



汉语二语学习者学术汉语写作能力研究

亓海峰 丁安琪 张艳莉

摘要:通过对汉语国际教育专业56名本科留学生、70名硕士留学生学术汉语写作的词汇复杂度、语法复杂度和准确度进行多维量化分析及定性研究,发现:在与母语者对比分析的基础上,二语者与母语者语法复杂度差别不显著,在学术语体中都倾向于使用复杂的长句和T单位,词汇复杂度、句法准确度有显著差异;随着语言水平的提升,二语者学术汉语写作的词汇多样性有所发展,词汇密度和虚词运用发展不明显。

关键词:汉语二语学习者;学术汉语写作;语法复杂度;词汇复杂度;句法准确度;学术语体

DOI: 10.13734/j.cnki.1000-5315.2022.01.016

10.3969/j.issn.1000-5315.2022.01.016

收稿日期:2021-09-12

基金项目:本文系国家社会科学基金社科学术社团主题学术活动项目“疫后世界汉语教学学会创新助推国际中文教育发展策略的研究”(20STA014)的阶段性成果。

作者简介:亓海峰,女,山东泰安人,文学博士,上海外国语大学国际文化交流学院副教授,研究方向为二语发展、社会语言学,E-mail: zoeqihairfeng@126.com;

丁安琪,女,山东聊城人,文学博士,华东师范大学国际汉语文化学院教授;

张艳莉,女,山东曲阜人,文学博士,上海外国语大学国际文化交流学院教授。

一 引言

目前,接受高等学历教育的来华留学生占比逐步上升,发展与专业学习相关的学术汉语能力成为留学生培养中亟需重视和解决的问题,其中,学术汉语写作能力是学术汉语能力构成的基础和核心。

由于学术汉语写作能力普遍缺失,留学生在专业论文撰写中有较多困难,论文撰写质量不高成为来华留学生培养中的通病。已有不少研究指出了留学生学位论文写作中存在的问题,比如:亓华、范桂娟等都指出留学生学位论文在撰写规范、研究方法、语言表达等方面均不同程度地存在问题,其中,最为集中的问题就是不能恰当地运用学术语言进行论文写作^①。“硕士生、博士生的大量增加使得我们还要研究在不以汉语为学习内容、而是作为纯粹的学术工具的专业学习环境下,如何让外国学生有效地用汉语这个工具去学习和研究专业内容并用汉语写作学位论文。这是近些年才逐步凸显出来的问题,给我们的教学法研究提出了挑战”^②。对留学生学术汉语写作能力的培养和研究需要引起充分重视。

学术汉语写作能力是指用于学术研究和专业学习的书面语表达能力,与其他语体相比,学术写作语言更为简洁、严谨、正式,在二语写作中具有特殊性。目前对于汉语二语学习者学术汉语写作能力的研究还非常有限,二语学习者在学术汉语表达中词汇、语法、篇章等方面具有哪些特征都需要进行深入探讨。

基于此,本文在二语写作的CAF(复杂度、准确度、流利度)分析框架下,采用词汇、语法的复杂度、准确度以及虚词的运用特征作为分析指标,以汉语国际教育专业二语学习者为考察对象,对其学术汉语写作能力

① 亓华《留学生毕业论文的写作特点与规范化指导》,《云南师范大学学报(对外汉语教学与研究版)》2006年第1期,第6—11页;范桂娟《汉语国际教育硕士外国留学生的培养问题探讨》,《重庆师范大学学报(社会科学版)》2019年第6期,第62—68页。

② 吴勇毅《汉语作为第二语言/外语教学法研究四十年之拾穗》,《国际汉语教育》2018年第4期,第48—49页。

进行定性和定量的分析研究。

二 文献回顾

(一) 学术汉语的提出

英语二语教学中跟专业学习密切相关的专门用途英语(English for Specific Purposes, ESP)称为“学术英语”(English for Academic Purpose, EAP),学术英语以专业学习需求为驱动,用于专业学习和研究,学术英语教学为学习者用英语从事自己的专业学习和学术活动提供语言支撑^①。“学术汉语”借鉴自“学术英语”,其内涵可界定为“从事专业学习和学术研究所使用的汉语”^②。学术汉语能力是与学科专业领域内的同行进行口语交流和书面交流、进行专业阅读和写作的能力,在特定专业领域内需要掌握的语言能力即专门学术汉语能力。

