

A Study into Colloquial Style in English Majors' Writing: From Lexical Dimensions

Weiwei He

School of Foreign Languages, Zhaotong University, Zhaotong Yunnan
Email: he1303648@126.com

Received: Dec. 27th, 2018; accepted: Jan. 8th, 2019; published: Jan. 17th, 2019

Abstract

Based on a corpus of 120 argumentative English compositions on the same topic written by 30 Chinese English majors through their four-year study period, this study, from three dimensions of reader/writer visibility, lexical density and lexical frequency, explores the developmental features of colloquial style in English writing. The study finds that there exists colloquial style in students' compositions with no obvious improvement between year-to-year changes. The first person pronouns are found to be frequently used by students. Lexical density cannot be employed to precisely measure colloquial style in English writing. High-frequency words are overused by students.

Keywords

L2 Writing, Colloquial Style, Reader/Writer Visibility, Lexical Density, Lexical Frequency

英语专业学生写作口语化特征：词汇层面探究

何微微

昭通学院外国语学院，云南 昭通
Email: he1303648@126.com

收稿日期：2018年12月27日；录用日期：2019年1月8日；发布日期：2019年1月17日

摘要

本研究基于30名英语专业学生从大一到大四4年中所写的120篇同题议论文，运用纵贯研究方法从读者/作者显现度、词汇密度和词频三个维度，探讨了二语写作口语化特征的历时发展变化。研究发现：学习者存在写作口语化倾向，并且随着学习者写作水平的提高，写作口语化特征没有呈现明显减弱趋势；二语写作中频繁使用第一人称代词表达观点；词汇密度不能准确反映二语学习者写作口语化特征；学习者在写作中过度使用常见高频词汇。

关键词

二语写作, 口语化特征, 读者/作者显现度, 词汇密度, 词频

Copyright © 2019 by author(s) and Hans Publishers Inc.

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

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



Open Access

1. 引言

英语语体可分为口语语体和书面语语体。口语语体通常指亲密朋友之间进行交谈或邮件中使用的英语, 具有临时性、随意性; 书面语语体通常指在正式书信、正式会议、书籍中使用的英语, 具有严谨性、准确性、严密性[1]。英语议论文文体属于书面语中正式的语体, 强调准确性和逻辑性。

国外学者对书面语体中口语特征的研究相对较少, Ure (1971) [2]指出口语语体词汇密度远低于书面语体词汇密度, 提出词汇密度是口语语体与书面语语体最显著的区别。国内文秋芳等(2003) [3]对不同母语背景下英语学习者/读者/作者显现度、词汇分布进行对比分析, 结果显示中国高水平英语学习者书面语存在较强的口语化倾向, 但随着英语水平的提高有所弱化。吴瑾和邹青(2009) [4]采取中介语对比分析法、从词汇密度和词频两方面, 发现中国学生英语口语没有呈现出较强的语体特征差异。以上研究均是从词汇层面展开的横切面探究, 缺乏对学习者的大学四年的纵贯研究, 且研究维度各异, 研究结果有待进一步验证。

鉴于此, 本研究拟从读者/作者显现度、词汇密度和词频分布三个维度, 探究中国西部地区英语专业学生在 1~4 年级阶段的书面写作中是否存在口语化特征? 若有, 写作口语化特征随着英语水平的提高呈现怎样的发展趋势?

2. 研究设计

2.1. 语料来源

研究语料来源于中国西部地区某省属普通高校英语专业 30 名学生的写作。学生年龄为 18~23 岁, 其中男生 4 名, 女生 26 名。写作主题统一为 2012 年英语专业八级考试题目 *Comment on College Students' Attitude towards Phones Calls and Text Messaging*, 文体要求为议论文, 课堂 45 分钟内完成, 不能查阅词典和参考其它资料, 要求英语书写清晰, 词数不少于 150 词。

2014 年 9 月, 收集完成学生一年级的作文, 2015 年 9 月、2016 年 9 月和 2017 年 9 月分别完成学生二、三、四年级的作文, 历时 4 年收集了学生在同一写作要求下完成的 120 篇有效作文, 词数分别为 6369, 8384, 8601 和 7011, 共计 30,365 个形符。

