

中文社会科学引文索引 (CSSCI) 来源集刊

中国知网 (CNKI) 全文收录

DUIWAI HANYU YANJIU

对外汉语研究

上海师范大学《对外汉语研究》编委会 编

第十期



商务印书馆
The Commercial Press

中文社会科学引文索引(CSSCI)来源集刊

中国知网(CNKI)全文收录

对外汉语研究

第 十 期

上海师范大学

《对外汉语研究》编委会 编



创于1897

商务印书馆

The Commercial Press

2013年·北京

图书在版编目(CIP)数据

对外汉语研究. 第10期/上海师范大学《对外汉语研究》编委会编. —北京:商务印书馆, 2013
ISBN 978-7-100-10309-1

I. ①对… II. ①上… III. ①汉语—对外汉语教学—教学研究—文集 IV. ①H195-53

中国版本图书馆 CIP 数据核字(2013)第 234759 号

所有权利保留。
未经许可,不得以任何方式使用。

DUIWAI HANYŮ YANJIŪ

对外汉语研究

(第十期)

上海师范大学《对外汉语研究》编委会 编

商务印书馆出版

(北京王府井大街36号 邮政编码 100710)

商务印书馆发行

北京瑞古冠中印刷厂印刷

ISBN 978-7-100-10309-1

2013年12月第1版

开本 787×1092 1/16

2013年12月北京第1次印刷

印张 15¼

定价: 35.00 元

目 录

汉语教学研究

- 新加坡小学一年级华文口语诊断评量表的开发
.....〔新加坡〕陈之权 孙晓曦 1
- 论高级阶段修辞教学
——以比喻手法的输入与输出为例 肖奚强 颜 明 13
- 论汉语教学大纲本土化
——以法国《初中汉语教学大纲》为例 丁安琪 24
- “把”量构式形成的认知机制和语义格局
——基于汉语作为第二语言的量词教学策略研究
..... 高亚楠 吴长安 34
- 国际汉语基础教材中练习的设置与编写〔新加坡〕徐 峰 44
- 国际汉语教师培养模式考察:问题与对策 马国彦 58
- 第二语言教学的听说微技能训练及方法 高 红 72
- 汉语国际教育硕士教师话语微技能对比核查
——以教师话语中的标记语赘言及其纠正为个案 姜有顺 81
- 海峡两岸利用同语素类推构词情况研究 许 蕾 87
- 19 世纪西方人汉外词典编纂中的汉语“根字”研究 江 莉 95
- 外向型汉语学习词典的检字法及其创新 ... 段濛濛 付 娜 孙 菁 105

汉语本体研究

- 及物动作动词构成的句干句式 范 晓 110

新加坡小学一年级华文口语 诊断评量表的开发^{*}

〔新加坡〕陈之权 孙晓曦

摘 要:本文利用质化和量化的分析数据描述了“新加坡小学一年级华文口语诊断评量表”的研发过程,并验证了该评量表的效度、信度和可操作性。通过对数据的呈现和解析发现,“新加坡小学一年级华文口语诊断评量表”的建构由于主要依赖真实的口语语料,同时结合专家和一线教师的意见,其效度和信度是一个逐步建立的过程,而且最终都达到了很高的程度。除了科学性,数据证明该评量表也是一个操作性很强的口语诊断工具,而且适用于课堂的实时评估。

关键词:新加坡小学一年级华文口语诊断评量表;效度;信度;可操作性

一、背景

新加坡自 1965 年独立后,出于对生存发展与地域政治的考量,政府在政策上一开始就规定以英语作为行政和商贸用语,但保障各族有使用本族语言的权利,并通过双语教育政策传承各族的语言文化(李光耀,2011)。以英语为行政兼商贸用语的政策,导致国民的社会经济地位在很大程度上与个人英语掌握能力相关的结局,英语成了社会的主流语言。现实的考量致使越来越多的家长把孩子送进英文学校受教育,以便日后得以谋得较好的职业。这批独立之后入学的孩子,目前已经成为新加坡经济人口的主要组成部分。从小就接受英文教育的结果,使他们习惯性地以英语思考和沟通,以英语为主要家庭用语。英语家庭在全国人口中的比例因此迅速扩大,其中尤以华族家庭的变化最大。

