



教育部人文社会科学重点研究基地  
北京外国语大学中国外语与教育研究中心  
大数据视野下的外语与外语学习研究系列丛书  
总主编 ◎ 梁茂成

# 基于衔接特征的英语 学习者书面语语篇 连贯自动评价研究

刘国兵 ◎ 著

WordNet metadata discourse  
lexical cohesion text  
chunk CLEC cluster lemma keywords semantic prosody  
link coherence word  
cohesive features  
idiom principle bond open-choice principle

非  
外  
信

外语教学与研究出版社  
FOREIGN LANGUAGE TEACHING AND RESEARCH PRESS



教育部人文社会科学重点研究基地  
北京外国语大学中国外语与教育研究中心  
大数据视野下的外语与外语学习研究系列丛书

总主编 ◎ 梁茂成

# 基于衔接特征的英语 学习者书面语语篇 连贯自动评价研究

Automatic Evaluation of Textual Coherence  
in Chinese EFL Learners' Written English:  
A Study Based on Cohesive Features

刘国兵 ◎ 著

外语教学与研究出版社

FOREIGN LANGUAGE TEACHING AND RESEARCH PRESS

北京 BEIJING

## 图书在版编目 (CIP) 数据

基于衔接特征的英语学习者书面语语篇连贯自动评价研究 / 刘国兵著. ——  
北京: 外语教学与研究出版社, 2017.12

(大数据视野下的外语与外语学习研究系列丛书 / 梁茂成总主编)

ISBN 978-7-5135-9698-5

I. ①基… II. ①刘… III. ①英语-写作-研究 IV. ①H315

中国版本图书馆 CIP 数据核字 (2017) 第 322363 号

出版人 徐建忠  
责任编辑 毕争  
责任校对 解碧琰 刘伟  
封面设计 彩奇风  
出版发行 外语教学与研究出版社  
社址 北京市西三环北路 19 号 (100089)  
网址 <http://www.fltrp.com>  
印刷 北京九州迅驰传媒文化有限公司  
开本 650×980 1/16  
印张 16.5  
版次 2018 年 1 月第 1 版 2018 年 1 月第 1 次印刷  
书号 ISBN 978-7-5135-9698-5  
定价 61.90 元

购书咨询: (010) 88819926 电子邮箱: [club@fltrp.com](mailto:club@fltrp.com)

外研书店: <https://waiyants.tmall.com>

凡印刷、装订质量问题, 请联系我社印制部

联系电话: (010) 61207896 电子邮箱: [zhijian@fltrp.com](mailto:zhijian@fltrp.com)

凡侵权、盗版书籍线索, 请联系我社法律事务部

举报电话: (010) 88817519 电子邮箱: [banquan@fltrp.com](mailto:banquan@fltrp.com)

法律顾问: 立方律师事务所 刘旭东律师

中咨律师事务所 殷斌律师

物料号: 296980001

2014年国家社科基金项目“英语学习者  
书面语语篇连贯自动评价系统研究”（项目  
编号：14BYY084）的阶段性成果

# 总 序

---

## 一、引言

科学研究方法大致有二：其一，归纳法。归纳法指根据一类事物的部分对象的属性推知该类事物的所有对象皆具有某种属性。比如，早期的人类在多次与狼邂逅的过程中，逐渐意识到这种体型匀称协调、四肢修长、头腭尖形、鼻端突出、耳尖直立、善于快速奔跑的野生动物具有极强的攻击性，不可为伍，需要敬而远之或群起而杀之。显然，人类是在经历了多次这样的邂逅之后才意识到了狼的危险性，每一次邂逅都为人类积累了经验、加深了印象，终于在总结若干次教训之后形成了结论：所有的狼都是危险的。诚然，人类在形成结论之前不可能邂逅了所有的狼，但照样可以得出正确的结论。其二，演绎法。演绎法指从一般性的（general）前提出发，通过推导得出具体的（specific）结论。比如，在人们把“所有的狼都是危险的”这一命题视作为一般性前提时，每次邂逅一匹狼，必然会立刻意识到眼前这匹狼是危险的。这其中包含了三个论断，即：所有的狼都是危险的；这是一匹狼；这匹狼是危险的。归纳法是由具体到一般的过程，而演绎法是由一般到具体的过程。

