

基于语料库的学术论文非正式语体特征对比研究

——以中美硕士研究生为例

杨 苑

上海海事大学外国语学院, 上海

收稿日期: 2024年3月9日; 录用日期: 2024年4月8日; 发布日期: 2024年4月15日

摘 要

学术语篇的产出是对研究生学术思维的高度凝练, 而英语作为当今国际学术界所通用的媒介语言, 在学术写作和提升中国研究者的国际学术影响力方面的重要性可见一斑。目前, 中国学者与西方学者在英语学术写作上的差距之一是语域意识薄弱, 主要表现为书面语中大量出现非正式性特征。因此, 探究学术语篇语体正式程度是十分必要的。本研究以中国与美国外国语言文学专业的硕士研究生为研究对象, 分别选取双方在2018~2022年间的硕士毕业论文为语料, 自建成两个语料库, 采用定量的研究方法, 参照Hyland & Jiang非正式语体特征的框架, 选取5个主要非正式性语体特征指标, 使用语料库分析软件, 对两个语料库在非正式语体特征方面的异同进行对比分析, 并尝试剖析其差异成因, 以为学术英语写作教学带来启示和借鉴意义。

关键词

非正式语体特征, 学术语篇, 中美硕士研究生, 语料库

A Corpus-Based Comparative Study of Informal Stylistic Features in Academic Thesis

—Taking Chinese and American Graduate Students as the Case

Yuan Yang

College of Foreign Languages, Shanghai Maritime University, Shanghai

Received: Mar. 9th, 2024; accepted: Apr. 8th, 2024; published: Apr. 15th, 2024

Abstract

On the one hand, the output of academic discourse is highly condensed to the academic thinking of graduate students. On the other hand, English, as a common medium language in today's international academia, plays a significant role in academic writing as well as enhancing the international academic influence of Chinese researchers. At present, one of the gaps between Chinese scholars and Western scholars in English academic writing is the weak register awareness, which is mainly manifested by the large number of informal features in written language. Therefore, there is a necessity to explore the formal degree of academic discourse style. In this study, master's theses of Chinese and American graduate students majoring in linguistics from 2018 to 2022 have been selected as the source, which were self-constructed into two corpora respectively. Quantitative research methods were adopted and the framework of informal stylistic features of Hyland & Jiang was referred to, from which five main informal stylistic characteristics were selected. By employing corpus analysis software, the similarities and differences between the two corpora in informal stylistic characteristics have been compared and analyzed, as well as the possible causes of the result being discussed, with a view to bringing enlightenment and reference significance to academic English writing teaching.

Keywords

Informal Stylistic Features, Academic Paper Writing, Chinese and American Graduate Students, Corpus

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着国际交流的频繁与深入,我国对“学术国际化”的要求也不断提高,为了让我国的科研成果顺利“走出去”,扩大国际学术影响力,科研人员有必要遵循现行的国际学术话语体系和规范。这无疑对未来科研工作者的学术英语写作能力提出了更高的要求。语体正式度作为学术写作一个至关重要的方面,一直是语言研究和语篇研究的焦点[1] [2] [3] [4],同时,把握学术语言的语体正式度,也是学生提高学术英语写作水平、确保学术语言表达得体性的关键。另外,美国作为综合实力顶尖的发达国家,英语母语使用者的数量也在全球范围内占比最多。因此,比较我国与美国硕士研究生的英文学术论文在语体正式度方面是否存在差异是十分必要的。

非正式语体特征作为衡量语体正式度的指标,也受到学者们广泛而热烈的关注。近年来,语体研究的一个重要结论是,在(多语域)写作文本中,非正式语体特征的使用正呈现增多的趋势,即写作文本逐渐从严谨正式的语体特征向个性化、大众化的语体特征发展[5] [6] [7]。为确认这一趋势在学术领域中是否存在,学者们纷纷开展了一系列历时研究。Hyland & Jiang [8]考察了四个学科的顶级期刊分别在1965年、1985年、和2015年的非正式性语言特征的使用情况,发现学术英语写作中的非正式性特征的使用在过去50年间有所增长;Praminatih *et al.* [9]调查了1992年至2016年印度尼西亚地区的英语学习者论文摘要部分中出现的非正式性特征,却发现非正式性特征的使用频率随着时间推移逐步降低;姜峰[10]运用多维分析的方法,从5个维度考察了学术语篇在1965~2015年的历时变化,也得出了与Hyland & Jiang [8]同

