

语言研究开放式概要语步的句法复杂度研究*

金 檀¹ 康梦超¹ 徐曼菲¹ 梁泽堃²

(1. 华南师范大学外国语言文化学院 广东广州 510631;

2. 中山大学外国语学院 广东广州 510275)

摘要:语言研究开放式概要(以下简称“概要”)是近年来多种国际语言研究期刊逐步采用的一种新兴体裁,已成为语言研究领域重要知识传播媒介。为探究该体裁的句法特征,本研究采用语料库研究方法,对概要语步进行了人工标注,并通过计算句法复杂度对概要语步的语言特征进行了量化分析。本研究采用的句法复杂度指标包括句子长度、名词化、复杂名词短语、限定性从属小句与非限定性从属小句。结果表明,概要不同语步的句法复杂度存在显著差异,特定语步与不同类型的句法复杂句之间存在一定对应关系。本研究有助于加深对概要这一新兴体裁的认识,也可为概要写作教学提供实证依据。

关键词:语言研究开放式概要;语步;句法复杂度

中图分类号:H0 文献标识码:A 文章编号:1004-6038(2024)02-0013-13

1. 引言

语步,也称修辞语步,由一个或多个语阶组成,是“实现连贯交际功能的修辞单位”(Swales 2004: 228)。语步分析是学术英语研究的常用方法之一。过往研究多关注学术论文的语步结构或语步的语言特征(Moreno & Swales 2018),并通过考察不同语步的语言特征来揭示学术论文中语言形式与功能之间的联系,进而为学术论文写作教学提供实证依据与参考范例。以往此类特征研究主要集中在两个层面:

一是短语层面。已有研究系统探究了学术论文摘要、引言、讨论以及结语部分语步的短语特征,例如:摘要语步的多词表达(Omidian et al. 2018)、引言语步的词块(Cortes 2013)与短语框架(Lu et al. 2021b)、讨论语步的词块(李梦骁、刘永兵 2017)与词簇(Le & Harrington 2015)以及结论语步的词块(Deng & Liu 2023)。这些研究表明,学术论文的不同语步具有其独特的短语使用特点,在短语层面上,学术语言形式与功能之间存在一定的对应关系。

二是句法层面。现有相关研究主要考察了不同学科学术论文引言或摘要中语步的句法复杂度特征。例如:Lu等(2020)与Zhou等(2023)分别考察社会科学与自然科学研究论文中引言语步的句法复杂度特征,发现不同类型的复杂句与特定的语阶之间存在一定对应关系。姜蕾和康梦超(2020)则对应用语言学论文摘要语步的句法复杂度特征进行了量化分析,发现不同语步间的句法复杂度特征存在显著差异。此外,Lu等(2021a)还发现引言语步的句法

* 基金项目:本文系2022年广东省本科高校教学质量与教学改革工程建设项目(项目编号:508)的阶段性研究成果。

作者简介:金檀,教授,博士,研究方向:语料库语言学,语言智能教学,语言能力测评;康梦超,博士生,研究方向:语料库语言学,学术英语,二语写作;徐曼菲(通信作者),教授,博士,研究方向:语料库语言学,英语教育,教师教育;梁泽堃,硕士生,研究方向:二语写作,计算机辅助语言学习,语料库语言学。通信作者邮箱:xumanfei@senu.edu.cn

复杂度特征存在学科差异。上述研究结果表明,学术论文不同语步也具有独特的句法复杂度特征,特定类型的复杂句与特定的语步之间也存在一定的对应关系。

总体而言,已有的语步语言特征研究主要以学术论文为考察对象,尚未对其他学术体裁开展深入探究。学术概要是不同于论文摘要的一种新兴学术体裁,要求论文作者使用通俗易懂的语言来概括学术论文的研究目标、研究方法与研究发现,可以让非专业读者快速了解论文的主要内容,从而更好地推动科学知识传播(Kuehne & Olden 2015; Liu & Li 2024)。在语言研究领域,最具影响力的学术概要是语言研究开放式概要(open accessible summaries in language studies, OASIS,以下简称“概要”)。概要篇幅在一页左右,已被众多语言研究领域的权威国际期刊采用,成为该领域重要的知识传播媒介之一(Alferink & Marsden 2023; Marsden et al. 2018)。

虽然概要在语言研究知识传播过程中扮演重要角色,但是尚未有实证研究对其语步及语言特征进行探究。句法复杂度是衡量二语水平(Kyle et al. 2021; Lu 2017; 朱慧敏 2021)、二语写作质量(Ortega 2015; Kyle & Crossley 2018; 张晓鹏、李雯雯 2022)以及学术写作质量的重要指标(Biber et al. 2011; Nasserri 2021; Staples et al. 2016)。有鉴于此,本研究以句法复杂度为切入点,考察概要语步的语言特征,尝试回答以下两个问题:

- (1) 概要中不同语步的句法复杂度特征是否存在差异?
- (2) 概要中的句法复杂句常用于实现哪些语步?