语言研究也不例外，其方法概括起来也不外乎有归纳法和演绎法。演绎法依据可靠的前提进行严密推导，常常可以直击结论。对这种研究方法的运作逻辑我们暂且不做讨论。对于归纳法，其中有若干要素需要考虑。首先，狼有很多特征，哪些特征才具有区别性？哪些属性才是狼的致命属性？比如说，狼嚎是否是我们应该考虑的特征？其次，人类需要与狼邂逅多少次，得出来的结论才是可靠的？返回到语言研究中，前一个问题是语言学家最为关注的问题。语言分析可以从多种语言特征入手，但哪些语言特征才是最有意义的？我们又该如何选择、提取和分析这些语言特征呢？后一个问题是实证研究中的样本问题，即，我们需要

观察多大的语言样本，才可以得出可靠的结论？

自20世纪后半叶语料库语言学问世以来，研究者越发对自然发生的语言数据产生了依赖，因而产生了“经验主义语言学”、“概率语言学”、“数据驱动语言学”等说法，语料库语言学也随之兴起。就其实质而言，语料库语言学采用的是典型的归纳法。语料库是大量自然语言样本的汇集，解决了以上的第二个问题，即实证研究中的样本问题。有了大样本，充分观察成为可能，归纳而得到的结果变得更为可靠甚至可以反复验证。此外，作为方法论的语料库语言学还包含一整套分析方法和分析工具，因而解决了以上第一个问题，即如何提取和分析语言特征的问题。关于选择何种语言特征进行分析，我们将在下面讨论。

总之，有了语料库，我们可能“邂逅”的语言事实更为真实、丰富、全面，这也使得通过归纳法得出的结论更为可靠、经得起验证，不需要像Edward Sapir那样亲力亲为地走入印第安部落之中去采集各式各样的语言数据，也不需要像Charles Fries那样随身携带录音机，甚至不需要像Otto Jespersen那样不失时机地以卡片形式随时记录阅读和日常生活中接触到的各种语言事实。

基于语料库数据进行语言研究，这种方法与演绎法最重要的区别之一在于，研究者在研究中所使用的所有数据均为实际发生的语言事实，而不是靠想象编造出来的句子：

The rat the cat the dog chased killed ate the malt.

Colorless green ideas sleep furiously.

Sincerity admires John.

Golf admires John.

显然，以依据研究者的直觉编造出来的句子作为研究数据，所得结果需要以语言事实来加以验证。正因为语料库语言学研究中的全部数据皆源于事实，结果也更为可靠，因而受到了越来越多研究者的青睐。在这一理念的主导下，我们近年来进行了若干项研究，目的在于利用语料库和语言大数据，对一些语言理论问题进行深入探讨，并试图解决中国外语教育中的一些现实问题。基于这些研究，我们编辑出版了这一套丛书。

## 二、语料分析中的语言特征选择

正如狼的所有特征并非同等重要一样，语言特征的选择在语言的量化研究中也至关重要。在前语料库时代，虽有研究者关注语言事实，但



大部分研究者常常根据自己的直觉选择一些特征进行研究。到了语料库时代，特征的选择方法发生了根本性变化。

在语料库时代，人们将语料库中的连续文本制作成词表或多词词表，甚至制作成词类（POS, part of speech）列表或词类序列（POS sequence）列表，然后对基于不同语料库制作而成的此类列表通过精巧的算法进行频率对比，进而有效地发现语料库中更为有意义的语言特征，特别是词语使用方面的特征。这种方法是语料库语言学研究常用的主题词分析（keywords analysis），研究中几乎总会使用到一个观察语料库和一个参照语料库，并将由这两个语料库析出的词表进行对比，差异较大的词语（即语言特征）会自动浮现出来。这种特征选择方法虽有人工参与，但研究者的主观性和偏好得到了有效控制，因而研究结果也更为可靠，研究也可以重复验证。在有些研究中，人们还在两个语料库中查询自己感兴趣的语言现象，然后对所得频数进行对比，以发现两语料库间的差异。此外，人们还可以编写复杂的正则表达式，从语料库中提取比词表更复杂的语言特征，如名词短语、介词短语、动宾结构、定中结构、关系从句等，甚至涉及意义单位。

