

向心理论应用于英语学习者书面语语篇局部连贯评价的有效性研究

刘国兵¹,常芳玲²

(1.河南师范大学 外国语学院,河南 新乡 453007;2.北京外国语大学 中国语言文学学院,北京 100089)

摘要:以中国学生英语口语语料库为语料来源,分析比较了向心理论的四项原则分析法用于评价议论文、记叙文和说明文三种不同体裁书面语语篇局部连贯的有效性.结果发现:(1)不同体裁二语写作人工评分与 Kibble 评分之间的相关性存在差异.议论文中人工评分与 Kibble 评分显著相关;记叙文和说明文人工评分与 Kibble 评分的相关性相对较弱.(2)除衔接性原则外,四项基本原则在二语写作中的比例分布因作文体裁不同而各有特点.议论文中延续性原则所占比例较高,显著性原则和低价性原则占比较低;记叙文中显著性原则和低价性原则占比较高,延续性原则占比较低;说明文各项原则占比多介于议论文和记叙文之间.(3)四项基本原则和人工评分之间的相关性因体裁而异.议论文中衔接性原则与人工评分显著相关;记叙文中各项原则与人工评分之间的相关性受评分员影响较大;说明文中各项原则与人工评分无显著相关关系.四项原则分析法对议论文语篇局部连贯的评价效度优于记叙文和说明文,其中衔接性原则与议论文语篇局部连贯质量相关度最高,可作为议论文二语写作语篇局部连贯评价指标纳入自动评分系统.

关键词:向心理论;局部连贯;四项原则分析法;二语写作;作文体裁

中图分类号:H08

文献标志码:A

语篇连贯性是衡量语篇质量的主要标准之一^[1-2].由于连贯本身具有抽象性和复杂性,学者们对语篇连贯的评价标准争议颇多,因此诸多着眼于不同视角评价语篇连贯的理论应运而生,如衔接与连贯理论、连贯指称理论、主位推进理论和修辞结构理论等.这些理论的出现也催生了一系列有关语篇连贯的应用性研究^[3-8],促进语篇连贯研究朝着更深、更广方向发展.尽管上述理论从不同视角出发剖析了语篇连贯现象,但它们普遍缺乏语篇连贯的量化评价指标,因此评价结果带有不同程度的主观性.向心理论是涵盖语篇词汇、语义、语法和语用等因素的一个综合性语篇连贯性认知推理模型^[9].该理论包括主流四分法、七分法和四项原则分析法三种分析模型,旨在解释说明自然产生的语篇中语句注意中心的变化,涉及注意状态、推理难度和指称形式三者之间关系的阐释^[10].与其他语篇连贯理论相比,向心理论最大的特点在于它的可计量性^[11].这一特点也使得该理论能为语篇连贯评价提供一些较为客观的指标,使评价结果更加科学可信.国外一些文献关注向心理论用于评价语篇连贯的实际应用价值,主要探讨了向心理论的连贯原则转化为语篇连贯评价指标的方法^[12-15].文献^[16-17]讨论了向心理论在自然语言生成领域的价值.近年来国内也有一些学者在向心理论的指导下研究语篇局部连贯现象,如洪明^[18]以学习者二语作文为语料来源,验证了四项原则分析法用于评价二语写作语篇连贯的有效性.她发现在向心理论的三种分析模型中,四项原则分析法对二语写作语篇连贯的评价效度最高.此外,虽然向心理论用于评价语篇连贯的研究已较为广泛,但这些研究所涉对象均以议论文为主,该理论用于评价不同体裁语篇局部连贯的效度仍鲜有涉及.因此,本文基于向心理论的四项原则分析法,以中国学生英语口语语料库笔语子库中不同体裁的二语作文为语料来源,试分析该理论对不同体裁二语写作语篇局部连贯的评价效度,促进二语写作语篇质量评价的科学式发展.

收稿日期:2019-08-06;**修回日期:**2020-03-11.

基金项目:河南省哲学社会科学规划项目(2017BYY008)

作者简介(通信作者):刘国兵(1978—),男,河南滑县人,河南师范大学教授,博士,研究方向为语料库语言学,E-mail:liuguobing55@163.com.

1 理论框架

向心理论是计算语言学学者提出的一种用于分析语篇局部连贯的实体连贯理论,在语篇局部连贯的量化分析中具有显著优势^[19],包括主流四分法、七分法和四项原则分析法三个分析模型.基于洪明^[18]对向心理论三种分析模型评价二语作文语篇连贯质量的研究成果,本研究以四项原则分析法为理论框架.四项原则分析法是 Kibble^[11]为弥补主流四分法缺陷而提出的语篇局部连贯分析模型,该模型定义了延续性、显著性、低价性和衔接性四种过渡方式用以分析语篇的局部连贯程度,具体如表 1 所示.

在四项原则分析法各过渡方式的公式化定义中, U_i 和 U_{i-1} 分别代表当前语句以及当前语句的前一语句; C_f (Forward-looking Center)指语句的前瞻中心,是某个语句中提及的所有语篇实体的集合; C_b (Backward-looking Center)代表语句的回指中心,是语句 U_i 中实现的、在 $C_f(U_{i-1})$ 中显著性最高的元素^[20]; C_p (Preferred Center)为语句的优选中心,指代某一语句中显著性最高的实体.由表 1 可知,如果相邻的两个语句的前瞻中心存在交集,则两句间的过渡是延续性的;如果某语句回指中心与其优选中心相同,那么该语句与上一语句间为显著性过渡;如果一个语句的回指中心与上一语句的优选中心相同,则两句之间的过渡符合低价性原则;如果相邻的两个语句回指中心相同,那么这两个语句就被认为是衔接性过渡.

