

形成性评估在口译课程中的应用有效性研究

刘 芹

(上海理工大学 外语学院, 上海 200093)

摘要: 通过英语专业两个三年级班级参与2014年为期16周的关于形成性评估应用研究项目, 探讨形成性评估应用于口译课程的有效性。两组被试在研究初期口译水平相当, 教学内容、课堂条件和任课教师都相同的条件下, 实验班实行形成性评估(包括课内表现、课外活动、期末考试和评估奖励四个部分), 对照班实行终结性评估(包括出勤率和期末考试两个部分)。研究结果显示, 学习动机和交流欲望更强的实验班学生比对照班学生收到更显著的学习效果。

关键词: 形成性评估; 口译课程; 学习动机; 交流欲望; 学习效果

中图分类号: H 319 **文献标志码:** A **文章编号:** 1009-895X(2015)04-0359-05

DOI: 10.13256/j.cnki.jusst.sse.2015.04.012

On the Effectiveness of Applying Formative Assessment into an Interpretation Course

Liu Qin

(College of Foreign Languages, University of Shanghai for Science & Technology, Shanghai 200093, China)

Abstract: This paper displays the effectiveness of applying formative assessment into an interpretation course of two classes of the third-year English majors through a 16-week study in 2014. The participants were at the same interpretation level at the beginning of the research, given the same teaching material, language laboratory and the same teacher during the study. The Experimental Class was evaluated by formative assessment (including In-class Performance, After-class Performance, Final Exam and Assessment Bonus) while the Control Class was evaluated by summative assessment (including Attendance and Final Exam). Results show statistically significant progress of the former over the latter. The better learning effect may be explained by higher motivation and willingness to communicate.

Keywords: *formative assessment; interpretation course; motivation; willingness to communicate; learning effect*

当今的国际交流形势对口译员的需求量日益增大。20世纪90年代开始, 国内诸多高校把口译课程作为英语专业必修课和非英语专业选修课加以开设。然而, 由于口译的学科地位尚未稳固, 目前国内的口译教学虽遍地开花, 但没有统一的教学大纲

和课程体系, 教学质量缺乏保证^[1]。

为了提高教学质量, 必须在课程设置、授课资料、教学方法和评估模式等诸方面加以研究。笔者以篇名为条件搜索1990年至2013年CNKI上发表的口译类论文发现, 上述四个方面的论文分别为

收稿日期: 2014-08-22

作者简介: 刘 芹 (1972-), 女, 教授。研究方向: 语言教学与评估、语料库语言学。E-mail: liulinda2004@163.com

44篇、38篇、688篇和57篇。而在口译评估领域的57篇论文中,只有10篇左右与口译课堂教学有关。其中有些是关于设计具体口译课程的评估模式(如李游子^[2],刘银燕^[3],刘银燕、张姗姗^[4],康晓芸^[5]),有些则是建议在口译评估中引入持续性评估、学生自评等多角度评估(如蔡小红^[6],肖桂芳、刘建珠^[7],王虹、俞宝红^[8],万宏瑜^[9])。

形成性评估的理念已开始在国内口译课程中显现,有必要对这一模式的实际应用效果加以深入研究。因此,笔者将其应用于实际口译课程教学,以检验其有效性。

一、研究理据

(一) 形成性评估及其教学效应

外语教学评估方式通常有两种:终结性评估(summative assessment)和形成性评估(formative assessment)。前者通过学习阶段结束后的统一考试来评价学生能力,后者则是经过充分设计的一种持续性评估模式^[10]。

Weir^[11]敦促教师在教学中使用形成性评估,以便根据学生学习情况的变化及时调整教学手段。Bachman & Palmer^[12]认为形成性评估是指导学生合理学习,并帮助教师监控学习过程的有效方式。Black & Wiliam^[10]指出形成性评估能有效促进学习自主性。Genesee & Upshur^[13]提出形成性评估最重要的效应是能反映目前学习情况和学习目标之间的差距。

众多学者指出形成性评估应基于学生需求设计,以加速教学模式的变革(如Genesee & Upshur^[13], Bachman^[14])。它使学生发现自身学习上的弱点,制订相应的学习目标,充分发挥学习潜能,达到最大的学习效果。另一方面,教师也能通过反馈发现教学过程中的不足之处,促进教学革新。