2. 研究设计

2.1 语料收集

我们从OASIS的合作期刊中选择 *Language Learning*、*The Modern Language Journal*、*TESOL Quarterly*、*Language Teaching Research* 和 *Studies in Second Language Acquisition* 作为本研究概要的来源期刊。这五本期刊是OASIS合作期刊中提供概要累计数量最多的期刊,且均为语言学领域具有高影响力的国际期刊。在选取语料时,我们仅选取来源论文为实证研究论文的概要。选择这种类型的概要主要有两个原因:一是不同类型的论文(例如理论研究、综述、通讯短文等)具有不同的语步结构(Swales 2004),概要作为论文的总结,其语步结构则会受到来源论文类型的影响;二是实证研究是语言研究领域的主流,具体包括量化研究、质化研究以及混合研究(参见 Gao et al. 2001)。我们从OASIS官网^①收集整理了2020—2022年间发表在这五本期刊上的实证性研究论文的概要文本,构建概要语料库。最终我们收集到463篇概要文本,总计258,841个形符,概要语料库的基本统计信息详见表1。

表1 概要语料库构成

期刊名称	文本数	形符数	平均文本长度	标准差
<i>Language Learning</i>	97	49,459	509.89	74.660
<i>The Modern Language Journal</i>	96	54,436	567.04	81.915
<i>TESOL Quarterly</i>	97	55,678	574.00	72.564
<i>Language Teaching Research</i>	126	72,108	572.29	91.786
<i>Studies in Second Language Acquisition</i>	47	27,160	577.87	90.926
总计	463	258,841	559.05	85.943

2.2 语步标注

语步标注分为三个阶段:(1)框架建立阶段;(2)信度检验阶段;(3)独立标注阶段。在框架建立阶段,两位标注者参考 Marsden 等(2023)与 Mizumoto 等(2017),首先标注从语料库中随机选取的 30 篇概要文本,再基于标注结果通过讨论建立语步标注框架(含语阶)。在信度检验阶段,两位标注者再独立标注从概要语料库中随机选取的 10% 的文本。标注完成后,对标注结果进行信度检验。统计结果表明,语步与语阶标注的 Cohen's Kappa 系数分别为 0.958 和 0.936,信度结果较好。在独立标注阶段,两位标注者各自标完剩余语料。最后,两位标注者交叉检查,协商达成一致。语步标注结果见表 2,其中句子数指实现该语步的句子总数,百分比指实现该语步的句子总数占概要语料库句子总数的比例,M 代表语步,S 代表语阶。

表 2 语步标注结果

部分	语步/语阶	功能描述	句子数	百分比(%)
研究内容 与意义	M1	建立学术领域	789	7.78
	M1_S1	建立中心议题	141	1.39
	M1_S2	概括议题内容	486	4.79
	M1_S3	总结现有研究	16	0.16
	M1_S4	以往研究推论/主要研究问题	146	1.44
	M2	设置合适的研究地位	416	4.10
	M2_S1	指出研究空缺	212	2.09
	M2_S2	补充已知信息	16	0.16
	M2_S3	提出正面论证	116	1.14
	M2_S4	提出研究问题	72	0.71
	M3	呈现当前研究	1,724	17.00
	M3_S1	描述当前研究和/或研究目的	306	3.02
	M3_S2	陈述研究目的	158	1.56
	M3_S3	呈现研究问题或假设	51	0.50
	M3_S4	阐释定义	113	1.11
	M3_S5	总结研究方法	359	3.54
	M3_S6	汇报重要结果	595	5.87
M3_S7	陈述研究意义	142	1.40	
研究方法	M4	描述研究样本	420	4.14
	M5	描述研究工具	127	1.25
	M6	描述研究过程	2,178	21.48
	M7	描述数据分析过程	198	1.95

(续表)

部分	语步/语阶	功能描述	句子数	百分比(%)
研究发现	M8	准备呈现研究结果	5	0.05
	M8_S1	(再次)陈述数据收集与分析过程	4	0.04
	M8_S2	再次陈述研究问题或假设	1	0.01
	M9	汇报具体研究结果	2,083	20.54
	M9_S1	结果位置	5	0.05
	M9_S2	汇报重要发现	2,073	20.45
	M9_S3	提供证据支持	2	0.02
	M9_S4	指出矛盾之处	3	0.03
	M10	评论解释结果	107	1.06
	M11	概括研究结果	60	0.59
	需要考虑的事项	M12	总结研究启示	1,343
M12_S1		提出未来研究建议	257	2.53
M12_S2		指出研究/教学启示	1,086	10.71
M13		总结当前研究	423	4.17
M14		评价当前研究	266	2.62
M14_S1		指出意义/优势	123	1.21
M14_S2		指出研究局限	72	0.71
M14_S3	评价研究方法	71	0.70	

