

doi:10.13582/j.cnki.1674-5884.2020.05.009

# 英语专业本科毕业论文结语口语化词汇特征研究

张金福

(九江学院 外国语学院,江西 九江 332005)

**摘要:**以 Paquot 的 24 个口语化词汇特征项为理论依据,基于自建语料库和孙亚真(2012)的研究数据,应用对比中介语分析方法,对英语专业本科毕业论文结语做对比分析,以了解其口语化特征。研究结果显示,就整体而言,英语专业本科毕业论文结语口语化词汇特征现象相当明显,但随着其二语写作者水平的提高,其口语化词汇特征呈衰减趋势。该研究给本科英语学术论文写作教学及教材的编写与应用均带来一些启示。

**关键词:**英语专业本科毕业论文;结语;语料库;对比中介语分析;口语化词汇特征

**中图分类号:**H3;G642.0      **文献标志码:**A      **文章编号:**1674-5884(2020)05-0052-08

结语(结论)通常是各类学术、学位论文正文的最后部分,是作者总结研究的精华所在,因此也是读者获取重要研究信息的关键部分,其重要性显而易见。然而笔者以“结语”“结论”“conclusion(s)”等为检索词,在 CNKI、万方、维普、Wiley、ProQuest、EBSCOhost、Taylor & Francis 等大型国内外数据库检索发现有关结语/结论/conclusion[s](以下统称为“结语”)的研究论文仅 12 篇(起止时间为 2000 年 1 月 1 日至 2019 年 12 月 31 日,国内只统计了 CSSCI 期刊论文)。进一步梳理发现,这些文献主要依据 Swales 的 CARS 模型(creating a research space)<sup>[1]</sup>对学术语篇结论部分的结构及语言特征进行分析,并表现出以下特点:(1)大多聚焦于非英语本族语学者与英语本族语学者结语比较,或者非本族语硕士、博士论文结语与本族语学者论文结语之间的比较,涉及非本族语学习者本科毕业论文(学士学位论文)层次极少。(2)国外研究起步早于国内,且做体裁分析较多(占 80%)。国内研究起步相对较晚,但发展较快,主要在国外研究基础上依据

Coxhead(2000)AWL 词表<sup>[2]</sup>做学术词汇特征研究,或者从体裁分析和元话语研究的视角做论文的首尾(引言和结论)呼应效果探究<sup>[3]</sup>。鲜见从口语化特征视角研究本科毕业论文,更少见以历时的视角对本科、硕士、博士论文做对比研究。因此,本文试图通过语料库来了解英语专业本科毕业论文结语口语化特征。

## 1 文献回顾及理论依据

自 Biber《口笔语差异》<sup>[4]</sup>出版以来,国内外有关语言变体研究层出不穷。在国外,Biber 建立了多维度/多特征(MD/MF)模型并提出区别从口头语到书面语的六个维度,即交互性/信息生成(interactive/information [D1])、叙述性/非叙述性(narrative/nonnarrative [D2])、所指明确/所指依赖环境(explicit/environment-dependent [D3])、劝诱/非劝诱(induce/noninduce [D4])、抽象/非抽象(abstract/nonabstract [D5]),另加即时性信息阐述(online informational elaboration)。Leech, Rayson & Wilson 首先发现

收稿日期:20200114

基金项目:江西省高校人文社会科学研究项目(YY1407)

