

目 录

第 1 章 绪论	1
1.1 两个重要的语言学思想	2
1.1.1 Firth 论受限语言	2
1.1.2 Harris 论子语言	4
1.2 来自应用语言学的分析模型	6
1.2.1 学术篇章组织: Swales 的体裁分析模型	6
1.2.2 元话语意义: Hyland 的分析模型	7
1.2.3 评价意义: Hunston 的事态、价值与关联模型	9
1.3 语料库语言学的理念与分析技术	10
1.3.1 成语原则与扩展意义单位	10
1.3.2 型式语法的分析技术	11
1.3.3 短语序列的分析技术	13
1.4 本书的研究方法	14
1.5 本书结构	15
第 2 章 学术话语研究的局部语法路径	17
2.1 引言	17
2.2 局部语法: 背景及概念表述	18
2.3 局部语法: 研究方法	20
2.4 “例举局部语法” 构建	23
2.5 结论	32
第 3 章 局部语法对比研究	33
3.1 引言	33
3.2 研究方法	34
3.2.1 语料库	34
3.2.2 概念界定	35
3.2.3 工作程序	35
3.3 结果分析	39

IV 学术话语探索：局部语法、话语行为与知识构建

3.3.1	评价型式频数分布对比	39
3.3.2	局部语法型式特征与频数分布对比	40
3.4	讨论	43
3.5	结论	45
第 4 章	概指名词共选及局部话语行为	47
4.1	引言	47
4.2	研究方法	49
4.2.1	语料库	49
4.2.2	工作步骤	50
4.3	结果	50
4.3.1	数据分布	50
4.3.2	语义组分类	51
4.4	局部话语行为	53
4.4.1	STUDY 组	53
4.4.2	RESULT 组	56
4.4.3	METHOD 组	59
4.4.4	EFFECT 组	63
4.4.5	CIRCUMSTANCE 组	65
4.5	讨论	66
4.6	结论	67
第 5 章	概指名词共选及双重功能	69
5.1	引言	69
5.1.1	概指名词的评价功能与篇章组构功能	69
5.1.2	语篇构建的小句关系路径与多种语篇型式	70
5.2	研究方法	72
5.3	数据与分析	73
5.3.1	概指名词语篇型式分布	74
5.3.2	“情景－评价－依据”型式	76
5.3.3	“概述－具体”型式	77
5.3.4	“问题－解决”型式	79
5.4	结论与启示	81

第 6 章 话语行为与话语策略.....	85
6.1 引言.....	85
6.2 研究设计.....	86
6.2.1 研究对象和问题.....	86
6.2.2 研究工具和方法.....	87
6.3 数据分布特征.....	88
6.4 词语序列的局部功能特征.....	90
6.4.1 数值 / 图表语义组.....	90
6.4.2 研究 / 发现语义组.....	92
6.4.3 模型 / 理论语义组.....	95
6.4.4 人称代词组.....	97
6.4.5 其他组.....	99
6.5 讨论.....	101
6.6 结论.....	103
第 7 章 评价型式与意义方式.....	105
7.1 引言.....	105
7.1.1 评价语言研究.....	105
7.1.2 it 评价型式研究.....	106
7.2 研究方法.....	107
7.2.1 研究问题.....	107
7.2.2 分析概念与框架.....	107
7.2.3 语料库与研究步骤.....	108
7.3 it 评价型式的数据分布.....	108
7.4 话语意义分析.....	110
7.4.1 it v-link ADJ to.....	110
7.4.2 it v-link ADJ that.....	112
7.4.3 it v-link ADJ for n to 和 it v-link ADJ wh-.....	115
7.5 讨论.....	116
7.6 结论.....	118
第 8 章 评价型式与知识构建.....	119
8.1 引言.....	119
8.2 Becher (1989) 学科分类框架.....	121