2.3 句法复杂度指标选取与统计

本研究选用了五个句法复杂度指标:句子长度、名词化、复杂名词短语、限定性从属小句及非限定性从属小句(详见表3),分别代表句子整体复杂度、短语复杂度、限定性从属关系及小句复杂度(clausal elaboration)(Norris & Ortega 2009)。

为统计这五个量化指标,我们首先使用斯坦福句法分析器(Klein & Manning 2003)对概要语料库中每个句子进行词性标注与句法拆分。然后,我们使用L2SCA(L2 Syntactic Complexity Analyzer; Lu 2010)中的Python代码获取句长、复杂名词短语、限定性从属小句及非限定性从属小句等四个句法特征指标(代码准确度F值分别为1.000、0.830、0.925、0.926)。最后,我们使用正则表达式筛选出语料库中以-tion、-sion、-ity、-ment或-ness为后缀的名词,再人工剔除非名词化的名词(如section),然后利用Python代码统计每个句子中名词化单词的数量。

表3 句法复杂度指标(改编自Lu(2010)与Lu等(2020))

类别	指标	描述
整体复杂度	句子长度	构成句子的单词数量
短语复杂度	名词化	句中含有-tion、-sion、-ity、-ment或-ness后缀的名词数量
	复杂名词短语	句中复杂名词短语的数量,具体包括:(1)名词与形容词、所有格、介词短语、关系小句、分词及同位语构成的短语;(2)名词性小句;(3)主语位置的动名词和不定式
限定性从属关系	限定性从属小句	句中限定性从属小句的数量,具体包括名词性小句、形容词性小句及状语小句
小句复杂度	非限定性从属小句	句中由动名词、不定式或过去分词引导的非限定性从属小句数量

2.4 统计分析

本研究采用Kruskal-Wallis H检验以确定不同语步/语阶之间的句法复杂度是否存在显著差异,采用Mann-Whitney U检验比较每个语步/语阶的句法复杂度与语料整体的句法复杂度,从而确定在每个句法复杂指标上与概要整体之间存在显著性差异的语步/语阶。

为了考察概要中的句法复杂句常用于实现哪些语步,我们需先确立句法复杂句的判定阈值。参照Lu等(2020),我们将每个指标的第三四分位数设为阈值,达到阈值的句子即被判定为句法复杂句(结果见表4)。随后,针对每个指标,我们以句法复杂句的数量为因变量,每个语步/语阶的句子总数为自变量,开展简单线性回归分析并生成拟合线图。在每个句法复杂度指标的拟合线图中,95%置信区间内的数据点能很好地适应回归线,符合模型预期。由此可知,这些数据点所代表的语步/语阶在该句法特征的复杂句使用上与其他语步/语阶之间没有显著区别。95%置信区间外的数据点则超出了模型预期,这表明这些数据点代表的语步/语阶在该句法特征的复杂句使用上显著高于或低于其他语步/语阶。

表4 句法复杂句数量

指标	平均值	标准差	第三四分位数	达到阈值的句子数
句子长度	23.23	10.406	29	2,620
名词化	0.93	1.101	1	5,610
复杂名词短语	3.01	1.867	4	3,395
限定性从属小句	0.74	0.926	1	5,044
非限定性从属小句	0.75	0.934	1	5,083

3. 结果

3.1 概要语步/语阶的句法复杂度差异

Kruskal-Wallis H检验结果显示:概要语步/语阶在五个句法复杂度指标上均存在显著性差异($p < 0.01$)(见表5)。

表5 不同语步/语阶之间句法复杂度差异

指标	χ^2	自由度	显著性	ε^2
句子长度	592.189	32	0.00	0.055
名词化	441.474	32	0.00	0.041
复杂名词短语	426.735	32	0.00	0.039
限定性从属小句	611.364	32	0.00	0.057
非限定性从属小句	615.230	32	0.00	0.058

Mann-Whitney U检验结果显示:在五个句法复杂度指标上,均有一些语步/语阶的均值与语料整体的均值间存在显著性差异($p < 0.05$)。

就句子长度而言,有11个语阶的均值显著高于语料整体均值($M = 23.23, SD = 10.406$),具体包括:M1_S3、M1_S4、M2_S1、M2_S3、M3_S1、M3_S2、M3_S3、M3_S5、M3_S6、M3_S7以及M12_S2;有7个语步/语阶的均值显著低于语料整体均值,具体包括M2_S4、M4、M5、M6、M9_S1、M9_S2以及M14_S3。

