

# 二语句法复杂度分析器 及其在二语写作研究中的应用

广东外语外贸大学/[美]宾夕法尼亚州立大学 陆小飞  
广东外语外贸大学 许琪

**提要:**句法复杂度是二语写作教学与研究中的一个重要概念,它与二语学习者的语言水平、语言发展过程及写作质量之间的关系成为众多二语写作研究的重点。本文首先简要回顾国内外二语写作句法复杂度研究的进展及其面临的挑战,进而介绍二语句法复杂度分析器的设计原理和功能,及其在二语写作句法复杂度研究中的应用。该分析器可用十四种句法复杂度测量指标对英语写作样本进行自动分析,从而解决了句法复杂度研究中的数据分析瓶颈,有效推动了二语写作句法复杂度前沿研究。最后本文对其在二语写作句法复杂度未来研究中的应用前景进行展望。

**关键词:**二语写作、句法复杂度、二语句法复杂度分析器

[中图分类号] H319 [文献标识码] A [文章编号] 1000-0429(2016)03-0409-12

## 1. 引言

句法复杂度指语言产出中句法结构的多样性及复杂性(Ortega 2003, 2012; Lu 2011; Bulté & Housen 2014)。作为二语写作教学与研究中的重要概念之一,句法复杂度与二语学习者的语言水平、语言发展过程及写作质量之间的关系成为众多二语写作研究的重点。与此同时,二语写作句法复杂度还受到学习者因素、写作任务和教学写作环境等相关变量的影响,对此,研究者也开展了深入研究。就测量方法而言,研究者已提出数十种分析二语写作句法复杂度的测量指标(Wolfe-Quintero *et al.* 1998)。经过近40年的发展,该领域研究者对二语学习者的句法复杂度发展过程及测量方法的理解日渐完善。然而,由于长期以来缺乏句法复杂度自动分析工具,直至近年,二语写作句法复杂度研究仍遇到双重瓶颈,一是数据分析规模有限,二是测量指标数量有限。这两大瓶颈在一定程度上影响了前人研究中一些结论的可靠性和进一步推广(Ortega 2003; Lu 2011)。

针对以上问题,本文首先简要回顾国内外二语写作句法复杂度研究的进展及其所面临的挑战,进而介绍二语句法复杂度分析器(Lu 2010)的设计原理和功能及其在二语写作句法复杂度研究中的应用。该分析器可用十四种常用句法复杂度测量指标对英语写作文本进行自动分析,从而解决了句法复杂度研究中的数据分析瓶颈,有效推动了二语写作句法复杂度前沿研究的进展。文末对其在二语写作句法复杂度未来研究中的应用前景和方向进行展望。

## 2. 二语写作句法复杂度研究进展及挑战

### 2.1 国外研究进展

近四十年来,国外研究者对二语写作句法复杂度的关注度不断提高。该领域的研究可分为以下三个主要路径:1)从理论构建和研究方法的角度出发,探寻二语写作句法复杂度的组成维度和可以有效衡量各维度的测量指标;2)从二语习得、发展或测试的角度出发,探寻句法复杂度与语言水平和写作质量之间的关系及句法复杂度的发展过程;3)从二语习得、发展或测试的角度出发,探寻与学习者背景、写作任务或教学写作环境相关的因素对二语写作复杂度的影响。

从理论构建和研究方法的角度出发,研究者逐渐认识到应将句法复杂度理解为一个多维度构念,并且每个维度都需要多种指标进行全面测量。Wolfe-Quintero *et al.* (1998)分类总结了二语写作研究中使用到的多种句法复杂度测量指标,不过,他们没有定义句法复杂度的多维度特征。在此基础上,Norris & Ortega(2009)明确提出应将句法复杂度理解为一个包括5个子维度的构念,包括从属结构使用量、句子整体复杂度、子句的短语扩展程度、并列结构使用量和句子结构的多样性、复杂程度及习得顺序。除了对句法复杂度在概念上的理解外,研究者对测量指标的效度也进行了探讨。例如,Biber *et al.* (2011)通过比较英语学术写作语料和口语语料,发现前者使用了更多的复杂名词性短语,而后者则使用了更多的限定性从属子句。他们因此认为二语写作句法复杂度研究不应使用基于T单位的测量指标,而应更关注复杂名词性短语的使用。

