (ed.), *The mathematics of language*, 87–114. Amsterdam: John Benjamins Publishing Co.
- Joshi, Aravind K. 1987b. Word-order variation in natural language generation. In *AAAI 87, Sixth National Conference on Artificial Intelligence*, 550–555. Seattle.
- Joshi, Aravind K., Tilman Becker & Owen Rambow. 2000. Complexity of scrambling: A new twist to the competence-performance distinction. In Abeillé & Rambow (2000b) 167–181.
- Joshi, Aravind K., Leon S. Levy & Masako Takahashi. 1975. Tree Adjunct Grammar. *Journal of Computer and System Science* 10(2). 136–163.
- Joshi, Aravind K. & Yves Schabes. 1997. Tree-Adjoining Grammars. In G. Rozenberg & A. Salomaa (eds.), *Handbook of formal languages*, 69–123. Berlin: Springer Verlag.
- Joshi, Aravind K., K. Vijay Shanker & David Weir. 1990. The convergence of mildly context-sensitive grammar formalisms. Tech. Rep. MS-CIS-90-01 Department of Computer and Information Science, University of Pennsylvania. http://repository.upenn.edu/cis_reports/539/.
- Jungen, Oliver & Horst Lohnstein. 2006. *Einführung in die Grammatiktheorie* (UTB 2676). München: Wilhelm Fink Verlag.
- Jurafsky, Daniel. 1996. A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science* 20(2). 137–194.
- Kahane, Sylvain. 1997. Bubble trees and syntactic representations. In Tilman Becker & Hans-Ulrich Krieger (eds.), *Proceedings of Mathematics of Language (MOL5) Meeting*, 70–76. Saarbrücken: DFKI.
- Kahane, Sylvain. 2003. The Meaning-Text Theory. In Ágel et al. (2003b) 546–570.
- Kahane, Sylvain. 2009. On the status of phrases in Head-Driven Phrase Structure Grammar: Illustration by a fully lexical treatment of extraction. In Alain Polguère & Igor A. Mel'čuk (eds.), *Dependency in linguistic description* (Studies in Language Companion Series 111), 111–150. Amsterdam: John Benjamins Publishing Co.
- Kahane, Sylvain, Alexis Nasr & Owen Rambow. 1998. Pseudo-projectivity: A polynomially parsable non-projective Dependency Grammar. In Isabelle (1998) 646–652. DOI:10.3115/980845.980953. <http://www.aclweb.org/anthology/P98-1106>.
- Kallmeyer, Laura. 2005. Tree-local Multicomponent Tree Adjoining Grammars with shared nodes. *Computational Linguistics* 31(2). 187–225.
- Kallmeyer, Laura & Aravind K. Joshi. 2003. Factoring predicate argument and scope semantics: Underspecified semantics with LTAG. *Research on Language and Computation* 1(1–2). 3–58. DOI:10.1023/A:1024564228892.
- Kallmeyer, Laura, Timm Lichte, Wolfgang Maier, Yannick Parmentier, Johannes Dellert & Kilian Evang. 2008. TuLiPA: Towards a multi-formalism parsing environment for grammar engineering. In Stephen Clark & Tracy Holloway King (eds.), *Coling 2008: Proceedings of the Workshop on Grammar Engineering Across Frameworks*, 1–8. Manchester, England: Association for Computational Linguistics.

- Kallmeyer, Laura & Rainer Osswald. 2012. A frame-based semantics of the dative alternation in Lexicalized Tree Adjoining Grammars. In Piñón (2012) 167–184.
- Kallmeyer, Laura & Maribel Romero. 2008. Scope and situation binding in LTAG using semantic unification. *Research on Language and Computation* 6(1). 3–52.
- Kallmeyer, Laura & Sinwon Yoon. 2004. Tree-local MCTAG with shared nodes: An analysis of word order variation in German and Korean. *Traitement automatique des langues TAL* 45(3). 49–69.
- Kamp, Hans & Uwe Reyle. 1993. *From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and Discourse Representation Theory* (Studies in Linguistics and Philosophy 42). Dordrecht: Kluwer Academic Publishers.
- Kaplan, Ronald M. 1995. The formal architecture of Lexical-Functional Grammar. In Dalrymple et al. (1995) 7–27.
- Kaplan, Ronald M. & Joan Bresnan. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In Bresnan (1982b) 173–281. Reprint in: Dalrymple et al. (1995: 29–130).
- Kaplan, Ronald M. & John T. Maxwell III. 1996. LFG grammar writer’s workbench. Tech. rep. Xerox PARC.
- Kaplan, Ronald M., Stefan Riezler, Tracy Holloway King, John T. Maxwell III, Alexander Vasserman & Richard Crouch. 2004. Speed and accuracy in shallow and deep stochastic parsing. In *Proceedings of the Human Language Technology Conference and the 4th Annual Meeting of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL’04)*, Boston, MA: Association for Computational Linguistics.
- Kaplan, Ronald M. & Annie Zaenen. 1989. Long-distance dependencies, constituent structure and functional uncertainty. In Baltin & Kroch (1989) 17–42.
- Karimi, Simin. 2005. *A Minimalist approach to scrambling: Evidence from Persian* (Studies in Generative Grammar 76). Berlin, New York: Mouton de Gruyter.
- Karimi-Doostan, Gholamhossein. 2005. Light verbs and structural case. *Lingua* 115(12). 1737–1756.
- Karlgren, Hans (ed.). 1990. *COLING-90: Papers presented to the 13th International Conference on Computational Linguistics*. Helsinki: Association for Computational Linguistics.
- Karlsson, Fred (ed.). 1986. *Papers from the Fifth Scandinavian Conference of Computational Linguistics*. Helsinki.
- Karlsson, Fred. 1990. Constraint Grammar as a framework for parsing running text. In Karlgren (1990) 168–173.
- Karmiloff-Smith, Annette. 1998. Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences* 2(10). 389–398.
- Karmiloff-Smith, Annette, Julia Grant, Ioanna Berthoud, Mark Davies, Patricia Howlin & Orlee Udwin. 1997. Language in Williams Syndrome: How intact is ‘intact’? *Child Development* 68(2). 246–262.
- Karttunen, Lauri. 1986. Radical lexicalism. Report No. CSLI-86-68 Center for the Study of Language and Information.
- Karttunen, Lauri. 1989. Radical lexicalism. In Baltin & Kroch (1989) 43–65.

- Kasper, Robert T. 1994. Adjuncts in the Mittelfeld. In Nerbonne et al. (1994b) 39–70.
- Kasper, Robert T., Bernd Kiefer, Klaus Netter & Vijay K. Shanker. 1995. Compilation of HPSG to TAG. In Uszkoreit (1995) 92–99.
- Kathol, Andreas. 1995. *Linearization-based German syntax*: Ohio State University dissertation.
- Kathol, Andreas. 2000. *Linear syntax*. New York, Oxford: Oxford University Press.
- Kathol, Andreas. 2001. Positional effects in a monostratal grammar of German. *Journal of Linguistics* 37(1). 35–66.
- Kathol, Andreas & Carl J. Pollard. 1995. Extraposition via complex domain formation. In Uszkoreit (1995) 174–180.
- Kaufmann, Ingrid & Dieter Wunderlich. 1998. Cross-linguistic patterns of resultatives. SFB 282: Theorie des Lexikons 109 Heinrich Heine Uni/BUGH Düsseldorf.
- Kaufmann, Tobias. 2009. *A rule-based language model for speech recognition*: Computer Engineering and Networks Laboratory, ETH Zürich dissertation.
- Kaufmann, Tobias & Beat Pfister. 2007. Applying licenser rules to a grammar with continuous constituents. In Müller (2007d) 150–162. <http://csli-publications.stanford.edu/HPSG/2007/>.
- Kaufmann, Tobias & Beat Pfister. 2008. Applying a grammar-based language model to a broadcast-news transcription task. In Moore et al. (2008) 106–113. <http://www.aclweb.org/anthology/P/P08/P08-1013.pdf>.
- Kay, Martin. 1967. Experiments with a powerful parser. In *Proceedings of Conference Internationale Sur Le Traitement Automatique Des Langues (COLING 1967)*, Grenoble. <http://aclweb.org/anthology/C/C67/C67-1009.pdf>.
- Kay, Martin. 2000. David G. Hays. In William J. Hutchins (ed.), *Early years in machine translation* (Amsterdam Studies in the Theory and History of Linguistics Science Series 3), 165–170. Amsterdam: John Benjamins Publishing Co.
- Kay, Martin. 2011. Zipf’s law and *L’Arbitraire du Signe*. *Linguistic Issues in Language Technology* 6(8). 1–25. <http://journals.linguisticsociety.org/elanguage/lilt/article/view/2584.html>. Special Issue on Interaction of Linguistics and Computational Linguistics.
- Kay, Paul. 2002. An informal sketch of a formal architecture for Construction Grammar. *Grammars* 5(1). 1–19.
- Kay, Paul. 2005. Argument structure constructions and the argument-adjunct distinction. In Mirjam Fried & Hans C. Boas (eds.), *Grammatical constructions: Back to the roots* (Constructional Approaches to Language 4), 71–98. Amsterdam: John Benjamins Publishing Co.
- Kay, Paul & Charles J. Fillmore. 1999. Grammatical Constructions and linguistic generalizations: The What’s X Doing Y? Construction. *Language* 75(1). 1–33.
- Kay, Paul, Ivan A. Sag & Daniel P. Flickinger. 2015. A lexical theory of phrasal idioms. Ms. CSLI Stanford.
- Kayne, Richard S. 1994. *The antisymmetry of syntax* (Linguistic Inquiry Monographs 25). Cambridge, MA: MIT Press.
- Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8(1). 63–99.

参考文献

- Keller, Frank. 1994. German functional HPSG – An experimental CUF encoding. Tech. rep. Institut für Maschinelle Sprachverarbeitung Stuttgart.
- Keller, Frank. 1995. Towards an account of extraposition in HPSG. In Abney & Hinrichs (1995).
- Kettunen, Kimmo. 1986. On modelling dependency-oriented parsing. In Karlsson (1986) 113–120.
- Kiefer, Bernd, Hans-Ulrich Krieger & Mark-Jan Nederhof. 2000. Efficient and robust parsing of word hypotheses graphs. In Wahlster (2000b) 280–295.
- Kifle, Nazareth Amlesom. 2012. *Tigrinya applicatives in Lexical-Functional Grammar*: University of Bergen dissertation. <http://hdl.handle.net/1956/5730>.
- Kim, Jong-Bok & Peter Sells. 2008. *English syntax: An introduction* (CSLI Lecture Notes 185). Stanford, CA: CSLI Publications.
- Kim, Jong-Bok, Peter Sells & Jaehyung Yang. 2007. Parsing two types of multiple nominative constructions: A Constructional approach. *Language and Information* 11(1). 25–37.
- Kim, Jong-Bok & Jaehyung Yang. 2003. Korean phrase structure grammar and its implementations into the LKB system. In Dong Hong Ji & Kim Teng Lua (eds.), *Proceedings of the 17th Pacific Asia Conference on Language, Information and Computation*, 88–97. National University of Singapore: COLIPS Publications.
- Kim, Jong-Bok & Jaehyung Yang. 2004. Projections from morphology to syntax in the Korean Resource Grammar: Implementing typed feature structures. In Alexander Gelbukh (ed.), *Computational linguistics and intelligent text processing: 5th International Conference, CICLing 2004, Seoul, Korea, February 15-21, 2004, Proceedings* (Lecture Notes in Computer Science 2945), 13–24. Berlin: Springer Verlag.
- Kim, Jong-Bok & Jaehyung Yang. 2006. Coordination structures in a typed feature structure grammar: Formalization and implementation. In Tapio Salakoski, Filip Ginter, Sampo Pyysalo & Tapio Pahikkala (eds.), *Advances in natural language processing: 5th International Conference, FinTAL 2006 Turku, Finland, August 23-25, 2006 proceedings* (Lecture Notes in Artificial Intelligence 4139), 194–205. Berlin: Springer Verlag.
- Kim, Jong-Bok & Jaehyung Yang. 2009. Processing three types of Korean cleft constructions in a typed feature structure grammar. *Korean Journal of Cognitive Science* 20(1). 1–28.
- Kim, Jong-Bok, Jaehyung Yang, Sanghoun Song & Francis Bond. 2011. Deep processing of Korean and the development of the Korean Resource Grammar. *Linguistic Research* 28(3). 635–672.
- Kimball, John P. 1973. *The formal theory of grammar* (Foundations of Modern Linguistics). Englewood Cliffs, NJ: Prentice-Hall.
- King, Maghi (ed.). 1985. *Proceedings of the 2nd European Meeting of the Association for Computational Linguistics*. Geneva: Association for Computational Linguistics. <http://aclweb.org/anthology/E/E85/>.
- King, Paul. 1994. An expanded logical formalism for Head-Driven Phrase Structure Grammar. Arbeitspapiere des SFB 340 No. 59 Eberhard-Karls-Universität Tübingen. <http://www.sfs.uni-tuebingen.de/sfb/reports/berichte/59/59abs.html>.

- King, Paul. 1999. Towards truth in Head-Driven Phrase Structure Grammar. In Valia Kordoni (ed.), *Tübingen studies in Head-Driven Phrase Structure Grammar* (Arbeitsberichte des SFB 340 No. 132), 301–352. Tübingen: Universität Tübingen. <http://www.sfs.uni-tuebingen.de/sfb/reports/berichte/132/132abs.html>.
- King, Tracy Holloway & Emily M. Bender (eds.). 2007. *Grammar Engineering across Frameworks 2007*. (Studies in Computational Linguistics ONLINE) Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/GEAF/2007/>.
- King, Tracy Holloway & John T. Maxwell III. 2007. Overlay mechanisms for multi-level deep processing applications. In King & Bender (2007) 182–202. <http://csli-publications.stanford.edu/GEAF/2007/>.
- King, Tracy Holloway & Marianne Santaholma (eds.). 2009. *Proceedings of the 2009 Workshop on Grammar Engineering Across Frameworks (GEAF 2009)*. Suntec, Singapore: Association for Computational Linguistics. <http://www.aclweb.org/anthology/W/W09/#2600>.
- Kinyon, Alexandra, Owen Rambow, Tatjana Scheffler, SinWon Yoon & Aravind K. Joshi. 2006. The Metagrammar goes multilingual: A cross-linguistic look at the V2-phenomenon. In Laura Kallmeyer & Tilman Becker (eds.), *TAG+8: The Eighth International Workshop on Tree Adjoining Grammar and Related Formalisms: Proceedings of the workshop*, 17–24. Sydney, Australia: Association for Computational Linguistics.
- Kiparsky, Paul. 1987. Morphology and grammatical relations. Unpublished paper, Stanford University, Stanford.
- Kiparsky, Paul. 1988. Agreement and linking theory. Unpublished paper, Stanford University, Stanford.
- Kiparsky, Paul & Carol Kiparsky. 1970. Fact. In Manfred Bierwisch & Karl Erich Heidolph (eds.), *Progress in linguistics*, 143–173. The Hague/Paris: Mouton.
- Kiss, Katalin E. 2003. Argument scrambling, focus movement and topic movement in Hungarian. In Simin Karimi (ed.), *Word order and scrambling*, 22–43. London: Blackwell.
- Kiss, Tibor. 1991. The grammars of LILOG. In Herzog & Rollinger (1991) 183–199.
- Kiss, Tibor. 1992. Variable Subkategorisierung: Eine Theorie unpersönlicher Einbettungen im Deutschen. *Linguistische Berichte* 140. 256–293.
- Kiss, Tibor. 1993. Infinite Komplementation – Neue Studien zum deutschen Verbum infinitum. Arbeiten des SFB 282 No. 42 Bergische Universität Gesamthochschule Wuppertal.
- Kiss, Tibor. 1995. *Infinite Komplementation: Neue Studien zum deutschen Verbum infinitum* (Linguistische Arbeiten 333). Tübingen: Max Niemeyer Verlag.
- Kiss, Tibor. 2001. Configurational and relational scope determination in German. In Meurers & Kiss (2001) 141–175.
- Kiss, Tibor. 2005. Semantic constraints on relative clause extraposition. *Natural Language and Linguistic Theory* 23(2). 281–334.
- Kiss, Tibor & Artemis Alexiadou (eds.). 2015. *Syntax – theory and analysis: An international handbook*, vol. 42 (Handbooks of Linguistics and Communication Science). Berlin: Mouton de Gruyter 2nd edn.

- Kiss, Tibor & Birgit Wesche. 1991. Verb order and head movement. In Herzog & Rollinger (1991) 216–242.
- Klann-Delius, Gisela. 2008. *Spracherwerb*. Stuttgart: J.B. Metzler–Verlag 2nd edn.
- Klein, Wolfgang. 1971. *Parsing: Studien zur maschinellen Satzanalyse mit Abhängigkeitsgrammatiken und Transformationsgrammatiken*, vol. 2. Frankfurt a. M.: Athenäum Verlag.
- Klein, Wolfgang. 1985. Ellipse, Fokusgliederung und thematischer Stand. In Reinhard Meyer-Hermann & Hannes Rieser (eds.), *Ellipsen und fragmentarische Ausdrücke*, 1–24. Tübingen: Max Niemeyer Verlag.
- Klein, Wolfgang. 1986. *Second language acquisition* (Cambridge Textbooks in Linguistics). Cambridge, UK: Cambridge University Press.
- Klein, Wolfgang. 2009. Finiteness, Universal Grammar and the language faculty. In Jiansheng Guo, Elena Lieven, Nancy Budwig, Susan Ervin-Tripp, Keiko Nakamura & Seyda Ozcaliskan (eds.), *Cross-linguistic approaches to the study of language: Research in the tradition of Dan Isaac Slobin* (Psychology Press Festschrift Series), 333–344. New York: Psychology Press.
- Klenk, Ursula. 2003. *Generative Syntax* (Narr Studienbücher). Tübingen: Gunter Narr Verlag.
- Kluender, Robert. 1992. Deriving island constraints from principles of predication. In Helen Goodluck & Michael Rochemont (eds.), *Island constraints: Theory, acquisition, and processing*, 223–258. Dordrecht: Kluwer Academic Publishers.
- Kluender, Robert & Marta Kutas. 1993. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8(4), 573–633.
- Knecht, Laura. 1985. *Subject and object in Turkish*: M.I.T. dissertation.
- Kobele, Gregory M. 2008. Across-the-board extraction in Minimalist Grammars. In *Proceedings of the Ninth International Workshop on Tree Adjoining Grammar and Related Formalisms (TAG+9)*, 113–128.
- Koenig, Jean-Pierre. 1999. *Lexical relations* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.
- Koenig, Jean-Pierre & Karin Michelson. 2010. Argument structure of Oneida kinship terms. *International Journal of American Linguistics* 76(2), 169–205.
- Koenig, Jean-Pierre & Karin Michelson. 2012. The (non)universality of syntactic selection and functional application. In Piñón (2012) 185–205.
- Kohl, Dieter. 1992. Generation from under- and overspecified structures. In Zampolli (1992) 686–692.
- Kohl, Dieter, Claire Gardent, Agnes Plainfossé, Mike Reape & Stefan Momma. 1992. Text generation from semantic representation. In Gabriel G. Bes & Thierry Guillotin (eds.), *The construction of a natural language and graphic interface: Results and perspectives from the ACORD project*, 94–161. Berlin: Springer Verlag.
- Kohl, Karen T. 1999. *An analysis of finite parameter learning in linguistic spaces*. Massachusetts Institute of Technology MA thesis. <http://karentkohl.org/papers/SM.pdf>.

- Kohl, Karen T. 2000. Language learning in large parameter spaces. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence*, 1080. AAAI Press / The MIT Press.
- Kolb, Hans-Peter. 1997. GB blues: Two essays on procedures and structures in Generative Syntax. Arbeitspapiere des SFB 340 No. 110 Eberhard-Karls-Universität Tübingen.
- Kolb, Hans-Peter & Craig L. Thiersch. 1991. Levels and empty categories in a Principles and Parameters based approach to parsing. In Haider & Netter (1991) 251–301.
- Konieczny, Lars. 1996. *Human sentence processing: A semantics-oriented parsing approach*: Universität Freiburg Dissertation. IIG-Berichte 3/96.
- König, Esther. 1999. LexGram: A practical Categorical Grammar formalism. *Journal of Language and Computation* 1(1). 33–52.
- Koopman, Hilda & Dominique Sportiche. 1991. The position of subjects. *Lingua* 85(2–3). 211–258.
- Kordoni, Valia (ed.). 1999a. *Tübingen studies in Head-Driven Phrase Structure Grammar* (Arbeitspapiere des SFB 340, No.132, Volume 1). Tübingen: Eberhard-Karls-Universität Tübingen.
- Kordoni, Valia (ed.). 1999b. *Tübingen studies in Head-Driven Phrase Structure Grammar* (Arbeitspapiere des SFB 340, No.132, Volume 1). Tübingen: Eberhard-Karls-Universität Tübingen.
- Kordoni, Valia. 2001. Linking experiencer-subject psych verb constructions in Modern Greek. In Daniel P. Flickinger & Andreas Kathol (eds.), *Proceedings of the HPSG-2000 Conference, University of California, Berkeley*, 198–213. CSLI Publications. <http://csli-publications.stanford.edu/HPSG/1/>.
- Kordoni, Valia & Julia Neu. 2005. Deep analysis of Modern Greek. In Su et al. (2005) 674–683.
- Kornai, András & Geoffrey K. Pullum. 1990. The X-bar Theory of phrase structure. *Language* 66(1). 24–50.
- Koster, Jan. 1975. Dutch as an SOV language. *Linguistic Analysis* 1(2). 111–136.
- Koster, Jan. 1978. *Locality principles in syntax*. Dordrecht: Foris Publications.
- Koster, Jan. 1986. The relation between pro-drop, scrambling, and verb movements. *Groningen Papers in Theoretical and Applied Linguistics* 1. 1–43.
- Koster, Jan. 1987. *Domains and dynasties: The radical autonomy of syntax*. Dordrecht: Foris Publications.
- Kratzer, Angelika. 1984. On deriving syntactic differences between German and English. TU Berlin, ms.
- Kratzer, Angelika. 1996. Severing the external argument from its verb. In Johan Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon*, 109–137. Dordrecht: Kluwer Academic Publishers.
- Krauwert, Steven, Michael Moortgat & Louis des Tombe (eds.). 1993. *Sixth Conference of the European Chapter of the Association for Computational Linguistics. Proceedings of the conference*. Utrecht: Association for Computational Linguistics.
- Krieger, Hans-Ulrich & John Nerbonne. 1993. Feature-based inheritance networks for computational lexicons. In Ted Briscoe, Ann Copestake & Valeria de Paiva (eds.),

参考文献

- Inheritance, defaults, and the lexicon*, 90–136. Cambridge, UK: Cambridge University Press. A version of this paper is available as DFKI Research Report RR-91-31. Also published in: Proceedings of the ACQUILEX Workshop on Default Inheritance in the Lexicon, Technical Report No. 238, University of Cambridge, Computer Laboratory, October 1991.
- Kroch, Anthony S. 1987. Unbounded dependencies and subadjacency in a Tree Adjoining Grammar. In Alexis Manaster-Ramer (ed.), *Mathematics of language*, 143–172. Amsterdam: John Benjamins Publishing Co.
- Kroch, Anthony S. & Aravind K. Joshi. 1985. The linguistic relevance of Tree Adjoining Grammar. Tech. Rep. MS-CIS-85-16 University of Pennsylvania. http://repository.upenn.edu/cgi/viewcontent.cgi?article=1706&context=cis_reports.
- Kroch, Anthony S. & Aravind K. Joshi. 1987. Analyzing extraposition in a Tree Adjoining Grammar. In Huck & Ojeda (1987) 107–149.
- Kropp Dakubu, Mary Esther, Lars Hellan & Dorothee Beermann. 2007. Verb sequencing constraints in Ga: Serial verb constructions and the extended verb complex. In Müller (2007d) 99–119. <http://csli-publications.stanford.edu/HPSG/2007/>.
- Kruijff-Korbayová, Ivana & Mark J. Steedman. 2003. Discourse and information structure. *Journal of Logic, Language and Information: Special Issue on Discourse and Information Structure* 12(3). 249–259.
- Kübler, Sandra, Ryan McDonald & Joakim Nivre. 2009. *Dependency Parsing* (Synthesis Lectures on Human Language Technologies 2). San Rafael, U.S.A: Morgan & Claypool Publishers.
- Kuhn, Jonas. 1995. Information packaging in German: Some motivation from HPSG-based translation. Universität Stuttgart, ms. <ftp://ftp.ims.uni-stuttgart.de/pub/papers/kuhn/Info-Pack.pdf>.
- Kuhn, Jonas. 1996. An underspecified HPSG representation for information structure. In Tsuji (1996) 670–675.
- Kuhn, Jonas. 2007. Interfaces in constraint-based theories of grammar. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford handbook of linguistic interfaces*, 613–650. Oxford: Oxford University Press.
- Kuhn, Jonas & Christian Rohrer. 1997. Approaching ambiguity in real-life sentences: The application of an Optimality Theory-inspired constraint ranking in a large-scale LFG grammar. In *Proceedings of DGfS/CL 97*, Heidelberg.
- Kuhns, Robert J. 1986. A PROLOG implementation of Government-Binding Theory. In Alan W. Biermann (ed.), *Proceedings of the Twenty-Fourth Annual Meeting of the Association for Computational Linguistics*, 546–550. Columbia University, New York: Association for Computational Linguistics.
- Kunze, Jürgen. 1968. The treatment of non-projective structures in the syntactic analysis and synthesis of English and German. *Computational Linguistics* 7. 67–77.
- Kunze, Jürgen. 1975. *Abhängigkeitsgrammatik* (studia grammatica 12). Berlin: Akademie Verlag.
- Kunze, Jürgen. 1991. *Kasusrelationen und semantische Emphase* (studia grammatica XXXII). Berlin: Akademie Verlag.

- Kunze, Jürgen. 1993. *Sememstrukturen und Feldstrukturen* (studia grammatica 36). Berlin: Akademie Verlag. Unter Mitarbeit von Beate Firzlaß.
- Labelle, Marie. 2007. Biolinguistics, the Minimalist Program, and psycholinguistic reality. *Snippets* 14. 6–7. <http://www.ledonline.it/snippets/>.
- Laczkó, Tibor, György Rákosi & Ágoston Tóth. 2010. HunGram vs. EngGram in ParGram: On the comparison of Hungarian and English in an international computational linguistics project. In Irén Hegedűs & Sándor Martsa (eds.), *Selected papers in linguistics from the 9th HUSSE Conference*, vol. 1, 81–95. Pécs: Institute of English Studies, Faculty of Humanities, University of Pécs.
- Laenzlinger, Christoph. 2004. A feature-based theory of adverb syntax. In Jennifer R. Austin, Stefan Engelberg & Gisa Rauh (eds.), *Adverbials: The interplay between meaning, context, and syntactic structure* (Linguistik Aktuell/Linguistics Today 70), 205–252. Amsterdam: John Benjamins Publishing Co.
- Lai, Cecilia S. L., Simon E. Fisher, Jane A. Hurst, Faraneh Vargha-Khadem & Anthony P. Monaco. 2001. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413(6855). 519–523. DOI:10.1038/35097076.
- Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: The University of Chicago Press.
- Lang, Ewald, Claudia Maienborn & Cathrine Fabricius-Hansen (eds.). 2003. *Modifying adjuncts* (Interface Explorations 4). Berlin: Mouton de Gruyter.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*, vol. 1. Stanford, CA: Stanford University Press.
- Langacker, Ronald W. 2000. A dynamic usage-based model. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 1–63. Stanford, CA: CSLI Publications.
- Langacker, Ronald W. 2008. *Cognitive Grammar: A basic introduction*. Oxford: Oxford University Press.
- Langacker, Ronald W. 2009. Cognitive (Construction) Grammar. *Cognitive Linguistics* 20(1). 167–176.
- Lappin, Shalom, Robert D. Levine & David E. Johnson. 2000a. The revolution confused: A response to our critics. *Natural Language and Linguistic Theory* 18(4). 873–890.
- Lappin, Shalom, Robert D. Levine & David E. Johnson. 2000b. The structure of unscientific revolutions. *Natural Language and Linguistic Theory* 18(3). 665–671.
- Lappin, Shalom, Robert D. Levine & David E. Johnson. 2001. The revolution maximally confused. *Natural Language and Linguistic Theory* 19(4). 901–919.
- Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19(3). 335–391.
- Lascarides, Alex & Ann Copestake. 1999. Default representation in constraint-based frameworks. *Computational Linguistics* 25(1). 55–105.
- Lasnik, Howard & Mamoru Saito. 1992. *Move α : Conditions on its application and output* (Current Studies in Linguistics 22). Cambridge, MA: MIT Press.
- Lasnik, Howard & Juan Uriagereka. 2002. On the poverty of the challenge. *The Linguistic Review* 19(1–2). 147–150.