学术汉语属于专门用途汉语,学界很早就开始关注专门用途汉语的教学实践研究。张道一提出加强特殊目的汉语教学并呼吁应重视培养预科留学生开展专业学习需要打下的语言基础,汉语教学领域在80年代也开展了科技汉语、商务汉语等专门用途汉语教材的编撰和研究,但当时未明确提出学术汉语的概念^③。单韵鸣以科技汉语教材为例探讨了专门用途汉语教材的编写,该研究首次提出学术汉语的概念并探讨了其与通用汉语的区别^④。李泉扩大了专门用途汉语的内涵,将专门用途汉语分为专业汉语和业务汉语^⑤,但尚未明确指出学术汉语在专门用途汉语中的地位。高增霞、刘福英提出学术汉语教学应以“训练学生运用汉语从事专业学习和学术活动”为目标,该研究明确了学术汉语的内涵和外延,将学术汉语分为特殊学术用途汉语和通用学术用途汉语^⑥。张赫等进一步厘清了学术汉语的概念,学术汉语“可区分为通用学术汉语与专门学术汉语。熟练运用通用学术汉语是来华攻读学位的留学生必须掌握的语言能力之一,而对专门学术汉语运用的熟练程度则是区别某专业领域内新手与专家的重要依据之一”^⑦。上述研究关于学术汉语内涵和外延的定位从源头上推动了学术汉语相关的教学实践和习得研究。

(二) 二语学术写作的研究

英语作为第二语言/外语教学的研究在国内外均有较多二语学术写作的研究成果,可分为教学实践和文本语言特征分析两个分支领域。教学实践研究包括EAP的教学策略、教学过程、学习者的写作过程、写作信念等。关于英语二语学术写作文本特征的研究多基于语料库对学习者的词汇、短语或语块、句法、语步等多种特征开展实证研究。比如:杜季兵在自建语料库的基础上研究发现国内英语学习者学术语体的平均词长和句长都大于一般文本,词块结构和功能类型比较丰富^⑧;徐昉对中国英语专业本科、硕士研究生和博士研究生论文中的学术词块与国际学者常用词块进行对比,发现其大量使用的学术词块与国际学者常用词块之间存在明显差异^⑨。上述研究所分析的文本多为二语者的学术论文,针对论文的题目、摘要、引言、结论等不同部分进行语言特征多维度的量化分析,有助于认识二语者学术英语写作的特点及影响因素,对二语者学术汉语写作的研究有较多启发。

相对于学术英语写作,关于学术汉语写作的研究,特别是针对汉语二语者开展的实证研究则较为薄弱。20世纪80年代汉语学界就开始了关于专门用途汉语语言特征的分析:杜厚文从语体的角度针对专门用途汉语进行了定性研究,发现科技汉语具有大量使用专业术语、抽象词语、长定语和部分文言词语的特点,从词

①蔡基刚《学术英语再认识:学术英语或通用英语?》,《西安外国语大学学报》2019年第1期,第7—11页。

②汲传波《韩国学生汉语学术论文中文言结构使用初探》,《汉语学习》2016年第6期,第78页。

③张道一《建设对外汉语教学这个新兴学科》,第一届国际汉语教学讨论会组织委员会编《第一届国际汉语教学讨论会论文集》,北京语言学院出版社1986年版,第522—528页。

④单韵鸣《专门用途汉语教材的编写问题——以〈科技汉语阅读教程〉系列教材为例》,《暨南大学华文学院学报》2008年第2期,第31—37页。

⑤李泉《论专门用途汉语教学》,《语言文字应用》2011年第3期,第110—117页。

⑥高增霞、刘福英《论学术汉语在对外汉语教学中的重要性》,《云南师范大学学报(对外汉语教学与研究版)》2016年第2期,第44—45页。

⑦张赫、李加堃、申盛夏《学术汉语的词汇使用特征研究》,《语言教学与研究》2020年第6期,第20页。

⑧杜季兵《基于语料库的中国学术英语词块结构和功能特征研究》,《外语电化教学》2016年第5期,第9—13页。

⑨徐昉《中国学生英语学术写作中身份语块的语料库研究》,《外语研究》2011年第3期,第57—63页。

语和句式的特征上考察了其特点^①;汲传波通过自建学术论文语料库考察了韩国留学生本科论文文言结构的使用特征,发现韩国留学生语体能力不足、存在少用汉语文言结构的倾向,从语体特征的角度对学术汉语写作进行了实证研究^②;李海燕等从本科留学生学术汉语写作课的需求和课程建设角度对学术汉语写作进行探讨,并基于教学实践呼吁重视留学生学术汉语写作^③;张赫等在自建学术汉语论文语料库的基础上比较了6类不同同学科学术语体的词汇特征,对学术汉语词汇特征进行了定性和定量相结合的分析并提取了高频学术汉语虚词词表^④,但未对二语者学术汉语语言特征进行考察。

上述研究对学术汉语写作研究的开展打下了基础,均不同程度地体现出学术汉语写作在语体方面的特殊性。二语者学术汉语写作应具有与通用汉语写作不同的要求及特征,但目前学界对二语学习者学术汉语写作语言特征的探索不论是理论研究还是实证研究都非常匮乏,尚无法对二语者和母语者学术汉语写作的差异及影响因素、二语者学术汉语写作能力的发展过程等问题形成明晰的认识,因此本文在自建学术汉语论文语料库的基础上对汉语国际教育专业本科留学生和硕士留学生学术汉语写作能力进行初步考察。主要研究如下问题:(1)与母语者相比,二语者学术汉语写作中词汇、语法的复杂度、准确度有何特征?(2)与本科生相比,经过两到三年的专业训练之后,硕士留学生的学术汉语写作能力有何发展变化?(3)与母语者相比,二语者学术汉语写作中高频介词和连词的使用有何差异?