新加坡教育部于 2010 年公布的数据显示,过去 20 年来,以英语作为家庭主要用语

^{*} 感谢新加坡教育部“新加坡小一华文口语能力诊断工具”项目对本研究的支持。

的小学一年级华裔新生的比例,从1991年的28%提高到2010年的59%,且有继续增加的趋势(新加坡教育部,2010)。虽然如此,依然有相当部分的家庭,继续以母语为主要的家庭用语。与此同时,近年来移民新加坡的外国人口也在增加,其中尤以中国移民的人数为最,带来了纯讲高水平华语的家庭。此外,来自非华语地区的移民人口也在上升。这批为数不少的非华语地区的移民,因为中国崛起为重要经济体的关系,纷纷要求孩子在学校里修读华文。本地华裔国民家庭用语的迅速“脱华入英”、华语地区移民补充了原已萎缩的华语家庭、非华语地区移民孩子华语的低起点,这三个本地语言生态环境中的重要元素,导致了新加坡学校华语教学背景的多样性和复杂性的特点,也无可避免地决定了差异性教学将是华文教学的主要特色。

语言的学习规律是先听说后读写,以听说能力的发展来带动读写能力的发展(McGuinness, 2006)。新加坡当前社会语言环境下这三类背景的学生入学之后,华文教学上面对的首要问题便是华语口语能力的明显差距。华文教师需要更专业地诊断学生的口语起点,并针对不同能力的学生设置最优化的口语教学。新加坡华文教师在过去很长的时间里,对学生口语能力的判断纯凭个人的专业知识,各校之间没有统一的诊断标准。因此新加坡需要一个能客观有效评鉴学生口语能力的工具,既能为全国不同学校评价学生口语能力提供参照,又能对学生口语表现的强弱面进行诊断,以协助教师根据学生口语能力的优势与不足进行差异性教学。

基于这样的情况,南洋理工大学新加坡华文教研中心,于2010年向新加坡教育部申报了“新加坡小一华文口语能力诊断工具”项目,以协助华文教师科学、有效地诊断学生的口语能力,并进行有针对性的口语教学。这套工具已经于2011年11月产出,其中的“新加坡小一华文口语能力诊断评量表”是核心部分。它不仅能够在课堂口语活动过程中诊断学生的口语强弱面、描述学生的口语表现,还能记录学生的口语能力发展以及他们在课堂上使用华语进行交际的信心,能够给教学提供及时的反馈,让教师能够根据学生的需求调整教学策略和内容,进行针对性的差异教学,进而提高每一个学生的华文口语能力。

二、第二语言口语评量表的研发概况

第二语言口语测量研究的发展要归因于二战后对外语人才的极度需求和外语教学的蓬勃发展。为选拔合格的外语人才,并评估不同外语教学方法的效果,很多第二语言口语能力评量表相继问世。其中不乏一些对后来的口语测量研究意义深远的量表(见表1)。

表 1 主要第二语言口语能力评量表

描述 量表名称	评分方式	级别	研发方法	对象	目标语言
FSI(Foreign Service Institute)	综合	11	直觉与经验	成人	第二语言
ILR (Interagency Language Roundtable Scale)	综合	13	直觉与经验	成人	第二语言
ACTFL(the American Council on the Teaching of Foreign Languages)	综合	9	直觉与经验	成人(主要为大学生)	英语为第二语言
ISL PR (International Second Language Proficiency Ratings)	综合	12	直觉与经验	成人及青少年	第二语言
FNC(Finnish National Certificate)	综合	6	直觉与经验	成人	10 种语言
CLB (the Canadian Language Benchmarks)	综合	12	直觉与经验	成人	英语为第二语言
ALTE (Association of Language Testers in Europe)	综合	6	直觉与经验	成人	27 种语言
CEFR(Common European Framework of Reference for languages)	综合、分项	6	直觉、经验与实证数据	成人及年轻人	欧洲语言
TSE (Test of Spoken English)	综合、分项	5	直觉与经验	成人	英语
Melbourne medical students' diagnostic speaking scales	综合、分项	6	直觉与经验	成人(大学生)	英语
HSK(Hanyu Shuiping Kaoshi)	综合	11	直觉与经验	成人	汉语为第二语言
HKC (Hanyu Kouyu Shuiping Ceshi)	综合	9	直觉与经验	成人	汉语为第二语言
TOP (Test of Proficiency - Hanyu)/TOCFL (Test of Chinese as a Foreign Language)	综合	4	直觉与经验	成人	汉语为第二语言
CCCC(Children's Chinese Competency Certificate)	综合	3	直觉与经验	儿童	汉语为第二语言