- Lavoie, Benoit & Owen Rambow. 1997. RealPro—A fast, portable sentence realizer. In *Proceedings of the Conference on Applied Natural Language Processing (ANLP'97)*, .
- Le, Hong Phuong, Thi Minh Huyen Nguyen & Azim Roussanaly. 2008. Metagrammar for Vietnamese LTAG. In *Proceedings of the Ninth International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+9)*, 129–132. Tübingen.
- Legate, Julie & Charles D. Yang. 2002. Empirical re-assessment of stimulus poverty arguments. *The Linguistic Review* 19(1–2). 151–162.
- Lehtola, Aarno. 1986. DPL: A computational method for describing grammars and modelling parsers. In Karlsson (1986) 151–159.
- Leiss, Elisabeth. 2003. Empirische Argumente für Dependenz. In Ágel et al. (2003b) 311–324.
- Leiss, Elisabeth. 2009. *Sprachphilosophie* (de Gruyter Studienbuch). Berlin: Walter de Gruyter.
- Lenerz, Jürgen. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen* (Studien zur deutschen Grammatik 5). Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.
- Lenerz, Jürgen. 1994. Pronomenprobleme. In Brigitta Haftka (ed.), *Was determiniert Wortstellungsvariation? Studien zu einem Interaktionsfeld von Grammatik, Pragmatik und Sprachtypologie*, 161–174. Opladen: Westdeutscher Verlag.
- Lenneberg, Eric H. 1964. The capacity for language acquisition. In Fodor & Katz (1964) 579–603.
- Lenneberg, Eric H. 1967. *Biological foundations of language*. New York: John Wiley & Sons, Inc.
- Levelt, Willem J. M. 1989. *Speaking: From intonation to articulation* (ACL-MIT Press Series in Natural Language Processing). Cambridge, MA: MIT Press.
- Levin, Beth. 1993. *English verb classes and alternations: A preliminary investigation*. Chicago, Illinois: University of Chicago Press.
- Levin, Beth & Malka Rappaport Hovav. 2005. *Argument realization*. Cambridge University Press.
- Levin, Lori S., Malka Rappaport & Annie Zaenen (eds.). 1983. *Papers in Lexical Functional Grammar*. Indiana: Indiana University Linguistics Club.
- Levine, Robert D. 2003. Adjunct valents, cumulative scopings and impossible descriptions. In Jongbok Kim & Stephen Mark Wechsler (eds.), *The proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*, 209–232. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/HPSG/3/>.
- Levine, Robert D. & Thomas E. Hukari. 2006. *The unity of unbounded dependency constructions* (CSLI Lecture Notes 166). Stanford, CA: CSLI Publications.
- Levine, Robert D. & Walt Detmar Meurers. 2006. Head-Driven Phrase Structure Grammar: Linguistic approach, formal foundations, and computational realization. In Brown (2006) 237–252.
- Lewis, Geoffrey L. 1967. *Turkish grammar*. Oxford: Clarendon Press.
- Lewis, John D. & Jeffrey L. Elman. 2001. Learnability and the statistical structure of language: Poverty of Stimulus arguments revisited. In Barbora Skarabela, Sarah Fish

- & Anna H.-J. Do (eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development*, 359–370. <http://crl.ucsd.edu/~elman/Papers/BU2001.pdf>.
- Li, Charles N. & Sandra A. Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley and Los Angeles: University of California Press.
- Li, Wei. 1996. Esperanto inflection and its interface in HPSG. Working papers of the linguistics circle University of Victoria.
- Lichte, Timm. 2007. An MCTAG with tuples for coherent constructions in German. In Laura Kallmeyer, Paola Monachesi, Gerald Penn & Giorgio Satta (eds.), *Proceedings of the 12th Conference on Formal Grammar 2007*, Dublin, Ireland.
- Lieb, Hans-Heinrich. 1983. *Integrational linguistics: Vol. I: General outline* (Current Issues in Linguistic Theory 17). Amsterdam: John Benjamins Publishing Co.
- Lightfoot, David W. 1997. Catastrophic change and learning theory. *Lingua* 100(1). 171–192.
- Link, Godehard. 1984. Hydras: On the logic of relative constructions with multiple heads. In Fred Landmann & Frank Veltman (eds.), *Varieties of formal semantics*, 245–257. Dordrecht: Foris Publications.
- Lipenkova, Janna. 2009. *Serienverbkonstruktionen im Chinesischen und ihre Analyse im Rahmen von HPSG*. Institut für Sinologie, Freie Universität Berlin MA thesis.
- Liu, Gang. 1997. *Eine unifikations-basierte Grammatik für das moderne Chinesisch – dargestellt in der HPSG*: FG Sprachwissenschaft, Universität Konstanz dissertation. <http://www.ub.uni-konstanz.de/kops/volltexte/1999/191/>.
- Liu, Haitao. 2009. *Dependency Grammar: From theory to practice*. Beijing: Science Press.
- Liu, Haitao & Wei Huang. 2006. Chinese Dependency Syntax for treebanking. In *Proceedings of the Twentieth Pacific Asia Conference on Language, Information and Computation*, 126–133. Beijing: Tsinghua University Press.
- Lloré, F. Xavier. 1995. *Un Método de ‘Parsing’ para Gramáticas Catoriales Multimodales*: I.C.E. de la Universidad Politécnica de Catalunya dissertation.
- Lobin, Henning. 1993. *Koordinationssyntax als strukturelles Phänomen* (Studien zur Grammatik 46). Tübingen: Gunter Narr Verlag.
- Lobin, Henning. 2003. Dependenzgrammatik und Kategorialgrammatik. In Ágel et al. (2003b) 325–330.
- Löbner, Sebastian. 1986. In Sachen Nullartikel. *Linguistische Berichte* 101. 64–65. <http://user.phil-fak.uni-duesseldorf.de/~loebner/publ/Nullartikel.html>.
- Lohndal, Terje. 2012. Toward the end of argument structure. In María Cristina Cuervo & Yves Roberge (eds.), *The end of argument structure?*, vol. 38 (Syntax and Semantics), 155–184. Bingley, UK: Emerald Group Publishing.
- Lohnstein, Horst. 1993. *Projektion und Linking: Ein prinzipienbasierter Parser fürs Deutsche* (Linguistische Arbeiten 287). Tübingen: Max Niemeyer Verlag.
- Lohnstein, Horst. 2014. Artenvielfalt in freier Wildbahn: Generative Grammatik. In Hagemann & Staffeldt (2014) 165–185.
- Longobardi, Giuseppe & Ian Roberts. 2010. Universals, diversity and change in the science of language: Reaction to “The myth of language universals and cognitive science”. *Lingua* 120(12). 2699–2703.

- Lorenz, Konrad. 1970. *Studies in human and animal behavior*, vol. I. Cambridge, MA: Harvard University Press.
- Lötscher, Andreas. 1985. Syntaktische Bedingungen der Topikalisierung. *Deutsche Sprache* 13(3). 207–229.
- Loukam, Mourad, Amar Balla & Mohamed Tayeb Laskri. 2015. Towards an open platform based on HPSG formalism for the Standard Arabic language. *International Journal of Speech Technology* DOI:10.1007/s10772-015-9314-4.
- Lüdeling, Anke. 2001. *On particle verbs and similar constructions in German* (Dissertations in Linguistics). Stanford, CA: CSLI Publications.
- Lüdeling, Anke & Merja Kytö (eds.). 2009. *Corpus linguistics: An international handbook*, vol. 29 (Handbücher zur Sprach- und Kommunikationswissenschaft). Berlin: Mouton de Gruyter.
- Luuk, Erkki & Hendrik Luuk. 2011. The redundancy of recursion and infinity for natural language. *Cognitive Processing* 12(1). 1–11.
- Maas, Heinz Dieter. 1977. The Saarbrücken Automatic Translation System (SUSY). In Eric James Coates (ed.), *Proceedings of the Third European Congress on Information Systems and Networks: Overcoming the Language Barrier*, vol. 1, 585–592. München: Verlag Dokumentation.
- Maché, Jakob. 2010. Towards a compositional analysis of verbless directives in German. Paper presented at the HPSG 2010 Conference.
- Machicao y Priemer, Antonio. 2015. SpaGram: An implemented grammar fragment of Spanish. Ms. Humboldt Universität zu Berlin. In Preparation.
- MacWhinney, Brian. 1995. *The CHILDES project: Tools for analyzing talk*. Hillsdale, NJ: Erlbaum 2nd edn.
- Maess, Burkhard, Stefan Koelsch, Thomas C. Gunter & Angela D. Friederici. 2001. Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience* 4(5). 540–545.
- Marantz, Alec. 1984. *On the nature of grammatical relations* (Linguistic Inquiry Monographs 10). Cambridge, MA: MIT Press.
- Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *U. Penn Working Papers in Linguistics* 4(2). 201–225. <http://www.ling.upenn.edu/papers/v4.2-contents.html>.
- Marantz, Alec. 2005. Generative linguistics within the cognitive neuroscience of language. *The Linguistic Review* 22(2–4). 429–445.
- Marcus, Gary F. 1993. Negative evidence in language acquisition. *Cognition* 46(1). 53–85.
- Marcus, Gary F. & Simon E. Fisher. 2003. FOXP2 in focus: What can genes tell us about speech and language? *TRENDS in Cognitive Sciences* 7(6). 257–262.
- Marcus, Mitchell P. 1980. *A theory of syntactic recognition for natural language*. London, England/Cambridge, MA: MIT Press.
- Marimon, Montserrat. 2013. The Spanish DELPH-IN grammar. *Language Resources and Evaluation* 47(2). 371–397. DOI:10.1007/s10579-012-9199-7.
- Marslen-Wilson, William. 1975. Sentence perception as an interactive parallel process. *Science* 189(4198). 226–228.

- Masuichi, Hiroshi & Tomoko Ohkuma. 2003. Constructing a practical Japanese parser based on Lexical-Functional Grammar. *Journal of Natural Language Processing* 10. 79–109. In Japanese.
- Masum, Mahmudul Hasan, Muhammad Sadiqul Islam, M. Sohel Rahman & Reaz Ahmed. 2012. HPSG analysis of type-based Arabic nominal declension. In *The 13th International Arab Conference*, 272–279.
- Mayo, Bruce. 1997. Die Konstanzer LFG-Umgebung. Arbeitspapier 82 des Fachbereichs Sprachwissenschaft der Universität Konstanz Universität Konstanz.
- Mayo, Bruce. 1999. *A computational model of derivational morphology*: Universität Hamburg dissertation. <http://www.sub.uni-hamburg.de/opus/volltexte/1999/386/>.
- Meinunger, André. 2000. *Syntactic aspects of topic and comment* (Linguistik Aktuell/Linguistics Today 38). Amsterdam: John Benjamins Publishing Co.
- Meisel, Jürgen. 1995. Parameters in acquisition. In Paul Fletcher & Brian MacWhinny (eds.), *The handbook of child language*, 10–35. Oxford: Blackwell Publishing Ltd.
- Mel'čuk, Igor A. 1964. *Avtomatičeskij sintaksičeskij analiz*, vol. 1. Novosibirsk: Izdatel'stvo SO AN SSSR.
- Mel'čuk, Igor A. 1981. Meaning-Text Models: A recent trend in Soviet linguistics. *Annual Review of Anthropology* 10. 27–62.
- Mel'čuk, Igor A. 1988. *Dependency Syntax: Theory and practice* (SUNY Series in Linguistics). Albany, NY: SUNY Press.
- Mel'čuk, Igor A. 2003. Levels of dependency description: Concepts and problems. In Ágel et al. (2003b) 188–230.
- Melnik, Nurit. 2007. From “hand-written” to computationally implemented HPSG theories. *Research on Language and Computation* 5(2). 199–236.
- Mensching, Guido & Eva-Maria Remberger. 2011. Syntactic variation and change in Romance: A Minimalist approach. In Peter Siemund (ed.), *Linguistic universals and language variation* (Trends in Linguistics. Studies and Monographs 231), 361–403. Berlin: Mouton de Gruyter.
- Menzel, Wolfgang. 1998. Constraint satisfaction for robust parsing of spoken language. *Journal of Experimental & Theoretical Artificial Intelligence* 10(1). 77–89.
- Menzel, Wolfgang & Ingo Schröder. 1998a. Constraint-based diagnosis for intelligent language tutoring systems. In *Proceedings of the ITF & KNOWS Conference at the IFIP '98 Congress*, Wien/Budapest.
- Menzel, Wolfgang & Ingo Schröder. 1998b. Decision procedures for Dependency Parsing using graded constraints. In Alain Polguère & Sylvain Kahane (eds.), *Processing of dependency-based grammars: Proceedings of the workshop at COLING-ACL '98*, 78–87. Association for Computational Linguistics. <http://www.aclweb.org/anthology/W/W98/#0500>.
- Meurer, Paul. 2009. A computational grammar for Georgian. In Peter Bosch, David Gabelaia & Jérôme Lang (eds.), *Logic, language, and computation: 7th International Tbilisi Symposium on Logic, Language, and Computation, TbiLLC 2007, Tbilisi, Georgia, October 2007, revised selected papers* (Lecture Notes in Artificial Intelligence 5422), 1–15. Berlin: Springer Verlag.

参考文献

- Meurers, Walt Detmar. 1994. On implementing an HPSG theory. In Erhard W. Hinrichs, Walt Detmar Meurers & Tsuneko Nakazawa (eds.), *Partial-VP and split-NP topicalization in German – An HPSG analysis and its implementation* (Arbeitspapiere des SFB 340 No. 58), 47–155. Tübingen: Eberhard-Karls-Universität. <http://www.sfs.uni-tuebingen.de/~dm/papers/on-implementing.html>.
- Meurers, Walt Detmar. 1999a. German partial-VP fronting revisited. In Webelhuth et al. (1999) 129–144.
- Meurers, Walt Detmar. 1999b. *Lexical generalizations in the syntax of German non-finite constructions*. Tübingen: Eberhard-Karls-Universität dissertation.
- Meurers, Walt Detmar. 1999c. Raising spirits (and assigning them case). *Groninger Arbeiten zur Germanistischen Linguistik (GAGL)* 43. 173–226. <http://www.sfs.uni-tuebingen.de/~dm/papers/gagl99.html>.
- Meurers, Walt Detmar. 2000. Lexical generalizations in the syntax of German non-finite constructions. Arbeitspapiere des SFB 340 No. 145 Eberhard-Karls-Universität Tübingen. <http://www.sfs.uni-tuebingen.de/~dm/papers/diss.html>.
- Meurers, Walt Detmar. 2001. On expressing lexical generalizations in HPSG. *Nordic Journal of Linguistics* 24(2). 161–217.
- Meurers, Walt Detmar, Kordula De Kuthy & Vanessa Metcalf. 2003. Modularity of grammatical constraints in HPSG-based grammar implementations. In Bender et al. (2003) 83–90. <http://www.sfs.uni-tuebingen.de/~dm/papers/meurers-dekuthy-metcalf-03.html>.
- Meurers, Walt Detmar & Tibor Kiss (eds.). 2001. *Constraint-based approaches to Germanic syntax* (Studies in Constraint-Based Lexicalism 7). Stanford, CA: CSLI Publications.
- Meurers, Walt Detmar & Stefan Müller. 2009. Corpora and syntax. In Lüdeling & Kytö (2009) chap. 42, 920–933.
- Meurers, Walt Detmar, Gerald Penn & Frank Richter. 2002. A web-based instructional platform for constraint-based grammar formalisms and parsing. In Dragomir Radev & Chris Brew (eds.), *Effective tools and methodologies for teaching NLP and CL*, 18–25. Association for Computational Linguistics. Proceedings of the Workshop held at 40th Annual Meeting of the Association for Computational Linguistics. Philadelphia, PA.
- Micelli, Vanessa. 2012. Field topology and information structure: A case study for German constituent order. In Luc Steels (ed.), *Computational issues in Fluid Construction Grammar* (Lecture Notes in Computer Science 7249), 178–211. Berlin: Springer Verlag.
- Michaelis, Jens. 2001. *On formal properties of Minimalist Grammars*: Universität Potsdam dissertation.
- Michaelis, Laura A. 2006. Construction Grammar. In Brown (2006) 73–84.
- Michaelis, Laura A. & Josef Ruppenhofer. 2001. *Beyond alternations: A Constructional model of the German applicative pattern* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.
- Miller, George A. & Kathryn Ojemann McKean. 1964. A chronometric study of some relations between sentences. *Quarterly Journal of Experimental Psychology* 16(4). 297–308.

- Mittendorf, Ingo & Louisa Sadler. 2005. Numerals, nouns and number in Welsh NPs. In Butt & King (2005) 294–312. <http://csli-publications.stanford.edu/LFG/10/>.
- Miyao, Yusuke, Takashi Ninomiya & Jun'ichi Tsujii. 2005. Corpus-oriented grammar development for acquiring a Head-Driven Phrase Structure Grammar from the Penn Treebank. In Su et al. (2005) 684–693.
- Miyao, Yusuke & Jun'ichi Tsujii. 2008. Feature forest models for probabilistic HPSG parsing. *Computational Linguistics* 34(1). 35–80.
- Moeljadi, David, Francis Bond & Sanghoun Song. 2015. Building an HPSG-based Indonesian resource grammar. In Bender et al. (2015) 9–16.
- Moens, Marc, Jo Calder, Ewan Klein, Mike Reape & Henk Zeevat. 1989. Expressing generalizations in unification-based grammar formalisms. In Somers & Wood (1989) 174–181.
- Montague, Richard. 1974. *Formal philosophy*. New Haven: Yale University Press.
- Moore, Johanna D., Simone Teufel, James Allan & Sadaoki Furui (eds.). 2008. *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. Columbus, Ohio: Association for Computational Linguistics. <http://www.aclweb.org/anthology/P/P08/>.
- Moortgat, Michael. 1989. *Categorical investigations: Logical and linguistic aspects of the Lambek Calculus* (Groningen Amsterdam Studies in Semantics 9). Dordrecht/Cinnaminson, U.S.A.: Foris Publications.
- Moortgat, Michael. 2011. Categorical type logics. In van Benthem & ter Meulen (2011) 95–179.
- Moot, Richard. 2002. *Proof nets for linguistic analysis*: University of Utrecht dissertation.
- Morgan, James L. 1989. Learnability considerations and the nature of trigger experiences in language acquisition. *Behavioral and Brain Sciences* 12(2). 352–353.
- Morin, Yves Ch. 1973. A computer tested Transformational Grammar of French. *Linguistics* 116(11). 49–114.
- Morrill, Glyn V. 1994. *Type Logical Grammars: Categorical logic of signs*. Dordrecht: Kluwer Academic Publishers.
- Morrill, Glyn V. 1995. Discontinuity in Categorical Grammar. *Linguistics and Philosophy* 18(2). 175–219.
- Morrill, Glyn V. 2012. CatLog: A Categorical parser/theorem-prover. In *Logical aspects of computational linguistics: System demonstrations*, 13–16. Nantes, France: University of Nantes.
- Müller, Gereon. 1996a. A constraint on remnant movement. *Natural Language and Linguistic Theory* 14(2). 355–407.
- Müller, Gereon. 1996b. On extraposition and successive cyclicity. In Uli Lutz & Jürgen Pafel (eds.), *On extraction and extraposition in German* (Linguistik Aktuell/Linguistics Today 11), 213–243. Amsterdam: John Benjamins Publishing Co.
- Müller, Gereon. 1998. *Incomplete category fronting: A derivational approach to remnant movement in German* (Studies in Natural Language and Linguistic Theory 42). Dordrecht: Kluwer Academic Publishers.

- Müller, Gereon. 2000. *Elemente der optimalitätstheoretischen Syntax* (Stauffenburg Linguistik 20). Tübingen: Stauffenburg Verlag.
- Müller, Gereon. 2009a. There are no Constructions. Handout Ringvorlesung: *Algorithmen und Muster: Strukturen in der Sprache*. Freie Universität Berlin, 20. Mai.
- Müller, Gereon. 2011a. Regeln oder Konstruktionen? Von verblosen Direktiven zur sequentiellen Nominalreduplikation. In Engelberg et al. (2011) 211–249.
- Müller, Gereon. 2014a. Syntactic buffers. Linguistische Arbeitsberichte 91 Institut für Linguistic Universität Leipzig. <http://www.uni-leipzig.de/~muellerg/mu765.pdf>.
- Müller, Natascha & Beate Riemer. 1998. *Generative Syntax der romanischen Sprachen: Französisch, Italienisch, Portugiesisch, Spanisch* (Stauffenburg Einführungen 17). Tübingen: Stauffenburg Verlag.
- Müller, Stefan. 1995. Scrambling in German – Extraction into the *Mittelfeld*. In Benjamin K. T'sou & Tom Bong Yeung Lai (eds.), *Proceedings of the Tenth Pacific Asia Conference on Language, Information and Computation*, 79–83. City University of Hong Kong.
- Müller, Stefan. 1996c. The Babel-System—An HPSG fragment for German, a parser, and a dialogue component. In *Proceedings of the Fourth International Conference on the Practical Application of Prolog*, 263–277. London.
- Müller, Stefan. 1996d. Yet another paper about partial verb phrase fronting in German. In Jun-ichi Tsuji (ed.), *Proceedings of COLING-96: 16th International Conference on Computational Linguistics (COLING96). Copenhagen, Denmark, August 5–9, 1996*, 800–805. Copenhagen, Denmark: Association for Computational Linguistics.
- Müller, Stefan. 1999a. *Deutsche Syntax deklarativ: Head-Driven Phrase Structure Grammar für das Deutsche* (Linguistische Arbeiten 394). Tübingen: Max Niemeyer Verlag.
- Müller, Stefan. 1999b. An HPSG-analysis for free relative clauses in German. *Grammars* 2(1). 53–105.
- Müller, Stefan. 1999c. Restricting discontinuity. In *Proceedings of the 5th Natural Language Processing Pacific Rim Symposium 1999 (NLPRS'99)*, 85–90. Peking.
- Müller, Stefan. 1999d. Restricting discontinuity. Verbmobil Report 237 Deutsches Forschungszentrum für Künstliche Intelligenz Saarbrücken. This report is also published in the Proceedings Proceedings of GLDV 99 (Frankfurt/Main). The initial part is contained in the proceedings of the 5th Natural Language Processing Pacific Rim Symposium 1999 (NLPRS'99).
- Müller, Stefan. 2001. Case in German – towards an HPSG analysis. In Walt Detmar Meurers & Tibor Kiss (eds.), *Constraint-based approaches to Germanic syntax* (Studies in Constraint-Based Lexicalism 7), 217–255. Stanford, CA: CSLI Publications.
- Müller, Stefan. 2002a. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford, CA: CSLI Publications.
- Müller, Stefan. 2002b. Multiple frontings in German. In Jäger et al. (2002) 113–124.
- Müller, Stefan. 2002c. Syntax or morphology: German particle verbs revisited. In Dehé et al. (2002) 119–139.
- Müller, Stefan. 2003a. Mehrfache Vorfeldbesetzung. *Deutsche Sprache* 31(1). 29–62.

- Müller, Stefan. 2003b. Object-to-subject-raising and lexical rule: An analysis of the German passive. In Stefan Müller (ed.), *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar*, Michigan State University, East Lansing, 278–297. Stanford, CA: CSLI Publications.
- Müller, Stefan (ed.). 2003c. *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar*, Michigan State University, East Lansing. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/HPSG/2003/>.
- Müller, Stefan. 2003d. Solving the bracketing paradox: An analysis of the morphology of German particle verbs. *Journal of Linguistics* 39(2). 275–325.
- Müller, Stefan. 2004a. An analysis of depictive secondary predicates in German without discontinuous constituents. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar*, Center for Computational Linguistics, Katholieke Universiteit Leuven, 202–222. Stanford, CA: CSLI Publications.
- Müller, Stefan. 2004b. Complex NPs, subadjacency, and extraposition. *Snippets* 8. 10–11.
- Müller, Stefan. 2004c. Continuous or discontinuous constituents? A comparison between syntactic analyses for constituent order and their processing systems. *Research on Language and Computation, Special Issue on Linguistic Theory and Grammar Implementation* 2(2). 209–257.
- Müller, Stefan. 2004d. An HPSG analysis of German depictive secondary predicates. In Lawrence S. Moss & Richard T. Oehrle (eds.), *Proceedings of the joint meeting of the 6th Conference on Formal Grammar and the 7th Conference on Mathematics of Language* (Electronic Notes in Theoretical Computer Science 53), 233–245. Helsinki: Elsevier Science Publisher B.V. (North-Holland). DOI:10.1016/S1571-0661(05)82585-X.
- Müller, Stefan (ed.). 2004e. *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar*, Center for Computational Linguistics, Katholieke Universiteit Leuven. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/HPSG/2004/>.
- Müller, Stefan. 2005a. Resultative Constructions: Syntax, world knowledge, and collocational restrictions: Review of Hans C. Boas: A Constructional approach to resultatives. *Studies in Language* 29(3). 651–681.
- Müller, Stefan. 2005b. Zur Analyse der deutschen Satzstruktur. *Linguistische Berichte* 201. 3–39.
- Müller, Stefan. 2005c. Zur Analyse der scheinbar mehrfachen Vorfeldbesetzung. *Linguistische Berichte* 203. 297–330.
- Müller, Stefan. 2006. Phrasal or lexical Constructions? *Language* 82(4). 850–883.
- Müller, Stefan. 2007a. The Grammex CD Rom: A software collection for developing typed feature structure grammars. In King & Bender (2007) 259–266. <http://csli-publications.stanford.edu/GEAF/2007/>.
- Müller, Stefan. 2007b. *Head-Driven Phrase Structure Grammar: Eine Einführung* (Stauffenburg Einführungen 17). Tübingen: Stauffenburg Verlag 1st edn.
- Müller, Stefan. 2007c. Phrasal or lexical Constructions: Some comments on underspecification of constituent order, compositionality, and control. In Stefan Müller (ed.),

- Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 373–393. Stanford, CA: CSLI Publications.
- Müller, Stefan (ed.). 2007d. *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/HPSG/2007/>.
- Müller, Stefan. 2007e. Qualitative Korpusanalyse für die Grammatiktheorie: Introspektion vs. Korpus. In Gisela Zifonun & Werner Kallmeyer (eds.), *Sprachkorpora – Datenmengen und Erkenntnisfortschritt* (Institut für Deutsche Sprache Jahrbuch 2006), 70–90. Berlin: Walter de Gruyter.
- Müller, Stefan. 2008a. Depictive secondary predicates in German and English. In Christoph Schroeder, Gerd Hentschel & Winfried Boeder (eds.), *Secondary predicates in Eastern European languages and beyond* (Studia Slavica Oldenburgensia 16), 255–273. Oldenburg: BIS-Verlag.
- Müller, Stefan (ed.). 2008b. *Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar*. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/HPSG/2008/>.
- Müller, Stefan. 2009b. A Head-Driven Phrase Structure Grammar for Maltese. In Bernard Comrie, Ray Fabri, Beth Hume, Manwel Mifsud, Thomas Stolz & Martine Vanhove (eds.), *Introducing Maltese linguistics: Papers from the 1st International Conference on Maltese Linguistics (Bremen/Germany, 18–20 October, 2007)* (Studies in Language Companion Series 113), 83–112. Amsterdam: John Benjamins Publishing Co.
- Müller, Stefan. 2009c. On predication. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 213–233. Stanford, CA: CSLI Publications.
- Müller, Stefan (ed.). 2009d. *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*. Stanford, CA: CSLI Publications.
- Müller, Stefan. 2010a. *Grammatiktheorie* (Stauffenburg Einführungen 20). Tübingen: Stauffenburg Verlag.
- Müller, Stefan. 2010b. Persian complex predicates and the limits of inheritance-based analyses. *Journal of Linguistics* 46(3). 601–655.
- Müller, Stefan (ed.). 2011b. *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar, University of Washington*. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/HPSG/2011/>.
- Müller, Stefan. 2012. On the copula, specificational constructions and type shifting. Ms. Freie Universität Berlin.
- Müller, Stefan. 2013a. The CoreGram project: A brief overview and motivation. In Denys Duchier & Yannick Parmentier (eds.), *Proceedings of the workshop on high-level methodologies for grammar engineering (HMGE 2013), Düsseldorf*, 93–104.
- Müller, Stefan. 2013b. *Grammatiktheorie* (Stauffenburg Einführungen 20). Tübingen: Stauffenburg Verlag 2nd edn.
- Müller, Stefan. 2013c. Unifying everything: Some remarks on Simpler Syntax, Construction Grammar, Minimalism and HPSG. *Language* 89(4). 920–950.