形成性评估方式多样,例如课堂交流、评分、自我评估(self-assessment)、同伴评估(peer-assessment)、终结性考试的形成性使用等(Black & Wiliam^[15])。其中,自我评估和同伴评估是最重要的方式。前者要求学生评价自己的学习情况,并据此制订学习计划。后者要求学生对有相似教育背景的同伴做出评价(Falchikov^[16], Topping^[17])。

Black & Wiliam^[15]提出形成性评估应符合课堂教学模式。Clark^[18]亦指出剑桥大学评估改革小组(ARG)和经济合作与发展组织(OECD)等国际研究组织非常强调课堂教学通过促进学习动机和自

我尊重而加强自我学习规范的正面作用,因此形成性评估对改变学生学习动机和自我尊重意识的重要性不容忽视。

由此可见,形成性评估与学习动机之间有着紧密联系。教师应经常与学生交流,了解他们的学习动机,并与他们一同设定教学计划。同时,教师应尽力激发学生的潜力,使他们能全身心地融入到自我评估和同伴评估活动中,以确保学习过程的有效性。

(二) 形成性评估在翻译课程中的应用

虽然形成性评估在教学中的作用已有共识,但讨论其运用于翻译课程的研究却十分匮乏,以下列举一些这方面的研究成果。Sidman-Taveau^[19]在一门计算机辅助成人英语课程中,将团队协作策略和合作翻译方法运用于形成性评估中,发现其有益效应。Elorza^[20]探索了在学习跨文化交际能力的笔译活动中运用形成性评估的潜在有效性。在我国翻译教学过程中,多使用终结性评估,很少运用形成性评估(庄智象^[21])。因此,这方面的研究成果较少,且多限于理论探索(如冷静^[22],那彩霞^[23],解晓丽^[24])和初步应用探析(如曹荣平、陈亚平^[25])。

(三) 学习动机和交流欲望

在二语/外语教学中可以发现,有些学生非常乐于交流,而有些则很被动。即使他们与主动的学生拥有相似的交流能力,但学习效果会有不同。Drnyei^[26]指出在诸多影响二语学习效率和成绩的因素中,动机是关键因素。在MacIntyre等^[27]提出的交流欲望(willingness to communicate, WTC)模式中,动机位于中心位置。它以组间交际的情感和认知环境为基础,最终可以引出自信心,激发与别人交流的愿望。学习和使用第二语言的愉悦感和满足感会使学习者对学习过程付出更积极和周密的努力。基于此,Peng & Woodrow^[28]构建了中国英语课堂教学环境中的WTC模式。在此模式中,学习英语的动机引发用英语交流的信心,并最终形成用英语交流的欲望。

那么,中国学生学习英语的动机如何?高一虹等^[29]通过探索性因子分析方法(exploratory factor analysis, EFA)研究了中国不同高校2 278名本科生后发现了七大因素,分别是内部兴趣、即时成就、学习情境、出国需要、社会责任、个人发展和信息媒介。其中,内部兴趣和即时成就最为重要。他们进一步指出“内部兴趣”与Atkinson等1953年提出的经典动机模型相符,而“即时成就”属于

Gardner & Lambert 1972 年提出的工具性动机。可见，第二个主要动机因素与评估紧密相关。

在中国英语教学中，一般只在课堂学习中使用英语，教学基本与考试挂钩（Peng & Woodrow^[28]）。但成功的学习者需要有意识地寻找自己的学习目标并规划英语学习，充分运用有效的英语学习策略，监控并评价自己的学习过程（Liu^[30]）。既然学界已认为形成性评估是促进学生学习动机和学习效果的最强大途径之一（Clark^[18]），则有必要探索其在口译课程中如何激发学生的学习动机和交流欲望，进而促进学习效果。

