

大数据方法与高职英语写作教学创新研究

——以“2017年百万同题英文写作研究”为例

曾艳¹

[摘要] 本文在总结现有大数据方法应用于高职英语写作研究的基础上,以“2017年百万同题英文写作研究”大数据的分析与应用为例,探讨在大数据技术背景下,如何从教师角色、写作训练方式和教学方法及教学手段等多个方面,创新高职英语写作教学,提升教学效能和教学质量,以此促进高职生的英语语言综合应用能力的提高。

[关键词] 大数据; 英语同题写作; 高职英语写作; 教学创新

一、前言

近五年来,“大数据”一词越来越多地在各行各业被提及,人们用它来描述和定义信息爆炸时代产生的海量数据,并命名与之相关的技术发展与创新。它具有如下三个典型特征:一是整体数据,而不是随机采样;二是大体方向,而不是精确制导;三是相关关系,而不是因果关系。这就意味着大数据时代的到来必将带来现代教育的巨大变革,传统的教学方式、学习方式和评价方式将无法满足大数据时代人才培养的需求。

二、大数据方法应用于高职英语写作教学研究现状

随着现代教育技术的快速发展,互联网技术对外语教学的辅助作用日渐凸显。其中,作为典型代表的在线自动写作评分系统,国外的有:ETS应用于TOEFL、GMAT考试的e-rater写作评估系统,美国普林斯顿开发的WordNet;国内的有:北京词网科技研发的句酷批改网,高等教育出版社与清华大学合作开发的TRP数字化写作教学资源平台等,均频频走进广大师生的视野。

根据社会关注度的不同,笔者运用文献分析法进行文献阅读,对内容进行质性分析,对基于句酷批改网平台进行高职英语写作教学研究的现状进行了归纳整理,以中国学习期刊网(CNKI)教育类数据库为基础,为确保文献搜集的全面性和可操作性,研究采用“主题检索”方式,选取关键词为“大数据”、“高职英语写作教学”进行检索,但受CNKI数据采集选项功能、论文发表周期等的限制,仍会有文献被遗漏。自2015年1月至2017年8月共检索到48篇有效论文,其中23篇仅针对大数据时代高职英语写作教学改革给出了宏观层面的建议;杨继红的《基于大数据的批改网在高职英语写作教学中的应用》和侯芬的《在线写作自动评估系统在大学英语写作教学中的应用》基于句酷批改网平台进行了高职英语写作的实证研究;以百万同题英文写作大数据分析报告为依据来指引高职英语

¹ 曾艳(1982-),女,汉族,讲师,研究方向为英语教学法。

写作教学的研究还是空白，笔者以“2017 年百万同题英文写作研究”大数据的分析与应用为例，探讨在大数据技术背景下，创新高职英语写作教学的思路。

三、2017 百万同题写作活动

（一）活动概况

中国学生“百万同题写作”是中国高校英语写作联盟（由清华大学、北京大学、复旦大学和中山大学于 2013 年联合成立）开展的一项重要活动。活动由写作联盟和北京语言智能协同研究院主办，句酷批改网具体组织实施，此类活动的连续推进，旨在构建与挖掘中国学生英语写作大数据，协同创新高效可行的信息化英语教学模式，切实提升中国学生的英文写作能力。自 2014 年 4 月首届以“Let Classics Be Classics”为题的同题英文写作活动举办至今，已成功举办 4 年。以“Are We More Connected or More Alone”为题的第四届英文同题写作活动于 2017 年 4 月 10 日拉开序幕，截止于 5 月 31 日，来自全国各地 4,139 所学校的师生参与其中，教师引用布置题目 17,158 次，收到学生作文共计 1,408,626 篇，学段覆盖大学、高职、高中和初中。

（二）参与学生地域分布情况

活动得到来自全国 32 个省市地区的师生积极响应，其中参与人数最多的地区分别为四川省（140,840 人），广东省（112,455 人）以及山东省（107,301 人）。