表 1 所列的量表中,FSI,ILR 与 ACTFL 量表一脉相承,且前两者的问世源于当时的军事历史背景,后者则主要针对外语教学领域。FSI 量表是美国在 20 世纪 50 年代为了评量军事外交人才的口语能力而制定的标准。它的能力标准完全依赖人们对受良好教育的本族语者的定义,并无任何客观证据。基于 FSI 量表开发的 ILR 和 ACTFL 量表用途更加广泛、标准更加细化、能力级别更加科学,至今还在世界上广泛使用。

ISL PR 量表由澳大利亚开发,其内容在很大程度上与 ILR 和 ACTFL 量表相似,但在使用上则具体分为特定职业专用型和通用型。

FNC 量表是芬兰 1994 年出台的国家外语水平考试的口语能力评量工具。除了口语,该考试还考核阅读、写作、句型词汇和听力。FNC 量表将口语能力分为六级,并将其平均设在三个水平的考试(初级、中级和高级)中进行评估。

CLB 量表是 20 世纪末加拿大政府针对外来移民制定的英语作为第二语言的语言能力评估标准。它的语言能力描述跨越 12 个级别,并涵盖听、说、读、写四个方面。较之于其他量表,CLB 量表更明确地描述了在不同交际场合中,应试者可以完成的交际任务。

ALTE 和 CEFR 量表都于 20 世纪末源自欧洲。前者的开发团队为后者的开发提供了有力帮助,而后者的语言能力级别又为前者提供了参考。尽管如此,二者的开发目的却不尽相同。ALTE 量表的目的是为欧洲各国之间外语考试证书的相互认证提供统一标准。CEFR 量表则主要是为全欧洲地区语言教学的大纲编写、课程设置、教材编写和考试设计提供普遍框架(韩宝成,2006)。对于语言能力描述,ALTE 量表与 CLB 量表比较接近,分别从社会生活情境和旅游情境两个角度描述应试者的听、说、读、写之技能。

TSE 量表也是现今最为广泛使用的英语口语能力评量表之一。该量表将口语交际能力分为五个等级,并按照 Bachman 和 Palmer(1996)提出的交际能力构成框架对每个级别的能力表现进行描述。该量表于 2001 年由美国教育考试服务中心(ETS)推出,是专门评估口语交际能力的评量表。

“墨尔本医科学生口语诊断量表(Melbourne medical students' diagnostic speaking scales)”较之于其他评量表,是明确以诊断为目的的口语能力评量表,它能够提供应试者现有交际水平与实际要求之间差距的详细信息。

针对汉语为第二语言的语言能力评量表也为数不少,其中 HSK、HKC、TOP/TOCFL 和 CCCC 量表的使用颇为普遍。HSK 从设计到使用至今已有 20 多年,不仅在中国大陆和香港,而且在澳大利亚等地也多为使用。HKC、TOCFL 和 CCCC 量表都是用来评估口语能力的评量表。前者主要在中国大陆使用,后两者则是在中国台湾设计和使用,值得指出的是,CCCC 量表的测试对象集中为 7 到 12 岁的儿童。

以上所述评量表主要针对第二语言口语能力的评量,而且按照不同标准将口语能力分成不同级别,这对准确区别不同能力水平学习者的口语表现及详细描述其口语表现的特点大有裨益。尽管如此,根据表 1 所示,这些评量表仍有以下局限:第一,大部分评量表都以综合评分的方式对口语水平进行评定,只有少数几个评量表(如 CEFR 和

