

1.1 引言

基于电子语料库开展的语言研究，经过数十年的发展，如今已渐成主流，并且还呈现出语料库研究与邻近学科广泛融合的局面。在国际范围内就成果数量和影响而论，语料库在与话语研究的结合方面，成绩最为突出。图1.1所示为2006—2016年的十年间在国际期刊发表的语料库研究热点。其中discourse居于画面正中，其所在的知识节点圆面积数倍于其他热点关键词，这反映出话语研究在语料库研究中的影响和地位。其他基于语料库的研究还包括语法(grammar)研究、语法化(grammaticalization)研究和语言习得(language acquisition)研究。这些研究的数据和方法论基础都是频率(frequency)。

若将观察的镜头推近，我们可以深入了解更为丰富的语料库研究世界。图1.2的中观研究热点显示，除了图1.1中的宏观热点外，还可看到在过去十年间，约有20项其他热点研究话题。例如，conversation、speech、discourse marker、identity、politeness、genre、variation、conversation analysis和gender等(详见表1.1)。这些热点词明白无误地将语料库研究指向了对真实会话(conversation、speech、conversation analysis)的探究，具体包括会话中的话语标记(discourse marker)研究、

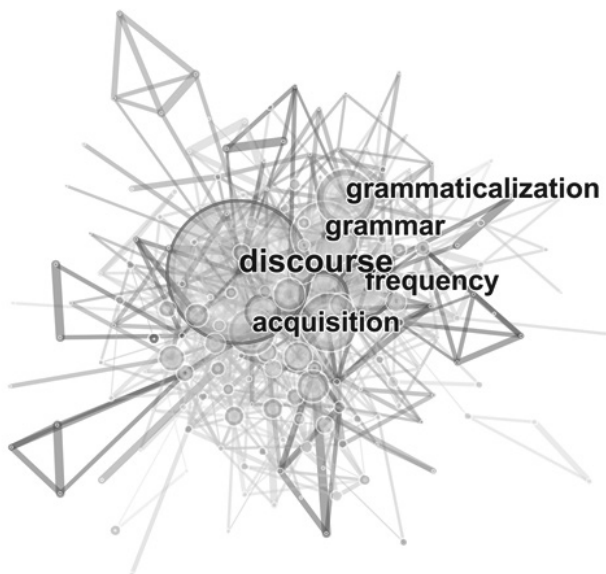


图 1.1 2006—2016年间发表的语料库相关英文论文热点关键词（宏观）

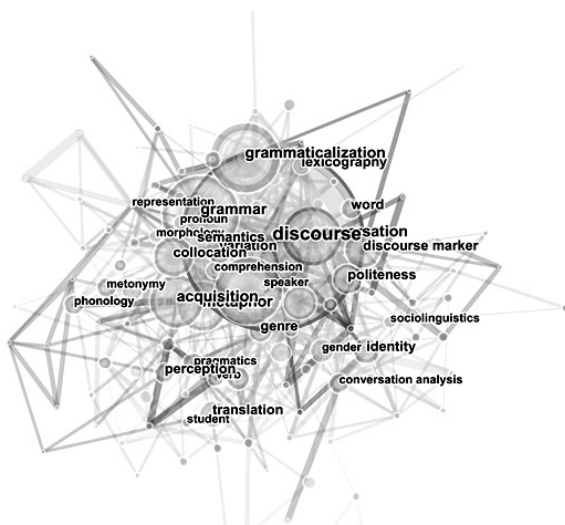


图 1.2 2006—2016年间发表的语料库相关英文论文热点关键词（中观）

身份认同和建构 (identity) 研究、礼貌现象 (politeness) 研究以及性别 (gender) 和体裁 (genre) 等方面的变异 (variation) 研究。上述这些热点词构成了英文语料库研究的主体, 其数量明显高于隐喻 (metaphor)、句法 (syntax)、翻译 (translation) 和语义 (semantics) 等方面的研究。类似的知识热点分析, 还可以深入到更为微观的层面。

表 1.1 2006—2016年间发表的语料库相关英文论文热点关键词

频数	热点关键词	频数	热点关键词
177	discourse	43	translation
85	grammar	43	politeness
83	grammaticalization	42	semantics
76	frequency	39	genre
72	acquisition	38	variation
65	conversation	38	word
60	metaphor	34	lexicography
51	speech	34	perception
51	collocation	33	conversation analysis
49	discourse marker	33	gender
46	syntax	32	pragmatics
44	identity	31	phonology

我们对这些研究热点稍作分析便不难发现, 英文发表的语料库研究成果集中在传统意义上的话语分析、社会语言学和语用学等语言学分支学科。事实上, 当前“话语”这一概念的内涵和外延都已拓展至原本属于社会语言学和语用学的领域。这几个方面的研究已经很难截然分开。对于

“话语”概念的界定，后文会专门讨论。

以上通过文献计量方法观察到的语料库研究热点，与我们日常所接触到的相关文献和学术走势较为吻合。比如，在语料库领域中的代表性期刊*International Journal of Corpus Linguistics*和*Corpora*上总能见到较多的话语研究论文，而在话语语用研究期刊*Discourse Studies*和*Journal of Pragmatics*上，也有相当比重的研究采用的是语料库方法。

语料库领域新近(拟)创办的三本学术期刊，*Journal of Corpora and Discourse Studies*(《语料库与话语研究学刊》，2018年出版创刊号，英国卡迪夫大学出版社)、*Register Studies*(《语域研究》，将于2019年出版创刊号，John Benjamins出版社)、*Corpus Pragmatics*(《语料库语用学》，2017年创刊，Springer出版社)也都与话语研究直接相关。*Journal of Corpora and Discourse Studies*杂志代表的是英国及欧洲大陆语料库与批判话语研究相结合的传统。*Register Studies*则是美国语料库研究有关语域变异的新阵地。*Corpus Pragmatics*刊载的文章多半属于话语研究范畴。其他一些刊物，如*Applied Linguistics*(《应用语言学》)、*English for Specific Purposes*(《专门用途英语》)、*Journal of English for Academic Purposes*(《学术用途英语学刊》)、*Journal of Language and Politics*(《语言与政治学刊》)也时常刊登语料库话语研究论文。

此外，国际上出版的与语料库研究相关的论著，着眼于话语研究的结果数量也在不断增加。在近十多年的国际语料库学术会议上，话语研究热度不减。比如，2004年美国应用语料库语言学研究会(The American Association for Applied Corpus Linguistics)主办了“第五届北美语料库语言学研讨会”(The Fifth North American Symposium on Corpus Linguistics)，会后出版的名为*Corpus Linguistics Beyond the Word: Corpus Research from Phrase to Discourse*(2007)的论文集，旗帜鲜明地指出：语料库研究需要拓展至话语层面。2008年第29届ICAME(The 29th International Computer Archive of Modern and Medieval English)会议

更是将大会主题明确定为“Corpora: Pragmatics and Discourse”。目前国际上影响最大的“国际语料库语言学大会”(The International Corpus Linguistics Conference),最近几届的大会主旨发言所请专家超半数都是话语语用学者,比如2011年的Susan Hunston(主攻话语评价)和Paul Baker(主攻话语与性别建构);2013年的Karin Aijmer(主攻话语标记)、Michael Hoey(主攻衔接和词汇触发理论)和Ute Römer(主攻学术话语分析及短语学);2015年的Douglas Biber(主攻语域变异)、Michaela Mahlberg(主攻语料库文体学)和Alan Partington(主攻批判话语研究);2017年的Susan Hunston(主攻话语评价)和Dan McIntyre(主攻语料库文体学)。在这一趋势之下,Alan Partington于2016年创办了一个新的语料库会议系列,名为“语料库与话语国际会议”(The Corpora and Discourse International Conference)。首次会议于2016年在意大利锡耶纳举行,会议每两年一次。会议的六大主题为:话语组织(discourse organisations);政治、机构和媒体话语(political, institutional and media/social media texts);社会科学及社会政策议题话语研究(social science and social policy issues);文化及跨文化话题(cultural and cross-cultural topics);翻译研究的话语含义(discourse implications in translation studies);文学文本中的话语效应(discourse effects in literary texts)。