四项原则分析法在评价语篇局部连贯程度时可采用两种方法:一是确定四项原则的强弱顺序,违反最弱原则的语篇理解或生成方式即为最优;二是将上述四项原则等同看待,满足数量最多的语篇理解或生成方式最佳^[21].由于各原则之间的强弱顺序无法得到客观衡量,本文拟采用第二种方法评价语篇的局部连贯质量,具体算法如下:将延续性原则、显著性原则、低价性原则和衔接性原则中每项原则分值设为 1,标记语篇中各语句满足上述四项原则的情况,最后通过 Kibble 评分(Kibble 评分 = 某语篇中四项基本原则的数量/(该语篇语句总数 - 作为语段分界的语句数))计算该语篇的局部连贯程度,Kibble 评分越高,表明该语篇局部连贯程度越高.

2 研究设计

2.1 研究问题

本文旨在考察四项原则分析法对不同体裁二语写作语篇局部连贯的评价效度,拟回答以下问题:

- (1)不同体裁二语写作人工评分和 Kibble 评分的相关性是否存在差异?如存在,具体表现如何?
- (2)不同体裁二语写作语句间过渡方式在四项原则分析法的理论上呈现什么特点?
- (3)四项原则分析法用于分析不同体裁二语写作语篇局部连贯的评价效度是否相同?

2.2 语料选择

为回答上述研究问题,本研究以中国学生英语口语笔语语料库的笔语子库为语料来源,通过系统抽样方式抽取议论文、说明文及记叙文各 50 篇,将其以 ARG-1, ARG-2, ..., ARG-49, ARG-50; NAR-1, NAR-2, ..., NAR-49, NAR-50; EXP-1, EXP-2, ..., EXP-49, EXP-50 的形式重命名,并保存为.txt 格式.

2.3 参数设定

参数设定影响向心理论评价语篇局部连贯的有效程度^[22].即使对于同一项研究,若设定参数不同,研究结果也可能存在差异.因此,在借助向心理论开展应用性研究之前,参数设定的工作必不可少.本研究所涉及的参数主要包括语篇片段、语句、实现和 C_f 排序.

(1)语篇片段.设定语篇片段是基于向心理论开展研究的基础^[18].语篇片段的划分以意图结构为依据,若干具有相同意图的语句序列被看成是同一个语篇片段^[23].本研究以二语学习者英语作文为研究对象,作文中的语句从理论上讲应服务于各自的作文题目,具有共同的语篇意图.因此,本研究将语篇片段设定为整篇

表 1 四项原则分析法

Tab.1 The four principle model of centering theory

过渡方式	定义
延续性原则	$C_f(U_i) \cap C_f(U_{i-1}) \neq \emptyset$
显著性原则	$C_b(U_i) = C_p(U_i)$
低价性原则	$C_b(U_i) = C_p(U_{i-1})$
衔接性原则	$C_b(U_i) = C_b(U_{i-1})$

作文。换言之,Kibble评分计算公式中作为语段分界的语句数设为1。

(2)语句。将T单位作为语句切分的主要标准。T单位指符合语法规则、语句能够被切分的最小单位。文献[24-25]认为,分析口语或书面语最理想的单位不是句子,而是T单位。梁茂成、文秋芳^[26]在研究中也证明了T单位用于语篇分析的有效性。

(3)实现。如果实体 a 是语句 U 所描述情景的一个成分,或 a 是语句 U 语义解释的组成部分,那么就认为语句 U 实现了实体 a 。作为向心理论中的一个重要概念,语句中“实现”的定义条件直接影响Cf的确定。一般而言,“实现”可以分为直接实现和间接实现。直接实现指上文中提到的某个实体在当前语句中重复出现;间接实现指当前语句中某个名词或名词短语与上文中出现过的某个词语存在关联但不完全对等。前者对语篇连贯的促进作用外显且易于辨认,而对后者的界定具有模糊性。不过,间接实现也是促进语篇连贯的一个重要手段,符合研究者对连贯概念的定义^[13,22,27]。本文参照文献[22]的标注方法,不仅对语句中的直接实现加以标注,而且标注了成员关系、一般所有关系和次成分关系三种间接实现。需要说明的是这些标注只涉及名词或名词短语之间的实体关系,不包括事件或者动作的回指形式、小句中的命题关系、语篇指示词以及动名词形式。此外,由于第一人称的使用对语篇连贯的促进作用有待验证,本研究在人称代词的实现问题上仅标注了语句中出现的第二人称和第三人称。

(4)Cf排序。语句中Cf的排列次序因语言不同而存在差异。本文采用文献[28]对英语中Cf显著性大小依次为:主句主语,主句间接宾语,主句宾语,主句其他成分,从句主语,从句间接宾语,从句宾语,从句其他成分。

2.4 研究步骤

如图1所示,本文以向心理论的四项原则分析法为分析框架,首先从中国学生英语口语语料库笔语子库中抽取议论文、记叙文和说明文各50篇作为语料样本,将其重命名后以.txt格式保存以便开展进一步研究。