8.3	研究方法	122
8.3.1	研究问题	122
8.3.2	分析工具与框架	122
8.3.3	语料库与研究步骤	124
8.4	数据结果	124
8.4.1	it 评价型式与学科的对对应关系	124
8.4.2	it 评价型式的跨学科对比	128
8.5	讨论	136
8.6	结论	138
第 9 章	事态意义与学术知识构建	139
9.1	引言	139
9.1.1	事态、价值与关联模型	139
9.1.2	事态意义、事态名词与知识构建	141
9.2	研究方法	143
9.2.1	研究问题	143
9.2.2	分析工具与框架	143
9.2.3	语料库与研究步骤	144
9.3	研究结果	145
9.3.1	N_{status} that 型式的频数分布	145
9.3.2	N_{status} that 型式的意义与知识构建功能	146
9.3.3	事态意义的跨学科对比分析	153
9.4	讨论	157
9.5	结论	158
第 10 章	语义序列与知识构建	159
10.1	引言	159
10.2	研究设计	160
10.2.1	研究对象和问题	160
10.2.2	研究工具和方法	161
10.3	数据分布与语义序列分析框架	162
10.3.1	基本数据分布	162
10.3.2	语义序列分析框架	164
10.4	语义序列的学科分布特征	168

10.5 结论	171
第 11 章 结论与建议	175
11.1 主要的思想与方法特征	175
11.2 主要的研究发现及观点	176
11.3 讨论	178
11.4 未来发展.....	183
参考文献.....	185
附录 1 Beijing CARE 语料库信息统计.....	201
附录 2 英汉概念索引	203

第1章 绪论

过去30年见证了学术话语研究持久不衰的勃勃生气与纵深发展。毫不夸张地说，学术话语研究已经成为话语分析乃至整个应用语言学最为兴盛的探索领域之一。多维向和多路径下的研究成果展现出学术话语的多面和多重特征。同时，特定研究路径又将学术话语研究带向特定的纵深，不同程度地触及话语的构成与运作机制。概而言之，林林总总的研究相得益彰，使我们得以从多视域、多层面审视与理解学术话语，并催生新的认知，开辟新的疆域，构建新的知识。

正是在这种学术语境下，怀着这样的意图，本书尝试从另一侧面，依循我们特有的路径和工作方法，探索学术话语的构成与运作。我们聚焦学术话语的局部语法（local grammar）、话语行为（discourse act）、话语策略（discourse strategy）与知识构建（knowledge construction）功能这几个尚未被研究社团问津或鲜有系统探讨的话题。我们试图在本书所采取的理论路径和分析框架下，通过对相关词语实体的结构、意义、功能共选的语料库探究，呈现多个层面上的细微粒度话语事实，提出新解，贡献新的分析技术和新的知识。

本章系统阐述全书所采用的语言学思想、分析模型以及语料库分析技术，并揭示全书结构及内容特征。

1.1 两个重要的语言学思想

学术话语研究的渊源可追溯至英国语言学的重要奠基人 J. R. Firth 的受限语言说和美国结构主义的重要旗手 Zellig Harris 的子语言说。这两个理论构成了学术话语研究原初的底层思想，也是我们在本书中采取的基本理论参照。

1.1.1 Firth 论受限语言

多年来对 Firth 论著的研读使我们得以重新发现受限语言说的重要价值。Firth (1956a, 1956b, 1956c) 相继提出并阐述了受限语言说。他将受限语言界定为“主要语言中界定良好的有限类别或形式”，具有“清晰界定的材料”“专门化的词汇、语法和风格”“适宜建立权宜性结构和系统，描写其微观语法和微观词汇”等特征。他的具体论述可见于下面的英文表述(划线部分)：

(i) Descriptive linguistics of the structural kind is at its best when applied to a well defined limited type or form of a major language, let us say of English. (Firth 1956b: 98)

(ii) Restricted languages function in any form of speech or writing with specialized vocabulary, grammar and style. (Firth 1956b: 112)

(iii) The material is clearly defined, the linguist knows what is on his agenda and the field of application is sufficiently circumscribed for him to set up ad hoc structures and systems... A restricted language can be said to have a micro-grammar and a micro-glossary. (Firth 1956b: 106)

在这些论述中，“有限的类别或形式”“专门化的词汇、语法和风格”“微观语法和微观词汇”等说法均突显受限语言不同于普通语言的显著特征。的确，受限语言意味着特殊性很强的一类语言使用。Firth 认为，它涵盖了描写语言学之描写所用语言 (language of description)、被描写的语言 (language under description) 和翻译语言 (language of translation) 等三个方面内容。描写所用语言指的是特定专门领域常规性使用的公式、术语、