作者简介:张金福(1963-),男,湖北襄阳人,教授,博士,主要从事应用语言学、语料库语言学等研究。

口语与书面语在词汇上的轻微差异<sup>[5]</sup>。Lorenz 报告了德语为母语的学习者写作中对副词 so 和连词 because 的超用问题<sup>[6]</sup>。Nartia & Sugiura 发现在口语中高频使用的 of course(表示可能和确定)在瑞典语、法语和日语为母语的学习者写作中超用<sup>[7]</sup>。Granger 调查了法语为母语的二语学习者对口语词汇特征项“I think”的超用现象<sup>[8]</sup>。在国内,杨惠中依据 MD/MF 模型,利用《上海交通大学科技英语语料库》和《LOB 语料库》进行语言变异研究。他发现学术英语与小说之间在 D3 和 D4 上没有显著差异。他认为学术英语与小说相比信息量更大、更抽象且叙述性更低<sup>[9]</sup>。马广惠采用 Biber 的 MD/MF 模式,比较中国大学生和美国大学生作文,发现中国大学生在作文中更多使用第二人称代词、语篇虚词、连词和形容词,而美国大学生更偏爱写长文章、because 引导状语从句、that 引导定语从句、that 引导宾语从句和说服性动词<sup>[10]</sup>。Chen 讨论了台湾学习者在写作中对 besides 出现于句首的超用现象<sup>[11]</sup>。文秋芳等比较中国大学生与英语为母语者使用第一人称单复数(I, me, my, mine, we, our, ours)、第二人称单复数(you, your, yours)、语境代词(here, now, this essay)、强调小品词(just, really)和模糊词(kind, sort of, and so on 等),发现中国大学生英语学习者作文中表现出明显的口语化特征趋势。他们还发现中国大学生英语学习者随着二语能力的提高,其口语化趋势减弱<sup>[12]</sup>。

以上国内外研究为丰富二语习得理论作出了巨大贡献,但遗憾的是关于中国学习者英语学术论文的口语化特征研究极少,且不系统。

Paquot 应用中介语对比分析法(CIA Approach)比较 ICLE 语料库专家写作与学习者写作,她把中介语特征分为六类,认为学习者缺乏语域意识是典型特征,并列出了学习者写作中 24 个超用口语化特征项目<sup>[13]</sup>。该研究是目前学界公认比较系统地考察学术论文口语化特征的方法。

本研究基于自建的《中国高等学校英语专业本科毕业论文结语语料库》(以下简称为“本科结语”)和《英语学者学术论文结语语料库》(以下简称为“学者结语”),通过 Paquot 的 24 个口语特征词<sup>[13]</sup>频率的对比对大陆高校英语专业本科毕业论文结语进行对比研究,并借用孙亚真基于硕士论文语料库(Chinese MA Theses Corpus, 以下简

称为“CM”)和博士论文语料库(Chinese Doctoral Dissertations Corpus, 以下简称为“CD”)的研究数据<sup>[14]</sup>作历时考察,以探索英语专业“本科结语”中口语化词汇发展特征。

## 2 研究设计

### 2.1 研究问题

本研究的总体目标是对汉语为母语以英语为外语的英语学习者的本科毕业论文结语的口语化词汇特征研究。研究问题如下:

1) 与学者结语相比,本科结语是否存在口语化特征?

2) 如果存在口语化特征,那么这种口语化特征是什么?程度如何?

3) 与 CM 和 CD 对比,本科结语的口语化词汇特征是否有变化?变化趋势是怎样的?

### 2.2 研究对象及语料库采集原则

语料库建设是本研究的基础。为保证本科结语语料的平衡以提高研究的效度,我们遵循以下原则:一是区域平衡。即保证中国知网划分的全国七大地区高校都有语料。二是兼顾“双一流”和非“双一流”高校,并考虑两类学校的占比。其中“双一流”大学 10 所,非“双一流”大学 40 所,共 50 所,约占全国有外语专业学校数的 5%。三是注意语料撰写年份的平衡。所采集语料撰写时间是 2017~2019 年(共 3 年)。四是控制每所学校论文的总数。每所学校最多不超过 2 篇/年。五是应用 CNKI 科研诚信管理系统研究中心开发的“学术不端文献检测系统”进行检测,剔除有抄袭行为的结语。最终从搜集的 2 500 多篇论文中抽取 300 篇符合要求的结语,建立《本科结语》语料库。另外建立《学者结语》语料库,其语料全部来自近 5 年 *Cognition* 杂志上高被引率的 100 篇文章的结语。在语料加工处理时,我们保存了 3 种不同的文本(即生文本、词性赋码文本、词形还原文本),以适应不同研究的需要。