三 研究设计

(一)语料来源

学术论文摘要是对论文内容进行概述的独立的语篇,“作为学术论文内容的高度概括,具有指引读者、便于检索和查阅等多种功能”^⑤,因此我们采用学术论文摘要进行文本分析。所收集的语料分别来自二语者和母语者的学术论文并自建小型语料库,其中二语者语料共32 850字,母语者对比语料共27 655字。二语者为上海某大学汉语国际教育专业的本科生和硕士研究生,本科生汉语学习时间为4年左右,均已通过HSK6级,成绩在205分以上;硕士生汉语学习时间为6—7年,均已通过HSK6级,成绩在220分以上。二者均属于高级汉语水平学习者。受试者主要来自泰国、巴基斯坦、尼泊尔、乌兹别克斯坦、哈萨克斯坦、俄罗斯、韩国、日本。本研究收集了56位本科留学生毕业论文初稿的中文摘要和70位硕士留学生学期论文初稿的中文摘要。论文初稿基本未经指导教师修改,能较真实地体现出受试者的语言表达水平。母语者语料来自汉语国际教育领域的核心期刊《世界汉语教学》、《语言教学与研究》,采用了两家期刊2017—2020年间引用率较高的84篇论文的摘要作为母语者语料参考数据,以更清晰地呈现二语者学术汉语写作的特点。

(二)分析维度和指标

我们对语料进行了定量分析和定性分析。由于学术汉语语体方面具有特殊性,相较于其他文本,其语言更为精准、正式、词汇密度更高。我们从词汇、语法的复杂度、准确度三方面进行多维度的量化分析,并针对介词和连词的使用情况进行定性分析。

对语料准确度和复杂度的分析参照亓海峰、廖建玲采用了多项分析指标^⑥。各指标的测量方法见表1。

词汇复杂度是写作中反映词汇质量的重要维度,我们采用词汇的多样性、词汇密度、虚词词类比例三种指标综合考察词汇的复杂度。汉语二语写作的多项研究中采用多样性、密度等指标对学习者的写作中的词汇能力进行研究,比如吴继峰在研究中用上述分析指标对二语者汉语写作的词汇特征进行考察,发现上述指标能有效反映不同汉语水平学习者书面语表达词汇的复杂性^⑦。词汇多样性采用类符/形符(type-token ratio)

① 杜厚文《汉语科技文体的语言特点》,《语言教学与研究》1981年第2期,第87—101、160页;杜厚文《科技汉语中的长定语》,《语言教学与研究》1988年第3期,第47—58页。

② 汲传波《韩国学生汉语学术论文中文言结构使用初探》,《汉语学习》2016年第6期,第77—85页。

③ 李海燕、张文贤、辛平《本科留学生学术汉语写作课需求调查与课程建设——以北京大学本科留学生学术汉语写作通选课为例》,《国际汉语教育(中英文)》2020年第1期,第51—60页。

④ 张赫、李加堡、申盛夏《学术汉语的词汇使用特征研究》,《语言教学与研究》2020年第6期,第19—27页。

⑤ 刘辉《学术论文摘要的目的连贯模式研究》,《外语学刊》2013年第1期,第71页。

⑥ 亓海峰、廖建玲《基于记叙文和议论文的汉语二语写作发展研究》,《世界汉语教学》2019年第4期,第567页。

⑦ 吴继峰《英语母语者汉语写作中的词汇丰富性发展研究》,《世界汉语教学》2016年第1期,第129—142页。

o), 即不重复词数/总词数, 词汇密度采用实词量占总词汇量的比例^①。杜厚文指出科技文体中常使用介词、连词等虚词^②, 在学术汉语的词汇复杂度分析中, 我们将虚词词类比率作为分析指标之一, 各虚词类占总词数的比例构成虚词词类比率。