- Müller, Stefan. 2014b. Artenvielfalt und Head-Driven Phrase Structure Grammar. In Jörg Hagemann & Sven Staffeldt (eds.), *Syntaxtheorien: Analysen im Vergleich* (Stauffenburg Einführungen 28), 187–233. Tübingen: Stauffenburg Verlag.
- Müller, Stefan. 2014c. Elliptical constructions, multiple frontings, and surface-based syntax. In Paola Monachesi, Gerhard Jäger, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2004, Nancy*, 91–109. Stanford, CA: CSLI Publications.
- Müller, Stefan. 2014d. Kernigkeit: Anmerkungen zur Kern-Peripherie-Unterscheidung. In Antonio Machicao y Priemer, Andreas Nolda & Athina Sioupi (eds.), *Zwischen Kern und Peripherie* (studia grammatica 76), 25–39. Berlin: de Gruyter.
- Müller, Stefan. 2015a. The CoreGram project: Theoretical linguistics, theory development and verification. *Journal of Language Modelling* 3(1). 21–86. DOI:10.15398/jlm.v3i1.91.
- Müller, Stefan. 2015b. *German sentence structure: An analysis with special consideration of so-called multiple fronting* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. Submitted.
- Müller, Stefan. 2015c. HPSG – A synopsis. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – Theory and analysis: An international handbook* (Handbooks of Linguistics and Communication Science 42.2), 937–973. Berlin: Walter de Gruyter 2nd edn.
- Müller, Stefan (ed.). 2015d. *Proceedings of the 22nd International Conference on Head-Driven Phrase Structure Grammar, Nanyang Technological University (NTU), Singapore*. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/HPSG/2015/>.
- Müller, Stefan. 2016a. *Germanic syntax* (Textbooks in Language Sciences). Berlin: Language Science Press. In Preparation.
- Müller, Stefan. 2016b. Satztypen: Lexikalisch oder/und phrasal. In Rita Finkbeiner & Jörg Meibauer (eds.), *Satztypen und Konstruktionen im Deutschen* (Linguistik – Impulse und Tendenzen 65), 72–105. Berlin, Boston: de Gruyter.
- Müller, Stefan & Masood Ghayoomi. 2010. PerGram: A TRALE implementation of an HPSG fragment of Persian. In *Proceedings of 2010 IEEE International Multiconference on Computer Science and Information Technology – Computational Linguistics Applications (CLA'10)*. Wisła, Poland, 18–20 October 2010, vol. 5, 461–467. Polish Information Processing Society.
- Müller, Stefan & Martin Haspelmath. 2013. Language Science Press: A publication model for open-access books in linguistics. Grant Proposal to the DFG.
- Müller, Stefan & Walter Kasper. 2000. HPSG analysis of German. In Wolfgang Wahlster (ed.), *Verbmobil: Foundations of speech-to-speech translation* (Artificial Intelligence), 238–253. Berlin: Springer Verlag.
- Müller, Stefan & Janna Lipenkova. 2009. Serial verb constructions in Chinese: An HPSG account. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 234–254. Stanford, CA: CSLI Publications.
- Müller, Stefan & Janna Lipenkova. 2013. ChinGram: A TRALE implementation of an HPSG fragment of Mandarin Chinese. In Huei ling Lai & Kawai Chui (eds.), *Proceedings of the 27th Pacific Asia Conference on Language, Information, and Computation*

- (PACLIC 27), 240–249. Taipei, Taiwan: Department of English, National Chengchi University.
- Müller, Stefan & Janna Lipenkova. 2016. *Mandarin Chinese in Head-Driven Phrase Structure Grammar* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. In Preparation.
- Müller, Stefan & Bjarne Ørsnes. 2011. Positional expletives in Danish, German, and Yiddish. In Stefan Müller (ed.), *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar, University of Washington, U.S.A.*, 167–187. Stanford, CA: CSLI Publications.
- Müller, Stefan & Bjarne Ørsnes. 2013a. Passive in Danish, English, and German. In Stefan Müller (ed.), *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin*, 140–160. Stanford, CA: CSLI Publications.
- Müller, Stefan & Bjarne Ørsnes. 2013b. Towards an HPSG analysis of object shift in Danish. In Glyn Morrill & Mark-Jan Nederhof (eds.), *Formal Grammar: 17th and 18th International Conferences, FG 2012, Opole, Poland, August 2012, revised selected papers, FG 2013, Düsseldorf, Germany, August 2013: Proceedings* (Lecture Notes in Computer Science 8036), 69–89. Berlin: Springer Verlag.
- Müller, Stefan & Bjarne Ørsnes. 2015. *Danish in Head-Driven Phrase Structure Grammar* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. In Preparation.
- Müller, Stefan & Stephen Mark Wechsler. 2014a. Lexical approaches to argument structure. *Theoretical Linguistics* 40(1–2). 1–76.
- Müller, Stefan & Stephen Mark Wechsler. 2014b. Two sides of the same slim Boojum: Further arguments for a lexical approach to argument structure. *Theoretical Linguistics* 40(1–2). 187–224.
- Muraki, Kazunori, Shunji Ichiyama & Yasutomo Fukumochi. 1985. Augmented Dependency Grammar: A simple interface between the grammar rule and the knowledge. In King (1985) 198–204. <http://aclweb.org/anthology/E/E85/>.
- Musso, Mariacristina, Andrea Moro, Volkmar Glauche, Michel Rijntjes, Jürgen Reichenbach, Christian Büchel & Cornelius Weiller. 2003. Broca’s area and the language instinct. *Nature Neuroscience* 6(7). 774–781.
- Müürisep, Kaili. 1999. Determination of syntactic functions in Estonian Constraint Grammar. In Henry S. Thompson & Alex Lascarides (eds.), *Ninth conference of the European Chapter of the Association for Computational Linguistics*, 291–292. Bergen, Norway.
- Müürisep, Kaili, Tiina Puolakainen, Kadri Muischnek, Mare Koit, Tiit Roosmaa & Heli Uibo. 2003. A new language for Constraint Grammar: Estonian. In *International Conference Recent Advances in Natural Language Processing*, 304–310.
- Muysken, Peter. 1982. Parameterizing the notion of “head”. *Journal of Linguistic Research* 2. 57–75.
- Mykowiecka, Agnieszka, Małgorzata Marciniak, Adam Przepiórkowski & Anna Kupść. 2003. An implementation of a Generative Grammar of Polish. In Peter Kosta, Joanna Błaszczak, Jens Frasek, Ljudmila Geist & Marzena Żygis (eds.), *Investigations into formal Slavic linguistics: Contributions of the Fourth European Conference on Formal De-*

- scription of Slavic Languages – FDSL IV held at Potsdam University, November 28–30, 2001*, 271–285. Frankfurt am Main: Peter Lang.
- Nagao, Makoto (ed.). 1986. *Proceedings of COLING 86*. University of Bonn: Association for Computational Linguistics.
- Nagao, Makoto (ed.). 1994. *Proceedings of COLING 94*. Kyoto, Japan: Association for Computational Linguistics.
- Naumann, Sven. 1987. Ein einfacher Parser für generalisierte Phrasenstrukturgrammatiken. *Zeitschrift für Sprachwissenschaft* 6(2). 206–226.
- Naumann, Sven. 1988. *Generalisierte Phrasenstrukturgrammatik: Parsingstrategien, Regelorganisation und Unifikation* (Linguistische Arbeiten 212). Tübingen: Max Niemeyer Verlag.
- Neeleman, Ad. 1994. *Complex predicates*. Utrecht: Onderzoeksinstituut voor Taal en Spraak (OTS) dissertation.
- Nelimarkka, Esa, Harri Jäppinen & Aarno Lehtola. 1984. Two-way finite automata and Dependency Grammar: A parsing method for inflectional free word order languages. In Wilks (1984) 389–392.
- Nerbonne, John. 1986a. ‘Phantoms’ and German fronting: Poltergeist constituents? *Linguistics* 24(5). 857–870.
- Nerbonne, John. 1986b. A phrase-structure grammar for German passives. *Linguistics* 24(5). 907–938.
- Nerbonne, John. 1993. A feature-based syntax/semantics interface. *Annals of Mathematics and Artificial Intelligence* 8(1–2). 107–132. Special issue on Mathematics of Language edited by Alexis Manaster-Ramer and Wlodek Zadrozny, selected from the 2nd Conference on Mathematics of Language. Also published as DFKI Research Report RR-92-42.
- Nerbonne, John, Klaus Netter & Carl J. Pollard (eds.). 1994a. *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46). Stanford, CA: CSLI Publications.
- Nerbonne, John, Klaus Netter & Carl J. Pollard (eds.). 1994b. *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46). Stanford, CA: CSLI Publications.
- Netter, Klaus. 1991. Clause union phenomena and complex predicates in German. DYANA Report, Deliverable R1.1.B University of Edinburgh.
- Netter, Klaus. 1992. On non-head non-movement: An HPSG treatment of finite verb position in German. In Günther Görz (ed.), *Konvens 92. 1. Konferenz „Verarbeitung natürlicher Sprache“*. Nürnberg 7.–9. Oktober 1992 (Informatik aktuell), 218–227. Berlin: Springer Verlag.
- Netter, Klaus. 1993. Architecture and coverage of the DISCO Grammar. In Stephan Busemann & Karin Harbusch (eds.), *DFKI Workshop on Natural Language Systems: Re-Usability and Modularity, October 23* (DFKI Document D-93-03), 1–10. Saarbrücken, Germany: DFKI.
- Netter, Klaus. 1994. Towards a theory of functional heads: German nominal phrases. In Nerbonne et al. (1994b) 297–340.
- Netter, Klaus. 1996. *Functional categories in an HPSG for German*. Saarbrücken: Universität des Saarlandes Dissertation.

- Netter, Klaus. 1998. *Functional categories in an HPSG for German* (Saarbrücken Dissertations in Computational Linguistics and Language Technology 3). Saarbrücken: Deutsches Forschungszentrum für Künstliche Intelligenz Universität des Saarlandes.
- Neville, Anne & Patrizia Paggio. 2004. Developing a Danish grammar in the GRASP project: A construction-based approach to topology and extraction in Danish. In Lawrence S. Moss & Richard T. Oehrle (eds.), *Proceedings of the joint meeting of the 6th Conference on Formal Grammar and the 7th Conference on Mathematics of Language* (Electronic Notes in Theoretical Computer Science 53), 246–259. Helsinki: Elsevier Science Publisher B.V. (North-Holland).
- Nevins, Andrew Ira, David Pesetsky & Cilene Rodrigues. 2009. Pirahã exceptionality: A reassessment. *Language* 85(2). 355–404.
- Newmeyer, Frederick J. 2004a. Against a parameter-setting approach to language variation. *Linguistic Variation Yearbook* 4. 181–234.
- Newmeyer, Frederick J. 2004b. Typological evidence and Universal Grammar. *Studies in Language* 28(3). 527–548.
- Newmeyer, Frederick J. 2005. *Possible and probable languages: A Generative perspective on linguistic typology*. Oxford: Oxford University Press.
- Newmeyer, Frederick J. 2010. On comparative concepts and descriptive categories: A reply to Haspelmath. *Language* 86(3). 688–695.
- Newport, Elissa L. 1990. Maturational constraints on language learning. *Cognitive Science* 14(1). 11–28.
- Ng, Say Kiat. 1997. *A double-specifier account of Chinese NPs using Head-Driven Phrase Structure Grammar*. University of Edinburgh, Department of Linguistics MSc speech and language processing.
- Nivre, Joakim. 2003. An efficient algorithm for projective dependency parsing. In van Noord (2003).
- Nolda, Andreas. 2007. *Die Thema-Integration: Syntax und Semantik der gespaltenen Topikalisierung im Deutschen* (Studien zur deutschen Grammatik 72). Tübingen: Stauffenburg Verlag.
- Noonan, Michael. 1994. A tale of two passives in Irish. In Barbara Fox & Paul J. Hopper (eds.), *Voice: Form and function* (Typological Studies in Language 27), 279–311. Amsterdam: John Benjamins Publishing Co.
- van Noord, Gertjan (ed.). 2003. *Proceedings of the 8th International Workshop on Parsing Technologies (IWPT 03)*. Nancy.
- van Noord, Gertjan & Gosse Bouma. 1994. The scope of adjuncts and the processing of lexical rules. In Nagao (1994) 250–256.
- Nordgård, Torbjørn. 1994. E-Parser: An implementation of a deterministic GB-related parsing system. *Computers and the Humanities* 28(4–5). 259–272.
- Nordlinger, Rachel. 1998. *Constructive case: Evidence from Australia* (Dissertations in Linguistics). Stanford, CA: CSLI Publications.
- Nowak, Martin A., Natalia L. Komarova & Partha Niyogi. 2001. Evolution of Universal Grammar. *Science* 291(5501). 114–118.
- Nozohoor-Farshi, R. 1986. On formalizations of Marcus' parser. In Nagao (1986) 533–535.

- Nozohoor-Farshi, R. 1987. Context-freeness of the language accepted by Marcus' parser. In Sidner (1987) 117–122.
- Nunberg, Geoffrey. 1995. Transfers of meaning. *Journal of Semantics* 12(2). 109–132.
- Nunberg, Geoffrey, Ivan A. Sag & Thomas Wasow. 1994. Idioms. *Language* 70(3). 491–538.
- Nunes, Jairo. 2004. *Linearization of chains and Sideward Movement* (Linguistic Inquiry Monographs 43). Cambridge, MA/London, England: MIT Press.
- Ochs, Elinor. 1982. Talking to children in Western Samoa. *Language and Society* 11(1). 77–104.
- Ochs, Elinor & Bambi B. Schieffelin. 1985. Language acquisition and socialization: Three developmental stories. In Richard A. Shweder & Robert A. LeVine (eds.), *Culture theory: Essays in mind, self and emotion*, 276–320. Cambridge, UK: Cambridge University Press.
- O'Donovan, Ruth, Michael Burke, Aoife Cahill, Josef van Genabith & Andy Way. 2005. Large-scale induction and evaluation of lexical resources from the Penn-II and Penn-III Treebanks. *Computational Linguistics* 31(3). 328–365.
- Oehrle, Richard, Emmon Bach & Deirdre Wheeler (eds.). 1988. *Categorial Grammars and natural language structures*. Dordrecht: D. Reidel Publishing Company.
- Oepen, Stephan & Daniel P. Flickinger. 1998. Towards systematic grammar profiling: Test suite technology ten years after. *Journal of Computer Speech and Language* 12(4). 411–436. <http://www.delph-in.net/itsdb/publications/profiling.ps.gz>. (Special Issue on Evaluation).
- Özkaragöz, İnci. 1986. Monoclausal double passives in Turkish. In Dan I. Slobin & Karl Zimmer (eds.), *Studies in Turkish linguistics* (Typological Studies in Language 8), 77–91. Amsterdam: John Benjamins Publishing Co.
- Oliva, Karel. 1992. Word order constraints in binary branching syntactic structures. CLAUS-Report 20 Universität des Saarlandes Saarbrücken.
- Oliva, Karel. 2003. Dependency, valency and Head-Driven Phrase-Structure Grammar. In Ágel et al. (2003b) 660–668.
- O'Neill, Michael & Randall Wood. 2012. The grammar of happiness. Essential Media & Entertainment / Smithsonian Networks. <https://www.youtube.com/watch?v=er2VAk4uXUs#t=1519>.
- Oppenrieder, Wilhelm. 1991. *Von Subjekten, Sätzen und Subjektsätzen* (Linguistische Arbeiten 241). Tübingen: Max Niemeyer Verlag.
- Orgun, Cemil Orhan. 1996. *Sign-based morphology and phonology*: University of California, Berkeley dissertation.
- Osborne, Miles & Ted J. Briscoe. 1997. Learning Stochastic Categorial Grammars. In T. Mark Ellison (ed.), *CoNLL97: Computational Natural Language Learning: Proceedings of the 1997 meeting of the ACL Special Interest Group in Natural Language Learning*, 80–87. Madrid: Association for Computational Linguistics. <http://www.aclweb.org/anthology-new/signll.html#1997-0>.
- Osborne, Timothy. 2014. Dependency Grammar. In Carnie et al. (2014) 604–626.
- Osborne, Timothy. 2016. Ellipsis in Dependency Grammar. In Jeroen van Craenenbroek (ed.), *Oxford handbook of ellipsis*, Oxford: Oxford University Press. To appear.

- Osborne, Timothy & Thomas M. Groß. 2012. Constructions are catenae: Construction Grammar meets Dependency Grammar. *Cognitive Linguistics* 23(1). 165–216.
- Osborne, Timothy & Sylvain Kahane. 2015. Translators' introduction. In *Elements of structural syntax*, xxix–lxxiii. Amsterdam: John Benjamins Publishing Co. Translated by Timothy Osborne and Sylvain Kahane.
- Osborne, Timothy, Michael Putnam & Thomas M. Groß. 2011. Bare Phrase Structure, label-less trees, and specifier-less syntax: Is Minimalism becoming a Dependency Grammar? *The Linguistic Review* 28(3). 315–364.
- Osenova, Petya. 2010a. Bulgarian Resource Grammar – efficient and realistic (BURGER). Tech. rep. LingoLab, CSLI Stanford. <http://www.bultreebank.org/BURGER/BURGER3.pdf>.
- Osenova, Petya. 2010b. *Bulgarian Resource Grammar: Modeling Bulgarian in HPSG*. Saarbrücken: VDM Verlag Dr. Müller.
- Osenova, Petya. 2011. Localizing a core HPSG-based grammar for Bulgarian. In Hedeland et al. (2011) 175–182.
- Ott, Dennis. 2011. A note on free relative clauses in the theory of Phases. *Linguistic Inquiry* 42(1). 183–192.
- Ørsnes, Bjarne. 1995. *The derivation and compounding of complex event nominals in Modern Danish: An HPSG approach with an implementation in Prolog*: University of Copenhagen dissertation.
- Ørsnes, Bjarne. 2002. Case marking and subject extraction in Danish. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2002 conference*, 333–353. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/LFG/7/>.
- Ørsnes, Bjarne. 2009a. Das Verbalfeldmodell: Ein Stellungsfeldermodell für den kontrastiven DaF-Unterricht. *Deutsch als Fremdsprache* 46(3). 143–149.
- Ørsnes, Bjarne. 2009b. Preposed negation in Danish. In Müller (2009d) 255–275.
- Ørsnes, Bjarne & Jürgen Wedekind. 2003. Parallele datamatiske grammatikker for Norsk og Dansk [parallel computational grammars for Norwegian and Danish]. In Henrik Holmboe (ed.), *Årbog for nordisk sprogteknologisk forskningsprogram 2000–2004*, 113–130. Copenhagen: Museum Tusculanums Forlag.
- Ørsnes, Bjarne & Jürgen Wedekind. 2004. Parallele datamatiske grammatikker for Norsk og Dansk: Analyse og disambiguering af modalverber [parallel computational grammars for Norwegian and Danish: Analysis and disambiguation of modal verbs]. In Henrik Holmboe (ed.), *Årbog for Nordisk Sprogteknologisk forskningsprogram 2000–2004*, 165–182. Copenhagen: Museum Tusculanums Forlag.
- Pafel, Jürgen. 1993. Ein Überblick über die Extraktion aus Nominalphrasen im Deutschen. In Franz-Josef d'Avis, Sigrid Beck, Uli Lutz, Jürgen Pafel & Susanne Trissler (eds.), *Extraktion im Deutschen I* (Arbeitspapiere des SFB 340 No. 34), 191–245. Tübingen: Eberhard-Karls-Universität Tübingen.
- Paggio, Patrizia. 2005. Representing information structure in a formal grammar of Danish. In *Proceedings of the 2nd International Workshop on Logic and Engineering of Natural Language Semantics (LENLS2005)*. Kitakyushu, Japan. June 13–14, .

- Parmentier, Yannick, Laura Kallmeyer, Wolfgang Maier, Timm Lichte & Johannes Dellert. 2008. TuLiPA: A syntax-semantics parsing environment for mildly context-sensitive formalisms. In *Proceedings of the Ninth International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+9)*, 121–128. Tübingen. <http://www.sfs.uni-tuebingen.de/~lk/papers/tag+9-parmentier-et-al.pdf>.
- Partee, Barbara H. 1987. Noun phrase interpretation and type-shifting principles. In Jeroen A. G. Groenendijk, Dick de Jongh & Martin J. B. Stokhof (eds.), *Studies in Discourse Representation Theory and the theory of generalized quantifiers*, 115–143. Dordrecht: Foris Publications.
- Patejuk, Agnieszka & Adam Przepiórkowski. 2012. Towards an LFG parser for Polish: An exercise in parasitic grammar development. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation, LREC 2012*, 3849–3852. Istanbul, Turkey: ELRA.
- Paul, Hermann. 1919. *Deutsche Grammatik. Teil IV: Syntax*, vol. 3. Halle an der Saale: Max Niemeyer Verlag. 2nd unchanged edition 1968, Tübingen: Max Niemeyer Verlag.
- Paul, Soma. 2004. *An HPSG account of Bangla compound verbs with LKB implementation*. Hyderabad, India: CALTS, University of Hyderabad, India dissertation.
- Peeters, Bert (ed.). 2000. *The lexicon–encyclopedia interface*. Amsterdam: Elsevier.
- Penn, Gerald. 2004. Balancing clarity and efficiency in typed feature logic through delaying. In Donia Scott (ed.), *Proceedings of the 42nd Meeting of the Association for Computational Linguistics (ACL'04), main volume*, 239–246. Barcelona, Spain.
- Penn, Gerald & Bob Carpenter. 1999. ALE for speech: A translation prototype. In Géza Gordos (ed.), *Proceedings of the 6th Conference on Speech Communication and Technology (EUROSPEECH)*, Budapest, Hungary.
- Perlmutter, David M. 1978. Impersonal passives and the Unaccusative Hypothesis. In *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society*, 157–189. Berkeley Linguistic Society.
- Perlmutter, David M. (ed.). 1983. *Studies in relational grammar*, vol. 1. Chicago: The University of Chicago Press.
- Perlmutter, David M. (ed.). 1984. *Studies in relational grammar*, vol. 2. Chicago: The University of Chicago Press.
- Perlmutter, David M. & John Robert Ross. 1970. Relative clauses with split antecedents. *Linguistic Inquiry* 1(3). 350.
- Pesetsky, David. 1996. *Zero syntax: Experiencers and cascades*. Cambridge, MA: MIT Press.
- Peters, Stanley & R. W. Ritchie. 1973. On the generative power of Transformational Grammar. *Information Sciences* 6(C). 49–83.
- Petrick, Stanley Roy. 1965. *A recognition procedure for Transformational Grammars*: Massachusetts Institute of Technology. Dept. of Modern Languages dissertation. <http://hdl.handle.net/1721.1/13013>.
- Phillips, Colin. 2003. Linear order and constituency. *Linguistic Inquiry* 34(1). 37–90.
- Phillips, John D. 1992. A computational representation for Generalised Phrase Structure Grammars. *Linguistics and Philosophy* 15(3). 255–287.

- Phillips, John D. & Henry S. Thompson. 1985. GPSGP – A parser for Generalized Phrase Structure Grammar. *Linguistics* 23(2). 245–261.
- Piattelli-Palmarini, Massimo (ed.). 1980. *Language and learning: The debate between Jean Piaget and Noam Chomsky*. Cambridge: Harvard University Press.
- Pickering, Martin & Guy Barry. 1993. Dependency Categorical Grammar and coordination. *Linguistics* 31(5). 855–902.
- Pienemann, Manfred. 2005. An introduction to Processability Theory. In Manfred Pienemann (ed.), *Cross-linguistic aspects of processability theory*, 1–60. Amsterdam: John Benjamins Publishing Co.
- Piñango, Maria Mercedes, Jennifer Mack & Ray S. Jackendoff. 2006. Semantic combinatorial processes in argument structure: Evidence from light-verbs. In *Proceedings of the 32nd Annual Meeting of the Berkeley Linguistics Society: Theoretical approaches to argument structure*, vol. 32, Berkeley, CA: BLS.
- Pineda, Luis Alberto & Iván V. Meza. 2005a. A computational model of the Spanish clitic system. In Alexander Gelbukh (ed.), *Computational linguistics and intelligent language processing*, 73–82. Berlin: Springer Verlag.
- Pineda, Luis Alberto & Iván V. Meza. 2005b. The Spanish pronominal clitic system. *Procesamiento del Lenguaje Natural* 34. 67–103.
- Pinker, Steven. 1984. *Learnability and cognition: The acquisition of argument structure*. London/Cambridge, MA: MIT Press.
- Pinker, Steven. 1994. *The language instinct: How the mind creates language*. New York: William Morrow.
- Pinker, Steven & Ray S. Jackendoff. 2005. The faculty of language: What’s special about it? *Cognition* 95(2). 201–236.
- Piñón, Christopher (ed.). 2012. *Empirical issues in syntax and semantics*, vol. 9. Paris: CNRS.
- Pittner, Karin. 1995. Regeln für die Bildung von freien Relativsätzen: Eine Antwort an Oddleif Leirbukt. *Deutsch als Fremdsprache* 32(4). 195–200.
- Plank, Frans & Elena Filimonova. 2000. The universals archive: A brief introduction for prospective users. *Sprachtypologie und Universalienforschung* 53(1). 109–123.
- Poletto, Cecilia. 2000. *The higher functional field: Evidence from Northern Italian Dialects*. Oxford: Oxford University Press.
- Pollard, Carl J. 1984. *Generalized Phrase Structure Grammars, Head Grammars, and natural language*: Stanford University dissertation.
- Pollard, Carl J. 1988. Categorical Grammar and Phrase Structure Grammar: An excursion on the syntax-semantics frontier. In Oehrle et al. (1988) 391–415.
- Pollard, Carl J. 1994. Toward a unified account of passive in German. In Nerbonne et al. (1994b) 273–296.
- Pollard, Carl J. 1996a. The nature of constraint-based grammar. Paper presented at the Pacific Asia Conference on Language, Information, and Computation, Kyung Hee University, Seoul, Korea. <http://lingo.stanford.edu/sag/L221a/pollard-96.txt>.

- Pollard, Carl J. 1996b. On head non-movement. In Harry Bunt & Arthur van Horck (eds.), *Discontinuous constituency* (Natural Language Processing 6), 279–305. Berlin: Mouton de Gruyter. Published version of a Ms. dated January 1990.
- Pollard, Carl J. 1999. Strong generative capacity in HPSG. In Webelhuth et al. (1999) 281–298.
- Pollard, Carl J. & Andrew M. Moshier. 1990. Unifying partial descriptions of sets. In Philip P. Hanson (ed.), *Information, language and cognition* (Vancouver Studies in Cognitive Science 1), 285–322. Vancouver: University of British Columbia Press.
- Pollard, Carl J. & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.
- Pollard, Carl J. & Ivan A. Sag. 1992. Anaphors in English and the scope of Binding Theory. *Linguistic Inquiry* 23(2). 261–303.
- Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.
- Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar and the structure of IP. *Linguistic Inquiry* 20(3). 365–424.
- Popowich, Fred & Carl Vogel. 1991. A logic based implementation of Head-Driven Phrase Structure Grammar. In Charles Grant Brown & Gregers Koch (eds.), *Natural Language Understanding and Logic Programming, III. The 3rd International Workshop, Stockholm, Sweden, 23–25 Jan., 1991*, 227–246. Amsterdam: Elsevier, North-Holland.
- Porzel, Robert, Vanessa Micelli, Hidir Aras & Hans-Peter Zorn. 2006. Tying the knot: Ground entities, descriptions and information objects for Construction-based information extraction. In *Proceedings of the OntoLex Workshop at LREC, May 2006. Genoa, Italy*, 35–40.
- Postal, Paul M. 2004. *Skeptical linguistic essays*. Oxford: Oxford University Press.
- Postal, Paul M. 2009. The incoherence of Chomsky’s ‘Biolinguistic’ ontology. *Biolinguistics* 3(1). 104–123.
- Postal, Paul M. & Geoffrey K. Pullum. 1986. Misgovernment. *Linguistic Inquiry* 17(1). 104–110.
- Prince, Alan & Paul Smolensky. 1993. Optimality Theory: Constraint interaction in Generative Grammar. RuCCS Technical Report 2 Center for Cognitive Science, Rutgers University, Piscataway, N.J., and Computer Science Department, University of Colorado, Boulder. <http://roa.rutgers.edu/files/537-0802/537-0802-PRINCE-0-0.PDF>.
- Przepiórkowski, Adam. 1999a. *Case assignment and the complement-adjunct dichotomy: A non-configurational constraint-based approach*: Eberhard-Karls-Universität Tübingen dissertation. <https://publikationen.uni-tuebingen.de/xmlui/handle/10900/46147>.
- Przepiórkowski, Adam. 1999b. On case assignment and “adjuncts as complements”. In Webelhuth et al. (1999) 231–245.
- Przepiórkowski, Adam & Anna Kupść. 2006. HPSG for Slavists. *Glossos* 8. 1–68.
- Przepiórkowski, Adam, Anna Kupść, Małgorzata Marciniak & Agnieszka Mykowiecka. 2002. *Formalny opis języka polskiego: Teoria i implementacja*. Warsaw: Akademicka Oficyna Wydawnicza EXIT.