综上所述,话语研究是近年语料库研究中当仁不让的主角。事实上,从更长的历史时期来看,话语研究一直是语料库研究中最重要领域之一(Hoey & Shao 2015: 15; 马晓雷、陈颖芳 2016: 44-50)。语料库与话语研究的结合有其内外因,后文我们会专门探讨。

需要指出的是,我国学者用中文撰写的学术成果,其研究热点和趋势呈现出相当不同的面貌。三十多年来,我国语料库学者并未跟风,而是立足本土研究选题,运用语料库资源和方法,着重解决语言习得和翻译中存在的问题(见图1.3,参阅刘霞等 2014: 75)。

在我国,基于语料库的批判话语研究、学术英语话语研究、教学与教



图 1.3 1984—2016年间 CSSCI 期刊语料库相关论文热点关键词

师话语研究、中介语话语研究、多模态话语研究等，虽然体现出了自己的特色，但在整体数量上并不占优势。

以下我们将分别就语料库研究的核心特征、话语研究的核心特征以及两者融合研究的现状及特色加以阐述。

1.2 语料库研究的核心特征

语料库研究的基本内涵指的是针对大规模口语笔语真实语言材料，通过统计分析手段，在一定的理论视角下，发掘语言在形式和意义上的典型性规律。这种语言研究理念在国内外早已存在，其历史至少在百年以上 (Xu 2015; 许家金 2017a)。它反映的是一种描写语言观，即尊重真实的

语言使用，主张语言的基本规律蕴藏在语言使用之中。当然，在计算机产生之后，如今谈论语料库研究，则专指借助计算机软件，对电子语料库进行分析，并揭示语言运作模式的研究领域。

综合以往文献，我们将语料库研究的核心特征概括为“用”“量”“器”“聚”（许家金 2014: 35-36, 2017a: 52-53）四个方面。“用”即尊重语言事实和关注用法。语料库研究以对语言使用的充分描写为根本。其次，语料库研究在“量”上的要求指的是通过“量化”分析手段研究语言。“量”在统计上，主要通过频率得以体现。语料库研究主张语言是概率性现象（Halliday 1991; Bod *et al.* 2003）。Halliday (1992) 指出概率性既是对语言实例（instance）的描摹，同时也是一种理论构念（construct）。在很大程度上这种构念体现为一种“聚合”或是“共现”关系，我们称之为“聚”。弗斯语言学的“语境论”和社会语言学都是“聚”的理论体现。新弗斯语料库语言学学者们有关共现的研究思路，反映的是词汇、语法范畴、语义和语用等不同语言层面之间的共选关系；而基于语料库的语域变异、批判话语研究，反映的是语言同社会语言学变量之间的共选关系。不论是语言成分之间的关联，还是语言成分与社会变量之间的关系，最终都要通过概率统计以及语料库分析工具来揭示，即“器”。语料库研究的这四项设计特征是我们区分语料库研究和非语料库研究的基本要素。其中，“用”是基础，“量”是关键，“器”是途径，“聚”是语言学理论归宿。如果某项研究全部满足上述四项特征，则可以被认为是典型的语料库研究，可称为“大写的语料库语言学”（big C Corpus Linguistics，即狭义语料库研究）。如果只满足了上述四项特征中的部分特征，我们不妨称之为“小写的语料库语言学”（little c corpus linguistics，即广义的语料库研究）。前者是典型的语料库研究，后者是宽泛意义上的语料库研究。

概而言之，“用”“量”“器”“聚”前三个字涉及语料采集、语言特征的量化统计分析以及语料库工具的运用，它们都与研究方法紧密相

关。可见，语料库研究总体上是“具有很强方法论导向的语言学分支” (a methodologically-oriented branch of linguistics) (Leech 2011: 158)。

语料库研究的四项核心特征，其语言学价值，或者说理论落脚点在于“聚”，即挖掘语言成分之间以及语言与社会变量之间的共选或共聚关系。用语境论的术语来说，“聚”即“语境化”(contextualization)，是对真实语料的逐层分析，尽力还原语言实际发生时的语境含义。具体来说，语料库研究在很多情况下是由字词检索开始，进而通过索引分析(concordancing)获得字词的句级上下文语境(又称“共文”，co-text)，还可以扩展到段落、篇章和整个作品或会话场景。语料库方法还能够观察到语言特征的文本间复现情况，即所谓互文现象。语料文本附带的社会文化语境信息，如性别、说话人年龄、场景等，往往通过元信息(metadata)的形式随同语料库文本一起存储。这种将语料库研究视为“语境化”的观点，在Scott & Tribble(2006: 9)中，以语境范围(contextual scope)的提法得以更条理化地展示，如下所示：

Contextual Scope(语境范围)

SCOPE 1: a few words to left and right(范围一：检索词左右若干词)

SCOPE 2: the whole sentence(范围二：整句)

SCOPE 3: the paragraph(范围三：段落)

SCOPE 4: the story so far(范围四：当前故事片段)

SCOPE 5: the section or chapter(范围五：章节)

SCOPE 6: the whole text(范围六：整个文本)

SCOPE 7: the colony of texts to which this one belongs(范围七：当前文本与所属文本群落)

SCOPE 8: other related texts(范围八：其他相关文本)

SCOPE 9: the context of culture(范围九：文化语境)

EXTRA-LINGUISTIC SCOPE: where you are when you meet the text (非语言范围: 接触该文本时的场景)

Scott & Tribble (2006) 的表述非常详细, 但通常不是每个文本分析都需要经历这些步骤; 再者, 有些语境范围层级可以合并。比如范围七和范围八。我们把上述语境化的分析思路称为“点—线—面语境扩展法”。这种由点及面不断挖掘语境意义的做法, 是基于语料库的话语研究中的常见思路。

语境是话语研究中不可忽视的重要特征。Malinowski (1923: 310-311) 将人类学研究中的语境分为情境语境 (context of situation) 和文化语境 (context of culture)。而在语言分析中, 这一语境二分法还要包括语言单位间的上下文语境、更大的语言单位 (段落与段落、前后篇章、不同文本等) 之间的互文性 (intertextuality) 等。

语料库研究方法和技术擅长词汇短语层面的分析, 但作为话语研究, 词汇检索和分析只是研究的切入点, 最终还应回归到词汇所在的上下文和社会文化语境中。在这个过程中, 对于词汇、短语和一些句法结构的检索结果以及他们在不同文本中的分布情况, 往往体现为量化数据。因此, 通常语料库研究方法被视作是一种典型的量化研究方法。然而, 通过语料库检索得到的大量例证, 仍然需要研究人员通过思辨或相关语言学理论进行定性分析。可见, 完整的语料库研究应能综合运用定量和定性的研究方法。

1.3 话语研究的核心特征

1.3.1 话语及话语研究

话语 (discourse) 这一概念, 何时从一个通用名词发展成语言学术

语，已很难考证。从话语研究发展之初，与之相生相伴的另一英文术语是text。在中文文献中，text被更多地译成“语篇”或“篇章”。部分文献倾向于将text理解为书面语概念，将discourse理解为口语概念；将text理解为语言结构特征，而将discourse理解为语言结构特征的同时，还附带有社会文化属性。text和discourse相对应的中文术语“语篇/篇章”和“话语”，在文献中混用的情况比较普遍。就英文术语而言，时至今日，discourse已逐渐占据优势。text(及其形容词形式textual)虽然时常见到，但多限于表达语篇衔接等语言结构特征。而discourse已成为话语研究领域更为通用的术语。其形容词形式有两种，一个是discoursal，另一个是discursive。前者常与lexical、grammatical和syntactic并用，指的是话语的结构维度；而discursive则更多表示话语的社会维度，常出现在discursive construction(话语建构)这样的短语中。在中文文献中，有人将discourse译为“话语”，有人译为“语篇”。本书统一采用“话语”这一表述。