句法复杂度采用 T 单位长度作为主要分析指标, T 单位长度采用总词数占 T 单位总数的比例。T 单位是独立主句及它所嵌入或附带的一个或多个句子成分或从句^③。在英语二语写作研究中 T 单位是一项应用较为广泛的有效分析指标, 沈金花、鲍贵在英语二语写作的研究中发现文体对英语写作的 T 单位长度有显著影响^④, 吴继峰在汉语二语写作研究中也发现 T 单位长度对文体比较敏感, 在汉语二语写作中这一指标逐步被采用并证明了这一分析指标具有适用性^⑤。关于准确度本研究采用 T 单位正确率(正确的 T 单位/T 单位总数)作为单一指标, 由于有误的 T 单位会受到词汇及句法的共同影响, T 单位准确率可以反映出词汇和语法是否准确, 所以在准确度的分析中我们仅采用 T 单位正确率作为分析指标。

表 1 语料分析指标

维度	分析指标	测量方法
准确度	T 单位正确率	正确使用的 T 单位/T 单位总数
词汇复杂度	词汇多样性	不重复词数/总词数(TTR)
		不重复词数/总词数平方根(\sqrt{TTR})
	词汇密度	实词总数/总词数
	虚词词类比率	介词、连词、助词、代词词数/总词数
句法复杂度	T 单位长度	总词数/T 单位总数

语料的定性分析中我们采用高频介词和高频连词考察二语者学术汉语的语体特点。张赅等基于小型语料库对母语者学术汉语核心虚词进行了考察, 发现“学术汉语高频使用一小部分虚词”^⑥。对二语者学术论文的语料分析我们也着重对这类比较封闭的词类进行考察, 对篇章中具有介引、连接作用的介词及连词进行归类和分析。

(三) 数据处理

通过“语料库在线”平台^⑦对语料进行分词并标注词性, 再进行人工核对, 计算出每篇语料的总词数、词频和实词、虚词各词类的词数及不重复词数。

针对研究问题(1)、(2), 本研究对二语者和母语者学术汉语词汇、语法的各项指标进行了描述性统计, 在此基础上进行了单因素方差分析, 考察语言水平对各项指标的影响。

针对研究问题(3), 本研究对母语者、二语者文本中出现的介词、连词进行了穷尽式考察, 包括介词、连词出现的次数、使用频率和出现的规律, 并对语料中高频出现的介词和连词进行了质性分析。

四 分析与讨论

(一) 二语者学术汉语写作复杂度、准确度的分析

1. 词汇复杂度

表 2 汇报了汉语国际教育专业二语者和母语者词汇复杂度各项指标的描述性统计结果。从二语者与母

① Jean Ure, "Lexical Density and Register Differentiation," in *Applications of Linguistics: Selected Papers of the Second International Congress of Applied Linguistics, Cambridge 1969*, eds. G. E. Perren and J. L. M. Trim (Cambridge: Cambridge University Press), 1971: 443-452.

② 杜厚文《汉语科技文体的语言特点》,《语言教学与研究》1981年第2期,第94页。

③ Wenying Jiang, "Measurements of Development in L2 Written Production: The Case of L2 Chinese," *Applied Linguistics* 34, no. 1 (February 2013): 1-24.

④ 沈金花、鲍贵《语言水平和体裁对英语学习者作文句法长度的影响》,《南京工业大学学报(社会科学版)》2010年第4期,第73-76页。

⑤ 吴继峰《韩国学生不同文体写作中的语言特征对比研究》,《语言教学与研究》2019年第5期,第1-12页。

⑥ 张赅、李加璠、申盛夏《学术汉语的词汇使用特征研究》,《语言教学与研究》2020年第6期,第26页。

⑦ 语料库在线网址: <http://www.aihanyu.org/cncorpus/CpsWParser.aspx>。

语者词汇各指标的对比看,除虚词词类比率外,母语者其他各项指标的均值都高于二语者。

表2 词汇复杂度的描述性统计结果

分析指标	母语者 N=84		本科留学生 N=56		硕士留学生 N=70	
	均值 M	标准差 SD	均值 M	标准差 SD	均值 M	标准差 SD
词汇多样性(TTR)	.6333	.09337	.5864	.07340	.6550	.08980
词汇密度	.7607	.04506	.7148	.05084	.7358	.06653
虚词词类比率	.2069	.04116	.2390	.04789	.2152	.04431
介词比率	.0468	.02072	.0455	.02076	.01851	.00221
代词比率	.0192	.01890	.0313	.02431	.0268	.02138
连词比率	.0476	.02057	.0545	.02479	.0453	.02352
助词比率	.0933	.02495	.1077	.02792	.1043	.02713

母语者词汇多样性均值为63.33%,词汇密度均值为76.07%,多样性和词汇密度较高。我们认为,在学术语篇中特别是在学术论文摘要中为避免口语化倾向,需要使用大量专业词汇和规范正式的表达,二语者在词汇运用方面上述指标均明显低于母语者,较之本科生,研究生学术论文中的上述指标均值更接近母语者,显示出词汇密度和词汇多样性指标对区分不同水平二语者学术汉语写作水平具有有效性。