样的结论，并将其与专业语篇“大众化”的社会情境演变联系在一起。研究所采用的非正式性定义、考察的语言特征和使用的参数各异是现有研究结论互相矛盾的主要原因。然而，横向对比研究的数量相比历时的研究显得寥寥无几，有部分学者通过对比本族语者与外语学习者的学术语篇[11][12]，或对比国际学术语言规定与实际语言使用的差异[13]来考察非正式语体特征使用的差异，但关注某一特定学者群体，将其学术语篇进行横向对比的研究可谓凤毛麟角，还有待丰富。

鉴于此，本文聚焦于中国和美国语言学专业的硕士研究生，搜集他们 2018~2022 年间的相关英文硕士论文，加以比较，以 5 项非正式语体特征为指标，采用定量的研究方法，探究语体正式程度的异同，并探索其可能成因，以期对学术英语写作教学带来启示、为提升学生英语学术写作能力提供有价值的参照。

2. 非正式语体特征指标

由前文所述，非正式性的定义发挥着决定性的作用。非正式语体区别于正式语体，但从现有文献来看，(非)正式性是很难界定的。站在语用学的视角，正式性与“消极礼貌”和“距离行为”紧密相关，目的是为维护他人的面子，以示尊重[14]；Richards & Platt [15]考虑了社会语境的因素把非正式性特征定义为“话者在发音、用词和句子结构的选择上非常小心地使用的言语特征”；Lillis [16]、Barton [17]把书面语和口语视为对立概念，认为口语特征就是非正式性特征，包括交互性、省略、冗杂等特点。本文采用 Hyland & Jiang [8]的定义，即学术写作中的非正式性指更具个人化的语体风格，意味着与读者之间更亲密的关系，表达了一种协商的意愿，且对主体性持积极态度。另外，正式性与非正式性并非语体的两个极端，而应看作一个连续体。

近几十年来，对于非正式语体特征的指标，研究者们从多样的角度进行了不同的划分。Aijmer [18]着重研究了瑞典语学习者书面语中的情态动词使用情况；Alipour & Nooreddinmoosa [19]在比较二语学习者和母语者学术写作语体正式性差异时主要关注第一人称代词、无人参与的指称代词、初始连词、初始连词副词和指示词；Hyland & Jiang [8]在 Chang & Swales [20]对非正式性语言特征的研究基础上，考察了应用语言学、社会学、电子工程和生物学四个学科的 10 个非正式语体特征指标(第一人称代词、无人参与的指称代词、分裂不定式、句首连词或连词副词、句尾介词、列举表达、第二人称代词/限定词、省略形式、直接疑问句和感叹句)。如下表 1 所示。

Table 1. Hyland & Jiang's informal stylistic characteristics

表 1. Hyland & Jiang 的非正式语体特征指标

Index	Example
First person pronouns to refer to the author(s) (I and we)	e.g., "I will approach this issue in a roundabout way."
Unattended anaphoric pronouns (this, these, that, those, it) that can refer to antecedents of varying length	e.g., "This is his raw material."
Split infinitives an infinitive that has an adverb between to and the verb stem	e.g., "The president proceeded to sharply admonish the reporters."
Sentence initial conjunctions or conjunctive adverbs	e.g., "And I will blame her if she fails in these ways."
Sentence final preposition	e.g., "A student should not be taught more than he can think about."
Listing expressions ("and so on", "etc.", "and so forth" used when ending a list)	e.g., "These semiconductors can be used in robots, CD players, etc."

续表

Second person pronouns/determiners to refer to the reader (you and your)	e.g., "Suppose you are sitting at a computer terminal which assigns you role R."
Contractions	e.g., "Export figures won't improve until the economy is stronger."
Direct questions	e.g., "What can be done to lower costs?"
Exclamations	e.g., "This is not the case!"