TSE 量表)针对口语能力的不同侧面评定分项分数。例如,较之于 ALTE 和 CLB 量表,CEFR 量表对语言能力和语言活动的划分和描述更为细致。除了综合性语言能力描述,CEFR 量表还从范围、准确性、流利度、互动性和连贯性五个方面进行分项评估。第二,除 CEFR 量表之外的所有量表对口语能力的描述都主要基于量表设计者或语言专家们的直觉或经验,缺乏来源于第二语言学习者的客观证据的支撑。相反,CEFR 量表的设计并不是依赖单纯的直觉与经验,而是将其与量化数据相结合,从而提高评量表的准确性和可靠性。第三,大部分评量表的测评对象为成人,针对儿童口语能力评估的工具还很有限。第四,还有待开发评估中国以外地区的汉语为第二语言学习者口语能力的评量表。第五,除“墨尔本医科学生口语诊断量表”外,其他量表都以衡量学习者口语表现特点、界定他们的口语水平为目的。这种评量信息对于了解学习者的口语程度确有价值,但其教学意义并不明显。测试的主要目的之一是服务教学,只有将评量信息有效融入教学实践,使教学有据可依,才能保证教学的针对性和有效性。从这个角度而言,能够传递学习者口语能力不同方面强弱项信息的诊断性测试则更具开发和实践价值。

新加坡是一个多种族、多语言社会。虽然华人人口众多,但英语作为官方和通用语言的地位决定了华文的教与学带有第二语言之特点,华文的使用和发展也明显带有新加坡本地之特色。因此,华文教师和专家学者们都意识到建立符合新加坡特色的华文口语评估工具的必要性和迫切性。以上综述的这些普遍使用并深受大众认可的口语能力评量表不仅为“新加坡小一华文口语能力诊断评量表”的构建提供参考,还为对以往局限性的突破提供了方向。

本研究中“新加坡小一华文口语能力诊断评量表”的开发历经两个阶段:评量表语言能力指标的确立和评量表效度、信度、可操作性的建立。

具体来讲,本研究的研究目的有三:

第一,如何确立“新加坡小一华文口语能力诊断评量表”的语言能力指标?

第二,“新加坡小一华文口语能力诊断评量表”的效度和信度如何?

第三,“新加坡小一华文口语能力诊断评量表”的可操作性如何?

三、研究方法

3.1 研究对象

不同阶段由于研究目的不同,所参与的对象也不尽相同。具体研究对象的信息如表 2 所示。

表2 研究对象信息

阶段	学校数量	班级数量	教师数量	学生数量	被评估学生数量
阶段一	4	16	16	400	184
阶段二	12	20	36	500	210

3.2 数据收集与分析过程

本研究的第一阶段于2010年2月开始进行。首先,研究组分别对新加坡四所小学一年级的16个班的学生的华文课学习进行观察,同时随机抽取和收集了对其中184名学生一对一的访谈语料,并从不同角度对这些学生华文口语表现的特征进行分析、评估和总结,从而为建构评量表的语言能力指标提供依据。其次,研究组开放网上平台调查新加坡小学一年级教师对学生华文口语能力表现的看法(见表3),以期为评量表的语言能力指标的确定提供参考。通过将这两种渠道获得的数据汇总,初步建构评量表的雏形。

表3 教师在线问卷调查

您认为以下哪些方面是学生口语能力的表现?	参与教师人数(共44人)
学生在说话或回答问题时能运用课堂上所学的句型、词汇等	35(79.5%)
学生能用正确的华语在课堂中与同学进行交谈、讨论	41(93.2%)
学生能以正确的发音与声调朗读或念出课文	23(52.3%)
学生能用正确的华语在课堂上回应老师	39(88.6%)
学生能在日常生活中与家人或其他人以华语进行交谈	31(70.5%)
学生能对设定的物品、图片或情景进行描述	38(86.4%)