至于 CM 和 CD,我们借用孙亚真(2012)研究论文中的数据,其研究对象为 2008~2010 年大陆高校以汉语为母语学习者的英语硕士和博士论文,专业为应用语言学。其中,CM 总计 30 篇,库容量(Tokens)为 416 711。CD 总计 20 篇,库容量(Tokens)为 1 110 515。为增强可比性,其频率均做标准化处理。

### 2.3 研究方法及工具

主要通过本科结语与学者结语对比来回答第一、二两个问题。通过本科结语与 CM、CD 对比来回答第三个问题。用 TreeTagger 做词性标注,用 WordSmith Tool(7.0)做口语特征词项频数统计并检索出相关语境共现行,用 SPSS 等进行数据统计。

### 2.4 研究过程

数据收集与分析过程如下:

1)为了回答第一个问题,我们分别把本科结语和学者结语导入检索工具 Wordsmith Tool 7.0,用 Paquot 的 24 个口语特征词<sup>[13]</sup>作为关键词得到其在两个语料库的原始频率和相应的语境共现行。这些语境共现行将以 Excel 电子表格形式保存以备做进一步分析。

在做关键词检索时,我们考虑到部分特征词可能出现的派生和曲折变化及大小写。如在检索 look like 时,我们使用“look | looks | looked | looking \* like”作为关键词以确保能检索到所有符合要求的语境共现行。除句子结尾副词 though 和句子起始词 And 外,我们关注了所有其他超用词的大写和小写形式。

2)手工详细检查所有口语特征词项的语境共现行,剔除了不能履行其相应修辞功能的共现行。以口语特征词 like 为例,它在本研究中有例证、比较和对比两个不同功能。在检索统计时,所有的共现行都做了功能归类,剔除了与本研究不相关的共现行。

3)为了保证不同库容的语料库所检索出词频具有可比性,将检索出的词项词频做了标准化处理。

4)用卡平方检测对口语词在两个语料库里的总标准化频率进行比较,测试其不同语料库的标准化频率是否有显著差异。通过逐项比较口语特征词在两个语料库中的频率找出本科结语写作中口语特征词,用卡平方检验法来测量得知与学者结语相比,本科结语中哪些口语特征词超用、哪些少用。

5)为了测试学术论文中口语特征词是否会随着学习者二语水平提高而有所发展变化,我们再用卡平方检验法,通过比较本科结语和 CM、CD 口语特征词使用变化来测量其发展变化的显著性。

6)通过语料库观察和对个别毕业论文撰写者访谈,对典型的和异常的共现行做更具体深入的探究,以了解这些口语特征词在本科结语中超用或者少用的深层原因。

## 3 结果与讨论

本部分通过本科结语与学者结语比较来回答第一、二两个问题。然后通过本科结语与 CM、CD 口语特征词的历时比较研究,了解口语特征词在学习者学术论文写作中的发展特征,以回答第三个问题。

### 3.1 本科结语是否存在口语化特征

通过比较本科结语和学者结语,高级英语学习者在学术写作中的口语化特征得以揭示。两个语料库的整体概貌见表 1。

表 1 本科结语与学者结语的整体概貌数据表

语料类别	本科结语	学者结语
篇数	300	100
形符数(Tokens)	104 057	94 415
类符数(Types)	9 104	9 272
形符/类符比(TTR)	8.75	9.82
标准 TTR	48.60	49.00
平均词长	5.15	5.33
平均句长	21.60	26.28
平均每篇结语长度	346.86	944.15

表 1 显示,本科结语平均篇长(346.86 Tokens)比学者结语(944.15 Tokens)短很多,只有学者结语的 36.7%。其标准形符/类符比、平均词长和平均句长都比学者结语低,这说明本科结语在用词的正式(formal)程度、句子复杂度等方面都比学者结语低,表现出较明显的口语化特征。