上文中描述的基于语料库的语言研究是当今最为常见的语言研究方法之一，其源头至少可以追溯到20世纪八九十年代，也有研究者将此种研究范式视为盛行于20世纪50年代的美国结构主义的延续和发展，甚至也有研究者将语料库之源头追溯到更为久远的时代。笔者认为，基于语料库的研究最早也只能追溯到电子语料库问世之日。正是随着电子语料库的问世，语言研究所需的研究素材在量（quantity）和质（quality）（即语言的真实性）两方面才有了真正的突破。基于语料库的语言研究是时代发展的必然，也为语言研究带来了新视野和新维度。在研究过程中，文本的质和量是研究的基础，而文本分析技术和对比算法起到了关键的作用，可以帮助我们发现最有意义的语言特征。

到了当今的大数据时代，情况又有了新的变化。计算机技术的发展推进了网络技术和互联网的普及，而网络的普及就意味着越来越多的人会花费更多的时间浏览越来越多的网页、上传越来越多的内容，发帖、回帖、发表评论，等等，这一切几乎无时无刻不在发生。智能手机的出现和普及更加推进了这一进程，登录网络、发表言论不再受时间和空间的限制。而所有这一切活动中最为常见的媒介正是我们研究的对象——语言。如此发展下去，网络上的语言资源会越来越多，沉淀也会越来越深，长尾效应也越来越明显。在这一背景之下，语言学家自然不应该满

足于原来规模的语料库，他们与计算机领域的专家联手，设计出了各种工具（常称为网络爬虫），可以从网络上获取大量的文本，彻底颠覆了传统语料库的概念。如今，语料库规模已经由原来的百万词级增大到动辄几千万词或数亿词级，甚至达到几十亿或百亿词级。如此规模的语料库，其优势自然毋庸置疑，长尾效应更扩展了研究维度，基于这样的语料库所得到的研究结果也更为可靠、更为多样化，对语言变化的预测能力也更强。然而，在这样的语料库中查询语言特征或由如此规模的语料库生成词语、词类、各类序列或结构列表变得不再那么容易，对这些海量语料库通过主题词分析法进行对比则更加困难。在大数据时代，我们所面临的问题已经不再是语言研究素材的不足。恰恰相反，数据量过于庞大为语言特征的提取带来了新的挑战，原来的文本分析技术和对比算法不再适用。研究者不得不另辟蹊径。

### 三、大数据时代的语言研究

大数据给语料库语言学者带来了新问题和新的挑战。数据量（volume）庞大是大数据时代最为显著的特征，但这并不是大数据的唯一特征。数据传输和变化之快，即大数据的速度（velocity）使得研究所依赖的数据几乎没有确定的形态，时刻处于变化之中，体量也不断增大，这也是我们必须面对的另一问题。除此之外，大数据的庞杂性（variety）也是一个棘手的问题。以上三个V被公认是大数据的典型特征。在大数据时代，语料库的创建、语言分析工具的开发、统计分析方法的更新和完善、统计结果的呈现等多个问题都将面临一场革命性的变化。

在语料库创建方面，巨量语料库的提纯是一个至关重要的问题。由于网络文本的多样性，粗暴而盲目地堆砌文本、追求语料库的大容量，会使得语料库变得十分地异质、庞杂，因而是不可取的。为此，人们汲取了网络爬虫技术，并加以改造，推出了Web as Corpus技术并开发了专用软件，依据网络页面中的关键词快速创建各种专题语料库。这种技术必将成为大数据时代语言研究中的重要技术。另外，专题语料库固然重要，但对于语言研究者而言，语体差异性、文本的时代性等问题也是语言研究中必须考虑的因素。与语体差异性、文本时代性等密切相关的问题之一是，我们应该如何通过各种途径有效获取文本的外部属性（即元信息），这也是大数据时代的语言研究中面临的又一重大挑战。只有挖掘网络文本的元信息特征，研究者才可以利用文本的各种社会属性（如语