本研究的第二阶段始于2011年4月,另外12所新加坡小学一年级的20个班级和教师参与试用初始版本的“新加坡小一华文口语能力诊断评量表”。研究组通过问卷收集有关评量表效度的反馈信息,并对评量表的语言能力指标进行相应调整。评量表的信度共经历了两轮测试。在评量表的试用阶段,研究组比较了20位参与教师分别与项目组评分员之间对196名学生课堂华文口语表现评分的一致性。在评量表经过调整之后,项目组又比较了另外16位教师通过观看录像对14名学生的课堂华文口语表现评分的一致性。同时,通过比较每位教师的评分与标准分的一致性,评量表的可操作性也得以证明。

本研究采用了定性和定量分析相结合的方法。其中,评量表效度的建立主要依靠定性分析的方法;而第二阶段对评量表效度、信度和可操作性的验证则主要借助定量的分析手段。

四、结果与讨论

4.1 效度的建立与验证

通过对第一阶段 184 名学生访谈语料的分析,研究组从语音、词汇、句法、交际表达、偏误等角度^①对新加坡小一学生的华文口语表现特征进行归纳和总结后将其划分为四个发展期——萌芽期、展叶期、开花期和幼果期。

教师在线问卷调查所得的数据清楚显示,小一华文教师认为口语能力基本涵盖词汇、句型、发音及回应、描述、讨论等不同交际功能。

据此,评量表中口语能力评量的主要内容基本锁定在词汇、语法句型、语音语调和交际表达^②。这四个层面的语言能力指标项(indicator)和描述项(descriptor)^③也得以初步建构。表 4 概括了评量表的主要内容。从中不难发现,除了上文提到的将口语能力发展划分为四个发展期、将口语表现分割为四个层面之外,各个层面的指标项在每个发展期保持不变。例如,对词汇表现的评量在四个发展期中都是从用词的丰富度角度进行的。而四个发展期的口语表现特点则是由描述项体现^④。另外,为了便于教师细化操作,四个发展期又分别涵盖下、中、上三个阶段。这样的设计又为教师鉴别同一发展期学生口语表现的强弱面提供了更可靠的依据。

表 4 新加坡小一华文口语能力诊断评量表(简表)

发展期 构成	萌芽期 (下—中—上)	展叶期 (下—中—上)	开花期 (下—中—上)	幼果期 (下—中—上)
词汇	丰富度	丰富度	丰富度	丰富度
语法句型	语法、词序 句型变化	语法、词序 句型变化	语法、词序 句型变化	语法、词序 句型变化
语音语调	语音语调句调	语音语调句调	语音语调句调	语音语调句调
交际表达	听、回应 停顿与话轮 提问 语意连贯	听、回应 停顿与话轮 提问 语意连贯	听、回应 停顿与话轮 提问 语意连贯	听、回应 停顿与话轮 提问 语意连贯

① 这些角度的选取参照了现有主要口语评量表的结构,如 ACTFL、CEFR 等。

② 在此,“词汇”“语法句型”和“语音语调”被界定为语言能力;“交际表达”被界定为交际能力。

③ 指标项是指能明显标示和观测口语能力发展特点的要素。这些要素以不同量值标注,量值的不同范围界定被评估者的语言能力程度;描述项是指对指标项在四个发展期中的解释。

④ 例如词汇层面的“丰富度”指标项在四个发展期的描述分别以“贫乏”“不丰富”“较丰富”“丰富”等表述进行区别。

在教师试用评量表后,研究组再次调查了教师对评量表内容效度的态度,即“您觉得诊断工具在词汇、语法句型、语音语调和交际表达层面的各个指标项是不是能够清楚区分和描述各等级学生的能力表现?”。如表 5 所示,教师们分别三次在不同主题活动中试用评量表,每一次绝大多数教师都认可评量表在四个层面上的语言能力指标能够明确显示和区别各等级学生的口语能力表现。具体来说,在第一个主题活动中试用后,分别有 85%(17 人)、75%(15 人)、90%(18 人)和 80%(16 人)的教师肯定评量表在词汇、语法句型、语音语调和交际表达四个方面的有效性;在第二个主题活动中试用后,同意评量表的四个层面能有效反映和显示学生口语能力差异的教师的比例分别为 94.1%(16 人),88.2%(15 人),94.1%(16 人)和 82.4%(14 人);同样,通过第三个主题活动的试用,调查结果表明赞成评量表在四个层面上的内容有效性的教师比例分别为 80%(16 人),80%(16 人),85%(17 人)和 75%(15 人)。

