

[DOI]10.16164/j.cnki.22-1062/c.2018.02.020

反馈方式对写作修改过程及质量的影响研究

白云,王俊菊

(山东大学外国语学院,山东 济南 250100)

[摘要] 探讨不同反馈方式对大学生英语写作修改的作用,重点对比教师反馈和系统反馈对修改过程和文本质量的影响具有重要价值。通过对教师反馈、自动评分系统反馈、修改记录和访谈的分析,发现基于两种反馈的修改过程在多维度上存在显著差异。其中,教师反馈引发更多的内容修改、整体修改、增加和自我修改;系统反馈导致较多的校订修改、单词及以下的局部修改和替代。对修改后的文本质量,两种反馈均能提高词汇复杂性,而教师反馈对提高语言流利性亦有显著作用。据此,在写作教学中,运用不同的策略将两种反馈方式加以融合,提升文本质量和写作水平。

[关键词] 教师反馈;系统反馈;修改过程;文本质量

[中图分类号] H315

[文献标志码] A

[文章编号] 1001-6201(2018)02-0134-06

一、引言

反馈是英语写作教学中至关重要的环节,也是写作研究中的持续热点,传统上以教师反馈为主要形式。近年来随着自动作文评分技术的开发与完善,基于网络的系统反馈在教学中得到越来越广泛的应用。而学生根据不同形式的反馈所做的修改,与提高作文质量和学生写作水平密切相关。

20世纪90年代以来,教师反馈对修改的影响成为研究热点,研究大致分为三类:一是探讨书面反馈的特征和范围,以证实教师反馈的有效性和可行性;二是关注学生的修改类型、修改策略和修改质量,旨在发现反馈和写作水平的关系;三是考查学生对教师反馈的态度和感受,分析学习者对各类反馈的接受程度。其中王俊菊^{[1]24-30}和杨苗^{[2]293-301}发现,在重视教师权威的教学环境中,学生更愿意接受教师反馈来修改作文,极少就反馈内容与教师进行互动。

有大量研究关注了评分系统反馈在二语写作课堂中的应用,主要集中在三方面:一是关注

计算机系统评分与人工评分的一致性,从而证实其作为测试工具的效度和信度^{[3]66-78};二是关注系统反馈对学生修改的影响,指出可使学生关注更大的写作模块,以进行大规模修改^{[4]217-235};三是研究自动反馈对作文质量的影响^{[5]51-65}。国内研究探讨评分系统对学生写作能力产生的影响^{[6]53-59},最新研究聚焦系统反馈对修改的影响^{[7]45-49}^{[8]81-88},分别关注在线修改的行为、过程和成果,对新的反馈方式下的修改行为和信念进行探讨。

亦有学者开展了教师反馈和系统反馈的对比研究^{[9]1-17},发现教师能提供更多高质量的反馈,而系统反馈有冗长多余之嫌。周一书^{[10]87-96}认为教师反馈能明显增强学生的篇章布局能力,而同伴反馈和网络反馈能提高学生的语言表达能力和写作兴趣。周丽^{[11]26-29}则发现系统反馈对学生的句法能力提高有显著作用。

不难发现,前人研究多是比较教师反馈和系统反馈的性质、类别及学生的接受程度,较少涉及不同反馈方式对二语学习者修改过程的影响。其次,大多数研究关注教师的书面反馈,缺少针

[收稿日期] 2017-08-15

[基金项目] 山东省社科规划基金项目(11DWZJ05,11CWZJ33)。

[作者简介] 白云,女,山东大学外国语学院讲师;王俊菊,女,山东大学外国语学院教授,博士生导师。

对以计算机和网络为介质的教师在线反馈的研究。鉴于此,本研究旨在对比教师在线反馈和系统即时反馈对英语写作修改过程和修改成果的影响,主要探讨如下问题:

(1)教师反馈和系统反馈对学生的作文修改过程中的修改目标、修改范围、修改行为、修改动机等方面有何影响?

(2)两种反馈方式对学生作文的文本质量有何影响?