从二语习得、发展或测试的角度出发,研究者发现句法复杂度的各维度与二语学习水平和写作质量有不同程度的相关;各维度并非同时均衡线性地发展,而是在不同的语言发展阶段可能有不同的发展模式。这一路径的研究包括大量的横向和纵向研究。横向研究主要通过比较代表不同学习水平或写作质量的二语写作语料,探究可有效区别不同学习水平或写作质量的句法复杂度测量指标(Ortega 2003; Lu 2011; Verspoor *et al.* 2012; Ai & Lu 2013; Taguchi *et al.* 2013)。纵向研究则通过分析同一组学习者在不同时间点的多次写作语料,探寻句法复杂度的发展模式(Ortega 2003; Stockwell & Harrington 2003; Vyatkina

2013; Bulté & Housen 2014; Crossly & McNamara 2014)。Norris & Ortega (2009)总结了该路径的研究成果,指出句法复杂度与学习水平有如下关系:初级水平学习者较多使用并列结构;到中级阶段,学习者开始使用更多的从属子句;高水平学习者则更多使用复杂短语。

最后,在第三个路径的研究中,研究者发现多种与学习者背景、写作任务或教学写作环境相关的因素对二语写作句法复杂度有不同程度的影响。这些因素包括作文是否限时(Lu 2011)、限时的长短(Ellis & Yuan 2004)、写作是否在线进行(Sotillo 2000)、写作文体类别(Way *et al.* 2000; Lu 2011)、母语或非母语教学环境(Ortega 2003),等等。同时,研究者也发现有些因素,如任务复杂度对二语写作复杂度没有显著影响(Kuiken & Vedder 2008)。

## 2.2 国内研究进展

通过查阅 2005-2015 年 10 年间国内出版的常用外语类期刊,我们发现国内二语写作句法复杂度的相关研究主要集中在上节讨论的三个研究路径中的部分子路径上:1)句法复杂度的测量指标效度;2)句法复杂度与学习者语言水平的关系;3)写作任务对句法复杂度的影响。其中,很多研究都将 T 单位作为衡量句法复杂度的重要指标,且研究者大多依靠人工标注统计的方式测量句法复杂度。

在考察不同测量指标效度的研究中,陈慧媛(2010)通过对比 T 单位长度和每个 T 单位中的 S 节数量,发现 T 单位长度更能有效反映句法复杂度。赵俊海、陈慧媛(2012)考察了每个 T 单位中的从属句数量、从属句、非限定动词和衔接成分,发现这些指标均能有效反映句法复杂度,但非限定动词的作用优于其他三个指标,而从属句的作用较弱。这些结果对部分测量指标的效度提供了有用信息,不过这一路径研究的系统性和全面性尚显缺乏。

在句法复杂度与学习者语言水平关系的相关研究中,研究使用的测量方法不同,产生的结果也有较大差异。纪小凌(2009)、王静萍(2013)分别把句法复杂度操作化为每个 T 单位中的小句数量和从句数量,二者研究的结果均显示学习者的语言水平与写作中的句法复杂度无显著关联。鲍贵(2009)将句法复杂度分为单位长度和子句密度两类,发现学习者写作的长度指标增长速度快于密度指标。徐晓燕等(2013)发现学生写作的单位长度和密度与年级和写作水平呈线性增长趋势,但就句式类型而言,简单句、状语、宾语和表语从句的使用随年级升高而减少,而紧缩子句和被动句的使用则随写作水平提高而增多。总体而言,该路径的相关研究表明句法复杂度各个维度与学习者语言水平之间的关系可能不尽相同。然而,由于现有研究并未深入探讨句法复杂度的所有维度,因此对句法复杂度与学习者语言水平关系的理解还缺乏系统性和全面性。

国内学者在写作任务对句法复杂度影响方面的研究相对较少。在我们查阅

的相关文献中,研究者通常采用单一测量指标衡量句法复杂度。刘春燕(2009)、蔡艳玲(2012)通过测量 T 单位长度发现不同的写作任务完成方式对句法复杂度没有显著影响。王静萍(2013)通过测量 T 单位中的从句数量发现任务复杂度对句法复杂度也没有影响。值得一提的是,还有一些研究使用 T 单位以外的测量指标调查了写作任务对作文语言复杂度的影响(曾路、伍忠杰 2010;张新玲、周燕 2014),例如作文类符形符比、小句分词数、平均词长等。总体而言,与国外研究相比,国内学者对其他与学习者因素和教学写作环境相关变量对二语写作复杂度影响的研究尚显缺乏。