就名词化而言,有6个语阶的均值显著高于语料整体均值($M = 0.93, SD = 1.101$),具体包括M3_S1、M3_S6、M3_S7、M8_S1、M12_S以及M14_S1;有4个语步/语阶的均值显著低于语料整体均值,具体包括M2_S4、M4、M5以及M6。

就复杂名词短语而言,有11个语步/语阶的均值显著高于语料整体均值($M = 3.01, SD = 1.867$),具体包括M1_S1、M1_S3、M1_S4、M2_S1、M3_S1、M3_S2、M3_S6、M3_S7、M12_S1、M12_S2以及M13;有5个语步/语阶的均值显著低于语料整体均值,具体包括M2_S4、M4、M5、M6以及M9_S1。

就限定性从属小句而言,有12个语步/语阶的均值显著高于语料整体均值($M = 0.74, SD = 0.926$),具体包括M1_S4、M2_S1、M2_S3、M3_S1、M3_S2、M3_S3、M3_S6、M3_S7、M9_S4、M10、M12_S2以及M13;有4个语步/语阶的均值显著低于语料整体均值,具体包括M4、M5、M6以及M9_S2。

就非限定性从属小句而言,有8个语阶的均值显著高于语料整体均值($M = 0.75, SD = 0.934$),具体包括M1_S1、M2_S3、M3_S1、M3_S2、M3_S7、M8_S1、M12_S1以及M12_S2;有8个语步/语阶的均值显著低于语料整体均值,具体包括M2_S1、M2_S4、M4、M5、M6、M7、M9_S2以及M14_S2。

3.2 句法复杂句的交际功能

我们分别对五个句法复杂度指标开展了简单线性回归分析。结果表明,每个指标回归模型拟合效果良好(见表6)。

表6 线性回归结果

指标	残差标准差	R^2	R^2 (调整)	F	回归模型显著性
句子长度	30.135	0.942	0.940	504.57	0.000
名词化	43.4139	0.977	0.977	1338.35	0.000

(续表)

指标	残差标准差	R^2	R^2 (调整)	F	回归模型显著性
复杂名词短语	37.1321	0.944	0.942	525.44	0.000
限定性从属小句	49.4747	0.956	0.954	665.92	0.000
非限定性从属小句	51.0013	0.959	0.958	733.29	0.000

随后,我们分析了每个指标的拟合线图,分析结果见表7。如表7所示,不同语步倾向于使用不同类型的句法复杂句。

表7 句法复杂句所实现的语步/语阶

指标	句法复杂句数量显著高于预期水平			句法复杂句数量显著低于预期水平		
句子长度	M2_S1	M3_S1	M3_S5	M1_S2	M2_S4	M4
	M3_S6	M12_S2		M5	M6	M12_S1
名词化	M3_S1	M3_S6	M3_S7	M4	M5	M6
	M12_S2					
复杂名词短语	M2_S1	M3_S1	M3_S5	M2_S4	M4	M5
	M3_S6	M12_S1	M12_S2	M6	M14_S2	
限定性从属小句	M1_S4	M2_S1	M3_S1	M4	M5	M6
	M3_S6	M12_S2	M13			
非限定性从属小句	M2_S3	M3_S1	M7	M4	M5	M6
	M12_S1	M12_S2	M13			

下面我们将针对每个指标,逐一进行分析。就句子长度而言,有5个语阶的句法复杂句数量显著高于预期平均水平,这表明这些语阶通常由长句实现,具体包括:M2_S1、M3_S1、M3_S5、M3_S6以及M12_S2。有6个语步/语阶的句法复杂句数量显著低于预期的平均水平,这表明这些语步/语阶则通常由短句实现,具体包括:M1_S2、M2_S4、M4、M5、M6以及M12_S1。

通常,由长句实现的语阶具有概括性。在实现这些语阶时,需要总结提炼多个方面的信息。具体来说,在实现M2_S1时,作者通常先总结评价前人研究再指出研究空缺,如例(1)。在实现M3_S1时,作者需要凝练研究目的、情境、样本或研究意义等信息,如例(2)。在实现M3_S5时,作者需要总结参与者、工具及步骤等信息,如例(3)。在实现M3_S6时,作者需要提炼重要的研究发现,如例(4)。在实现M12_S2时,作者通常会先回顾研究发现,再呈现启示,如例(5)。

- (1) While some have approached this linguistic diversity as a problem to be solved, more recent research has focused on how multi-lingualism among students can be a valuable resource for teaching and learning languages. (M2_S1)