符号、语音、词汇以及语法形式，等等。它既包括我们今天常说的元语言 (meta-language)、元符号 (meta-symbol)、技术术语等，又涉及具有特定或专门化意义的词汇、语法特征。被描写的语言指专门领域里的文本材料。翻译语言则是一种宽泛的说法，既包括一般意义上翻译活动涉及的源语言 (source language) 和目的语 (target language)，也涵盖同一语言内所操作的转换或转述语言 (Firth 1956b: 112)。

然而，受限语言并不仅指特殊的语言，它更意味着一条语言描写路径。受限语言说是Firth语境论思想的扩展和延伸。从其语境论的三要素来看，任何语言事件都是受限的，包括受限的事件参与者及其话语行为和话语行为、受限的环境因素和受限的行为后效。那么，重要的是用受限语言的视角和方法来观察语言，用受限语言路径在不同层面上研究语言，如词汇层面、语法层面、搭配层面以及语篇层面。Firth进行的受限语言研究包括文学、翻译、语言教学、科学、技术、政治等领域（参见Firth 1956a, 1956b, 1956c, 1957）。他甚至提出，应采用受限语言路径进行语言教学。例如，为促进欧洲统一而举行的欧洲各国专业人员的英语学习课程最好采用受限语言路径，而非所谓的普通语言路径 (Firth 1956b: 106)。第二次世界大战中，Firth受命为英国皇家空军举办飞行员日语会话课程，成效显著。毋庸置疑，他采用的是受限语言路径。

受限语言概念后来发生了重要演进。在1956年之后的一系列论文中¹，Firth将这个概念扩展，包括了语言变体研究。比如，作为重要世界语言的英语和法语，在不同时空和民族文化场景下的使用会发生重要变化；亚洲、非洲、美洲和欧洲等地使用的英语或法语因不同的使用者、地理环境、文化环境都会发生内部结构变化，等等。不过，这些在不同国际会议上宣读的论文直到1968年才正式付印出版 (Firth 1956a, 1956b, 1956c)。而Halliday等人很可能并不了解这些发展，故在1964年提出了语类说 (register) (Halliday et al. 1964)。其实，语类说所述的域 (field)、旨

1 1956年后，Firth论及受限语言的文章有《语言学与翻译》《描写语言学与英语研究》和《语法的新路径》。这些文章分别在伦敦、柏林等地的国际会议上宣读，于1968年正式发表于Frank R. Palmer编辑的《弗斯论文集 (1952—59)》。

(tenor)、式 (mode) 都已囊括在受限语言说中。

受限语言说给我们提供了观察和审视学术话语的独特视角和方法。其中，“受限”二字至为关键。它意味着学术话语的构成及运作都受制于一系列关键的因素，包括其受限或特有的交际目的、专业化的话语事件参与者及其交流内容、独特的意义与功能方式、有限范围的词汇与语法特征，等等。这些都是本书重要的思想之源。在受限语言说的指引下，本书探讨学术话语的局部语法、学术话语行为的局部功能性质以及独特的话语策略等问题。

1.1.2 Harris 论子语言

另一个给了本书写作以重要启示的是美国后结构主义代表人物 Zellig Harris 的子语言说。Harris 所谓的子语言，实际上专指科学或专业文本的语言。由于科学文本具有独特的专业化内容，其语法结构、词汇和意义均显示有别于普通文本的语言特征，因而务须专门研究。根据 Harris (1968, 1988)，子语言的主要界定特征包括 (1) 形成封闭子集的句子；(2) 或多或少封闭的专业内容；(3) 有限范围的词汇；(4) 语法结构横切于整体语言语法而非包括其内。见下述（划线部分）：

(i) Certain proper subsets of the sentences of a language may be closed under some or all of the operations defined in the language, and thus constitute a sub-language of it. (Harris 1968: 152)

(ii) ...composed of sentences which deal with a more or less closed subject matter – one in which a limited vocabulary is used and in which the occurrence of other words is rare. (Harris 1988: 278)