二、研究方法

本研究的受试者是一所综合性大学的大一学生,其英语水平相当,入学时英语分级考试的总分及作文得分均相近。研究者所教授的两个平行班分别接受教师反馈和系统反馈,从每班各选30人,作为教师反馈组和系统反馈组。

研究数据分别来自教师反馈、系统反馈、修改记录和访谈。教师反馈是教师针对学生网上提交的作文所做出的在线反馈;系统反馈是批改网通过计算学生作文和标准语料库之间的差异,即时生成的作文得分和每句话的修改建议;修改记录包括学生每一稿所做的具体改动和作文的文本质量要素;而访谈为半结构式的,内容包括学生对每一稿具体修改过程的反思,修改时的心理活动及对不同反馈方式的态度及认识。

本研究选用一次写作练习作为研究素材。要求学生就目前年轻人大多依赖父母来照看小孩这一社会现象发表自己的见解,并阐明将来自己的做法。长度要求为120—180词,作文题目自拟。

两组作文的布置、提交、反馈和修改均在网上进行。针对系统反馈组,研究者先在网上布置好作文题目,只规定两节课后是提交的最后期限而不限定学生的提交次数。学生将初稿提交到评分系统后,系统会即时给出分数和针对每句话的修改建议。学生可根据反馈对初稿修改后再提交,可多稿修改,直到对自己的作文满意为止。对于教师反馈组,教师使用批改网的教师批改功能,对学生提交的作文进行在线批改,提供对单词拼写、语句表达、衔接连贯和篇章结构等的修改建议,并对整篇文章给出评语,之后学生进行在线修改并提交。评分系统可保存学生所有稿次的作文和每一稿的修改记录。点击除第一稿外的所有作文,会以不同颜色显示出添加、删除和替换的内容。

在修改完成后,研究者对学生分别进行访谈。参照每一稿的修改记录,询问每处修改的心理活动以及对反馈的接受和感知。每位学生的访谈时长为10到30分钟。为方便分析,研究者对所有访谈进行了全程录音。

本研究的数据分析包括对修改进行标记、衡量文本质量、用社科统计软件包SPSS处理数据和对访谈录音进行转写并分析四部分。

研究者对学生所做的每一次修改进行标记。本研究使用多维度的分类框架,参照Stevenson, et al.^{[12]201-233}的研究,将在线修改按修改目标、修改范围、修改行为和修改动机等四个维度进行分类(见表1)。

表1 在线修改的多维度分类体系

维度	类别
修改目标	内容修改、语言修改、校订修改
修改范围	单词及以下、短语、句子、篇章
修改行为	增加、删除、替代、移位、其他
修改动机	反馈引发、自我修改

修改目标包括内容修改、语言修改和校订修改(typing revisions)。其中的校订修改是在线修改所特有的类型,如当写作者本来掌握某个单词的拼写,却无意中击错了键,随后对之进行的修改就属于校订修改。校订修改与拼写修改容易混淆,辨别标准为校订错误是失误,而拼写错误是错误。修改范围分为单词及以下、短语、句子(如句子的时态、语态)和篇章(如句群、段落)等层面。修改行为为维度,参照Faigley & Witte^{[13]95-108}的分类体系,共有增加、删除、替代、移位和其他五类。修改动机分为反馈引发和自我修改。具体标记时,除校订修改不做修改行为的标记外,每次修改均都在各维度上进行标记。

此外,为消除不同作文长度对修改数据的影响,对以上四个维度的修改次数进行标准化,计算出修改频次。采用Stevenson, et al.^{[12]201-233}的计算方法,即修改频次指每一百个单词的平均修改次数,为总修改次数除以终稿的总单词数再乘以一百。

回答第二个研究问题,传统的二语写作研究一般以人工评定作文成绩测量产出。随着自动文本测量工具的出现,越来越多研究使用语言特征分析写作产出^{[14]102-114}。而文本质量一般应涉及语言、内容和篇章结构三个方面,对其语言质量的评价往往从流利性、准确性和复杂性三方面入手^{[15]5}。本研究选取的语言特征涉及以上三大