- (2) The study explored the role of emerging inhibitory control skills in language acquisition in typically developing toddlers by administering five inhibitory control tasks, a language test, and a parent-report language questionnaire. (M3_S1)
- (3) To test this, adult native Hebrew speakers, with English as their second-language, learned words in German either with Hebrew (native language) translations, or with English translations (a second-language, that is more similar to German). (M3_S5)
- (4) The researchers found that learners benefited from recasts to a greater extent when they performed simple in comparison to complex tasks and that aptitude was linked to learners' L2 gains only under complex task conditions. (M3_S6)
- (5) However, given that bilingual learners displayed immediate learning of novel words introduced on Day 2, our results indicate that learning novel words is faster when one has existing experience of word learning. (M12_S2)

就名词化而言,有4个语阶的句法复杂句数量显著高于预期的平均水平,具体包括:M3_S1、M3_S6、M3_S7及M12_S2。可见,概要作者倾向于使用名词化词汇较多的句子实现这些语阶。M4、M5与M6的句法复杂句数量显著低于预期的平均水平。这表明这三个语步一般由名词化词汇数量较少的句子来实现。

在分析了该指标上的复杂句索引行后,我们发现在实现M3_S1、M3_S6与M12_S2时,作者使用的名词化词汇多与研究主题相关,如例(6);在实现M3_S7时,作者使用的名词化词汇不仅包括与研究主题相关的词汇,还包括诸如 *usefulness*、*implications* 等阐述价值取向的词汇,如例(7)。

- (6) **Aviation** trainees should also develop **accommodation** skills to gauge and adjust to others' communicative repertoires. (M12_S2)
- (7) Results also revealed the **usefulness** of predicate-related **combinations** in assessing L2 Chinese phraseological diversity, sophistication, and complexity. (M3_S7)

就复杂名词短语而言,有7个语步/语阶的句法复杂句数量显著高于预期的平均水平,具体包括:M2_S1、M3_S1、M3_S5、M3_S6、M12_S1、M12_S2与M13。这表明在实现这些语步/语阶时,作者会频繁地使用复杂名词短语。有5个语步/语阶的句法复杂句数量显著低于预期的平均水平,具体包括:M2_S4、M4、M5及M6与M14_S2。这表明在实现这些语步/语阶时,作者较少使用复杂名词短语。

在由包含更多复杂名词短语的句子实现的7个语步/语阶中,有5个语阶(M2_S1、M3_S1、M3_S5、M3_S4、M12_S2)同时与长句相关联。总体而言,在实现这些语阶时,作者需要凝练总结多方面的信息。使用更多的复杂名词短语有助于凝练大量信息(Biber & Gray 2010)。因此实现这些语阶的句子一般句长较长,包含更多的复杂名词短语,信息密度较高。然而,实现M12_S1与M13的句子虽然句长较短但却包含了较多复杂名词短语。文本分析显示,这两个语步/语阶虽无需整合多方面的信息,但却要呈现细节信息。具体而言,在实现M12_S1时,作者常使用复杂名词短语描述未来研究的具体信息,如例(8);在实现M13时,作者常使用复杂名词短语提炼当前研究的重要发现,如例(9)。

- (8) Future research could gauge **learning outcomes in the target language in multilingual ESL classrooms through pre- and posttests or other assessment measures.** (M12_S1)
- (9) In this study, the students benefited more from **the form of individual writing** when it comes to **their motivation for English learning.** (M13)

就限定性从属小句而言,有6个语步/语阶的句法复杂句数量显著高于预期的平均水平,具体包括:M1_S4、M2_S1、M3_S1、M3_S6、M12_S2与M13。这表明实现这些语步/语阶的句子通常包含较多的限定性从属小句。M4、M5与M6的句法复杂句数量显著低于预期的平均水平。这说明在实现这三个语步时,作者较少使用限定性从属小句。

文本分析显示,在实现M2_S1时,限定性从属小句具有两个功能:引出研究问题(例10)和回顾前人研究(例11)。在实现M3_S1时,限定性从属小句则主要被用于引出研究问题。在实现M1_S4、M3_S6、M12_S2与M13时,作者会倾向使用“动词+that”句式的宾语从句。具体而言:在实现M1_S4时,作者多使用 Previous research (has) suggested/showed/indicated that;在实现M3_S6时,作者多使用 We/The researchers found that 与 The findings/analysis/results suggested/indicated/showed/revealed that;在实现M12_S2与M13时,作者多使用 The findings suggested/showed that。

- (10) However, it is still not known whether the same pattern of language loss is attested in conversational data for heritage speakers. (M2_S1,功能1:引出研究问题)
- (11) Although research on L1-L2 relations dates back as early as the late 1970s, the impact of this line of research on language learning research so far has been rather minimal. (M2_S1,功能2:回顾前人研究)