### 2.3 数据分析瓶颈造成的挑战与自动标注工具的出现

长期以来,句法复杂度自动分析工具的缺乏和人工标注巨大的工作量给二语写作句法复杂度研究的数据分析规模及数据分析中使用的测量指标数量带来了双重瓶颈。例如,Ortega(2003)综述了 20 余篇国外期刊论文,其中,除 4 篇论文使用了 4 或 5 个测量指标外,其他论文均用到 1~3 个测量指标,研究分析的样本大小在 16~300 之间(均值为 84)。Lu(2011)指出,这种数据分析规模的局限性在 2003-2010 年间的国外研究中依然持续。同时,在我们查阅到的 2005-2015 十年间发表于国内常用外语类期刊的相关文献中,研究者使用到的测量指标基本都在 1~4 个之间,而样本大小则在 36~480 之间(均值为 183)。这种数据分析的瓶颈对二语写作句法复杂度研究至少有三方面负面影响。首先,数据分析规模偏小会直接影响所得结论的可靠性和可推广性。第二,由于不同研究使用的测量指标和分析的句法复杂度维度不尽相同,且研究方法的其他方面(如对学生语言水平的操作化)各异,导致不同研究所汇报的结果难以直接对比,甚至互相冲突。第三,大量研究只使用少量句法复杂度测量指标或者只涉及部分句法复杂度维度,这也使得该领域的研究结果缺乏全面性和系统性。

随着语料库语言学的发展,近几年可自动分析句法复杂度的软件逐渐出现并趋于成熟。例如,Coh-Matrix(Graesser *et al.* 2004)作为分析语篇连贯性的工具,也加入了一些测量句法复杂度(如名词短语中的修饰成分数量)和句型密度(如每千词中的名词短语数量)的指标,因此近年也开始在二语写作句法复杂度的研究中得到应用。Crossley & McNamara(2014)利用 Coh-Matrix 中的 9 种指标分析了 57 个英语学习者在四个月期间的作文,结果显示,随着时间的推移,学习者作文的句法复杂度呈显著上升趋势,但只有从句数量的减少与学生作文得分的提高有显著关联,而其他测量指标的变化(如名词短语中修饰成分数量的增多)则不能有效预测学生作文得分的变化。国内部分学者也开始使用该软件进行句法复杂度研究。例如,李雪莲(2014)以全国英语专业四级作文为语料,通过 Coh-Matrix 分析发现,学校类型和作文水平对句法复杂度没有显著影响。

值得注意的是,尽管 Coh-Metrix 提供了大量有用的测量语篇连贯性的指标,其作为二语写作句法复杂度的研究工具尚有很大的局限性。首先,Coh-Metrix 所包含的句法复杂度测量指标有限,不能全面系统地反映句法复杂度的组成维度。其次,作为一个测量语篇连贯性的工具,Coh-Metrix 所包含的句法复杂度测量指标并非直接来源于二语写作句法复杂度的研究。因此,在这些指标衡量二语写作句法复杂度的有效性得到完全验证之前,我们仍需谨慎对待相关的研究结果。再次,该工具目前只能通过在线分析界面逐篇处理 15,000 字符以内的文本,这对数据处理的规模和效率都有一定的影响。

另一个日益受到二语写作句法复杂度研究者关注的工具是二语句法复杂度分析器(L2 Syntactic Complexity Analyzer,简称 L2SCA; Lu 2010)。与 Coh-Metrix 包含的句法复杂度测量指标不同,该分析器涵盖的指标直接来自二语写作句法复杂度研究。下一节将重点介绍该分析器的设计原理和功能。