语料准备工作就绪后,进入不同体裁二语写作的语料评分阶段。语料评分由人工评分和Kibble评分两部分构成。人工评分由两名长期从事英语教学的英语教师按照刘国兵^[8]提出的语篇局部连贯评价标准完成。为保证人工评分效度,对两名评分员给定的人工评分会进行相关分析。如人工评分1与人工评分2显著相关,则可开展下一步工作;若两者相关性不显著,语料样本需进行二次评分或多次评分,直至人工评分1和人工评分2之间呈显著相关方可结束。Kibble评分是在语料样本语句切分、语料标注以及数据统计录入的基础上,利用本文理论框架部分提及的四项原则分析法中有关四项原则的定义以及Kibble评分公式(Kibble评分=某语篇中四项基本原则的数量/(该语篇语句总数-作为语段分界的语句数)计算得出的作文分数。假设某篇语料被切分为25个语句,延续性、显著性、低价性和衔接性四项原则共计出现41次,再者本文已根据语料实际情况设定每个语篇作为语段分界的语句数为1,那么该语篇的Kibble评分的结果为1.71(保留两位小数)。

最后,将不同体裁语料样本的人工评分和Kibble评分、Kibble评分和四项基本原则进行相关分析,依据统计学结果解释Kibble评分用于评价学习者不同体裁书面语语篇局部连贯的效度以及四项原则在不同体裁二语写作中的分布特征。

3 结果与分析

3.1 不同体裁人工评分和Kibble评分相关分析

在150篇语料样本人工评分1和人工评分2显著相关($r=0.588, p=0.000<0.05$)的基础上,进一步考察了人工评分和Kibble评分在不同体裁样本中的相关性,结果如下。

3.1.1 议论文人工评分和Kibble评分相关分析

据表2可知,议论文语料样本人工评分1和人工评分2显著相关($r=0.652, p=0.000<0.05$),与样本总体人工评分相关分析结果($r=0.588, p=0.000<0.05$)相比,两名评分员在议论文的人工评分中更为一致,对议论文二语写作语篇局部连贯质量有非常统一的认识。从两名评分员给出的人工评分和Kibble评分的相

关情况来看,人工评分 1 和人工评分 2 在 0.05 的水平上均与 Kibble 评分显著相关.

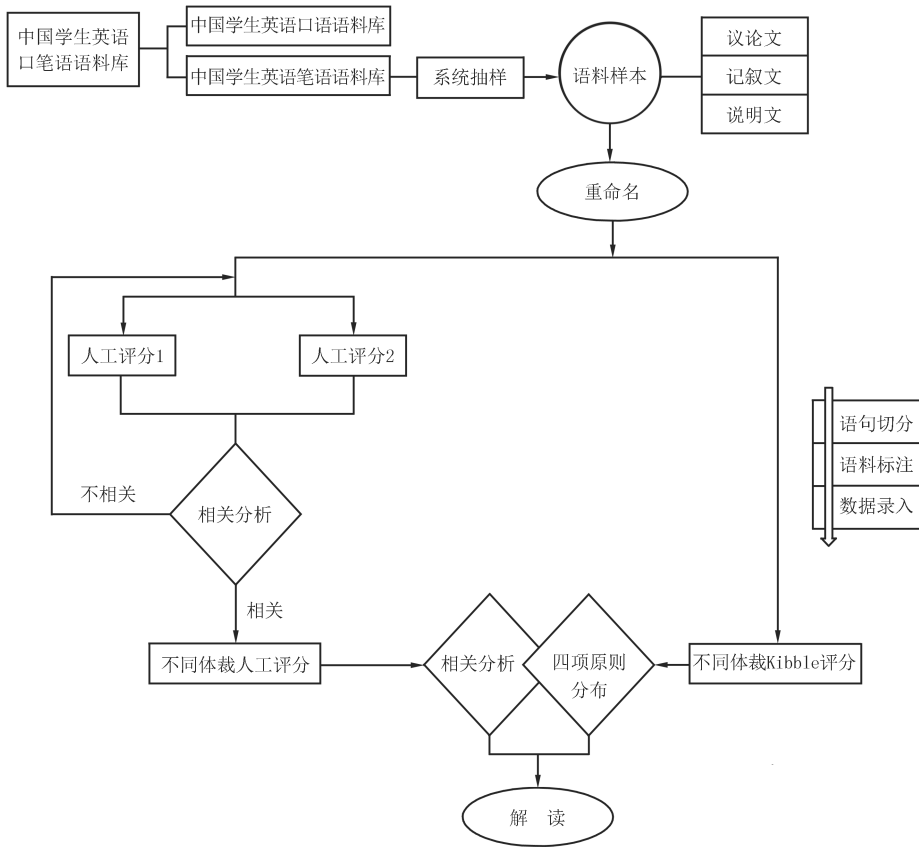


图1 研究步骤流程图

Fig.1 Research procedures of the current study

表 2 议论文人工评分和 Kibble 评分相关分析

Tab.2 Correlation analysis of manual scores and Kibble scores in argumentations

评分类别	Pearson 相关分析	人工评分 1	人工评分 2	Kibble 评分
人工评分 1	Pearson 相关性	1	0.652 **	0.484 **
	显著性(双侧)	—	0.000	0.000
	N	50	50	50
人工评分 2	Pearson 相关性	0.652 **	1	0.360 *
	显著性(双侧)	0.000	—	0.010
	N	50	50	50
Kibble 评分	Pearson 相关性	0.484 **	0.360 *	1
	显著性(双侧)	0.000	0.010	—
	N	50	50	50

** 在 0.01 水平(双侧)上显著相关; * 在 0.05 水平(双侧)上显著相关.

3.1.2 记叙文人工评分和 Kibble 评分相关分析

记叙文语料样本人工评分和 Kibble 评分的相关分析结果如表 3 所示.