本文的非正式语体特征的指标主要借鉴了以上 Hyland & Jiang [8] 的分类，原因有二。一是这些特征在学术写作教材中都有提及[21]，且在多个期刊指南中被禁止使用[20]。二是已有研究证明这十类特征在学术文本的非正式性分析中最为常见，尤其是第一人称代词、指示代词和句首连词/连接副词这三种形式 [8] [22] [23]。然而考虑到学术写作应力求精简，适当地使用无人参与的指称代词会使表达更集中和简洁。因此，本文没有采用无人参与的指称代词作为非正式语体特征的指标参考。此外，学术写作应倾向于特定的、具体的描述。而一些例举的内容却通常具有模糊和抽象的含义，在阅读过程中也很容易被忽视。鉴于此，本文将列举表达也纳入到非正式语体特征的指标当中。

如表 2 所示，本文筛选出的 5 项非正式语体特征指标分别为：第一人称代词、第二人称代词、句首连词/连接副词、列举表达及缩略词。

Table 2. Indexes of informal stylistic characteristics

表 2. 非正式语体特征指标

指标	内容
第一人称代词	I, me, my, mine, we, us, our, ours
第二人称代词	you, your
句首连接词/连接副词	and, but, or, so, yet, again, also, besides, however, indeed, still, thus
列举表达	and so forth, and so on, etc.
缩略词	can't, won't, it's, wouldn't, couldn't, shouldn't, hasn't, haven't, isn't, aren't, weren't, don't, doesn't, let's, what's...

3. 研究设计与方法论

3.1. 研究问题

本文的研究问题归结起来有以下两个：

- 1) 从非正式语体特征的各项指标上看，中国和美国外国语言文学专业的硕士毕业论文分别有怎样的差异？
- 2) 造成这些差异的可能原因是什么？

3.2. 研究工具

本文的研究工具有两种，分别为 Lancsbox6.0 和北外语料库分析工具 Log-likelihood and Chi-square Calculator1.0。

Lancsbox6.0 是一款多功能语料分析工具，适用于语料库语言学、翻译学、外语教学等领域的研究者。它可以用于计算与文本的频次结构相关的计量指标，除此之外，它还提供了统计功能和增加了可视化显

示节点词的搭配关系。本文主要用它来计算语料库中的形符和类符数量以及非正式性特征的各项指标中每个单词或短语在语料库中出现的频率。

另外,为了判断非正式性特征的各项指标是否存在显著差异,本文使用北外语料库分析工具 Log-likelihood and Chi-square Calculator1.0 计算对数似然函数值(LL),对数似然值和显著性成正比,对数似然值越大,显著性越高,研究结果更加具有统计学上的意义。

3.3. 语料来源及筛选

首先,本文自建了两个语料库,命名为“中国硕士论文(CMT) Chinese Master Thesis”和“美国硕士论文(AMT) American Master Thesis”,语料库的内容分别包含了中国硕士研究生的论文摘要和美国硕士研究生的论文摘要。之所以选取摘要部分,是因为摘要是论文不可或缺、也是独立存在的一部分。摘要是论文的精髓,能准确地反映整篇论文的中心思想,精简地总结出整个研究的来龙去脉,包括绪论、方法论、结论和讨论。鉴于摘要在研究论文中的重要作用,它也可以被用来探究文章文体的正式程度。

CMT 收录了 2018 至 2022 年间三所大学的外国语言学和应用语言学的硕士论文摘要,这三所大学分别是浙江大学(ZU)、大连海事大学(DMU)和长春理工大学(CUST)。浙江大学属于 985 项目(中国 985 计划旨在建设世界一流大学)和 211 项目(中国 211 项目旨在加强约 100 所高等教育机构和重点学科)。大连海事大学是 211 计划的成员之一。长春理工大学不属于这两个项目。其中,来自浙江大学的摘要有 30 篇,大连海事大学有 40 篇,长春理工大学的有 20 篇,共 90 篇。

为了尽量与 CMT 的语料容量保持一致,加强两个语料库的可比性,同时考虑到美国硕士毕业论文的资源开放性问题,AMT 包含了来自亚利桑那大学、加利福尼亚大学、印第安纳大学等共 120 篇论文摘要,2018 至 2022 年每年 24 篇。

接着,把两个数据库分别导入 Lancsbox6.0 中,计算出两个语料库中的形符与类符数量。其中 CMT 的形符(文本中出现的所有词的个数)共计 38,716 个,类符(文本中出现的不重样的词的个数)4397 个;AMT 中形符 40,165,类符 6363 个。两个数据库的形符总量为 78,881。两个数据库中的形符数量差距不大,类符数量相近,因此本研究的数据是比较合适的。具体数目如下表 3 所示:

