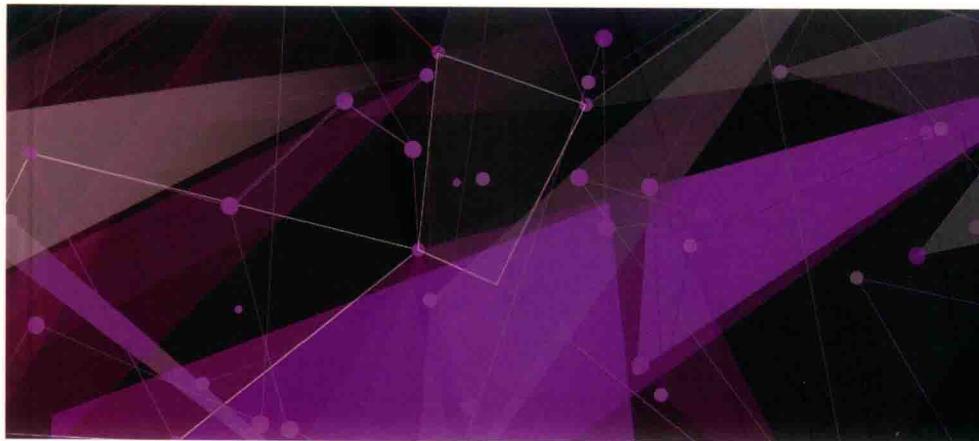


基于多模态语料库的语力研究： 多模态语用学新探索

A Multimodal Corpus-Based Study of Illocutionary Force:

An Exploration of Multimodal Pragmatics from a New Perspective



黄立鹤 著



上海外语教育出版社

社 SHANGHAI FOREIGN LANGUAGE EDUCATION PRESS

www.sflp.com

基于多模态语料库的语力研究： 多模态语用学新探索

A Multimodal Corpus-Based Study of Illocutionary Force:

An Exploration of Multimodal Pragmatics from a New Perspective

黄立鹤 著

图书在版编目(CIP)数据

基于多模态语料库的语力研究：多模态语用学新探索/黄立鹤著.

—上海：上海外语教育出版社，2018

ISBN 978-7-5446-5334-3

I . ①基… II . ①黄… III. ①语料库—语言学 IV. ①H0

中国版本图书馆 CIP 数据核字(2018)第 103459 号

出版发行：上海外语教育出版社

(上海外国语大学内) 邮编：200083

电 话：021-65425300 (总机)

电子邮箱：bookinfo@sflp.com.cn

网 址：<http://www.sflp.com>

责任编辑：许进兴

印 刷：常熟高专印刷有限公司

开 本：700×1000 1/16 印张 23.75 字数 362千字

版 次：2018 年 9 月第 1 版 2018 年 9 月第 1 次印刷

印 数：1100 册

书 号：ISBN 978-7-5446-5334-3 / H

定 价：75.00 元

本版图书如有印装质量问题,可向本社调换

质量服务热线：4008-213-263 电子邮箱：editorial@sflp.com

序 (一)

十分高兴获悉立鹤的书稿即将付梓。受邀为之作序，欣然应允。

本书有一个副标题，“多模态语用学新探索”。本序主要花点笔墨交代一下学术背景，为读者阅读此书提供一个历史线索。

汉语“模态”一词可以用来翻译英语的 mode、modality 或 module 三个词。“多”字是用来翻译英语前缀 multi-。因而，英语词 multimodality 被翻译为“多模态”。我们知道，在语言学中，mode 很早就成为术语，如在以韩礼德为代表的功能语言学中，语境由三元素组成：(1) field，场景；(2) tenor，人物及其关系；(3) mode，指在某个语境里，是选用口头交流还是通过书面文字，即作“语言方式”的选择。Hodge & Kress (1988) 在拓展韩礼德的社会符号学理论时，一方面承袭这个用法，另一方面又把它用来指非语言方式 (nonverbal mode)。后来，Kress (2009; 2010) 用它来泛指能够产生意义的任何符号资源，如文字、图像、颜色、版式、手势、音乐、动漫画，甚至包括发型、化妆等，这些都可以用来传达意义，因此都是 mode。当今的图书页面、网页、手机触屏设计等，往往同时混合使用多个这样的 mode，如此呈现方式被称之为 multimodal，普遍译为“多模态”，对混合使用的作品用话语分析的方法作系统研究，称之为 multimodal discourse analysis，译为“多模态话语分析”。