表 3 显示,记叙文人工评分 1 和人工评分 2 显著相关($r=0.532, p=0.000<0.05$),其相关系数略低于样本总体人工评分($r=0.588, p=0.000<0.05$),但二者差异不大.不过,两名人工评分员对记叙文的人工评分与 Kibble 评分分析结果存在差异.人工评分 1 与 Kibble 评分显著相关($r=0.307, p=0.03<0.05$),而人工评

分2与Kibble评分不存在显著相关关系($r=0.275, p=0.053>0.05$).由于人工评分1和人工评分2显著相关,因此可推测这种差异可能是由于Kibble评分用于评价记叙文局部连贯质量效度不高而导致的.

表3 记叙文人工评分和Kibble评分相关分析

Tab.3 Correlation analysis of manual scores and Kibble scores in narrations

评分类别	Pearson 相关分析	人工评分 1	人工评分 2	Kibble 评分
人工评分 1	Pearson 相关性	1	0.532 **	0.307 *
	显著性(双侧)	—	0.000	0.030
	N	50	50	50
人工评分 2	Pearson 相关性	0.532 **	1	0.275
	显著性(双侧)	0.000	—	0.053
	N	50	50	50
Kibble 评分	Pearson 相关性	0.307 *	0.275	1
	显著性(双侧)	0.030	0.053	—
	N	50	50	50

** 在 0.01 水平(双侧)上显著相关; * 在 0.05 水平(双侧)上显著相关.

3.1.3 说明文人工评分和 Kibble 评分相关分析

与样本总体人工评分相关分析结果($r=0.588, p=0.000<0.05$)相比,说明文人工评分之间相关度较低($r=0.419, p=0.002<0.05$),表明评分员对说明文语篇局部连贯的评价一致性低于议论文和记叙文.但人工评分1和人工评分2之间仍显著相关($r=0.419, p=0.002<0.05$),因此人工评分1和人工评分2在对说明文的语篇局部连贯评价上仍具有统计学意义.从表4人工评分和Kibble评分的相关数据来看,说明文中Kibble评分与人工评分的相关性不能一概而论,存在一定的不确定性,该情况与记叙文中人工评分和Kibble评分相关分析结果大体相同.

表4 说明文人工评分和Kibble评分相关分析

Tab.4 Correlation analysis of manual scores and Kibble scores in expositions

评分类别	Pearson 相关分析	人工评分 1	人工评分 2	Kibble 评分
人工评分 1	Pearson 相关性	1	0.419 **	0.287 *
	显著性(双侧)	—	0.002	0.043
	N	50	50	50
人工评分 2	Pearson 相关性	0.419 **	1	0.065
	显著性(双侧)	0.002	—	0.653
	N	50	50	50
Kibble 评分	Pearson 相关性	0.287 *	0.065	1
	显著性(双侧)	0.043	0.653	—
	N	50	50	50

** 在 0.01 水平(双侧)上显著相关; * 在 0.05 水平(双侧)上显著相关.

3.2 四项基本原则在不同体裁语篇中的分布特征

上一小节借助Kibble评分从总体上讨论了四项原则分析法与人工评分的相关性,本节尝试从微观层面入手探讨延续性、显著性、低价性和衔接性四项原则用于评价不同体裁二语写作语篇局部连贯质量的效度.

3.2.1 四项基本原则在不同体裁语篇中的比例分布

根据四项基本原则在不同体裁语料样本中的出现频数,绘制成如图2所示的柱状图.从宏观层面来看,四项基本原则在语篇中的分布比例未表现出体裁差异.无论在何种体裁中,延续性原则所占比例最高,低价

性原则和显著性原则次之,衔接性原则最低.四项基本原则的这一分布规律与各原则的操作定义密切相关.依据表1各项原则公式蕴含的实际意义,延续性原则规定的相邻语句间的语篇连贯关系最为广泛,且涵盖了其他三项原则所传达的语篇局部连贯关系.换言之,某语句只有满足了延续性原则的定义,才有可能满足通过其他三项原则实现语篇局部连贯的条件.延续性原则的这一特点保证了该原则在语料样本中的出现比例最高.此外,与显著性和低价性两个原则相比,衔接性原则在语句中的实现难度较高.显著性原则和低价性原则的公式分别为 $Cb(U_i) = Cp(U_i)$ 和 $Cb(U_i) = Cp(U_{i-1})$,由公式可知,显著性原则和低价性原则至少涉及相邻两个语句局部注意焦点的延续;而衔接性原则涉及三个或三个以上相邻语句之间的关系($Cb(U_i) = Cb(U_{i-1})$).因此,衔接性原则对学习者的语言综合水平要求很高,这在一定程度上降低了衔接性原则在二语作文中的出现频率.

从微观层面来看,各项基本原则所占比例因体裁不同而有差异.延续性原则在议论文中所占比例最高,说明文中次之,记叙文中所占比例最低;显著性原则在记叙文中出现比例最高,其次是说明文,最后是议论文;低价性原则在议论文和说明文中占比相同,均低于该原则在记叙文中的所占比例;衔接性原则在三种作文体裁中占比相同,未表现出体裁差异,因此该原则可作为一个理想评价指标衡量语篇

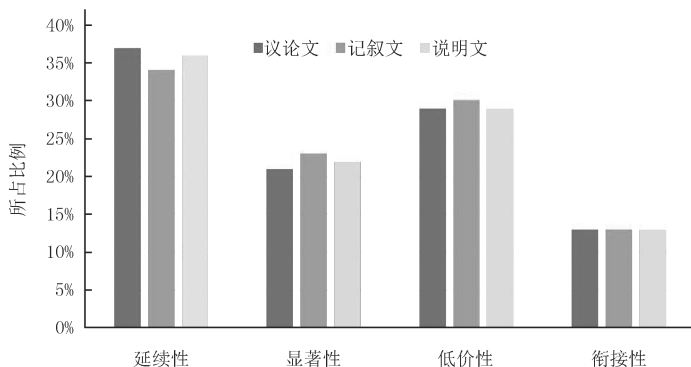


图2 四项基本原则在不同体裁二语写作语篇中的比例分布
Fig.2 Distributions of the four principles in second language writing with different genres