参考文献

- Pullum, Geoffrey K. 1977. Word order universals and grammatical relations. In Peter Cole & Jerrold M. Sadock (eds.), *Grammatical relations* (Syntax and Semantics 8), 249–277. New York, San Francisco, London: Academic Press.
- Pullum, Geoffrey K. 1982. Free word order and phrase structure rules. In James Pustejovsky & Peter Sells (eds.), *Proceedings of the 12th Annual Meeting of the Northeast Linguistic Society*, 209–220. Amherst: Graduate Linguistics Student Association.
- Pullum, Geoffrey K. 1983. How many possible human languages are there? *Linguistic Inquiry* 14(3). 447–467.
- Pullum, Geoffrey K. 1984. Stalking the perfect journal. *Natural Language and Linguistic Theory* 2(2). 261–267.
- Pullum, Geoffrey K. 1985. Assuming some version of X-bar Theory. In *Papers from the 21st Annual Meeting of the Chicago Linguistic Society*, 323–353.
- Pullum, Geoffrey K. 1986. Footloose and context-free. *Natural Language and Linguistic Theory* 4(3). 409–414.
- Pullum, Geoffrey K. 1988. Citation etiquette beyond thunderdome. *Natural Language and Linguistic Theory* 6(4). 579–588.
- Pullum, Geoffrey K. 1989a. Formal linguistics meets the Boojum. *Natural Language and Linguistic Theory* 7(1). 137–143. DOI:10.1007/BF00141350.
- Pullum, Geoffrey K. 1989b. The incident of the node vortex problem. *Natural Language and Linguistic Theory* 7(3). 473–479.
- Pullum, Geoffrey K. 1991. *The great Eskimo vocabulary hoax and other irreverent essays on the study of language*. Chicago: The University of Chicago Press.
- Pullum, Geoffrey K. 1996. Learnability, hyperlearning, and the Poverty of the Stimulus. In J. Johnson, M. L. Juge & J. L. Moxley (eds.), *Proceedings of the 22nd Annual Meeting of the Berkeley Linguistics Society: General session and parasession on the role of learnability in grammatical theory*, 498–513. Berkeley, CA: Berkeley Linguistic Society. <http://users.ecs.soton.ac.uk/harnad/Papers/Py104/pullum.learn.html>.
- Pullum, Geoffrey K. 2003. Learnability: Mathematical aspects. In William J. Frawley (ed.), *Oxford international encyclopedia of linguistics*, 431–434. Oxford: Oxford University Press 2nd edn.
- Pullum, Geoffrey K. 2007. The evolution of model-theoretic frameworks in linguistics. In Rogers & Kepsar (2007) 1–10.
- Pullum, Geoffrey K. 2009. Response to Anderson. *Language* 85(2). 245–247.
- Pullum, Geoffrey K. 2013. The central question in comparative syntactic metatheory. *Mind and Language* 28(4). 492–521.
- Pullum, Geoffrey K. & Barbara C. Scholz. 2001. On the distinction between Generative-Enumerative and Model-Theoretic syntactic frameworks. In Philippe de Groote, Glyn Morrill & Christian Retor (eds.), *Logical Aspects of Computational Linguistics: 4th International Conference* (Lecture Notes in Computer Science 2099), 17–43. Berlin: Springer Verlag.
- Pullum, Geoffrey K. & Barbara C. Scholz. 2002. Empirical assessment of stimulus poverty arguments. *The Linguistic Review* 19(1–2). 9–50.

- Pullum, Geoffrey K. & Barbara C. Scholz. 2010. Recursion and the infinitude claim. In van der Hulst (2010) 113–138.
- Pulman, Stephen G. 1985. A parser that doesn't. In King (1985) 128–135. <http://aclweb.org/anthology/E/E85/>.
- Pulvermüller, Friedemann. 2003. *The neuroscience of language: On brain circuits of words and serial order*. Cambridge, UK: Cambridge University Press.
- Pulvermüller, Friedemann. 2010. Brain embodiment of syntax and grammar: Discrete combinatorial mechanisms spelt out in neuronal circuits. *Brain & Language* 112(3). 167–179.
- Pulvermüller, Friedemann, Bert Cappelle & Yury Shtyrov. 2013. Brain basis of meaning, words, constructions, and grammar. In Hoffmann & Trousdale (2013) 397–416.
- Quaglia, Stefano. 2014. On the syntax of some apparent spatial particles in Italian. In Butt & King (2014) 503–523.
- Radford, Andrew. 1990. *Syntactic theory and the acquisition of English syntax*. Cambridge, MA: Blackwell Publishing Ltd.
- Radford, Andrew. 1997. *Syntactic theory and the structure of English: A Minimalist approach* (Cambridge Textbooks in Linguistics). Cambridge, UK: Cambridge University Press.
- Rákosi, György, Tibor Laczkó & Gábor Csernyi. 2011. On English phrasal verbs and their Hungarian counterparts: From the perspective of a computational linguistic project. *Argumentum* 7. 80–89.
- Rambow, Owen. 1994. *Formal and computational aspects of natural language syntax*: University of Pennsylvania dissertation.
- Ramchand, Gillian. 2005. Post-Davidsonianism. *Theoretical Linguistics* 31(3). 359–373.
- Randriamasimanana, Charles. 2006. Simple sentences in Malagasy. In Henry Y. Chang, Lillian M. Huang & Dah ah Ho (eds.), *Streams converging into an ocean: Festschrift in honor of Professor Paul Jen-kuei Li on his 70th birthday*, 71–96. Taipei, Taiwan: Institute of Linguistics, Academia Sinica.
- Raposo, Eduardo & Juan Uriagereka. 1990. Long-distance case assignment. *Linguistic Inquiry* 21(4). 505–537.
- Rappaport, Malka. 1983. On the nature of derived nominals. In Levin et al. (1983) 113–42.
- Rauh, Gisa. 2013. Linguistic categories and the syntax-semantics interface: Evaluating competing approaches. Ms. Universität Wuppertal.
- Reape, Mike. 1991. Word order variation in Germanic and parsing. DYANA Report Deliverable R1.1.C University of Edinburgh.
- Reape, Mike. 1992. *A formal theory of word order: A case study in West Germanic*: University of Edinburgh dissertation.
- Reape, Mike. 1994. Domain union and word order variation in German. In Nerbonne et al. (1994b) 151–198.
- Reape, Mike. 2000. Formalisation and abstraction in linguistic theory II: Toward a radical Linearisation Theory of German. unpublished paper.
- Redington, Martin, Nick Chater & Steven Finch. 1998. Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science* 22(4). 425–469.

- Reis, Marga. 1974. Syntaktische Hauptsatzprivilegien und das Problem der deutschen Wortstellung. *Zeitschrift für Germanistische Linguistik* 2(3). 299–327.
- Reis, Marga. 1980. On justifying topological frames: ‘Positional field’ and the order of nonverbal constituents in German. *Documentation et Recherche en Linguistique Allemande Contemporaine* 22/23. 59–85.
- Reis, Marga. 1982. Zum Subjektbegriff im Deutschen. In Abraham (1982) 171–211.
- Remberger, Eva-Maria. 2009. Null subjects, expletives and locatives in Sardinian. In Georg A. Kaiser & Eva-Maria Remberger (eds.), *Proceedings of the workshop Null-Subjects, Expletives, and Locatives in Romance* (Arbeitspapier 123), 231–261. Konstanz: Fachbereich Sprachwissenschaft, Universität Konstanz.
- Resnik, Philip. 1992. Probabilistic Tree-Adjoining Grammar as a framework for statistical natural language processing. In Zampolli (1992) 418–424.
- Reyle, Uwe. 1993. Dealing with ambiguities by underspecification: Construction, representation and deduction. *Journal of Semantics* 10(2). 123–179.
- Richards, Marc. 2015. Minimalism. In Kiss & Alexiadou (2015) 803–839.
- Richter, Frank. 2004. *A mathematical formalism for linguistic theories with an application in Head-Driven Phrase Structure Grammar*. Eberhard-Karls-Universität Tübingen Phil. Dissertation (2000). <https://publikationen.uni-tuebingen.de/xmlui/handle/10900/46230>.
- Richter, Frank. 2007. Closer to the truth: A new model theory for HPSG. In Rogers & Kepser (2007) 101–110.
- Richter, Frank & Manfred Sailer. 1999a. A lexicalist collocation analysis of sentential negation in French. In Kordoni (1999b) 231–300.
- Richter, Frank & Manfred Sailer. 1999b. Lexicalizing the left periphery of German finite sentences. In Kordoni (1999b) 116–154.
- Richter, Frank & Manfred Sailer. 2004. Basic concepts of lexical resource semantics. In Arnold Beckmann & Norbert Preining (eds.), *ESSLLI 2003 – Course material I* (Collegium Logicum 5), 87–143. Wien: Kurt Gödel Society.
- Richter, Frank & Manfred Sailer. 2009. Phraseological clauses as Constructions in HPSG. In Müller (2009d) 297–317.
- Riehemann, Susanne. 1993. *Word formation in lexical type hierarchies: A case study of bar-adjectives in German*. Eberhard-Karls-Universität Tübingen MA thesis. Also published as SfS-Report-02-93, Seminar für Sprachwissenschaft, University of Tübingen.
- Riehemann, Susanne Z. 1998. Type-based derivational morphology. *Journal of Comparative Germanic Linguistics* 2(1). 49–77.
- Riemsdijk, Henk van. 1978. *A case study in syntactic markedness: The binding nature of prepositional phrases*. Lisse: The Peter de Ridder Press.
- Riezler, Stefan, Tracy Holloway King, Ronald M. Kaplan, Richard Crouch, John T. Maxwell III & Mark Johnson. 2002. Parsing the Wall Street Journal using a Lexical-Functional Grammar and discriminative estimation techniques. In Isabelle (2002) 271–278. <http://aclanthology.info/events/acl-2002>.

- Rizzi, Luigi. 1982. Violations of the *wh* island constraint and the Subjacency Condition. In Luigi Rizzi (ed.), *Issues in Italian syntax* (Studies in Generative Grammar 11), 49–76. Dordrecht: Foris Publications.
- Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17(3). 501–577.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer Academic Publishers.
- Rizzi, Luigi. 2009a. The discovery of language invariance and variation, and its relevance for the cognitive sciences. *The Behavioral and Brain Sciences* 32(5). 467–468.
- Rizzi, Luigi. 2009b. Language variation and universals: Some notes on N. Evans and S. C. Levinson (2009) “The myth of language universals: Language diversity and its importance for cognitive science”. In Paola Cotticelli-Kurras & Alessandra Tomaselli (eds.), *La Grammatica tra storia e teoria. Studi in onore di Giorgio Graffi*, 153–162. Alessandra: Edizioni dell’Orso.
- Roberts, Ian F. & Anders Holmberg. 2005. On the role of parameters in Universal Grammar: A reply to Newmeyer. In Hans Broekhuis, N. Corver, Riny Huybregts, Ursula Kleinhenz & Jan Koster (eds.), *Organizing grammar: Linguistic studies in honor of Henk van Riemsdijk* (Studies in Generative Grammar 86), 538–553. Berlin: Mouton de Gruyter.
- Robins, Robert Henry. 1997. *A short history of linguistics* (Longman Linguistics Library). London: Routledge 4th edn.
- Rogers, James. 1994. Obtaining trees from their descriptions: An application to Tree-Adjoining Grammars. *Computational Intelligence* 10(4). 401–421.
- Rogers, James. 1997. “Grammarless” Phrase Structure Grammar. *Linguistics and Philosophy* 20. 721–746.
- Rogers, James. 1998. *A descriptive approach to language-theoretic complexity* (Studies in Logic, Language and Information). Stanford, CA: CSLI Publications.
- Rogers, James & Stephan Kepser (eds.). 2007. *Model-theoretic syntax at 10 – Proceedings of the ESSLLI 2007 MTS@10 Workshop, August 13–17*. Dublin: Trinity College Dublin.
- Rohrer, Christian. 1996. Fakultativ kohärente Infinitkonstruktionen im Deutschen und deren Behandlung in der Lexikalisch Funktionalen Grammatik. In Gisela Harras & Manfred Bierwisch (eds.), *Wenn die Semantik arbeitet: Klaus Baumgärtner zum 65. Geburtstag*, 89–108. Tübingen: Max Niemeyer Verlag.
- Rohrer, Christian & Martin Forst. 2006. Improving coverage and parsing quality of a large-scale LFG for German. In *Proceedings of the Language Resources and Evaluation Conference (LREC-2006)*, Genoa, Italy.
- Ross, John Robert. 1967. *Constraints on variables in syntax*: MIT dissertation. <http://files.eric.ed.gov/fulltext/ED016965.pdf>. Reproduced by the Indiana University Linguistics Club and later published as Ross (1986).
- Ross, John Robert. 1986. *Infinite syntax!* Norwood, New Jersey: Ablex Publishing Corporation.

参考文献

- Rothkegel, Annely. 1976. *Valenzgrammatik* (Linguistische Arbeiten 19). Saarbrücken, Germany: Sonderforschungsbereich Elektronische Sprachforschung, Universität des Saarlandes.
- Sabel, Joachim. 1999. Das Passiv im Deutschen: Derivationale Ökonomie vs. optionale Bewegung. *Linguistische Berichte* 177. 87–112.
- Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–484.
- Sag, Ivan A. 2000. Another argument against *Wh*-trace. Jorge Hankamer Webfest. <http://ling.ucsc.edu/Jorge/sag.html>.
- Sag, Ivan A. 2007. Remarks on locality. In Müller (2007d) 394–414. <http://csli-publications.stanford.edu/HPSG/2007/>.
- Sag, Ivan A. 2010. English filler-gap constructions. *Language* 86(3). 486–545.
- Sag, Ivan A. 2012. Sign-Based Construction Grammar: An informal synopsis. In Boas & Sag (2012) 69–202.
- Sag, Ivan A., Hans C. Boas & Paul Kay. 2012. Introducing Sign-Based Construction Grammar. In Boas & Sag (2012) 1–29.
- Sag, Ivan A., Philip Hofmeister & Neal Snider. 2007. Processing complexity in subjacency violations: The Complex Noun Phrase Constraint. In Malcolm Elliott, James Kirby, Osamu Sawada, Eleni Staraki & Suwon Yoon (eds.), *Proceedings of the 43rd Annual Meeting of the Chicago Linguistic Society*, 215–229. Chicago: Chicago Linguistic Society.
- Sag, Ivan A. & Carl J. Pollard. 1991. An integrated theory of complement control. *Language* 67(1). 63–113.
- Sag, Ivan A. & Thomas Wasow. 2011. Performance-compatible competence grammar. In Borsley & Börjars (2011) 359–377.
- Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic theory: A formal introduction* (CSLI Lecture Notes 152). Stanford, CA: CSLI Publications 2nd edn.
- Sag, Ivan A. & Tom Wasow. 2015. Flexible processing and the design of grammar. *Journal of Psycholinguistic Research* 44(1). 47–63.
- Sailer, Manfred. 2000. *Combinatorial semantics and idiomatic expressions in Head-Driven Phrase Structure Grammar*: Eberhard-Karls-Universität Tübingen Dissertation. <https://publikationen.uni-tuebingen.de/xmlui/handle/10900/46191>.
- Samarin, William J. 1984. Socioprogrammed linguistics. *The Behavioral and Brain Sciences* 7(2). 206–207.
- Sampson, Geoffrey. 1989. Language acquisition: Growth or learning? *Philosophical Papers* 18(3). 203–240.
- Samvelian, Pollet. 2007. A (phrasal) affix analysis of the Persian Ezafe. *Journal of Linguistics* 43. 605–645.
- Sarkar, Anoop & Aravind K. Joshi. 1996. Coordination in Tree Adjoining Grammars: Formalization and implementation. In Tsuji (1996) 610–615.
- Sato, Yo. 2006. Constrained free word order parsing with Lexicalised Linearisation Grammar. In *Proceedings of 9th Annual CLUK Research Colloquium*, Open University, UK.
- Sato, Yo. 2008. *Implementing Head-Driven Linearisation Grammar*: King's College London dissertation.

- Sauerland, Uli & Paul Elbourne. 2002. Total reconstruction, PF movement, and derivational order. *Linguistic Inquiry* 33(2). 283–319.
- Sauerland, Uli & Hans-Martin Gärtner (eds.). 2007. *Interfaces + recursion = language? Chomsky's Minimalism and the view from syntax-semantics* (Studies in Generative Grammar 89). Berlin: Mouton de Gruyter.
- Savin, Harris B. & Ellen Perchonock. 1965. Grammatical structure and the immediate recall of English sentences. *Journal of Verbal Learning and Verbal Behavior* 4(5). 348–353.
- Schein, Barry. 1993. *Plurals and events* (Current Studies in Linguistics 23). Cambridge, MA: MIT Press.
- Scherpenisse, Wim. 1986. *The connection between base structure and linearization restrictions in German and Dutch* (Europäische Hochschulschriften, Reihe XXI, Linguistik 47). Frankfurt/M.: Peter Lang.
- Schluter, Natalie & Josef van Genabith. 2009. Dependency parsing resources for French: Converting acquired Lexical Functional Grammar f-structure annotations and parsing f-structures directly. In Kristiina Jokinen & Eckhard Bick (eds.), *Nodalida 2009 conference proceedings*, 166–173.
- Schmidt, Paul, Sibylle Rieder & Axel Theofilidis. 1996a. Final documentation of the German LS-GRAM lingware. Deliverable DC-WP6e (German) IAI Saarbrücken.
- Schmidt, Paul, Axel Theofilidis, Sibylle Rieder & Thierry Declerck. 1996b. Lean formalisms, linguistic theory, and applications: Grammar development in ALEP. In Tsuji (1996) 286–291. DOI:10.3115/992628.992679.
- Scholz, Barbara C. & Geoffrey K. Pullum. 2002. Searching for arguments to support linguistic nativism. *The Linguistic Review* 19(1–2). 185–223.
- Schubert, K. 1987. *Metataxis: Contrastive Dependency Syntax for machine translation*. Dordrecht: Foris Publications.
- Schumacher, Helmut, Jacqueline Kubczak, Renate Schmidt & Vera de Ruiter. 2004. *VALBU – Valenzwörterbuch deutscher Verben*. Tübingen: Gunter Narr Verlag.
- Schütz, Jörg. 1996. The ALEP formalism in a nutshell. Tech. rep. IAI Saarbrücken. <http://www.iai-sb.de/docs/alep-nutshell.pdf>.
- Schwarze, Christoph & Leonel de Alencar. 2016. *Lexikalisch-funktionale Grammatik: Eine Einführung am Beispiel des Französischen, mit computerlinguistischer Implementierung* (Stauffenburg Einführungen 30). Tübingen: Stauffenburg Verlag.
- Seiss, Melanie & Rachel Nordlinger. 2012. An electronic dictionary and translation system for Murrinh-Patha. *The EUROCALL Review: Proceedings of the EUROCALL 2011 Conference* 20(1). 135–138.
- Sengupta, Probal & B. B. Chaudhuri. 1997. A delayed syntactic-encoding-based LFG parsing strategy for an Indian language—Bangla. *Computational Linguistics* 23(2). 345–351.
- Seuren, Pieter A. M. 1984. The Bioprogram Hypothesis: Facts and fancy. *The Behavioral and Brain Sciences* 7(2). 208–209.
- Seuren, Pieter A. M. 2004. *Chomsky's Minimalism*. Oxford: Oxford University Press.
- Shieber, Stuart M. 1985. Evidence against the context-freeness of natural language. *Linguistics and Philosophy* 8(3). 333–343.

- Shieber, Stuart M. 1986. *An introduction to unification-based approaches to grammar* (CSLI Lecture Notes 4). Stanford, CA: CSLI Publications.
- Shieber, Stuart M. & Mark Johnson. 1993. Variations on incremental interpretation. *Journal of Psycholinguistic Research* 22(2). 287–318.
- Shieber, Stuart M., Hans Uszkoreit, Fernando Pereira, Jane Robinson & Mabry Tyson. 1983. The formalism and implementation of PATR-II. In *Research on interactive acquisition and use of knowledge*, 39–79. Menlo Park, CA: Artificial Intelligence Center, SRI International.
- Shtyrov, Y., E. Pihko & F. Pulvermüller. 2005. Determinants of dominance: Is language laterality explained by physical or linguistic features of speech? *Neuroimage* 27(1). 37–47.
- Sidner, Candy (ed.). 1987. *25th Annual Meeting of the Association for Computational Linguistics*. Stanford, CA: Association for Computational Linguistics.
- Siegel, Melanie. 2000. HPSG analysis of Japanese. In Wahlster (2000b) 264–279.
- Siegel, Melanie & Emily M. Bender. 2002. Efficient deep processing of Japanese. In *Proceedings of the 3rd Workshop on Asian Language Resources and International Standardization at the 19th International Conference on Computational Linguistics*. Taipei, Taiwan, <http://www.aclweb.org/anthology-new/W/W02/W02-1210.pdf>.
- Simov, Kiril, Petya Osenova, Alexander Simov & Milen Kouylekov. 2004. Design and implementation of the Bulgarian HPSG-based treebank. *Research on Language and Computation* 2(4). 495–522.
- Simpson, Jane. 1983. Resultatives. In Levin et al. (1983) 143–157. Reprint: Simpson (2005b).
- Simpson, Jane. 2005a. Depictives in English and Warlpiri. In Nikolaus P. Himmelmann & Eva Schultze-Berndt (eds.), *Secondary predication and adverbial modification: The typology of depictives*, 69–106. Oxford: Oxford University Press.
- Simpson, Jane. 2005b. Resultatives. In Miriam Butt & Tracy Holloway King (eds.), *Lexical semantics in LFG*, 149–161. Stanford, CA: CSLI Publications.
- Singleton, Jenny L. & Elissa L. Newport. 2004. When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive Psychology* 49(4). 370–407.
- Sleator, Daniel D. K. & Davy Temperley. 1991. Parsing English with a Link Grammar. CMU-CS-TR-91-126 School of Computer Science, Carnegie Mellon University.
- Smith, Carlota S. 1970. Jespersen’s “move and change” class and causative verbs in English. In Peeters (2000) 101–109.
- Smith, Carlota S. 1972. On causative verbs and derived nominals in English. *Linguistic Inquiry* 3(1). 136–138.
- Snyder, William. 2001. On the nature of syntactic variation: Evidence from complex predicates and complex word-formation. *Language* 77(2). 324–342.
- Soehn, Jan-Philipp & Manfred Sailer. 2008. At first blush on tenterhooks: About selectional restrictions imposed by nonheads. In Jäger et al. (2008) 149–161. <http://csli-publications.stanford.edu/FG/2003/soehn.pdf>.

- Somers, Harold & Mary McGee Wood (eds.). 1989. *Proceedings of the Fourth Conference of the European Chapter of the Association for Computational Linguistics*. Manchester, England: Association for Computational Linguistics.
- Son, Minjeong. 2007. Directionality and resultativity: The cross-linguistic correlation revisited. *Tromsø University Working Papers on Language & Linguistics* 34. 126–164. <http://hdl.handle.net/10037/3191>.
- Son, Minjeong & Peter Svenonius. 2008. Microparameters of cross-linguistic variation: Directed motion and resultatives. In Natasha Abner & Jason Bishop (eds.), *Proceedings of the 27th West Coast Conference on Formal Linguistics*, 388–396. Somerville, MA: Cascadilla Proceedings Project.
- Song, Sanghoun, Jong-Bok Kim, Francis Bond & Jaehyung Yang. 2010. Development of the Korean Resource Grammar: Towards grammar customization. In *Proceedings of the 8th Workshop on Asian Language Resources, Beijing, China, 21–22 August 2010*, 144–152. Asian Federation for Natural Language Processing.
- Sorace, Antonella. 2003. Near-nativeness. In Catherine J. Doughty & Michael H. Long (eds.), *The handbook of second language acquisition* (Blackwell Handbooks in Linguistics), 130–151. Oxford, UK/Cambridge, USA: Blackwell Publishing Ltd.
- Stabler, Edward P. 1987. Restricting logic grammars with Government-Binding Theory. *Computational Linguistics* 13(1–2). 1–10.
- Stabler, Edward P. 1991. Avoid the pedestrian’s paradox. In Robert C. Berwick, Steven P. Abney & Carol Tenny (eds.), *Principle-based parsing: Computation and psycholinguistics*, 199–237. Dordrecht: Kluwer Academic Publishers.
- Stabler, Edward P. 1992. *The logical approach to syntax: Foundations, specifications, and implementations of theories of Government and Binding* (ACL-MIT Press Series in Natural Language Processing). Cambridge, MA: MIT Press.
- Stabler, Edward P. 2001. Minimalist Grammars and recognition. In Christian Rohrer, Antje Rossdeutscher & Hans Kamp (eds.), *Linguistic form and its computation* (Studies in Computational Linguistics 1), 327–352. Stanford, CA: CSLI Publications.
- Stabler, Edward P. 2011a. After Government and Binding Theory. In van Benthem & ter Meulen (2011) 395–414.
- Stabler, Edward P. 2011b. Computational perspectives on Minimalism. In Cedric Boeckx (ed.), *The Oxford handbook of linguistic Minimalism* (Oxford Handbooks in Linguistics), chap. 27, 616–641. Oxford: Oxford University Press.
- Starosta, Stanley. 1988. *The case for Lexicase* (Open Linguistics Series). London: Pinter Publishers.
- Starosta, Stanley & Hirosato Nomura. 1986. Lexicase parsing: A lexicon-driven approach to syntactic analysis. In Nagao (1986) 127–132. DOI:10.3115/991365.991400.
- von Stechow, Arnim. 1979. Deutsche Wortstellung und Montague-Grammatik. In Jürgen M. Meisel & Martin D. Pam (eds.), *Linear order and Generative theory*, 317–490. Amsterdam: John Benjamins Publishing Co.
- von Stechow, Arnim. 1989. Distinguo: Eine Antwort auf Dieter Wunderlich. *Linguistische Berichte* 122. 330–339.

- von Stechow, Arnim. 1996. The different readings of *wieder* ‘again’: A structural account. *Journal of Semantics* 13(2). 87–138.
- von Stechow, Arnim & Wolfgang Sternefeld. 1988. *Bausteine syntaktischen Wissens: Ein Lehrbuch der Generativen Grammatik*. Opladen/Wiesbaden: Westdeutscher Verlag.
- Steedman, Mark J. 1985. Dependency and coordination in the grammar of Dutch and English. *Language* 61(3). 523–568.
- Steedman, Mark J. 1989. Constituency and coordination in a Combinatory Grammar. In Baltin & Kroch (1989) 201–231.
- Steedman, Mark J. 1991. Structure and intonation. *Language* 67(2). 260–296.
- Steedman, Mark J. 1997. *Surface structure and interpretation* (Linguistic Inquiry Monographs 30). Cambridge, MA/London, England: MIT Press.
- Steedman, Mark J. 2000. *The syntactic process* (Language, Speech, and Communication). Cambridge, MA/London, England: MIT Press.
- Steedman, Mark J. 2011. Romantics and revolutionaries. *Linguistic Issues in Language Technology* 6(11). 1–20. <http://journals.linguisticsociety.org/ellanguage/lilt/article/view/2587.html>. Special Issue on Interaction of Linguistics and Computational Linguistics.
- Steedman, Mark J. & Jason Baldridge. 2006. Combinatory Categorical Grammar. In Brown (2006) 610–621.
- Steels, Luc. 2003. Evolving grounded communication for robots. *Trends in Cognitive Science* 7(7). 308–312.
- Steels, Luc (ed.). 2011a. *Design patterns in Fluid Construction Grammar* (Constructional Approaches to Language 11). Amsterdam: John Benjamins Publishing Co.
- Steels, Luc (ed.). 2011b. *Design patterns in Fluid Construction Grammar* (Constructional Approaches to Language 11). Amsterdam: John Benjamins Publishing Co.
- Steels, Luc (ed.). 2012. *Computational issues in Fluid Construction Grammar* (Lecture Notes in Computer Science 7249). Berlin: Springer Verlag.
- Steels, Luc. 2013. Fluid Construction Grammar. In Hoffmann & Trousdale (2013) 153–167.
- Steels, Luc. 2015. *The Talking Heads experiment: Origins of words and meanings* (Computational Models of Language Evolution 1). Berlin: Language Science Press.
- Steels, Luc & Joachim De Beule. 2006. A (very) brief introduction to Fluid Construction Grammar. Paper presented at the Third International Workshop on Scalable Natural Language Understanding (ScaNaLU 2006) June 8, 2006, following HLT/NAACL, New York City.
- Steels, Luc & Remi van Trijp. 2011. How to make Construction Grammars fluid and robust. In Steels (2011b) 301–330.
- Stefanowitsch, Anatol. 2008. Negative entrenchment: A usage-based approach to negative evidence. *Cognitive Linguistics* 19(3). 513–531.
- Stefanowitsch, Anatol & Kerstin Fischer (eds.). 2008. *Konstruktionsgrammatik II: Von der Konstruktion zur Grammatik* (Stauffenburg Linguistik 47). Tübingen: Stauffenburg Verlag.
- Stefanowitsch, Anatol & Stephan Th. Gries. 2009. Corpora and grammar. In Lüdeling & Kytö (2009) chap. 43, 933–952.