表 5 评量表内容效度教师态度调查

诊断层面 试用次数	词汇			语法句型			语音语调			交际表达			教师人数
	是	否	—	是	否	—	是	否	—	是	否	—	
第一次	17	2	1	15	4	1	18	1	1	16	3	1	20
第二次	16	1	0	15	2	0	16	1	0	14	3	0	17
第三次	16	2	2	16	2	2	17	0	3	15	3	2	20

表 5 还显示,综观三次不同主题活动的试用反馈,评量表四个语言层面中“语音语调”和“词汇”的各个指标项对帮助教师评量学生的口语能力表现更加有效。如,三次试用分别有 90%(18 人),94.1%(16 人)和 85%(17 人)的教师赞成“语音语调”各个指标项的有效性;同时也有 85%(17 人),94.1%(16 人)和 80%(16 人)的教师认为“词汇”层面的指标项更能有效帮助他们诊断学生的口语能力。持否定态度的教师的意见主要是描述项的表述不够详细,建议提供实例进行说明。

4.2 信度的建立与验证

如前所述,评量表的信度历经两次测试,第一次在评量表试用期间进行。20 位教师分别与一位项目组评分员在课堂活动中使用评量表对学生的口语表现进行实时诊断。参与诊断的项目组评分员已事先接受了针对评量表使用的专门培训。并且他们之间在词汇、语法句型、语音语调、交际表达和综合表现方面的评分一致性分别为 0.849、0.862、0.820、0.868、0.882。

表 6 显示,教师与项目组评分员之间评分的组内相关系数的显著性比率在词汇、语法句型、语音语调、交际表达和综合表现方面分别是 65%、70%、70%、80%、80%。在

评分显著一致的范围,大部分 ICC 值都在 0.700 以上^①。例如,在词汇方面,11 组教师—评分员的评分一致性在 0.700 以上,甚至有 5 组处在 0.900 以上,只有两组低于 0.600。在语法句型、语音语调、交际表达和综合表现方面,评分一致性达到 0.700 以上的组数分别是 11、12、14 和 16。而低于 0.700 的组数分别是 3、2、2 和 0。

表 6 教师—项目组评分员评分一致性归总

教师数量	ICC	词汇	语法句型	语音语调	交际表达	综合
		显著比率 65%	显著比率 70%	显著比率 70%	显著比率 80%	显著比率 80%
20	<0.600	2	0	2	1	0
	0.600~0.699	0	3	0	1	0
	0.700~0.799	4	1	5	2	2
	0.800~0.899	2	4	4	7	9
	0.900~1	5	6	3	5	5

ICC = 组内相关系数

试用阶段的测试表明,教师—项目组评分员的评分在交际表达和综合评估两项的一致性很高,在词汇、语法句型和语音语调三项的一致性也很明显。这一点不仅可以从 ICC 值的显著性比率看出,达到 0.800 以上的 ICC 值数量也显示教师—项目组评分员在交际表达和综合表现方面的评分一致性很高。这主要归结于很多教师在理解评量表中发展期尤其是“幼果期”的语言能力描述时存在分歧。例如,他们在访谈中有的表示对某些指标项的描述还不够熟悉、把握得不够准确;有的则表示很多交际表达和综合表现都很好的学生在用词、语法准确性或语音语调上不一定就达到很高的程度。所以,这一阶段的信度测试显示,评量表的信度虽然已经得到较好的建立,但还有进一步提升的空间。

表 7 教师评分一致性归总

教师数量	口语样本数量	词汇 ICC	语法句型 ICC	语音语调 ICC	交际表达 ICC	综合 ICC
16	14	0.724	0.729	0.705	0.737	0.752

ICC = 组内相关系数

* ICC 值都显著 ($P < 0.001$)