局部连贯质量.这些原则在不同体裁作文中的比例差异与各种作文体裁的特点存在密切关系.议论文通常围绕某论点展开讨论,各个语句均服务于同一个话题,因此该体裁中符合延续性原则的语句相对较多,符合显著性原则和低价性原则的语句较少;记叙文注重对事物动态过程的描写,往往需要通过注意焦点的不断变化来推进记叙对象的发展,在四项原则分析法框架下表现为该类作文体裁显著性原则和低价性原则所占比例较高,延续性原则占比例较低;说明文注重对说明对象进行客观科学的阐释,它不像议论文那样每个语句都服务于同一个论点,同时其说明对象的动态变化速度也不及记叙文.因此,四项基本原则在说明文中的分布比例多介于议论文和记叙文之间.

3.2.2 不同体裁四项基本原则与 Kibble 评分相关分析

(1) 议论文四项基本原则与人工评分相关分析.为进一步了解议论文中四项基本原则的分布与人工评分的相关性,对四项基本原则和人工评分进行了相关分析,结果如表5所示.从四项基本原则与人工评分的相关分析来看,无论是人工评分1还是人工评分2,与之显著相关的原则只有衔接性原则,其他三项原则与人工评分之间不存在显著相关关系.

综合表5、表2和图2相关数据可知,向心理论的四项原则分析法可以用来评价议论文二语写作语篇局部连贯质量.虽然衔接性原则在议论文语料样本中所占比例最少,但该原则反映语篇局部连贯质量的效度最高.除四项基本原则和 Kibble 评分的相关分析结果以外,表5还呈现了议论文中人工评分1和人工评分2以及各项原则之间的相关分析结果.其中,四项基本原则中每两项原则间显著相关,表明四项基本原则在二语写作语篇中相互关联、相互影响.人工评分1和人工评分2的相关性已在前文提及,此处不再赘述.

(2) 记叙文四项基本原则与人工评分相关分析.与表5相同,表6各变量间统计分析结果大体包括三类:人工评分之间的相关性、四项基本原则之间的相关性及四项基本原则和人工评分之间的相关性.首先,人工评分之间显著相关,该结果与表3中的人工评分相关分析结果一致.其次,在对记叙文的评价上,四项基本原则两两之间也呈现较高相关性.不过,四项基本原则与人工评分间的相关性不稳定.在人工评分1和四项基本原则的相关分析中,延续性原则、低价性原则和衔接性原则与人工评分1显著相关,显著性原则与其无显

著相关关系;人工评分2和任一原则都不存在显著相关关系.因此,记叙文中四项基本原则与人工评分间的相关性不甚明确,四项原则分析法对记叙文语篇局部连贯质量的评价效度尚不能定论.

表5 议论文四项基本原则与人工评分相关分析

Tab.5 Correlation analysis of the four principles and manual scores in argumentations

评分类别	Pearson 相关分析	人工评分 1	人工评分 2	延续性	显著性	低价性	衔接性
人工评分 1	Pearson 相关性	1	0.652 **	0.241	0.245	0.262	0.331 *
	显著性(双侧)	—	0.000	0.092	0.087	0.066	0.019
	N	50	50	50	50	50	50
人工评分 2	Pearson 相关性	0.652 **	1	0.245	0.173	0.189	0.296 *
	显著性(双侧)	0.000	—	0.086	0.229	0.189	0.037
	N	50	50	50	50	50	50
延续性	Pearson 相关性	0.241	0.245	1	0.711 **	0.914 **	0.580 **
	显著性(双侧)	0.092	0.086	—	0.000	0.000	0.000
	N	50	50	50	50	50	50
显著性	Pearson 相关性	0.245	0.173	0.711 **	1	0.866 **	0.707 **
	显著性(双侧)	0.087	0.229	0.000	—	0.000	0.000
	N	50	50	50	50	50	50
低价性	Pearson 相关性	0.262	0.189	0.914 **	0.866 **	1	0.628 **
	显著性(双侧)	0.066	0.189	0.000	0.000	—	0.000
	N	50	50	50	50	50	50
衔接性	Pearson 相关性	0.331 *	0.296 *	0.580 **	0.707 **	0.628 **	1
	显著性(双侧)	0.019	0.037	0.000	0.000	0.000	—
	N	50	50	50	50	50	50

** 在 0.01 水平(双侧)上显著相关; * 在 0.05 水平(双侧)上显著相关.

表6 记叙文四项基本原则与人工评分相关分析

Tab.6 Correlation analysis of the four principles and manual scores in narrations

评分类别	Pearson 相关分析	人工评分 1	人工评分 2	延续性	显著性	低价性	衔接性
人工评分 1	Pearson 相关性	1	0.532 **	0.347 *	0.191	0.310 *	0.389 **
	显著性(双侧)	—	0.000	0.014	0.184	0.029	0.005
	N	50	50	50	50	50	50
人工评分 2	Pearson 相关性	0.532 **	1	0.277	0.174	0.248	0.245
	显著性(双侧)	0.000	—	0.052	0.226	0.083	0.087
	N	50	50	50	50	50	50
延续性	Pearson 相关性	0.347 *	0.277	1	0.880 **	0.970 **	0.855 **
	显著性(双侧)	0.014	0.052	—	0.000	0.000	0.000
	N	50	50	50	50	50	50
显著性	Pearson 相关性	0.191	0.174	0.880 **	1	0.943 **	0.907 **
	显著性(双侧)	0.184	0.226	0.000	—	0.000	0.000
	N	50	50	50	50	50	50
低价性	Pearson 相关性	0.310 *	0.248	0.970 **	0.943 **	1	0.902 **
	显著性(双侧)	0.029	0.083	0.000	0.000	—	0.000
	N	50	50	50	50	50	50
衔接性	Pearson 相关性	0.389 **	0.245	0.855 **	0.907 **	0.902 **	1
	显著性(双侧)	0.005	0.087	0.000	0.000	0.000	—
	N	50	50	50	50	50	50

