



现代汉语框架语义网

刘开瑛 由丽萍 著



科学出版社

国家科学技术学术著作出版基金资助出版

现代汉语框架语义网

刘开瑛 由丽萍 著

科学出版社

北 京

内 容 简 介

本书面向自然语言理解的需要,以框架语义理论为基础,从资源建设、自动标注到应用技术,对汉语框架语义网的相关工作进行探索和实践。介绍汉语框架语义资源建设工作,包括框架语义理论、框架语义词典和语料库的构建以及该知识库的语义 Web 语言表示体系等;通过大量语料分析和统计,分析汉语语言的规则形式和特点;展示汉语框架语义网本体表示、推理、集成及其应用方面的研究工作;包含汉语词元识别模块、未知框架检测模块、框架消歧模块和语义角色标注模块等技术内容。

本书可供人工智能、自然语言理解、机器翻译、情报学、语言学等相关专业的科研人员和技术开发人员学习、参考。

图书在版编目(CIP)数据

现代汉语框架语义网/刘开瑛,由丽萍著. —北京:科学出版社, 2015.6

ISBN 978-7-03-044558-2

I. ①现… II. ①刘… ②由… III. ①汉语-语义网络 IV. ①TP18

中国版本图书馆 CIP 数据核字 (2015) 第 124808 号

责任编辑:张艳芬 余 丁 董素芹/责任校对:桂伟利

责任印制:张 倩/封面设计:陈 敬

科学出版社出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

三河市燕喜印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2015 年 6 月第 一 版 开本: 720 × 1000 1/16

2015 年 6 月第一次印刷 印张: 16 1/2

字数: 320 000

定价: 95.00 元

(如有印装质量问题,我社负责调换)

前 言

自然语言理解是计算机科学领域与人工智能领域中的一个重要方向，其研究目标是实现人与计算机之间用自然语言进行有效通信。一个自然语言理解系统，就是将自然语言句子分解为机器能够处理的结构化数据，这需要一部机器词典和一个包含语言规则的语言解析器。早期对自然语言处理的研究，其分析深度主要为词语和句法层面，包括对词法、句法自动分析器的研究和句法标注语料库的构建等，这类研究现在已经相对成熟。而语义层面的研究尚处于起步阶段，但它们对于自然语言理解更关键，也更困难，目前是制约自然语言处理研究向前发展的主要瓶颈。在考察了诸多语义理论和语义知识库成果的基础上，本书选择菲尔墨的框架语义学为理论支撑，开展一个汉语框架语义知识库构建工程——现代汉语框架语义网（以下简称汉语框架语义网）。以框架为单位对词语进行语义分类，分类细致、合理，类与类之间的联系多样，能够支持复杂的推理，不限定论元角色的种类和数量，以细化的角色类型描述词语的论元结构，能够精准地描述句子的语义关系，能够满足计算机处理复杂语义信息的需要。工作主要包括词汇的框架语义描述、语义标注语料库建设和语言规则的研究等，本书对此进行系统的介绍和分析。

第1~4章介绍汉语框架语义知识库构建方面的相关工作。首先，介绍框架语义理论，然后分析和确定汉语框架语义知识库的构建原则和工作路线，并展示具体的构建结果；最后对知识库的语义 Web 语言表示体系进行设计和实现，使之成为机器可读、可理解的语义资源。汉语框架语义知识库的建设是工作的重点，包括框架库、词元库和例句库三部分。在框架库中，以“框架”为单位对汉语词汇进行分类汇编，细化定义每一类词语的论元角色体系，并建立框架与框架之间的关系网络。词元库以词语为描述单元，明确词语的语义内容，并根据语料标注结果提取其论元角色的句法结构模式。从深层语义结构到表层句法结构之间的转换关系正是语义学一直关注的议题，词元库使之在语料库和计算机技术的帮助下得以精确、细致地展示。例句库建设的主要任务是将框架库定义的语义分类信息和语义角色信息标注在真实语句上，但是，为达到计算机自动语义分析的需求，还必须对语义角色的句法信息进行标注，因此，实际上是短语类型、句法功能和框架语义信息三层标注。语义知识库作为机器可读词典和语料库，为汉语的自动语义分析提供基础资源。