### 3.2 本科结语口语化特征如何

语料库检索显示,在 24 个口语特征词中,有 5 个在两个语料库里都没有出现,它们的修辞功能分别用于表达个人意见(to my mind, from my point of view, it seems to me)、引介话题或者观点(I would like to / want / am going to talk about)和在句末表达让步功能(though),因此无法对比分析。为确保数据的有效性,我们又删除了在两个语料库中频率都低于 5 的口语词项(thanks to, that/this is why, look like, by the way)。其他 15 个口语词项被用于两个语料库的分析和比较,以探究本科结语中口语化词汇特征具体情况,详见表 2。

表2 本科结语和学者结语标准词频统计表(15项)

序号	修辞功能	词项	本科结语词频/10万	学者结语词频/10万
1	例证	like [E] *	32	35
2	因果	so	202	40
3		because	95	100
4	比较和对比	like [C] *	36	19
5	添加信息	And	120	24
6		besides	40	1
7	表达个人意见	I think	7	0
8	表达可能性 和确定性	really	19	6
9		of course	10	17
10		absolutely	10	2
11		maybe	12	1
12	引介话题和观点	thing	19	7
13	列出项目	first of all	24	1
14	引用和报告信息	say	20	22
15	概括和归纳总结	all in all	14	1
总计		15	660	276

注:表中标有“\*”的两项方括号为作者所添加。因为 like 在表中出现 2 次,分别代表不同的修辞功能。为了避免混淆和下文叙述方便,在本文中用“like [E]”来代表其“例证”功能,用“like [C]”来代表其“比较和对比”功能。

从表 2 得知,本科结语总词频为 660,而学者结语词频为 276,明显低于本科结语,这说明与学者结语相比,本科结语表现出较明显的口语化特征。仔细观察各个词项,发现有 11 个特别明显,有 4 个词项不明显(要么基本持平,要么略高于本科结语)。详见图 1 更直观。

图 1 呈现了 15 个口语特征词每个词项在本科结语和学者结语的比较结果。在图 1 中,这些

词项在本科结语中的超用和少用现象都列出来了。口语特征词 so, like, and, besides, I think, really, absolutely, maybe, thing, first of all 和 all in all 在本科结语中都超用了。这 11 个词项的超用现象与 Lorenz, Chen 和 Paquot 的发现吻合<sup>[6][11][13]</sup>。下面我们以后 I think 为例来进一步说明。

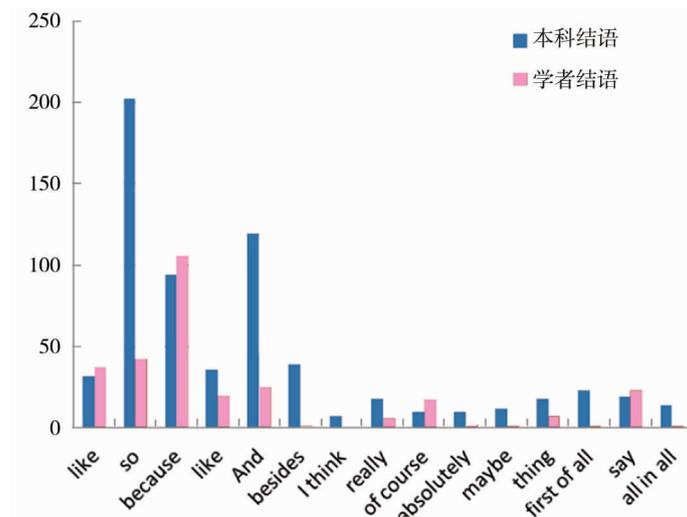


图 1 本科结语和学者结语口语化特征词频率(十万分比)