** 在 0.01 水平(双侧)上显著相关; * 在 0.05 水平(双侧)上显著相关.

(3)说明文四项基本原则与人工评分相关分析.在抽取的50篇说明文语料样本中,四项基本原则与人工评分相关分析结果如表7所示.

表7 说明文四项基本原则与人工评分相关分析
Tab.7 Correlation analysis of the four principles and manual scores in expositions

评分类别	Pearson 相关分析	人工评分 1	人工评分 2	延续性	显著性	低价性	衔接性
人工评分 1	Pearson 相关性	1	0.419 **	0.000	0.009	0.027	0.147
	显著性(双侧)	—	0.002	0.995	0.951	0.851	0.307
	N	50	50	50	50	50	50
人工评分 2	Pearson 相关性	0.419 **	1	0.151	0.105	0.182	0.116
	显著性(双侧)	0.002	—	0.294	0.466	0.206	0.423
	N	50	50	50	50	50	50
延续性	Pearson 相关性	0.000	0.151	1	0.845 **	0.955 **	0.703 **
	显著性(双侧)	0.995	0.294	—	0.000	0.000	0.000
	N	50	50	50	50	50	50
显著性	Pearson 相关性	0.009	0.105	0.845 **	1	0.920 **	0.865 **
	显著性(双侧)	0.951	0.466	0.000	—	0.000	0.000
	N	50	50	50	50	50	50
低价性	Pearson 相关性	0.027	0.182	0.955 **	0.920 **	1	0.817 **
	显著性(双侧)	0.851	0.206	0.000	0.000	—	0.000
	N	50	50	50	50	50	50
衔接性	Pearson 相关性	0.147	0.116	0.703 **	0.865 **	0.817 **	1
	显著性(双侧)	0.307	0.423	0.000	0.000	0.000	—
	N	50	50	50	50	50	50

** 在 0.01 水平(双侧)上显著相关.

由表7可知,尽管说明文语料样本人工评分与各项基本原则内部均显著相关,但人工评分1、人工评分2与四项基本原则间的相关性并不显著.换言之,在说明文二语写作语篇局部连贯质量的评价问题上,人工评分和向心理论的四项原则分析法存在较大的差异.因此就本文语料而言,四项原则分析法目前尚不能用于说明文二语写作语篇局部连贯质量的自动评价.

4 结束语

本研究以中国学生英语口语语料库的笔语子库为语料来源,分析比较了四项原则分析法应用于议论文、记叙文和说明文三种不同体裁二语书面语语篇局部连贯评价的有效性.研究发现:(1)不同体裁二语写作人工评分与 Kibble 评分之间的相关性存在差异.议论文中人工评分与 Kibble 评分显著相关;记叙文和说明文的人工评分与 Kibble 评分的相关性相对较弱.(2)除衔接性原则外,四项基本原则在二语写作中的比例分布因作文体裁不同而各有特点.议论文中延续性原则所占比例较高,显著性原则和低价性原则占比较低;记叙文中显著性原则和低价性原则占比较高,延续性原则占比较低;说明文各项原则占比多介于议论文和记叙文之间.(3)四项基本原则和人工评分之间的相关性因作文体裁而异.议论文中衔接性原则与人工评分显著相关;记叙文中各项原则与人工评分之间的相关性受评分员影响而有波动;说明文中各项原则与人工评分无显著相关关系.

纵观本文研究结果,四项原则分析法对议论文语篇局部连贯评价效度优于记叙文和说明文,其中衔接性原则与议论文语篇局部连贯质量相关度最高,可作为议论文二语写作语篇局部连贯评价指标纳入自动评分系统.不过,由于本研究的语料限制以及其他主客观因素的影响,此研究结果是否具有普适性仍有待进一步验证.