英语词 multimodality，还有一种意义，来自 multi- + modality。在脑神经科学里，modality 指感官，如人们日常说的“五官”，在现代脑神经科学里称之为 sensory modalities (即感官模态系统)。以视觉为例，眼睛器官加上处理视觉信号的神经系统组成一个 visual modality，即视觉

模态。多模态指多个感觉器官加上处理各自信号的神经系统。

II 神经科学里的“多模态感官系统”是人(包括动物)跟外部世界互动的渠道。哲学、心理学研究的感知(perception)正是涉及这样的多模态。计算机科学、人工智能早就关注多模态感官系统在人机互动中的重要作用：触摸输入、语音输入、打字输入、机器人模拟人类多模态互动等，都是非常基本的高科技课题。

概言之，多模态话语分析中的 multimodality，源自英文的 multi + mode，指多种符号系统；多模态语用学里的 multimodality，源自英文的 multi + modality。由于人脑在处理由多符号系统构成的文本时，需要调用多模态感官系统，多模态话语分析因此跟本文的多模态感官系统研究是有内在联系的，为了避免混淆，本人曾建议把 multimodal (= multi + mode) 译为“多符号”，把 multimodal (= multi + modality) 译为“多模态感官”，把视觉、听觉、触觉、嗅觉和味觉统称为“多模态感官系统”。多模态感官系统和语言研究有着曲折的历史，充满了争议。尽管如此，研究者的观点愈来愈趋同，即从基于单模态的转向基于多模态的。

语用学在学界公认的定义是研究语言使用的学问。作为一门语言学的分支，公认肇始于 Morris 做出的三分野：(1) 语法学，研究语言成分之间的句法关系；(2) 语义学，研究词语与句子的语义关系；(3) 语用学，研究语言使用跟使用者之间的关系。非常有趣的是，Morris 在做这些分工时，他针对的不是语言单位，而是更加普遍的人类“造义行为”(meaning-making behavior, semiosis)。上面说的多模态感官系统都参加造义活动。Morris 在 *Foundations of the Theory of Signs* (1951 [1938], Chicago: The University of Chicago Press, p. 30) 给语用学下的定义原文引用如下：

By ‘pragmatics’ is designated the science of the relation of signs to their interpreters. [...] signs have as their interpreters *living organisms*, it is a sufficiently accurate characterization of pragmatics to say that it deals with the biotic aspects of *semiosis*, that is, with *all the psychological, biological, and sociological phenomena which occur in the functioning of signs.* (斜体为引者加)

living organisms 指鲜活的生命体，包括人类和动物；semiosis 指造义活动，即涵盖符号功能所涉及的符号解读者的心灵，以及生物、社会的所有方面。Morris 把参与造义活动的多模态感知系统分为远距离的如听觉、视角、嗅觉，和近距离的如触觉、味觉等。

当初 Morris 这个语用学初心，所构想的正是我们现在要研究的多模态语用学。

在我国思想史和中医实践上，不乏多模态语用学的思想。荀子把多模态感官系统称之为“天官”，天官具有自然而然的知觉“知识”，《荣辱》指出：

目辨白黑美恶，耳辨声音清浊，口辨酸咸甘苦，鼻辨芬芳腥臊，骨体肤理辨寒暑疾养，是又人之所常生而有也，是无待而然者也，是禹桀之所同也。

“目”“耳”等即多模态感官系统，“辨”如此如此，用现代的术语说，即感知功能。这些感知功能是人类所共有的，是高一级认知的基础。中医的四诊——望、闻、问、切，正是医生通过多模态感官系统获取求诊者信息的方法。

概言之，多模态语用学在思想渊源上非常古老，我们现在能做的是利用当今高科技手段，把此项研究推向深入。

需要指出的是，立鹤在博士论文研究期间，锲而不舍，精神可嘉。论文用多模态语料库语言学的方法研究文盲人的言语行为，虽然有些问题仍待继续深入探究，但毫无疑问，立鹤的这个多模态语用学研究的确是一个值得做的尝试。

顾曰国
2017年5月20日于北京知行斋

序

(二)

序 (二)

欣悉黄立鹤副教授的专著《基于多模态语料库的语力研究：多模态语用学的新尝试》付梓，十分高兴，为学术界新增了一部把语用学和多模态研究相结合来探讨话语语力(Illlocutionary Force)的新著。由此，特向他表示祝贺。

立鹤具有很好的创新意识、科研意识，善于探索和研究，无论从事什么工作，都坚持学术研究，持之以恒，从不间断，在读博士之前就表现出很好的科研潜力和发展前途。特别值得提出的是，他师从我国话语分析领域的领军人物——中国社科院的顾曰国教授，从他那里学到了科学的研究的真谛。