2.2. 语料分析步骤

1) **语料处理**: 本研究文本语料只计入词形、词意和搭配均无误的词汇, 删除文本中出现的人名、地名等。

2) **读者/作者显现度**: 运用 AntConc 软件统计处理后文本各特征频率。按照标准化公式统计出每 10,000 词的相对频率: 标准化频率(每千词) = 观察频数 ÷ 总体频数 × 10,000 (注: 观察频数即检索词项实际出现的次数; 总体频数即语料库的大小或总形符数) [5]。

3) **词汇密度**: 将英语中的功能词(限定词、代词、介词、连接词、助动词、感叹词等)设为停用词(stop list), 再运用 AntConc 软件统计出每篇处理后文本中实词形符(token)量(即名词、动词、形容词和副词)占各文本总字数之比。为便于统计, 将所得数据去掉百分号。

4) **词频分布**: 使用词频分析工具 RANGE 分析处理后文本词频分布情况——包括最常用 1000 词, 次常用 1000 词, 学术词汇和范围外词汇, 并分别计算各年级文本词类(Type)占总词数的比率。

以上研究步骤得出的所有数据, 分别使用语料库卡方检测器和 SPSS 13.0 进行统计分析。

3. 结果与讨论

3.1. 读者/作者显现度

基于 Petch-Tyson (转引自文秋芳等, 2003 [3])读者/作者显现度的划分, 本研究主要从第一/二人称代词、情境指代词/小品词/模糊词两个主要参数进行分析。

表 1 所示, 学习者第一人称代词(单复数)的使用频率最为显著, 与此同时, 学习者在 1~4 年级阶段第一人称代词的使用呈非线性发展趋势, 书面语写作中存在显著口语化倾向, 该研究结果与文秋芳等(2003) [3]、王莉青(2009) [6]研究基本一致。学习者倾向于使用第一人称代词表达文章观点和立场, 议论文写作以客观陈述为主, 第一人称代词的过度使用突显出其论述具有很强的主观性。为深入考察第一/二人称代词情况, 运用 SPSS 软件采取对应分析方法进行分析, 结果表明受试在 1~4 年级阶段没有显著性差异($X^2 = 6.86, p = 0.334$), 说明受试在大学本科各阶段的书面语写作中均广泛使用第一/二人称代词。

Table 1. Relative frequency for the first/second person pronouns

表 1. 第一/二人称代词相对频率

特征维度 \ 年级	一	二	三	四	合计
第一人称代词(单数): I, me, my, mine	363.6/万词	161.9/万词	270.8/万词	201.7/万词	998.0/万词
第一人称代词(复数): We, us, our, ours	276.9/万词	206.0/万词	330.9/万词	343.6/万词	1157.4/万词
第二人称代词(单复数): You, your, yours	150.1/万词	119.1/万词	156.4/万词	91.8/万词	517.4/万词
合计	790.6/万词	487.0/万词	758.1/万词	637.1/万词	2672.8/万词

为具体探究第一/二人称代词使用情况, 对其各特征类型分别进行语料库卡方检测, 统计结果表明受试仅在第一人称代词单数类别存在显著差异($X^2_{1\sim 2 \text{ 年级}} = 60.19, p = 0.000, X^2_{2\sim 3 \text{ 年级}} = 23.45, p = 0.000; X^2_{3\sim 4 \text{ 年级}} = 7.79, p = 0.005$)。观察检索行, 发现第一人称单数中 *I* 的使用高达 69.67%, 用于作者向读者表达看法或作出论断、提出建议, 以期唤起读者思考, 如 “*I think...*”, “*I support...*”, “*I disagree...*”, “*I approve...*” 等。

观察表 2 可以发现, 学习者在 1~4 年级阶段写作中强调小品词使用频率最高, 学习者将口语交际中常用的 just/really 广泛用于书面英语写作中, 而情境指代词和模糊词的使用频率相对不高。运用 SPSS 软件采取对应分析方法进行分析, 结果表明受试在 1~4 年级阶段写作中情境指代词/小品词/模糊词不存在显著性差异($X^2 = 16.294, p = 0.061$)。为深入探究使用情况, 对相邻年级分别进行语料库卡方检测, 统计结果表明各特征类别在相邻年级中均不存在显著性差异($X^2_{1\sim 2 \text{ 年级}} = 0.34, p = 0.56; X^2_{2\sim 3 \text{ 年级}} = 0.14, p = 0.71; X^2_{3\sim 4 \text{ 年级}} = 0.81, p = 0.37$)。