参 考 文 献

- 18(1):91-126.
- [2] 秦晓晴,文秋芳.中国大学生英语写作能力发展规律与特点研究[M].北京:中国社会科学出版社,2007.
- [3] 陈海庆,张绍杰.语篇连贯:言语行为理论视角[J].外语教学与研究,2004,48(6):420-426.
CHEN H Q,ZHANG S J.Coherence in Discourse:A Perspective from Speech Act Theory[J].Foreign Language Teaching and Research, 2004,48(6):420-426.
- [4] 冯德正,张艳,王艳.修辞结构理论在多模态语篇分析中的应用[J].当代修辞学,2016,35(5):49-58.
FENG D Z,ZHANG Y,WANG Y.The Application of Rhetorical Structure Theory in The Analysis of Multimodal Discourse[J].Contemporary Rhetoric,2016,35(5):49-58.
- [5] 刘国兵.词汇衔接理论视角下的英语学习者书面语篇连贯评价研究[J].中国外语教育,2016,9(2):33-43.
LIU G B.The Evaluations of Discourse Coherence:A Perspective from Theory of Lexical Cohesion[J].Foreign Language Education in China,2016,9(2):33-43.
- [6] 刘国兵.基于 WordNet 语义知识库的英语学习者作文局部连贯自动评价[J].河南师范大学学报(自然科学版),2016,57(6):149-158.
LIU G BAutomatic Evaluation of Local Coherence in Chinese EFL Learners' Essays with WordNet[J].Journal of Henan Normal University(Natural Science Edition),2016,57(6):149-158.
- [7] 刘国兵.基于衔接特征的英语学习者书面语篇连贯自动评价研究[M].北京:外语教学与研究出版社,2017.
- [8] 刘国兵.大学英语学习者书面语篇连贯自动评价模型的创建[J].解放军外国语学院学报,2017,40(5):89-95.
LIU G BAutomatic Evaluation of Textual Coherence in Chinese EFL Learners' Written English[J].Journal of PLA University of Foreign Languages,2017,40(5):89-95.
- [9] 洪明,徐艳霞.中国高校英语专业学生写作局部连贯能力的发展特征——一项基于向心理论的参数化研究[J].浙江师范大学学报(社会科学版),2015,40(1):117-124.
HONG M,XU Y X.On Developmental Features of University English Majors' Ability in Achieving Local Coherence in English Writing:A Parametric Study Based on Centering Theory[J].Journal of Zhejiang Normal University(Social Sciences),2015,40(1):117-124.
- [10] 苗兴伟.语篇向心理论述评[J].当代语言学,2003,31(2):149-157.
MIAO X W.A Survey of Centering Theory in Discourse[J].Contemporary Linguistics,2003,31(2):149-157.
- [11] KIBBLE R.A Reformulation of Rule 2 of Centering Theory[J].Computational Linguistics,2001,27(4):579-587.
- [12] MILTSTAKAKI E,KUKICH K.Automated Evaluation of Coherence in Student Essays[C]//In Proceedings of the Workshop on Language Resources and Tools in Educational Applications.Massachusetts:MIT Press,2000:1-8.
- [13] HASLER L.An Investigation into the Use of Centering Transitions for Summarization[C]//In Proceedings of the 7th Annual CLUK Research Colloquium.Birmingham:University of Birmingham Press,2004:100-107.
- [14] KARAMANIS N.Entity Coherence for Descriptive Text Structuring[D].Edinburgh:University of Edinburgh,2003.
- [15] KARAMANIS N.Supplementing Entity Coherence with Local Rhetorical Relations for Information Ordering[J].Journal of Logic, Language and Information,2007,16(4):445-464.
- [16] KIBBLE R,POWER R.Using Centering Theory to Plan Coherent Texts[C]//In Proceedings of the 12th Amsterdam Colloquium.Amsterdam:ILLC,1999:187-192.
- [17] KARAMANIS N.Exploring Entity-based Coherence[C]//In Proceedings of the 4th Computational Linguistics UK Research Colloquium.Edinburgh:University of Edinburgh Press,2001:18-26.
- [18] 洪明.向心理论的局部连贯模式与二语写作质量评价[D].上海:上海外国语大学,2009.
HONG M.Application of Centering Local Coherence Modal in Second Language Writing Evaluation[D].Shanghai International Studies University,2009.
- [19] 洪明.向心局部连贯模式的发展及其仍需解决的问题[J].浙江师范大学学报(社会科学版),2011,36(1):77-83.
HONG M.Latest Development in the Centering Local Coherence Model and Its Problems[J].Journal of Zhejiang Normal University(Social Sciences),2011,36(1):77-83.
- [20] 王德亮.语篇连贯性的量化测量——基于向心理论的研究[C]//中国计算语言学研究前沿进展(2007-2009).北京:清华大学出版社,2009:289-294.
WANG D L.Quantified Measurement of the Discourse Coherence:A Study Based on Centering Theory[C]//Advances of Computational Linguistics in China(2007-2009).Beijing:Tsinghua University Press,2009:289-294.
- [21] 许余龙.向心理论的参数化研究[J].当代语言学,2008,36(3):225-236.
XU Y L.A Parametric Approach to Centering Theory[J].Contemporary Linguistics,2008,36(3):225-236.
- [22] POESIO M,STEVENSON R,EUGENIO B D,et al.Centering:A Parametric Theory and Its Instantiations[J].MIT Press,2004,30(3):309-363.
- [23] GROSZ B J,SIDNER C L.Attention,Intentions,and the Structure of Discourse[J].Computational Linguistics,1986,12(3):175-204.
- [24] POLIO C G.Measures of Linguistic Accuracy in Second Language Writing Research[J].Language Learning,1997,47(1):101-143.

- [25] FOSTER P, TONKYN A, WIGGLESWORTH G. Measuring Spoken Language: A Unit for All Reasons[J]. *Applied Linguistics*, 2000, 21(3): 354-375.
- [26] 梁茂成, 文秋芳. 国外作文自动评分系统评述及启示[J]. *外语电化教学*, 2007, 29(5): 18-24.
LIANG M C, WEN Q F. A Critical Review and Implications of Some Automated Essay Scoring Systems[J]. *Computer-Assisted Foreign Language Education*, 2007, 29(5): 18-24.
- [27] HAHN S, STRUBE M. Incremental Centering and Center Ambiguity[C]//In Proceedings of the 18th Annual Conference of the Cognitive Science Society. New Jersey: Lawrence Erlbaum Associates, 1996: 568-573.
- [28] MILTSAKAKI E. The Syntax-Discourse Interface: Effects of the Main Subordinate Distinction on Attention Structure[D]. Philadelphia: Philadelphia University of Pennsylvania, 2003.