- Sternefeld, Wolfgang. 1985a. Deutsch ohne grammatische Funktionen: Ein Beitrag zur Rektions- und Bindungstheorie. *Linguistische Berichte* 99. 394–439.
- Sternefeld, Wolfgang. 1985b. On case and binding theory. In Toman (1985) 231–285.
- Sternefeld, Wolfgang. 1995. Voice phrases and their specifiers. *FAS Papers in Linguistics* 3. 48–85.
- Sternefeld, Wolfgang. 2006. *Syntax: Eine morphologisch motivierte generative Beschreibung des Deutschen* (Stauffenburg Linguistik 31). Tübingen: Stauffenburg Verlag.
- Sternefeld, Wolfgang & Frank Richter. 2012. Wo stehen wir in der Grammatiktheorie? – Bemerkungen anlässlich eines Buchs von Stefan Müller. *Zeitschrift für Sprachwissenschaft* 31(2). 263–291.
- Stiebels, Barbara. 1996. *Lexikalische Argumente und Adjunkte: Zum semantischen Beitrag verbaler Präfixe und Partikeln* (studia grammatica 39). Berlin: Akademie Verlag.
- Stowell, Timothy. 1981. *Origins of phrase structure*: MIT dissertation. <http://hdl.handle.net/1721.1/15626>.
- Strunk, Jan & Nil Snider. 2013. Subclausal locality constraints on relative clause extraposition. In Gert Webelhuth, Manfred Sailer & Heike Walker (eds.), *Rightward movement in a comparative perspective* (Linguistik Aktuell/Linguistics Today 200), 99–143. Amsterdam: John Benjamins Publishing Co.
- Su, Keh-Yih, Oi Yee Kwong, Jn'ichi Tsujii & Jong-Hyeok Lee (eds.). 2005. *Natural language processing IJCNLP 2004* (Lecture Notes in Artificial Intelligence 3248) Berlin: Springer Verlag.
- Suchsland, Peter (ed.). 1992. *Biologische und soziale Grundlagen der Sprache* (Linguistische Arbeiten 280). Tübingen: Max Niemeyer Verlag.
- Suchsland, Peter. 1997. *Syntax-Theorie: Ein zusammengefaßter Zugang* (Konzepte der Sprach- und Literaturwissenschaft 55). Tübingen: Max Niemeyer Verlag. Deutsche Bearbeitung von Borsley (1991) durch Peter Suchsland.
- Sulger, Sebastian. 2009. Irish clefting and information-structure. In Butt & King (2009) 562–582. <http://csli-publications.stanford.edu/LFG/14/>.
- Sulger, Sebastian. 2010. Analytic and synthetic verb forms in Irish – An agreement-based implementation in LFG. In Manfred Pinkal, Ines Rehbein, Sabine Schulte im Walde & Angelika Storrer (eds.), *Semantic approaches in natural language processing: Proceedings of the Conference on Natural Language Processing 2010*, 169–173. Saarbrücken: Saarland University Press (universaar).
- Svenononius, Peter. 2004. Slavic prefixes inside and outside VP. *Nordlyd. Special Issue on Slavic Prefixes* 32(2). 205–253.
- Takami, Ken-ichi. 1988. Preposition stranding: Arguments against syntactic analyses and an alternative functional explanation. *Lingua* 76(4). 299–335.
- Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard & Julie C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268(5217). 1632–1634.
- Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard & Julie C. Sedivy. 1996. Using eye movements to study spoken language comprehension: Evidence for visually mediated incremental interpretation. In Toshio Inui & James L. McClell-

- land (eds.), *Information integration in perception and communication* (Attention and Performance XVI), 457–478. Cambridge, MA: MIT Press.
- ten Hacken, Pius. 2007. *Chomskyan linguistics and its competitors*. London: Equinox Publishing Ltd.
- Tesnière, Lucien. 1959. *Eléments de syntaxe structurale*. Paris: Librairie C. Klincksieck.
- Tesnière, Lucien. 1980. *Grundzüge der strukturalen Syntax*. Stuttgart: Klett-Cotta. Translated by Ulrich Engel.
- Tesnière, Lucien. 2015. *Elements of structural syntax*. Amsterdam: John Benjamins Publishing Co. Translated by Timothy Osborne and Sylvain Kahane.
- Thiersch, Craig L. 1978. *Topics in German syntax*: M.I.T. Dissertation.
- Thompson, Henry S. 1982. Handling metarules in a parser for GPSG. D.A.I. Research 175 University of Edinburgh.
- Timberlake, Alan. 1982. The impersonal passive in Lithuanian. In Monica Macaulay, Orin D. Gensler, Claudia Brugmann, Inese Čivkulis, Amy Dahlstrom, Katherine Krile & Rob Sturm (eds.), *Proceedings of the Eighth Annual Meeting of the Berkeley Linguistics Society*, 508–524. Berkeley: University of California.
- Toman, Jindřich (ed.). 1985. *Studies in German grammar* (Studies in Generative Grammar 21). Dordrecht: Foris Publications.
- Tomasello, Michael. 1995. Language is not an instinct. *Cognitive Development* 10(1). 131–156.
- Tomasello, Michael. 2000. Do young children have adult syntactic competence? *Cognition* 74(3). 209–253.
- Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Tomasello, Michael. 2005. Beyond formalities: The case of language acquisition. *The Linguistic Review* 22(2–4). 183–197.
- Tomasello, Michael. 2006a. Acquiring linguistic constructions. In Deanna Kuhn & Robert Siegler (eds.), *Handbook of child psychology*, vol. 2, New York: John Wiley & Sons, Inc. 6th edn.
- Tomasello, Michael. 2006b. Construction Grammar for kids. *Constructions Special Volume 1*. <http://www.constructions-journal.com/>.
- Tomasello, Michael. 2006c. Konstruktionsgrammatik und früher Erstspracherwerb. In Kerstin Fischer & Anatol Stefanowitsch (eds.), *Konstruktionsgrammatik: Von der Anwendung zur Theorie* (Stauffenburg Linguistik 40), 19–37. Tübingen: Stauffenburg Verlag.
- Tomasello, Michael. 2009. Universal Grammar is dead. *The Behavioral and Brain Sciences* 32(5). 470–471.
- Tomasello, Michael, Nameera Akhtar, Kelly Dodson & Laura Rekau. 1997. Differential productivity in young children’s use of nouns and verbs. *Journal of Child Language* 24(2). 373–387.
- Tomasello, Michael, Malinda Carpenter, Josep Call, Tanya Behne & Henrike Moll. 2005. Understanding and sharing intentions: The origins of cultural cognition. *The Behavioral and Brain Sciences* 28(5). 675–735.

- Travis, Lisa. 1984. *Parameters and effects of word order variation*. Cambridge, MA: M.I.T. Dissertation.
- Trosterud, Trond. 2009. A Constraint Grammar for Faroese. In Bick et al. (2009) 1–7.
- Tseng, Jesse (ed.). 2000. *Aspekte eines HPSG-Fragments des Deutschen* (Arbeitspapiere des SFB 340 No. 156). Tübingen: Eberhard-Karls-Universität Tübingen. <http://www.sfs.uni-tuebingen.de/sfb/reports/berichte/156/156abs.html>.
- Tseng, Jesse L. 2003. LKB grammar implementation: French and beyond. In Bender et al. (2003) 91–97. <http://w3.erss.univ-tlse2.fr/textes/pagespersos/tseng/Pubs/mgd03.pdf>.
- Tseng, Jesse L. 2007. English prepositional passive constructions. In Müller (2007d) 271–286. <http://csli-publications.stanford.edu/HPSG/2007/>.
- Tsuji, Jun-ichi (ed.). 1996. *Proceedings of COLING-96. 16th International Conference on Computational Linguistics (COLING96). Copenhagen, Denmark, August 5–9, 1996*. Copenhagen, Denmark: Association for Computational Linguistics.
- Umemoto, Hiroshi. 2006. Implementing a Japanese semantic parser based on glue approach. In *Proceedings of The 20th Pacific Asia Conference on Language, Information and Computation*, 418–425. <http://dspace.wul.waseda.ac.jp/dspace/handle/2065/29076>.
- Uszkoreit, Hans. 1986a. Categorical Unification Grammars. In Nagao (1986) 187–194. <http://aclweb.org/anthology-new/C/C86/C86-1045.pdf>.
- Uszkoreit, Hans. 1986b. Linear precedence in discontinuous constituents: Complex fronting in German. Report No. CSLI-86-47 Center for the Study of Language and Information Stanford, CA.
- Uszkoreit, Hans. 1987. *Word order and constituent structure in German* (CSLI Lecture Notes 8). Stanford, CA: CSLI Publications.
- Uszkoreit, Hans. 1990. Extraposition and adjunct attachment in Categorical Unification Grammar. In Werner Bahner, Joachim Schildt & Dieter Viehweger (eds.), *Proceedings of the Fourteenth International Congress of Linguists, Berlin/GDR, August 10–15, 1987*, vol. 3, 2331–2336. Berlin: Akademie Verlag.
- Uszkoreit, Hans (ed.). 1995. *33rd Annual Meeting of the Association for Computational Linguistics. Proceedings of the conference*. Cambridge, MA: Association for Computational Linguistics.
- Uszkoreit, Hans, Rolf Backofen, Stephan Busemann, Abdel Kader Diagne, Elizabeth A. Hinkelman, Walter Kasper, Bernd Kiefer, Hans-Ulrich Krieger, Klaus Netter, Günter Neumann, Stephan Oepen & Stephen P. Spackman. 1994. DISCO—An HPSG-based NLP system and its application for appointment scheduling. In Nagao (1994) 436–440.
- Uszkoreit, Hans, Rolf Backofen, Jo Calder, Joanne Capstick, Luca Dini, Jochen Dörre, Gregor Erbach, Dominique Estival, Suresh Manandhar, Anne-Marie Mineur & Stephan Oepen. 1996. The EAGLES formalisms working group: Final report Expert Advisory Group on Language Engineering Standards. Technical Report LRE 61–100. <http://www.coli.uni-sb.de/publikationen/softcopies/Uszkoreit:1996:EFW.pdf>.
- Valian, Virginia. 1991. Syntactic subjects in the early speech of American and Italian children. *Cognition* 40(1–2). 21–81.
- van Benthem, Johan F. A. K. & G. B. Alice ter Meulen (eds.). 2011. *Handbook of logic and language*. Amsterdam: Elsevier 2nd edn.

- van der Hulst, Harry (ed.). 2010. *Recursion in human language* (Studies in Generative Grammar 104). Berlin: Mouton de Gruyter.
- Van Eynde, Frank. 2015. Sign-Based Construction Grammar: A guided tour. *Journal of Linguistics* DOI:10.1017/S0022226715000341.
- Van Langendonck, Willy. 1994. Determiners as heads? *Cognitive Linguistics* 5. 243–259.
- van Trijp, Remi. 2011. A design pattern for argument structure constructions. In Steels (2011b) 115–145.
- van Trijp, Remi. 2013. A comparison between Fluid Construction Grammar and Sign-Based Construction Grammar. *Constructions and Frames* 5(1). 88–116.
- van Trijp, Remi. 2014. Long-distance dependencies without filler–gaps: A cognitive-functional alternative in Fluid Construction Grammar. *Language and Cognition* 6(2). 242–270.
- Van Valin, Robert D. Jr. (ed.). 1993. *Advances in Role and Reference Grammar*. Amsterdam: John Benjamins Publishing Co.
- Van Valin, Robert D. Jr. 1998. The acquisition of *wh*-questions and the mechanisms of language acquisition. In Michael Tomasello (ed.), *The new psychology of language: Cognitive and functional approaches to language structure*, 221–249. Hillsdale, NJ: Lawrence Erlbaum.
- Vancoppenolle, Jean, Eric Tabbert, Gerlof Bouma & Manfred Stede. 2011. A German grammar for generation in Open CCG. In Hedeland et al. (2011) 145–150.
- Vargha, Dénes (ed.). 1988. *Proceedings of COLING 88*. University of Budapest: Association for Computational Linguistics.
- Vargha-Khadem, Faraneh, Kate E. Watkins, Katie Alcock, Paul Fletcher & Richard Passingham. 1995. Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. In *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, 930–933.
- Vasishth, Shravan & Richard L. Lewis. 2006. Human language processing: Symbolic models. In Brown (2006) 410–419.
- Vasishth, Shravan, Katja Suckow, Richard L. Lewis & Sabine Kern. 2010. Short-term forgetting in sentence comprehension: Crosslinguistic evidence from verb-final structures. *Language and Cognitive Processes* 25(4). 533–567.
- Vater, Heinz. 2010. Strukturalismus und generative Grammatik in Deutschland. In Hans-Harald Müller, Marcel Lepper & Andreas Gardt (eds.), *Strukturalismus in Deutschland: Literatur- und Sprachwissenschaft 1910–1975* (Marbacher Schriften. Neue Folge 5), 125–160. Göttingen: Wallstein Verlag.
- Veenstra, Mettina Jolanda Arnoldina. 1998. *Formalizing the Minimalist Program*: Rijksuniversiteit Groningen Ph.d. thesis.
- Vennemann, Theo & Ray Harlow. 1977. Categorical Grammar and consistent basic VX serialization. *Theoretical Linguistics* 4(1–3). 227–254.
- Verhagen, Arie. 2010. What do you think is the proper place of recursion? Conceptual and empirical issues. In van der Hulst (2010) 93–110.
- Verspoor, Cornelia Maria. 1997. *Contextually-dependent lexical semantics*: University of Edinburgh dissertation. <ftp://ftp.cogsci.ed.ac.uk/pub/kversp/thesis.ps.gz>.

- Vierhuff, Tilman, Bernd Hildebrandt & Hans-Jürgen Eikmeyer. 2003. Effiziente Verarbeitung deutscher Konstituentenstellung mit der Combinatorial Categorical Grammar. *Linguistische Berichte* 194. 213–237.
- Vijay-Shanker, K. & Aravind K. Joshi. 1988. Feature structures based Tree Adjoining Grammars. In Vargha (1988) 714–719. <http://www.aclweb.org/anthology-new/C/C88/C88-2147.pdf>.
- Villavicencio, Aline. 2002. The acquisition of a unification-based Generalised Categorical Grammar. UCAM-CL-TR-533 University of Cambridge Computer Laboratory.
- Vogel, Ralf. 2001. Case conflict in German free relative constructions: An Optimality Theoretic treatment. In Gereon Müller & Wolfgang Sternefeld (eds.), *Competition in syntax*, 341–375. Berlin: Mouton de Gruyter.
- Vogel, Ralf & Markus Steinbach. 1998. The dative – An oblique case. *Linguistische Berichte* 173. 65–91.
- Volk, Martin. 1988. *Parsing German with GPSG: The problem of separable-prefix verbs*. University of Georgia MA thesis.
- Voutilainen, Atro, Juha Heikkilä & Arto Anttila. 1992. *Constraint Grammar of English: A performance-oriented introduction* (Publications of the Department of General Linguistics 21). Helsinki: University of Helsinki.
- Wada, Hajime & Nicholas Asher. 1986. BUILDERS: An implementation of DR Theory and LFG. In Nagao (1986) 540–545.
- Wahlster, Wolfgang (ed.). 2000a. *Verbmobil: Foundations of speech-to-speech translation* (Artificial Intelligence). Berlin: Springer Verlag.
- Wahlster, Wolfgang (ed.). 2000b. *Verbmobil: Foundations of speech-to-speech translation* (Artificial Intelligence). Berlin: Springer Verlag.
- Walther, Markus. 1999. *Deklarative prosodische Morphologie: Constraint-basierte Analysen und Computermodelle zum Finnischen und Tigrinya* (Linguistische Arbeiten 399). Tübingen: Max Niemeyer Verlag.
- Webelhuth, Gert. 1985. German is configurational. *The Linguistic Review* 4(3). 203–246.
- Webelhuth, Gert. 1990. Diagnostics for structure. In Grewendorf & Sternefeld (1990) 41–75.
- Webelhuth, Gert. 1995. X-bar Theory and Case Theory. In Gert Webelhuth (ed.), *Government and Binding Theory and the Minimalist Program: Principles and Parameters in syntactic theory* (Generative Syntax), 15–95. Oxford, UK & Cambridge, USA: Blackwell Publishing Ltd.
- Webelhuth, Gert. 2011. Paradigmenwechsel rückwärts: Die Renaissance der grammatischen Konstruktion. In Engelberg et al. (2011) 149–180.
- Webelhuth, Gert, Jean-Pierre Koenig & Andreas Kathol (eds.). 1999. *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1). Stanford, CA: CSLI Publications.
- Weber, Heinz J. 1997. *Dependenzgrammatik: Ein interaktives Arbeitsbuch* (Narr Studienbücher). Tübingen: Gunter Narr Verlag 2nd edn.
- Wechsler, Stephen Mark. 1991. *Argument structure and linking*: Stanford University dissertation.

- Wechsler, Stephen Mark. 1995. *The semantic basis of argument structure* (Dissertations in Linguistics). Stanford, CA: CSLI Publications.
- Wechsler, Stephen Mark. 1997. Resultative predicates and control. In Ralph C. Blight & Michelle J. Moosally (eds.), *Texas Linguistic Forum 38: The syntax and semantics of predication: Proceedings of the 1997 Texas Linguistics Society Conference*, 307–321. Austin, Texas: University of Texas, Department of Linguistics.
- Wechsler, Stephen Mark. 2005. What is right and wrong about little v. In *Grammar and beyond—Essays in honour of Lars Hellan*, 179–195. Oslo, Norway: Novus Press.
- Wechsler, Stephen Mark. 2008a. A diachronic account of English deverbal nominals. In Charles B. Chang & Hannah J. Haynie (eds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics*, 498–506. Somerville, MA: Cascadilla Proceedings Project.
- Wechsler, Stephen Mark. 2008b. Dualist syntax. In Müller (2008b) 294–304. <http://csli-publications.stanford.edu/HPSG/2008/>.
- Wechsler, Stephen Mark & Bokyung Noh. 2001. On resultative predicates and clauses: Parallels between Korean and English. *Language Sciences* 23(4). 391–423.
- Wegener, Heide. 1985. *Der Dativ im heutigen Deutsch* (Studien zur deutschen Grammatik 28). Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.
- Weir, Morton W. 1964. Developmental changes in problem-solving strategies. *Psychological Review* 71(6). 473–490.
- Weissgerber, Monika. 1983. *Valenz und Kongruenzbeziehungen: Ein Modell zur Vereindeutigung von Verben in der maschinellen Analyse und Übersetzung*. Frankfurt a. M.: Peter Lang.
- Weisweber, Wilhelm. 1987. Ein Dominanz-Chart-Parser für generalisierte Phrasenstrukturgrammatiken. KIT-Report 45 Technische Universität Berlin Berlin.
- Weisweber, Wilhelm & Susanne Preuss. 1992. Direct parsing with metarules. In Zampolli (1992) 1111–1115.
- Welke, Klaus. 1988. *Einführung in die Valenz- und Kasus-theorie*. Leipzig: Bibliographisches Institut.
- Welke, Klaus. 2009. Konstruktionsvererbung, Valenzvererbung und die Reichweite von Konstruktionen. *Zeitschrift für Germanistische Linguistik* 37(3). 514–543.
- Welke, Klaus. 2011. *Valenzgrammatik des Deutschen: Eine Einführung* (De Gruyter Studium). Berlin: de Gruyter.
- Wells, Rulon S. 1947. Immediate constituents. *Language* 23(2). 81–117.
- Werner, Edeltraud. 1993. *Translationstheorie und Dependenzmodell: Kritik und Reinterpretation des Ansatzes von Lucien Tesnière* (Kultur und Erkenntnis: Schriften der Philosophischen Fakultät der Heinrich-Heine-Universität Düsseldorf 10). Tübingen: Francke Verlag.
- Wetta, Andrew C. 2011. A Construction-based cross-linguistic analysis of V2 word order. In Müller (2011b) 248–268. <http://csli-publications.stanford.edu/HPSG/2011/>.
- Wexler, Kenneth. 1998. Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua* 106(1–4). 23–79.

- Wexler, Kenneth & Peter W. Culicover. 1980. *Formal principles of language acquisition*. Cambridge, MA/London: MIT Press.
- Weydt, Harald. 1972. „Unendlicher Gebrauch von endlichen Mitteln“: Mißverständnisse um ein linguistisches Theorem. *Poetica* 5(3/4). 249–267.
- Wharton, R. M. 1974. Approximate language identification. *Information and Control* 26(3). 236–255.
- White, Mike & Jason Baldridge. 2003. Adapting chart realization to CCG. In Ehud Reiter, Helmut Horacek & Kees van Deemter (eds.), *Proceedings of the 9th European Workshop on Natural Language Generation (ENLG-2003) at EACL 2003*, 119–126.
- Wijnen, Frank, Masja Kempen & Steven Gillis. 2001. Root infinitives in Dutch early child language: An effect of input? *Journal of Child Language* 28(3). 629–660.
- Wiklund, Anna-Lena, Gunnar Hrafn Hrafnbjargarson, Kristine Bentzen & Þorbjörg Hróarsdóttir. 2007. Rethinking Scandinavian verb movement. *Journal of Comparative Germanic Linguistics* 10(3). 203–233.
- Wilcock, Graham. 2001. Towards a discourse-oriented representation of information structure in HPSG. In *13th Nordic Conference on Computational Linguistics, Uppsala, Sweden*, <http://www.ling.helsinki.fi/~gwilcock/Pubs/2001/Nodalida-01.pdf>.
- Wilder, Chris. 1991. Small clauses and related objects. *Groninger Arbeiten zur Germanistischen Linguistik* 34. 215–236.
- Wilks, Yorick (ed.). 1984. *Proceedings of the 10th International Conference on Computational Linguistics and 22nd Annual Meeting of the Association for Computational Linguistics*. Stanford University, California: Association for Computational Linguistics.
- Williams, Edwin. 1984. Grammatical relations. *Linguistic Inquiry* 15(4). 639–673.
- Winkler, Susanne. 1997. *Focus and secondary predication* (Studies in Generative Grammar 43). Berlin, New York: Mouton de Gruyter.
- Wittenberg, Eva, Ray S. Jackendoff, Gina Kuperberg, Martin Paczynski, Jesse Snedeker & Heike Wiese. 2014. The processing and representation of light verb constructions. In Asaf Bachrach, Isabelle Roy & Linnaea Stockall (eds.), *Structuring the argument* (Language Faculty and Beyond 10), 61–80. Amsterdam: John Benjamins Publishing Co.
- Wittenberg, Eva & Maria Mercedes Piñango. 2011. Processing light verb constructions. *The Mental Lexicon* 6(3). 393–413.
- Wöllstein, Angelika. 2010. *Topologisches Satzmodell* (Kurze Einführungen in die Germanistische Linguistik 8). Heidelberg: Universitätsverlag Winter.
- Wunderlich, Dieter. 1987. Vermeide Pronomen – Vermeide leere Kategorien. *Studium Linguistik* 21. 36–44.
- Wunderlich, Dieter. 1989. Arnim von Stechow, das Nichts und die Lexikalisten. *Linguistische Berichte* 122. 321–333.
- Wunderlich, Dieter. 1992. CAUSE and the structure of verbs. Arbeiten des SFB 282 No. 36 Heinrich Heine Uni/BUGH Düsseldorf/Wuppertal.
- Wunderlich, Dieter. 1997. Argument extension by lexical adjunction. *Journal of Semantics* 14(2). 95–142.
- Wunderlich, Dieter. 2004. Why assume UG? *Studies in Language* 28(3). 615–641.

- Wunderlich, Dieter. 2008. Spekulationen zum Anfang von Sprache. *Zeitschrift für Sprachwissenschaft* 27(2). 229–265.
- Wurmbrand, Susanne. 2003a. *Infinitives: Restructuring and clause structure* (Studies in Generative Grammar 55). Berlin: Mouton de Gruyter.
- Wurmbrand, Susanne. 2003b. Long passive (corpus search results).
- XTAG Research Group. 2001. A lexicalized Tree Adjoining Grammar for English. Tech. rep. Institute for Research in Cognitive Science Philadelphia. <ftp://ftp.cis.upenn.edu/pub/xtag/release-2.24.2001/tech-report.pdf>.
- Yamada, Hiroyasu & Yuji Matsumoto. 2003. Statistical dependency analysis with support vector machines. In van Noord (2003).
- Yamada, Jeni. 1981. Evidence for the independence of language and cognition: Case study of a “hyperlinguistic” adolescent. UCLA Working Papers in Cognitive Linguistics 3 University of California, Los Angeles.
- Yampol, Todd & Lauri Karttunen. 1990. An efficient implementation of PATR for Categorical Unification Grammar. In Karlgren (1990) 419–424.
- Yang, Charles D. 2004. Universal Grammar, statistics or both? *Trends in Cognitive Sciences* 8(10). 451–456. DOI:10.1016/j.tics.2004.08.006.
- Yasukawa, Hideki. 1984. LFG System in Prolog. In Wilks (1984) 358–361.
- Yip, Moira, Joan Maling & Ray S. Jackendoff. 1987. Case in tiers. *Language* 63(2). 217–250.
- Yoshinaga, Naoki, Yusuke Miyao, Kentaro Torisawa & Jun’ichi Tsujii. 2001. Resource sharing amongst HPSG and LTAG communities by a method of grammar conversion between FB-LTAG and HPSG. In *Proceedings of ACL/EACL workshop on Sharing Tools and Resources for Research and Education*, 39–46. Toulouse, France.
- Zaenen, Annie & Ronald M. Kaplan. 1995. Formal devices for linguistic generalizations: West Germanic word order in LFG. In Dalrymple et al. (1995) 215–239.
- Zaenen, Annie, Joan Maling & Höskuldur Thráinsson. 1985. Case and grammatical functions: The Icelandic passive. *Natural Language and Linguistic Theory* 3(4). 441–483.
- Zampolli, Antonio (ed.). 1992. *14th International Conference on Computational Linguistics (COLING ’92)*, August 23–28. Nantes, France: Association for Computational Linguistics.
- Zappa, Frank. 1986. Does humor belong in music? EMI Music Germany GmbH & Co.KG.
- Zucchi, Alessandro. 1993. *The language of propositions and events: Issues in the syntax and the semantics of nominalization* (Studies in Linguistics and Philosophy 51). Berlin: Springer Verlag.
- Zwart, C. Jan-Wouter. 1994. Dutch is head-initial. *The Linguistic Review* 11(3–4). 377–406.
- Zweigenbaum, Pierre. 1991. Un analyseur pour grammaires lexicales-fonctionnelles. *TA Informations* 32(2). 19–34.
- Zwicky, Arnold M., Joyce Friedman, Barbara C. Hall & Donald E. Walker. 1965. The MITRE syntactic analysis procedure for Transformational Grammars. In *Proceedings – FALL Joint Computer Conference*, 317–326. DOI:10.1109/AFIPS.1965.108.