种、产生年代、作者身份、作者性别、语体特征、领域特征等),使语言研究特别是文本差异(text variation)研究得以深入。

在语言分析工具方面,由于大量文本都存储于网络或云端,加之语料库规模不断扩大,原先广泛使用的WordSmith Tools、AntConc等单机版的文本分析工具逐渐会变得不再适用,基于网络或云端的工具或许将会成为技术开发的重点之一。此外,在语料库加工方面,基于大数据和深度学习(Deep Learning)技术设计的系统(如谷歌公司开发的句法标注工具SyntaxNet)将代表主流的研究方向,标注的准确率也会有明显提高。

从标注语料库中提取和统计语言特征时,原先广泛使用的统计方法不再适用,主题词分析方法随着语料库规模的增大也必将变得越来越困难,逐渐取而代之的是更为复杂的数据科学(Data Science),聚类、因子分析、复杂回归分析等成为语言分析的常用方法,分析工具也由原来常用的SPSS等工具变成R等更为复杂的系统。R软件的优势不仅在于可以分析大数据,还将编程和统计融合起来,使研究者可以定制各式各样的分析手段。

在统计结果呈现方面,语料库研究常见的图表呈现方式仍然会被广泛使用,但与此同时,随着数据量的增大,数据的可视化将成为呈现研究结果的重要方式,这种呈现方式将更为直观、便于理解。相信在不远的未来,语料库研究的结果将会使越来越多的人受益。

#### 四、结语

随着大数据时代的到来,语料库语言学必将得到更多研究者的重视和青睐,大数据时代的特点将在语言研究中逐渐显现。我们希望通过本系列丛书的出版推进语言研究的不断科学化,推动我国外语与外语教育研究的发展。

本套丛书是教育部人文社会科学重点研究基地北京外国语大学中国外语与教育研究中心“十三五”规划重大项目“大数据视野下的外语与外语学习研究”(编号:17JJD740003)的研究成果,特此鸣谢。

梁茂成  
二〇一七年三月

# 前 言

---

在大规模外语考试中，写作是测试学习者语言综合运用能力的一个重要手段，也是目前多数外语考试的一种常见题型。近些年来，不管是国际性的外语考试（如IELTS、TOEFL、GRE等），还是国内的高考，全国大学英语四、六级，英语专业四、八级，全国英语等级考试以及2016年开始的国际人才英语考试等，这些大规模考试都将写作作为必考项目，而且在整个试卷分值结构中占有突出权重。已有的二语写作研究表明，能够有效反映语篇质量的语言特征主要包括语篇连贯性、词汇复杂性、句法复杂性以及语法准确性等（Ellis 1996；Polio 2001；Wolfe-Quintero et al. 1998；秦晓晴、文秋芳 2007）。因此，语篇连贯性与词汇、句法特征成为评价二语写作质量的重要观测点，其中连贯的语篇是高质量作文的核心特征之一。

目前对于二语写作中词汇复杂性、句法复杂性以及语法准确性的评价研究已经较为成熟，其研究成果不仅能为人工二语作文批改提供较为客观的量化指标，而且还成功运用于作文自动评分系统的研发。但由于语篇连贯性一直被人们认为是一个较为主观的概念，意义指向也不够具体，因此目前研发的多数作文自动评分系统在设计内部评价参数时，均没有明确涉及语篇连贯性这一重要指标。我们认为，缺少连贯性指标的评分体系是不完整的，它会对评分效果、特别是构念效度造成直接影响；而建立在不完整指标体系上的评分系统，其评价结果也不够客观。