(iii) ...the sub-language grammar contains rules which the language violates and the language grammar contains rules which the sub-language never meets. It follows that while the sentences of such science object-languages are included in the language as a whole, the grammar of these sub-languages intersects (rather than is included in) the grammar of the language as a whole. (Harris 1968: 155)

在上述论述中，Harris认为是语言的某些运作实践导致了封闭性子集句子的出现。而这些封闭性子集句子专事封闭性的专业内容，往往使用有限的词汇，而极少使用别的词汇。尤为重要的是，Harris独具慧眼，敏锐洞察并尖锐指出了子语言语法与普通语言语法的关系：子语言的一些语法规则常被普通语言违反，而普通语言的一些语法规则从不为子语言所用。那么，尽管子语言的句子都包含于作为整体的普通语言，但子语言语法却是横切于整体语言语法，而非包含其内。这个“子语言语法横切于整体语言语法”的论断一语中的，直击当时语言研究重普通语言、忽视专业语言的偏颇要害，揭示了后者研究的必要性与合法性。

由上述可知，Harris关注的是语言整体中存在的一部分专门表达特殊专业内容的一套特殊句子，它们的语法特征有别于普通语言，应予专门表述，称之为子语言语法。实际上，所谓子语言，大体上就是今天学界所谓“科学语言”或“科技语言”。Harris本人所进行的子语言研究涉及生物医学、免疫学等领域，采用了一套严谨的科学概念和方法，如分布数据(distributional data)、依存关系(dependency relations)等等(Harris 1968: 152-155, 1988: 278; 见卫乃兴 2017: 10)。

如果说Firth的受限语言说更多地意味着一种语言研究路径，Harris的子语言说则专指专业性强的科技语言及其研究。前者的涵盖范围较后者要更加广泛。但子语言说突显了专业文本或科技文本语言的独特性及其研究价值。“子语言语法”的概念尤其具有开辟新研究领域的引领作用，给予其后应用语言学开展的科技英语研究重要启示和指引。也正是子语言说的启示与引领，计算语言学开始了局部语法研究。

本书探讨的学术话语涉及大量的子语言运作。我们研究的局部语法本质上是子语言语法的体现，所研究的话语行为本质上是一种受限语言或子语言环境下的局部话语功能，其话语策略也具有局部性的特征。尤其是局部语法的构建，包括其结构范畴与功能范畴的建立，思想与方法启迪都来自Harris的子语言说。

理论语言学之外，应用语言学给我们提供重要的工作模型。本书采用的模型有Swales的体裁分析模型、Hyland的元话语模型和Hunston的事态、价值与关联模型。

1.2 来自应用语言学的分析模型

Swales (1990) 的体裁分析模型、Hyland (2005a; Hyland & Tse 2004) 的元话语模型与 Hunston (1989) 的事态、价值与关联模型被广泛应用于学术话语分析。但三个模型各有焦点：体裁分析模型旨在探讨学术文本的篇章组织模式，元话语模型聚焦于人际或态度意义，而事态、价值与关联模型关注命题信息的认识意义 (epistemic meaning) 作用。

1.2.1 学术篇章组织：Swales 的体裁分析模型

Swales 基于研究论文，构建了解构学术信息配置方式的体裁分析 (genre analysis) 模型。这种信息配置方式已经约定俗成，成为国际话语社团共同认可的话语方式。遵循社团公认或广泛接受的信息配置方式和话语方式，是合格学术论文的门槛要求，也是合格社团成员必须接受的基本训练。随着 ESP 教学，尤其是学术论文研究与教学在过去多年中渐成显学，学术篇章组织及其信息配置方式愈益重要，Swales (1990) 的体裁分析模型也一直发挥着持久不衰的广泛影响。

Swales (1990, 2004) 将体裁界定为“具体话语社团实施的、由交际目的驱动的结构化交际事件”。他建构了学术论文信息配置方式的层级模型。在最高层级上，论文呈现 IMRD (Introduction-Method-Results-Discussion) 的结构模式。并基于论文的引言，抽象出 CARS (Creating a Research Space) 模型。CARS 模型又体现为具体的语步-语阶模型。语步是联系交际目的与命题内容的单位，语阶则是语步包含的更小的话语单位。位于第三层级上的是语步和语阶的词语实现。