我们对母语者、本科留学生和硕士留学生学术汉语文本中词汇多样性和词汇密度进行了单因素方差分析,以进一步考察本科、硕士留学生学术汉语文本中各项指标的发展变化。方差分析结果显示,语言水平对词汇多样性的主效应显著 $[F(2,207)=9.862, p<.01]$ 。多重比较显示,本科阶段二语者和硕士阶段二语者词汇多样性呈显著差异 $(p<.01)$,这一结果说明随着语言水平的提高,词汇多样性有所发展。

Hyland的研究指出学术英语语篇具有实词高密度的特点(high lexical density)^①,我们对母语者文本词汇密度的考察有相同发现,方差分析的结果则显示语言水平之间在词汇密度性上有明显差异 $[F(2,207)=12.240, p<.01]$ 。多重比较显示,本科阶段二语学习者和硕士阶段二语学习者的差异 $(p=.84)$ 不显著,母语者、二语者之间的差异以及不同水平二语者在词汇密度上的差异显示出二语者学术汉语的词汇密度发展较慢。

虚词词类比率作为词汇复杂度的一项指标,与其他指标相比,呈现出相反的趋势,二语者的虚词词类比率高于母语者,其中本科阶段二语者虚词词类比率(23.90%)>硕士阶段二语者(21.52%)>母语者(20.69%)。具体到代词、连词、助词等几种主要词类,可以发现代词、连词和助词的比率在三者中都呈现出二语者使用比率高于母语者的情况,特别是本科阶段二语者学术论文摘要中几类虚词使用的比率都较高,分别为4.55%(介词)、10.77%(助词)、5.45%(连词)。这一结果可能与二语者在学术文本中语言不够简练,过多地运用助词、代词、连词等有关。

我们分别对母语者和两组受试介词比率、连词比率、助词比率、代词比率进行了单因素方差分析以比较语言水平对虚词词类使用的影响。母语者、二语者代词比率差异显著 $[F(2,207)=5.835, p=.003]$,多重比较显示,本科阶段二语学习者和硕士阶段二语学习者的差异 $(p=.467)$ 不显著;母语者、二语者助词比率差异显著 $[F(2,207)=5.848, p=.003]$,多重比较显示,本科和硕士留学生助词比率差异不显著 $(p=.763)$;母语者、二语者连词差异 $[F(2,207)=2.696, p=.070]$ 呈边缘显著,多重比较显示本科、硕士留学生连词比率无明显差异 $(p=.065)$;母语者、二语者介词差异不显著 $[F(2,207)=3.345, p=.037]$,多重比较显示,本科、硕士留学生介词比率差异不显著 $(p=.148)$ 。在虚词的使用比率上,本科阶段二语者与硕士阶段二语者之间无明显差异,反映出二语者在学术汉语写作中发展虚词的使用需要较长时间。

^①Ken Hyland, "Directives: Argument and Engagement in Academic Writing," *Applied Linguistics* 23, no. 2 (June 2002): 215-239.

2. 句法复杂度

表 3 是母语者、二语者句法复杂度的描述性统计结果。母语者语料的 T 单位长度均值为 14.7188, 本科留学生 T 单位长度均值为 13.80, 硕士留学生 T 单位长度均值为 13.89, 二语者句法复杂度略低于母语者, 同时, 随着学习阶段的提升, 学术汉语文本中 T 单位更长。方差分析结果显示, 语言水平对平均 T 单位长度主效应不显著 [$F(2, 207) = 1.459, p = .235$], 母语者、二语者在 T 单位长度上差异不显著, 多重比较显示, 本科留学生和硕士留学生的 T 单位长度差异也不显著 ($p = .990$)。上述结果反映出母语者、二语者在学术语体中句法复杂度都较高, 这一点与英语二语学术写作的研究结果一致^①。“大量使用长句是学术汉语的重要语法特征之一, 长句包括长单句与长复句。长单句的重点在于中心语有多层定语或状语修饰”^②, 句子、T 单位长度有密切关系, 本研究对 T 单位长度的分析结果也印证了母语者、二语者学术汉语文本具有大量使用复杂长句的特征。

表 3 句法复杂度的描述性统计结果

分析指标 \ 统计结果	母语者 N=84		本科留学生 N=56		硕士留学生 N=70	
	均值	标准差	均值	标准差	均值	标准差
	M	SD	M	SD	M	SD
平均 T 单位长度	14.7188	3.05794	13.8055	4.42866	13.8949	3.46622

3. 准确度

表 4 是母语者、二语者句法准确度的描述性统计结果。如前所述, 由于 T 单位正确率受到词汇正确率和句法正确率的影响, 准确度的分析采用 T 单位正确率作为单一分析指标。