### 3. 二语句法复杂度分析器的设计原理及功能

表 1. L2SCA 句法复杂度测量指标

类别	指标
第一类 单位长度	1) 平均子句长度 2) 平均句子长度 3) 平均 T 单位长度
第二类 句子复杂度	每个句子中的子句数量
第三类 从属子句使用量	1) 每个 T 单位中的子句数量 2) 复杂 T 单位比率(即每个 T 单位中的复杂 T 单位数量) 3) 从属子句比率(即每个子句中的从属子句数量) 4) 每个 T 单位中的从属子句数量
第四类 并列结构使用量	1) 每个子句中的并列短语数量 2) 每个 T 单位中的并列短语数量 3) 并列句比率(即每个句子中的 T 单位数量)
第五类 特定短语结构	1) 每个子句中的复杂名词性短语数量 2) 每个 T 单位中的复杂名词性短语数量 3) 每个 T 单位中的动词短语数量

二语句法复杂度分析器的设计动机是为了解决二语写作句法复杂度研究中的数据分析瓶颈,以帮助二语写作研究者对句法复杂度的相关问题展开更有效、更系统的研究。Lu(2010)从 Wolfe-Quintero *et al.* (1998)和 Ortega(2003)的综述中提到的二语写作句法复杂度测量指标中选取了 14 种实现自动化分析。其

中的 11 种指标均经前人研究显示与学习水平、二语发展或写作质量有显著关系,而其他 3 种则为 Wolfe-Quintero *et al.* (1998)向二语写作研究者推荐使用的指标。这 14 种指标可分为五类(见表 1)。

L2SCA 要求待分析的英语作文为纯文本文件,且文件中没有除作文本身以外的其他标注信息。句法复杂度的分析分三步。L2SCA 首先使用斯坦福句法剖析器(Stanford Parser; Klein & Manning 2003)对作文中的句子逐一进行句法结构分析并生成句法树。斯坦福句法剖析器兼有自动断句、分词和词性赋码功能,因此待分析的文件无需其他预处理。例如,对例(1)中的句子,斯坦福句法剖析器生成例(2)中的句法树。

(1) Mary is nice and smart.

(2) (ROOT

(S

(NP (NNP Mary))

(VP (VBZ is)

(ADJP (JJ nice)

(CC and)

(JJ smart)))

(. .)))

(Root=句法树根,S=简单陈述句,NP=名词短语,NNP=单数专有名词,VP=动词短语,VBZ=动词现在式,ADJP=形容词短语,JJ=形容词,CC=并列连词。)

句法剖析结束后,L2SCA 先根据句法树中所包含的分词和词性赋码信息统计出作文中除标点符号以外的单词数量,然后使用 Tregex(Levy & Andrew 2006)对句法树逐一检索以统计出以下语言单位在作文中的出现频率:句子、子句、T 单位、从属子句、复杂 T 单位、并列短语、复杂名词性短语和动词短语。Lu (2010)对这些语言单位均给出了清晰的定义以明确检索目标。作为一个句法树检索工具,Tregex 可根据用户提供的符合其特定句法要求的正则表达式找出句法树中所有满足条件的句法树节。L2SCA 为上述每个语言单位均设计了一个或多个正则表达式,以涵盖其定义中所包含的各种结构。由于 Tregex 的正则表达式句法较为复杂,我们在此仅举一例简单说明,详细讨论可参见 Lu(2010, 2014)。例如,Lu(2010)将并列短语定义为并列形容词、副词、名词或动词短语;L2SCA 使用例(3)中的正则表达式来检索这些短语。使用该表达式,Tregex 会在例(2)的句法树中检索到一个匹配的并列形容词短语。

(3) ADJP|ADVP|NP|VP<CC

(ADVP=副词短语,“|”表示“或”,“<”表示“直接包含”。)

在完成所有语言单位的检索和统计后, L2SCA 利用相关单位的数量计算出上述 14 种句法复杂度测量指标值。每篇作文最终的分析结果包括了上述 9 种语言单位(即单词、句子等)的出现频率和 14 种句法复杂度测量指标的数值。