第5章介绍面向中文阅读理解而开展的汉语框架语义标注文本库建设工作,即选取特定领域文本,进行全文标注。框架语义知识库中的例句标注针对给定目标词,仅标注句子中与该目标词有句法从属关系的成分,是以该词语所在的论元结构为单元的标注。而文本标注以文本中每一个句子、句子中的每一个框架承担词为标注目标,表示出句子所含的全部论元结构。一般来说,用户关心的事件和关系往往散布于文本的不同位置,其中涉及的实体通常可以有多种不同的表达方式,并且还有许多事实信息隐含于文本之中。为了准确而没有遗漏地从文本中抽取相关信息,信息抽取系统必须能够识别文本中的共指现象,进行必要的推理,以合并描述同一事件或实体的信息片段。这样的标注结果能够真实反映标注信息的分布情况,具有统计意义。

第6章和第7章研究汉语句法分析规则描述方面的相关工作。汉语框架语义网的研究最终是要实现汉语语义分析,而语义分析一定要以句法分析为基础,诸多中文处理的应用课题最终都归结到汉语分析的基本问题上,那就是句法和语义的分析。因此,语言规则是语言解析器不可缺少的部分,也是影响分析结果的关键。为实现自动语义分析,需要汉语的句法规则和语义规则。本书通过大量语料分析和统计,分析了汉语语言规则形式和特点。

第8章和第9章研究汉语框架网本体的构建、查询技术及其应用。本体作为一种能在语义和知识层次上描述信息系统的概念模型建模工具,在计算机的许多领域得到了广泛应用。语言学本体是指关于语言、词汇、语法、语义等的本体,它独立于任何一个领域,对人们通常使用的一些概念构造出一个通用的知识框架结构。本书的研究集中在汉语框架语义网这一语言学本体的表示、推理、集成及其旅游应用方面。应用系统的目标是获取、描述和表示相关旅游领域的知识,提供对该领域知识的共同理解,确定该领域内共同认可的词汇,并从不同层次的形式化模式上给出这些词汇(术语)和词汇间相互关系的明确定义。

第10章对汉语框架语义网自动识别技术进行初步探索。语义分析研究的最终目标是实现大规模真实文本语义信息的自动标注,汉语框架语义分析器主要包含汉语句法结构分析器和词汇语义标注器两部分内容。首先构建汉语框架语义分析模型,遍历词法和句法分析、现代汉语框架语义网语义角色标注和核心依存图分析三大流程阶段。包含汉语词元识别模块、未知框架检测模块、框架消歧模块和语义角色标注模块等技术内容。

刘开瑛 由丽萍

2014年10月

目 录

前言

第 1 章 框架语义学及汉语框架语义网	1
1.1 框架语义学简介	1
1.1.1 基于框架的语义理解	1
1.1.2 框架元素的句法实现方式	4
1.1.3 框架语义学和其他语义学的关系	7
1.2 框架语义网	8
1.2.1 框架语义网概述	8
1.2.2 FrameNet 的内容	11
1.2.3 FrameNet 的特色	16
1.2.4 FrameNet 的应用价值	17
1.3 汉语框架语义网构建路线	17
1.3.1 自底向上的构建路线	17
1.3.2 语料库建设	19
参考文献	19
第 2 章 汉语框架语义网框架库的构建	21
2.1 认知领域框架	21
2.1.1 认知领域的界定及词语采集	21
2.1.2 汉语认知领域框架示例	22
2.1.3 认知领域框架-框架关系	32
2.1.4 认知领域框架解析	34
2.1.5 认知领域框架元素解析	36
2.2 位移领域框架举例	38
2.3 名物类框架举例	40
2.4 修饰性框架举例	42
参考文献	43
第 3 章 汉语框架语义网例句标注规范	44
3.1 标注范围	44
3.1.1 目标词 tgt 和框架元素标注	44