Table 2. Relative frequency for situational pronouns/particles/fuzzy terms
表 2. 情境指代词/小品词/模糊词相对频率

特征维度 \ 年级	一	二	三	四	合计
阅读/写作情境指代词: Here, now	7.7/万词	2.4/万词	4.5/万词	8.3/万词	22.9/万词
强调小品词: Just, really	26.3/万词	36.9/万词	28.3/万词	25.0/万词	116.5/万词
模糊词: kind/sort of, and so on, etc.	10.2/万词	8.3/万词	3.4/万词	19.5/万词	41.4/万词
总计	44.2/万词	47.6/万词	36.2/万词	52.8/万词	180.8/万词

在读者/作者显现度总体特征方面,通过合计表 1 和表 2 可以发现,学习者在 1~4 年级阶段总体特征分别为 834.8/万词, 530.46/万词, 790.43/万词, 680.99/万词;学习者写作词汇中读者/作者显现度并没有随着语言水平的提高而降低,而是呈现出非线性发展趋势。对照文秋芳(2003) [3]研究:美国大学生读者/作者显现度总体特征为 119/万词,发现中国西部地区英语专业学习者在 1~4 年级阶段每万词的使用数量约高达美国大学生的 4~7 倍不等。

3.2. 词汇密度

图 1 显示,学习者在 1~4 年级阶段词汇密度的发展呈现波动性。为纵向考察相邻年级是否存在显著性,进行单因素方差分析。方差齐性检验结果($p = 0.640 > 0.05$)表明组间方差齐性,满足单因素方差检验要求条件。单因素方差检验结果显示英语专业学习者在 1~4 年级阶段写作词汇密度发展不显著($F = 1.372, p = 0.255 > 0.05$)。词汇密度是文本中信息量蕴含大小的指标,该结果表明随着学习者英语水平提高,文本信息量却没有明显增多,究其原因可能在于学习者在写作过程中缺乏透过现象的深入思考,也有可能是其词汇知识的发展无法满足思维表达的需要。

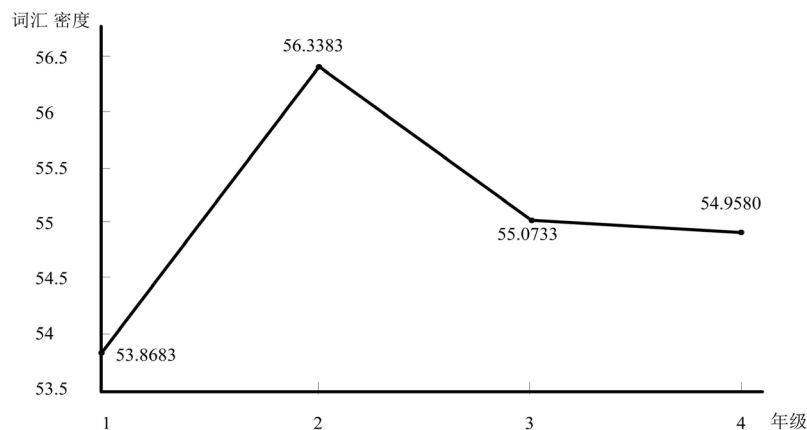


Figure 1. Mean development curve for lexical density

图 1. 词汇密度均值发展曲线

进一步观察图 1 可以发现,学习者在大学四年阶段写作词汇密度值均大于 50%。Waller (1993) [7]曾提出实词密度大于或等于 50%的作文可能被判定为母语写作,但结合本研究表 1 结果可知,词汇密度无法有效区分同一实词(如 I, we, you 等人称代词)在文本中的重复出现率,这样即便词汇密度大于 50%,

也无法说明文章涵盖较大的信息量, 该研究结果与朱慧敏和王俊菊(2013) [8]等一致: 词汇密度不能很好地反映二语写作中的词汇发展水平, 进一步说明词汇密度无法准确反映学习者写作口语化特征。