Lu(2010)使用中国学生英语口语笔语语料库(文秋芳等 2005)中的 40 篇作文对 L2SCA 分析的准确性进行了评估。与人工标注比较的结果显示, 该分析器对计算上述 14 种测量指标需要用到的 8 种语言单位(单词数量除外)的 F 值在 .83~1.000 之间, 而自动计算与人工计算得出的 14 种测量指标值的相关度在 .834~1.000 之间。Lu(2010)发现分析结果中的误差主要来源于句法剖析的错误, 而后者又主要来源于作文中的连写句、不完整句子和部分有结构歧义的句子。同时, 中国学生英语笔语语料库代表较高水平学习者的写作, 因此其中的语言问题大多集中在语言使用的地道性(如搭配错误)和准确性(如介词使用错误)两方面, 而并非句子的完整性。这两类错误对句法结构分析的影响很小, 从而保证了 L2SCA 分析结果的准确性。其他一些研究者对 L2SCA 的准确性也进行了评估。例如, Yoon & Polio(forthcoming)随机选择了 30 篇成人英语学习者的作文, 对除每个句子中的子句数量和复杂 T 单位比率外的其他 12 种测量指标进行了人工标注。他们发现, 除并列句比率外( $r=.74$ ), L2SCA 自动计算与人工计算得出的测量指标值的相关度均在 .81 以上。基于这些结果, Lu(2010)和 Yoon & Polio(forthcoming)均认为 L2SCA 可以很好地分析较高水平学习者的作文。在使用 L2SCA 分析口语语料或初级学习者的作文时, 研究者应该对文件中的连写句和不完整的句子等做必要的更正, 以确保分析的准确度。

#### 4. 二语句法复杂度分析器在二语写作研究中的应用

对于二语写作句法复杂度研究而言, L2SCA 与其他句法复杂度分析工具相比具有理论依托和大规模数据处理能力的双重优势, 因此, 自发布以来日益受到该领域研究者的关注和使用。L2SCA 解决了句法复杂度研究的数据分析瓶颈, 使研究者可以用更多的测量指标分析更大规模的写作数据。这种新的数据分析能力不仅有效提高了数据分析的效率和研究结果的丰富度与可靠性, 更重要的是, 它还可以帮助研究者回答一些之前难以解答的重要研究问题。

首先, 从研究方法的角度出发, 研究者曾经提出句法复杂度测量指标繁多, 而有些指标提供的信息可能互相重复(Norris & Ortega 2009)。L2SCA 的数据自动分析功能可以帮助研究者直接验证哪些测量指标互相重复。如 Lu & Ai(2015)用 L2SCA 的 14 种测量指标比较了 8 组不同母语背景的大学生英语作文, 结果显示, 复杂 T 单位比率、从属子句比率及每个 T 单位中的从属子句数量所提供的关于组间差异的信息完全一致; 另外, 每个子句中的复杂名词性短语数

量与每个 T 单位中的复杂名词性短语数量所提供的信息也完全一致。因此,作者建议在需要选取少量指标进行研究时,从这三种从属子句使用量指标和两种名词性短语复杂度指标中各选一种即可。

其次,从二语习得的角度出发,L2SCA 可帮助研究者更系统地比较不同测量指标与学习水平之间的关系。如 Lu(2011)使用 L2SCA 中的 14 种指标分析了中国学生英语口语语料库中四个年级的大学生作文,研究结果系统显示了哪些指标可显著区别不同年级的作文,以及各指标从低年级到高级的变化趋势与相互作用。另外,在二语发展的研究中,研究者可以利用 L2SCA 分析大量的语言发展数据以系统观察学习者句法复杂度的动态发展过程。如 Chan *et al.* (2015)用 L2SCA 分析了一对双胞胎英语学习者在 8 个月时间内产生的大量口语语料,结果显示这两位学习者的口语句法复杂度的发展均先于笔语句法复杂度的发展;不过,在随后的发展中,其中一位的口语句法复杂度一直持续高于笔语,而另一位笔语句法复杂度则逐渐超越了口语。Mazgutova & Kormos (2015)利用 L2SCA 分析了学习者在参加为期一个月的学术英语课程开始和结束阶段的作文,结果显示虽然该课程并未明确强调句法教学,但学习者在课程结束时的作文依然呈现了更高的句法复杂度,尤其是学习水平较低的学习者使用了更多的复杂名词性短语。