Name index

- Abbott, Barbara, 558
 Abeillé, Anne, 110, 146, 147, 378, 513, 514, 596
 Abney, Steven P., 26, 107–109, 438, 456, 471, 493, 639
 Abraham, Werner, 11, 139
 Abzianidze, Lasha, 238
 Ackerman, Farrell, 107, 490, 624, 638
 Adams, Marianne, 422
 Ades, Anthony E., 477
 Adger, David, xvi, 115, 119–121, 123–129, 134, 150, 155, 156, 159, 510, 539, 541, 543
 Ágel, Vilmos, 335
 Aguado-Orea, Javier, 448, 494
 Ahmed, Mohamed Ben, 377
 Ahmed, Reaz, 237
 Ajdukiewicz, Kasimir, 144, 147, 219, 335
 Akhtar, Nameera, 496
 Alcock, Katie, 433, 434
 Alsina, Alex, 35, 282, 411, 549, 554, 562
 Altmann, Rebecca J., 476
 Ambridge, Ben, 420–422, 441, 446, 453, 500
 Anderson, John M., 335
 Anderson, Stephen R., 206
 Andrews, Avery, 196
 Anttila, Arto, 336
 Aoun, Joseph, 109
 Arad Greshler, Tali, 237, 238
 Aras, Hidir, 309
 Arends, Jacques, 431
 Arka, I Wayan, 196
 Arnold, Doug, 519
 Arnold, Jennifer E., 476
 Asher, Nicholas, 195
 Askedal, John Ole, 40
 Asudeh, Ash, 205, 217, 277, 282, 549, 562–566
 Attardi, Giuseppe, 336
 Attia, Mohammed A., 195
 Avgustinova, Tania, 238
 Bach, Emmon, 58, 93
 Backofen, Rolf, 237
 Bahrani, Mohammad, 161
 Baker, Carl Lee, 434, 439
 Baker, Mark C., 407, 411, 415, 484
 Baldridge, Jason, xvi, 147, 219, 220, 225, 228, 232, 236, 267, 528
 Balla, Amar, 237
 Ballweg, Joachim, 502
 Baltin, Mark, 414, 415, 423
 Bangalore, Srinivas, 376
 Bannard, Colin, 283, 518
 Bar-Hillel, Yehoshua, 519, 521
 Barry, Guy, 236
 Bartsch, Renate, 94, 145
 Barwise, Jon, 251
 Barz, Irmhild, 12, 21–23, 39, 45, 66
 Baschung, K., 219
 Bates, Elizabeth A., 429, 430, 432, 458
 Baumgärtner, Klaus, 335, 339, 367
 Bausewein, Karin, 142, 143, 257
 Bayer, Josef, 91, 107
 Beavers, John, 146, 147, 219, 275, 310, 598, 601
 Bech, Gunnar, 43, 383, 396
 Becker, Tilman, 88, 383, 385, 394–397
 Beermann, Dorothee, 238
 Beghelli, Filippo, 133
 Behaghel, Otto, 509
 Behne, Tanya, 431, 437
 Behrens, Heike, 334, 429, 498, 598
 Bellugi, Ursula, 433

人名索引

- Bender, Emily, 301, 512
Bender, Emily M., xvi, 6, 25, 108, 238, 239, 274, 278, 297, 298, 478, 512, 515, 520
Bentzen, Kristine, 133
Bergen, Benjamin K., 88, 282, 291, 292, 309, 311, 312, 496, 584
Berman, Judith, 35, 91, 93, 195, 196, 208–213, 216, 218, 394, 411, 412, 665, 688
Berthoud, Ioanna, 433
Berwick, Robert C., 144, 155, 158, 441, 485, 486, 501
Bes, G. G., 219
Bever, Thomas G., 473
Bialystok, Ellen, 430
Bick, Eckhard, 336, 337
Bickerton, Derek, 431
Bierwisch, Manfred, 6, 77, 93, 106–108, 426
Bildhauer, Felix, xvi, 16, 134, 239, 240, 279, 292, 319, 363
Bird, Steven, 240, 307
Bishop, Dorothy V. M., 432–434
Bjerre, Tavs, 604
Blackburn, Patrick, 75, 461, 528
Błaszczak, Joanna, 133, 159
Blevins, James P., 289
Block, Hans-Ulrich, 195
Blom, Corrien, 591, 603
Bloom, Paul, 485, 487
Blutner, Reinhard, 527
Boas, Hans C., 297, 325, 333, 550
Bobaljik, Jonathan, 134
Bod, Rens, 441, 448, 606
Bögel, Tina, 196
Bohnet, Bernd, 336
Bolc, Leonard, 237
Bond, Francis, 237, 238
Booij, Geert E., 541, 564, 591, 603, 605
Borer, Hagit, 140, 549, 556, 568, 574, 577, 582, 583
Börschinger, Benjamin, 195
Borsley, Robert D., 111, 113, 145, 146, 154, 159, 263, 308, 334, 370, 618, 739
Bos, Johan, 75, 528
Bouchard, Lorne H., 161
Boukedi, Sirine, 237
Boullier, Pierre, 195
Bouma, Gerlof, 219
Bouma, Gosse, 153, 154, 219, 238, 255, 272, 291, 304, 321, 349, 354, 461, 519, 520, 534, 544, 588
Bragmann, Sascha, 374, 621
Braine, Martin D. S., 423, 455
Branco, António, 238
Brants, Sabine, 417
Bredt, Thomas H., 109
Bresnan, Joan, 91, 113, 142, 146, 154, 195, 199, 202, 206–209, 211, 218, 289, 319, 409, 411, 501, 512, 516, 520, 533, 558, 560, 562, 563, 571
Brew, Chris, 239, 330
Briscoe, Ted, 554
Briscoe, Ted J., 111, 161, 220, 227, 236, 331, 423, 535, 553, 567, 638
Bröker, Norbert, 342
Brosziewski, Ulf, 119
Brown, Roger, 456
Bruening, Benjamin, 27, 156
Bryant, John, 283
Büchel, Christian, 432, 633
Budde, Monika, xvi, 613
Bühler, Karl, 424
Bungeroth, Jan, 239
Burke, Michael, 197
Burzio, Luigi, 86, 103
Buscha, Joachim, 335
Busemann, Stephan, 180, 237
Butt, Miriam, 195, 196, 601

Cahill, Aoife, 197
Calder, Jo, 237, 638
Calder, Jonathan, 219
Call, Josep, 431, 437
Callmeier, Ulrich, 237
Candito, Marie-Hélène, 286, 378, 380, 388, 584
Canon, Stephen, 197
Cappelle, Bert, 291, 316, 602–605
Capstick, Joanne, 237
Cargo, Martha B., 433
Carlson, Gregory N., 599
Carpenter, Bob, 181, 194, 219, 237, 461
Carpenter, Malinda, 431, 437
Carroll, John, 161

- Casenhiser, Devin, 552
 Cavar, Damir, 471
 Çetinoğlu, Özlem, 196
 Chang, Nancy, 88, 282, 283, 291, 292, 309, 311, 312, 496, 584
 Chang, Nancy Chih-Lin, 312
 Chater, Nick, 455
 Chatterjee, Sudipta, 219, 220
 Chaudhuri, B. B., 195
 Chaves, Rui P., 305, 354, 520
 Chi, Zhiyi, 197
 Choi, Hye-Won, 209, 210
 Chomsky, Noam, 4, 5, 70, 77–80, 83, 85, 86, 89, 101, 108, 109, 111, 115, 117, 121, 130–132, 139–146, 149, 153–155, 157–161, 281, 289, 326, 394, 405, 407, 408, 410, 411, 413–415, 423–429, 431, 432, 434, 437, 438, 441, 454, 456, 462, 469, 472, 473, 475, 478, 479, 483, 484, 489, 491–493, 501, 506, 520, 560, 566, 572, 573, 637, 674
 Chouinard, Michelle M., 456
 Christiansen, Morten, 422
 Chrupala, Grzegorz, 197
 Chung, Sandra, 422
 Church, Kenneth, 639
 Cinque, Guglielmo, 131–134, 137, 157, 423, 634
 Citko, Barbara, 142, 144
 Clark, Alexander, 455
 Clark, Eve V., 456
 Clark, Herbert H., 475
 Clark, Stephen, 219, 220
 Clément, Lionel, 195, 286, 584
 Clifton, Charles, 479
 Clifton, Charles Jr., 471
 Coch, Jose, 336, 337
 Cole, Jennifer, 109, 471
 Comrie, Bernard, 143, 147, 257
 Cook, Philippa Helen, xvi, 16, 36, 195, 319, 363, 421
 Coopmans, Peter, 519
 Copestake, Ann, 237, 238, 244, 253, 285, 297, 327, 331, 332, 382, 528, 535, 553, 554, 567
 Corluy, A., 219
 Correa, Nelson, 109
 Costa, Francisco, 238
 Covington, Michael A., 336, 337, 366
 Crabbé, Benoît, 378
 Crain, Stephen, 413, 422, 446, 453, 477
 Cramer, Bart, 239
 Crocker, Matthew Walter, 109, 111
 Croft, William, 282, 411, 423, 496, 506, 549, 550, 555, 583, 586, 587, 600, 632
 Crouch, Richard, 195, 197
 Crysmann, Berthold, 16, 95, 146, 147, 238, 240, 275, 310, 598, 603
 Csernyi, Gábor, 196
 Culicover, Peter W., 39, 77, 145, 150, 412, 428, 474, 492, 521, 549, 588, 589, 601
 Culy, Christopher, 179, 377, 501
 Curran, James, 219
 Curtiss, Susan, 430
 Czuba, Krzysztof, 237
 Da Sylva, Lyne, 161
 Dąbrowska, Ewa, 281, 334, 432–434, 445, 495, 549
 Dahl, Østen, 422
 Dahllöf, Mats, 237, 238
 Dale, Robert, 195
 Dalrymple, Mary, 35, 91, 196, 198, 203, 205, 209, 217, 218, 277, 282, 342, 411, 414, 506, 526, 549, 562, 564, 566
 Davidson, Donald, 568
 Davies, Mark, 433
 Davis, Anthony R., 253, 286, 584
 de Alencar, Leonel, 195, 196, 218
 De Beule, Joachim, 282
 De Kuthy, Kordula, 14, 16, 108, 154, 238, 292, 319, 421, 502, 503
 de Saussure, Ferdinand, 3, 425
 Declerck, Thierry, 237
 Dellert, Johannes, 377
 Delmonte, Rodolfo, 195, 196
 Demberg, Vera, 326
 Demske, Ulrike, 27
 den Besten, Hans, 106, 107, 145
 Deppermann, Arnulf, 333, 435, 564
 Derbyshire, Desmond C., 427
 Devlin, Keith, 251
 Dhonnchadha, E. Uí, 337

人名索引

- Diagne, Abdel Kader, 237
Diesing, Molly, 92
Dini, Luca, 237
Dione, Cheikh Mouhamadou Bamba, 196
Dipper, Stefanie, 195, 196, 417
Dodsen, Kelly, 496
Donati, C., 141
Donohue, Cathryn, 275
Doran, Christine, 377
Doran, Robert W., 109
Dorna, Michael, 237
Dörre, Jochen, 237
Dowty, David, 220
Dowty, David R., 5, 85, 145, 219, 224, 289, 316, 527, 533, 556, 562, 568, 573, 576, 595
Dras, Mark, 195
Drellishak, Scott, 239
Drosdowski, Günther, 23, 38
Dryer, Matthew S., 93, 408, 632
Dürscheid, Christa, 106, 502
Dyvik, Helge, xvii, 196

Eberhard, Kathleen M., 475, 476
Egg, Markus, 527–529
Eikmeyer, Hans-Jürgen, 219
Eisele, Andreas, 195
Eisenberg, Peter, xv, 12, 21–24, 26, 30, 32, 34, 39, 45, 59, 66, 145, 417, 437
Elbourne, Paul, 106, 158, 481, 637
Ellefson, Michelle R., 422
Elman, Jeffrey L., 430, 432, 441, 446, 458
Embick, David, 568
Emirkanian, Louise, 161
Engdahl, Elisabet, 319
Engel, Ulrich, 335, 338–343, 348, 365, 370, 375, 502, 521
Epstein, Samuel David, 144, 155
Erbach, Gregor, 237
Ernst, Thomas, 138
Eroms, Hans-Werner, 88, 335, 339, 342, 344, 346, 348, 353, 360, 361, 374, 521
Erteschik-Shir, Nomi, 422
Estigarribia, Bruno, 447
Estival, Dominique, 237
Evang, Kilian, 377
Evans, Nicholas, 408, 411, 414, 423, 428, 635
Evans, Roger, 161
Everett, Daniel L., 427
Evers, Arnold, 107, 108, 582
Evert, Stefan, xvi

Faaß, Gertrud, 196
Fabregas, Antonio, 144
Fabricius-Hansen, Cathrine, 12, 21–23, 39, 45, 66
Fagnano, Maria, 476
Falk, Yehuda N., 110
Fan, Zhenzhen, 237, 238
Fang, Ji, 196
Fanselow, Gisbert, xvi, 77, 101, 105–109, 113, 139, 153, 154, 219, 228, 267, 268, 365, 399, 411, 414, 421, 431, 471, 502, 545, 635
Feldhaus, Anke, 93
Feldman, Jerome, 436
Felix, Sascha W., 77, 101, 113
Fiehler, Reinhard, 12, 21–23, 39, 45, 66
Filimonova, Elena, 407
Fillmore, Charles J., 63, 73, 110, 155, 281–285, 293, 295, 296, 309, 445, 514, 521, 554, 555, 584
Finch, Steven, 455
Fischer, Ingrid, 513
Fischer, Klaus, 335
Fisher, Simon E., 432–434
Fitch, W. Tecumseh, 80, 131, 414, 423, 424, 426–428, 432, 434, 454
Fletcher, Paul, 433, 434
Flickinger, Daniel P., 108, 238, 239, 244, 253, 278, 285, 297, 301, 316, 382, 512, 528, 536
Fodor, Janet Dean, 470, 484, 485, 489, 490, 493
Fodor, Jerry A., 473
Fokkens, Antske, 238, 239
Fong, Sandiway, xvi, 109, 158
Fordham, Andrew, 109, 111
Forst, Martin, 35, 196, 197, 411
Fortmann, Christian, 502
Fourquet, Jean, 93
Fouvry, Frederik, 239
Fox Tree, Jean E., 475
Fraj, Fériel Ben, 377

- Frank, Anette, 93, 195, 196, 528, 665, 688, 689
 Frank, Robert, 378
 Franks, Steven, 444, 483
 Frazier, Lyn, 470, 471, 479
 Freidin, Robert, 109, 161, 415, 635
 Freudenthal, Daniel, 448, 494, 495
 Frey, Werner, 104, 106, 107, 111, 139, 144, 150, 195, 203, 277, 502
 Fried, Mirjam, 281, 283
 Friederici, Angela D., 430, 432
 Friedman, Joyce, 109, 111
 Fries, Norbert, 257
 Fritz, Thomas A., 12, 21–23, 39, 45, 66
 Fukui, Naoki, 92
 Fukumochi, Yasutomo, 336, 337
- Gaifman, Haim, 339, 366, 367
 Gallmann, Peter, xvii, 12, 21–23, 39, 45, 66, 112
 Gardent, Claire, 195, 461
 Gardner, R. Allen, 432
 Garrett, Merrill F., 473
 Gärtner, Hans-Martin, 11, 133, 144, 150, 159, 502
 Gazdar, Gerald, 109–111, 113, 147, 154, 161–163, 167–169, 171–173, 177, 179, 180, 278, 289, 321, 459, 461, 520
 Geach, Peter Thomas, 154, 178
 Geißler, Stefan, 238
 Geman, Stuart, 197
 George, Marie St., 433
 Gerdemann, Dale, 237
 Gerdes, Kim, xvi, 337, 341, 377
 Gerken, LouAnn, 487
 Ghayoomi, Masood, 238, 628
 Gibson, Edward, 267, 395, 408, 422, 470, 471, 484–486, 582
 Gillis, Steven, 493, 494
 Ginsburg, Jason, 109, 158
 Ginzburg, Jonathan, 110, 147, 152, 155, 244, 251, 276, 368, 373, 445, 448, 492, 618
 Giorgolo, Gianluca, 565
 Gipper, Helmut, 38
 Glauche, Volkmar, 432, 633
 Gobet, Fernand, 448, 494, 495
 Gold, Mark E., 435
- Goldberg, Adele, xvi
 Goldberg, Adele E., 281–283, 291, 315–317, 326, 333, 334, 420–422, 456, 458, 493, 496, 500, 518, 521, 549–552, 555, 556, 562, 564, 591, 599–602, 635
 Gopnik, Myrna, 433
 Gordon, Peter, 434, 437
 Gosch, Angela, 433
 Götz, Thilo, 237
 Grant, Julia, 433
 Grebe, Paul, 38
 Green, Georgia M., 498, 499
 Grewendorf, Günther, 77, 86, 87, 99, 103, 106–108, 113, 115, 131, 132, 139, 159, 257, 412, 416, 502, 519, 589
 Grewendorf, Güther, 131, 139
 Gries, Stephan Th., 606, 608, 609
 Grimshaw, Jane, xv, 142, 422, 601
 Grinberg, Dennis, 336
 Grohmann, Kleantes K., 136, 159, 426, 427
 Groos, Anneke, 142
 Groß, Thomas M., 143, 176, 342, 346, 348, 359, 364, 372, 376, 521, 545, 546, 604
 Groß, Thomas Michael, 32
 Grosu, Alexander, 422
 Grover, Claire, 161
 Grubačić, Emilija, 604
 Guillotin, T., 219
 Gunji, Takao, 545
 Gunter, Thomas C., 432
 Günther, Carsten, 319
 Guo, Yuqing, 197
 Guzmán Naranjo, Matías, 330
- Haddar, Kais, 237
 Haegeman, Liliane, 113, 134, 505, 506, 509, 540
 Haftka, Brigitta, 87, 139
 Hagen, Kristin, 337
 Hahn, Michael, 237
 Haider, Hubert, 77, 91, 94, 100, 101, 103, 106, 107, 111, 134, 137, 154, 209, 217, 256, 259, 260, 267, 277, 388, 411, 412, 416, 483, 502, 515, 519, 545, 575, 635, 637
 Hajičová, Eva, 335

人名索引

- Hakuta, Kenji, 430
Hale, Kenneth, 120, 427, 561, 568
Hall, Barbara C., 109, 111
Han, Chung-hye, 378
Hanlon, Camille, 456
Hansen-Schirra, Silvia, 417
Harley, Heidi, 572, 573
Harlow, Ray, 26
Harman, Gilbert, 75, 161
Harris, Zellig S., 4
Hasan, K. M. Azharul, 237
Haspelmith, Martin, xvi, xviii, 88, 405, 423, 484, 632
Haugereid, Petter, 238, 256, 568, 576, 577, 579, 581, 638
Hauser, Marc D., 80, 131, 414, 423, 424, 426, 428, 432, 434, 454
Hausser, Roland, 477
Hawkins, John A., 267, 408, 419, 422
Hays, David G., 336, 337, 339, 367
Heikkilä, Juha, 336
Heinecke, Johannes, 337
Heinz, Wolfgang, 99, 217, 575
Helbig, Gerhard, 335
Hellan, Lars, 26, 238, 638
Hellwig, Peter, 335–337, 341, 342, 366, 372, 376, 501, 637
Her, One-Soon, 195, 196
Heringer, Hans Jürgen, 344, 346, 348
Herzig Sheinfux, Livnat, 237, 238
Higginbotham, James, 278, 516
Higinbotham, Dan, 195, 196
Hildebrandt, Bernd, 219
Hinkelman, Elizabeth A., 237
Hinrichs, Erhard W., 108, 154, 260, 278, 306, 342, 396, 536, 544, 548, 582, 698
Hinterhölzel, Roland, 139
Hoberg, Ursula, 397, 543
Hockenmaier, Julia, 219, 220
Hockett, Charles F., 427
Hockey, Beth Ann, 377
Hoeksema, Jack, 591
Hoffman, Beryl Ann, 220, 228, 267
Hofmeister, Philip, 422
Höhle, Tilman N., 40, 47, 91, 93, 104, 227, 240, 262, 305, 465, 544, 699, 700
Holler, Anke, 244
Holmberg, Anders, 484
Hornstein, Norbert, 136, 157, 159, 423, 426, 427, 429
Howlin, Patricia, 433
Hrafnbjargarson, Gunnar Hrafn, 133
Hróarsdóttir, Þorbjörg, 133
Huang, Wei, 337
Hudson, Carla L., 432
Hudson Kam, Carla L., 432
Hudson, Richard, xvi, 27, 335, 336, 338, 342, 344, 370, 372, 373, 428
Hudson, Richard A., 336, 338, 342–344, 348–350, 353, 359, 521
Hukari, Thomas E., 305, 461, 519, 520
Humboldt, Wilhelm von, 424
Hunze, Rudolf, 195
Hurford, James R., 429, 446
Hurskainen, Arvi, 337
Hurst, Jane A., 433

Ibbotson, Paul, 634
Ichiyama, Shunji, 336, 337
Ingram, David, 495
Iordanskaja, L., 336, 337
Islam, Md. Asfaqul, 237
Islam, Muhammad Sadiqul, 237

Jackendoff, Ray S., 39, 70, 77, 86, 88, 108, 130, 145, 147, 150, 259, 282, 327, 334, 374, 409, 412, 423, 428, 446, 474, 478, 480, 481, 492, 521, 535, 549, 562, 588, 589, 591, 601, 611, 620, 621, 624
Jacobs, Joachim, 93, 173, 225, 226, 334, 610, 624
Jacobson, Pauline, 173, 179, 263
Jaeggli, Osvaldo A., 540
Jäger, Gerhard, 527
Jäppinen, H., 336, 337
Jäppinen, Harri, 337
Johannessen, Janne Bondi, 337
Johnson, David E., 109, 130, 426, 461
Johnson, Jacqueline S., 430
Johnson, Kent, 436, 437
Johnson, Mark, 111, 176, 178, 181, 194, 195, 197, 326, 401, 477–479, 501
Johnson, Mark H., 430, 432, 458
Jones, Wendy, 433

- Joshi, Aravind K., 88, 310, 376–378, 380, 382, 383, 385, 387, 388, 394–397, 400, 401, 576, 596
- Jungen, Oliver, xv
- Jurafsky, Daniel, 283, 518
- Kahane, Sylvain, xvi, 335, 337, 339, 341, 348, 353, 356, 359, 368, 374, 380
- Kallmeyer, Laura, xvi, 286, 377, 378, 380, 382, 400, 401, 520, 584, 585
- Kamp, Hans, 203
- Kanerva, Jonni M., 208
- Kaplan, Ronald M., 111, 113, 195, 197, 202, 206, 210, 213, 215, 217, 277, 293, 411, 426, 461, 501, 533
- Karimi, Simin, 134
- Karimi-Doostan, Gholamhossein, 134
- Karlsson, Fred, 336
- Karmiloff-Smith, Annette, 430, 432–434, 458
- Karttunen, Lauri, 211, 219
- Kasper, Robert T., 94, 170, 265, 292, 367, 400, 507
- Kasper, Walter, 237, 238, 330
- Kathol, Andreas, 176, 260, 265, 275, 310, 318, 325, 365, 370, 592, 603, 604
- Kaufmann, Ingrid, 563
- Kaufmann, Tobias, 237, 238, 639
- Kay, Martin, 109, 111, 336, 639
- Kay, Paul, xvi, 281–285, 292, 293, 295–297, 301, 309, 333, 445, 514, 553, 555, 563, 584, 587
- Kayne, Richard S., 133, 145, 146, 267, 623, 633
- Keenan, Edward L., 143, 147, 257
- Keil, Martina, 513
- Keller, Frank, 154, 238, 324, 326
- Kempen, Masja, 493, 494
- Kern, Sabine, 471
- Kettunen, Kimmo, 336
- Keyser, Samuel Jay, 120, 561, 568
- Khлentzos, Drew, 413, 422
- Kiefer, Bernd, 237, 332, 400
- Kifle, Nazareth Amlesom, 196
- Kim, Jong-Bok, 4, 113, 238, 425, 426
- Kim, M., 336, 337
- Kim, Nari, 378
- Kimball, John P., 434, 438
- King, Paul, 181, 194, 426, 461, 600
- King, Tracy Holloway, 195–197, 217, 277
- Kinyon, Alexandra, 195, 286, 378, 388, 584
- Kiparsky, Carol, 420, 421
- Kiparsky, Paul, 420, 421, 552
- Kiss, Katalin E., 134
- Kiss, Tibor, xvi, 75, 93, 105, 106, 108, 227, 238, 256, 257, 410, 416, 545, 643
- Kittredge, R., 336, 337
- Klann-Delius, Gisela, 500
- Klein, Ewan, 109, 110, 147, 161–163, 167–169, 171–173, 179, 219, 240, 278, 289, 307, 459, 461, 463, 520, 638
- Klein, Wolfgang, 145, 257, 336, 351, 428, 429, 437
- Klenk, Ursula, 80
- Kliegl, Reinhold, 471
- Kluender, Robert, 422, 429
- Knecht, Laura, 289
- Kobele, Gregory M., 154
- Koelsch, Stefan, 432
- Koenig, Jean-Pierre, 253, 274, 282, 286, 288, 331, 584, 588, 615, 630
- Kohl, Dieter, 195
- Kohl, Karen T., 486
- Koit, Mare, 337
- Kolb, Hans-Peter, 109, 111, 130
- Komarova, Natalia L., 4, 425, 484
- Konieczny, Lars, 322, 326
- König, Esther, 219, 417, 520
- Koopman, Hilda, 92
- Kordoni, Valia, 238, 286, 584
- Kornai, András, 72, 109, 409, 426
- Kornfilt, Jaklin, 91, 107
- Koster, Jan, 93, 95, 103, 104, 106, 111, 277, 415, 416
- Kouylekov, Milen, 237
- Kratzer, Angelika, 106, 502, 510, 568–571
- Krieger, Hans-Ulrich, 237, 290, 332, 564, 585
- Krifka, Manfred, 429
- Kroch, Anthony S., 378, 387, 401, 520, 576
- Kropp Dakubu, Mary Esther, 238
- Kruijff, Geert-Jan M., 528
- Kruijff-Korbayová, Ivana, 199
- Kubczak, Jacqueline, 335
- Kübler, Sandra, 336
- Kuhn, Jonas, xvi, 113, 159, 196, 206, 207, 217, 319, 480, 481, 492

人名索引

- Kuhns, Robert J., 109
Kunze, Jürgen, 335–337, 348, 359, 576
Kuperberg, Gina, 601
Kutas, Marta, 422
- Labelle, Marie, 473
Laczkó, Tibor, 196
Laenzlinger, Christoph, 129, 133, 135, 139, 267, 623
Lafferty, John D., 336
Lai, Cecilia S. L., 433
Lai, Zona, 433
Lakoff, George, 281, 460
Lamping, John, 203, 342
Langacker, Ronald W., 281, 312, 549, 550
Lappin, Shalom, 109, 130, 419, 421
Lareau, François, 195
Larson, Richard K., 101, 120, 121, 144, 150, 519, 539
Lascarides, Alex, 535
Laskri, Mohamed Tayeb, 237
Lasnik, Howard, 443, 479
Lavoie, B., 336, 337
Lavoie, Benoit, 336
Le, Hong Phuong, 378
Lee-Goldmann, Russell R., 63, 73
Legate, Julie, 443–445
Lehtola, A., 336, 337
Lehtola, Aarno, 336, 337
Leiss, Elisabeth, 4, 370, 469, 470, 517
Lenerz, Jürgen, 11, 101, 106, 108, 129, 543
Lenneberg, Eric H., 430
Levelt, Willem J. M., 218
Levin, Beth, 576
Levine, Robert D., xvi, 109, 130, 279, 305, 327, 330, 353, 354, 461, 519, 520
Levinson, Stephen C., 408, 411, 414, 423, 428, 635
Levy, Leon S., 377, 401
Lewin, Ian, 109, 111
Lewis, Geoffrey L., 289, 585
Lewis, John D., 441, 446
Lewis, Richard L., 471
Lezius, Wolfgang, 417
Li, Charles N., 617
Li, Wei, 238
Liakata, Maria, 196
- Lichte, Timm, xvi, 377, 385, 399
Lichtenberger, Liz, 433
Lieb, Hans-Heinrich, xv
Lieven, Elena, 283, 518
Lieven, Elena V. M., 500
Lightfoot, David W., 109, 493
Link, Godehard, 410
Lipenkova, Janna, 238, 334, 617, 618, 628
Liu, Gang, 238
Liu, Haitao, xvi, 337
Lloré, F. Xavier, 219
Lobin, Henning, 236, 335, 375
Löbner, Sebastian, 519
Lødrup, Helge, 35, 411
Lohndal, Terje, 568, 569, 576
Lohnstein, Horst, xv, 103, 109, 111, 112, 473, 520, 540
Longobardi, Giuseppe, 428
Lorenz, Konrad, 430
Lötscher, Andreas, 515
Loukam, Mourad, 237
Lüdeling, Anke, 605
Luuk, Erkki, 426, 428
Luuk, Hendrik, 426, 428
- Maas, Heinz Dieter, 337
Maché, Jakob, xvi, 613, 617
Machicao y Priemer, Antonio, 629
Mack, Jennifer, 601
Mackie, Lisa, 196
MacWhinney, Brian, 440
Maess, Burkhard, 432
Maienborn, Claudia, 319
Maier, Wolfgang, 377
Maling, Joan, 34, 108, 170, 209, 259, 575
Malouf, Robert, 153, 238, 255, 272, 291, 304, 321, 349, 354, 519, 520, 534, 588
Manandhar, Suresh, 237
Manshadi, Mehdi Hafezi, 161
Marantz, Alec, 140, 473, 475, 561, 568–570, 572, 612
Marciniak, Małgorzata, 237, 238
Marcus, Gary F., 432, 434, 456
Marcus, Mitchell P., 109
Marimon, Montserrat, 239
Marslen-Wilson, William, 475
Martinet, André, 424

