

基于动态评估的来华留学生句子改写教学干预研究

陈香 何漪叶 李友敏

中央财经大学, 北京市, 100101;

摘要:为解决来华留学研究生在学术写作中普遍存在的句子层面表达瓶颈,本研究旨在检验以“动态评估”为核心的教学干预,对其句子改写能力的提升成效。研究对42名受试者进行了为期三周、结合对话式反馈与同伴互评的动态评估教学干预,并采用单一组前后测设计及配对样本t检验进行分析。结果显示,干预后学生的句子改写总分得到了统计上的极度显著提升,且在“内容准确性”“表达连贯性”“语言正确性”与“学术规范性”四个维度上均有显著进步,其中以“学术规范性”与“内容准确性”的进步幅度最为突出。本研究结论表明,动态评估是提升学习者学术句子改写能力的有效路径,其教学设计与实证结果可为高级汉语学术写作的教学实践提供有价值的参考。

关键词: 动态评估; 句子改写; 学术写作; 来华留学生

DOI: 10.64216/3080-1494.25.03.012

引言

学术写作是一种不间断的再写作过程,在此过程中,句子改写是写作中至关重要的环节(Ito等,2020)。句子改写是指用自己的话来改写其他人表达的观点,同时还需保持原义(Ardelia&Tiyas,2019)。句子改写不仅体现了作者对原文有深入的了解,还是作者写作可读性的一种技能,也是作者在学术交流中综合处理及运用外部信息的必备写作技能(Harris,2006)。

句子改写能力直接影响了二语习得者是否能科学地引用他人的观点,对学生的学术成功带来明显的影响(Liu&Deng,2019)。

这一挑战在来华研究生中更为显著。他们一方面需要了解与接受中国学术写作规范与引用文化,另一方面则是其自身学术背景多样导致对“原创性”与“剽窃”定义不同,并因此导致其利用外部文献时的剽窃与焦虑风险增大。如何实现提升来华留学生群体学术句子改写能力,是亟需解决、迫切而又至关重要的问题。

虽然当前关于句子改写技能教学方面的研究受到关注(Chandler&Sayeski,2021),但目前教师的指导还存在一些偏差。一些教师将注意力转移到句子的语法错误、词汇使用不当的问题(Reynolds&Teng,2021)上,而忽略了如何引导学生学习构建学术风格的句子(Truong&Do,2021),课堂上对于学生的训练也缺乏针对性,因此相应的教学干预措施较少。这种教学实践带来两种结果,一是学生“替换同义词”式改写;二是学生论文与原文太“像”,构成剽窃(Rets&Ilya,2018)。

动态评估理论为解决此类问题提供了相应的解决方案。动态评估的理论来源是维果茨基(1978)的社会

文化理论,其基本概念是“最近发展区”的思想(Dane shfar&Moharami,2018)。“最近发展区”是指儿童解决问题时的具体能力与潜在能力,当在经验丰富的成人指导和在能力更强的同伴的支持下,儿童实际水平和潜在水平之间的差距。动态评估旨在评估者在指导学生解决问题的时候帮助他们不断地接近最近发展区。

评价方式的动态性强调师生互动纳入评价过程中,使教师能够在了解学生水平的基础上给予支架式指导性教学,它与句子改写技能的教学要求十分一致,其匹配度很高主要体现在以下两点:

第一,动态评估诊断性的方式有助于解决教学指导的盲点。传统改错教学并不能准确地判断学生遇到难点是由于词汇问题、语法问题还是语体问题引起的,而动态评估通过教师的询问和学生的回答,可以直接找到学生出现难点的原因,从而做到精准指导。

第二,动态评估中的支架式教学能够给予学生学习支持,引导其超越“替换同义词”,提高技能层级,将形式模仿过渡至意义重构。

因此,本研究将运用动态评估理论设计教学方案,探讨其在提升学生句子改写能力方面的有效性,并试图回答以下两个问题:

(1) 动态评估教学法能否显著提升来华留学生句子改写能力?

(2) 动态评估教学法对句子改写能力的四个不同维度——“内容准确性”“表达连贯性”“语言正确性”和“学术规范性”——是否具有差异化效果?

