

目 录

第一章 引 言	1
1.1 研究背景	1
1.2 研究意义	4
1.2.1 理论意义	4
1.2.2 实践意义	5
1.3 研究概述	5
1.3.1 研究目的	6
1.3.2 研究对象与问题	6
1.3.3 研究步骤	7
1.4 本书结构	8
1.5 小结	8
第二章 英语学习者自动语法纠错研究综述	9
2.1 自动语法纠错研究概述	9
2.2 学习者自动语法纠错研究方法	11
2.2.1 通用型学习者语法纠错研究	11
2.2.2 适用型学习者语法纠错研究	23
2.3 学习者自动语法纠错研究中的语料库	29
2.3.1 语料库使用方法与模型构建	29
2.3.2 本族语者语料库	31

2.3.3 学习者标注语料库	33
2.3.4 大规模合成数据	36
2.3.5 学习者测试集	38
2.4 学习者自动语法纠错研究中的错误类型	39
2.5 学习者自动语法纠错模型评估指标	41
2.5.1 Precision、Recall与F值	42
2.5.2 MaxMatch	43
2.5.3 I-measure	43
2.5.4 GLEU	44
2.5.5 ERRANT	45
2.6 小结	46
<hr/>	
第三章 中国学习者英语语法错误平行语料库建设	47
3.1 概述	47
3.2 语料库设计	49
3.2.1 语料库组成结构	49
3.2.2 语法错误界定	51
3.2.3 文本预处理原则	53
3.3 语料库建设	55
3.3.1 训练语料库	55
3.3.2 测试语料库	64
3.4 小结	73
<hr/>	
第四章 基于深度学习的中国学习者语法纠错模型构建	74
4.1 连接主义	74
4.1.1 连接主义概述	74
4.1.2 神经网络的特征	75
4.1.3 神经网络对语法规则的学习	77

4.1.4 深度学习与语言学研究	80
4.2 深度学习模型构建方法	82
4.2.1 算法与工具	82
4.2.2 模型架构	86
4.2.3 实验数据	88
4.2.4 模型参数	89
4.3 语法规则编写方法	93
4.3.1 规则编写工具	93
4.3.2 规则编写方法	95
4.3.3 规则编写步骤	98
4.4 模型评估指标	99
4.4.1 整体性能评估	99
4.4.2 错误类型评估	99
4.5 小结	100

第五章 基于深度学习的语法纠错模型性能评估与讨论 101

5.1 实验过程数据	101
5.1.1 模型训练过程性数据	101
5.1.2 模型训练过程数据汇总	105
5.2 实验结果数据	106
5.2.1 中国学习者词形、词-码模型	107
5.2.2 国际学习者随机词形、词-码模型	111
5.2.3 国际学习者词形、词-码模型	114
5.2.4 组合语料库模型	117
5.2.5 语法纠错模型性能对比	119
5.3 分析与讨论	123
5.3.1 构建中国学习者语料库的重要性	123
5.3.2 语料库数据组合的优势	128
5.3.3 语言学特征的重要性	129

5.3.4 深度学习模型与母语迁移理论的互动	131
5.3.5 深度学习模型存在的问题	132
5.4 小结	135
<hr/>	
第六章 语法规则补充深度学习模型	136
6.1 语法错误类型	136
6.2 规则示例	137
6.2.1 形容词形式	137
6.2.2 流水句	138
6.2.3 无主句	139
6.2.4 主谓一致	142
6.3 语法规则测试	145
6.3.1 测试方法	145
6.3.2 测试结果	146
6.3.3 结果分析	147
6.4 小结	149
<hr/>	
第七章 结论	150
7.1 主要发现	150
7.2 研究启示	153
7.3 研究不足	155
7.4 后续研究计划	156
7.5 小结	157
<hr/>	
参考文献	158
<hr/>	
附录	181