以 CARS 模型为例，它由 3 个语步及多个必要或可选语阶组成。语步 1 “建立研究领地” (Establishing a territory) 通过“声称核心价值” (Claiming centrality)、 “概述研究话题” (Making topic generalization)、 “评述研究文献” (Reviewing previous research) 等语阶表述研究话题的重要价值和意义。语步 2 “建立研究空间” (Establishing a niche) 通过“揭示研究空白” (Indicating a gap)、 “提出研究问题” (Question-raising)、 “延续研究传统” (Continuing a tradition) 等语阶表述所选领域的尚存问题。语步 3 “占领研究空间” (Occupying the niche) 由“宣布本文研究” (Announcing

present research)、“宣布主要发现”(Announcing principal findings)、“揭示论文结构”(Indicating research article structure)等语阶构成。

每个语步都有其典型的词语序列实现手段。例如,“声称核心价值”常常涉及 a growing awareness has been given to、numerous studies have focused on、...has been investigated most frequently等词语序列的使用;“评述研究文献”通常由 previous studies show that、past work suggests that、a recent study has shown that等高频序列实现;“宣布本文研究”通常由 this paper describes、the purpose of this paper is to、the present study aims to、in this paper we will discuss等序列实现。

Swales的语步-语阶模型勾画出交际目的、话语行为和词语实现之间的层级制约关系,揭示出篇章信息配置的内在逻辑。自问世以来,该模型被广泛应用于ESP教学研究,尤其是学术论文的篇章研究,显示出极强的描述力和解释力。ESP教学研究在过去多年不断在广度和深度上的拓进,极大程度上是得益于这个模型的启迪与引领。

本书基于学术论文语料库数据,探究学术话语的构建,Swales的体裁分析模型具有直接的指引作用。这个模型所阐述的从篇章结构到语步-语阶、再到词语实现的各个层级制约,形成我们探讨话语行为、话语策略以及知识构建方式时隐性或显性的分析机制。全书通过考察大量词语共现型式和高频词语序列,讨论话语行为与知识构建方式,是Swales体裁分析模型的拓展应用。

1.2.2 元话语意义:Hyland的分析模型

“元话语”(metadiscourse)的概念有广义和狭义之分。狭义的元话语是“关于话语的话语,与主题的关联不大”(Williams 1981: 211)。也就是说,它主要实施篇章组织功能,并不直接相关于思想、论辩和事实表达。广义的元话语既包括篇章组织功能,又包括作者与读者的互动。过去多年的学术话语研究使人们认识到,科学知识构建与信息交流不仅需要客观命题信息的传递,还涉及大量的主观态度表达,比如评述前人研究,解释研究结果,评判各种命题,等等。这些主观态度在林林总总的文献中被称为“评价”(evaluation, Thompson & Hunston 2000; appraisal, Martin & White

2005)、“态度意义”(attitudinal meaning, Halliday 1994)、“立场”(stance, Biber & Finegan 1988; Hyland 2005a)等等。其中, Hyland提出的广义元话语概念及其分析模型被广泛使用, 影响深远。

Hyland和Tse(2004)将元话语分为两类: 引导型元话语(interactive metadiscourse)与互动型元话语(interactional metadiscourse)。前者为传统意义上的篇章信息组织语。作者利用语言资源引导读者理解命题信息, 涉及过渡语、框架标记语、文内标记语等词语手段的使用; 后者指态度意义表达, 包括模糊语、加强语、介入标记语等等。之后, Hyland(2005a)进一步提出了互动型元话语新模型。新模型包括“立场”(stance)与“介入”(engagement), 皆属于交互型语言的使用。“立场”指表达作者声音、态度的修辞手段, 包括模糊语、强化语、态度表达语、自我提及等类别; “介入”是与读者互动的词语资源, 涵盖读者代词、指令语、疑问句、共享知识、个人旁白等手段。