图 2 是本科结语中含有口语特征词 I think 来表达个人观点的语境共现行。I think 词项在本

科结语中的例证印证了 Paquot 的发现,即 14 种不同母语背景的英语学习者都超用 I think 这个

词项<sup>[13]</sup>,因此,不论母语背景如何,词项 I think 可能是中介语发展特征。

works, it is a popular market winner, but also a classic. I think this novel can be so successful and then become a
end is the transmission of positive energy or negative energy, I think that we cannot t make a positive answer, as long as we
do not know the world and who need for love are hurting mostly. I think Ian's humor is permeated through a strong humanistic
where their life value, is also the life of the writer's novel. I think the moon and six pence is a pair of contradictions, like
. The creation of the vampire image is all based on reality. I think people are under great pressure now. They need to wreak
is almost few research on this analysis. Through my research, I think that through these four aspects of the study, ordinary
easily abandon and escape the reality of our shackles. So I think the best explanation of this book for most of us is that

图2 I think 在本科结语中的语境共现行

表3为两库口语特征词频数卡方检验值。检验数据说明,在0.05的显著性水平上,本科结语结果显示,本科结语与学者结语存在整体差异。检验表现出明显的口语化特征。

表3 本科结语和学者结语中口语特征词频数卡方检验值

Numer of obs = 30R-squared = 0.737 8					
Root MSE = 33.464 5Adj R-squared = 0.456 9					
Source	Partial SS	df	MS	F	Prob > F
Model	44 118.6	15	2 941.24	2.63	0.0 394
dum	5 386.8	1	5 386.8	4.81	0.0 457
s	38 731.8	14	2 766.557 1	2.47	0.0 510
Residual	15 678.2	14	1 119.871 4		
Total	59 796.8	29	2 061.958 6		

为探究其原因,我们分别对10所学校共计10位本科结语撰写者进行了电话访谈。其中典型回复内容及关键词如下:(1)学生甲说,他并不十分清楚这些词属于口语化特征词汇。写作中当他需要表述自己观点时,首先想到的就是“I think”“In my opinion”这类短语。(2)学生乙说,他只记得学术论文写作教材上强调论文不同部分写作内容不同、目的不同,文体要“正式一点”,被动语态要多用,要用第三人称来写,写作课老师也给他们介绍过一些常用词串,但动笔写时还是把握不好文体特点。(3)学生丙说,让他写300~500词的小作文还可以,写3000词以上的论文根本把控不好。因为要注意的问题太多了,所以他首先考虑的是如何保证不出语法错误和拼写错误。(4)当问及高年级是否撰写过“课程论文”时,学生丁说,写过,但不多,课程老师也只是强调要注意格式,不能抄袭,最终老师只是给一个期末总成绩,写得到底如何,自己也不清楚。(5)当问及阅读文献数量及英汉语文献比例时,学生戊说,因为做毕业论文时还要考研、考专八,时间紧张,他只读了10篇左右与毕业论文特别相关的论文,为节约时间以读汉语文献为主,没有认真阅读相关专著。

从学生的回答中我们可以发现,教材对口笔语差异的讲解不够,学术论文写作课教师强调也不够,更缺乏深入细致的对比分析,学生普遍缺乏基础英语写作向学术论文写作发展的中间环节训练以及学生的语类和语域意识缺失。

### 3.3 口语特征词发展趋势如何

最后一个问题是:与CM、CD相比,本科结语中的这些口语化特征词词频是否有变化,变化趋势是怎样的?要回答这个问题,我们需要查看这15个口语特征词在本科结语、CM和CD中的频率。

由表4得知,在15个口语特征词项中,有11项表现出明显的下降趋势、有3项表现出先降后略升态势,但总体来看还是保持下降趋势,因为虽然与CM相比,CD有微升,但与本科结语相比,CD还是呈现下降趋势。细致观察会发现与硕士论文相比,CD中表现上升的3项变化幅度都不大。这可能与语料库不够大有关。通过表4,我们也观察到有一项口语词汇特征词与众不同,表现出从本科、硕士到博士逐级上升趋势。经过对比和思考,我们认为这可能是为了有说服力,无论哪级学位论文都需要例证支撑,而硕士和博士论文撰写者更重视例证对其观点的支撑作用。图3