二、研究实施

笔者在 2014 年上半年进行了一学期的教学实验。在这之前，对 180 名大学生和 50 名口译教师展开问卷调查，询问形成性评估是否比终结性评估更适合口译课程，及应该如何设计具体评估模式。根据回收的 159 份有效学生问卷和 48 份有效教师问卷可知，形成性评估更受欢迎。关于具体评估模块，35.8% 的学生和 62.5% 的教师认为课堂出勤、课内表现、课后活动和阶段反馈最为重要（有关该调查的详细信息，请参见刘芹、熊秋婷^[31]）。这些信息为本研究具体评估模块的设计提供了参照。

（一）研究被试

上海某高校两个英语专业班级，共 58 名学生是本研究的被试，笔者取其中一个班级为实验班，另一个班级为对照班，两个班级人数相当。

（二）研究过程

研究从 2014 年 2 月进行至 2014 年 6 月。每个班级每周上 4 学时的口译课，该学期一共 16 周，总

计 64 学时。第一周和最后一周用于前测和后测，14 周用于实验。两次测试的试题都是 2013 年 12 月英语专业八级口试（TEM8-Oral）的口译部分。

为了确保实验结果的可信度，两个班级的授课资料、授课教师和授课地点完全相同。对照班采用终结性评估，由 30% 的出勤率和 70% 的期末大考成绩组成学期成绩；实验班采用形成性评估（详见表 1）。

由表 1 可知，实验班的学期成绩由课内表现（占 30%）、课外活动（占 40%）、期末考试（占 20%）和评估奖励（10%）四部分组成。其中，课内表现又分为自我评估、同伴评估和教师评估三部分（学生在学期中的所有课堂活动都按以下表 2 评分）。

表 1 实验班成绩构成

Tab. 1 Final score configuration of the experimental class

评估构成	成绩占比
课内表现	自我评估 10%
	同伴评估 10%
	教师评估 10%
课外活动	四次录像作业 40%
期末考试	20%
评估奖励	10%
总分	100%

关于课外活动，要求学生以学习小组为单位完成，在期末上交四个录像作业，主题分别为“礼仪祝辞”、“陪同参观”、“媒体访谈”和“商务谈判”。每位学生在每个作业里扮演客户和译员等不同角色，录像可以反复进行，只需上交最好版本。课后活动的评分标准与课内表现相同。每次作业 10 分，共 40 分。

评估奖励由教师根据学生是否认真参与自我评估和同伴评估判定给分，占学期总分的 10%。

表 2 课堂表现评估表

Tab. 2 Assessment form of the in-class performance

Overall Grade	Comprehension of SL material	Reconstruction of TL material	Use of interpretation skills	Professionalism	Progress
5 Exceptional	Totally correct	Totally correct	Excellent	Perfect	Remarkable
4 Outstanding	Almost correct	Almost correct	Suitable	Excellent	Great
3 Average	Just-so-so	Just-so-so	Average	Good	Certain
2 Having Potential	Somewhat correct	Somewhat correct	Unsuitable	Average	No
1 Needing to Catch up	Totally incorrect	Totally incorrect	None	None	Setback

（注：SL 为原语，TL 为译语）

三、数据分析与讨论

笔者从质化和量化两方面对口译课程引入形成性评估的有效性进行分析。

(一) 质化分析

为了检验形成性评估是否能促进学生的学习动机和交流欲望,每月进行一次课堂观察和课后访谈。

1. 学习动机

由于内部兴趣是二语学习的最佳动机,笔者在实验开始一个月后询问两个班的学生是否对口译课程真正感兴趣。实验班的大多数学生表现出了浓厚的兴趣,因为该课程某种程度上激发了他们当口译员的潜质。相比之下,对照班的学生并没有完全对口译课程感兴趣。

为了探究实验班学生比对照班学生是否在课后付出更多的努力,笔者在实验中期对所有学生进行了访谈,内容涉及:是否制订自己的学习目标和计划;是否遵循学习目标和计划;是否形成课后锻炼口译技能的习惯;每周花多少小时练习口译等。

实验班的 23 名学生和 20 名学生对前两个问题分别给出肯定回答。相比之下,对照班只有 7 名学生和 4 名学生对前两个问题分别给出肯定回答。可以看出,大多数实验班同学能够自我制订学习目标和计划并自觉遵循,而只有很少一部分对照班学生可以做到这一点。