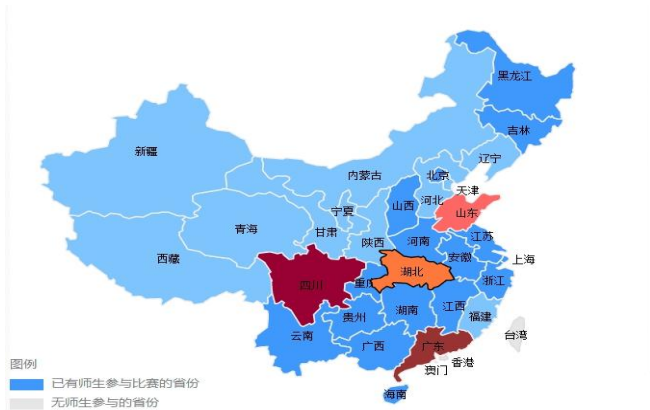


图 1 全国参与省市分布图

（三）学段分布

本次活动参与者覆盖学段包括大学本科、高职、高中及初中，各学段学生同写一道题目。其中，本科学生参与人数最多，占总参与人数的 89.61%，高职生其次，占总参与人数 7.05%。

（四）自主学习行为数据对比

1. 修改行为表现



表 1 同题写作活动学生修改作文次数分布表

统计结果显示，每篇作文的平均修改次数为 7.97 次，约 75% 的学生都对自己的文章进行过多次修正，超过 10% 的学生修改作文 20 次以上，学生整体修改表现良好。

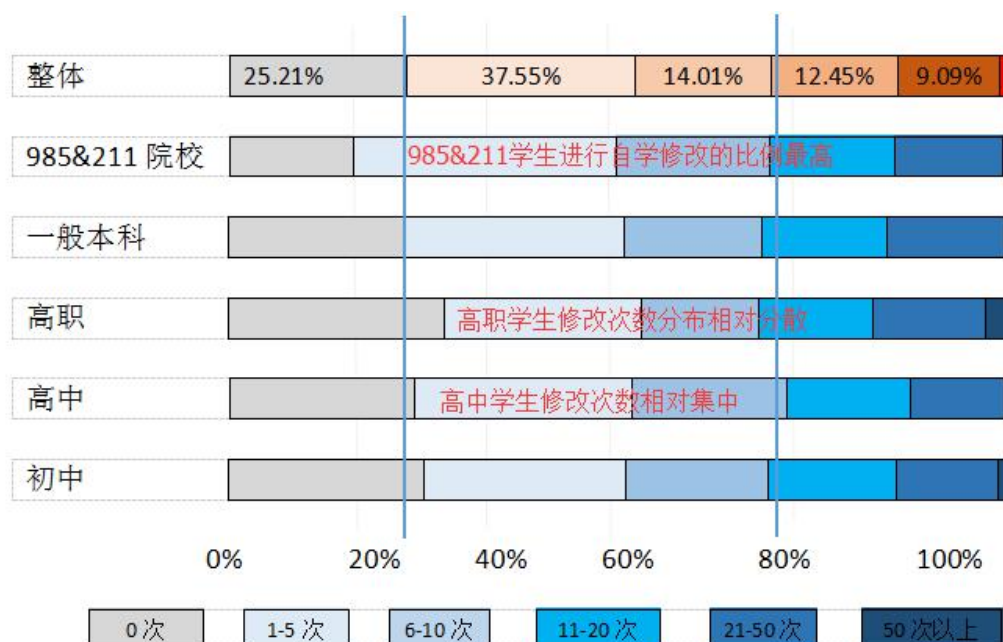


表 2 同题写作活动的学生修改作文次数院校分布表

整体上看，约 75% 的学生都对自己的作文进行了修正，学生自学能动性普遍较好。其中，985&211 院校学生的自主修改比例最高，80.2% 的学生都在写作搓成中进行了自主修改，自学能动性相对较强；高职学生未修改及修改 20 次以上区间内占比均最高，修改次数分布相对分散，后面修改过程分析环节的数据变化能清楚地阐释这一现象；高中学生修改 10 次以上的学生占比最小，学生修改次数普遍相对集中。