这部新著是上海外语教育出版社通过层层筛选确定出版的。本书既反映了世界语言学以及一些其他学科，如符号学、人类学、社会学、生理学、医学、人工智能、教育学等发展的一个大趋势，同时，也具有自身的理论和方法的创新性。

语言的多模态性是人类很早就发现，并且一直实践着的，如说话时打手势，身体摆一定的姿势，有一定的面部表情，用一定的腔调、口气、语调等。这些都是表达意义的模态。说同样的话，用不同的口气、腔调、语调，就会有不同的解读。但语言学研究一直没有把这种现象作为研究的重点，认为语言学的主要任务是研究语言的结构、语音、意义模式等，这些特征表达的意义都是在了解了语言的意义后，可以自动解读的。但对于非语言和伴语言的研究也一直没有间断过。在现代，从多模态话语分析角度研究多模态话语最早的研究者之一是 R. Barthes，他在 1977 年发

表的论文《图像的修辞》(Rhetoric of the image)中探讨了图像在表达意义上与语言的相互作用，并且发现了图像和文字之间三种主要关系，分别是锚定(anchor)、接续(relay)和阐释(illustration)。但当时并没有引起足够的重视。直到20世纪90年代，才在西方有一系列相关著作和论文陆续出版和发表。其中影响较大的是Kress & van Leeuwen (1996)的*Reading Images: The Grammar of Visual Design*《读图：视觉设计的语法》。另外，他们(Kress & van Leeuwen, 2001, 2003)还研究了模态与媒体的关系，专门探讨了多模态现象规则地表达意义的现象，包括视觉图像、颜色语法以及报纸的版面设计和不同媒介的作用等；O'Halloran (2004)不仅研究多模态的理论建构，还专门研究了数学语篇中的多模态现象；Royce (2002)研究了不同符号在多模态话语中的互补性以及在第二语言课堂教学中多模态的协同性等。与此同时，在国内多模态话语研究也在21世纪初如火如荼地发展起来。除了引进外，还进行了卓有成效的研究(李战子, 2003；胡壮麟, 2007；朱永生, 2007；顾曰国, 2007；张德禄, 2009)。

那么，为什么多模态话语的研究几千年来没有受重视，而现在却突出迅速发展起来呢？本人认为主要有两个原因：一是语言学理论研究的发展趋势使然；二是现代科学技术的发展，以及全球化、信息化、多样化的发展趋势使然。从20世纪70年代开始出现话语分析理论，到了90年代已经蓬勃发展起来。但话语是在语境中产生的，语境中的其他模态也参与交际事件，所以，只从语言的角度不能窥见社会交际的全貌，不能很好地分析语篇的意义，语境依赖强的意义自然需要通过其他模态来解释，这样话语分析就扩展为多模态话语分析了。

而第二个理由更加重要。现代科学技术的发展大大扩大了交际的方式，使交际方式发生了革命性的变化。以前的以口语和书面语为主的交际现在扩展为多种交际方式并行的模式。从口语来说，以前主要通过面对面谈话，后来扩展为电话通讯，但现在可以通过虚拟空间、视频对话进行。另外，微信、QQ、Facebook等多种交际渠道都兼有口语和书面语的特点。其中，语言之外的模态起了很大的作用。从书面语来说，以前书面语的主要载体是页面：作者只能从一页页纸上表达自己的意义。文字成为主要的媒介，因为它的产出相对于图像、绘画等要容易些。但现在的主要交际媒介是屏幕。屏幕的形状虽然像页面，但设计的逻辑和方

式不同。屏幕可以很容易地展示文字、图像、动画；屏幕是动态的，可以迅速和连续性地切换页面，可以播放声音等，是多模态交际的理想媒介。科学技术上的革命需要新的理论来探讨、分析、解释和评判，多模态话语分析理论应运而生。

Jewitt(2009)区分了多模态研究的三个视角：社会符号学分析(Kress & van Leeuwen, 1996; van Leeuwen, 2005)、系统功能语法分析(O'Toole, 1994; O'Halloran, 2005)和社会互动分析(Norris, 2004)。这三个视角虽然略有差别，但都属于社会功能分析(Bateman & Schmidt, 2012)。同时，Jewitt也忽略了另外两个重要研究视角，即以认知隐喻理论为基础的多模态隐喻研究(Forceville, 1996)和语料库语言学视角的多模态文本分析(Gu, 2006; 2009)。另外，还有神经语言学、人机交互及学习研究中的多模态理论模式。