话语研究在相当长的时间内被称为“话语分析”(discourse analysis)。近二三十年来，随着其跨学科性质越来越突出，话语研究已远远超出一种语言学分析方法的范畴，俨然成为一个独立的语言学分支学科。于是，部分学者，特别是批判话语分析学者，倡导使用“话语研究”(discourse studies)这一概念。本书将采用这一新的学科名称。当然，在谈及旧有文献时，为尊重史实，还会采用话语分析这一表述。在新近的文献中，话语分析和话语研究两者仍然并存，但两者并无高下、优劣之分。

话语研究之所以内涵越来越丰富，外延越来越广阔，原因在于其理论源头众多。因此，如今话语研究的分支和流派众多，各家对话语研究的认识也不尽相同。

文献中一般将Zellig Harris于1952年在*Language*杂志上发表的“Discourse analysis”一文作为现代话语研究的起点。在此之前的西方经

典修辞学和我国古代辞章学可视作古典意义上的话语研究。Harris是美国结构主义语言学的集大成者。他在Bloomfield等学者奠定的结构主义描写语言学基础上，努力拓展语言结构描写的层面和维度。但传统的结构主义考察范围仅限于句以下单位。一个完整的理论应当能够对更长单位内的语言现象加以解释。有关“话语分析”的倡议，正是在这一背景下产生的。在Harris(1952)的文章中，他指出话语是超出句子边界的结构单位，可以是一个段落或一本书(Harris 1952: 5)。话语分析能够解决两大问题(Harris 1952: 4)：一是有助于分析句子和句子间的分布组合关系(distributional relations among sentences)；二是可以解决语言和社会情境之间的相关性(correlation between language and social situation)。在Harris的分析示例中，他并未涉及有关社会情境的分析。Harris旨在提出一套描写连续话语中具有相似上下文的对等结构的方法。比如在前文句子中出现的the middle of autumn和后文出现的the end of October，它们的使用语境相似，因此属于对等结构。Harris通过一套形式化的方式，分析话语中出现的此类对等结构，从而描写比句子更大的语言单位。虽然这种话语分析方式跟后来的话语照应和主位推进等概念有相通之处，但仍然具有明显的结构主义特点，是Harris“分布假说”(distributional hypothesis, Harris 1954)在话语层面的体现。总体而言，Harris提出的话语分析方法并未受到太多重视，也没有得到继承，与后来蓬勃发展的话语研究传统也并无直接联系。Harris对于话语研究的贡献在于：他作为当时语言学界的旗帜人物，明确地指出了研究超句语言现象的重要性。

20世纪70年代在英国及欧洲大陆，有学者开始系统开展话语研究。De Beaugrande和Dressler于1981年出版的*Introduction to Text Linguistics* (《篇章语言学引论》)一书集中总结了欧洲篇章语言学的相关理论认识。其中概括了话语的若干核心特征，例如，形式衔接(formal cohesion)、主题连贯(thematic coherence)、意向性(intentionality)、可接受性

(acceptability)、信息性(informativity)、情境性(situationality)以及互文性(intertextuality)。而Dressler(1978)的*Current Trends in Textlinguistics*(《篇章语言学的当前趋势》)一书则涵盖了欧洲篇章语言学传统之外世界范围内话语研究各流派的研究动向,包括符号学(代表学者为Nöth和Petöfi)、会话分析(代表学者为Schegloff)、认知心理研究(代表学者为Van Dijk和Kintsch)、叙事分析(代表学者为Grimes)、文体学(代表学者为Enkvist)、系统功能语言学(代表学者为Hasan)等。20世纪60、70年代蓬勃发展的其他一些学科领域对话语研究的发展也发挥了关键的作用,例如,与言语行为、合作原则和面子等相关的语用学议题,Foucault关于话语和权力的讨论,Fowler等人倡导的批评语言学(critical linguistics),韩礼德关于语域(register)的讨论,Labov倡导的变异社会语言学(variational sociolinguistics)以及Goffman倡导的互动社会语言学(interactional sociolinguistics)等。这些领域更多关注的是话语的社会属性,而20世纪70年代的话语研究重点更多在于语言内部特征,因此Dressler(1978)并没有将Fowler、Foucault、Labov、Goffman等人的研究纳入考察视野。

综上所述,话语研究从结构主义语言学后期开始萌芽,在20世纪70年代博采众多学科之长,蓄势发展,成了如今具有跨学科性质的热门研究领域。然而,如此众多的理论滋养也给话语研究的学科界定带来了挑战。

在相当长时间里,Stubbs(1983: 1)对话语研究的界定是本领域较为通行的认识。他对话语研究的定义如下:

[Discourse analysis]...refer[s] mainly to the linguistic analysis of naturally occurring connected speech or written discourse. Roughly speaking, it refers to attempts to study the organisation of language above the sentence or above the clause, and therefore to study larger linguistic units, such as

conversational exchanges or written texts. It follows that discourse analysis is also concerned with language use in social contexts, and in particular with interaction or dialogue between speakers. (Stubbs 1983: 1)

译文：话语分析主要指的是对自然发生的连续口语和书面语的语言学分析。笼统来说，话语分析旨在研究超越句子或小句单位的语言组织情况，因此它研究的是更大的语言单位，比如会话交流或书面语篇。另外，话语分析关注的是社会语境中的语言使用，特别是说话人之间的互动和对话。

这一定义包含三层核心含义：1) 话语研究的对象是真实交际中的语言；2) 其研究重点是超句语言单位的前后关联机制；3) 话语、个人和社会之间的关系。

在过去二十年左右的时间里，Schiffrin *et al.* (2001: 1) 在 *The Handbook of Discourse Analysis* (《话语分析手册》) 一书的《导言》中对话语研究的概括被更多地引述。Schiffrin 将以往种种话语及话语研究的定义凝练后得到如下认识：话语研究是 1) 有关超句单位的研究；2) 有关语言使用的研究；3) 有关与语言和非语言交际相关联的广泛社会实践行为方面的考察。原文如下：

[The definitions of discourse and discourse analysis], however, fall into the three main categories...: 1) anything beyond the sentence, 2) language use, and 3) a broader range of social practice that includes nonlinguistic and nonspecific instances of language. (Schiffrin *et al.* 2001)

The Handbook of Discourse Analysis (第二版) (Tannen *et al.* 2015: 1) 并未提出新的话语研究定义，而是沿用了第一版对话语研究的三点认识。然而，在内容编排上，*The Handbook of Discourse Analysis* (第二版)

中近半数章节都作了替换，未替换的其他章节内容也都有所更新。新补充的章节大多跟社会实践和意识形态有关。可见近二十年来，增长最快的话语研究领域正是话语的社会维度。这也是当前的话语研究定义不同于Stubbs (1983) 的最主要方面。如今话语研究着眼于社会交际的方方面面，如机构话语、政治话语、(新)媒体话语、法律话语、医疗话语、教育话语和职场话语等。在新增的话语研究选题中，讨论较多的议题是话语中表现出的权力关系和身份认同。

可见，话语研究的领域，在语言之外更多关切的是与语言相关的社会问题。这使得话语研究与相关领域有了更多的交融，比如语用学中的很多议题也都从话语视角加以探究。社会语言学中的性别、身份认同、权势、语域变异以及跨文化交际中的很多现象如今也都大量出现在话语研究文献中。

The Handbook of Discourse Analysis (第二版) 所概括的三类话语研究焦点中，第一类研究的关注点是话语的结构特征，简单地说就是话语涉及的长度和跨度，即话语不限于一词一句，而更强调词句间的连缀协同，从而完成意义的连贯表达。第二类强调话语研究考察的是自然发生、真实交际中的口语和书面语。比如，随着时代的发展，新媒体话语研究也逐渐占有一席之地。值得一提的是，在语言使用这一点上，话语研究和语用学有共同的研究焦点，它们都关注使用中和语境中的言语交际，都关注言语行为、指称、情态、评价等研究选题，所以如今话语研究和语用学两者已很难截然分开。第三类研究取向在话语研究中相对较新，它强调话语在更广的社会文化情境中解读才充分完整。另一方面，一些学者也主张通过话语来探究社会文化的现状及其变化发展，甚至寄希望于改变不合理的社会现状，即所谓的“批判话语分析”(critical discourse analysis)。