有关 like[E] 的索引行可以为证。

表 4 本科结语、CM 和 CD 中口语化词项频率变化情况 (单位:十万)

词项	本科结语	CM	CD	变化趋势
like [E]	31.71	41.75	42.41	↗
so	201.81	70.31	42.86	↘
because	94.18	57.83	67.45	↘↗
like [C]	35.56	13.20	17.02	↘↗
And	119.17	55.19	30.35	↘
besides	39.40	21.12	18.01	↘
I think	6.73	1.20	0.00	↘
really	18.26	9.84	5.22	↘
of course	9.61	6.48	3.06	↘
absolutely	9.61	2.88	0.36	↘
maybe	11.53	2.64	0.81	↘
thing	18.26	11.52	5.85	↘
first of all	23.06	7.20	4.59	↘
say	19.22	0.48	0.99	↘↗
all in all	13.45	1.20	0.27	↘

注: ↘ 表示下降趋势; ↗ 表示上升趋势; ↘↗ 表示先降后升

图 4 更加直观地显示本科结语、硕士论文 (CM) 和博士论文 (CD) 中口语化词汇特征的总体变化趋势。图 4 中的三条线分别代表了本科结语、硕士和博士论文中口语化特征词的标准化频率。代表本科结语口语化频率的线处于图的上部, 而代表博士论文的线处于底部, 代表硕士论文口语化特征词频率的线处于中间。这三条线的分布揭示出随着学习者二语水平的提高, 其学术论文里的口语特征词总体上处于下降趋势。除 I think (表示个人意见) 在博士论文中没有出现外, 其余 14 个词项在三个语料库里都出现了。

表 5 显示了本科结语、CM 和 CD 卡方检验结果。自由度为 34, 卡方值为 2.925 E2, 至少在 .000 水平上是显著的。显著水平比对照意义水平 .05 低很多。15 个口语特征词项在本科结语的平均频率是 1.604 0E2, 高于 CM 平均频率 (1.273 5E2) 和 CD 平均频率 (1.199 8E2)。

be found in the translations of terms like "barrister", "counsel", "attorney", and "Inn"  
 Twilight. And some of us are unconfident like Bella. We need a chance to liberate  
 is argued that network phrases own features like changing fast, popularation, nature of recrea  
 between languages. In translation, many factors like culture, economy, education, background of ag  
 reflect to the need of social development. Like culture, language can not be in a  
 organization and messages related to the products like date, quantity and so on. And the  
 , students are able to get the feelings, like disappointment, criticalness, misgiving, happ  
 her hold the positive attitude. Some scholars, like Dong(2000) hold the opinion that the influenc  
 to wreak their stress by some ways like drug, alcohol and so on. This is  
 munication, talker ignore the space background like festival occasion with bad words and funeral  
 still needs to be perfect, other aspects, like grammar, rhetoric device and etc, haven't  
 that exists not only among the wicked like Hitler, but also within the spiritual traditi  
 Interpersonal Metaphor, but also on other scopes like ideational meaning and textual meaning.  
 ciency. As mentioned above, translating methods like literal translation, liberal translation and  
 rences. Generally speaking, translating methods like literal translation, liberal translation and  
 a tragedy to have such a personality like Mr. Hooper. Hawthorne wanted to tell us  
 (DCT) can be used as a method, like open questionnaire. In the DCT, investigators  
 of humor besides the usual viewing angle like pragmatics and rhetoric. We get the results:  
 , it is inevitable to violate the maxims like quality maxim and manner maxim. From the  
 good for English writing, other English studies like reading might be used the method for  
 ls translation needs further theoretical support like skopos theory and adaptation theory. And it  
 actors whereas the analysis of other characters like Stella, Eunice is not sufficient, which may  
 learners are inclined to ignore the issues like tense and subject-predicate agreement especia  
 Nature. In today's society, some issues, like the depletion of energy resources and environ  
 on one type of the commercial advertisement like the skin care advertisement or the car