相较于母语者, 二语者文本的准确度偏低, 本科留学生学术汉语文本中的正确率为 73.13%, 硕士留学生语料中的正确率为 79.29%。单因素方差分析的结果显示, 语言水平对 T 单位正确率的主效应非常显著 [$F(2, 207) = 69.694, p = .000$], 多重比较显示, 母语者与本科留学生、硕士留学生的差异均非常显著 ($p = .000$), 本科留学生和硕士留学生之间的差异呈边缘显著 ($p = .047$), 反映出二语者随着学习阶段的提升, 学术汉语的准确度有所发展。

表 4 句法准确度的描述性统计结果

分析指标 \ 统计结果	母语者 N=84		本科留学生 N=56		硕士留学生 N=70	
	均值	标准差	均值	标准差	均值	标准差
	M	SD	M	SD	M	SD
平均 T 单位正确率	1.0000	.00000	.7313	.17614	.7929	.19357

以下是本科留学生和硕士留学生论文摘要中的片段:

片段 1(本科留学生):

本文通过问卷调查和访谈, 分析泰国学生在课堂怎么学习汉语。在课堂活动中泰国学生用哪些方法来了解汉语, 和他们在学习中哪些方法对他们有效果呢, 比如说重复操练、看电视剧、背诵等等。本文的调查对象是高中的泰国学生。结果显示, 大部分同学认为录音和表演对他们课堂学习最有效果。

片段 2(硕士留学生):

在一定程度上, 汉语学习者对汉语成语的掌握程度能反映出他们的汉语水平。因为汉语成语的庞大数量和客观难度, 所以对大多数意大利汉语学习者来说, 成语很难。基于此问题, 本文将以问卷调查的方式从意大利汉语中高级学习者对象, 试图了解他们学习汉语成语的困难和动机有哪些, 并提出相应的教学对策。希望本文能对意大利的对外汉语教育有贡献。

片段 1 为一位本科留学生的论文摘要, 片段 1 平均 T 单位长度为 11.5, 总词数为 69 个, 不重复词数 45

① Viviana Solange Cortes, "Lexical Bundles in Academic Writing in History and Biology," (ph. D Thesis, Northern Arizona University, 2002), 98-223, http://www.pqdtcn.com/thesis_Details/521BE69EF61651496728D8EBEDC913B7.

② 张赓、李加堃、申盛夏《学术汉语的词汇使用特征研究》,《语言教学与研究》2020 年第 6 期, 第 22 页。

个,词汇密度为 69.56%,各项指标接近本科生总体语料的均值。片段 1 甲级词 41 个,乙级词 12 个,丙级词 8 个,丁级词及超纲词 8 个,60%左右的词语为难度较低的甲级词,是简单常用的词语,词汇的复杂度不高。片段 1 中还使用了语气词“呢”,带有明显的口语色彩。

片段 1 中 2 个 T 单位存在偏误:(1)“分析泰国学生在课堂怎么学习汉语”偏误为词语遗漏,应为“分析泰国学生在课堂中怎么学习汉语”;(2)“和他们在学习中哪些方法对他们有效果呢”偏误为成分冗余,应为“在学习哪些方法对他们有效果呢”。

片段 2 为一位硕士留学生撰写的论文摘要,仍有明显偏误,比如“本文将以问卷调查的方式从意大利汉语中高级学习者对象”,存在句式杂糅的偏误,但相较于片段 1,片段 2 中句式和词语的复杂度都有所提高,T 单位正确率也有了一定提高。

(二)介词、连词的定性研究

张赅等对母语者学术汉语高频虚词的特征进行了研究,发现连词、介词、副词等封闭性词类是学术汉语的核心词类,“人文社科类学术汉语在高频虚词使用上更具丰富性”^①。我们在对母语者、二语者词汇、句法复杂度及准确度的多个指标量化分析的基础上,选择介词和连词做进一步的定性分析。之所选择介词和连词进行考察,主要在于学术语体中语言的复杂度较高,在较长的单句和复句中带有介引和连接作用的介词、连词必不可少,这两类词类在篇章中具有语法功能更能体现学术汉语的语体色彩。

1. 介词的使用情况

母语者文本中使用频率最高的是“在”,在全部文本中累计出现频率为 25.18%,使用频率最低的是“沿”,仅出现了 1 次,使用频率为 0.15%。按照使用频率由高到低排列,前 15 位的介词分别为:“在”(25.18%)、“对”(18.85%)、“从”(9.87%)、“以”(6.63%)、“于”(6.04%)、“通过”(5.60%)、“为”(4.86%)、“基于”(4.71%)、“把”(3.09%)、“由”(1.91%)、“对于”(1.33%)、“被”(1.18%)、“根据”(1.18%)、“向”(1.18%)、“依据”(0.88%)。其中,表示方式依据的介词最多,比如“以”、“根据”、“依据”、“由”、“通过”、“基于”等介词都具有相同的语法意义,这可能与母语者在学术论文摘要中对某一现象或对象进行客观分析时,较为注重交代研究的具体方法、分析的依据有关,所以摘要的文本中会频繁使用表示方式、依据的介词。