关于第三个问题,实验班的 20 名学生和对照班的 10 名学生给出肯定回答。关于第四个问题,实验班的 27 名学生回答他们每周花在课后活动的时间为 6 小时以上,其中有 11 名学生每周花 11 小时以上练习口译。相对而言,对照班只有 12 名学生每周花 6 小时以上练习口译,其中 6 名学生花 9 小时以上练习。

由此可见,拥有更强学习动力的实验班学生投入了更多的时间和精力学习,而对照班学生付出的努力相对较少。

2. 交流欲望

笔者收集到的 4 次课堂观察录像显示,实验班学生对学习更感兴趣,有较高的学习热情和参与意识。他们考虑和讨论的内容都与课内口译话题有关。当回答教师提问和在同学面前展示模拟口译练习时,

他们信心十足,嗓音响亮清晰。然而,对照班学生很少主动在模拟口译练习中发言。即使按教师要求回答时,内容也很简短,声音很轻。该班学生交流欲望不强,有些甚至在课上闲聊。

以 2014 年 5 月 8 日(学期第 12 周)为例,当教师要求学生模仿中国领导人进行新年演讲和口译训练时,实验班的所有学生立即开始与同桌练习,而对照班的大多数学生显得不愿参与。在 10 分钟准备期后,实验班的一半学生举手要求在全班同学面前模拟,而对照班无一人要求上台。

可以看出,交流欲望更强的实验班学生的课堂表现更佳。但是否因此达到更好的学习效果呢?下文的系列量化研究可以回答这一问题。

(二) 量化分析

为确保评分信度,两个班级的前测和后测都由两名教师评分。由于这些考试的评分员间信度都超过 0.7 的理论标准(实验班前测和后测分别为 0.931 和 0.921,对照班前测和后测分别为 0.923 和 0.907,均在 0.01 水平上显著),笔者取两名教师评分的平均分作为学生的最后得分。

1. 描述性统计

表 3 展示实验班和对照班前测和后测成绩的描述性统计情况。实验班前测的平均分为 63.862 1,标准差为 8.918 2;后测的平均分为 70.069 0,标准差为 8.413 1。对照班前测的平均分为 61.189 7,标准差为 12.262 3;后测的平均分为 63.706 9,标准差为 11.474 8。倾斜度(Skewness)和峰值(Kurtosis)均在 ± 2 之间,符合正态分布理论标准,可以进行下一步的对比分析。

表 3 两个班前测和后测成绩描述性统计
Tab. 3 Descriptive statistics of the pre-test and post-test of the two classes

描述项目	实验班前测	实验班后测	对照班前测	对照班后测
人数	29	29	29	29
最小值	40.00	44.50	39.00	39.00
最大值	77.50	83.00	89.50	84.50
平均分	63.862 1	70.069 0	61.189 7	63.706 9
标准差	8.918 2	8.413 1	12.262 3	11.474 8
倾斜度	-0.654	-1.035	0.258	-0.384
峰值	0.497	1.702	-0.076	-0.050

2. 对比分析

表4~7显示对比分析结果。

表4 前测独立样本 *t* 检验

Tab.4 Independent samples *t* test of the pre-test

	<i>t</i> 值	自由度	显著性(<i>p</i>)	平均分差异
实验班-对照班	-0.949	56	0.347	-2.672 4

表5 后测独立样本 *t* 检验

Tab.5 Independent samples *t* test of the post-test

	<i>t</i> 值	自由度	显著性(<i>p</i>)	平均分差异
实验班-对照班	2.408	56	0.019	6.362 1

表6 实验班配对样本 *t* 检验

Tab.6 Paired samples *t* test of the experimental class

	<i>t</i> 值	自由度	显著性(<i>p</i>)	平均分差异
后测-前测	8.363	28	0.000	6.206 9

表7 对照班配对样本 *t* 检验

Tab.7 Paired samples *t* test of the control class

	<i>t</i> 值	自由度	显著性(<i>p</i>)	平均分差异
后测-前测	3.263	28	0.003	2.517 2

表4显示,实验班和对照班的前测成绩之间不存在显著性差异($t(56) = -0.949, p = 0.347 > 0.05$),即两个班级学生在研究初期的口译能力相当。表5显示,实验班和对照班的后测成绩平均分相差6.362 1(在0.05水平上显著)。表6显示,实验班后测比前测平均分高出6.206 9分(在0.01水平上显著)。表7显示,对照班后测比前测平均分高出2.517 2分(在0.01水平上显著)。由此可见,在经过一个学期的学习后,实验班学生比对照班学生的进步更大。形成性评估通过激发学生的学习动机和交流欲望促进了学习效果。