在根据试用教师和专家的意见对评量表的期别划分,尤其是“幼果期”的界定进一步调整之后,项目组对其信度进行第二轮测试,即,由另外 16 名教师利用量表对录像中

^① 根据国际研发惯例,ICC 值在 0.700 以上即可证明一致性很高(Novak, Herman, & Gearhart, 1996)。

14 个口语样本进行实时评价。表 7 明确显示所有 ICC 值都显著($P<0.001$),这与信度测试一相比,是一个明显的提高。除此之外,教师们在词汇、语法句型、语音语调、交际表达和综合表现方面评分的一致性都达到 0.700 以上。由此可见,调整后的评量表能有效地帮助教师诊断学生的口语表现在各个层面上的强弱。

4.3 可操作性的验证

“新加坡小一华文口语能力诊断评量表”的研发原则之一便是“易用”,因此,项目组通过比较以上 16 名教师的评分与标准分之间的一致性对评量表的可操作性进行了验证。从表 8 中不难发现,首先,所有 ICC 值都具有显著性。其次,16 名教师中大部分(12 人、12 人、15 人和 14 人)的评分在词汇、语法句型、交际表达和综合表现上与标准分的一致性达到 0.700 以上。这表明在经过了将近两天的培训之后,大部分教师都能很好地使用评量表对口语表现进行有效评估。但同时数据还显示,在词汇、语法句型和语音语调层面分别有 4 名、4 名和 10 名教师的评分与标准分的一致性低于 0.700。尤其是在“语音语调”这一层面,虽然根据试用阶段教师的反馈,这个层面的能力指标能够清楚区分和描述各发展期学生的能力表现(见表 5),但在实践中对很多教师而言似乎还无法准确把握。这个结果一方面与很多以往的研究相近,如在 Hendricks 等人的研究(1980;cf. Fulcher, 2003)中评分员对语音语调评分的一致性仅为 0.43;另一方面也说明,相对于交际表达,教师们对评量表中语言能力各个层面指标的熟练掌握还需要更多时间和操练,还需要更多专业性的指导。

表 8 教师评分与标准分一致性归总

教师数量	ICC	词汇	语法句型	语音语调	交际表达	综合
16	<0.600	0	2	4	0	0
	0.600~0.699	4	2	6	1	2
	0.700~0.799	4	4	3	5	3
	0.800~0.899	6	8	3	8	8
	0.900~1	2	0	0	2	3

ICC = 组内相关系数 * ICC 值都显著 ($P<0.001$)

五、结束语

“新加坡小一华文口语能力诊断评量表”对新加坡华文教学具有深远的实践意义。

首先,该诊断评量表能辅助教师更客观有效地了解学生华文口语能力不同侧面的强弱。这不仅可以让教师对学生的起点和学习进程一目了然,并且能够促使教师对教学过程做更现实的反思,从教学内容到教学策略上进行更具有针对性的调整,从而提高学生的华文口语能力。其次,通过使用该诊断评量表,教师将能够对口语表现的方方面面、口语能力的发展轨迹乃至口语教学具有更深刻的认识。这对新加坡华文教师的专业提升无疑大有裨益。

作为华文口语的评量工具,“新加坡小一华文口语能力诊断评量表”为评量华文为第二语言学习者,尤其是初级水平学习者的口语能力提供了有效参考和可靠依据。其研发过程也对华文口语能力评估工具的进一步开发具有借鉴意义。

第一,就其性质而论,本评量表是诊断性评估的有效工具。它超越了绝大部分现存的仅止步于界定学习者口语能力级别的评量表,更加深入地挖掘学习者语言使用中的优势和缺陷。这种诊断信息的价值在于它能够使语言教育工作者意识到不同背景的第二语言学习者语言学习成果的差异性,并进一步思考如何在第二语言的教学乃至教育政策的制定和实施中进行调整。因此,本评量表的研发不仅是诊断性评估实施和推广的前提,也为今后诊断性评估工具的研发抛砖引玉。

第二,从评分方式上来讲,该评量表采用综合评分和分项评分相结合的方式。不仅提供学习者口语能力总体水平的信息,更反映了他们在词汇、语法、发音、交际表达各方面的具体表现。对教师而言,这种评分方式既方便他们了解全体学生口语能力水平的比例和分布,也有助于他们认清每个学生口语能力不同层面的强弱表现。较之于大多数仅限于综合评分的评量表而言,本评量表无疑更具有教育价值。