3.1.2 几类不用标注的情况	47
3.2 支撑词标注	48
3.2.1 标注范围	48
3.2.2 动词目标词的支撑词	48
3.2.3 名词目标词的支撑词	50
3.2.4 形容词目标词的支撑词	50
3.2.5 支撑词标记标注时应注意的问题	51
3.3 透明成分标记	51
3.4 框架元素的几种特殊标记	52
3.4.1 零形式框架元素	52
3.4.2 框架元素融合	53
3.4.3 框架元素分裂	53
3.5 句法功能标注体系	53
3.5.1 主语的标注规则	55
3.5.2 宾语的标注规则	56
3.5.3 外部论元的分析与标注规则	57
3.5.4 定语的标注规则	59
3.5.5 状语的标注规则	59
3.5.6 补语的标注规则	59
3.5.7 中心语的标注规则	60
3.5.8 兼语与兼语补语的标注规则	60
3.5.9 连谓结构的标注规则	60
3.5.10 插入语和被插入语的标注规则	62
参考文献	63
第4章 汉语框架语义网词元库的构建	64
4.1 汉语框架语义网的词元	64
4.1.1 名词词元	64
4.1.2 动词词元	65
4.1.3 形容词词元	68
4.2 多义词与框架的对应关系	68
4.3 词元的标注报告	72
第5章 汉语阅读理解系统的构建	79
5.1 阅读理解系统研究现状	79
5.2 语料库的主题分类与指称指代标注	80
5.2.1 文章的筛选	80

5.2.2 文献的主题分类标引	81
5.2.3 指称指代关系标注	82
5.3 目标词的确定和框架语义标注	83
5.3.1 目标词的确定	83
5.3.2 框架语义标注	84
5.3.3 目标词和指称指代关系标注举例	85
5.4 编写篇章问句和标注问句答案	87
5.5 知识获取和知识库构建方法	90
5.5.1 阅读理解技术对语料库的需求分析	90
5.5.2 篇章标注	91
5.5.3 语义知识库的构建	91
5.6 阅读理解任务评测机制	93
参考文献	95
第 6 章 汉语短语结构分析及实例库	96
6.1 复杂特征集与功能合一语法	96
6.2 短语结构规则形式分析与表达	97
6.2.1 短语结构规则的组成	97
6.2.2 短语结构规则的整体性质	98
6.2.3 短语规则的约束条件	100
6.2.4 短语规则的各项属性汇总	100
6.2.5 短语规则实例解析	101
6.2.6 短语结构规则统计分析	103
6.3 短语组合模式与实例库	106
6.3.1 np 短语组合模式	107
6.3.2 vp 短语组合模式	109
6.3.3 ap 短语组合模式	112
6.3.4 mp 短语组合模式	114
6.3.5 pp 短语组合模式	115
6.3.6 sp 短语组合模式	116
6.3.7 tp 短语组合模式	117
6.3.8 dp 短语组合模式	118
参考文献	119
第 7 章 汉语句法分析及实例库	120
7.1 事件描述单元	120
7.1.1 事件描述单元的定义和特征	120

7.1.2 事件描述单元的句法结构统计	121
7.2 单句型短语的组合模式及实例库	124
7.2.1 由体词性成分做谓语	124
7.2.2 由形容词性成分做谓语	125
7.2.3 由动词性成分做谓语	126
7.2.4 由单句型短语做谓语	127
7.3 单句的组合模式及实例库	128
7.4 复句型短语的组合模式及实例库	129
7.5 语料库中复句的统计与分析	132
7.5.1 语料库中复句的统计	132
7.5.2 复句关系举例	133
7.6 语料库中出现的几类句法类型及其标注	135
7.6.1 整句	135
7.6.2 直接引语	136
7.6.3 独立成分	136
7.6.4 句群	136
参考文献	137
第 8 章 汉语框架语义网本体和语义网标记语言	138
8.1 本体	138
8.1.1 本体定义及其要素	138
8.1.2 本体的类型	139
8.1.3 语言学本体	141
8.1.4 领域本体的作用	141
8.1.5 构建本体的准则	142
8.2 语义 Web 标记语言	142
8.2.1 语义 Web 标记语言概述	142
8.2.2 可扩展标记语言	144
8.2.3 资源描述框架语言	147
8.2.4 语义网描述语言 OWL	153
8.3 汉语框架语义网本体语义 Web 表示方法	159
8.3.1 框架库语义 Web 表示	159
8.3.2 词元库的语义 Web 表示	161
8.3.3 句子库的语义 Web 表示	162
8.4 本体开发环境与工具	162