四、结束语

蔡小红^[6]指出:“学习与动机、信心、责任心和积极性紧密相关,因此在设计考评类型与方法时,还要充分考虑其鼓励和推进的作用,通过科学的训练评估,帮助学生明确动机,坚定信心,培养兴趣,调动积极性,建立责任心。”基于此,本文从学习动机和交流欲望两方面对形成性评估运用于口译课

程的有效性进行了研究。

对比分析结果显示,与对照班学生相比,采用该评估方式的实验班学生取得了更大的进步。由于学生对自己和学习同伴有评估的权力,激发了他们的学习动机和交流欲望。另一方面,教师也能从期末考试以外的其他途径对学生有较全面的了解。同时,课内表现、课外活动、期末考试和评估奖励四部分组成的形成性评估模式避免了学生某次成功或失败导致的不公正评分。由于要求学生在整个学期中坚持训练,他们获得持续进步,口译能力稳步提高。

本文只是对形成性评估应用于口译课程的有效性进行了初步研究,提出的形成性评估模式还有待论证。笔者将开展进一步的研究以验证其效度和可行性。

参考文献:

- [1] 詹成. 中国口译教学三十年:发展及现状[J]. 广东外语外贸大学学报,2010,21(6):89-92.
- [2] 李游子. 口译学习指标和测试评估[J]. 外语研究,2007(2):69-70.
- [3] 刘银燕. 英语专业本科口译教学测试与评估探索[J]. 西南民族学院学报(人文社科版),2009(S1):69-72.
- [4] 刘银燕,张珊珊. 英语专业本科口译教学结业测试设计与评估方法探索[J]. 外语研究,2009(4):74-78.
- [5] 康晓芸. 论口译教学中的测试评估[J]. 林区教学,2010(4):42-43.
- [6] 蔡小红. 论口译教学训练评估[J]. 中国翻译,2005(6):58-61.
- [7] 肖桂芳,刘建珠. 我国口译教学评估的现状与探讨[J]. 江西理工大学学报,2007,28(5):77-80.
- [8] 王虹,俞宝红. 英语本科口译模拟功能评估模式实证研究[J]. 无锡商业职业技术学院学报,2010(3):91-92.
- [9] 万宏瑜. 口译形成性评估案例分析[J]. 黔南民族师范学院学报,2012(5):80-87.
- [10] Black P, Wiliam D. Assessment and classroom learning [J]. Assessment in Education Principles Policy & Practice,1998,5(1):7-74.
- [11] Weir C J. Understanding and Developing Language Tests [M]. London:Prentice Hall,1993.
- [12] Bachman L F, Palmer A S. Language Testing in Practice [M]. Oxford:Oxford University Press,1996.
- [13] Genesee F, Upshur J A. Classroom-based Evaluation in Second Language Education [M]. Beijing: Foreign Language Teaching and Research Press,2001.

(下转第374页)