第三,在句法复杂度与二语写作质量关系的研究中,L2SCA 可使用多个测量指标对大量数据进行分析,这样一来,研究者便可以系统地比较不同测量指标对写作质量的解释力。如 Yang *et al.* (2015)使用 L2SCA 分析了 190 个研究生英语学习者的 380 篇作文(每个学习者两篇不同题目的议论文),结果发现反映句子整体复杂度的测量指标(包括句子长度和 T 单位长度)与两个不同题目的作文得分均有显著关联,而其他反映局部结构复杂度的指标与不同命题作文的得分关系则有所不同。邵银波等(2012)利用 L2SCA 统计出英语专业八级作文的复杂度数据,主要考察了 T 单位长度、从句长度和每个 T 单位中的复杂名词性短语数量三个指标,结果表明,作文得分与句法复杂度呈正相关,其中,T 单位长度和从句长度最能反映高质量作文的复杂度。除此以外,研究者还可以通过因子分析更好地发现不同测量指标之间的关系及不同因子对写作质量的解释力。张艳敏等(2015)使用 L2SCA 和可读性分析器(Readability Analyzer)中的多个测量指标分析了中国学生口语语料库中的限时作文,结果显示这些指标构成四个因子;这些因子可解释的作文成绩的方差递减,而累计解释作文成绩的方差则达到 91.13%。

最后,部分与学习者因素、写作任务和教学写作环境相关的变量可有大量不同取值,因此,如要全面系统地了解这些变量及各变量之间的相互作用对二语写



作句法复杂度的影响,需进行大规模数据处理。例如,要系统研究学习者的母语背景对二语写作句法复杂度的影响,需要使用大量不同的测量指标分析大规模来自不同母语的 learners 数据,因此这方面的研究在 L2SCA 出现之前基本缺失。Lu & Ai(2015)使用 L2SCA 分析了 1,600 篇来自 8 种不同母语背景的大学生英语议论文写作,发现前人仅基于学习水平的理论(如 Norris & Ortega 2009)不能全面阐释数据中出现的句法复杂度模式,而学习者的母语背景对句法复杂度的各个维度均有显著影响。

国内使用二语句法复杂度分析器的研究相对较少。除邵银波等(2012)和张艳敏等(2015)使用 L2SCA 研究句法复杂度和写作质量的关系外,其他方面的研究尚不多见。

## 5. 二语写作句法复杂度研究的前景

在今后的二语写作句法复杂度相关研究中,研究者可以充分利用 L2SCA (或其他有效工具)的数据分析能力,收集并分析更为全面和丰富的语料,更深入系统地研究句法复杂度的多种维度,多种相关构念(包括语言水平、语言发展和写作质量),以及与学习者因素、写作任务和教学写作环境相关的多种变量之间的相互关系和动态影响。此外,研究者对句法复杂度与语言发展和写作质量所涉及的其他重要维度之间的关系也可展开更全面的研究。在句法复杂度测量指标效度方面的研究上,研究者可以通过分析不同类型的二语写作数据系统地比较不同工具中(如 Coh-Metrix 和 L2SCA)所包含的不同测量指标对学习水平、语言发展和写作质量的预测与解释能力。这不仅是从研究方法的角度出发对比不同工具的效度,也是从理论的角度出发对比来自不同领域的测量指标对二语写作句法复杂度分析的效度。

为了更深入了解二语写作句法复杂度不同维度与学习水平和写作质量之间的关系,研究者可以同时研究句法复杂度与这两者的关系。这类研究需要搜集大量不同水平学习者不同写作质量的语料,然后对这些语料的句法复杂度进行多维度分析。这种更丰富的分析层次可帮助我们了解句法复杂度测量指标是否可以同时区分学习水平和写作质量,或者句法复杂度的不同维度与写作质量之间的关系会如何受到学习者学习水平的影响。这种研究结果不仅有重要的理论贡献,对写作测试也有实际贡献。

在研究与学习者因素、写作任务和教学写作环境相关的变量对二语写作句法复杂度的影响时,研究者可注重研究多种变量之间的相互作用。例如,前人研究过写作文体、写作题目和写作任务复杂度对句法复杂度的影响,然而很少有研究者对这些变量之间可能产生的相互作用和影响进行研究。以上这些问题都需

要通过对信息丰富的语料进行更加系统全面的分析才能得出有效的结论。

最后,从语言发展和测试的角度,研究者可系统分析句法复杂度与其他语言子系统(如词汇复杂度、语义复杂度和语篇连贯性等)之间的相互关系。从动态系统理论(Ellis & Larsen-Freeman 2006; Verspoor *et al.* 2011) 的角度出发,语言发展是一个动态变化的过程。在这个过程中的不同阶段,各语言子系统之间可能互相竞争,也可能共同发展。同时,二语写作质量的评估亦涉及除句法复杂度之外的众多维度。系统探究语言发展过程和二语写作评估中句法复杂度与其他语言维度之间的关系,需要使用丰富的测量指标分析大量的语言使用数据,而这种分析将会帮助我们更深入地了解句法复杂度在二语发展过程和二语写作评估中的地位和作用。