8.4.1	Java 语言	162
8.4.2	Jena 工具包	165
8.4.3	JavaScript 脚本语言	166
8.4.4	本体构建工具 Protégé	167
	参考文献	168
第 9 章	旅游领域本体语言查询技术	169
9.1	旅游领域本体知识库的构建	169
9.1.1	基于《中国分类主题词表》的旅游领域知识描述	169
9.1.2	旅游语料收集	171
9.1.3	山西旅游本体库构建	172
9.2	中文问句的句法与语义分析	176
9.2.1	问句的句法分析	176
9.2.2	问句的语义解析和语义扩展	179
9.3	中文问答系统答案库构建	182
9.3.1	问句与答案句的内在联系	182
9.3.2	句式转换	184
9.4	中文问答句形式化描述	185
9.4.1	语义网查询语言 (OWL-QL)	185
9.4.2	中文问答句形式化描述举例	187
9.5	复杂问题分解的本体汉语框架语义网	189
9.6	基于本体的语义网检索模型	191
9.6.1	基于语义网的语义检索模型	191
9.6.2	模型实现的关键技术	194
	参考文献	197
第 10 章	汉语框架语义网自动识别技术	199
10.1	汉语框架语义分析模型	199
10.2	目标词识别技术	202
10.2.1	未登录目标词识别流程	202
10.2.2	未登录目标词识别理论依据	203
10.2.3	最大熵模型特征模板设计	204
10.2.4	实验结果与分析	206
10.3	汉语框架歧义消解	209
10.3.1	框架歧义消解研究现状	209
10.3.2	最大熵模型在框架识别中的应用	210

10.3.3 汉语框架消歧模型的实验	211
10.3.4 实验结果与分析	213
10.4 汉语的语义角色自动标注	216
10.4.1 汉语的语义角色自动标注现状	216
10.4.2 汉语框架语义角色标注任务描述	217
10.4.3 条件随机场模型标注	218
10.4.4 实验语料处理方案	222
10.5 汉语框架语义依存图分析	226
参考文献	228
附录 1 汉语框架语义网框架汇总表	231
附录 2 通用非核心框架元素	241
附录 3 语料库汉语词类标记集	244
索引	247
后记	251

第1章 框架语义学及汉语框架语义网

1.1 框架语义学简介

框架语义学 (frame semantics)^[1-4]是由美国语言学家菲尔墨 (Fillmore) 提出的研究词语意义和句法结构意义的一种理论方法, 即试图用经验主义方法, 寻找语言和人类经验之间的紧密关系, 并研究一种可行的描述方式, 表示语言和经验之间的关系。词语代表了经验的分类 (范畴化), 范畴中的每一个词语背后都有一定的基于知识和经验的激活性情境, 框架语义学试图理解一个言语社团为何会创建一个范畴并以词语来表达它; 同时通过对背景动因进行描述, 来解释这个词语的意义。该理论出自经验主义语义学传统, 而不是形式语义学, 它和形式语义学的差别主要在于, 前者强调语言与经验之间的连续性 (continuity), 而后者是非连续的。但它与形式主义又并不矛盾, Fillmore 表示, 框架语义学应该说是“先于形式化的 (pre-formal)”, 而不是“非形式化的 (non-formalist)”。

1.1.1 基于框架的语义理解

框架语义学把词义、句子意义和文本意义统一用框架 (frame) 进行描述, 框架是和一些激活性语境 (motivating context) 相一致的一个结构化的范畴系统, 是存储在人类经验中的图式化情境 (scenario), 这种范畴系统所描述的既可能是一个实体, 也可能是一种行为实践模式, 甚至是一些社会制度、习俗等。Fillmore 使用框架这个术语来涵盖以往讨论自然语言理解问题的文献中提过的一系列概念, 如场景 (scene)、图式 (schema)、脚本 (script)、情境、观念框架 (ideational scaffolding)、认知模式 (cognitive model)、民俗理论 (folk theory) 等, 但是这些术语属于人类认知或经验范畴, 框架则是语言学术语, 一个框架虽然表示的是一个认知结构, 但它的意义在于这个认知结构的各个部分是和词语挂钩的, 为语言理解服务。

Fillmore 用场景、透视域 (perspective) 等概念来阐释其框架语义分析系统。场景指语言之外的真实世界, 如物、事件、状态、行为、变化和人们对真实世界的记忆、感觉、知觉等。场景可以是视觉形象, 也可以是任何一个独立的知觉、记忆、经历、行动或物体等。句子描述的是场景, 语言中的一个词、短语、句子