- Martner, Theodore S., 109
 Masuichi, Hiroshi, 196
 Masum, Mahmudul Hasan, 237
 Matiassek, Johannes, 99, 217, 575
 Matsumoto, Yuji, 336
 Maxwell III, John T., 195, 197
 Mayo, Bruce, 195, 196
 McCloskey, James, 422
 McDonald, Ryan, 336
 Mchombo, Sam A., 206, 562
 McIntyre, Andrew, xvi
 McKean, Kathryn Ojemann, 471
 Meinunger, André, 106, 129, 133, 533
 Meisel, Jürgen, 409, 431, 483, 500
 Mel'čuk, Igor A., 335, 337, 342, 372, 520
 Melnik, Nurit, 237, 238
 Mensching, Guido, xvi, xvii, 491
 Menzel, Wolfgang, 336, 337
 Metcalf, Vanessa, 238
 Meurer, Paul, 196
 Meurers, Walt Detmar, 16, 93, 108, 154, 237, 238, 258–260, 265, 278, 279, 305, 316, 327, 330, 370, 396, 417, 512, 515, 535, 591, 638
 Meyer-Viol, Wilfried, 461
 Meza, Iván V., 239
 Micelli, Vanessa, 283, 290, 309
 Michaelis, Jens, xvi, 148, 151, 501, 502
 Michaelis, Laura, xvi
 Michaelis, Laura A., 217, 282, 286, 296, 297, 333, 521, 584, 637
 Michelson, Karin, 274, 331, 630
 Miller, George A., 394, 469, 471
 Mineur, Anne-Marie, 237
 Mistica, Meladel, 196
 Mittendorf, Ingo, 196
 Miyao, Yusuke, 239, 400
 Moeljadi, David, 238
 Moens, Marc, 638
 Mohanan, KP, 35, 411
 Mohanan, Tara, 35, 411
 Mok, Eva, 283
 Moll, Henrike, 431, 437
 Momma, Stefan, 195
 Monaco, Anthony P., 433
 Montague, Richard, 167, 219
 Moortgat, Michael, 219, 220
 Moot, Richard, 219
 Morgan, James L., 437
 Morin, Yves Ch., 109
 Moro, Andrea, 432, 633
 Morrill, Glyn V., 219, 220, 235
 Moshier, Andrew M., 288
 Motazedi, Yasaman, 195
 Muischnek, Kadri, 337
 Müller, Gereon, xv, 87, 107, 108, 129, 154, 156, 377, 502, 612–614, 622
 Müller, Max, 428
 Müller, Natascha, 429
 Müller, Stefan, xv, xviii, 6, 14, 27, 32, 39, 93, 100, 108, 140, 142–144, 147, 148, 151, 153–155, 159, 160, 194, 217, 225, 237–239, 243, 246, 252, 256, 257, 259, 260, 263, 265, 272, 275, 277–279, 281–283, 288–290, 293, 299, 304, 305, 308, 310, 315, 318, 324, 325, 328, 330–332, 334, 363–365, 370, 374, 388, 396, 399, 405, 411, 414, 415, 417, 426, 466, 467, 496, 502, 512, 515, 519, 522, 533, 535, 536, 544, 550, 552, 555, 558, 562–564, 566, 568, 571, 575, 576, 580–582, 587–593, 599–601, 603–606, 616–619, 627–629, 633, 638, 643, 653
 Muraki, Kazunori, 336, 337
 Musso, Mariacristina, 432, 633
 Muysken, Peter, 72, 146
 Mykowiecka, Agnieszka, 237, 238
 Nakayama, Mineharu, 446, 453
 Nakazawa, Tsuneko, 108, 154, 260, 278, 306, 342, 396, 536, 544, 548, 582, 698
 Nasr, Alexis, 348, 359
 Naumann, Sven, 161
 Nederhof, Mark-Jan, 332
 Neeleman, Ad, 282
 Nelimarkka, Esa, 337
 Nerbonne, John, 176, 178, 251, 290, 411, 528, 564, 585
 Netter, Klaus, 26, 75, 93, 225, 226, 237, 238, 396, 400, 633, 642
 Neu, Julia, 238
 Neumann, Günter, 237

人名索引

- Neville, Anne, 237
Nevins, Andrew Ira, 427
Newmeyer, Frederick J., 107, 134, 408, 409,
414, 415, 423, 429, 484, 492, 493,
519, 632
Newport, Elissa L., 430, 432
Ng, Say Kiat, 238
Nguyen, Thi Minh Huyen, 378
Niño, María-Eugenia, 196
Ninomiya, Takashi, 239
Nivre, Joakim, 336
Niyogi, Partha, 4, 425, 484–486
Noh, Bokyoung, 535, 563, 581
Nøklestad, Anders, 337
Nolda, Andreas, xv
Nomura, Hirosato, 336
Noonan, Michael, 289, 584
Nordgård, Torbjørn, 109
Nordhoff, Sebastian, xvi
Nordlinger, Rachel, 196, 278, 514
Nowak, Martin A., 4, 425, 484
Noyer, Rolf, 572, 573
Nozohoor-Farshi, R., 109
Nübling, Damaris, 12, 21–23, 39, 45, 66
Nunberg, Geoffrey, 492, 553, 571, 587, 603
Nunes, Jairo, 136, 154, 159, 426, 427
Nygaard, Lars, 336

Ochs, Elinor, 446, 456
O'Connor, Mary Catherine, 281, 555
Odom, Penelope, 471
O'Donovan, Ruth, 197
Oepen, Stephan, 108, 237, 239
Özkaragöz, İnci, 289, 584
Ofłazer, Kemal, 196
Ohkuma, Tomoko, 196
Oliva, Karel, 93, 263, 366, 372, 375
O'Neill, Michael, 4, 425
Oppenrieder, Wilhelm, 106, 544
Orgun, Cemil Orhan, 240
Ørsnes, Bjarne, xvi
Osborne, Miles, 220
Osborne, Timothy, xvi, 143, 176, 342, 346,
348, 356, 359, 364, 372, 374, 376,
521, 545, 546, 604
Osenova, Petya, 237
Osswald, Rainer, 286, 584, 585

Ott, Dennis, 143

Paczynski, Martin, 601
Pafel, Jürgen, 502
Paggio, Patrizia, 237, 319
Palmer, Alexis, 219, 220
Palmer, Martha, 378
Pankau, Andreas, xvi, xvii
Pankau, Rainer, 433
Parisi, Domenico, 430, 432, 458
Parmentier, Yannick, 377
Partee, Barbara H., 536
Passingham, Richard, 433, 434
Patejuk, Agnieszka, 196
Paul, Hermann, 43
Paul, Soma, 237
Pembrey, Marcus E., 433
Penn, Gerald, 237, 327
Pentheroudakis, Joseph, 195, 196
Perchonock, Ellen, 471
Pereira, Fernando, 309
Perles, Micha A., 519, 521
Perlmutter, David M., 86, 108, 127, 256, 410
Perry, John, 251
Pesetsky, David, 427, 572, 574, 634
Peters, Jörg, 12, 21–23, 39, 45, 66
Peters, Stanley, 79
Petrick, Stanley Roy, 109
Pfister, Beat, 238, 639
Phillips, Colin, 326, 471, 478
Phillips, John D., 161
Pickering, Martin, 236
Pienemann, Manfred, 218
Pietroski, Paul, 158, 441
Pietsch, Christian, xvi
Pihko, E., 605
Piñango, Maria Mercedes, 601, 602
Pine, Julian M., 441, 446, 448, 453, 494, 495
Pineda, Luis Alberto, 239
Pinker, Steven, 130, 218, 406–409, 411, 422,
423, 426, 428, 432, 458, 611
Pittner, Karin, 257
Plainfossé, Agnes, 195
Plank, Frans, 407
Plunkett, Kim, 430, 432, 458
Poletto, Cecilia, 133, 138
Polguère, A., 336, 337

- Pollack, Bary W., 109
Pollard, Carl J., 27, 32, 141, 147, 153, 154, 180,
181, 231, 233, 237, 239, 244, 246,
251–253, 255, 257, 260, 267, 270,
272, 274, 276, 278, 285, 288, 289,
297, 298, 304, 308, 325, 365, 374,
382, 426, 480, 490, 501, 506, 512,
514, 515, 520, 526, 528, 534, 575,
600
Pollock, Jean-Yves, 132, 134, 206
Ponvert, Elias, 601
Popowich, Fred, 237
Porzel, Robert, 309
Postal, Paul M., 109, 426, 461, 519
Poulson, Laurie, 239
Preuss, Susanne, 161, 170, 171
Prince, Alan, xv
Przepiórkowski, Adam, 108, 196, 237, 238,
259, 278, 279, 512, 515
Pullum, Geoffrey K., xvi, xviii, 4, 72, 86, 109,
110, 143, 147, 161–163, 165, 167–
169, 171–173, 179, 180, 257, 278,
289, 377, 409, 426, 427, 434, 436,
437, 439–441, 445, 448, 456, 458–
463, 483, 520
Pulman, Stephen G., 477
Pulvermüller, F., 605
Pulvermüller, Friedemann, 4, 446, 458, 602–
605
Puolakainen, Tiina, 337
Putnam, Michael, 143, 144, 376

Quaglia, Stefano, 196

Radford, Andrew, 136, 159, 407, 493
Rahman, M. Soheli, 237
Rahman, Md. Mizanur, 237
Rákosi, György, 196
Rambow, Owen, 88, 195, 336, 348, 359, 376–
378, 383, 385, 388, 392–397, 399–
401, 520
Ramchand, Gillian, 536
Randriamasimanana, Charles, 196
Raposo, Eduardo, 444, 483
Rappaport Hovav, Malka, 576
Rappaport, Malka, 573
Rauh, Gisa, 140

Reape, Mike, 88, 195, 265, 275, 292, 310, 318,
325, 332, 368, 502, 603, 638
Redington, Martin, 455
Reichenbach, Jürgen, 432, 633
Reis, Marga, 32, 33, 40, 45, 47, 93, 95, 106, 412
Rekau, Laura, 496
Remberger, Eva-Maria, 444, 491
Resnik, Philip, 378
Reyle, Uwe, 195, 203, 528
Rhomieux, Russell, 63, 73
Richards, Marc, 130, 131, 473, 476
Richter, Frank, xvi, 160, 181, 194, 237, 251,
275, 299, 301, 303, 309, 327, 330,
359, 426, 512, 516, 635
Rieder, Sibylle, 237
Riehemann, Susanne, 177, 541
Riehemann, Susanne Z., 177, 541
Riemer, Beate, 429
Riemsdijk, Henk van, 88
Riezler, Stefan, 195, 197
Rijntjes, Michel, 432, 633
Ritchie, R. W., 79
Rizzi, Luigi, 131, 132, 134, 139, 157, 413–415,
423, 483, 634
Roberts, Ian, 428
Roberts, Ian F., 484
Robins, Robert Henry, xv
Robinson, Jane, 309
Rodrigues, Cilene, 427
Rogers, James, 111, 426, 461
Rohrer, Christian, 196, 197, 417
Romero, Maribel, 401
Roosmaa, Tiit, 337
Rosén, Victoria, 196
Ross, John Robert, 108, 177, 233, 321, 410, 414,
419, 733
Roth, Sebastian, 196
Rothkegel, Annely, 336
Roussanaly, Azim, 378
Rowland, Caroline F., 441, 446, 453
Ruiter, Vera de, 335
Ruppenhofer, Josef, 217, 282, 286, 296, 521,
584, 637

Sabel, Joachim, 502
Sadler, Louisa, 196

人名索引

- Sag, Ivan A., xvi, 25, 27, 109, 110, 138, 141, 146, 147, 152–155, 161–163, 167–169, 171–173, 179, 181, 233, 237, 238, 244, 246, 251–253, 255, 257, 267, 270, 272, 274–276, 278, 282, 285, 288, 289, 291, 297, 298, 301, 304–310, 319, 321, 326, 330, 333, 334, 349, 354, 368, 373, 374, 382, 422, 445, 448, 459, 471, 475, 477, 478, 480, 481, 490, 492, 506, 512, 514, 515, 519, 520, 528, 534, 562, 571, 575, 587, 588, 592, 598, 600, 603, 606, 618, 619
- Sagot, Benoît, 195
- Sailer, Manfred, 251, 275, 299, 301, 309, 512, 513, 516
- Saito, Mamoru, 479
- Samarin, William J., 431
- Sameti, Hossein, 161
- Sampson, Geoffrey, 429, 441, 443
- Samvelian, Pollet, 410
- Saraswat, Vijay, 203, 342
- Sarkar, Anoop, 310, 377, 596
- Sato, Yo, 237, 275
- Sauerland, Uli, xvi, 106, 158, 481, 637
- Savin, Harris B., 471
- Schabes, Yves, 401, 513, 514
- Schäfer, Roland, xvii, 224
- Scheffler, Tatjana, 378, 388
- Schein, Barry, 568, 569
- Schenkel, Wolfgang, 335
- Scherpenisse, Wim, 99, 106, 175, 346
- Schieffelin, Bambi B., 446, 456
- Schlesewsky, Matthias, 471
- Schluter, Natalie, 197
- Schmidt, Paul, 237
- Schmidt, Renate, 335
- Scholz, Barbara C., 4, 426, 427, 434, 436, 437, 439–441, 445, 448, 456, 458, 460–463, 483
- Schopp, Andrea, 319
- Schröder, Ingo, 336, 337
- Schubert, K., 335
- Schumacher, Helmut, 335
- Schütz, Jörg, 237
- Schwarze, Christoph, xvi, 195, 218
- Sedivy, Julie C., 475, 476
- Segond, Frédérique, 196
- Seiffert, Roland, 237
- Seiss, Melanie, 196
- Sells, Peter, 4, 113, 238, 425, 426
- Sengupta, Probal, 195
- Sethuraman, Nitya, 552
- Seuren, Pieter A. M., 130, 431
- Sgall, Petr, 335
- Shamir, Eliahu, 519, 521
- Shanker, K. Vijay, 383
- Shanker, Vijay K., 400
- Shieber, Stuart M., 179, 181, 194, 309, 326, 377, 401, 477, 478, 501
- Shtyrov, Y., 605
- Shtyrov, Yury, 602–605
- Siegel, Melanie, 238
- Simov, Alexander, 237
- Simov, Kiril, 237
- Simpson, Jane, 196, 282, 515, 563, 736
- Singleton, Jenny L., 432
- Sleator, Daniel D. K., 336
- Sleator, Daniel Dominic, 336
- Smith, Carlota S., 572, 573
- Smith, George, 417
- Smolensky, Paul, xv
- Snedeker, Jesse, 601
- Snider, Neal, 422
- Snider, Nil, 419
- Snyder, William, 484
- Soehn, Jan-Philipp, 309
- Son, Minjeong, 484, 491
- Song, Sanghoun, 237, 238
- Spackman, Stephen P., 237
- Speas, Margaret, 92
- Spencer, Andrew, 519
- Spivey-Knowlton, Michael J., 475, 476
- Sportiche, Dominique, 92, 109
- Srinivas, Bangalore, 377
- Stabler, Edward P., 109, 111, 131, 144, 148, 151, 158, 159, 278, 412, 477, 479, 480, 501, 516, 637
- Ståding, Gabriele, 433
- Stark, Elisabeth, xvii
- Starosta, Stanley, 336, 520
- Stede, Manfred, 219
- Steedman, Mark J., 144, 145, 147, 151, 199, 219, 220, 222, 225, 228, 230–233, 236,

- 267, 291, 463, 477, 480, 594, 595,
607, 639
- Steels, Luc, 282, 283, 313, 317, 318, 332
- Stefanowitsch, Anatol, 457, 606, 608, 609
- Steinbach, Markus, 11, 107, 502
- Sternefeld, Wolfgang, 77, 91, 106, 107, 113,
131, 132, 136, 159, 160, 209, 359,
411, 476, 520, 635
- Stiebels, Barbara, 603
- Stowell, Timothy, 133, 420
- Stroik, Tom, 144
- Strunk, Jan, xvi, 418, 419
- Suchsland, Peter, 502
- Suckow, Katja, 471
- Sulger, Sebastian, 196
- Svenonius, Peter, 484, 491
- Svenonius, Peter, 134
- Tabbert, Eric, 219
- Takahashi, Masako, 377, 401
- Takami, Ken-ichi, 421
- Tanenhaus, Michael K., 475, 476, 599
- Temperley, Davy, 336
- ten Hacken, Pius, 464
- Tesnière, Lucien, 28, 335, 341, 351, 352, 356,
374, 376
- Theofilidis, Axel, 237
- Thiersch, Craig L., 93, 106, 109, 111
- Thomas, James, 470, 471
- Thompson, Henry S., 113, 161, 170
- Thompson, Sandra A., 617
- Thompson, William, 495
- Thornton, Rosalind, 413, 422
- Thráinsson, Höskuldur, 34, 170, 209, 575
- Timberlake, Alan, 289, 584
- Toivonen, Ida, 217, 277, 282, 549, 562–566
- Tomasello, Michael, 80, 107, 283, 334, 406,
423, 428, 431, 433, 437, 456, 458,
493, 495, 496, 500, 518, 549–552,
555, 634
- Torisawa, Kentaro, 400
- Tóth, Ágoston, 196
- Travis, Lisa, 408
- Trosterud, Trond, 337
- Tseng, Jesse L., 238, 256
- Tsujii, Jun'ichi, 239, 400
- Turpin, Myfany, 195
- Tyson, Mabry, 309
- Udwin, Orlee, 433
- Uibo, Heli, 337
- Ulinski, Morgan, 195
- Umemoto, Hiroshi, 196
- Uriagereka, Juan, 443, 444, 483
- Uszkoreit, Hans, 88, 161, 162, 169, 170, 173,
174, 179, 180, 219, 227, 237, 309,
382, 395, 417, 418, 470, 520
- Valian, Virginia, 487–489
- Valkonen, K., 336, 337
- Vallduvi, Enric, 319
- Van Eynde, Frank, 308
- van Genabith, Josef, 197, 337
- Van Langendonck, Willy, 27, 338
- van Noord, Gertjan, 219, 238, 291, 304, 461,
544
- van Riemsdijk, Henk, 142
- van Trijp, Remi, 313–319, 321, 322, 324–327,
333
- Van Valin, Robert D. Jr., 421
- Vancoppenolle, Jean, 219
- Vargha-Khadem, Faraneh, 433, 434
- Vasishth, Shravan, xvi, 471
- Vasserman, Alexander, 197
- Vater, Heinz, 77
- Veenstra, Mettina Jolanda Arnoldina, 109,
111, 149, 150, 158, 181, 412, 461, 520,
637
- Velupillai, Viveka, 422
- Vennemann, Theo, 26, 94, 145
- Verhagen, Arie, 518
- Verspoor, Cornelia Maria, 535, 563
- Vierhuff, Tilman, 219
- Vijay-Shanker, K., 388
- Villavicencio, Aline, 219, 220, 227, 236, 423,
445, 638
- Vogel, Carl, 237
- Vogel, Ralf, 107, 143, 502
- Volk, Martin, 161
- von Stechow, Arnim, 77, 107, 113, 133, 219,
519, 527, 533
- Voutilainen, Atro, 336
- Wada, Hajime, 195
- Walker, Donald E., 109, 111

人名索引

- Walther, Markus, 240, 307
Wang, Haifeng, 197
Wasow, Thomas, 25, 237, 239, 274, 278, 297,
298, 326, 471, 475, 477, 478, 492,
520, 571, 587, 603
Wasow, Tom, xvi, 237, 478, 481
Watkins, Kate E., 433, 434
Way, Andy, 197
Webelhuth, Gert, 103, 106, 107, 133, 490, 618,
624, 638
Weber, Heinz J., 335, 351, 355
Wechsler, Stephen Mark, xvi, 140, 144, 253,
282, 290, 315, 325, 334, 388, 496,
535, 550, 552, 560, 561, 563, 566,
570–572, 574, 581, 601
Wedekind, Jürgen, 195
Wegener, Heide, 398
Weiller, Cornelius, 432, 633
Weir, David, 383
Weir, Morton W., 432
Weissgerber, Monika, 336
Weisweber, Wilhelm, 161, 170, 171
Welke, Klaus, 335, 338, 626
Wells, Rulon S., 310, 603
Werner, Edeltraud, 335
Wesche, Birgit, 93, 227
Wetta, Andrew C., 176, 275, 592, 604
Wexler, Kenneth, 145, 408, 484–486, 493
Weydt, Harald, 424
Wharton, R. M., 436
White, Mike, 219
Wiese, Heike, 601
Wijnen, Frank, 493, 494
Wiklund, Anna-Lena, 133
Wilcock, Graham, 319
Wilder, Chris, 106
Wiley, Edward, 430
Williams, Edwin, 217, 411
Wing, Ben, 219, 220
Winkler, Susanne, 107, 278, 318, 516
Wintner, Shuly, 237, 238
Wittenberg, Eva, xvi, 601, 602
Wöllstein, Angelika, 40
Wood, Randall, 4, 425
Wunderlich, Dieter, 80, 282, 423, 519, 521,
563
Wurmbrand, Susanne, 156, 544
Xia, Fei, 377
XTAG Research Group, 378
Yamada, Hiroyasu, 336
Yamada, Jeni, 433
Yampol, Todd, 219
Yang, Charles D., 443–445, 486, 491
Yang, Jaehyung, 238
Yankama, Beracah, 158, 441
Yasukawa, Hideki, 195
Yatsushiro, Kazuko, 106
Yip, Moira, 108, 259
Yoon, Juntae, 378
Yoon, SinWon, 378, 388
Yoshinaga, Naoki, 400
Zaenen, Annie, 34, 170, 195, 206, 209, 213,
215, 293, 575, 689
Zalila, Ines, 237
Zappa, Frank, 406
Zeevat, Henk, 219, 638
Zhang, Yi, 238, 239
Ziehe, T. W., 336, 337
Zorn, Hans-Peter, 309
Zribi, Chiraz, 377
Zucchi, Alessandro, 560
Zwart, C. Jan-Wouter, 134
Zweigenbaum, Pierre, 195
Zwicky, Arnold M., 109, 111

Language index

- Abaza, 414
Akan, 419
Akkadian, 427
Arabic, 195, 220, 237, 377
Arrernte, 195

Basque, 484, 527, 533
Bengali, 195, 237
Bulgarian, 237

Cantonese, 237
Chamorro, 272
Czech, 484

Danish, 195, 237, 419, 445
Dutch, 219, 225, 238, 415n13, 493–495, 544, 582, 591
Dyirbal, 409, 427

English, 25, 80, 195, 202, 207, 219, 228, 238, 283, 301, 378, 388, 407, 413, 415, 422, 430, 437–440, 448, 453, 456, 457, 485, 487, 490, 493–495, 512, 515, 518, 519, 534, 542, 594, 611, 615, 618–620, 625, 629
Esperanto, 238
Ewe, 272

Finnish, 211n10, 219
French, 25, 195, 218, 219, 238, 272, 378, 456, 629

Ga, 238
Galician, 444, 483n1
Georgian, 196, 238
German, 453, 495, 627, 628

Greek, 238
Guugu Yimidhirr, 414

Hausa, 238
Hawaiian Creole English, 431
Hebrew, 238, 484
Hindi, 196
Hixkaryana, 427
Hungarian, 196, 409

Icelandic, 272, 544
Indonesian, 196, 238
Irish, 196, 272, 584
Italian, 196, 378, 432, 483n1, 487, 487

Jakaltek, 409
Japanese, 80, 196, 228, 238, 407, 419, 432, 484
Javanese, 484
Jiwarli, 409

Kikuyu, 272
Korean, 196, 238, 378, 419, 484

Latin, 228
Lithuanian, 584

Malagasy, 196
Malayalam, 409, 484
Maltese, 238, 628
Mandarin Chinese, 196, 238, 422, 453, 487, 617–618, 628
Moore, 272
Murrinh-Patha, 196

Northern Sotho, 196
Norwegian, 196, 238, 419, 445

语言索引

Oneida, 630

Palauan, 272

Persian, 238, 628

Pirahã, 427

Polish, 196, 238

Portuguese, 196, 238

Proto-Uralic, 427

Russian, 211n10, 238

Sahaptin, 239

sign language, 431–432

American (ASL), 430, 484

British, 239

French, 239

German, 239

Greek, 239

South African, 239

Slavic, 335

Sorbian

Lower, 483n1

Upper, 483n1

Spanish, 196, 239, 272, 292, 494–495, 629

Swedish, 419, 562n19

Tagalog, 220, 228, 409

Tamil, 419

Thompson Salish, 272

Tigrinya, 196

Turkish, 196, 220, 228, 239, 290, 584

Urdu, 196

Vietnamese, 378

Wambaya, 239, 409, 515

Warlpiri, 409, 427

Welsh, 196

Wolof, 196

Yiddish, 239, 272, 378, 629

世界语, 336

丹麦语, 336

俄语, 337

加泰罗尼亚语, 337

古冰岛语, 337

巴斯克语, 337

德语, 337

意大利语, 337

拉丁语, 337

挪威语, 337

斯瓦希里语, 337

日语, 337

法罗语, 337

法语, 25, 337

爱尔兰语, 337

爱沙尼亚语, 337

现代汉语, 337

瑞典语, 337

芬兰语, 337

英语, 25, 336

萨米语, 337

葡萄牙语, 337

西班牙语, 337

Subject index

- + , 375, 622
- \ , 221, 228
- , 368
- ↓ , 202, 235, 378
- ∃ , 84
- ∀ , 84
- λ-calculus, 56
- ⊕ , 245
- , 49
- ↑ , 202, 230, 235
- ↑↑ , 231
- ∨ , 186
- | , 212, 228
- * , 9n2, 59, 70, 122, 170, 215, 378
- / , 220–221, 228
- # , 9n2
- § , 9n2
- ⇒ , 245, 249
- ¬ , 204

- a-structure, 208
- accent, 240
- accomplishment, 554
- acquisition, 429–430, 434–457, 480, 483–500, 625
 - speed, 429
- Across the Board Extraction, 154, 177, 321, 604n49
- actant, 28
- activity, 554
- adjacency, 51
- adjective, 17–19, 22, 86, 254, 423
 - depictive, 38
 - predicative, 38
- adjunct, 28–32, 59–61, 70, 85, 92, 139, 169–170, 206, 223–224, 246, 253–256, 285, 378–379, 420, 507
 - head, 70
- adjunction, 379
 - ban, 514
 - obligatory, 388, 391–392
- adjuncts, 205
- adposition, 21, 423
- adverb, 17, 21, 86
 - pronominal-, 23
 - relative, 23
- adverbial, 32n14, 36–37
- adverbs, 20
- agent, 28, 85, 208
- Agree, 118, 138, 158, 510, 543n1
- agreement, 5, 32, 52–54, 158, 187
 - object, 527, 533
- ambiguity, 396, 397
 - spurious, 66, 72, 467, 567
- analogy, 454
- animacy, 85
- antipassive, 612–616
- apparent multiple fronting, 176n14
- Arc Pair Grammar, 426
- argument, 28–32, 84, 220, 246, 252, 420
 - designated, 217, 256, 286n5, 637
 - external, 85, 91, 217
 - internal, 85, 217
 - position, 85
- argument structure, 84
- article, 17, 49
- aspect, 554
- attribute-value matrix (AVM), 181
- attribute-value structure, 181
- automaton
 - finite, 456