Table 3. Number of tokens and type for CMT and IRA
表 3. CMT 和 IRA 的形符和类符数量

文本	形符	类符
CMT	38,716	4397
AMT	40,165	6363
Total	78,881	10,760

4. 研究结果与讨论

综上所述,本文的非正式语体特征分为 5 大指标,接下来将依次分析 5 类指标的数据。首先使用软件 Lancsbox6.0 计算指标中每个单词或短语分别在两个语料库中出现的频率。接着为了探究这些指标之间是否存在显著差异,本文利用北外语料库分析工具 Log-likelihood and Chi-square Calculator1.0 计算它们的对数似然函数值(LL),根据 Lee *et al.* (2019),在 5%水平下,LL \geq 3.84 时,意味着 $p < 0.05$; 1%水平下,LL \geq 6.63,则 $p < 0.01$; 在 0.1%水平下,LL \geq 10.83,相当于 $p < 0.001$; 在 0.01%水平下,LL \geq 15.13 代表 $p < 0.0001$ 。除此之外,该计算器同时显示了具体 p 值,若 $p < 0.05$,则两个数据库的此项指标呈现显著差异,若 $p > 0.05$,则无显著差异。

本文将每项指标计算得到的结果以表格形式分别呈现在下文。

4.1. 第一人称代词

由表 4 可知, 比起中国, 美国硕士研究生明显更偏好使用第一人称代词, 并且在 AMT 语料库中, “I”、“my”“we”、“our”四个词项的出现频次显著多于 CMT, 其中, “we”和“I”在 AMT 中出现频次最高, 同时, “we”在 CMT 中使用的数量也最多。这一发现表明, 尽管学术写作教材、词典和期刊指南禁止学术文本中使用第一人称代词[24], 中西方学者似乎均认可这些语言特征在语篇连贯、作者声音构建和引导读者上的作用。

Table 4. Word frequency and statistical significance of first person pronouns

表 4. 第一人称代词的词频和相关显著性

	CMT	AMT	LL	p-value	Significant
I	3	127	146.97	0.0000	√
me	0	1	1.35	0.2453	×
my	0	24	32.40	0.0000	√
mine	0	1	1.35	0.2453	×
we	18	113	73.30	0.0000	√
us	5	11	2.09	0.1482	×
our	5	28	16.84	0.0000	√
ours	0	0	0	/	×
Total	31	305	249.04	0.0000	√

注: LL = log-likelihood value; √ = $p < 0.05$; × = $p > 0.05$ 。

根据 Hyland [25], 论文作者使用第一人称代词是作者投射个人权威的表现, 也是寻求观点认同的有效手段, 说明作者对研究的独特价值充满信心。中国硕士研究生倾向于少用或尽力避免使用第一人称代词, 即使在学术论文中有使用, 也多半使用其复数形式(we), 徐防[26]推测, 这是由于英语学习者对研究缺乏足够的信心, 担心承担声言责任。声言与援引的重要区别在于前者意味着直接承担责任, 后者间接承担责任[27]。声言鲜明地体现作者对所传递的信息承担直接责任, 同时表明研究观点的主观性和可探讨性。第一人称代词(特别是单数)的使用增加了作者的显现度, 是作者的直接介入。传统观点认为, 论文写作应强调客观、中立, 避免作者个人介入, 但“事实并非不言自明, 需要进行个人解读”[28]。第一人称代词“I”的使用是学术语篇中声言责任策略的重要实现手段, 直接体现作者对所传递的信息承担责任, 也由此显现作者对该研究的个体贡献。“we”的使用一方面弱化作者在研究过程中的作用, 但同时又将其话语社团成员的身份呈现在读者面前, 既体现了作者的谦虚, 又隐含地揭示了自己的贡献。

4.2. 第二人称代词

如表 5 所示, 第二人称代词在两个语料库中出现的频率都极低, 且没有表现出显著差异。学术论文里的第二人称代词是对一般或作为个体的读者的显性指称, 也是对读者身份和存在的显性认同[29], 从而强调了读者与作者的互动, 而在正式文体中这种互动关系应该避免, 因此第二人称代词的过度使用会使文章显得不正式。

Table 5. Word frequency and statistical significance of second person pronouns
表 5. 第二人称代词的词频和相关显著性

	CMT	AMT	LL	p-value	Significant
you	1	3	0.97	0.3236	×
your	1	0	1.42	0.2329	×
Total	2	3	0.16	0.6834	×