都是对场景的描述。场景中的各参与者承担各角色，通过透视域的选择，一部分参与者进入透视域，成为核心成分，其他参与者不一定能进入句子，即使出现在句子中，也只能成为句子的外围成分。

语义联系着场景，但是场景并不等于语义，每一个句子或者每一段话语，都有一个特定的透视域，从这个特定的视角去考虑一个场景，当整个场景都在考虑之中时，人们一般也只是注意场景的某一部分，因此，透视域在框架语义学中是个比较重要的概念，下面的论述经常会涉及，所以首先介绍这个概念，例句如下：

Carla bought the computer from Sally for \$100.

(Carla 以 100 美元从 Sally 那里买了电脑。)

尽管这句话提及了买者 (Carla)、卖者 (Sally)、钱 (\$100) 和商品 (the computer)，但是表达的时候基本上还是站在买者的视角上进行的。同样：

Sally sold the computer to Carla for \$100.

(Sally 以 100 美元卖给 Carla 电脑。)

这句则是从卖者的视角进行表达的。下面可以看包含 *buy* 和 *sell* 的另外两个句子：

I bought this thing[in order to test it out].

([为了检验一下]，我把那个东西买了下来。)

At least you are selling them[in order to give the money to cancer research].

(至少你是[为了筹钱做癌症研究]才卖了它们。)

第一句所表达的目的是和买者有关的，而第二句的目的是和卖者有关的，可见，理解 *buy* 和 *sell* 两个词语所支撑的句子，虽然都需要在整个商品交易这个大背景下，但是视角是不同的。

文本中的所有语言成分都可以激活框架，理解一个文本，就是搜索或感知文本内容所激活的框架，并对它们进行综合的过程。先来看 Fillmore^[3]列举的几个以实体、社会制度或风俗为框架的词语的理解过程，最后再看它们是如何综合的。

(1) *decedent* (死者)。理解该词语不能仅理解为“死去的人”，也不能限于一般的法律或新闻语境，而是要将视角置于法律子领域遗产法中去理解。一般的法律语境的死者可以用 *the deceased* (死者) 来表达，而 *decedent* 用在讨论一个死者的遗产时。

(2) *land* (陆地)。和 *ground* (地面) 相比，*land* 用来指“陆地”，与 *sea* (海洋) 相对；*ground* 用来指“地面”，与 *air* (空中) 相对。因此，如果描写某人 *on land* (在岸上)，就是将视角定位于一个航海事件中，只有在海上行驶过程中，才会这样描述某人所在的位置。

(3) *mufti* (便衣)。*mufti* 是指军队中通常穿制服的人穿的普通衣服，如果看到一个军官和一个普通百姓穿同样的衣服，不会说那个百姓 *in mufti* (穿便衣)，

这个词语只能用来描述那个军官, *mufti* 显然是限于和军队有关的语境的。

(4) *weekend* (周末)。理解这个词语需要以一定的社会制度和实践经验为背景, 即日历七天为一周, 一周内有五天连续时间是公众工作时间, 而有连续两天是私人生活时间。如果只有一个休息日, 可能不需要 *weekend* 这个词, 因为人们可以简单地使用表示那一天的名词 (如 *Sunday* (星期天)) 来指称它。如果人们三天工作, 四天休息, 似乎也不可能用这个名称来表达人们用来过私人生活的这段时间。