A study on the efficiency of centering theory in evaluating local coherence of english learners' written language

Liu Guobing¹, Chang Fangling²

(1. Faculty of International Studies, Henan Normal University, Xinxiang 453007, China;

2. School of Chinese Language and Literature, Beijing Foreign Studies University, Beijing 100089, China)

Abstract: The thesis extracted texts from Spoken and Written English Corpus of Chinese Learners and then respectively analyzed validities of Four Principle Model in evaluating argumentation, narration and exposition. The results are shown as follows: (1) There are discrepancies in the correlation between manual score and Kibble score of second language writing with different genres. In argumentation, Kibble score and manual score are significantly correlative; nevertheless, the correlation between the two variables is not significant in narration and exposition. (2) Except the principle of cohesion, the other principles in the Four Principle Model have different percentages with the variations of genres. The principle of continuity has higher proportions in argumentations than narrations while the principle of salience and the principle of cheapness have lower proportions in argumentations than narrations. In addition, the proportions of the three principles in expositions are mostly between the proportions of them in argumentations and narrations. (3) The degree of correlation between the four basic principles and the manual score varies with the changes of compositional genres. The principle of cohesion and manual score are significantly correlative in argumentations; in narrations, the degree of correlation between principles and manual scores changes with the variations of scorers; and in expositions, the correlation between the two variables is not significant. In conclusion, the Four Principle Model is more effective in evaluating local coherence of argumentations than that of narrations and expositions. What's more, the principle of cohesion is closely connected to qualities of argumentations and it can be regarded as an index of automatic essay scoring system to evaluate local coherence of argumentation.

Keywords: centering model; local coherence; Four Principle Model; second language writing; compositional genre

[责任编辑 陈留院 赵晓华]

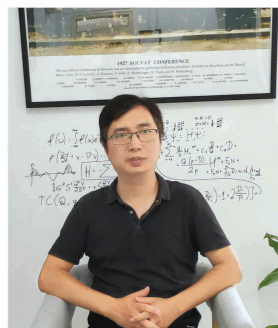


本期专家介绍



刘国兵,河南师范大学教授,硕士生导师,外国语学院副院长.毕业于北京外国语大学,获外国语言学及应用语言学专业博士学位.研究方向为语料库语言学、计算语言学.中国语料库语言学会常务理事、国际应用语言学学会会员,全国中学通用教材《英语》(新标准初一至高三必修 CD-ROM)核心编委,河南省教师教育专家,河南省中等职业教育师资培训专家,河南省高级翻译人才库副召集人.主持国家社科基金项目、教育部人文社科项目、河南省哲学社科项目等省部级以上项目 10 项;在国内外期刊发表学术论文 40 余篇;出版专著 3 部、译著 2 部、教材 8 部.2014 年入选“河南省科技创新人才支持计划”,同年被授予河南省青年五四奖章.2018 年被评为全国教育硕士优秀教师.

吴健聘,扬州大学教授,博士生导师.2012 年毕业于北京师范大学获博士学位,2012 年 8 月至 2013 年 9 月在韩国西江大学量子时空中心从事博士后研究.2012 年 7 月至 2018 年 4 月在渤海大学从事科研和教学工作.2018 年 4 月起,进入扬州大学引力与宇宙学研究中心,从事科研与教学工作,入选扬州大学拔尖人才成长计划.主要从事引力的全息性质、规范引力对偶及其应用、量子引力、黑洞物理学等研究.在 *Phys Rev Lett*, *Phys Rev D*, *JHEP* 等国际物理期刊发表 SCI 学术论文 70 余篇,论文总引用达 1 400 多次,H 指数为 22,扩展版 ESI 高被引论文(扩展至前 3%) 3 篇.先后入选辽宁省杰出青年学者成长计划,辽宁省“百千万人才工程”千人层次,获中国引力与相对论天体物理学会 2013 年度青年学者优秀论文奖并列三等奖.主持国家自然科学基金项目 2 项.



杨清香,河南师范大学特聘教授,博士生导师,河南师范大学生态学 and 微生物学学科带头人,河南省农业微生物生态与技术国际联合实验室主任,河南省黄河流域生态工程技术研究中心主任,黄淮水环境与污染防治省部共建教育部重点实验室学术带头人,资源微生物与功能分子河南省科技创新型团队带头人,教育部“新世纪优秀人才支持计划”获得者,河南省科技创新杰出青年计划获得者,河南省科技创新杰出人才计划获得者,河南省教育厅学术技术带头人,中国生态学会微生物生态专业委员会委员,河南省生物工程学会副理事长.国家自然科学基金的通讯评审专家,国际多家 SCI 期刊评审专家,河南省“师德标兵”获得者,河南省“五一”劳动奖章获得者.

杨清香博士 1988 年毕业于河南师范大学,1997 年在武汉大学获得硕士学位,2003 年在中国科学院生态环境研究中心环境水质学国家重点实验室获得博士学位.先后在德国亥姆霍兹环境健康研究中心慕尼黑中心、澳大利亚昆士兰大学高级水管理中心、日本独立行政法人高级产业技术研究院、美国肯塔基大学食品土壤与环境学院等单位从事合作研究和访问学者工作.目前主持完成和在研 NSFC-河南省联合基金重点项目 1 项、国家自然科学基金面上项目 4 项,主持完成教育部新世纪优秀人才支持计划、教育部博导基金、河南省杰出青年基金、河南省杰出创新人才计划、河南省高校创新团队项目以及科技部重点基础研究发展计划“973 计划”子课题等项目 10 余项,发表学术论文 100 余篇,其中在 SCI 一区、二区发表论文 40 余篇,发明国家专利 8 件(5 件授权),主持获得河南省科技进步二等奖 1 项,成果鉴定 3 项,编写著作 5 部.