方面,包括总词数、超高频词汇占比、次高频词汇占比、学术词汇占比、超纲词汇占比、动词数、动词短语、连词数、平均词长、平均词长方差、平均句长、词汇丰富度、句子数和从句数等 14 项。其中,参照 Laufer & Nation^{[16]307-322} 的研究,依据词频的高低将单词分为 3 个等级,最常见的 1 000 词构成基础词汇 baseword1,本研究称为“超高频词汇”;次常见的 1 000 词构成基础词汇 baseword2,称为“次高频词汇”;再次常见的 570 词构成基础词汇 baseword3,称为“学术词汇”。另外,词汇丰富度用来表示文本中词汇的使用密度,计算公式为 $c = \text{type} / \sqrt{2 * \text{num}}$,其中 c 为词汇丰富度, type 为类符数(不重复计算的形符数), num 为形符数^{[7]45-49}。

对统计出的各维度修改频次和文本质量数据,用 SPSS17.0 进行处理,主要采用独立样本 t 检验和配对样本 t 检验两种方法。

分析访谈数据是采用 Yin^{[17]176-232} 的质性数据分析方案,包括数据的编辑、分解、再汇编、解析、得出结论等步骤,最终归纳出学生修改策略的特点。

三、结果与讨论

数据分析显示,教师在线反馈和系统即时反馈对学生的修改过程和修改质量都产生了不同程度的影响,具体涉及修改目标、修改范围、修改行为、修改动机和文本质量等方面。

(一)修改目标

表 2 修改目标的独立样本 t 检验结果

	教师反馈组 ($n=30$)		系统反馈组 ($n=30$)		MD	t	P
	M	SD	M	SD			
内容修改	0.94	0.84	0.00	0.00	0.940	3.530	0.006*
语言修改	4.06	1.47	5.06	1.19	-0.996	-1.668	0.113
校订修改	0.45	0.76	1.62	1.43	-1.169	-2.280	0.039*

注: * 在 0.05 水平上有显著意义,下同。

(二)修改范围

两种反馈方式在修改范围维度上也存在明显差异(见表 3)。首先在单词及以下层面的修改上有显著差异($t = -2.589, P < 0.05$)。教师反馈组的单词及以下修改显著低于系统反馈组($MD = -2.470$)。这一结果与钟彩顺^{[8]81-88}的发现类似,在基于自动反馈的修改中,“学习者有效修改的幅度很小。学习者大多只会关注单词拼

写、措辞和简单语法等层面,无法深入语篇和内容。”这也支持了赵蔚彬^{[19]2}的研究结论,在线写作中的“修改决策大多发生在个别单词内部,少有对词组和句子的改动”。鉴于目前的自动评分系统对语法层面的反馈较多,加之大量的对键入错误的反馈,不难理解系统反馈组的小范围修改显著增多。

独立样本 t 检验结果显示,两组在内容修改上存在显著差异($t = 3.530, P < 0.05$)。教师反馈组的内容修改显著高于系统反馈组($MD = 0.940$)(见表 2)。这主要是由两种反馈的特点所决定,评分系统“在内容、衔接、篇章和结构方面仅有总体评价,没有针对性地提供详细的反馈和修改建议”^{[18]43},导致该组受试未做任何内容修改($M = 0.00$);而教师反馈更注重作文内容的补充和完善。两组在语言修改上未显示出显著差异,说明两种反馈引发的语言层面的修改基本相当。而在校订修改上有显著差异($t = -2.280, P < 0.05$)。教师反馈组的校订修改显著低于系统反馈组($MD = -1.169$)。Stevenson, et al.^{[12]201-233} 的研究发现,在计算机写作过程中产生大量键入错误,校订修改占总修改的比例接近 60%。本研究访谈中,学生称键入错误为“手残”打错的词,如将 their 打成 thire,将 because 打成 becasue。由于学生熟悉纸笔写作而缺乏在线写作的经验,导致即使会拼写的单词也常出现键入错误。虽然两组受试均大量出现键入错误,但系统自动反馈对空格、缩进等格式和单词拼写的要求精准、反馈迅速,且做出相应修改后评分会有提高,学生都会依据反馈进行校订修改;而教师因需关注语言、内容和结构等各方面,实际操作中又常有时间的限制,很难对所有的键入错误做出反馈,因此教师反馈组的校订修改明显少于系统反馈组。

表3 修改范围的独立样本 *t* 检验结果

	教师反馈组		系统反馈组		MD	<i>t</i>	P
	M	SD	M	SD			
单词及以下	2.26	1.29	4.73	2.73	-2.470	-2.589	0.019*
短语	1.70	0.96	1.71	0.91	-0.001	-0.001	0.999
句子	1.17	0.65	0.24	0.26	0.927	4.198	0.001*
篇章	0.42	0.57	0.00	0.00	0.419	2.331	0.045*

