



目 录

绪论	1
0.1 研究背景	1
0.2 研究问题	2
0.3 研究概述	2
0.4 研究意义	4
0.5 本书结构	5
<hr/>	
第一章 隐喻研究综述	6
1.1 隐喻的界定	6
1.2 隐喻的认知视角	7
1.2.1 转喻	8
1.2.2 明喻	10
1.2.3 常规隐喻与新生隐喻	10
1.2.4 复杂隐喻与基本隐喻	11
1.3 隐喻的语料库视角	12
1.4 人体隐喻研究综述	14
1.4.1 人体隐喻的认知视角	14
1.4.2 人体隐喻的语言视角	18
1.4.3 已有研究对构建模型的启示	23
1.5 小结	25



第二章 隐喻自动识别研究述评	26
2.1 基于规则的方法	26
2.2 基于统计的方法	27
2.2.1 聚类方法	28
2.2.2 有监督的机器学习方法	30
2.2.3 语言学假设与统计方法相结合	31
2.3 对已有研究的总体评价	33
2.4 小结	34
<hr/>	
第三章 本研究使用的资源与研究算法	35
3.1 本研究使用的资源	35
3.1.1 知识库与语料库等资源	35
3.1.1.1 WordNet	35
3.1.1.2 WordNet扩展概念域	40
3.1.1.3 顶层本体	41
3.1.1.4 英语国家语料库	42
3.1.1.5 英语词典	44
3.1.2 第三方研究工具	45
3.1.2.1 Editpad Pro与PowerGrep	45
3.1.2.2 NLTK与最大熵工具包	45
3.2 研究算法	46
3.2.1 用于知识库模块的算法	46
3.2.1.1 N元序列搭配强度算法	46
3.2.1.2 限定人体词序列长度的算法——K值	48
3.2.2 用于机器学习模块的算法	49
3.2.2.1 抽象值算法	49
3.2.2.2 最大熵原理	50
3.3 小结	53



第四章 研究流程	54
4.1 语料的前期准备	54
4.1.1 人体词的选取	54
4.1.2 研究语料的收集与抽样	56
4.1.3 人体隐喻的人工标注	59
4.1.3.1 人工识别流程	59
4.1.3.2 人工标注的难点及解决办法	61
4.1.3.3 人工标注的信度报告	62
4.2 语言特征的提取	63
4.2.1 知识库模块的语言特征	64
4.2.1.1 人体隐喻搭配序列	64
4.2.1.2 人体隐喻始源域对应的词语	70
4.2.2 机器学习模块的语言特征	71
4.2.2.1 语言特征的选取范围	71
4.2.2.2 词语的抽象值	74
4.3 数据分析	75
4.3.1 构建模型	75
4.3.2 模型验证阶段	77
4.4 小结	78
<hr/>	
第五章 英语人体隐喻的语言分析	79
5.1 英语人体隐喻的语义分布	79
5.1.1 英语人体域的语义分布	79
5.1.2 英语人体词的语义分布	81
5.2 英语人体隐喻的语言特征	84
5.2.1 英语人体隐喻语言特征的总体分布	84
5.2.2 英语人体隐喻语言特征的具体分析	86
5.2.2.1 人体转喻的语言特征	86
5.2.2.2 人体域为始源域的隐喻的语言特征	90
5.2.2.3 人体域为目标域的隐喻的语言特征	94
5.3 小结	98



第六章 人体隐喻自动识别模型的构建与验证	99
6.1 检验语言特征应用方式的先导研究	99
6.1.1 测试机器学习语言特征的先导试验	100
6.1.2 测试知识库与机器学习结合方式的先导试验	103
6.2 人体隐喻自动识别模型的构建	104
6.2.1 人体隐喻语言特征知识库	104
6.2.2 机器学习模块	107
6.3 模型的验证结果	111
6.3.1 模型的整体验证结果	112
6.3.2 知识库模块的验证结果	113
6.3.2.1 误判语句分析	113
6.3.2.2 漏判语句分析	114
6.3.3 机器学习模块的验证结果	115
6.3.3.1 模块识别过程的示例	115
6.3.3.2 模块的验证情况	116
6.3.3.3 模块识别的隐喻类型	117
6.4 与其他研究的对比	117
6.5 小结	119
<hr/>	
第七章 结论	120
7.1 研究发现	120
7.1.1 人体隐喻的主要语言特征	120
7.1.2 人体隐喻自动识别模型的主要构成	121
7.1.3 人体隐喻自动识别模型的验证结果	122
7.2 研究的不足之处	122
7.3 后续的研究计划	123
<hr/>	
参考文献	124
<hr/>	
附录	135