首先，第二人称代词一般使用在上级对下级的场合，而在学术写作中，第二人称代词一般是专家对学生使用的，对硕士研究生来说，相较于持有各种身份的读者，如前辈、老师、主编等，他们都更倾向于把自己的地位降低，以体现谦逊的态度；其次，抒发观点、提出意见或者给出研究计划时使用第二人称代词会显得咄咄逼人，很容易给读者带来不悦；再次，在论文观点的呈现过程中，作者和读者主要属于“单方阐述，另一方阅读和理解”的关系，不同于现实情境中双方面对面交谈的口语形式，因此缺乏使用第二人称的特殊语境，如果一味使用第二人称，则会使论文带有很强的口语性，其用语规范性和学术严谨性也就会被大大削减。这同时适用于中国和美国的硕士研究生。

4.3. 句首连接词/连接副词

学术文本中副词常位于句中[21]，而非正式语体副词常位于句首或句尾[21]。中国硕士研究生倾向于将连接副词如“and”、“so”“besides”和“still”置于句首，并在与 AMT 的对比中表现出显著差异。这种前置倾向，一方面可能在于中国硕士研究生同其他二语学习者一样，忽视了连接副词句中位置的语体得体性，另一方面可能源于他们构建语篇连贯和衔接的方式。他们倾向使用显性的衔接手段——句首连词/连接副词，将两个语段连接起来，在句子间构建连贯性。AMT 中出现的这种衔接方式大部分起指示的作用，强调学术论文中的逻辑。这与我国传统英语写作教学息息相关，在课堂中，我们常常被教导使用连接词来体现英语中的逻辑关系，这也是与汉语写作的差异之一。

Table 6. Word frequency and statistical significance of sentence initial conjunctions or conjunctive adverbs
表 6. 句首连词/连接副词的词频和相关显著性

	CMT	AMT	LL	p-value	Significant
and	20	1	21.78	0.0000	√
but	3	2	0.24	0.6244	×
or	0	0	0	/	×
so	3	0	4.27	0.0388	√
yet	0	3	4.05	0.0442	√
again	0	1	1.35	0.2453	×
also	1	3	0.97	0.3236	×
besides	14	1	13.93	0.0001	√
however	41	35	0.72	0.3961	×
indeed	2	1	0.38	0.5389	×
still	16	1	16.52	0.0000	√
thus	8	4	1.51	0.2191	×
total	108	52	22.13	0.0000	√

另外,值得一提的是,由表 6 可知,句首连词“however”分别在两个语料库中出现的频率均为最高,这与 Lee *et al.* [23]、Alipour & Nooreddinmoosa [19]的研究发现——在母语者和非母语者的学术论文中,出现频率最高的句首连接词是“however”——相一致。“however”呈现一种清晰而强烈的转折关系,可用于转换话题、提示局限性或例外情况、以及强调对立观点等等,在学术写作中被看作一种重要的语篇衔接手段。另外,中美学者似乎均认可句首连接词类的语言特征在语篇连贯、作者声音构建和引导读者上的作用,在不同学科论文中均被高频使用。

4.4. 列举表达

表 7 数据显示,AMT 中没有使用列举表达,而在 CMT 中有少量使用,并表现出与 AMT 的显著差异。具体来说,囿于有限的语言手段,在表达列举时,中国硕士研究生倾向于使用 *and so on* 和 *etc.* 两个常用表达式,且并未意识到该类表达用于句尾的非正式性。这一发现与 Petch-Tyson [30]和 McCrostie [31]的研究一致,他们通过定量研究分析发现二语学习者文本中列举表达的出现频率显著多于母语者。本族语学习者清楚此类表达的模糊性,学术文本中甚少使用或不使用[23]。

Table 7. Word frequency and statistical significance of listing expressions

表 7. 列举表达的词频和相关显著性

	CMT	AMT	LL	p-value	Significant
<i>and so forth</i>	0	0	0	/	×
<i>and so on</i>	6	0	8.54	0.0035	√
<i>etc.</i>	6	0	8.54	0.0035	√
Total	12	0	17.08	0.0000	√