其次两组在短语层面的修改无显著差异,而对句子($t=4.198, P<0.05$)和篇章($t=2.331, P<0.05$)的修改有显著差异。教师反馈组中较大范围的修改显著高于系统反馈组,分别为句子修改(MD=0.927)和篇章修改(MD=0.419)。这可能是因为教师反馈更多地涉及内容完整,谋篇布局和语篇连贯等方面,学生就会尝试在这些方面进行修改。

(三)修改行为

两组的“增加”行为有显著差异($t=4.777,$

$P<0.05$)。教师反馈组的增加显著多于系统反馈组(MD=1.290)(见表4)。如前所述,教师反馈组的内容修改和大范围修改较多,导致增加行为相应增多。两组的替代也有显著差异($t=-2.762, P<0.05$)。教师反馈组的替代显著少于系统反馈组(MD=-1.722)。且系统反馈组的替代,因不涉及内容修改,所以均为不改变意义的替代。这一结果支持了 Stevenson, et al.^{[12]201-233}的结论,在母语和外语的修改中,替代均为最常用的修改动作。

表4 修改行为的独立样本 *t* 检验结果

	教师反馈组		系统反馈组		MD	<i>t</i>	P
	M	SD	M	SD			
增加	1.65	0.75	0.36	0.41	1.290	4.777	0.000*
删除	0.25	0.43	0.14	0.30	0.115	0.699	0.493
替代	2.87	1.32	4.59	1.47	-1.722	-2.762	0.013*
移位	0.30	0.37	0.04	0.14	0.261	2.106	0.058
其他	0.13	0.28	0.06	0.19	0.066	0.615	0.546

(四)修改动机

两组中反馈引发的修改无统计学上的差异,而自我修改存在显著差异($t=2.555, P<0.05$)。教师反馈组的自我修改显著多于系统反馈组(MD=0.444)(见表5)。原因可能是教师反馈比系统反馈更具有启发性,更能引导学生主动发现问题并解决问题;其次,学生受益于系统即时反

馈的便捷性,每次修改的范围较小,短时间内多次提交,以验证自己修改是否正确;而基于教师反馈的修改,因为人工批改在时间上相对滞后,时效性不好,学生对作文修改后常会通读全文,发现新问题会自发进行修改。可见教师反馈更能培养学生自发修改、自主学习的意识。

表5 修改动机的独立样本 *t* 检验结果

	教师反馈组		系统反馈组		MD	<i>t</i>	P
	M	SD	M	SD			
反馈引发	5.01	1.60	6.67	2.36	-1.669	-1.850	0.081
自我修改	0.44	0.55	0.00	0.00	0.444	2.555	0.031*

(五)文本质量

反馈方式对修改成果的影响具体体现在作文文本质量的变化上。本研究首先使用独立样本 *t* 检验,比较两组终稿的文本质量,发现各项指标均无显著差异。这表明虽然介入不同形式的反馈,学生修改后的终稿的文本质量并无显著差异。

然后使用配对样本 *t* 检验,对两组各自的初稿和终稿的文本质量进行比较,考察在修改前后是否有显著性差异;如有差异,在哪些质量要素上存在差异。如表6所示,通过比较教师反馈组的初稿和终稿,发现14项质量要素中有多项呈现明显差异。结果显示,文本质量在总词数、动词数、动词短语和词汇丰富度等要素上存在显著差异。

表 6 教师反馈组初终稿的文本质量配对样本 *t* 检验结果

	初稿		终稿		MD	<i>t</i>	<i>P</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
总词数	209.10	59.64	233.20	52.84	-24.100	-3.510	0.007*
动词数	44.60	13.95	48.80	11.14	-4.200	-2.294	0.047*
动词短语	18.90	6.42	21.40	6.33	-2.500	-2.466	0.036*
词汇丰富度	5.57	0.51	5.74	0.46	-0.171	-2.429	0.038*