本研究综合国内外英语作文自动评分与语篇连贯性研究两个领域的知识，利用语料库语言学、计算语言学、信息检索技术、统计学以及自然语言处理等学科中的研究方法，基于已有的连贯理论，通过提取学生作文中与语篇连贯相关的文本特征，尝试构建中国英语学习者书面语语篇连贯评价模型。本研究使用的语料选自中国学生英语口语语料库。语料选取采用随机抽样的方式，最终确定以“Education as a Lifelong Process”为题的同题议论文80篇。其中50篇作为训练集，用于构建语篇

连贯评价模型；30篇作为验证集，用于验证模型。在人工评分阶段，我们选定12名具有多年作文评分经验的大学英语教师作为评分员，评分之后选取信度最高的3名评分员，将其评分结果作为本研究的有效因变量。所有评分员均为在读博士或已获得博士学位。在变量提取阶段，我们基于已有连贯理论，结合中国大学英语学习者作文语料的实际情况，通过集成开发环境Delphi自行设计程序，最终实现了全部变量的自动检索与提取。在模型创建与验证阶段，本研究共创建了语篇整体连贯与局部连贯两个评价模型，验证结果表明，这两个模型的预测分值与人工评分结果之间具有良好的一致性。模型预测力较强，具有统计学意义。

本研究的主要发现可以概括为三个方面：1) 语篇连贯性预测因子。提取的44个变量中，通过相关分析，共发现23个语篇整体连贯预测因子和18个语篇局部连贯预测因子。2) 模型构建结果。把语篇连贯性预测因子作为自变量，人工评分为因变量，采用逐步回归的方法进行多元线性回归分析。从模型构建的结果来看，语篇主题相关连结数、中心句百分比、相邻连系数、二元单位复现及连结密度等5个预测因子作为有效自变量进入了整体连贯方程模型。方程的复相关系数为0.782，决定系数为0.612。从语篇局部连贯评价模型的构建结果来看，非中心句百分比、语篇主题连系数、句子数、词元复现以及二元单位复现等5个预测因子作为有效自变量进入了局部连贯方程模型。该模型的复相关系数为0.604，决定系数为0.365。以上数据表明，进入方程模型的自变量与语篇连贯性之间高度相关，构建的模型整体上能够解释较多方差。3) 模型验证结果。本研究从机器评分与人工评分的相关性以及二者的相对一致性两个方面来对模型进行验证。从相关性分析结果来看，语篇整体连贯机器评分与人工评分之间的相关系数为0.785，局部连贯机器评分与人工评分的相关系数为0.708，语篇整体连贯评价模型与局部连贯评价模型均能有效预测学生作文的连贯性。从机器评分与人工评分的相对一致性表现来看，语篇整体连贯机器评分与人工评分等级相同或相邻的作文共有71篇，占作文总数的88.75%；而在语篇局部连贯性评价方面，机器评分与人工评分相对一致的段落数为286，占研究语料总段数的80.56%。这些数据表明，语篇整体连贯与局部连贯两个评价模型完全能够应用于学习者书面语语篇连贯性的自动评价，机器评分与人工评分在相对一致性方面已经达到令人满意的效果。