以下将进一步阐释话语所具备的核心特征。我们将常见的话语研究的切入点，按前述话语研究核心内容分为三方面，即话语的结构特征、话语的语义特征以及话语的社会属性。在一些文献中(如Van Dijk

2017), 也能见到从认知角度讨论话语现象的成果。这种认知观念很大程度上属于社会认知, 主要是基于语言产出考察知识、态度、思维方式或认识观念等。我们倾向于把这类认知现象也归入话语的社会属性。

1.3.2 话语的结构特征

话语的结构特征主要涉及话语成分的组织构成, 其中包括小的语言单位如何组合成更大的话语单位, 大的话语单位如何切分成小的语言单位以及各种语音、词汇、语法形式所起到的话语层面的前后关联效果。与这些相关的话题有: 衔接连贯、韵律特征、会话结构、叙事结构、体裁特征、元话语等。

1.3.2.1 衔接连贯

衔接主要指话语中有助于意义传达的词汇语法手段。英语衔接研究最有影响的是Halliday & Hasan (1976), 其中归纳了五种常见类型: 照应、替代、省略、连接和词汇衔接。对于Halliday & Hasan (1976) 有关衔接手段的概括和分类, 后来有学者作了厘清和扩展(胡壮麟 1994; 张德禄 2000, 2001, 2002)。但应该说, 英语衔接手段在词汇语法层面大致离不开Halliday & Hasan (1976) 五分法的框架。各种衔接手段的运用使得话语能够意义连贯。连贯是话语的核心概念之一。任何话语都应具备这一特性。我们将通过下面的段落简析英语中常见衔接手段的工作机制。

The *optimist* feels in the control of his own life. If things are going badly, *he*₁ acts quickly, looking for solutions, forming a new plan of action, and reaching out for advice. The *pessimist* feels like a toy of fate and moves slowly. *He*₂ doesn't seek advice, since *he*₃ assumes nothing can be done.

上面段落中的两个主要名词之一 *optimist* 与其后的 *he₁* 同指；另一个主要名词 *pessimist* 与其后的 *He₂* 和 *he₃* 也具有同指关系。这样前后句子或小句间的名词和代词，一显一隐，相互照应，达到前后意义连贯的效果。

这一话语片段涉及的 *optimist* 和 *pessimist* 是一对相反概念，因此前半部分采用 *quickly*，后半部分采用 *slowly*，对两类人的行事方式进行描述。反义词在话语上起到了相反相成的效果。

再比如，上文中的 *looking for*、*forming*、*reaching out*、*seek* 这些近义词语和短语，相互呼应，也有助于话语前后的意义连贯。

以上仅简单示意衔接和连贯的大致含义和效果。在话语实践中，连贯的实现可能会借助语言或非语言手段，包括语境特征等的协同作用。

1.3.2.2 韵律特征

在口头交流中，语调、节奏、重音、音高、时长、停顿等韵律特征可以对话语的组织及意义传达起到重要作用 (Wennerstrom 2001)。比如语调可以起到实施特定言语行为的作用；重音可以突显话语的信息焦点；停顿可引起听话人的注意，并引导话语走向。除了像停顿这样的韵律特征通常单独使用外，多数韵律特征都是伴生语言资源。在口头话语中，韵律特征通常都是与词汇、语法，乃至体态、手势等非语言手段共同使用以完成话语功能。可以说韵律特征在口头话语中不可或缺。甚至在具有口头话语特征的网络及新媒体语言中，话语使用者会采用特殊的拼写 (如全大写字母)、字符的重叠使用 (如 *soooooooooo*)、标点符号的异常使用 (如 *。。。* 表示长停顿或无语)、笑脸符号或表情包等多种方式，起到类似口语中的重音、拖长、停顿等韵律效果。

1.3.2.3 会话结构

会话结构研究中最重要议题之一是话轮结构，它是 Sacks *et al.*

(1974) 基于对口语会话中人们的互动方式概括出来的说话人话语轮次交替转换的机制。互动会话通常遵循14条原则，比如，说话人之间会轮换交替；大多数情况下，每次一人说话；同时多人说话的情形也时常发生，但一般持续时间很短等等。这些原则看似稀松平常，但却不可以轻易打破。后来很多类似的会话研究都会关注访谈或辩论双方是如何策略性地利用和破坏有关原则，获得话语控制权的。在Sacks等人开创的会话分析研究领域中，学者们致力于考察和探讨不同情境下的“话轮构建单位”(turn-construction unit) 和“话轮分配单位”(turn-allocation unit) 以及如何开启、索取、维持、接续和终结会话。比如Sinclair & Coulthard (1975) 在研究课堂教学话语的基础上，概括出了IRF(Initiation-Response-Follow up, 发话—回应—接续) 会话结构模式(见表1.2)，即发话人A提出某种想法观点，这一轮次称为“发话”，紧接着听话人B会有一个“回应”，通常A会再有一个接续轮次，对B的回应加以评价。这种结构在课堂会话中十分常见。不同场景的会话既遵守一般性的话轮转换原则，也存在特定的会话结构。

表 1.2 IRF会话结构模式

发话	A: Mm. Obviously it's not easy to eat a little and often when you're a busy man driving around the country.
回应	B: Yeah, that's true.
接续	A: Mm. Right.

1.3.2.4 叙事结构

叙事是一种重要的话语类型，通俗来说，叙事即是“(讲)故事”。叙事在不同学科含义并不相同。语言学领域所谓的叙事可以是口语或书面

语形式，可长可短，可以单独出现，也可能嵌套于会话等其他话语类型中。叙事话语形式多样，但也有结构上的规律。其中Labov & Waletzky (1967) 概括的个人经历型口头叙事结构模型影响最大，见表 1.3。

表 1.3 Labov & Waletzky (1967) 叙事结构

口头叙事要素	功能	示例
概要 (Abstract)	叙事人表示他开始讲故事，并简述故事的核心内容	Three weeks ago I had a fight with this other dude outside
发生 (Orientation)	介绍人物、时间、地点及场景	I was sitting on the corner and smoking my cigarette, you know
矛盾激化 (Complicating action)	叙述核心事件	I put that cigarette down, and...I beat the shit out of that motherfucker
评判 (Evaluation)	示意故事的关键之处，即为值得谈起这个故事	But it was quite an experience, I was shaking like a leaf
结局 (Result or resolution)	指明故事的结局	After all that I gave the dude a cigarette, after all that
尾声 (Coda)	示意故事结束，有时会进一步提示其他话题	And that was that

以下是作家王蒙在访谈节目《锵锵三人行》中的叙事片段。通过下面的分析，我们可以看出，这段叙事也大致符合Labov和Waletzky提出的叙事结构。我们把这一小段叙事命名为《你以为你是谁》。

表 1.4 《你以为你是谁》叙事结构简析

好像是池莉写过一篇文章，说《你以为你是谁》	概要
说那个叶，叶兆言呐，((说))托什么一个人儿啊，办个什么事儿	发生
就去找那个有关，有关管事者	
为了，就当敲门砖嘛	
就带着自己的好几本书	矛盾激化 + 结局
说一到那儿就把书送去了	
送上	
然后那位小领导	
说，一看，这个，咱，就往上	
说这一天不知给我送多少	评判
@@@@@@@@	
叶兆言，这还算是来有来历的	
他是叶圣陶的孙子	
是不是啊	
你连这么有来历的作家	
所以池莉就写了篇文章	尾声
就《你以为你是谁》	

注：@代表笑声；((说))代表“说”字在录音中发音不清晰。

1.3.2.5 体裁特征

体裁 (genre) 一词原本是文学艺术领域的概念，指的是小说、戏剧、诗歌等文艺形式。在话语研究领域，体裁这一概念使用十分广泛，含义也比较丰富。从 genre 一词的词源来看，它表示“类型”。整个语言可以裁定成不同的话语类型，某话语大类也可裁定出子类，还可因场景和社会功能之不同，裁定出相应的话语类型。比如，有时体裁可以指口语体、书面