其中,初稿和终稿的总词数有显著差异($t = -3.510, P < 0.05$)。初稿总词数显著低于终稿词数($MD = -24.100$)。这表明教师反馈组的终稿长度明显增加,而文本长度是检测作文流利性的重要指标,这就说明基于教师反馈的修改使作文流利性有所增加。两稿的动词数和动词短语也有显著差异,分别为($t = -2.294, P < 0.05$)和($t = -2.466, P < 0.05$)。词汇丰富度亦有显著

差异($t = -2.429, P < 0.05$),初稿显著低于终稿($MD = -0.171$)。词汇丰富度是衡量词汇复杂性的一个重要指标,这表明终稿的词汇复杂性显著提高。

此外,系统反馈组的初稿和终稿的文本质量在总词数、学术词汇占比和平均词长等方面呈现显著差异(见表 7)。

表 7 系统反馈组初终稿的文本质量配对样本 *t* 检验结果

	初稿		终稿		MD	<i>t</i>	<i>P</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
总词数	233.40	58.31	230.50	56.99	2.900	4.106	0.003*
学术词汇占比	0.05	0.02	0.05	0.02	-0.005	-3.015	0.015*
平均词长	4.48	0.30	4.51	0.29	-0.030	-3.360	0.008*

初稿和终稿的总词数有显著差异($t = 4.106, P < 0.05$),初稿总词数多于终稿词数($MD = 2.900$)。系统反馈组的终稿长度不升反降,这可能是因为本组的修改集中于单词及以下,而在反复修改不成功时,学生常采用回避策略,删除修改不成功的表达以提高分数。

学术词汇占比也有显著差异($t = -3.015, P < 0.05$),初稿的学术词汇占比显著低于终稿($MD = -0.005$)。访谈发现,学生修改时倾向于用大词、长词来替代常用词汇,如用 atavistic affection 替代 every generation is closer 来表示“隔代亲”,认为这样作文显得高大上,更容易得高分。这不仅导致学术词汇占比的显著增加,还使初稿的平均词长显著低于终稿($MD = -0.030$),说明终稿使用的长词明显增多。而平均词长也是词汇复杂性的一个指标,表明系统反馈引发的修改对提高词汇复杂性亦有帮助。

以上结果与胡学文^{[7]45-49}的结论不尽相同。他发现基于系统反馈的修改,使作文长度、低频词汇使用、词汇丰富度、平均词长、从句数量等指标均有显著提高,但对平均句长不产生明显影响。周丽^{[11]26-29}则发现系统反馈后的修改对句法流利度和复杂度有提高,而对其他句法能力的

影响不显著。上述研究结果间的差异可能是由于实验设计不同造成,时间跨度不一致、语言产出评价标准不统一^{[20]161-184}等都会导致结论缺乏可比性,需要更统一、更深入的研究。

四、结 论

本研究通过对比教师反馈组和系统反馈组的在线修改过程和成果,发现两种反馈方式对修改过程的影响体现在不同维度,对文本质量的影响也不尽相同。主要包括:教师反馈引发更多的内容修改,系统反馈导致较多的校订修改,而两种反馈在语言修改上并无显著差异;教师反馈带来更多的整体修改,而系统反馈更集中于单词及以下层面的局部修改;教师反馈导致更多的增加,而系统反馈组使用更多的替代;教师反馈引发更多的自我修改。此外,修改后的文本质量,教师反馈组的终稿在总词数、动词数、动词短语和词汇丰富度等要素上有显著提高,而系统反馈组在学术词汇占比和平均词长等方面大幅增长。

上述发现表明,如何把教师反馈和系统反馈进行融合以达到最佳的反馈效果,是值得进一步探索与实践的课题。首先,在反馈时机上,可尝试先教师、后系统的反馈方式。教师反馈会带来

较大范围、注重内容的修改,也更能培养学生自主修改的意识,之后的系统反馈则帮助学生进行语言层面的小范围修改。其次,可采用系统全批、教师精批精讲的模式。在师生比例过大的情况下,利用自动评分系统即时反馈的特点,对每位学生的每一稿做出即时、公正的反馈;教师则可以选取有代表性的作文,进行精批和精讲,提高作文批改和讲评的效率。最后,充分利用评分系统提供的资源来促进写作教学,比如利用系统提供的作文相似度检测,可有效防止学生的抄袭行为;利用共性分析的功能,可找出学生普遍存在的问题进行讲评和练习;利用每篇作文的体检报告,可提出针对性的修改建议。