图 3 本科结语中用 LIKE[E] 用来引出例证的索引行 (抽样)

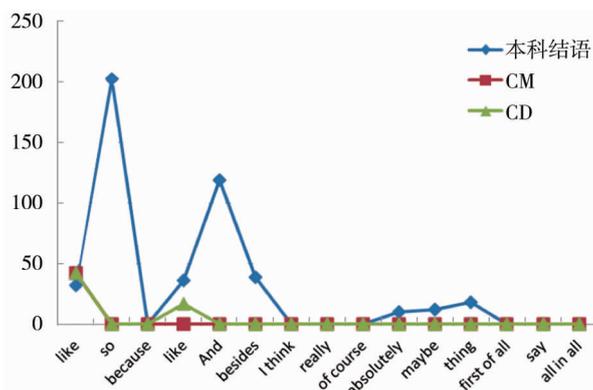


图 4 15 个口语特征词在本科结语、CM 和 CD 中的频数对比 (单位:十万)

表5 本科结语、CM和CD中口语特征词频卡方检验值

	N	Mini	Max	Mean	Std. Deviation	Value	Chi-Square df	Tests
								Asymp. Sig. (2-sided)
本科结语	15	.00	703.13	1.604 0E2	225.51 963			
CM	15	.00	674.46	1.273 5E2	193.44 309	2.925 E2	34	.000
CD	15	.00	596.71	1.1998 E2	171.25 943			

这一研究结果与文秋芳等的研究发现吻合。即读者/作者显现度(口语化特征的代表)随着学生英语水平的提高(大学一、二年级学生与大学三、四年级学生的对比)在降低。另外,随着学生英语水平的提高,他们在写作中倾向于更多使用学术词汇(AWL),更少使用基础词表词汇(GSL)<sup>[12]</sup>。本研究扩展了这一研究结论:随着二语水平的提高,中国英语学习者在他们的学术写作中趋向于更少使用口语特征词。结合 Shaw & Liu 的研究发现(二语写作的语言变化与二语写作发展关系密切)<sup>[15]</sup>,本研究证实了这个结论,即二语水平越高,他们在学术论文写作中使用口语特征词就越少。

#### 4 结语

本文基于自建语料库,采用中介语对比分析法,对英语专业本科毕业论文结语的口语化词汇特征进行了分析。主要研究发现及其对教学的启示归纳如下。

##### 4.1 主要发现

与英语为母语专家学术论文结语相比,英语专业本科毕业论文结语不仅存在口语化词汇特征,而且非常明显。本科毕业论文结语、硕士和博士论文写作者随着其二语写作水平的提高,写作中口语化词汇特征呈衰减趋势。但作为例证引导词的“like”却出现异常情况,即随着二语学习者水平的提高出现递增趋势。我们推测这是因为写作者为了使论文具有说服力,喜欢用 Like 引出例证来支撑自己的观点所致,且高级写作者更擅长这种技巧的应用。

##### 4.2 启示

鉴于本科毕业论文学术写作有明显口语化用词特征现象,我们应多管齐下共同解决这一影响毕业论文质量的“瓶颈”问题。首先教材编写者应高度重视对学生语类和语域意识的培养,要专门开辟章节对口语词汇和学术词汇、口语句型和

典型学术句型进行对比分析,以便引起学生的高度重视。教师要多给学生提供英文范文,并引导学生关注学术论文用词的特点。越来越多的研究证明意义单位常常由半固定短语实现,储存大量常用的短语(多词序列,四词词串最为理想)是母语口语地道流畅表达的关键。高频短语有组织语篇及维持社交关系等语用功能。因为学术性短语具有较为完整的意义,所以有助于提高学习者学术语言输出的自然性和流畅性。现代语料库语言学技术和方法已能比较容易地对其现成的结构整体提取和考察。“内因是变化的根据”,学生要平衡专八考试、研究生考试与毕业论文撰写的关系,增加英文学术论文阅读的数量,在阅读实践中提高语类和语域意识,提高学术英语表达的得体性。