4.5. 缩略词

表 8 显示,同列举表达一样,缩略词在 IRA 中没有出现,CMT 中少量使用了“*what’s*”、“*don’t*”和“*aren’t*”,单个词项并没有表现出差异,而在各词项总数的对比中呈现出显著差异,CMT 中使用的缩略词显著多于 AMT。学术写作的目的是精准、客观地传达信息。英语缩略词体现了语言使用的省力原则,形式简洁且表意丰富,然而,鉴于中英文化差异以及不同语言的不对等性,很多时候,英语缩略词很难做到百分之百将意义表达到位,有时甚至引起歧义或混淆,因为它们可能有多种解释或受到不同变体的影响。在类似学术论文的正式文体中,使用单词的完整形式有助于作者以一种更超然、更公正的方式呈现信息,能减少潜在的误解,有利于建立文本的可信度和权威性,让读者专注于对研究内容和研究方法的思辨和分析,而非作者个人的声音。

Table 8. Word frequency and statistical significance of contractions

表 8. 缩略词的词频和相关显著性

	CMT	FMT	LL	p-value	Significant
<i>what’s</i>	2	0	2.85	0.0916	×
<i>don’t</i>	1	0	1.42	0.2329	×
<i>aren’t</i>	1	0	1.42	0.2329	×
Total	4	0	5.69	0.0170	√

4.6. 讨论

逐个分析完单项指标，现将以上 5 个指标的计算总结果汇总于表 9 中，加以讨论。

Table 9. Statistical significance of the indexes of informal stylistic characteristics

表 9. 非正式语体特征指标的相关显著性

	CMT	FMT	LL	p-value	Significant
第一人称代词	31	305	249.04	0.0000	√
第二人称代词	2	3	0.16	0.6834	×
句首连接词/连接副词	108	52	22.13	0.0000	√
列举表达	12	0	17.08	0.0000	√
缩略词	4	0	5.69	0.0170	√

总体上看，五项指标中，除第二人称代词外，其余四项指标在中美语言学硕士研究生毕业论文的对比中均存在显著差异。进一步说，除第一人称代词在 AMT 中出现频率更高外，句首连接词/连接副词、列举表达及缩略词都在 CMT 中体现出更频繁的使用。同时，这也说明，在第一人称代词的使用上，美国语言学硕士研究生的非正式程度高于中国语言学硕士研究生，而中国硕士研究生在句首连接词/连接副词、列举表达及缩略词方面的非正式程度均高于美国硕士研究生。造成这些差异的可能原因复杂多样，本文主要讨论以下三点：

第一，英汉语言特征存在差异。中国硕士研究生少用第一人称代词的可能原因在于汉语中习惯“隐去”人称代词，根据汉语习惯，前句出现一个主语，后句如仍为同一主语，就不必重复出现。而在英语句子中，通常每句都有主语，必须由名词或者代词充当句子的施事与受事，不可省略。此外，英语重形合，英语的词语、分句以及句子之间借助语言形式手段(关联词、介词)来实现词语或者句子的连接，通过词汇本身的形态变化来表达语法意义合逻辑关系，词缀丰富。而汉语重意合，呈“隐形”，按照事情发展的先后顺序(时序、因果、时空)逻辑顺序，以短句的形式行文推进，句子之间无过多的形式连接，叙事从容不迫，层层展开。因此中国硕士研究生尤其注重句首连接词或连接副词的使用，尽管有时衔接并不自然，甚至牵强，也要达到表面上的“形合”。

第二，语体意识的差异。语体意识类型化不是一个瞬时产物，而是一个由不确定态到范式化的历时沉淀过程，是具体的、个人的话语或语篇经反复运用所形成的特定语言社群全体成员约定俗成的语用范式[32]。作为以英语为母语的学者，美国硕士研究生掌握了更系统的语言知识和更丰富的英语写作经验，且对英语的敏感度更高，通常能对某些有歧义或语义含混的缩略词快速作出反应，因此在学术写作中会有意识地避免使用类似列举表达、缩略词等标志着语体非正式性的特征。而中国硕士研究生作为二语学习者，语体意识较之母语者有着先天性差距，显得略为薄弱，对文体的把握缺少熟练度和精确度，在表 8 后三项指标中的非正式语体程度偏高，说明该群体还需不断加强语体意识的培养。