2. 分数变化

表 3 同题写作活动中学生修改作文分数变化

类别	初版分数（平均值）	终版分数（平均值）	分数增长
整体	74.59	79.07	4.48

985&211 院校	76.28	81.07	4.79
一般本科	74.89	79.26	4.37
高职	72.31	76.48	4.17
高中	70.08	73.52	3.44
初中	63.74	67.03	3.29

通过上图可以看到学生作文在修改过程中有较为显著的分数的提升。通过平均每人 7.97 次的修改，学生作文分数在整体上由初版作文均分 74.59 分上升到终版作文均分 79.07 分，分数提升 4.48 分。在各类院校中，985&211 院校学生在修改中分数提升幅度最大，一般本科次之，高职在高校系列中排最后，这与学生的整体英语水平差异事实吻合。

3. 错误修正情况

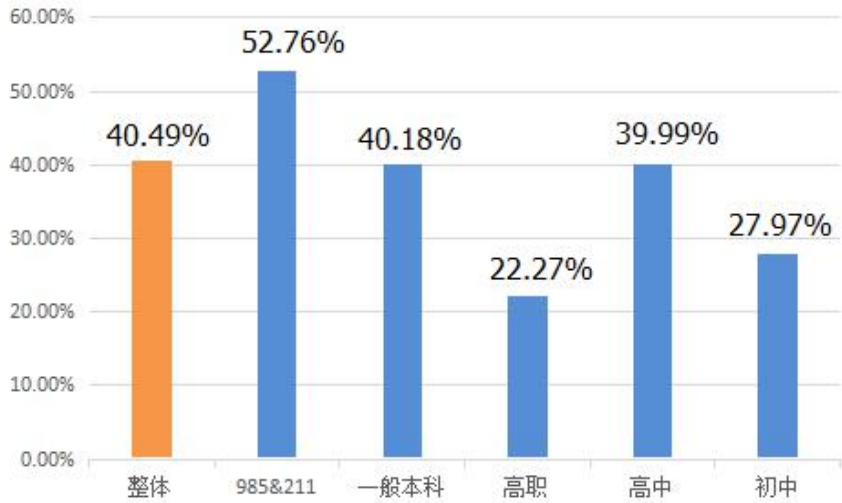


表 4 各类院校学生作文语法错误修正率 (不计拼写错误)

表 4 为各类院校学生作文语法错误修改情况。学生在自主修改过程中，累计修改错误超过 40%，985&211 院校语法错误修正率 52.76% 位居最高，一般本科与高中修正率分别为 40.18% 和 39.99%，初中修正率为 27.97%，高职院校语法错误修正率为 22.27%，位于末位。这说明高职生在自主学习状态下，语法的自我修正能力较弱。

4. 修改过程分析

以下图表内容以五类学校各抽取 10,000 篇作文为分析样本，对学生作文错误修改过程的推进情况进行了详细分析：

表 5 第一次修改的部分高频错误修正率

第一次修改后	985&211	一般本科	高职	高中	初中
拼写	21.22%	14.06%	7.11%	16.03%	9.57%
词性误用	10.48%	6.57%	-0.11%	7.24%	4.81%
动词	11.62%	6.03%	2.23%	12.33%	7.89%
冠词	13.36%	8.95%	1.10%	12.42%	9.92%
名词	14.93%	8.80%	1.26%	12.90%	12.23%
介词	15.83%	9.98%	2.87%	11.99%	10.14%
搭配	11.46%	5.16%	-1.24%	9.66%	5.42%
主谓一致	15.53%	9.74%	1.70%	13.69%	9.76%

成分冗余	15.33%	10.93%	1.41%	12.48%	8.92%
成分缺失	8.52%	3.02%	-1.90%	8.36%	0.91%
整体（除标点错误）	14.88%	9.01%	2.73%	12.14%	7.03%