本文共五部分。第一部分是引言部分,主要介绍基于动态评估理论指导下的句子改写教学模式的应用可

能性。第二部分是研究方法，介绍所研究的对象、标准及教学设计过程。第三部分是研究结果，分析相关实验数据。第四部分是研究的讨论，解释实验数据相关结果。最后是结论部分，阐述研究成果的局限性和后续的研究发展。

1 研究方法

1.1 研究对象

本研究的对象为中国北京 A 大学 42 名来华留学生研究生一年级学生，学生来自越南、泰国等 15 个国家，他们使用不同的母语，均通过 HSK 五级考试，处于中高级汉语水平。教学实验在学生的学术论文写作课程里进行。

1.2 前测与后测材料

我们设计了两组同等难度的句子改写任务，分别为教学前和教学后的测试中使用。每套测试卷设置了 5 个需要改写的句子。

通过鲁东大学文本评分平台 (LDU-TGP) (http://120.27.70.114:8000/analysis_a)，对两个材料进行了等值性检验，检验结果如下。

表 1. 前测后测文本相关数据

指标	前测文本	后测文本
难度等级	五级	五级
易读性得分	2.3305	2.3802
总字数	225 字	240 字
平均句长	45 字	48 字
字种数	104 字	112 字
词种数	74 词	79 词

表 2. 专家咨询的权重赋值统计结果

评估维度	统计指标	第一轮	第二轮	第三轮	最终权重
内容准确性	均值	34	31	30.4	30%
	标准差	8.22	4.18	2.61	
表达连贯性	均值	22	24	24.8	25%
	标准差	5.7	2.74	1.94	
语言正确性	均值	16	16	15	15%
	标准差	4.18	2.24	1.58	
学术规范性	均值	28	29	29.8	30%
	标准差	6.71	3.16	2.59	

具体每轮统计结果 (均值和标准差) 以及意见收敛过程如表 2 所示。如表所示，专家意见的分散程度 (标准差) 随着每轮的进行而逐渐减小，到第三轮时，所有维度的标准差均低于我们预设的 5% 共识标准。最终采用

两个文档被判定为同难度等级 (HSK5 级)，前测文档易读性分数是 2.3305，后测文档是 2.3802；从微观指标词数、平均句长、汉字丰富指数和词汇量来看，两个文档差值不大，因而两个文档可以用来对比。

1.3 评估量表

本次研究采用的量表主要基于二语习得的传统理论，Canale&Swain (1980) 指出，真正有意义的言语技能应该不仅仅包括语法的正误，同时还要有足够的流畅度。Jacobs 等 (1981) 研制了一个细化的具体的评估标准，将写作表现细化成一些量化的标准，比如内容、结构、词汇等。这二者的结合为我们量表的制定提供了思路。

为保持量表的有效性和适用性，本研究拟定的四维量表在开发过程中也参考了近年来相关领域的最新研究成果，保障了量表的科学有效和及时性。一方面，自动写作评估和以来源为中心的写作研究均表明，客观语言特征 (如表达流畅性、语言正确性) 与忠实反映原文所要表达的意义 (如内容正确性) 是评估改写效果的决定性因素 (Deane, 2014; Crossley, 2020)。另一方面，有关人工智能辅助写作的研究认为，由于人类与机器协作现象越来越普遍，对学术语体规范性的考察尤其具现实意义 (Rabbianty 等, 2023)。上述最新研究为本量表中的“内容准确性”“语言形式” (含表达连贯性和语言正确性) “学术规范性”等维度的确定提供了重要依据，保障了量表研究工具的有效性。

同时，为了反映上述各维度的重要性，本研究采用专家共识法确定了各维度的权重。请参见表 2。

的加权方案如下：“内容准确性”占 30%，“表达连贯性”占 25%，“语言正确性”占 15%，“学术规范性”占 30%。

在确定评估标准后，我们开始在实际教学情境中对

学生进行动态评估。

1.4 教学流程

步骤 1：开展前测。在正式讲解句子改写内容以前，先让学生进行一次前测（时长 45 分钟）。前测旨在测量学生句子改写方面的真实水平。

步骤 2：教学干预阶段。紧接着为期 3 周的核心教学干预阶段，以“建立标准—支架式教学—技能内化”的完整过程，分析和解决学生句子改写技能的“最近发展区”和系统性地促使其改写技能的“内化”。具体如下。

讲解评价标准：为了将较为抽象的“改写能力”具象化，我们设计了四大维度的评估结构和 1~5 分的量表说明。

进行有效的支架式教学：对于学生知其然而不知其所以然，我们采用正反案例结合的方式，帮助学生得出“怎样才能避免此类问题”的问题诊断。如在选取典型不合格案例与优秀案例后，把改进过程进一步细分为四个维度技能，指导学生分别分析并自我反思。

同伴互评，促进能力内化：为激励学生将外部标准转换为自我调节能力，在干预策略中安排了同伴评估。

该策略的核心在于借助该框架引导学生对同伴进行相互评估讨论，通过彼此间的讨论加强认识，实现元认知的发展。

步骤 3：开展后测。教学干预三次结束后，进行了第二次测试。与前测一样，执行的方式及测试耗时一样，用来考察教学干预后学生学习结果的内化及迁移。

1.5 评分与数据分析

我们邀请了两位从事多年国际中文教学的教师，经过“双盲”评阅，即不知道学生是谁，也不清楚试卷是前测还是后测卷，最终，两份评分结果相差很小，两位教师评分的 Cohen's Kappa 系数为 0.82。对于两份教师的评分如果不同维度上相差 1 分，则由 2 位专家进行研讨，直至达成一致。