第三，学术英语教学的要求。在国内学术英语论文写作教学中，通常要求学生在论文中少用第一人称代词，并且大部分期刊也明文规定文章中尽量不要出现第一人称单数作为作者自我指称语，而可以使用第三人称指示语来代替，认为“I”的使用显得论文主观色彩浓厚，而“we”的指代不明晰。所以，中国硕士研究生在完成学位论文发表自己的观点时总会不自觉地少用第一人称单数，避免给读者留下过于强调论文完成者的印象，而少部分偏向于使用“we”，这背后深层次的动因可能源于中西文化的差异——西方文化中强调个人主义，独立性自我观念强烈，人际关系较为松散；而中国文化中推崇集体主义，

自古以来受家族式社会结构和儒家文化思想的熏陶下, 人际关系较为紧密, 强调个人依附于集体而存在[33], 强调集体的力量, 而不是单独突显某一个体所发挥的作用。除此之外, 国内学术英语教学将文本间紧密的逻辑性看作英语鲜明的特点, 往往强调在句子衔接处使用句首连接词或连接副词将其表现出来。

5. 结论

本研究聚焦学术语篇中的非正式语体特征, 以中国和美国语言学专业的硕士研究生的硕士毕业论文为语料, 建成两个数量相当的语料库, 通过语料库分析软件, 经对比研究发现, 总体上, 在语言学领域, 美国硕士研究生在第一人称代词方面的非正式程度高于中国硕士研究生, 中国硕士研究生在句首连接词/连接副词、列举表达及缩略词方面的非正式程度均高于美国硕士研究生。此外, 双方均很少使用第二人称代词, 并在数量上没有表现出显著差异。基于以上分析结果, 本文从英汉差异、语体意识及学术英语教学三方面探讨了差异产生的原因。

这些差异应该引起我们对英语教学的反思, 语体这一重要因素常常容易被遗忘在角落, 我们在教学中应结合教学内容对英汉思维方式的不同进行比较, 让学生尽可能多地了解英汉语言本体的特征, 从而体会英汉思维的差异和相似之处, 在两种语言的碰撞和对比中增强语体意识和跨文化意识, 培养对英语文化的敏感性, 在英语学习中克服汉语思维的影响, 以期助力提升我国研究生的学术英语写作能力, 为我国国际学术影响力的提高贡献绵薄之力。

致 谢

本本作者对所有提供文献、资料、研究思想和设想的前辈及计算机软件的开发致以深切的谢意和祝福!

参考文献

- [1] 丁金国. 语体风格分析纲要[M]. 广州: 暨南大学出版社, 2009.
- [2] 王桂亮, 潘世松. 新文科背景下语体学和语言风格学的知识建构与发展规划——第十四届语体与语言风格学术研讨会述评[J]. 当代修辞学, 2023(4): 94-95.
- [3] 王用源, 陈宇豪. 语体语法视角下语体正式度的自动测量[J]. 语言教学与研究, 2023(4): 79-90.
- [4] 冯胜利. 《诗经》语体与《诗经》训诂[J]. 社会科学战线, 2023(7): 139-147.
- [5] Fairclough, N. (2001) *Language and Power*. 2nd Edition, Longman, Essex.
- [6] Foster, J. (2005) *Effective Writing Skills for Public Relations*. 3rd Edition, Kogan Page, London.
- [7] Biber, D. and Gray, C. (2016) *Grammatical Complexity in Academic English: Linguistic Change in Writing*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511920776>
- [8] Hyland, K. and Jiang, F. (2017) Is Academic Writing Becoming More Informal? *English for Specific Purposes*, 45, 40-51. <https://doi.org/10.1016/j.esp.2016.09.001>
- [9] Praminatih, G.A., Kwary, D.A. and Ardaniah, V. (2018) Is EFL Students' Academic Writing Becoming More Informal? *Journal of World Languages*, 5, 101-112. <https://doi.org/10.1080/21698252.2019.1570664>
- [10] 姜峰. 基于多维分析的学术语篇语体特征的历时考察[J]. 外语教学与研究, 2020, 52(5): 663-673, 798.
- [11] 文秋芳, 丁言仁, 王文宇. 中国大学生英语书面语中的口语化倾向: 高水平英语学习者语料对比分析[J]. 外语教学与研究, 2003, 35(4): 268-274.
- [12] Gao, X. (2020) A Comparable-Corpus-Based Study of Informal Features in Academic Writing by English and Chinese Scholars across Disciplines. *Ibérica*, No. 39, 119-140. <https://doi.org/10.17398/2340-2784.39.119>
- [13] Yang, Y. and Pan, F. (2023) Informal Features in English Academic Writing: Mismatch between Prescriptive Advice and Actual Practice. *Southern African Linguistics and Applied Language Studies*, 41, 102-119. <https://doi.org/10.2989/16073614.2022.2088579>
- [14] Brown, P. and Levinson, S. (1978) *Politeness: Some Universals in Language Usage*. Cambridge University Press,