(5) *breakfast* (早餐)。这也是一个以社会制度和实践经验为背景的词语, 理解这个词, 就是要理解三方面内容: ①人们的文化生活中的一个行为 (实践活动), 即一日三餐; ②早餐是在睡完一觉之后的一天中较早的时候进行的; ③它有独特的菜单 (这视民族的不同而有许多细节差异)。关于 *breakfast*, 最有趣的是, 这三个条件中的任何一个都可以单独缺席, 仍然允许说话者使用这个词, 这就涉及“原型 (prototype)”的概念, 一个人可以工作一个通宵, 然后在日出时吃几个鸡蛋、烤面包片, 喝咖啡、橘子汁, 并把这些称为 *breakfast*, 显然, “睡觉后”这个范畴特征不是强制性标准; 一个人也可以从早晨开始睡, 到下午三点起床, 起来吃鸡蛋、烤面包片, 喝咖啡和橘子汁, 也可以把这称为 *breakfast*, 显然, “早晨”这个范畴特征也不是一个强制性标准; 一个人还可以晚上睡觉, 早晨起床, 喝点卷心菜汤, 吃点巧克力派, 作为 *breakfast*, 显然, “早餐菜单”也不是一个强制性标准。当观察这些现象时, 不是没有抓住这个词语意义的核心, 而是词语给了人们一个范畴, 该范畴可以用在很多不同的语境中, 这些语境由一个典型用法的多方面来决定, 其用法或多或少和定义的原型相匹配, 但不一定完全匹配。

(6) *flip strength* (狂热强度)。理解这个词组, 需要有特殊的背景, 即色情文学。一些色情小说出版商用这个词组告诉作者, 每一页中都要包含一定数量的高“性”趣词, 这样, 一个潜在的顾客在书店中翻 *flip* 书时, 不管他从哪里翻开, 都会发现这本书中充满了令他兴奋的东西。一本书中, 如果每页淫秽词语的比例都比较高, 那么它有比较高的 *flip strength*, 反之则 *flip strength* 比较低。

以上分别用单个的词或词组, 解释了背景框架的知识。文本中的词汇在理解者大脑中会唤起如上认知框架, 认知框架提供理解期待, 带来角色、目的、事件类型的自然的和惯例性的结果, 理解者把这些通过这种预期的抽象结构知识加以综合, 得以正确理解语言。Fillmore 自造了一个句子 (Fillmore 称“这个自造的句子多少有些可笑的意味”), 说明人们是如何理解语言的:

The decedent while on land and in mufti last weekend ate a typical breakfast and read a novel high in flip strength.

(死者于上周末在陆地上, 穿着便衣, 吃了顿普通的早晨, 读了一本狂热度高的小说。)

通过以上背景知识,理解者就可以知道这是在谈论一个已经死亡的海军公务人员,他在上周六到周日这段时间读了一本色情小说,理解者还知道关于他的其他一些事情,如他是如何打发时间的等。这个句子并没有直接给人们这些信息,必须通过对以上背景框架进行综合,才能真正理解其意义。

1.1.2 框架元素的句法实现方式

Fillmore 和后来的研究者^[5]用这一理论比较多的情况是描述那些由动词激活的过程性框架,关注词语激活的活动场景是什么,这个活动场景中有哪些参与者,即框架元素(frame element, FE),各个框架元素又是如何在句子中实现的,下面具体进行说明。

要理解词语的意义,首先必须具备语义框架的知识,框架为词语在语言和言语中的使用提供背景和动因。例如,动词“写”描写的是这样一种场景:一个人在某个物体的表面握着一个顶部尖锐的工具使其进行移动,在物体表面留下痕迹。在这个场景中有4个实体(4个框架元素):发出这个行为的人、实施这个行为所凭借的工具、承受这个行为的物体表面、这个行为在物体表面留下的痕迹。这是在没有上下文时,单独一个动词“写”所描述的全部场景,也就是当没有遇到任何其他上下文条件时,一个单独的动词“写”所产生的全部想象。句子的功能在于突出框架的某些方面,如“小王正在写”突出了书写人小王这一个实体,但是理解者根据经验,仍然会完整地理解其意义,即小王正在握着一支笔,在某一物体表面移动,并且在物体表面留下痕迹,这个场景仍然有4个实体:书写人(小王)、书写工具(笔)、物体表面(纸)、在物体表面留下的痕迹(字)。同样,如“小王正在写信”,突出了书写人(小王)和在物体表面留下的痕迹(信)两个实体,但是仍然可以引出背景场景中的其他两个实体:书写工具和物体表面。可见,词语在理解者大脑中激活了一个完整的认知框架,使理解者带着这个预期的抽象结构,填入惯例性的内容,得以完整理解句子表达的事件,这种隐含语义的理解,正是框架语义学适合智能化的语言信息处理技术的一个重要原因。