术语索引

- auxiliary inversion, 90, 109, 263n16, 448, 448–456, 493
- back-formation, 93
- backward application, 151, 221
- base generation, 106
- benefactive, 565–567
- beneficiary, 85, 208, 209
- β -reduction, 57, 222
- bias, 494
- binary, 50
- Binding Theory, 83, 130, 252, 257n14, 414, 506, 507, 519
- Biolinguistics, 326
- bit vector, 332
- blocking, 456
- bounding node, 414
- branching
 - binary, 50, 144, 427n27, 448, 454, 505–508, 579, 632
 - unary, 51
- Broca's area, 432
- Burzio's Generalization, 103, 612
- c-command, 267n20, 413
- c-structure, 170n2, 198, 202
- capacity
 - generative, 148, 170, 179, 400, 501–504
- cartography, 131n6, 423
- case, 18, 19, 30, 54, 217, 396
 - absolutive, 627
 - accusative, 19, 35, 37, 39, 627
 - agreement, 39, 99n17
 - dative, 19, 100
 - ergative, 627
 - filter, 101
 - genitive, 19, 100, 259
 - lexical, 99–208
 - nominative, 19, 32, 100, 627
 - semantic, 37
 - structural, 99–411
- Case Theory, 130, 217
- Categorical Grammar (CG), 148–155, 159, 173, 178, 179, 219–235, 240, 260, 263n16, 276–278, 282, 285, 396, 401, 410, 411, 427, 445, 459, 461, 463, 471, 476, 478, 481n13, 506, 520, 528, 533, 536n5
- category
 - functional, 86, 131, 139, 423
 - Pol, 133
 - μ , 133
 - +Pol, 133
 - %Pol, 133
 - Adverb, 133, 137
 - Agr, 207
 - AgrA, 133
 - AgrIO, 133, 411
 - AgrN, 133
 - AgrO, 132, 133, 411, 527, 533
 - AgrS, 132, 133, 411, 527
 - AgrV, 133
 - Asp, 133, 134
 - Aux, 133
 - C, 88, 423
 - Clitic Voices, 133
 - Color, 133n9
 - D, 88
 - Dist, 133, 134
 - Fin, 132
 - Foc, 132, 139
 - Force, 132
 - Gender, 133
 - Hearer, 133, 138
 - Honorific, 133
 - I, 88, 91, 101, 423
 - Intra, 134
 - Kontr, 139
 - Mod, 133
 - Mood, 133
 - Nationality, 133n9
 - Neg, 132, 133, 207, 360
 - Number, 133
 - Obj, 134
 - OuterTop, 133
 - Pass, 541
 - Passive, 128
 - PassP, 539
 - PathP, 134
 - Perf, 541
 - Perfect(?), 133
 - Person, 133
 - PlaceP, 134
 - Predicate, 133
 - Quality, 133n9

- Shape, 133n9
- Share, 133
- Size, 133n9
- Speaker, 133, 138
- Subj, 134
- T, 88, 133, 207, 527
- Tense, 121, 133
- Top, 132, 134, 139
- Tra, 134
- v, 119–121, 539, 569–571
- Voice, 133, 510, 527
- Z, 133
- lexical, 25, 86, 423
- syntactic, 25
- causative construction, 290
- CAUSE, 527
- change of state, 554
- CHILDES, 440, 441, 442n35, 444, 454
- chimpanzee, 430
- classifier, 423
- clitic, 495
- CoGETI, 239
- Cognitive Grammar, 281, 405, 445, 495, 521
- coherence, 210, 214, 216
- coindexation, 251
- comparative, 20, 22
- competence, 325–327, 394–399, 425, 469–481, 487, 498
- complement, 32, 70, 85, 144–149
- complementizer, 96, 227, 262
- completeness, 210, 214, 216
- Complex NP Constraint, 414, 419
- complexity class, 79, 377, 456, 501–504
- composition, 229, 437
 - backward, 228, 229, 595
 - forward, 228, 229
- compositionality, 253, 429
- computer science, 108
- conjunction, 20, 21, 49, 140n14
- constituent, 7, 12
 - discontinuous, 40, 88, 292, 310, 598
- constituent order, 106, 166, 210, 213, 227–267, 292, 382–385
 - fixed, 228, 267
 - free, 228, 267
- constituent position, 265
- Constraint Language for Lambda-Structures (CLLS), 528
- constraint-based grammar, 426
- construction
 - Active Ditransitive, 311
 - Caused-Motion, 317, 318, 556, 562, 566, 602
 - Determiner Noun, 310
 - linking, 285
 - N-P-N, 374–375, 377n2, 620–622
 - passive, 286
 - resultative, 282, 518, 535–536, 563–581
 - subject, 286
 - transitive, 286
 - verb phrase, 284
- Construction Grammar (CxG), 88, 106, 140, 155, 158n38, 173, 178, 179, 211n10, 217, 251, 253, 277, 281–334, 377, 405, 445, 471n3, 478, 481n13, 495–498, 506, 514, 517, 521, 541, 550, 564n22, 591, 622, 637
 - Fluid, 290
 - Sign-Based, 309, 325–333
- Construction Grammar(CxG), 586
- context, 446
- context-free grammar, 179, 377, 435, 501
 - probabilistic (PCFG), 283, 518
- context-sensitive grammar, 377, 435, 501
- contraction, 83
- contrast, 476
- control, 32, 199
- Control Theory, 84, 130
- conversion, 25
- cooperativeness, 396
- coordination, 15–16, 44n22, 61, 73, 110, 150–151, 176n14, 232, 244, 245, 310n26, 427, 478n11, 497
 - test, 10, 15
- copula, 41, 520
- Copy Theory of Movement, 140n13, 141, 154
- core grammar, 86, 281, 483, 492
- CoreGram, 239
- corpus linguistics, 457, 638
- coverb, 408, 423
- creole language, 431–432
- critical period, 429–430
- cycle

术语索引

- in feature description, 189, 265, 312, 369
- transformational, 474, 509
- D-structure, 81, 130, 277
- Deep Structure, *see* D-Structure, 275, 278
- Definite Clause Grammar (DCG), 75, 194
- deletion, 473
- Dependency Categorical Grammar, 236
- Dependency Grammar (DG), 88, 106, 155, 173, 401, 470n1, 597n45, 604n49
- Dependency Unification Grammar (DUG), 372, 501, 637
- depictive predicate, 257n14, 278n26, 515–517
- derivation, 25, 111, 290, 627
- derivation tree, 380
- Derivational Theory of Complexity (DTC), 471n3, 471–475
- descriptive adequacy, 405
- determiner, 22, 49
 - as head, 26
- directive, 610
- disjunction, 186–187, 212
- Distributed Morphology, 612
- do*-Support, 495
- dominance, 50
 - immediate, 50, 165, 240, 291
- economy
 - transderivational, 131
- elementary tree, 378
- ellipsis, 257n14, 473, 478n11, 506, 519, 612
- empty element, 14n7, 63, 103, 144, 150, 153, 265, 276, 282, 293, 304, 384, 394, 400, 409, 412, 506, 519–536, 606, 618
 - PRO, *see* PRO
- empty head, 611, 620
- endocentricity, 88
- entity, 167
- epsilon, 521
- epsilon production, 63
- escape hatch, 416
- evokes operator, 312–313
- Exceptional Case Marking (ECM), 510
- experiencer, 28, 208
- explanatory adequacy, 405
- Extended Projection Principle (EPP), 411, 489
- external argument, 86
- extraction, 415, 422, 520
 - from adjuncts, 502
 - island, 419
 - subject, 483
- extraction path marking, 321
- extraposition, 95n12, 133n10, 154, 365, 396, 415, 419
- Extraposition, 415
- f-structure, 170n2, 198, 202, 506, 509, 514
- Faculty of Language
 - in the Broad Sense (FLB), 428
 - in the Narrow Sense (FLN), 428
- feature
 - checking, 117
 - deletion, 117
- feature checking, 131
- feature description, 181–192
- feature structure, 192
- feature-value structure, 181
- feral child, 430
- field
 - middle-, 41, 92
 - post-, 41
 - pre-, 41, 92
- filler, 174, 284
- focus, 131, 139, 198–199, 214
- Foot Feature Principle, 176
- foot node, 391
- formalization, 5–6, 282, 496
- forward application, 151, 220
- FoxP2, 434
- fronting, 9, 12–15
 - apparent multiple, 16n8, 139, 154, 363, 365, 592n39, 604n49
- function composition, 240, 278
- functional application, 167
- functional uncertainty, 215, 293, 514, 582
- functor, 220, 499
- future, 18, 494
- gap, 82
- gender, 19, 54, 251, 464–466, 509
- gene, 434, 635
- Generalized Phrase Structure Grammar (GPSG), 111, 194, 210n8, 211, 220, 235, 240, 246, 247, 260, 265,

- 275, 277, 292, 293, 297, 311, 349,
367n29, 377, 379, 383, 401, 426,
459, 461, 471, 478, 481n13, 520,
637
- Generative Grammar, 50, 426, 459
- genitive, 35, 518
- Gesetz der wachsenden Glieder, 509
- glue language, 203
- glue semantics, 203–205, 342
- goal, 208
- government, 30, 109
- Government and Binding (GB), 171, 172n5,
198, 200, 201, 203, 209, 210, 220,
224, 225, 228, 235, 240, 260, 275,
277, 282, 363, 379, 384, 396, 411,
413, 441, 471n3, 471, 474, 506, 509,
512, 516, 519, 527, 594n44, 612,
619, 638
- gradability, 461–462
- Grammar Matrix, 239
- grammatical function, 207, 256, 284, 411–
412, 638
 - governable, 198
- Greediness Constraint, 486n2, 486
- head, 26–28, 202, 246
- head domain
 - extended, 210
- head feature, 28
- Head Feature Convention (HFC), 162, 247
- Head-Driven Phrase Structure Grammar
(HPSG), 88, 131n6, 139, 140, 155,
170n2, 173, 176n14, 179, 217,
226n6, 228, 237–278, 282, 285,
286, 298, 310, 311, 325–333, 349,
350, 354, 364, 365, 374, 377, 388,
401, 411, 414, 426, 445, 461, 464,
471, 474, 478, 481n13, 481, 490,
506, 507, 511, 512, 515, 520, 522,
528, 534, 536n5, 596, 598, 604n49,
624, 629, 637
 - Constructional, 304
- Heavy-NP-Shift, 292
- Hole Semantics, 528
- hydra clause, 410
- hypotenuse, 312
- iambus, 487
- Icelandic, 259n15
- ID/LP grammar, 165, 210n8, 240, 382
- identification in the limit, 435–437
- ideophone, 423
- idiom, 309, 492, 512–517, 586
- imperative, 18, 32
- implication, 245, 249
- indicative, 18
- inflection, 18–20, 24, 82, 627
 - paradigm, 18
- inflectional class, 19
- information structure, 98, 139, 198, 319,
564n21
- inheritance, 217, 253, 284, 285, 290, 564, 586–
588, 620
 - default, 373, 586
 - multiple, 185, 388
- instrument, 208
- Integrational Linguistics, xv
- interface, 326
- interjection, 20, 21
- interrogative clause, 44–45, 244, 618–620
- intervention, 125
- inversion, 519
- IQ, 433
- Kleene star, 170
- label, 139–144, 152
- λ -abstraction, 57
- λ -conversion, 57
- language
 - formal, 108, 434
- language acquisition, 107, 121, 220, 283, 405,
414n11, 505
- language evolution, 283
- learnability, 436
- learning theory, 448
- left associativity, 221
- Lexical Decomposition Grammar, 282
- Lexical Functional Grammar (LFG), 139, 155,
159, 170n2, 173, 178, 194, 275–278,
282, 293, 341, 350, 394, 411, 412,
426, 459, 461, 471n3, 471, 478,
481n13, 481, 506, 507, 511, 512, 514,
520, 524, 526, 554, 564n22, 591,
596, 624, 637
- lexical integrity, 206, 282

术语索引

- Lexical Mapping Theory (LMT), 208–209, 260
- Lexical Resource Semantics (LRS), 251n9
- lexical rule, 172, 225, 256–536
verb-initial position, 263–265
- lexicon, 82, 251
- linear logic, 203
- linear precedence, 165, 240, 291
- Linear Precedence Rule, 165
- linearization rule, 153, 206, 311
- Linguistic Knowledge Builder (LKB), 237
- Link Grammar, 336
- linking, 86, 208–209, 252–253, 285–290
- list, 183, 648
difference, 649
- local maximum, 486
- locality, 138, 297–304, 309, 380, 400, 508–517
of matching, 124
- locative alternation, 600
- Logical Form (LF), 81, 84, 277, 401
- long-distance dependency, 96–99, 176, 213–216, 231, 263n16, 272, 293, 297, 388
- long-Distance dependency, 228–350
- LP-rule, 165
- macaque, 430
- macro, 217
- Mainstream Generative Grammar, 77
- Mandarin Chinese, 408, 437
- Markov model, 456
- matryoshka, 4, 6
- maturation, 486
- Meaning–Text Theory (MTT), 335, 336n2, 342, 372n34
- meaning constructor, 204
- memory, 473n4
- Merge, 131, 139, 410, 427n27, 431
External, 131, 632
Internal, 131
- metarule, 166–167, 171–173, 508
- metrical grid, 240
- metrics, 487
- Middle Construction, 288, 290
- middle field, *see* field
- mildly context-sensitive grammar, 377, 400
- Minimal Recursion Semantics (MRS), 244, 251, 255n10, 285, 297, 382, 528, 580
- Minimalist Grammar (MG), 131n6, 148–155, 158, 278, 501, 502
- Minimalist Program (MP), 115–157, 171, 326, 360, 409, 410, 427, 461, 463, 475, 536n5, 594n44, 619, 637, 638
- Missing VP effect, 470
- model, 192
- model-theoretic grammar, 77, 426, 459–467
- modifier, 32, 85
- modularity, 476
- module, 480
- modus ponens, 205
- mood, 18
- morphology, 82, 158, 206–624
- Move, 131
- movement
altruistic, 139
covert, 137
permutation, 12
- movement test, 8–9
- Move α , 82
- Multi-Component TAG, 382, 383, 385
- music, 428, 432
- nativism, 405
- negative evidence, 435, 436, 440, 443, 456–457
- Neo-Davidsonian semantics, 580
- neural network, 446
- neutral order, 104
- Nirvana, 488
- No Tampering Condition (NTC), 326n34, 478n11
- node, 50
daughter, 50
mother, 50
sister, 50
- nominalization, 100, 456
- Non-Tangling Condition, 88
- noun, 17, 18, 22, 49, 63, 86, 423
common, 309
mass, 64
relational, 61
- NP-split, 13

- number, 18, 19, 52, 54, 251
- numeration, 158
- o-command, 267n20
- object, 277, 411–412
 - direct, 35, 36
 - indirect, 35, 36
- obliqueness, 257n14, 411
- Obliqueness, 267n20
- observational adequacy, 405
- Optimality Theory (OT), xv, 471n3
- optional infinitive, 493
- optionality, 32
- order
 - unmarked, 104
- organ, 432
- paradigm, 24
- parameter, 80, 483–491, 625
 - default value, 485, 486, 489
 - head direction, 407–408
 - head position, 80
 - pro-drop, 444, 445, 483, 487–489
 - subjacency, 483
 - subject article drop, 487
 - SV, 484
 - topic drop, 487
 - V2, 484, 486
- parameters, 431
- parser, 502
- partial verb phrase fronting, 178, 305, 633
- participle
 - adjectival, 127
- particle, 20
- passive, 99, 100, 104, 127–129, 173, 206–209, 224–225, 256–290, 296, 342, 387–388, 473, 513, 533, 564n21, 566, 624
 - impersonal, 172, 259, 519
 - long, 277n25, 279
 - remote, 544
- path, 182
- path equation, 309
- patient, 28, 85, 208
- PATR-II, 194, 309
- performance, 283, 325–327, 394–399, 422, 425, 469–481, 487, 498, 582, 634
- periphery, 281, 492
- permutation test, 8–9
- person, 18, 52
- Person, 251
- phase, 326, 475
- Phenomenon, 192
- Phonetic Form (PF), 81, 83
- phrase, 6
- phrase structure grammar, 49–55, 220, 240, 242, 459, 519
- pidgin language, 431
- pied-piping, 233
- pivot schema, 495
- plural, 18, 64
- Polish, 512
- positional, 423
- postfield, *see* field
- postposition, 20
- Poverty of the Stimulus, 413, 431, 434–457
- predicate, 84
- predicate logic, 28
- predicate-argument structure, 428
- predicative, 38–39
- prefield, *see* field
 - ellipsis, 257n14
- preposition, 17, 20, 66–68, 86
- present, 18
- presupposition, 421
- preterite, 18
- principle
 - case, 627
 - Case, 259–260
 - Generalized Head Feature, 373
 - Head Feature, 248
 - NONLOC-, 270
 - Semantics, 253, 627
 - Sign, 298n15
 - Subject-Mapping, 208
- Principles & Parameters, 80, 483–491
- PRO, 384
- Probability Matching, 432
- progressive, 495
- projection, 27, 88
 - maximal, 27, 72, 89
 - of features, 28
- projectivity, 341, 383, 546
- pronominalization test, 7
- pronoun, 22–627

术语索引

- expletive, 10–11, 24, 28, 103, 412, 444, 488, 519
- reflexive, 251, 507
- relative, 520
- prosody, 232, 463, 476
- quantification, 205
- quantifier
 - existential, 84
 - universal, 84, 168
- question tag, 512
- raising, 387, 512, 581
- Random Step, 486
- Rangprobe, 43
- recursion, xvii, 45–46, 61, 400, 454, 469, 470n1, 517–518
- recursively enumerable language, 436
- reference, 10, 11
- regular language, 435, 501
- relation, 508
 - , 368
 - append*, 194, 245, 267
- Relational Grammar, xv, 461
- relative clause, 44–45, 244, 422, 469, 618–620
 - free, 142–143, 257n14
- remote passive, 342n11
- repetitive, 527
- representational model, 111
- restitutive, 527
- resultative construction, 484, 535
- rewrite grammar, 434
- Right Roof Constraint, 414
- right to left elaboration, 447
- rising, 545
- Role and Reference Grammar, xv, 140
- root, 625
- root Infinitive, 493
- rule-to-rule hypothesis, 58, 168
- S-structure, 81, 130, 277
- S-Structure, 401
- Satzglied, 12n3
- schema
 - Filler-Head, 271
 - head-adjunct, 255
 - head-argument, 245
- Schema
 - Specifier-Head-, 273
- scope, 94–106, 176n14, 206, 292, 382, 519, 604n49
- segmentation, 494
- selection, 28
 - restriction, 396
- self-embedding, 425n25, 469, 518
- semantic role, 28, 85, 208–209, 252, 285, 577
- semantics, 158, 253
- sentence bracket, 40, 92
- sentence symbol, 251
- set, 293–296
- sexus, 465
- Shortest Move Constraint (SMC), 502
- sideward movement, 154
- sign language, 432
 - American (ASL), 432
- signature, 191
- Single Value Constraint, 486
- singular, 18
- Situation Semantics, 251n9
- specifier, 70, 144–149, 502
- statistics, 197, 220, 283, 378, 446, 448–457, 606
- stem, 627
- Structure Preservation, 478n11
- structure sharing, 187–189, 260, 306, 309, 349, 474
- subcategorization, 84, 162
- subcategorization frame, 84
- subjacency, 414–422, 484
- Subjacency, 130
- subject, 32–35, 100, 277, 284, 411–412, 512
- Subject Condition, 411
- subjunctive, 18
- substitution, 378
- substitution test, 7
- subsumption, 295
- superlative, 20
- Surface Structure, *see* S-structure
- SVO, 485
- symbol
 - non-terminal, 435
 - terminal, 435
- syntax-semantics interface, 56
- T model, 81–82

- TAG
 - Free Order (FO-TAG), 383
- Tamagotchi, 429
- tense, 18, 22
- text, 427n26
- that-t, 483
- the*-clause, 619
- thematic grid, 84
- theme, 85, 208
- theory, 192
- Theta-Criterion, 85, 130, 201, 217
- θ -grid, 84, 198
- θ -role, 85
- θ -Theory, 111, 217, 479
- third construction, 633
- Tiger corpus, 417
- topic, 131, 139, 198–199, 214
- Topic Drop, 257n14
- topicalization, 98, 230, 231, 619
- topology, 40–46
- trace, 82
 - extraction trace, 270, 272
 - verb, 14n7, 263, 265
- TRALE, 327
- transderivational constraint, 143, 158
- transformation, 78–80, 275, 282, 400, 401, 445, 471–475, 478, 533–536
 - fact*-Deletion, 420
- Transformational Grammar, 470n1, 77–470
- Tree Adjoining Grammar (TAG), 88, 148, 151, 155, 172n5, 277, 286, 310n26, 377–401, 459, 461, 478, 481n13, 490, 501, 517, 520, 533, 596, 598, 624, 637, 638
 - Feature Structure-Based* (FTAG), 388
 - Synchronous* (STAG), 401
 - Feature Structure-Based* (FTAG), 389–392
 - Multi-Component* (MC-TAG), 383–385
 - Tree Tuple MC-TAG* (TT-MC-TAG), 385, 399
 - Vector* (V-TAG), 385, 392–394, 399, 401, 520
- trigger, 491
- trochee, 487
- truth value, 167
- type, 186, 187
 - atomic, 185
 - type hierarchy, 185–186, 253
 - type raising, 228, 229, 520, 536n5
 - backward, 595
- underspecification, 528
- Underspecified Discourse Representation Theory (UDRT), 528
- unification, 189–191, 222, 295
- uniformity, 412, 519
- universal, 429
 - implicational, 407n4
 - implicative, 635
- Universal Base Hypothesis, 637
- Universal Grammar (UG), 131, 133, 220, 405, 543, 614
 - falsifiability, 422n19, 423, 428
- unrestricted grammars, 501
- Unsupervised Data-Oriented Parsing (UDOP), 448–456, 606–608
- valence, 28–32, 53, 162, 198, 220–221, 240–241, 252, 378
 - change, 618
 - classes, 81
- valence frame, 84
- verb, 17–18, 22, 49, 86, 423, 625
 - final, 40
 - first, 40
 - second, 40
 - AcI, 99
 - auxiliary, 438
 - bivalent, 40
 - ditransitive, 39, 484
 - inherent reflexives, 15, 28
 - intransitive, 39
 - modal, 438, 495
 - monovalent, 149–150
 - particle, 473, 484
 - perception, 99n18
 - serial, 617–618
 - subjectless, 295
 - transitive, 39, 256n12
 - unaccusative, 86, 106, 127, 209, 256n12, 572, 626
 - unergative, 256n12, 626
- verb position, 96, 173, 209, 225–227, 260–292, 385–387, 613, 614

术语索引

- second-, 96
- Verb position, 173
- verb-final language, 95, 210
- verb-second language, 97
- verb-particle, 4
- verbal complex, 108, 582
- Verbmobil*, 158, 159, 239, 330n37
- visual perception, 428
- Vorfeld, 16n8
- Wernicke's area, 432
- wh*-exclamative, 619
- Williams Syndrome, 433
- word, 627
- Word Grammar (WG), 336, 338, 349, 372n34
- word sequence, 6
- \bar{X} theory, 68–72, 86–89, 107, 144–148, 223n3,
409–410, 479, 505, 506, 509, 621
- 一致关系, 353
- 前场, 16n8
- 前置
 - 显性的多项成分前置, 16n8
- 卫星结构, 338
- 句子成分测试, 12n3
- 屈折类型, 54
- 并列, 44n22
- 投射
 - 最大投射, 27
- 投影性, 346, 361
- 抽吸过程 (pied-piping), 374n36
- 提取路径标记 (extraction path marking),
349
- 新布拉格学派, 335
- 测谎仪 (polygraph), 356
- 状语, 32n14
- 空成分, 14n7
- 结, 338
- 行动元, 339
- 论元联接 (linking), 341
- 语义角色, 341
- 语迹
 - 动词, 14n7

Subject index Chinese

ϵ 生成式, 63

λ -演算, 56

\rightarrow , 49

$*$, 59, 70

β -还原, 57

比较级, 20, 22

变换测试, 8–9

宾语

 间接宾语, 35, 36

 直接宾语, 35, 36

并列, 15–16, 61, 73

 并列测试, 10, 15

补足语, 32, 70

场

 后-, 41

 前-, 41

 中-, 41

陈述语气, 18

成分

 非连续成分, 40

词类活用, 25

词类派生, 25

词语序列, 6

代词, 22

 -虚指代词, 24

 虚位成分, 10–11

 虚指代词, 28

代词化测试, 7

单数, 18

等级测试法, 43

递归, 45–46, 61

定子句文法, 75

动词, 17–18, 22, 49

不及物动词, 39

动词二位, 40

动词末位, 40

动词首位, 40

二价动词, 40

及物动词, 39

双及物动词, 39

 自反身代词, 15, 28

动词-助词, 4

短语, 6

短语结构语法, 49–55

俄罗斯套娃, 4, 6

二元, 50

范畴

 -词汇范畴, 25

 -句法范畴, 25

范式, 24

分支

 单叉, 51

 二叉, 50

副词, 17, 20, 21

 -代副词, 23

 -关系副词, 23

复数, 18, 64

感叹词, 20

格, 18, 19, 30, 54

 -宾格, 19

 -属格, 19

 -与格, 19

 -主格, 19

 宾格, 35, 37, 39

 一致关系, 39

 语义格, 37

主格, 32
构成成分, 7, 12
关系小句, 44-45
冠词, 17, 49
管辖, 30
规则到规则的假说, 58
过去式, 18

后场, 41

价, 28-32, 53
将来式, 18
节点, 50
 母节点, 50
 姊妹节点, 50
 子节点, 50
介词, 17, 20, 21, 66-68
经事, 28
句法分析器, 502
句法语义接口, 56

可选的, 32
空成分, 63
控制, 32
框架结构, 40

λ -变换, 57
 λ -抽象, 57
连词, 20, 21, 49
邻接, 51
论元, 28-32

名词, 17, 18, 22, 49, 63
 关系名词, 61
 集合名词, 64
命令语气, 18

NP-分裂, 13

派生, 25

歧义
 假歧义, 66, 72
祈使句, 32
前场, 41
前置, 9, 12-15
情态, 18
屈折

-范式, 18
屈折变化, 18-20, 24
屈折类型, 19, 54

人称, 18, 52

上下文无关语法, 501
上下文有关语法, 501
生成语法, 50
施事, 28
时态, 18, 22
受事, 28
受益者, 28
属格, 35
数, 18, 19, 52, 54
说明语, 28-32, 59-61, 70
 中心语-说明语, 70

拓扑, 40-46
叹词, 21
替换测试, 7
统制, 50
 直接统制, 50
投射, 27
 特征的投射, 28
 最大投射, 72

谓词逻辑, 28
谓语, 38-39
无限制语法, 501

\bar{X} 理论, 68-72
系词, 41
现在式, 18
限定词, 22, 49
限定语
 作为中心语, 26
形容词, 17-19, 22
 描述性形容词, 38
 谓形容词, 38

形式化, 5-6
行动元, 28
性, 19, 54
修饰语, 32
虚拟语气, 18
选择, 28

依存合一语法, 501

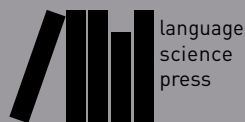
疑问句, 44-45
移位
 移位变换, 12
移位测试, 8-9
语义角色, 28

正则语言, 501
指称, 10, 11
指定语, 70, 502
中场, 41
中心语, 26-28
中心语特征, 28
主谓一致, 5, 32, 52-54
主语, 32-35
助词, 20
状语, 36-37
最短移位限制, 502
最高级, 20

Did you like this book?

This book was brought to you for free

Please help us in providing free access to linguistic research worldwide. Visit <http://www.langsci-press.org/donate> to provide financial support or register as a community proofreader or typesetter at <http://www.langsci-press.org/register>.



语法理论

This book introduces formal grammar theories that play a role in current linguistic theorizing (Phrase Structure Grammar, Transformational Grammar/Government & Binding, Generalized Phrase Structure Grammar, Lexical Functional Grammar, Categorical Grammar, Head-Driven Phrase Structure Grammar, Construction Grammar, Tree Adjoining Grammar). The key assumptions are explained and it is shown how the respective theory treats arguments and adjuncts, the active/passive alternation, local reorderings, verb placement, and fronting of constituents over long distances. The analyses are explained with German as the object language.

The second part of the book compares these approaches with respect to their predictions regarding language acquisition and psycholinguistic plausibility. The nativism hypothesis, which assumes that humans possess genetically determined innate language-specific knowledge, is critically examined and alternative models of language acquisition are discussed. The second part then addresses controversial issues of current theory building such as the question of flat or binary branching structures being more appropriate, the question whether constructions should be treated on the phrasal or the lexical level, and the question whether abstract, non-visible entities should play a role in syntactic analyses. It is shown that the analyses suggested in the respective frameworks are often translatable into each other. The book closes with a chapter showing how properties common to all languages or to certain classes of languages can be captured.

“With this critical yet fair reflection on various grammatical theories, Müller fills what has been a major gap in the literature.” Karen Lehmann, *Zeitschrift für Rezensionen zur germanistischen Sprachwissenschaft*, 2012

“Stefan Müller’s recent introductory textbook, *Grammatiktheorie*, is an astonishingly comprehensive and insightful survey of the present state of syntactic theory for beginning students.” Wolfgang Sternefeld und Frank Richter, *Zeitschrift für Sprachwissenschaft*, 2012

“This is the kind of work that has been sought after for a while. [...] The impartial and objective discussion offered by the author is particularly refreshing.” Werner Abraham, *Germanistik*, 2012