如果说Swales(1990)的体裁分析模型是篇章组织模型, Hyland的元话语模型则是关于态度意义的分析模型, 突显人际互动的特征, 从另一个方面探讨学术话语的篇章信息推进、主观态度表达、乃至话语策略使用。

不过, 上述元话语模型有囊括一切之嫌。用Hyland(1998b: 648)的话说, 这些语言现象看上去像是各种互不相干的语用符号的大杂烩。它既涉及Halliday(1985)所谓的概念意义(ideational meaning)和篇章意义(textual meaning), 更包括了人际意义(interpersonal meaning)。一切都囊括于元话语的名下。而从Sinclair(1991, 2004; Sinclair & Mauranen 2006)的观点来看, 话语以线性方式推进和展开, 将一部分语言手段称之为“元话语”意味着它们高出其他语言手段一个层级, 并不特别妥帖。从研究方法来看, 元话语分析单位主要是单个词和固定短语, 很少关注搭配、型式、构式以及其他较大的短语单位。根据Firth(1957)、Sinclair(1991, 1996)等人的思想, 单个的词并不表达太多的意义; 使用于不同的搭配或型式之中, 就参与不同意义的表达; 而出现在不同搭配或型式中的同一单词, 则可视作不同的词。

本书探讨的局部语法、话语行为、话语策略以及学术话语知识构建依赖对大量复现的搭配、词语型式以及词语序列的详尽分析。Hyland元话语

模型的分析技术并不特别适用。但元话语模型强调的作者-读者互动特征, 阐述的人际意义资源, 以及对具体态度意义的认定, 给我们以重要启示。

本书各章的讨论会经常性地与Hyland元话语思想与观点对接。

1.2.3 评价意义: Hunston 的事态、价值与关联模型

Hunston (1989) 基于对学术文本评价资源的分析, 构建了事态 (status)、价值 (value) 与关联 (relevance) 模型。根据该模型, 评价包含三个功能, 而评价研究也由三步构成, 分别是识别评价对象 (事态)、对评价对象赋值 (价值) 和识别信息的重要性 (关联)。具体而言, 事态是对命题或事物确定性程度的表述, 揭示命题或事件与客观世界的对应程度 (correspondence)。作者通过赋予命题一种认知状态而将其具体化。例如, 命题的事态可以是事实 (fact)、阐释 (interpretation)、评估 (assessment)、假设 (hypothesis) 等等。Hunston (1994, 2010) 认为, 所有小句均具有明确或隐含的事态信息, 以至于所有命题信息都依据它们与外部世界的对应程度而被评价。事态包括三个维度: 作者的表述方式、信息来源以及确定性程度。表达事态的语言资源有: 1) 动词、名词或形容词后接 that 引导的补语从句 (如 It has been suggested that...、The discovery that...、It is probable that...); 2) 包含上述动词、名词或形容词的其他语法结构 (如 As was suggested...、This discovery...、The probable cause); 3) 副词或状语 (如 probably、supposedly、in fact); 4) 情态助动词 (如 may、might、could); 5) 表示命题来源的结构 (如 According to Smith...、There is a general consensus that...、As is well known...) (参见 Hunston 2011: 28)。价值是对命题的价值或品质进行评价。基本的评价参数是“好”或“坏”, 具体价值分类与话语类型有关。比如, 学术话语中的价值可分为两类: 理论和实践的一致性和信息的效用性。价值可以通过显性的语言手段实现, 但大多数情况下, 价值是隐性的, 没有语言标记。另外, 价值评判与社团目的和评价对象的事态联系密切; 不同的事态会产生不同的价值标准。例如, 作为事实的研究结果可以被评价为“可靠”或“不可靠”, 但作者提出的假设无法依照可靠性这一维度进行评价。评价的第三个功能涉及篇章层面的信息组织, 是对文本的相关性或重要性进行标识, 借助关联标记语实现。关联标记语 (如

This...、These studies...、The data of Figure 3...、Therefore) 可以通过前瞻或回溯文本，表达后续或之前文本的重要性，具有组构篇章的作用。