所有得分输入计算机，利用 SPSS26.0 软件中“配对样本 t 检验”功能比较教学前后的两次测试成绩，考察动态评估法是否能够带给学生明显的学习进步。

2 结果

本研究首先对前测和后测的分数进行了描述性统计分析，结果如表 3 所示。

表 3. 前测后测各项得分

评量项目	前测平均分	前测标准差	后测平均分	后测标准差
总加权平均分数	2.24	0.56	3.00	0.36
内容准确性	2.36	0.85	2.93	0.51
表达连贯性	2.17	0.76	3.14	0.72
语言正确性	2.17	0.70	2.52	0.55
学术规范性	2.21	0.65	3.19	0.77

如表 3 所示，后测的平均分数高于前测，表明学生在接受动态评估教学干预后，句子的改写能力在所有维度上均有所提升。

接下来，本研究采用配对样本 t 检验对前测和后测的分数进行分析。详细结果如表 4 所示。

表 4. 配对样本 T 检验结果

比较项目	平均差异(后测-前测)	t 值	自由度 (df)	p 值	效果量(Cohen's d)
总加权平均分数	0.76	10.253	41	0.000	1.58 (极大)
内容准确性	0.57	4.459	41	0.000	0.69(较大)
表达连贯性	0.97	6.179	41	0.000	0.95 (极大)
语言正确性	0.35	3.186	41	0.003	0.49 (中等)
学术规范性	0.98	10.468	41	0.000	1.62 (极大)

如表 3、表 4 所示，后测得分(M=3, SD=0.36)明显优于前测得分(M=2.24, SD=0.56)， $t(41)=10.253$ ， $p<.001$ 。Cohen's d 效应量为 1.58，说明教学干预呈现

了较好的总体效果。

5 个子维度的后测分值均明显高于前测分值（所有 p 值均<.001）。其中，“学术规范性”（d=1.62）、“内

容准确性”(d=0.69)与“表达连贯性”(d=0.95)3个维度进步最大,其效应量为较大及以上水平。

3 讨论

在本研究中,动态评估教学法对来华研究生学术中文句子改写任务有显著性的影响($p<.001$,Cohen's $d=1.73$);这一显著的有效度是因为动态评估所具有的一种“协同性效应”,即动态评估的“建立标准—支架式教学—技能内化”三步相辅相成,形成一个高效的动态、个性化的学习闭环。

此外,我们还发现,学生不同维度的提升并不均衡。“学术规范性”和“内容准确性”方面的提升最大,而“语言准确性”方面的提升最小。我们认为这种不均衡的深层原因是因为我们的教学手段特别是引导式的对话式教学和同伴互评式教学激发了学生的“元认知”能力,让其进一步反思关于读者、语境和意义等更深层次的问题。研究表明,这种互动式教学方式是引导学生进一步反思的有效手段(Crossman&Kite,2012)。但是对于语言偏误(所谓的“化石化”)而言,三周专注于高阶思维的训练能够帮助学生“意识到”这些错误,但如若想达到更高层次的语言准确性,则要花更多的精力进行专项训练(Geng&Jin,2024)。

总的来说,本研究从另一个侧面印证了Canale&Swain(1980)的交际能力理论,证明了动态评估教学法能够使学生在高水平的学术写作训练中发展其语言知识(语言准确性)、话语构建(表达连贯性)及社会语言学能力(学术规范性)。

结合上述研究发现,本研究有以下教学实践和理论意义。就教学实践而言,本研究的建议是,国际中文教师需从“结果导向”式的评分方式转变为“过程导向”的对话式反馈,并有意识地结合相关活动在课程上嵌入结构化的同伴评价,以培养学生独立、基于批判性思维的写作能力。同时,本研究也为动态评估理论的应用提供了清晰、可复制的课内实践模型。就理论贡献而言,本研究为动态评估理论的适用范围提供了高级学术写作类语言课堂的实证支持。

本研究尚有一些不足之处,也为后续研究提供了思路。本研究所选样本较少,一定程度上限制了研究结果的推广性;三周的干预期稍短了些。因此后续研究还可扩充样本量,并进行纵向研究;也可关注人机反馈的不同方式,借助访谈等质性研究探究学生的学习体验,也