多模态语言研究的前景十分广阔，可以涉及语言学的多个研究领域。但从语用学的角度进行研究的还很鲜见，从语料库的角度研究的也不多。在这一背景下，本研究从日常会话的多模态交互本质出发，重新考察语用学等传统课题和经典理论，对其进行验证、修补和发展，尝试采用多模态语料库语言学(Multimodal Corpus Linguistics)范式对语用学课题进行研究。从这个角度讲，本书可以说填补了这个领域的一个研究空白。

在语用学中，言语行为(Speech Act)是经典研究课题，被视为语言交际的基本单位，相关研究在人机对话、人工智能等方面具有广阔的应用前景，由此成为目前语言学界重点关注的研究领域之一。但以前的研究集中在语句上，根据说的话来探讨语句的言语行为、语力和言后行为，特别是语力，需要从语言语句中推测出来。这种研究模式一直是语用学对语力研究的基本模式。后来，Mey(1993/2003)等把研究言语行为扩展到研究语用行为(Pragmatic Act)上，说明他已经注意到这个问题，并开始注重行动研究，但还是没有把多种模态的协同作用作为研究的对象。

本书一方面从理论上将多模态话语分析理论与语用学相结合，扩展了多模态话语分析的研究领域，更重要的是，本研究还起到把语用学进行“升级”的作用，一方面使语用学的研究领域扩展到多模态研究上，扩展了语用学的研究范围，同时，又对经典语用学研究的对象从多模态的

角度进行重新审视，使原来大多是通过推测得来的结论通过更多的多模态事实来证实。另一方面，从实践上，本研究把现时农村大众的话语进行了记录、标注和分析，建立了语料库，也具有比较强的现实意义和实践意义。

从研究思路和方法上，本书运用多模态语料库语言学的基本方法，关注说话人在现场即席话语(Situated Discourse)中如何运用多种表达手段表达语力这一问题，以贴真建模(Simulative Modeling)思想为基本研究思路，以汉语现场即席话语多模态语料库为数据来源，以包括多模态研究、语用学、语料库语言学、语音学等多个学科分支的知识和技术拓展研究视野，考察语力与情感状态、韵律特征、体貌表现之间的关系，分析说话人在不同情感状态下，韵律特征、体貌表现如何帮助说话人表达语力，描写说话人在表达不同语力时各种手段之间的互动协同规律，现场即席话语中语力显示项的特点、呈现规律及其影响因素。同时，他运用了定性和定量相结合的研究方法对语力进行了研究，对鲜活语力进行了建模。

本人认为，本研究具有以下几个突出的特点：

1. 创新性：作者从一开始就站在学科的前沿上，去开拓前人没有开拓或者很少开拓的领域。作者能够看到在多模态话语在话语分析领域轰轰烈烈开展的同时，与其关系十分密切，并同时苦于找不到解决自己理论中的诸多问题(如非语言现象和行动等表达语力的问题)的解决方案的语用学，却一直观望，没有参与到这项比较重要的语言学变革中，并选择了把语用学论题研究与多模态话语分析理论相结合的方法来探讨语力问题，从而在一定程度上对语用学进行“升级”尝试，使两个研究方向都得到了扩展。

2. 贴真性：本研究采用贴真建模的方法来对语料进行研究，使研究的范式尽量贴近客观实际，在实际的语境中进行信息采集，在此基础上选取一定的研究视角，挖掘数据，开展研究。贴真建模是科学研究中心常用的方法之一。这种方法得出的研究成果更加符合客观现实，更能反映人类交际的实际，说服力更强。

3. 实证性：与贴真性密切相关的是实证性。本文的论证都是建立在真实语料的基础上的，而不仅仅是理论推导和简单的实例分析。作者用了7个月的时间从农村采集12个文盲人的话语，长度达到28.98小时

(合计 1 739 分钟),然后进行整理、分类、标注,根据需要进行了定性和定量分析。

4. 跨学科性:本书研究的重点是语用学研究,探讨话语的语力,但该研究涉及的理论和方法是跨学科的。在语言学内部涉及语用学、多模态话语分析、语料库语言学、语音学和音系学等,在语言学之外涉及社会学、人类学、声学、资料学、人工智能、生理学、行为学等许多学科,这些不同学科的相关知识是对研究者的一个挑战,同时也说明了研究者个人的学术修养。