#### 参考文献

- Ai, H. & X. Lu. 2013. A corpus-based comparison of syntactic complexity in NNS and NS university students writing [A]. In A. Díaz-Negrillo, N. Ballier & P. Thompson (eds.), *Automatic Treatment and Analysis of Learner Corpus Data* [C]. Amsterdam: John Benjamins. 249-264.
- Biber, D., B. Gray & K. Poonpon. 2011. Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? [J]. *TESOL Quarterly* 45: 5-35.
- Bulté, B. & A. Housen. 2014. Conceptualizing and measuring short-term changes in L2 writing complexity [J]. *Journal of Second Language Writing* 26: 42-65.
- Chan, H., M. Verspoor & L. Vahtrick. 2015. Dynamic development in speaking versus writing in identical twins [J]. *Language Learning* 65: 298-325.
- Crossley, S. & D. McNamara. 2014. Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners [J]. *Journal of Second Language Writing* 26: 66-79.
- Ellis, N. & D. Larsen-Freeman. 2006. Language emergence: Implications for applied linguistics — Introduction to the Special Issue [J]. *Applied Linguistics* 27: 558-589.
- Ellis, R. & F. Yuan. 2004. The effects of planning on fluency, complexity, and accuracy in second language narrative writing [J]. *Studies in Second Language Acquisition* 26: 59-84.
- Graesser, C., D. McNamara, M. Louwse & Z. Cai. 2004. Coh-Metrix: Analysis of text on cohesion and language [J]. *Behavior Research Methods, Instruments, & Computers* 36: 193-202.
- Klein, D. & C. Manning. 2003. Fast exact inference with a factored model for natural language parsing [A]. In S. Becker & K. Obermayer (eds.), *Advances in Neural Information Processing Systems* [C]. Cambridge, MA.: The MIT Press. 3-10.
- Kuiken, F. & I. Vedder. 2008. Cognitive task complexity and written output in Italian and French as a foreign language [J]. *Journal of Second Language Writing* 17: 48-60.

- Levy, R. & G. Andrew. 2006. Tregex and Tsurgeon: Tools for querying and manipulating tree data structures [A]. In N. Calzolari, K. Choukri, A. Gangemi, B. Maegaard, J. Mariani, J. Odijk & D. Tapias (eds.). *Proceedings of the Fifth International Conference on Language Resources and Evaluation* [C]. Genoa: European Language Resources Association. 2231-2234.
- Lu, X. 2010. Automatic analysis of syntactic complexity in second language writing [J]. *International Journal of Corpus Linguistics* 15: 474-496.
- Lu, X. 2011. A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development [J]. *TESOL Quarterly* 45: 36-62.
- Lu, X. 2014. *Computational Methods for Corpus Annotation and Analysis* [M]. Dordrecht: Springer.
- Lu, X. & H. Ai. 2015. Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds [J]. *Journal of Second Language Writing* 29: 16-27.
- Mazgutova, D. & J. Kormos. 2015. Syntactic and lexical development in an intensive English for Academic Purposes programme [J]. *Journal of Second Language Writing* 29: 3-15.
- Norris, J. & L. Ortega. 2009. Towards an organic approach to investigating CAF in instructed LA: The case of complexity [J]. *Applied Linguistics* 30: 555-578.
- Ortega, L. 2003. Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing [J]. *Applied Linguistics* 24: 492-518.
- Ortega, L. 2012. Interlanguage complexity: A construct in search of theoretical renewal [A]. In B. Kortmann & B. Szmrecsanyi (eds.). *Linguistic Complexity: Second Language Acquisition, Indigenization, Contact* [C]. Berlin: Mouton de Gruyter. 127-155.
- Sotillo, S. 2000. Discourse functions and syntactic complexity in synchronous and asynchronous communication [J]. *Language Learning and Technology* 4: 82-119.
- Stockwell, G. & M. Harrington. 2003. The incidental development of L2 proficiency in NS-NNS email interactions [J]. *CALICO Journal* 20: 337-359.
- Taguchi, N., B. Crawford & D. Wetzel. 2013. What linguistic features are indicative of writing quality? A case of argumentative essays in a college composition program [J]. *TESOL Quarterly* 47: 420-430.
- Verspoor, M., K. de Bot & W. Lowie (eds.). 2011. *A Dynamic Approach to Second Language Development* [C]. Philadelphia: John Benjamins.
- Verspoor, M., M. Schmid & X. Xu. 2012. A dynamic usage based perspective on L2 writing [J]. *Journal of Second Language Writing* 21: 239-263.
- Vyatkina, N. 2013. Specific syntactic complexity: Developmental profiling of individuals based on an annotated learner corpus [J]. *The Modern Language Journal* 97: 11-30.
- Way, D., E. Joiner & M. Seaman. 2000. Writing in the secondary foreign language classroom: The effects of prompts and tasks on novice learners of French [J]. *The Modern Language Journal* 84: 171-184.
- Wolfe-Quintero, K., S. Inagaki & H. Kim. 1998. *Second Language Development in Writing: Measures of Fluency, Accuracy and Complexity* [M]. Honolulu: University of Hawaii, Second Language Teaching and Curriculum Center.