以下简要介绍本书的结构。开篇为绪论，主要介绍本研究的目的、理论及实践意义，并对研究步骤与流程做简要概述。第一章和第二章是

文献综述。第一章从主要范式与研究方法两个方面对国内外近40年来有关语篇连贯的研究进行综述，重点对 van Dijk 的宏观结构理论和 Hoey 的词汇衔接理论进行介绍。前者为本研究的理论基础，为本研究提供了理论框架；后者为本研究提供了连贯研究的操作原则。第二章从作文自动评分、语篇连贯性的测量两个方面对国内外英语书面语语篇质量评价文献进行述评，之后提出本研究对于语篇连贯的认识、所持基本观点以及评价语篇连贯性采用的方法。第三章与第四章主要报告了本书采用的研究方法。第三章介绍本研究的数据准备情况，包括使用的语料，人工评分标准、过程、结果信度以及训练集与验证集的形成过程等。第四章提出了本研究要解决的研究问题，并介绍所使用的研究工具、语篇连贯性特征的提取方法及过程、数据分析方法与研究流程等。其中专门设立一节介绍 WordNet 语义知识库，包括该数据库的设计原理、库中主要语义关系以及在本研究中的应用方法等。第五章报告了各种变量与语篇连贯性的相关分析过程及结果，之后根据相关分析结果确定语篇整体连贯性与语篇局部连贯性预测因子。所有预测因子都将作为自变量进入多元回归分析。第六章详细介绍了学习者书面语语篇整体连贯与局部连贯两个评价模型的构建情况，包括模型的拟合度、方差检验与 t 检验结果等。之后从人工评分与机器评分之间的相关性以及二者的绝对一致性和相对一致性两个方面详细报告评价模型的验证过程以及验证结果。第七章是本书的总结部分，重点总结了本研究的主要发现及研究的理论价值和实践意义，并分析研究存在的不足。最后探讨了今后可供拓展的研究方向。

在本书的撰写过程中，我的导师梁茂成教授给予了很多帮助。导师严谨的治学精神、宽广的学术视野、乐观的生活态度、直率的处世之道对我影响很大。从选题到书稿最终完成，梁老师给予了我悉心的指导和不懈的支持。还要感谢著名计算语言学家冯志伟教授、北京外国语大学中国外语与教育研究中心的文秋芳教授和李文中教授、北京师范大学外文学院的苗兴伟教授、北京邮电大学计算机科学学院的王小捷教授、对外经济贸易大学英语学院的王立非教授，他们百忙之中抽出时间审阅了本书初稿，并提出很多建设性意见。感谢北京外国语大学中国外语与教育研究中心的许家金教授和熊文新教授，本书成书过程中，他们也给予了大量理论与技术方面的支持。同时，本书作为2014年国家社科基金项目（编号：14BYY084）的阶段成果，离不开全国哲学社会科学规划办公室的大力支持，在此一并感谢。由于时间仓促、学识水平有限，本书难免存在纰漏之处，敬请各位读者不吝赐教，并批评指正。

# 目 录

---

<b>绪 论</b>	1
0.1 引言	1
0.2 本研究的理论意义及实践意义	3
0.2.1 理论意义	3
0.2.2 实践意义	5
0.3 本研究概述	8
0.4 本书结构	10
0.5 小结	11
<hr/>	
<b>第一章 语篇连贯研究文献回顾</b>	12
1.1 衔接与连贯：既有区别又相互联系的两个概念	12
1.1.1 衔接	13
1.1.2 连贯	14
1.1.3 衔接与连贯的关系	15
1.2 已有研究概述：研究范式与研究方法	16
1.2.1 研究概述	16
1.2.2 两种研究范式	17
1.2.3 三类研究方法	19
1.3 对连贯研究主要文献的分类综述	21
1.3.1 van Dijk的语义连贯理论	22
1.3.2 Hoey的词汇衔接理论	31
1.3.3 Halliday & Hasan的衔接连贯理论	42

1.3.4 Danés和Fries的主位推进理论	44
1.3.5 Hobbs的连贯指称理论	44
1.3.6 基于语言产出的其他研究	45
1.4 已有研究的不足之处	45
1.5 小结	47

## **第二章 书面语语篇质量评价文献回顾及本研究 所采用的连贯分析方法** 48

2.1 书面语语篇质量评价研究概述	48
2.2 书面语语篇质量评价的第一类研究：作文自动评分	48
2.2.1 作文自动评分的工作原理	48
2.2.2 作文自动评分的优越性	49
2.2.3 机器自动评分系统概述	51
2.2.4 对现有机器评分系统的评价	57
2.3 书面语语篇质量评价的第二类研究：语篇连贯性的测量	62
2.3.1 语篇连贯量化工具：Coh-Metrix	62
2.3.2 Mann与Thompson的修辞结构理论	64
2.3.3 Grosz等人的向心理论	65
2.3.4 语篇连贯测量的其他研究	65
2.4 已有文献对本研究的启示	66
2.4.1 作文自动评分对本研究的启示	66
2.4.2 语篇连贯测量对本研究的启示	67
2.5 本研究对连贯的认识及采用的分析方法	68
2.5.1 对连贯的认识	68
2.5.2 本研究采用的分析方法	74
2.6 小结	76