由于用于研究的语料库容量不够大,加上在做历时考察时,所借用的 CM 和 CD 数据是硕士和博士论文整篇文章的,而不是单独结语部分的,这些都可能对研究的信度和效度产生一定影响。因此今后的研究要扩大语料库容量并尽量保证对比语料库语篇的完整性,以增强可比性。

#### 参考文献:

- [1] SWALES J M. Aspects of Article Introductions [M]. Birmingham, UK: The University of Aston, Language Studies Unit, 1981.
- [2] COXHEAD A. A New Academic Word List [J]. TESOL Quarterly, 2000(2):213-238.
- [3] 郑新民,景飞龙. 我国外语类学术期刊实证研究论文首尾呼应效果探析:体裁分析和元话语研究的视角[J]. 外语与外语教学, 2017(4):42-51.
- [4] BIBER D. Variation across Speech and Writing [M]. Cambridge: Cambridge University Press, 1988.
- [5] LEECH G, RAYSON P, WILSON A. Word Frequencies in Written and Spoken English [M]. London: Pearson Education Limited, 2001.
- [6] LORENZ G. Learning to cohere: causal links in native

- vs. non-native argumentative writing [C]// In W. Bublitz, U. Lenk, & E. Ventola (eds.). *Coherence in Spoken and Written Discourse: How to Create It and How to Describe It*. Amsterdam and Philadelphia: John Benjamins, 1999.
- [7] NARITA M, SUGIURA M. The use of adverbial connectors in argumentative essays by Japanese EFL college students [J]. *English Corpus Studies*, 2006 (2): 23-42.
- [8] GRANGER S. Prefabricated patterns in advanced EFL writing: collocations and formulae [C]// In A. P. Cowie (ed.). *Phraseology: Theory, Analysis and Applications*. Oxford: Oxford University Press, 1998.
- [9] 杨惠中. 语料库语言学导论[M]. 上海:上海外语教育出版社, 2002.
- [10] 马广惠. 中美大学生英语作文语言特征的对比分析[J]. 外语教学与研究, 2002(5): 345-349.
- [11] CHEN C W. The use of conjunctive adverbials in the academic papers of advanced Taiwanese EFL learners [J]. *International Journal of Corpus Linguistics*, 2006(1): 113-130.
- [12] 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[J]. 外语教学与研究, 2003(4): 141-147.
- [13] PAQUOT M. *Academic Vocabulary in Learner Writing: From Extraction to Analysis* [M]. New York: Continuum Intl Pub Group, 2010.
- [14] 孙亚真. 中国高水平英语学习者学术写作中的口语化特征[D]. 武汉: 华中科技大学, 2012.
- [15] SHAW P, LIU. What develops in the development of second language writing? [J]. *Applied Linguistics*, 1998 (2): 225-254.

## Research on Colloquial Vocabulary in the Conclusion of Undergraduates' Theses of English Major

ZHANG Jinfu

(School of Foreign Languages, Jiujiang University, Jiujiang 332005, China)

**Abstract:** Based on self-built corpora and Sun Yazhen's data, employing CIA approach, and taking the twenty-four colloquial items identified by Paquot (2010) as theoretical basis, this paper explores the features of colloquial vocabulary in the conclusion part of undergraduates' theses. Results show that, on the whole, the characteristics of colloquial vocabulary in the conclusion of undergraduate's thesis are obvious. However, it has decreased as learners' L2 proficiency develops. This study has provided some references to the teaching of undergraduate's English academic writing, the compilation and application of teaching materials.

**Key words:** undergraduates' thesis of English major; conclusion; corpus; contrastive interlanguage analysis; features of colloquial vocabulary

(责任校对 朱正余)