- Cambridge.
- [15] Richards, C.J. and Platt, J. (2000) Dictionary of Language Teaching and Applied Linguistics. Foreign Language Teaching and Research Press, Beijing.
- [16] Lillis, T. (2013) The Sociolinguistics of Writing. Edinburgh University Press, Edinburgh.
<https://doi.org/10.1515/9780748637492>
- [17] Barton, D. (1994) Literacy: An Introduction to the Ecology of Written Language. Wiley-Blackwell, Oxford.
- [18] Aijmer, K. (2002) Modality in Advanced Swedish Learners' Written Inter-Language. In: Granger, S., Hung, J. and Petch-Tyson, S., Eds., *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*, John Benjamins Publishing Company, Amsterdam, 57-76. <https://doi.org/10.1075/llt.6.07ajj>
- [19] Alipour, M. and Nooreddinmoosa, M. (2018) Informality in Applied Linguistics Research Articles: Comparing Native and Non-Native Writings. *Eurasian Journal of Applied Linguistics*, **4**, 349-373. <https://doi.org/10.32601/ejal.464196>
- [20] Chang, Y.-Y. and Swales, J.M. (1999) Informal Elements in English Academic Writing: Threats or Opportunities for Advanced Non-Native Speakers? In: Candlin, C.N. and Hyland, K., Eds., *Writing: Texts, Processes and Practices*, Longman, London, 145-167. <https://doi.org/10.4324/9781315840390-8>
- [21] Swales, J. and Feak, C. (2012) Academic Writing for Graduate Students: Essential Tasks and Skills. 3rd Edition, University of Michigan Press, Ann Arbor, 22-25. <https://doi.org/10.3998/mpub.2173936>
- [22] Hundt, M. and Mair, C. (1999) "Agile" and "Uptight" Genres: The Corpus-Based Approach to Language Change in Progress. *International Journal of Corpus Linguistics*, **4**, 221-242. <https://doi.org/10.1075/ijcl.4.2.02hun>
- [23] Lee, T.B. and Maxwell, J. (2019) Breaking the Rules: A Corpus-Based Comparison of Informal Features in L1 and L2 Undergraduate Student Writing. *System*, **80**, 143-153. <https://doi.org/10.1016/j.system.2018.11.010>
- [24] Vandermeij, R., Meyer, V., Van Rys, J. and Sebranek, P. (2012) The College Writer. 4th Edition, Wadsworth, Boston.
- [25] Hyland, K. (2002) Authority and Invisibility: Authorial Identity in Academic Writing. *Journal of Pragmatics*, **34**, 1091-1112. [https://doi.org/10.1016/S0378-2166\(02\)00035-8](https://doi.org/10.1016/S0378-2166(02)00035-8)
- [26] 徐防. 中国学生英语学术写作中身份语块的语料库研究[J]. 外语研究, 2011(3): 57-63.
- [27] Sinclair, J. (1986) Talking about Text: Studies Presented to David Brazil on His Retirement. University of Birmingham, Birmingham.
- [28] Hyland, K. (1998) Boosting, Hedging and the Negotiation of Academic Knowledge. *Text*, **18**, 349-382.
<https://doi.org/10.1515/text.1.1998.18.3.349>
- [29] Hyland, K. (2005) Representing Readers in Writing: Student and Expert Practices. *Linguistics and Education*, **16**, 363-377. <https://doi.org/10.1016/j.linged.2006.05.002>
- [30] Petch-Tyson, S. (1998) Learner English on Computer. Longman, London, 107-118.
- [31] Mccrostie, J. (2018) Writer Visibility in EFL Learner Academic Writing: A Corpus-Based Study. *ICAME Journal*, **32**, 97-114.
- [32] 高霞. 中国学者英文学术论文中的非正式语体特征研究[J]. 外语与外语教学, 2021(2): 92-102, 150.
- [33] Howard, M. (2014) Writing Matters: A Handbook for Writing and Research. 2nd Edition, McGraw Hill, New York.