本研究只考察了两种反馈方式对一篇作文修改的影响,未来可对多篇作文的修改过程进行历时的研究;亦可增加受试样本,进行更大规模的系统研究。

[参考文献]

- [1] 王俊菊. 总体态度、反馈类型和纠错种类——对大学英语教师作文书面反馈的探讨[J]. 国外外语教学, 2006(3).
- [2] 杨苗. 中国英语写作课教师反馈和同侪反馈对比研究[J]. 现代外语, 2006(3).
- [3] Li, Z., Link, S., Ma, H., Yang, H. & Hegelheimer, V. The role of automated writing evaluation holistic scores in the ESL classroom [J]. *System*, 2014(44).
- [4] Tuzi, F. The impact of e-feedback on the revisions of L2 writers in an academic writing course [J]. *Computers and Composition*, 2004(21).
- [5] Stevenson, M. & Phakiti, A. The effects of computer-generated feedback on the quality of writing [J]. *Assessing Writing*, 2014(19).
- [6] 唐锦兰, 吴一安. 写作自动评价系统在大学英语教学中

- 的应用研究[J]. 外语与外语教学, 2012(4).
- [7] 胡学文. 在线作文自我修改对大学生英语写作结果的影响[J]. 外语电化教学, 2015(5).
- [8] 钟彩顺. 基于在线自动反馈的自主写作修改研究[J]. 解放军外国语学院学报, 2015(4).
- [9] Dikli, S. & Bleyer, S. Automated essay scoring feedback for second language writers: How does it compare to instructor feedback [J]. *Assessing Writing*, 2014(22).
- [10] 周一书. 大学英语写作反馈方式的对比研究[J]. 外语界, 2013(3).
- [11] 周丽. 英语在线写作平台对大学生写作句法能力的影响[J]. 外语电化教学, 2015(9).
- [12] Stevenson, M., Schoonen, R. & de Glopper, K. Revising in two languages: A multi-dimensional comparison of online writing revisions in L1 and FL [J]. *Journal of Second Language Writing*, 2006(15).
- [13] Faigley, L. & Witte, S. Measuring the effects of revisions on text structure [A]. In R. Beach & L. S. Bridwell (eds.). *New directions in composition research* [C]. New York: Guilford Press, 1984.
- [14] 刘兵, 王奕凯, Zhang, J. L. 任务类型对在线英语写作任务准备和产出的影响[J]. 现代外语, 2017(1).
- [15] 梁茂成. 大规模考试英语作文自动评分系统的研制 [M]. 北京: 高等教育出版社, 2012.
- [16] Laufer, B. & Nation, I. S. P. Vocabulary Size and Use: Lexical Richness in L2 Written Production [J]. *Applied Linguistics*, 1995(16).
- [17] Yin, R. K. *Qualitative Research from Start to Finish* [M]. New York: The Guilford Press, 2011.
- [18] 张荔, 盛越. 自动作文评阅系统反馈效果个案研究[J]. 外语电化教学, 2015(5).
- [19] 赵蔚彬. 英语写作修改模型 [M]. 开封: 河南大学出版社, 2007.
- [20] Ferris, D. R. & Roberts, B. Error feedback in L2 writing classes: How explicit does it need to be [J]. *Journal of Second Language Writing*, 2001(10).

Effects of Feedback Modes on EFL Students' Revision Process and Text Quality

BAI Yun, WANG Jun-ju

(School of Foreign Languages and Literature, Shandong University, Jinan 250100, China)

Abstract: This paper reports a comparative study of the effects of teacher feedback and AES (Automated Essay Scoring) feedback on revision process and revised product in a university setting, by focusing on revision orientation, revision domain, revision action, revision initiation and the quality of texts. The results indicate that revising processes initiated by two modes of feedback demonstrate striking differences in a multi-dimensional revision taxonomy. Teacher feedback brings more content revisions, large-scale revisions, addition and self-revisions, while AES feedback results in more typographic revisions, small-scale revisions and substitutions. And the effects on text quality are also different, that is, teacher feedback can improve language fluency and both can enhance vocabulary complexity.

Key words: Teacher Feedback; AES Feedback; Revising Process; Text Quality

[责任编辑: 哲文]