硕士留学生文本中使用频率最高的介词是“在”,使用频率最低的介词是“自从”。按照使用频率由高到低排列,前 15 位的介词分别为:“在”(31.93%)、“为”(11.23%)、“通过”(7.37%)、“以”(5.96%)、“从”(5.61%)、“关于”(5.26%)、“随着”(3.51%)、“对于”(3.16%)、“经过”(2.81%)、“把”(2.46%)、“为了”(2.46%)、“针对”(2.46%)、“对”(2.11%)、“让”(1.75%)、“有关”(1.40%)。其中表示对象内容类的介词最多,比如“对”、“对于”、“关于”、“有关”、“针对”。

本科留学生文本中使用频率最高的介词也是“在”,使用频率最低的介词是“由”。按照使用频率由高到低排列,前 15 位的介词分别为:“在”(33.53%)、“对”(14.37%)、“为”(8.69%)、“通过”(6.89%)、“从”(6.29%)、“以”(4.79%)、“随着”(2.99%)、“让”(2.69%)、“把”(2.40%)、“被”(2.10%)、“关于”(2.10%)、“为了”(2.10%)、“给”(1.20%)、“当”(0.90%)、“除了”(0.60%)。与硕士留学生文本分析结果一致,表示对象内容类的介词使用频次也最多,比如“对”、“关于”、“给”、“向”等都表示介引对象,体现出二语者在论文摘要中更为注重对研究对象、研究内容的介绍和说明,这一点符合学术语篇的表达需求,但也体现出二语者的摘要中对具体的研究手段、分析依据的说明不够充分。

冯胜利、吴春相等都指出介词具有体现语体特征的功能^②。我们选取母语者、二语者文本中使用频率较高的 15 个介词,依据吴春相的分类将其分为口语语体介词、中性语体介词和书面语体介词三类,用以比较母语者、二语者介词使用情况的异同。比较结果见表 5。

①张赅、李加燮、申盛夏《学术汉语的词汇使用特征研究》,《语言教学与研究》2020 年第 6 期,第 23 页。

②冯胜利《书面语语法及教学的相对独立性》,《语言教学与研究》2003 年第 2 期,第 53—63 页;吴春相《现代汉语介词结构的语体考察》,《当代修辞学》2013 年第 4 期,第 52—61 页。

表 5 母语者和二语学习者介词使用情况对比

类别 语体	母语者	本科留学生	硕士留学生
书面语体	以、通过、基于、由、对于、根据、依据、于	通过、以、关于	通过、以、关于、对于、针对、有关
口语体	对、为、把、被	对、为、随着、让、把、被、为了、除了、当	对、让、随着、把、为了、为
中性语体	在、从、向	在、从、给	在、从、经过

母语者文本中使用频率高的介词大部分带有书面语体特征,有些介词带有文言色彩,比如“于”、“由”、“基于”。二语者文本中使用频率高的介词多带有口语语体特征,比如“为了”、“当”、“随着”、“让”、“给”,特别是本科留学生的文本中大部分介词都带有口语语体特征,硕士留学生文本中书面语语体特征的介词比例有所增加,但仍带有较明显的口语特征。

2. 连词的使用情况

母语者文本中使用频率最高的连词是“和”,使用频率最低的连词是“无论”。按照使用频率由高到低排列,前 15 位的连词分别为:“和”、“与”、“而”、“并”、“及”、“以及”、“则”、“因此”、“但”、“或”、“此外”、“且”、“从而”、“因”、“进而”。从其所表示的连接关系看,主要表示并列关系、因果关系、转折关系和递进关系。摘要中一般需要说明研究条件、研究方法并用高度概括的语言归纳研究结果和研究发现,表并列关系、递进关系连词的高频使用与文本中需要较为充分地交代研究条件有关,因果关系和转折关系连词的高频使用与文本中需要依据研究条件进行因果论证有关。上述高频连词的使用符合学术论文写作的特点。

表 6 母语者和二语学习者使用高频连词情况对比

类别 连接关系	母语者	硕士留学生	本科留学生
并列关系	和、与、并、及、以及、或、此外	和、与、并、以及	和、并
因果关系	因此、从而、因	因此、所以、由于、因为	所以、因为、因此、由于、从而
转折关系	而、则、但	但是、而、虽然	但是、虽然
递进关系	且、进而	并且、而且	而且、不仅
承接关系	/	然后	然后
假设关系	/	如果	/
解说关系	/	/	比如、比如说
条件关系	/	/	不管