- Yang, W., X. Lu & S. Weigle. 2015. Different topics, different discourse: Relationships among writing topic, measures of syntactic complexity, and judgments of writing quality [J]. *Journal of Second Language Writing* 28: 53-67.
- Yoon, H. & C. Polio. forthcoming. The linguistic development of students of English as a second language in two written genres [J]. *TESOL Quarterly*.
- 鲍 贵, 2009, 英语学习者作文句法复杂性变化研究 [J], 《外语教学与研究》(4): 291-297。
- 蔡艳玲, 2012, 任务完成方式对英语专业学生写作的影响 [J], 《外语与外语教学》(2): 50-54。
- 陈慧媛, 2010, 英语写作表现测量指标的分类及特性研究 [J], 《现代外语》(1): 72-80。
- 纪小凌, 2009, 英语学习者书面语发展研究 [J], 《现代外语》(2): 178-185。
- 李雪莲, 2014, 英语专业学习者书面语文本特征多维度研究——以自然语言处理软件 Coh-Metrix 为研究工具 [J], 《南京工程学院学报》(社会科学版)(1): 41-46。
- 刘春燕, 2009, 语言产出任务对提高学习者书面中介语流利度、准确度和复杂度效果的研究 [J], 《中国海洋大学学报》(社会科学版)(6): 53-57。
- 邵银波、孙春园、桂 林、郭泉江, 2012, 英语专业八年级作文质量水平与句法复杂性的相关研究 [J], 《考试周刊》(9): 66-67。
- 王静萍, 2013, 资源指引型的任务复杂度对二语写作语言表现的影响 [J], 《外语教学》(4): 65-68。
- 文秋芳、王立非、梁茂成, 2005, 《中国学生英语口语笔语语料库》[M]。北京: 外语教学与研究出版社。
- 徐晓燕、王维民、熊燕宇、蒋 婧、潘小燕、孙念红, 2013, 中国英语专业学生英语议论文句法复杂性研究 [J], 《外语教学与研究》(2): 264-275。
- 曾 路、伍忠杰, 2010, 多媒体式写作任务测试难度特征及对 EFL 写作的影响 [J], 《外语电化教学》(1): 48-51。
- 张新玲、周 燕, 2014, 任务类型对中国英语学习者写作表现的影响 [J], 《现代外语》(4): 548-558。
- 张艳敏、王 涛、侯 旭, 2015, 基于语料库的二语写作句法复杂性测量因子研究 [J], 《天津外国语学院学报》(3): 56-62。
- 赵俊海、陈慧媛, 2012, 英语学习者书面语语法复杂度的测量研究 [J], 《外语教学理论与实践》(1): 27-33。

收稿日期: 2015-09-23; 修改稿, 2016-03-06; 本刊修订, 2016-04-12

通讯地址: 510420 广东省广州市 广东外语外贸大学英语语言文化学院(陆)

510420 广东省广州市 广东外语外贸大学外国语言学及应用语言研究中心(许)