总之,本书具有相当强的理论意义和实践价值,对于同类研究具有较好的借鉴和参考价值,期望得到广大读者的喜爱。

张德禄

2017 年 1 月 29 日于同济园

序

(二)

v

试读结束: 需要全本请在线购买: www.ertongbook.com

前言

如今，“多模态”已成为横跨自然科学、工程技术以及人文社科领域的热门词，涉及的学科也包括语言学、符号学、人类学、社会学、生理学、医学、人工智能、教育学等。

多模态语言研究认识到人类语言交际与发展均需要多模态感官系统支撑这一事实。人类的语言交际本质上是多模态的，除词句结构和话语内容之外，我们还可以通过包括韵律、表情、动作等非语言媒介进行意义构建。但限于以往的研究视野和技术手段，传统上对言语交际的研究只关注语言本身的使用，而忽略了意义是由多模态资源构建的事实。

多模态研究是一种整合了多种路径的研究范式。在语言研究及其相关的领域中，这些路径包括：符号学属性的多模态研究、多模态语料库语言学研究以及神经语言学、人机交互和学习科学中的多模态研究。多模态语言研究的前景十分广阔，可以涉及语言学的多个领域。其中就包括从人们日常会话的多模态交互本质出发，重新考察语用学等传统课题和经典理论，对其进行验证、修补和发展。

在我国，符号学属性的多模态研究成果较为丰富，涉及多模态话语分析、多模态隐喻研究、多模态文体学、多模态教学等；神经语言学、人机交互及学习科学中的多模态研究涉及多个学术领域，各类成果也相对突出。相比之下，基于多模态语料库的研究相对较少，而该领域的研究在国外方兴未艾。

在这一背景下，笔者多年前决定尝试采用多模态语料库语言学(Multimodal Corpus Linguistics)方法对语用学课题进行研究。在语用

学中,言语行为(Speech Act)是经典研究课题,被视为语言交际的基本单位,相关研究在人机对话、人工智能等方面具有广阔的应用前景,由此成为目前语言学界重点关注的研究领域之一。因此,笔者决定对言语行为进行多模态考察。

多模态语料库是指文字、音频和视频语料等多种信息集成,研究者可以通过多模态方式加工、检索和统计而进行研究的语料库,堪称“语料库4.0”。多模态语料库的语料采集、转写、切分、标注以及建库与传统文本语料库不尽相同。语言研究要以人们与外界的多模态互动作为重要的数据来源,面对具有充盈意义的现场即席话语,多模态语料库语言学是一种较为理想的研究方法。运用多模态语料库语言学方法对语用学课题进行研究具有重要意义,可以拓展语用研究视野,发展相关理论,并成为目前国际语用学界的前沿领域。

本书运用多模态语料库语言学的基本方法,关注说话人在现场即席话语(Situated Discourse)中如何运用多种手段表达语力(Illocutionary Force)这一问题,以贴真建模(Simulative Modeling)思想为基本研究思路,以汉语现场即席话语多模态语料库为数据来源,以包括多模态研究、语用学、语料库语言学、语音学等多个学科分支的知识和技术拓展研究视野,考察语力与情感状态、韵律特征、体貌表现之间的关系,分析说话人在不同情感状态下,韵律特征、体貌表现如何帮助说话人表达语力,描写说话人在表达不同语力时各种手段之间的互动协同规律,现场即席话语中语力显示项的特点、呈现规律及其影响因素。

本书认为,言语交际本质上是一种复杂系统,需要在复杂系统范式下加以审视,可以借用自然科学与工程领域的贴真建模作为核心思路,参考“言思情貌整一原则”(STFE-Match Assumption)构建建模视角,对现场即席话语中的鲜活语力进行建模。在混合研究法的启示下,本书对语力进行了定性和定量相结合的考察。

在构建多模态语料库的基础上,本书对不同语力类型进行了概念分析,采用集合形式对这些语力类型的本体性质、实施条件与规则等进行呈现,建立语力的概念模型;结合具体情境对各个类型对应的语力实例进行了个案分析,并对不同语力实例的差异性和共同性特征进行了展现;绘制了说话人在各种情境下语力表达手段的资源系统,并分析其相应功能承载;利用统计学方法描述说话人表达不同语力实例时,情感

状态与韵律特征、体貌表现之间的互动规律；对说话人语力表达的多模态手段和语力显示项呈现的影响因素进行了推测。

同时，本书初步讨论了多模态语用学的构建必要与研究逻辑，认为多模态研究可以在研究方法上对语用学研究进行升级，这对研究者从人们普通行为的高度更为全面、准确地解读语言使用意义十分关键。