- [10] 何莲珍. 自主学习及其能力的培养[J]. 外语教学与研究, 2003, 35(4): 287-289.
- [11] 徐锦芬, 占小海. 国内外“学习者自主”研究述评[J]. 外语界, 2004(4): 2-9.
- [12] Ellis R. Understanding Second Language Acquisition [M]. Oxford: Oxford University Press, 1985.
- [13] Bandura A. Self-efficacy: toward a unifying theory of behavioral change[J]. Psychological Review, 1977, 84(2): 191-215.
- [14] Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory [M]. Englewood Cliffs: Prentice-Hall, 1986: 99-102.
- [15] Pajares F, Miller M D. Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis[J]. Journal of Educational Psychology, 1994, 86(2): 193-203.
- [16] 周勇, 董奇. 学习动机、归因、自我效能感与学生自我监控学习行为的关系研究[J]. 心理发展与教育, 1994(3): 30-33.
- [17] 胡桂英, 许百华. 初中生学习归因、学习自我效能感、学习策略和学业成就关系的研究[J]. 心理科学, 2002, 25(6): 757-758.
- [18] 张庆宗. 自我效能与英语学习策略运用的调查研究[J]. 中国外语, 2004, 1(2): 45-48.
- [19] 卢洁. 外语学习自我效能和学习策略关系研究报告[J]. 山东外语教学, 2005(5): 83-87.
- [20] 王毅敏. 自我效能、动机、观念和策略对英语成绩的影响[J]. 西安外国语学院学报, 2006, 14(3): 51-53.
- [21] Bandura A, Schunk D H. Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation [J]. Journal of Personality and Social Psychology, 1981, 41(3): 586-598.
- [22] Gardner D, Miller L. Establishment Self-access from Theory to Practice [M]. Cambridge: Cambridge University Press, 1999: 11.

(编辑: 朱渭波)

(上接第363页)

- [14] Bachman L F. Building and supporting a case for test use[J]. Language Assessment Quarterly, 2005, 2(1): 1-34.
- [15] Black P, Wiliam D. Changing teaching through formative assessment: research and practice—The King's-Medway-Oxfordshire-Formative assessment project [C] // Formative Assessment-Improving Learning in Secondary Classrooms. London: Organization for Economic Cooperation and Development, 2005: 223-240.
- [16] Falchikov N. Peer feedback marking: developing peer assessment[J]. Innovations in Education and Training International, 1995, 32(2): 175-187.
- [17] Topping K. Peer assessment between students in colleges and universities [J]. Review of Educational Research, 1998, 68(3): 249-276.
- [18] Clark I. Formative assessment and motivation: theories and themes [J]. Prime Research on Education, 2011, 1(2): 27-36.
- [19] Sidman-Taveau R L. Computer-assisted project-based learning in second language: case studies in adult ESL [D]. Austin: The University of Texas at Austin, 2006.
- [20] Elorza I. Promoting intercultural competence in the FL/SL classroom: translations as sources of data [J]. Language and Intercultural Communication, 2008, 8(4): 261-277.
- [21] 庄智象. 我国翻译专业建设: 问题与对策 [M]. 上海: 上海外语教育出版社, 2007.
- [22] 冷静. 形成性评估在口译项目教学中的应用[J]. 考试周刊, 2009(24): 97-98.
- [23] 那彩霞. 英语专业翻译教学中形成性评价模式的探讨[J]. 辽宁经济管理干部学院学报, 2009(2): 104, 106.
- [24] 解晓丽. 论形成性评价在高职英语翻译教学中的应用[J]. 教育教学论坛, 2010(12): 147.
- [25] 曹荣平, 陈亚平. 形成性评估及其在口译教学中的应用探析[J]. 中国翻译, 2013(1): 45-50.
- [26] Dörnyei Z. Motivation in second and foreign language learning [J]. Language Teaching, 1998, 31(3): 117-135.
- [27] MacIntyre P D, Dörnyei Z, Clément R, et al. Conceptualizing willingness to communicate in a L2: a situational model of L2 confidence and affiliation [J]. The Modern Language Journal, 1998, 82(4): 545-562.
- [28] Peng J E, Woodrow L. Willingness to communicate in English: a model in the Chinese EFL classroom context [J]. Language Learning, 2010, 60(4): 834-876.
- [29] 高一虹, 赵媛, 程英, 等. 中国大学本科生英语学习动机类型[J]. 现代外语, 2003, 26(1): 28-38.
- [30] Liu H G. ELF motivations and autonomy in English learning: an investigation of Chinese non-English major undergraduates [J]. CELEA Journal, 2008, 31(5): 82-97.
- [31] 刘芹, 熊秋婷. 以学习为中心的科技英语口语课程设置[J]. 上海理工大学学报(社会科学版), 2011, 33(1): 62-68.

(编辑: 朱渭波)