语体，或者新闻、小说、学术这样的文体，有时也指法律文书、学术论文、求职信这样的话语类型，或者在学术话语内部，还可指论文摘要、引言、研究方法、讨论、结语以及致谢等微观话语类型，甚至有将体裁泛化的做法，即认为所有体现一定社会语用功能的话语类型都是体裁。这些不同的认识，对应于不同的体裁观，其中有专用英语/学术英语 (ESP/EAP) 研究传统、北美新修辞学派传统和系统功能语言学传统 (Hyon 1996)。北美新修辞研究传统主要关注英语母语中的修辞、作文法等。近年来 Christine Tardy 和 Paul Kei Matsuda 等北美学者则引领了二语写作领域的体裁研究。就体裁研究的整体情况看，基于语料库的体裁分析、语域变异和专用英语/学术英语研究成果更为丰富。

其中有关学术英语体裁风格的相关研究，突出反映了体裁的结构性特征。这方面的研究源自 John Swales 20 世纪 80 年代提出的“体裁分析法” (genre analysis)，后来在 Swales (1990) 中得以定型，其后引发了广泛而持久的学术英语体裁的理论及应用研究。表 1.5 以学术论文摘要简述学术话语的体裁结构特征。

表 1.5 学术论文摘要的体裁分析

学术论文摘要的体裁结构	话语功能
现有语料库翻译学研究，多数限于描述译文的词汇运用，少量涉及句法特征，对译文语篇特征的关注较少。	研究领域及研究空缺
本研究借助在线文本分析工具 Coh-Metrix，对比了汉译英英语译文和原创英语中的 25 项语篇衔接特征。	研究方法
研究发现，英语译文中有多项语法和词汇衔接特征与原创英语存在显著差异，呈现出衔接显化的特点。语法衔接突出表现为多用各类连词。词汇衔接突出表现为以语篇中句间实词重复。	研究发现

(待续)

(续表)

学术论文摘要的体裁结构	话语功能
英语译文的语法衔接显化,大致可从汉语重意合、英语重形合得到解释。而词汇衔接显化,可能与汉语源语特征以及英语译文中多用高频词和泛义词有关。本文还就衔接显化作为翻译普遍性特征进行了讨论。	研究结论

Swales等开创的体裁分析法旨在将话语功能与体裁结构关联起来,使我们能够了解特定体裁是话语形式与功能的统一体。鉴于体裁研究的跨学科性,不同学者在各自研究视角下,常常还会采用register(语域)和style(语体、文体)等近义术语。本书如无特别说明,则将genre、register和style等同视之。在涉及具体研究时,有时为尊重原作者的使用习惯,会选择使用语域、语体或文体的表达。

1.3.2.6 元话语

Harris(1959: 944)最早提出“元话语”(metadiscourse)这一概念。他指出在话语的核心内容(如We have found that这一英语结构后面引出的内容)之外(Harris 1959: 946),还有一些对话语本身进行谈论的语言成分(There are also metadiscourse kernels which talk about the main material.)。这里的We have found that这一英语结构即是一个元话语的例子。Hyland(2005a: 3)在有关元话语的专论中谈到,元话语的价值在于它让人们认识到,交际不仅是信息的传递或交易的达成,还包含个人特质、态度立场、认识理念等方面内容的表达。这些都在话语层面有相应体现。与其说元话语是主要话语内容之外的结构性成分,不如说元话语反映的是说话人/作者和听话人/读者之间的互动交流。在这一思路下,Hyland(2005a: 49)提出了书面语中元话语的人际模式,见表1.6。

表 1.6 Hyland提出的元话语人际模式

类别	功能	示例
交际引导 (interactive)	引导读者了解话语走向	
过渡标记	小句间关系的表达	in addition、but、thus、and
框架标记	提示语步、话语推进的次序或阶段	finally、to conclude、my purpose is
内指标记	指向话语中的相应部分	noted above、see Fig、in section 2
言据标记	注明信息来源	according to...、...states
语码注释语	对命题意义的注释	namely、e. g.、such as、in other words
互动交流 (interactional)	将读者纳入话语互动之中	
模糊限制语	缓和话语, 使话语保持开放	might、perhaps、possible、about
增强语	强化语势, 话语相应收拢	in fact、definitely、it is clear that
态度标记	表达作者对命题的态度	unfortunately、I agree、surprisingly
自我提及语	指称作者的表达	I、we、my、me、our
介入标记	明确建立与读者的关系	consider、note、you can see that

以上介绍了元话语的人际模式特征。事实上, 话语是形义兼具的概念。在理论上很难将其结构特征和语义特征分开讨论。然而, 在话语特征的研究实践中, 我们在考察衔接手段、韵律特征、学术论文体裁特征、元话语标记等话题时, 可以从具体语言形式入手, 进而解读其话语特征。而在讨论下面一些话语研究选题中, 则常常从语义或功能入手去观察话语特征。为行文方便, 本书将后一类话语特征统称为话语的语义特征。

1.3.3 话语的语义特征

这里所谈的话语语义特征，更准确地说，应该是话语的语用 (pragmatics) 特征，而非语义学 (semantics) 层面的特征。因此这里介绍的研究领域主要涉及言者含义 (speaker meaning) 或言外隐含意义 (implicated meaning)。虽说语用意义存在于言外，但在很多情况下，我们仍然可以借助语境推测得知。

话语的语义特征经常涉及的话题有：话语的主观评价维度 (包括情态、语气、立场、言据性等子话题) 以及话语中信息显隐与接续维度 (包括主题、话题、话题链、焦点、前景化等子话题)。

上述两大维度的讨论常见于各类功能语言学论述之中。譬如，对于话语的主观评价维度，Halliday (1973, 1985a) 就提出过系统解决方案。其系统功能语法三大元功能中，与概念功能 (ideational function) 和语篇功能 (textual function) 并列的人际功能 (interpersonal function) 主要针对的就是情态、语气、极性等主观性语言维度。后来，Martin 等人对 Halliday 的人际功能进行了扩展和细化，从而提出“评价理论” (The Appraisal Theory) (Martin 2000; Martin & White 2005; Martin & Rose 2003)。该理论将主观评价维度分为介入 (engagement)、态度 (attitude) 和级差 (graduation) 三大子系统，其下又分为若干次系统，见表 1.7。评价理论的这一体系被广泛应用于不同体裁的评价分析之中。

表 1.7 Martin 等人提出的评价理论框架

评价子系统	评价次系统	定义及英汉语典型表述
ENGAGEMENT 介入	entertain 接纳	个人主观意见的表达，承认所言只是一种可能的意见
		might、seem、probably、I guess、you can imagine、look like/看上去、好像、可能、肯定、看样子、显得、不太确定

(待续)

(续表)

评价子系统	评价次系统	定义及英汉语典型表述
ENGAGEMENT 介入	proclaim 宣言	将说话人自己的意见视作正当的意见, 否认其他意见的合理性
		make sure、actually、obviously、of course、anyway、we should say/ 确实、我们可以看出、显然、完全、基本上、确信
	disclaim 否认	通过语篇手段表明相对立的态度和立场
		without、anything、no、never、not at all、not any more/ 没有注意到、还不知道、没有想到任何事情
ATTITUDE 态度	affect 情感	正面或负面的情感(属个人情绪)
		happy、puzzled、angry、enjoy、reluctant/ 高高兴兴的、伤心的、惊讶的、无辜的
	judgment 判断	对特定行为的态度, 如崇敬、批评、褒扬、谴责; 道义判断; 行为缘起判断(属伦理性)
doubt、wonder、carefully、appreciation、unfortunately、properly/ 满腹狐疑、邪念、幸运地、小心翼翼地		
appreciation 鉴赏	对自然现象的评价; 审美价值判断(属审美层面)	
	tranquil、beautiful、pretty、lovely、bumpy、weighty/ 安静的、丰收的、有趣的、美丽的	
GRADUATION 级差	force 语势	涉及质性、可能性、过程等的强度以及数量、尺寸、重量、分布、远近等量化特征
		completely、slightly、just、all、near、nearby、suddenly、soon/ 非常、很、正好、整个突然地、飞快地、很多的

(待续)

(续表)

评价子系统	评价次系统	定义及英汉语典型表述
GRADUATION 级差	focus	根据典型性或精确性作出级差判断; 调整以表明典型性或边缘性
	聚焦	really、gradually、a little bit、about/差不多、有点、高点儿、胖点儿