就非限定性从属小句而言,有6个语步/语阶的句法复杂句数量显著高于预期的平均水平,具体包括:M2_S3、M3_S1、M7、M12_S1、M12_S2以及M13。由此可见,在实现这些语步/语阶时,作者倾向使用包含较多非限定性从属小句的句子。M4、M5与M6的句法复杂句数量显著低于预期的平均水平。这表明在实现这三个语步时,作者较少使用非限定性从属小句。

文本分析显示,由包含较多的非限定性从属小句的句子实现的6个语步/语阶中,非限定性从属小句常与介词搭配使用,作为方式、条件或目的状语,如例(12),或作为名词修饰语,此种情况常使用“名词+to do”结构,如例(13)。

- (12) We did this because previously researchers have documented college-level proficiency **without investigating** the students' backgrounds. (M2_S3)
- (13) The overall findings suggest that the **ability to** select the right piece of information and simultaneously manage misleading or confounding stimuli, is particularly relevant in early language acquisition. (M12_S2)

最后,我们发现在实现M4、M5与M6的句子中,5种类型的复杂句比例均显著低于预期平均水平。可见,实现这三个语步的句子的句法结构总体较为简单。原因有可能是概要写作指引中要求作者使用项目符号分点来呈现研究方法(Marsden et al. 2023)。因此,作者通常使用多个句子描述研究方法。由于每个句子只需呈现研究方法中的某一点信息,因而这些

句子的结构相对简单。值得一提的是,在分析了实现这些语步的句子后,我们发现研究样本和工具很少被独立呈现(M4占4.14%,M5占1.25%)。通常情况下,概要作者会在陈述研究过程时顺带提及样本或工具,如例(14)。可见,在概要中M4与M5通常会被嵌入M6。换言之,这三个语步出现了嵌入现象。

- (14) Small groups of bilingual Ojibwe youth (ages 9-12) and first speaker Elders were invited to go on short (i.e., approximately 15-minutes) walks in the woods together.
(M6)

4. 讨论

本研究首先考察了概要语步间的句法复杂度差异。研究结果表明,概要语步间在五个句法复杂度指标上均存在显著性差异。该研究结果与过往研究发现一致,说明实现不同的交际功能的句子所包含的复杂句法结构不同(Lu et al. 2020; Zhou et al. 2023; 姜蕾、康梦超 2020)。本研究也支持前人的研究发现,即在不同指标上,句法复杂度显著高于整体均值的语步/语阶与显著低于整体均值的语步/语阶之间基本没有交集(Lu et al. 2020; Zhou et al. 2023)。在本研究中,M2_S1是唯一的例外。该语阶在复杂名词短语指标上均值显著高于整体均值,而在非限定性从属小句指标上均值显著低于整体均值。

此外,结合本研究和以往研究的发现,我们认为在不同学术体裁中,实现相似交际功能的句子可展现出相似的句法复杂度特征。例如,本研究发现概要中用于“描述当前研究和/或研究目的”(M3_S1)和“陈述当前研究意义”(M3_S7)的句子,在五个指标上的均值均显著高于语料总体均值。这表明相较于概要整体,实现这两个语阶的句子句法结构相对复杂。Zhou等(2023)也有类似发现。他们的研究表明,在自然科学研究论文的引言中,相较于引言整体,“陈述研究目的或主要论点”的句子,其句法结构也更加复杂。本研究还发现,概要中描述研究方法的语步(M4、M5和M6)在五个指标上的均值均显著低于语料总体均值。由此可见,相较于概要整体,实现这三个语步的句子句法结构相对简单。这也与姜蕾和康梦超(2020)针对应用语言学国际期刊论文摘要语步的研究发现一致。他们也发现相对于摘要整体,实现方法语步的句子通常句长较短、句法结构较为简单。

在此基础上,本研究还进一步探究了概要中不同类型的句法复杂句常用于实现哪些语步。结果表明,不同类型的句法复杂句与特定的语步/语阶之间存在一定对应关系。这一结果也与前人研究发现一致(Lu et al. 2020; Zhou et al. 2023),再次证明了学术语篇中句法结构与交际功能之间存在一定对应关系。结合本研究和以往研究的成果,我们发现在不同学术体裁中,句法复杂句所实现的交际功能既有共性也有差异。就共性而言,Lu等(2020)和Zhou等(2023)发现在论文引言中,通常由长句来“陈述研究目的或主要论点”。本研究发现,在概要中长句常被用于“描述当前研究和/或研究目的”(M3_S1)。就差异而言,本研究与Lu等(2020)均发现,长句常被用于“指出研究空缺”,然而Zhou等(2023)却未有此发现。这种差异可能是由学科差异所引起的(Lu et al. 2021a),因为本研究与Lu等(2020)研究的是社会科学文本,而Zhou等(2023)则关注的是自然科学文本。据此,我们认为学科差异可能是影响句法复杂句与交际功能对应关系的重要因素之一。