通过表 5 可以看到，学生在进行第一次修改时，较多地对修改难度小或更为敏感的错误进行修正。除初中生外，其他各类院校学生作文第一次修改比例最小的错误均为修改难度最小的拼写类错误，从与其他类型错误的纠错情况对比来看，一般本科与高职院校最为突出。高职学生在第一次修改时，个别错误频次有不降反升的情况发生，另外，修正率也较低，说明修改的思路不够清晰合理。从错误难度上看，句子结构相关的错误是学生普遍感到修改难度较大的错误类别，成分缺失错误在第一次修改时修正率普遍较低，与其他错误类型相距较大。

表 6 修改过程前后段作文数据值变化

修改过程前后段 数据变化	985&211	一般本科	高职	高中	初中
拼写	28.78%	19.87%	15.32%	24.27%	15.16%
词性误用	10.24%	2.96%	-3.20%	8.25%	-6.84%
动词	15.11%	4.55%	-2.38%	11.56%	13.16%
冠词	14.64%	7.25%	1.23%	19.29%	12.72%
名词	15.00%	8.66%	4.25%	17.03%	24.82%
介词	17.29%	11.03%	7.26%	14.97%	11.52%
搭配	10.00%	1.10%	-8.95%	11.97%	-0.28%
主谓一致	18.60%	11.27%	0.74%	19.61%	8.15%
成分冗余	19.59%	12.61%	1.13%	16.42%	13.38%
成分缺失	1.07%	-5.26%	-8.24%	5.22%	-1.11%
整体（除标点错误）	16.24%	8.63%	3.71%	15.98%	7.04%

由表 6 可以看出，学生作文错误的修正大多集中在前半段修改过程中，特别是拼写类错误，修改过程前半段的修正率均高出后半段 15%以上。部分院校对于成分缺失、词性误用、搭配错误的修正更多地集中于修改过程的后半部分，这几类错误修改难度相对略大。

985&211 院校以及高中学生对于错误的修正较明显地集中于前半段，语法自我修正的进度较快，而高职院校及初中对于部分高频错误的修正相对集中于后半段，语法自我修正的进度相对慢一些。

综上所述我们可以得知，学生在自学过程的前半段进行的修改相对更为高效。学生在修改过程前期主要进行基础语法错误修正；在修改过程中后期，会对修改难度相对大的句子结构类错误给予更多关注和修正；对词汇与举行的调整会在整个自学过程中循序渐进地进行。

四、大数据方法与高职英语写作教学创新

从整个大数据分析报告可以看出，在线写作平台对于促进高职英语写作教学

是有一定作用的，但是高职生语法基础差，自我修正能力弱，需要设计大数据与高职生英语写作能力相结合的教学方案，改变了原有的英语写作教学理念，教学模式，教学流程。

（一）教师角色创新

学生自学能动性较低，错误自我修正能力较低，这就决定了教师既要做一名写作教学的主导者，做好学生英语写作水平提高过程中的向导，引领写作策略，也要根据学生学情进行适当的微观干预，引发自主性的前提下，少走弯路。

（二）写作训练方式创新

写作训练不一定是作文训练，一篇好的作文必须是思路与表达方式的完美展现。相对于篇章，句子和词块是较小的元素，可以先从词块的积累、句子的仿写抓起，进行一句多写，举一反三；另外，写作教学一定要加强美文赏析教学，多让学生体会语言的美感与用词的细腻。

（三）教学方法及教学手段创新

高职生的写作能力特点决定了写作教学要讲求教学的时效性，教师要擅用网络写作平台的智慧课堂等功能加强对写作过程的技术监管，辅助学生的写作薄弱环节，提高自主写作能力。

参考文献：

- [1]王嫣女. 计算机自动评估系统辅助课堂写作教学的研究述评[J]. 海外英语, 2012, (10):36-37.
- [2]黄文星. 基于教育大数据时代的高职高专英语教学改革研究[J]. 当代外语研究, 2015, (7):53-57.
- [3]侯芬. 在线写作自动评估系统在大学英语写作教学中的应用[J]. 当代外语研究, 2015, (16):24-27.