硕士留学生文本中使用频率最高的连词是“和”,使用频率最低的是“从而”,按照使用频率由高到低排列,前 15 位的连词分别为:“和”、“与”、“但是”、“并”、“因此”、“以及”、“所以”、“由于”、“而”、“因为”、“并且”、“而且”、“然后”、“如果”、“虽然”。上述高频连词表示因果关系的最多,与母语者不同的是表示假设关系和承接关系的连词也是高频使用的连词。

本科留学生文本中使用频率最高的连词也是“和”,使用频率最低的是“然而”。按照使用频率由高到低排列,前 15 位的连词分别为:“和”、“并”、“但是”、“所以”、“比如”、“因为”、“而且”、“因此”、“由于”、“不仅”、“虽然”、“然后”、“从而”、“不管”、“比如说”。其中表示并列关系和因果关系的连词数量最多。与母语者不同,本科留学生使用高频连词表示的连接关系更为多样,表解说关系、承接关系和条件关系的连词都成为高频连词。

五 结语

(一) 研究结论

本研究在 CAF 框架下,通过自建小型语料库,对汉语国际教育专业本科留学生和硕士留学生论文摘要进行考察并参照母语者学术论文摘要进行对比分析,初步得出以下结论。

第一,二语者学术汉语表达中语言的复杂度与母语者存在明显差异,其中,词汇复杂度差异非常显著,语法复杂度差异不显著,说明不同指标之间发展的不平衡性。这一方面反映出在学术汉语的书面表达中二语者有与母语者接近的方面,在学术语体中都倾向于使用复杂的长句进行表达;另一方面反映出相较于词汇复杂度,句法复杂度的发展优先于词汇复杂度。在词汇复杂度的各项指标中,随着语言水平和学习阶段的提升,词汇多样性有所发展,硕士留学生文本中词汇多样性更接近母语者,与本科留学生有显著差异,词汇密度和虚词比率二者无明显差异,词汇密度与文本的难度有一定关系,“词汇密度高的篇章蕴含的信息语域广,人们对词汇密度高的篇章的感知比对词汇密度低的篇章的感知要难”^①。从本研究的调查结果可知,二语者文本的阅读难度较低,文体特色不明显。

第二,本研究采用 T 单位正确率作为准确度的分析指标,结果显示,二语者学术汉语书面表达中语言的准确度与母语者有显著差异,反映出在学术语体中二语者语言的使用仍有较多问题。此外,硕士留学生学术汉语文本的正确度比本科留学生高,二者的差异处于边缘显著,说明随着语言水平的提高,准确度会有所发展。本研究结果还反映出 T 单位指标对二语写作用本分析具有敏感性,与亓海峰、廖建玲的研究发现相同^②。

第三,二语者虚词使用与母语者有显著差异,具体到不同的虚词词类又不尽相同。从质性分析的结果看,母语者学术汉语中常用虚词的分布和功能与二语者有差别,此外,二语者文本中虚词使用的总体数量虽然较多,但是重复使用的虚词多、口语色彩明显、冗余使用和过度使用的虚词多。

(二) 研究启示

来华留学接受高等教育的学历生已成为来华留学生的主流,学术汉语能力是学历留学生开展专业学习的关键,本研究对二语者学术论文写作的语言特征进行研究,对二语者学术汉语运用能力培养有以下启示。

第一,学术论文写作能力是学术汉语运用能力的核心,本研究发现汉语国际教育专业的二语者在论文摘要中语言的复杂度和准确度方面都与母语者有明显差异,需要在汉语二语写作教学中训练二语者运用难度高、复杂的词汇和句式进行学术表达。同时,要加强学术阅读,注重引导二语者关注母语者学术汉语表达的常用虚词和句式。

第二,学术汉语语体具有特殊性,母语者的语言正式、凝练,具有很强的书面语体色彩,母语者、二语者高频介词和连词的对比除了体现出分布和功能的差异,语体色彩上也有较大差异。二语学习者即使到了硕士学习阶段,学术论文写作中语言仍有明显的口语色彩,在二语写作教学中应加强语体知识的教学,注重训练二语者学术写作中语体风格的培养。

第三,本研究是对二语者学术汉语写作能力的初步考察,存在若干不足。受试者都来自汉语国际教育专业,语料有一定的局限,未来的研究可考虑扩大语料的范围,对比不同专业领域二语者学术汉语的语言特征。同时,也可考虑对二语者学术汉语能力的发展进行跟踪调查,对学术汉语纵向发展可以进行更为深入的研究。此外,对虚词的研究只考察了介词、连词两个词类,在进一步的研究中可扩大到虚词的其他词类,对二语者和母语者虚词主要词类的分布及功能进行更为全面、深入的考察。

[责任编辑:唐 普]

① 吕文涛、姚双云《词汇规制与立法语言的简明性》,《语言文字应用》2018年第4期,第67页。

② 亓海峰、廖建玲《基于记叙文和议论文的汉语二语写作发展研究》,《世界汉语教学》2019年第4期,第572页。