Hunston的事态、价值与关联模型提供了学术话语评价意义研究的另一视角和分析框架。不同于Hyland的元话语模型，该模型从命题信息与客观世界的对应以及认识价值 (epistemic value) 的角度探索不同层级的评价意义，同时也涉及组篇功能的探讨。它尤其有助于隐性评价意义的探讨，具有较强的描述力与解释力。另外，事态、价值与关联的实现并不局限于词汇、短语或单个语句，也可能跨越多个句子，是多种意义的累加。

本书对学术论文评价意义的探索吸收了事态、价值与关联等概念及其视角。

1.3 语料库语言学的理念与分析技术

语料库语言学为学术话语研究提供了独特的实证分析技术。语料库语言学秉持词汇与语法一体、结构与意义一体的基本立场，兼顾语言的组合关系与聚合关系，基于大规模真实、验证数据所作的探索，极大地推进了学术话语研究的发展。本书依循我们多年来一以贯之的研究方法，基于语料库数据进行所有问题探索。Sinclair (1991, 2004) 的成语原则和扩展意义单位模型、Hunston等人的型式语法 (Hunston & Francis 2000) 构成全书最为核心的分析技术。

1.3.1 成语原则与扩展意义单位

Sinclair (1991: 110) 的成语原则 (idiom principle) 认为：“语言使用者拥有大量的半预制词组，尽管这些词组可以被分成不同的部分，但却构成了单一的选择单位”。这就是说，交际过程中的语言形式选择并非根据语法规则逐词进行，而是一次选择两个或多个词形以实现意义。与成语原则相对的是开放选择原则 (open choice principle)。后者赋予语法制约统领一切的作用，所有的选择都依据语法结构的合法性做出，是一个空位接续一个空位的逐词选择 (slot by slot choice)。然而，Sinclair认为，成语原则是处于主导地位的模型，可以解释文本的绝大部分构成，开放选择原则是个次要的模型。后来，Sinclair将两个对立的模型表述为两个选择趋向，即短语趋向 (phraseological

tendency) 和术语趋向 (terminological tendency) (Sinclair 1996, 2004)。

成语原则及短语趋向来后来再被概括为扩展意义单位 (extended units of meaning) 模型 (Sinclair 1996)。扩展意义单位模型有5个构件: 节点词 (the node word)、搭配 (collocation)、类联接 (colligation)、语义趋向 (semantic preference) 和语义韵 (semantic prosody)。这实际上谈的是节点词在4个层面上的共选。搭配体现的是节点词与词汇的共选; 类联接是节点词与语法结构的共选; 语义趋向是节点词与语义特征的共选; 语义韵体现的是节点词与语用功能或态度意义的共选。

扩展意义单位模型将之前的成语原则、短语趋向、搭配和语义韵研究集成一个整体的语料库短语学模型, 是4个层面上的共选型式 (co-selection patterns) 相互制约, 共同生成交际中的意义单位。其中, 节点词和语义韵是扩展意义单位的必选项, 而语义韵处于核心地位, 是最高层级的制约。正如Sinclair后来所言, 意义的移变主要由于共选型式的改变, 而非单个词的改变 (Sinclair 2010)。

扩展意义单位模型是语料库语言学关于意义表达和文本构成的最重要分析模型。它颠覆了传统语言理论视单个词为自足意义单位的认知, 将词汇与语法、句法与语义、语义与语用这些长期以来被分割处理的问题综合一体。这个模型是对Firth建议的意义的韵律路径 (prosodic approach to meaning)、语法的韵律路径 (prosodic approach to grammar) (见Firth 1968各章) 的承继与发展。

扩展意义单位模型是本书最重要的分析机制和方法。在全书的研究中, 我们突破单个词的边界, 将意义载体扩展到以节点词为中心而产生的高频词语序列或型式, 构建局部语法, 分析话语行为, 概括话语策略, 探究知识构建。

1.3.2 型式语法的分析技术

Hunston和Francis (2000) 提出的型式语法 (pattern grammar) 是一种新型的词汇语法 (lexical grammar) 描写方法。型式由具体的词形、与其频繁共现的其他词或结构以及与之相联结的意义共同界定 (Hunston & Francis 2000)。型式语法实质上描述的是词的用法及意义方式。型式表征规避了传统语言描述常用的句法范畴和结构标签, 如主语、谓语、宾语、状语、及