其他一些学者, 由于各自不同的学科背景和特定的研究目的, 也提出了相应的评价分析理论。例如, Biber & Finegan (1988, 1989) 引入“立场”概念, 并在 Biber *et al.* (1999: 966-986) 作了更为系统的表述。他们将命题意义之外的主观评价概括为“认识型立场”(epistemic stance)、“态度型立场”(attitudinal stance) 和“言谈方式立场”(style of speaking stance) (Biber *et al.* 1999: 972-975)。若与 Martin 等人的评价理论体系进行对比, 可以看出 Biber 等人的三分法与“介入”“态度”“级差”较为接近。只是 Biber 等人的体系更为概括, 在体系的层级方面也未再作细分。作为语法描写的一个重要部分, Biber 等人的立场体系虽然收录在 *Longman Grammar of Spoken and Written English* 一书中, 但该理论体系从产生到其后的进一步应用, 都具有很强的语域变异倾向。他们所概括的立场类型是基于 Tannen (1982)、Halliday (1985b)、Chafe & Tannen (1987) 等人对口语和书面语差异的研究得出的语域区别性特征。Biber *et al.* (1999: 979-986) 在 *Longman Grammar of Spoken and Written English* 中提出立场分类体系后, 围绕口语、小说、新闻和学术四种语域类型, 对立场副词、状语从句、立场名词等分别作了量化描述。后期, Biber (2006) 还专门就大学语境中的课堂教学和教材等语料进行了认识型立场、态度型立场和言谈方式型立场的对比分析。概言之, Biber 的立场理论源自语域

变异研究，最终也服务于语域变异研究。另外，由于Biber旨在运用语料库方法开展立场研究，在他的相关论述中，他着重强调了立场与显性词汇语法手段之间的对应关系。

Hyland(1999, 2005b)对立场也有过系统论述和实证研究，他的研究焦点在于学术论述中的人际互动模式。综合Hyland(1998)有关学术论文中的模糊限制语、Hyland(2005a)有关学术交流中的元话语以及Hyland(2005b)中的立场与介入研究，可以看出，Hyland所讨论的立场与介入大体是其元话语体系中互动交流和交际引导的另外一种表述。他的元话语观似乎偏重结构性特征，而立场研究则更偏向于从人际功能出发去研究话语互动。

另一个有关评价研究的理论性探讨来自Hunston。她提出“事态评价”(status)(Hunston 1989: 106)、“价值评价”(value)(Hunston 1989: 172)、“关联评价”(relevance)(Hunston 1989: 236)三项评价功能，旨在考察话语中对相关命题的评价性表述。其中事态评价主要传递确定性(certainty)，这在一定程度上与其他评价体系中的介入和认识情态有相通之处；价值评价涉及好恶和善恶，与态度、义务情态类似。事态评价通常具有名词性，价值具有述谓性(Hunston 1989: 199)。例如在The hypothesis is supported by the data.这句话中，hypothesis表达的是事态，而support表达的是价值判断。至于关联评价则更具话语组织功能。例如：

(1) The electric dichroism of chromatin yields information relating to two structural features. (2) First, the average orientation...can be estimated from...(3) Second, the relaxation rate...can be measured...(4) These two parameters taken together form an exacting criterion for the structural state of chromatin. (Hunston 1989: 239)

在上述话语片段中，第(4)句话所起的评价功能即是关联。它将前文

(2)(3) 两句加以总结概述, 使前后呼应。Hunston 围绕事态、价值、关联开展的评价研究与 Hyland 的元话语研究有一点明显的区别, 前者更多关注话语核心元素所表达出的评价功能, 而后者关注的主要是话语核心成分之外的评注性要素。

Hunston 评价三分法中, 事态评价在后续研究中探讨得最多, 并且跟语料库理念, 特别是短语学和局部语法分析方法结合比较充分。例如,

There is an alternative **interpretation**... the position and momentum of the electron are defined all along. (Hunston 2011: 26)

该例中, 粗体的 **interpretation** (解释) 与后文 *the position and momentum of the electron are defined all along* (电子的位置和动量同时得以确定) 这一命题在话语层面形成照应。解释是对其后命题的预示, 也是对该命题的判断及评价。该句中的 **interpretation** 所发挥的评价功能即是事态。**interpretation** 也被称为事态名词 (*status noun*)。在此基础上, Hunston 将评价分析跟语法构型 (*grammar pattern*) 和短语联系起来, 并通过语料库手段辅助评价分析, 如以事态名词、名词 + *that* 小句、*it v-link ADJ* 作为“探针” (*probe*) (Hunston 2002: 62), 可以获得大量评价表述, 并可就某个特定的构型进行所谓局部语法的描述, 同时揭示其局部功能。与此相关的研究还包括“抽象名词” (*abstract noun*)、“概念名词” (*conceptual noun*)、“壳名词” (*shell noun*), 如 *problem*、*idea*、*fact* 等名词 (Schmid 2000); “泛指名词” (*general noun*), 如 *time*、*people*、*world*、*thing* 等 (Mahlberg 2005); “标识名词” (*signalling noun*), 如 *attitude*、*process*、*reason*、*result* 等 (Flowerdew 2003; Flowerdew & Forest 2015)。

有关主题、话题、话题链、焦点、前景化等子话题的研究较多反映在 Halliday (1985a) 的系统功能语言学、美国西海岸功能语言学 (Givón 1984; Hopper & Thompson 1984; Li & Thompson 1989;

Givón 1990; Bybee & Hopper 2001) 以及语言观相近的认知语言学 (Langacker 2008) 研究中。这一研究取向与句法研究关系紧密, 注重语法成分或结构的话语组织或语用功能, 例如张克定 (2001, 2007, 2011, 2012) 对 (方位) 倒装构式、句首空间附加语、主位化评述结构在话语组织和评价方面的探讨。然而总体来说, 这一研究传统只有少数研究成果 (苗兴伟、秦洪武 2010; 王义娜、李银美 2016)。目前少有研究基于大规模语料开展实证量化分析。

概言之, 本节介绍的话语评价特征研究, 在一定程度上都直接或间接受到 Halliday 功能语言学理论, 特别是人际功能相关论述的影响。这在 Biber、Martin、Hunston、Hyland 的文献中都有提及。另外, Biber 的立场观和语域变异研究很大程度上受到社会语言学的影响。Biber 的相关研究关注评价性语言和语域变异之间的互动关系。在研究方法上, 学者们注重将具有主观性特色的评价研究操作化, 从而可以更好地借助语料库方法开展实证研究。

正如话语结构特征与语义特征不能截然分开一样, 话语的语义特征及其社会属性也是紧密关联的。仅以上文探讨的立场研究为例, Hyland (2005b: 173) 就指出, 学者们在产出文本时, 不单是产出一种外在产品, 同时也建构和协商社会关系。Hyland & Guinda (2012) 汇集了有关立场的语言和社会视角的很多重要研究成果。

1.3.4 话语的社会属性

话语的社会属性常以社会认知、话语建构、形象建构、身份认同、权势关系等概念在文献中出现。前文谈过, Harris (1952) 提出“话语分析”这一概念时, 就将话语的社会维度考虑在内, 然而, 直到20世纪70年代之后, 话语的社会维度才逐渐受到重视, 近年来更成为话语研究的热点。这跟语言学和社会科学领域从结构主义到后结构主义/解构主义 (Angermuller 2014; Baker 2006: 5-9; 苏晓军 2009; 王宁 2009) 的思