## **第三章 数据准备** 77

3.1 语料	77
3.2 人工评分标准	78
3.2.1 制定人工评分标准的理论依据	78



3.2.2 局部连贯性人工评分标准	80
3.2.3 整体连贯性人工评分标准	81
3.2.4 人工评分权重的分配	83
3.3 人工评分过程	84
3.3.1 评分员选择	84
3.3.2 评分员培训	85
3.3.3 评分	86
3.4 人工评分信度报告	87
3.4.1 句对人工评分信度	89
3.4.2 段落人工评分信度	90
3.5 训练集与验证集的形成	92
3.6 小结	92

---

## **第四章 文本及数据分析** 93

4.1 研究问题	93
4.2 研究工具	93
4.2.1 文本预处理工具	94
4.2.2 文本检索与分析工具	95
4.2.3 数据分析工具	104
4.3 语篇连贯性特征提取	104
4.3.1 基础变量	105
4.3.2 语义关系变量	106
4.3.3 连系变量	109
4.3.4 连结变量	110
4.3.5 基于连系与连结关系的综合变量	112
4.4 数据分析	114
4.4.1 相关分析	114
4.4.2 多元线性回归分析	114
4.5 研究流程	115
4.5.1 模型构建	115
4.5.2 模型验证	116
4.6 小结	117

---

<b>第五章 学习者书面语语篇连贯性预测因子</b>	<b>118</b>
5.1 文本特征变量与语篇整体连贯性	118
5.1.1 基础/语义关系变量与整体连贯性	119
5.1.2 连系/连结变量与整体连贯性	122
5.1.3 综合变量与整体连贯性	125
5.2 文本特征变量与语篇局部连贯性	129
5.2.1 基础/语义关系变量与局部连贯性	129
5.2.2 连系/连结变量与局部连贯性	131
5.2.3 综合变量与局部连贯性	133
5.3 小结	135
<hr/>	
<b>第六章 学习者书面语语篇连贯评价模型的构建</b>	<b>136</b>
6.1 学习者书面语语篇整体连贯评价模型的构建	136
6.1.1 进入回归分析的语篇整体连贯相关变量	138
6.1.2 整体连贯评价模型概况	139
6.1.3 整体连贯评价模型方差检验结果	140
6.1.4 整体连贯评价模型系数的t检验结果	141
6.1.5 整体连贯评价模型的共线性情况	143
6.1.6 整体连贯评价模型回归方程	144
6.2 学习者书面语语篇局部连贯评价模型的构建	147
6.2.1 进入回归分析的语篇局部连贯相关变量	148
6.2.2 初次建模与二次建模情况	150
6.2.3 局部连贯评价模型的拟合情况	155
6.2.4 局部连贯评价模型方差检验结果	156
6.2.5 局部连贯评价模型 t 检验结果	156
6.2.6 局部连贯评价模型共线性情况	158
6.2.7 局部连贯评价模型回归方程	158
6.3 语篇整体连贯评价模型与局部连贯评价模型的验证	161
6.3.1 机器评分与人工评分之间的相关性及其一致性	162
6.3.2 机器评分和人工评分等级绝对一致与相对一致情况	162
6.4 小结	165
<hr/>	

<b>第七章 研究发现及价值</b>	167
7.1 研究发现	167
7.1.1 学习者书面语语篇连贯相关变量	167
7.1.2 构建的模型及性能	172
7.2 研究价值及创新之处	174
7.2.1 理论方面	174
7.2.2 实践方面	176
7.3 本研究的不足之处	178
7.4 本研究的拓展方向	180
<hr/>	
<b>参考文献</b>	182
<hr/>	
<b>附录</b>	203