书写活动,更概括一点,实际上是以人与人之间的交际、交流活动为背景的,交际活动可以是打电话、直接对话,也可以是广播、做报告、发通知等,因此,书写可以看成交际活动的一种。同时,从主体的特征来看,书写和楼房建造、生产活动、文艺创作等同属于一种有意识行为。这种抽象的框架和框架之间的联系,在人类大脑中构成了一个复杂的知识网络,解释人的推理过程。

一个框架涉及的不是个别的词,而是一类词语,它们都具有共同的认知结构,支配相同类型的语义角色。对于语言表达,最重要的就是描述框架元素在句子中与框架承担词的句法关系和框架元素自身的句法特点。下面作为示例,仔细地研

究一下[言谈交际框架 (speech communication frame)]。以下关于[言谈交际]框架的内容,用下划线标识框架元素名称。

[言谈交际]指的是信息传递者 (speaker) 把信息 (message) 传递给信息接收者 (addressee) 的活动。框架中的信息传递者、信息、信息接收者称为框架元素,相关的英语动词包括 say (说), speak (说、讲), utter (说), whisper (低语), talk (谈), tell (告诉), discuss (讨论), mention (提及), ask (问), promise (许诺), order (命令), plead (请求), confess (承认), warn (警告), threaten (恐吓) 等,这些都可以激活[言谈交际]框架。

描述动词的语义性质,就是要呈现其各种配价属性,也就是要描述各个框架元素的各种句法表现,通过研究这些框架元素所出现的短语和句子结构,来确定这些动词的句法和语义搭配可能性。在这些搭配模式中,关注的是各个框架元素如何或是否得到表达、记录框架元素在句法上是必须还是可选的和它们采用什么语法形式得以表达。

(1) 信息传递者一般是以名词性成分做主语的形式出现的;信息接收者,在有些情况下是必须表达的,例如,当 tell 做谓语中心语,并用 that 引导的子句或 to 引导的非限定动词短语来表达信息时,信息接收者就是必不可少的,例如:

They told<me>that she liked me.

I told<him>to leave the room.

但有时,信息接收者又是可选的,如:

He asked<a question>.

He told<a joke>.

信息接收者这个角色的句法实现方式是多种多样的,它可以是间接宾语:

She asked<me>why the meeting had been cancelled.

介词 to 的宾语:

He said it<to me>more than once.

Why don't you talk<to me>?

I've already spoken<to her>about that.

或者介词 with 的宾语:

They discussed that<with me>at some length.

I've already spoken<with him>about that.

除了信息传递者和信息接收者,还可能有关涉 (topic)。有时,关涉被表达为直接宾语,有时则作为介词 about 的宾语,甚至有时根本不表达:

We mentioned<your situation>.

They discussed<your situation>.

We talked<about your situation>.

He said a number of things <about your situation>.

We spoke <about your situation>.

此外，还有一系列相关的框架元素和信息相联系，首先是语音形式信息 (Message_as_phonological_form)，如可以说：

He said “<ouch>”.

但语音形式信息不能出现在以下表达中：

* *He told me “<ouch>”.*

* *He uttered “<ouch>”.*

但是，对于有些动词，可以用同位形式 (appositional form) 表达这种信息：

He said <the word “ouch” >.

He uttered <the word “ouch” >.

但有些动词不能这样表达：

* *He told me the word “ouch”.*

(2) 第二种信息是描写信息 (Message_as_description)，用各种有关语言或话语的概念来表达，例如：

He asked <a question>.

He told <a joke/a story>.

He recited <a poem>.

(3) 第三种信息是内容信息 (Message_as_content)，作为内容的信息，一般用 *that* 从句、*wh*-从句或非限定性短语来表达：

They said <that you weren't ready>.

They told me <that you weren't ready>.

They mentioned <that you weren't ready>.

They confessed <that they weren't ready>.

They warned me <that they weren't ready>.

但不能说：

* *They spoke (to me) that they weren't ready.*

* *They talked (to me) that they weren't ready.*

* *They uttered that they weren't ready.*

* *They discussed that they weren't ready.*

有时候，信息是以 *wh*-从句的形式出现的，但实际表达的内容则更像一个陈述，而不是真的有疑而问：

He said <what was on his mind>.

He told me <where they had hidden the gold>.

当然，信息也可以通过 *wh*-从句（包括 *whether* 从句或 *if* 从句）表达有疑