潮演变较为一致。后结构主义是对结构主义的背离。在结构主义理论中,包括语言在内的现象和观念,通常都有明确的核心以及与之相对立的边缘现象。一些重要观念都具有二元对立(binary opposition)的性质,例如,有长有幼、有尊有卑、非好即坏等等。以Derrida和Foucault为代表的后结构主义者则认为,人们认识中的这些二元对立和中心论等都应被解构(Angermuller 2014)。在此之外,可以作为后结构主义补充的理论是社会建构主义(social constructionism)(Burr 1995; Teubert 2007: 57, 67)。这一理论认为,社会现实和观念不是先验的存在,也并非固定不变,而是社会成员在长期互动基础上动态浮现(emerge)而成。这种社会建构会因时因地而有所区别。因此,我们也可将话语的社会属性研究称为社会建构(social construction/representation)或话语建构(discursive construction)研究,这同时也被Gee(2015)称为与话语的语言结构特性(discourse with a small d)相对的“大写的词语”(Discourse with a big D)。为讨论方便,我们将话语的社会建构常见选题概括为:1)具体事件建构;2)抽象观念建构;3)个体身份建构;4)群体形象建构。

对于话语的社会建构,有学者(如Teubert 2010)认为研究所建构的仍然是话语的语义特征,只不过是社会意义。这种认识并无不妥,只是语义更多局限于语言本体,社会意义指向语言符号的外部表征,甚至人们对某些话语现象的认识。语义和社会意义可理解为处于话语意义连续统的两端,一端为语言本体,一端为社会语境。话语和社会语境两者相辅相成,一定的话语根植于一定的社会情境,特定的话语也会塑造一定的社会现实。因此,我们可以通过研究话语探究社会现实,甚至推知社会现象背后的深层问题,同时也可能通过恰当的话语引导社会风气。通常,话语建构研究旨在揭示话语现实,某些批判话语研究致力于通过对话语的解析从而改变社会风气(Wodak & Chilton 2005)。在操作层面,对于话语的建构,也常被称为“所言之事”(aboutness)(Phillips 1985, 1989)。借用新闻写作的“5W + 1H”的六要素来看,话语建构就是什么人(Who)在什么

时间(When)什么地方(Where)做了什么事情(What),具体是如何做的(How),为何会出现这样的情形(Why),也可简化成“描述”和“阐释”两方面。有关事物和观念的建构也大体如此。话语产出(写作或会话)与话语建构研究的最大区别是,前者是建构话语事实,后者是反其道而行之,基于对话语的分析重构(re-construct或re-present)话语事实。以下将概述四类话语建构的常见研究选题。

1.3.4.1 具体事件建构

这类话语研究通常关注特定的社会、自然现象或突发事件。例如, Powers & Xiao(2008)关于非典(SARS)话语的研究、McEnery(2009)关于Mary Whitehouse维护电视节目内容洁净行动(Clean Up TV)的分析、Potts(2015)关于美国卡特里娜飓风报道的考察、Yang(2014)关于2008年奥运火炬传递中英媒体报道的研究。这里以网络上常雪松撰写的一篇题为《从语义分析看〈徐霞客游记〉的细节:他吃什么?住哪?买了什么纪念品?》的文章示例徐霞客旅游体验的话语建构。该文以朱惠荣所译约80万字(约60万词)的《徐霞客游记全译》作为语料,通过高频词(下文双引号中为语料分析高频词)分析,回答了这几个问题:1)徐霞客主要在哪里吃饭,和谁一同用餐,饭菜如何?2)徐霞客每天在哪里住宿?居住环境如何?3)徐霞客的主要交通工具是什么?4)徐霞客一路都有哪些购物及消费?上述问题是从旅游经济学的食、宿、行、游、购、娱(郑杨1998)几个方面提出的。文章尝试通过话语分析,重构徐霞客的经历,并加以适当评价。作者分析得出:徐霞客最主要吃饭的地方是在“寺庙”,常与“僧人”“法师”“挑夫”“脚夫”一起用餐。徐霞客所吃饭菜包括肉类、蔬菜、水果、主食等。其中肉类以鱼类为多,且种类丰富,包括“鲤鱼”“鲫鱼”“香鱼”“青鱼”“白鱼”等。还常以“松子”等作为零食。徐霞客主要住在“县城”“旅店”。除旅店外,“寺庙”也是徐霞客的落脚点之一,他偶尔还会“寄宿”在“当地人”的“茅屋”和“竹楼”内。作者常

雪松推断《徐霞客游记》的大部分应是完成于一个个旅店的二楼房间，因为通过高频词和上下文分析，徐霞客入住旅馆后，一般会挑选相对安静和舒适“楼上”“房间”入住，并写“日记”。

表 1.8 《徐霞客游记》中的交通出行方式

交通方式	所占比例	交通方式	所占比例
行	66.50%	辑	0.26%
舟	16.28%	轿	0.17%
马	8.95%	楫	0.17%
步	3.81%	滑竿	0.06%
走	3.09%	驴	0.04%
船	0.63%	骡	0.04%

如表1.8所示，出行方面，70%以上的行程，徐霞客是靠双腿走出来的(行、步、走所占比例为73.4%)。他在旅途中少有“娱乐”。除了游历山川，可能就是“被困”寺观，与“和尚”“道士”取经问道了。他购买最多的是“大米”“鱼”“柴火”“香蕉”“橘子”“衣服”等生活必需品。偶尔也会购买少量“石头”“扇子”作为纪念，这与其文人身份较为吻合。

这一案例说明话语研究可以有效重构几百年前，徐霞客长达24年，足迹遍及如今21个省、市、自治区的游历历程。

1.3.4.2 抽象观念建构

抽象观念话语建构可理解为大量具体事件整体呈现出的概括性话语现实。例如，偏见(Billig 1988)、恐怖主义(Qian 2010; 张立英、李可 2017)、移民问题(Hart 2013)、忧郁情绪(Wang 2013)、战争合法性(庞

超伟 2013)、悲伤情绪 (Zhang 2014)、中国梦 (邵斌、回志明 2014; 钱毓芳、黄晓琴 2016)、全球化 (Dye 2015)、气候和环境问题 (Dayrell & Urry 2015; Willis 2017)、低碳经济 (钱毓芳 2016)、恐惧与威胁 (Cap 2017)、中医药 (钱毓芳 & McEnery 2017) 等。这里我们以媒体中的幸福感建构 (廖卫民、钱毓英 2012) 为例说明此类话语建构研究的思路。通常相关研究中的抽象概念需加以操作化, 即根据某种理论框架或研究者的分类将抽象概念划分成若干维度或层级, 然后再从语言层面寻求对应的表述, 最终再对抽象概念的整体建构作出描述和评价。

廖卫民、钱毓英 (2012) 关注的是中国 GDP 高速增长背景下百姓的幸福感问题。这一选题看似俗套, 却有很强的现实意义。一般人们倾向于认为幸福感较为空泛, 但根据作者的文献综述, 幸福感研究有情绪模型、认知模型、主观幸福感、心理幸福感和整合模型。在该研究中, 作者提出了幸福的五大因素模型, 包括社会物质条件、工作事业发展、个体情感、个体认知和社会关系。这五个维度整合了前人的相关研究, 同时也考虑到了相关因素在文本分析中的可操作性。作者选取了浙江当地的 15 种报纸、五档电视节目作为分析语料, 分析结果见下表。

表 1.9 基于语料库数据的社会幸福感评估表 (精简自廖卫民、钱毓英 2012: 24)

社会幸福感构成因素	主要代表词汇	媒介呈现评估 (%)
社会物质条件	盛世、和谐、美满、腾飞、崛起、转型、小康、殷实、宜居、平安	37
工作事业发展	顺利、增长、工作、学习、劳动、忙碌、共建、共享、奋斗、创造	13
个体情感	爱、晒、自尊、喜欢、信任、快乐、感恩、安全、感动、愉悦	11

(待续)

(续表)

社会幸福感知因素	主要代表词汇	媒介呈现评估(%)
个体认知	认识、知晓、感悟、理解、豁达、乐观、宽容、满意、满足、放松	10
社会关系	婚姻、家庭、友爱、互助、帮扶、恩爱、和美、团圆、融洽、互谅	29

空泛的幸福感和离散的字词,通过概念模型得以条理化、体系化,从而帮助我们完成对抽象观念理论化的建构。此外,该研究在描述基础上作了很好的概括和评价。作者通过量化统计和文本解读得出:浙江媒体对社会幸福感总体上存在“重物质、轻精神,重集体、轻个体,重宏观、轻微观,重官方、轻民间,重规划、轻落实”的不平衡现象。同时,作者还尝试从“不幸福感”进行分析,挖掘所谓“被遗忘的幸福感”,说明在幸福的同时,还隐含一些诸如“健康”“忙碌”“亚健康”,对弱势群体关注不足以及部分人“炫”幸福等现象。