本研究可为概要写作教学提供实证依据与参考素材。一方面,本研究可为教学提供概要的语步框架,便于教师为学生讲解该体裁的语篇结构;另一方面,本研究揭示了概要语步/语阶的典型句法特征,可为教学提供实证依据。此外,本研究的发现也对学术英语写作与测评具有一定启示。尽管使用句法复杂句是高水平二语和学术写作的表现,但是语言复杂度越高并不等同于写作质量越高(Deng et al. 2021; 雷蕾 2017)。高质量的写作意味着作者不仅能够使用符合特定学科的词汇与句法结构(姜峰 2021),而且能够在特定的体裁中使用恰当的语言结构表达特定的交际功能(Lu et al. 2020)。由此可见,学术英语写作教学不应过于强调使用句法复杂句,而应引导学生关注句法复杂句的功能适切性。如果忽视语言形式与交际功能的关系,则容易产出与交际功能不相符的句法复杂句。学术写作评估也应考虑语言形式与交际功能的对应关系。虽然频繁使用复杂名词短语结构是现代学术文本的典型语言结构特征之一(Biber & Gray 2010),但只有在实现特定交际功能时才会这样使用。因此,学术写作质量评价也应关注句法复杂句的功能适切性。

5. 结语

本研究基于自建概要语料库,考察了概要语步的五个句法复杂度特征。研究结果显示,不同概要语步的句法复杂度存在显著差异,不同类型的句法复杂句常用于实现特定的语步/语阶。本研究为概要这一新兴学术体裁构建了语步框架,并揭示了其语步的句法特征。本研究发现加深了我们对该体裁语篇结构与语言特征的认识,并可为该体裁的写作教学提供实证依据与参考素材。然而,本研究也存在着一定的局限性。首先,本研究仅关注了语言研究领域的学术概要。鉴于学术写作存在学科差异(Lu et al. 2021a),未来研究可以考察其他学科的学术概要。其次,本研究仅采用了代表句法复杂度的五个测量指标,后续研究可考虑引入更多微观指标。最后,未来研究也可从其他视角考察概要的句法特征,如局部语法(参见刘国兵等 2021)等。

注释:

①参见 <https://oasis-database.org/?locale=en>。

参考文献:

- [1] Alferink, I. & E. Marsden. 2023. OASIS: One resource to widen the reach of research in language studies[J]. *Innovation in Language Learning and Teaching*, (5): 946-952.
- [2] Biber, D. & B. Gray. 2010. Challenging stereotypes about academic writing: Complexity, elaboration, explicitness[J]. *Journal of English for Academic Purposes*, (1): 2-20.
- [3] Biber, D., B. Gray & K. Poonpon. 2011. Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? [J]. *TESOL Quarterly*, (1): 5-35.
- [4] Cortes, V. 2013. The purpose of this study is to: Connecting lexical bundles and moves in research article introductions[J]. *Journal of English for Academic Purposes*, (1): 33-43.
- [5] Deng, L. & J. Liu. 2023. Move-bundle connection in conclusion sections of research articles across disciplines[J]. *Applied Linguistics*, (3): 527-554.