另外，本书还对多模态语言研究和相关领域中不同的多模态研究路径及其关系进行了论述，并对多模态语料库语言学的可能选题进行了展望。现场即席话语研究及对应的多模态研究方法，拓展了语言研究的视野，具有深刻的理论意义和广阔的应用前景，应当成为未来语言研究的新兴领域之一。

本书既介绍了相关领域的国内外研究现状，也全面论述了笔者本人的研究成果，适合希望在多模态、语用学、语料库等领域开展前沿研究的师生阅读。

黄立鹤

2017年2月6日于同济园

前

言

目 录

第1章 绪论.....	1
1.1 言语行为理论与语力研究回顾	1
1.1.1 言语行为理论的建立及分析哲学家的早期研究.....	2
1.1.2 言语行为研究中语言学家的加盟.....	6
1.1.3 基于语料库的言语行为研究	11
1.2 本书的研究内容与目标	17
1.2.1 主要研究内容	17
1.2.2 主要研究目标	18
1.3 本书的研究路径与方法	18
1.3.1 以贴真建模思想为基本研究思路	18
1.3.2 以多模态语料库为研究数据来源	19
1.3.3 以多学科背景拓展语力研究视野	19
1.4 本书的结构安排	19
1.5 本章小结	20
第2章 现场即席话语与多模态语料库.....	22
2.1 现场即席话语与多模态语言研究	22
2.1.1 现场即席话语	22

II	2.1.2 本研究中多模态的定义	24
	2.2 现场即席话语与当下亲历、当下认知及多模态数据	26
	2.3 多模态语料库简介	29
	2.3.1 多模态语料库基本情况	29
	2.3.2 多模态语料库的创建	31
	2.4 本章小结	42
	第3章 现场即席话语语力研究的核心思路与理论准备	43
	3.1 贴真建模思路的运用	43
	3.1.1 复杂系统观下的言语交际	43
	3.1.2 贴真建模的含义	46
	3.1.3 贴真建模的步骤及在本研究中的体现	48
	3.2 话语活动的研究思路与建模对象	51
	3.2.1 社会场景在本研究中的地位与处理方法	52
	3.2.2 体貌表现在本研究中的地位与处理方法	55
	3.3 言思情貌整一原则与现场即席话语语力的分析视角	56
	3.3.1 言思情貌整一原则与贴真建模	56
	3.3.2 言、思、情、貌四个视角的意义	59
	3.4 本研究语力实例的分类、依据及其意义	60
	3.4.1 分类及其依据	60
	3.4.2 分类效果与意义	61
	3.4.3 语力实例的利益关系分析	64
	3.5 本章小结	65
	第4章 现场即席话语语力的发现程序	67
	4.1 发现程序的制定原则、依据与特征	68
	4.1.1 根本原则	68
	4.1.2 基本依据	68
	4.1.3 整体特征	69
	4.2 发现程序的构架与内涵	69
	4.2.1 概念模型	70

4.2.2 情境因素	78
4.2.3 情感状态	81
4.2.4 韵律特征	92
4.2.5 体貌表现	99
4.3 本章小结	102
第5章 多模态语料的采集与加工	104
5.1 本研究语料的采集情况	104
5.1.1 采集对象与方法	104
5.1.2 采录设备	106
5.1.3 采录人员	107
5.1.4 关于语料采集的两个说明	107
5.2 原始语料资源及语料库的构成与加工	108
5.3 语料转写说明	112
5.4 本章小结	112
第6章 现场即席话语多模态语力库建设	114
6.1 现场即席话语语力的切分与标注方案及其实施	114
6.1.1 语力操作单元的工作定义、切分标准与标注方法	115
6.1.2 活动类型的工作定义、切分标准与标注方法	121
6.1.3 话轮的工作定义、切分标准与标注方法	122
6.1.4 情感状态的工作定义、切分标准与标注方法	124
6.1.5 韵律特征的工作定义、切分标准与标注方法	134
6.1.6 体貌表现的工作定义、切分标准与标注方法	143
6.1.7 意图状态的工作定义、切分标准与标注方法	146
6.1.8 依存性的工作定义、切分标准与标注方法	148
6.1.9 元数据与标注语言	149
6.2 概念建模、数据建模与鲜活整人建模视角的对应关系	149
6.2.1 鲜活整人建模视角与语力数据建模的关系	150
6.2.2 语力概念建模与数据建模的对应	151