1.3.4.3 个体身份建构

个体身份建构的对象通常包括两类。一类是具有突出社会影响的名人,比如政要、知名学者或网红等。例如有关特朗普(Knoblock 2017; Ott 2017)、Widdowson(De Beaugrande 2001)、John Swales和Deborah Cameron(Hyland 2010)、网络虚拟形象(Georgalou 2017)等身份的建构。另一类为数众多的个体身份建构研究关注的是话语中的自我建构。后一类研究虽然有时也是针对特定群体进行考察,但强调每个说话人或作者都有其独特的形象或声音(Hatch *et al.* 1993; Ivanič 1998; Matsuda & Tardy 2007; Matsuda 2015)。此类研究多见于二语写作和学

术写作研究文献中。话语中的身份建构概括起来有三类观察视角：1) 个体 (personal或individual) 身份视角，可以理解为一人一面，人人不同；2) 社会 (social或social constructionist) 身份视角，即众人一面，士农工商，君臣父子，人们各有其社会规约性身份，同时也会因社会场景的改变而切换成其他社会规约性身份；3) 互动 (interactive、dialogic或social constructivist) 身份视角，即身份无不由社会互动协商产生。作者总在有意识或无意识地设定一定的读者对象，从而呈现相应的个人形象。而读者事实上也并非完全被动，他们也是身份建构和协商的一部分 (参见Tardy 2012; Hatch *et al.* 1993)。前两类观察视角较容易理解，以下以Hatch *et al.* (1993) 为例说明互动身份研究的典型思路。

Hatch等人研究的副标题《读者印象中的作者个性》(“Readers’ impressions of writers’ personalities”)更能明示该研究的属性。这里的读者印象和作者个性就是我们谈的身份。Hatch等人基于几十份高中生撰写的大学入学申请信 (admissions essay) 设计了三项研究任务。第一项任务是请卡耐基梅隆大学的教师和助教一共三人分析61篇申请信，三人共概括出30项左右申请人的个性特征，见表1.10。三人对61份材料的作者身份分析一致性具有统计学上的显著意义。

表 1.10 Hatch *et al.* (1993: 579) 研究中 “读者印象中的作者个性”

Positive traits (积极身份特征)	Negative traits (消极身份特征)
likable(可爱)	flippant(无礼)
sensitive to people(人际敏感)	pretentious(矫揉造作)
mature(成熟)	arrogant(傲慢)
down-to-earth(脚踏实地)	immature(不成熟)
positive attitude(积极态度)	unlikeable(不可爱)

(待续)

(续表)

Positive traits (积极身份特征)	Negative traits (消极身份特征)
creative(有想法)	weird(怪异)
interesting(有趣)	socially awkward(不善社交)
broad(不狭隘)	negative attitude(负面态度)
socially adept(善于社交)	egocentric(自我中心)
self-sufficient(善于自理, 不求人)	naive(幼稚)
determined(有决心)	dull(无趣)
serious(严谨)	narrow(狭隘)
hardworking(勤勉)	
modest(谦逊)	
assertive(果决)	
ambitious(有抱负)	
self-confident(自信)	
extrovert(外向)	

第二项任务是请卡耐基梅隆大学的八位招生负责人(admissions counselor)评阅61份材料中的20份。请他们依据申请书的评阅结果录取20人中的10人,淘汰10人。研究结果显示,在任务一的分析中获得更多积极身份标签的申请人更容易受到青睐,从而被录取。

第三项任务是请23位研究生同学评阅九份申请信,圈点并标记出表1.10中提到的身份特点。这九份申请信中有些是经过润色的,有些是未经改动的原稿。相关改动润色着重考虑减少负面身份特征(比如过于华而不实的语言),而不改动申请信的核心内容。比如将原稿中的I could now carry on intelligent conversations where most of the students in the classroom understood what I was talking about.这一句改为I enjoyed

conversations with other students who shared knowledge and interests similar to my own. 第一句的语气有些傲慢, 改后的句子则要谦虚得多。测试结果显示, 修改后的申请信较原稿而言, 明显得到了更多的积极身份特点评分。

这一研究表明, 作者所呈现出的身份特征在读者一端是可以解读出来的, 并有较高的共识。良好的作者身份建构事关录取与否, 招生负责人(读者)对于话语身份也十分看重, 甚至是决定性的。话语层面的改善可以有效提高读者读取作者积极身份的概率。可见身份建构绝不是简单的单向度概念。

1.3.4.4 群体形象建构

常见的群体形象建构研究话题包括国家形象 (Boulding 1956, 1959; Wodak *et al.* 2009; 唐丽萍 2016; 梁云等 2017)、政党形象 (Fairclough 2000; L'Hôte 2014)、青年人形象 (Androustopoulos & Georgakopoulou 2003)、性别形象 (Baker 2005; Pearce 2008; Koller 2008 a; Koller 2013; Baker 2014)、商人形象 (彭焕萍 2008)、城市形象 (Koller 2008 b)、难民形象 (Gabrielatos & Baker 2008)、企业形象 (The PAD Research Group 2016) 等。其中国家形象建构应是研究较多的一项选题。

这里以唐丽萍 (2016) 为例, 示例此类研究的思路及方法。该课题利用 LexisNexis 媒体数据库, 采集了《纽约时报》《华盛顿邮报》《洛杉矶时报》自 2008 年 9 月 18 日至 2010 年 12 月 31 日的涉华报道, 建成中国形象语料库, 库容约 808 万词次。研究围绕 China 一词, 借助词语搭配、索引分析等方法, 并结合系统功能语言学中的及物性和评价分析以及语料库研究中的语义韵等理论视角, 对美国大报中的中国形象进行了描述。比如, 作者通过词语搭配分析发现: rising、Communist、assertive、authoritarian、hungry、prosperous 经常被用来修饰 China。而通过及物性和语义韵相结合的分析, 研究概括出中国“被评判者”“被规劝者”“受

益者”“被惩罚者”和“被遏制者”的形象。美国大报对中国国家形象的建构仍然是以妖魔化、边缘化、遏制为主要的的话语模式。虽然美国媒体也塑造出“经贸激增、富裕繁荣、国际影响力日益提升”这样的中国形象，但也存在有意贬低中国崛起、误导民众、蓄意遏制中国的特点。

本节从话语建构的常见选题，分四个方面对相关文献作了综述。不难看出，话语建构研究涉及面十分宽泛，可拓展性也非常强。随着新鲜事物的产生，新人物的出现，热点事件及新颖概念的不断形成，此类研究成果层出不穷。另外，从研究的理论视角看，话语建构受到传播学、社会学、人类学、心理学、政治学等多个学科的关注。在语言学内部，除话语研究之外，语用学、社会语言学、跨文化交际、翻译学等也都从各自理论立场，开展了很多话语建构研究。然而，无论上述研究涉及哪个领域、哪个学科分支，在开展话语建构研究时，都无法回避对话语的分析，即对真实使用中语言的考察。因为，语言与身份建构之间存在一种“指示性”(indexicality) (Silverstein 1976: 29)，即语言和身份(社会意义)之间存在一定的对应关系。例如，男性跟女性之间各有其说话方式；青少年和老年人也各有其不同的话语表达方式。然而，身份的社会规约性，也存在一定的变异性，相同的说话人会因时、因地、因事不同而对其说话方式有所调整或修饰。话语建构的动态性和变异性，甚至不可预期性，更增添了其研究价值。

1.3.5 结语

本章概述了本书的两大核心领域：语料库研究与话语研究。笔者在话语特征和话语研究的常见选题方面着墨更多，原因之一在于话语研究不断拓展的蓬勃势头；原因之二是在本书中笔者将“话语”视为主要研究对象或核心议题，而将“语料库”视为语言数据、分析方法和研究视角。换言之，除了语料库语言学这一研究角度外，还有很多其他类型的话语研究路径。语料库研究突出的方法论特色，我们将在下一章重点加以介绍。