- [6] Deng, Y., L. Lei & D. Liu. 2021. Calling for more consistency, refinement, and critical consideration in the use of syntactic complexity measures for writing[J]. *Applied Linguistics*, (5): 1021-1028.
- [7] Gao, Y., L. Li & J. Lv. 2001. Trends in research methods in applied linguistics: China and the West[J]. *English for Specific Purposes*, (1): 1-14.
- [8] Klein, D. & C. Manning. 2003. Fast exact inference with a factored model for natural language parsing[A]. In S. Becker, S. Thrun & K. Obermayer (eds.). *Advances in Neural Information Processing Systems*[C]. Cambridge: The MIT Press.
- [9] Kuehne, L. & J. Olden. 2015. Lay summaries needed to enhance science communication[J]. *Proceedings of the National Academy of Sciences*, (12): 3585-3586.
- [10] Kyle, K. & S. Crossley. 2018. Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices[J]. *The Modern Language Journal*, (2): 333-349.
- [11] Kyle, K., S. Crossley & M. Verspoor. 2021. Measuring longitudinal writing development using indices of syntactic complexity and sophistication[J]. *Studies in Second Language Acquisition*, (4): 781-812.
- [12] Le, T. & M. Harrington. 2015. Phraseology used to comment on results in the discussion section of applied linguistics quantitative research articles[J]. *English for Specific Purposes*, (39): 45-61.
- [13] Liu, Y. & T. Li. 2024. Comparing the syntactic complexity of plain language summaries and abstracts: A case study of marine science academic writing[J]. *Journal of English for Academic Purposes*, (68): 101350.
- [14] Lu, X. 2010. Automatic analysis of syntactic complexity in second language writing[J]. *International Journal of Corpus Linguistics*, (4): 474-496.
- [15] Lu, X. 2017. Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment[J]. *Language Testing*, (4): 493-511.
- [16] Lu, X., J. Casal & Y. Liu. 2020. The rhetorical functions of syntactically complex sentences in social science research article introductions[J]. *Journal of English for Academic Purposes*, (44): 100832.
- [17] Lu, X., J. Casal, Y. Liu, O. Kisselev & J. Yoon. 2021a. The relationship between syntactic complexity and rhetorical move-steps in research article introductions: Variation among four social science and engineering disciplines[J]. *Journal of English for Academic Purposes*, (52): 101006.
- [18] Lu, X., J. Yoon & O. Kisselev. 2021b. Matching phrase-frames to rhetorical moves in social science research article introductions[J]. *English for Specific Purposes*, (61): 63-83.
- [19] Marsden, E., I. Alferink, S. Andringa, C. Bolibaugh, L. Collins, C. Jackson, R. Kasprovicz, D. O'Reilly & L. Plonsky. 2018. Open accessible summaries in language studies (OASIS)[OL]. <https://www.oasis-database.org>, accessed 08/20/2023.
- [20] Marsden, E., S. Andringa, I. Alferink, C. Bolibaugh, L. Collins, C. Jackson, R. Kasprovicz, D. O'Reilly, & L. Plonsky. 2023. Guidelines for writing an open accessible summary (2nd Edition)[OL]. https://iris-database.org/assets/OASIS-DOI/3s1t-cp13/Guidelines_for_writing_OASIS_summaries_Ed2_jul2023.pdf, accessed 09/06/2023.
- [21] Mizumoto, A., S. Hamatani & Y. Imao. 2017. Applying the bundle-move connection approach to the development of an online writing support tool for research articles[J]. *Language Learning*, (4): 885-921.
- [22] Moreno, A. & J. Swales. 2018. Strengthening move analysis methodology towards bridging the function-form gap[J]. *English for Specific Purposes*, (50): 40-63.
- [23] Nasser, M. 2021. Is postgraduate English academic writing more clausal or phrasal? Syntactic complexifica-



- tion at the crossroads of genre, proficiency, and statistical modelling[J]. *Journal of English for Academic Purposes*, (49): 100940.
- [24] Norris, J. & L. Ortega. 2009. Towards an organic approach to investigating CAF in instructed SLA: The case of complexity[J]. *Applied Linguistics*, (4): 555-578.
- [25] Omidian, T., H. Shahriari & A. Siyanova-Chanturia. 2018. A cross-disciplinary investigation of multi-word expressions in the moves of research article abstracts[J]. *Journal of English for Academic Purposes*, (36): 1-14.
- [26] Ortega, L. 2015. Syntactic complexity in L2 writing: Progress and expansion[J]. *Journal of Second Language Writing*, (29): 82-94.
- [27] Staples, S., J. Egbert, D. Biber & B. Gray. 2016. Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline and genre[J]. *Written Communication*, (2): 1-35.
- [28] Swales, J. 2004. *Research Genres: Explorations and Applications*[M]. Cambridge: Cambridge University Press.
- [29] Zhou, W., Z. Li & X. Lu. 2023. Syntactic complexity features of science research article introductions: Rhetorical-functional and disciplinary variation perspectives[J]. *Journal of English for Academic Purposes*, (61): 101212.
- [30] 李梦骁 刘永兵. 2017. 中外学者期刊论文评论结果语步词块比较研究[J]. *外语与外语教学*, (5): 73-80.
- [31] 刘国兵 王凡瑜 郑冰寒. 2021. 中西方学者学术论文中评价局部语法型式使用特征研究[J]. *外语与外语教学*, (2): 62-70.
- [32] 姜峰. 2021. 中国学生学术话语能力发展与教学有效性研究[J]. *外语与外语教学*, (6): 34-44.
- [33] 姜蕾 康梦超. 2020. 国际期刊论文摘要语步的句法复杂度——以应用语言学为例[J]. *沈阳大学学报(社会科学版)*, (5): 569-574+585.
- [34] 雷蕾. 2017. 中国英语学习者学术写作句法复杂度研究[J]. *解放军外国语学院学报*, (5): 1-10+159.
- [35] 张晓鹏 李雯雯. 2022. 句法复杂度对中国大学生英语说明文写作质量的预测效应[J]. *现代外语*, (3): 331-343.
- [36] 朱慧敏. 2021. 二语书面语句法发展及其动态增长模型建构研究[J]. *外语与外语教学*, (2): 103-114.