将是富有价值的研究方向。

4 结论

本研究验证了动态评价理论在高级中文学术写作教学中的合理性与可操作性;其次,本研究给出可以模仿实践的教学案例,把理论、实践和评价结合在一起,为一线国际中文写作教师的教学实践提供直接参考。总而言之,本研究希望给出一条可以验证的高级中文学术写作教学及评价的有效路径。

参考文献

- [1]钱希,孟亚茹,&岳真.(2023).动态评估融入的混合式教学对学术英语转述能力发展的影响研究.外语教学,44(2),63-68.<https://doi.org/10.16362/j.cnki.cn61-1023/h.2023.02.014>
- [2]Ardelia, I., & Tiyas, Y. R. (2019). THE ACQUISITION OF PARAPHRASING AND ITS IMPACT ON TEACHING AND LEARNING QUOTATIONS TO AVOID PLAGIARISM. English Education Journal, 6. <https://ejournal.bbg.ac.id/geej/article/view/873>
- [3]Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. Applied Linguistics, 1(1), 1-47. <https://doi.org/10.1093/applin/I.1.1>
- [4]Chandler, B. W., & Sayeski, K. L. (2021). Teaching Writing Through Content: Six Instructional Activities to Improve Students' Sentence-Level Writing Skills. TEACHING Exceptional Children, 56(5), 336-344. <https://doi.org/10.1177/00400599211050087>
- [5]Crossley, S. A. (2020). Linguistic features in writing quality and development: An overview. Journal of Writing Research, 11(3), 415-443. <https://doi.org/10.17239/jowr-2020.11.03.01>
- [6]Crossman, J., & Kite, S. L. (2012). Facilitating improved writing among students through directed peer review. Active Learning in Higher Education, 13(3), 219-229. <https://doi.org/10.1177/1469787412452>
- [7]Daneshfar, S., & Moharami, M. (2018). Dynamic Assessment in Vygotsky's Sociocultural The

ory: Origins and Main Concepts. *Journal of Language Teaching and Research*, 9(3), 600-607. <http://dx.doi.org/10.17507/jltr.0903.20>

[8]Deane, P. (2014). Using writing process and product features to assess writing quality and explore how those features relate to other literacy tasks (ETS Research Report Series, RR-14-03). ETS. <https://doi.org/10.1002/ets2.12002>

[9]Geng, Z., & Jin, Z. (2024). A review of interlanguage fossilization in English learning. *Journal of Education, Humanities and Social Sciences*, 26, 153-160. <https://doi.org/10.54097/z4kvfr08>

[10]Harris, J. (2006). *Rewriting: How To Do Things With Texts*. University Press of Colorado. <http://www.jstor.org/stable/j.ctt4cgnf2>

[11]Ito, T., Kuribayashi, T., Kobayashi, H., Barrassard, A., Hagiwara, M., Suzuki, J., & Inui, K. (2020). Assisting authors to convert raw products into polished prose. *The Journal of Cognitive Science*, 21, 101-137.

[12]Jacobs, H. L., Zinkgraf, S. V. A., Wormuth, D. R., & Hartfiel, V. F. (1981). *Testing ESL composition: A practical approach*. Newbury House Publishers.

[13]Liu, Q., & Deng, L. (2019). Chinese MA Student Writers' Identity Construction in Citation Practices. *Chinese Journal of Applied Linguistics*, 42(4), 365-385. DOI:10.1515/CJAL-2019-

0022

[14]Rabbianty, E. N., Azizah, S., & Virdyna, N. K. (2023). AI in academic writing: Assessing current usage and future implications. *INSANIA: Jurnal Pemikiran Alternatif Kependidikan*, 28(1a), 14-35. <https://doi.org/10.24090/insania.v28ila.9278>

[15]Rets, I., & Ilya, A. (2018). Eliciting ELT students' understanding of plagiarism in academic writing. *Eurasian Journal of Applied Linguistics*, 4(2), 193-211. <https://doi.org/10.32601/ejal.464115>

[16]Reynolds, B. L., & Teng, M. F. (2021). Innovating teacher feedback with writing activities aimed at raising secondary school students' awareness of collocation errors. *Studies in Second Language Learning and Teaching*, 11(3), 423-444. <https://doi.org/10.14746/ssllt.2021.11.3.6>

[17]Truong, K. D., & Do, N. T. (2021). Selecting complex sentence structures to teach in academic writing: A corpus-based innovation. *English Australia Journal*, 37(2), 48-53. <https://files.eric.ed.gov/fulltext/EJ1324180.pdf>

作者简介：陈香，女（1981——），汉族，湖南湘潭，博士，副教授，研究方向：国际中文教育。

该论文获“世界汉语教学学会全球中文教育主题学术活动计划”资助，项目名称为“学术汉语写作教学中的动态评估模式研究”，项目号为 SH23Y16。