3.3. 词频

观察表 3 可以发现, 学习者在 1~4 年级阶段 1000 词频单词的使用比例分别为 63.87%~72.01%不等, 高频词的使用比例并没有随着其英语水平的提高而下降。为具体探究学习者词频使用情况, 运用 SPSS 软件采取对应分析方法, 结果表明受试在 1~4 年级阶段无显著性差异($X^2 = 2.374, p = 0.984$), 英语专业学习者在四年大学阶段英语写作词汇主要为 1000~2000 词频单词, 该结果与文秋芳等(2003) [3]、刘兴兵(2005) [9]等研究结果一致, 英语专业学生写作词汇过度使用中学阶段习得的常用高频词汇, 写作词汇口语化现象突出。课后交谈发现, 原因可能在于学习者存在畏难情绪, 恐出现单词拼写、单词搭配等词汇错误, 因此不敢自由使用大学阶段新学习学术词汇; 其次, 可能由于学习者大学阶段所接触的学术词汇依然停留在接受性阶段, 无法自由产出。

Table 3. Lexical frequency ratios

表 3. 各等级词频比率

单词级别	年级				
	一年级	二年级	三年级	四年级	
0~1,000	72.01%	63.87%	66.38%	64.92%	
1,000~2,000	13.00%	18.50%	17.96%	16.67%	
学术词汇	4.68%	4.36%	4.19%	4.65%	
其它	10.30%	13.26%	11.46%	13.76%	

4. 小结

本研究发现, 中国西部地区英语专业学生书面写作中的读者/作者显现度、词汇密度、各词频单词使用比率并没有随着英语总体水平的提高而有显著进步, 而是呈现出非线性发展趋势。学习者在英语议论文写作中频繁使用第一人称代词表达观点。词汇密度不能有效区分同一实词的重复出现率, 无法准确反映二语学习者写作口语化特征。学习者在写作中以常用词汇为主, 避免使用生僻单词。

基于以上发现, 本研究提出以下二语写作教学建议: 首先, 在英语专业写作课堂教学中, 教师可以将第一人称代词作为在英语写作中作者探求其身份构建的一种有效途径, 向读者渗透主体意识, 体现作者角色以及作者与读者之间的关系, 促进作者与读者之间的互动。其次, 教师应强调书面语体和口语语体的区别, 帮助学生树立语体意识, 避免写作中口语语体干扰, 以开展词汇竞赛等教学激励方式将学习者在大学阶段接触的学术词汇转化为产出性词汇, 并积极鼓励学生正确使用学术词汇。

当然, 本研究仅从写作词汇维度展开, 文体仅限于议论文, 未来研究可以从句法、语篇等视角进行多维度探究, 语料文本内容可选取如毕业论文、论文摘要等, 以期对二语写作口语化特征研究提供更为丰富、完整的结论。

基金项目

云南省教育厅科学研究基金项目“社会文化视阈下的英语写作概括能力探究”(编号 2018JS503), 昭通学院校级教改项目“地方本科院校英语专业学生写作现状及课程教学模式研究”(编号 Ztjx201814)阶段性成果。

参考文献

- [1] H. E. 帕麦尔, F. G. 布兰德弗. 英语口语语法[M]. 北京: 商务印书馆, 1964.
- [2] Ure, J.N. (1971) Lexical Density and Register Differentiation. In: Perren, G.E. and Trim, J.L.M., Eds., *Application of Linguistics: Selected Papers of the Second World Congress of Applied Linguistics*, Cambridge University Press, Cambridge.
- [3] 文秋芳, 丁言仁, 王文字. 中国大学生英语书面语中的口语化倾向——高水平学习者语料对比分析[J]. 外语教学与研究, 2003, 35(4): 268-274.
- [4] 吴瑾, 邹青. 中国学生英语口语语体特征研究: 词汇密度与词频[J]. 山东外语教学, 2009, 30(1): 8-13.
- [5] 梁茂成, 李文中, 许家金. 语料库应用教程[M]. 北京: 外语教学与研究出版社, 2010.
- [6] 王莉青. 学生在议论文中使用第一人称代词应用的研究[J]. 九江职业技术学院学报, 2009(1): 59-61.
- [7] Waller, T. (1993) Characteristics of Near Native Proficiency in Writing. In: Ringbom, H., Ed., *Near-Native Proficiency in English*, Abo Academy University, Abo, 183-294.
- [8] 朱慧敏, 王俊菊. 英语写作的词汇丰富性发展特征——一项基于自建语料库的纵贯研究[J]. 外语界, 2013(6): 77-86.
- [9] 刘兴兵. 中国大学英语学习者英语书面语中的口语化倾向[J]. 理工高教研究, 2005, 24(1): 105-112.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2330-1708, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: ml@hanspub.org