第三,从其自身有效性的验证来看,本评量表的构建主要立足于真实的口语语料,同时参考专家及一线教师的意见。这不仅在内容和架构上确保了评量表口语能力指标设立的理论性和完整性,还使其更符合新加坡本地华文第二语言初学者华文使用的特色。这种基于真实语料,而不单纯依赖经验和直觉建构的评量表才能为评估者和教师提供真实、客观、可信的评估信息。

第四,就其使用操作而言,本评量表专为教师于课堂内对学生口语产出进行实时诊断所用,即教师针对目标学生在特定口语活动^①中的口语产出进行诊断。毫无疑问,口语是交际中使用的语言,因此口语评估绝不应该脱离交际情景。实际上,很多口语测试都已设法融入交际情境元素,如FSI考试的OPI(Oral Proficiency Interview)、HKC考试的“问答”等等。正是基于这样的考虑,本评量表旨在帮助教师判断学生身处具体

^① 这里指根据课本教学内容设计的互动性活动。

情境、针对具体话题时口语产出的特点。该种操作方式不仅不占用教师额外的时间和精力,而且所生成的诊断信息对教师、对学生及家长而言都更客观、更具体、更可信。因此,这是对以往借助笔试、录音或访谈的口语测试形式的一次突破性尝试。

综上所述,“新加坡小一华文口语能力诊断评量表”的成功建立对新加坡乃至国际华文口语评量工具的研发及华文第二语言口语教学都具有不可低估的价值和启示意义。

参考文献

- 韩宝成 (2006) 国外语言力量表述评,《外语教学与研究》第 6 期。
- 李光耀 (2011) 《我一生的挑战——新加坡双语之路》,新加坡《联合早报》。
- 新加坡教育部 (2010) 《乐学善用——2010 年母语检讨委员会报告书》,新加坡教育部。
- Alderson, J. C., Clapham, C. M. and Wall, D. (1995) *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.
- Bachman, L. F. and Palmer, A. S. (1996) *Language Testing in Practice*. Oxford: Oxford University Press.
- Fulcher, G. (2003) *Testing Second Language Speaking*. Britain: Pearson Education Limited.
- McGuinness, D. (2006) *Language Development and Learning to Read*. MA: MIT Press.
- Novak, J. R., Herman, J. L. and Gearhart, M. (1996) *Issues in Portfolio Assessment: The Scorability of Narrative Collections*. CSE Technical Report 410. Los Angeles: National Center for Research on Evaluation, Standards and Student Testing (CRESST), University of California.

(279623 新加坡,新加坡南洋理工大学新加坡华文教研中心)

论高级阶段修辞教学*

——以比喻手法的输入与输出为例

肖奚强 颜 明^①

摘 要:文章以比喻手法的输入与输出为例讨论高级阶段汉语修辞教学问题。在介绍《多文体精泛结合高级汉语教程》修辞方面的编写理念与实践以及比喻的教学输入的基础上,分析外国留学生的输出。发现留学生对比喻次类及与其他辞格的兼用、连用的使用不仅仅是简单的模仿而是能动地创造性地产生输出。认为丰富有效的教学输入对学生的理解和输出有着正面、积极的引导和促进作用;学生的输出不等同于输入,往往大于输入;输出既是教学输入的证据,也是学生业已内化的语言知识的证据。

关键词:高级阶段;修辞教学;教材编写;比喻手法;教学输入;输出

缘 起

在众多的对外汉语教材以及教学实践中,修辞教学几近阙如。我们自新世纪立项建设高级汉语课程以来,在这方面进行了长期的探索与实践。此间,我们改革了课程教学内容和教学方法,立项编写并出版了配套的教材——《多文体·精泛结合·高级汉语教程》,并将该课程建设成校级精品课程。

虽然自20世纪90年代后期以来,陆续有人撰文呼吁要重视修辞教学;但迄今为止发表在核心以上期刊上的此类论文数量仅寥寥十数篇,且多为宏观讨论修辞教学的重

* 本文得到国家社科基金、教育部哲社基金、江苏省优势学科建设项目、江苏省研究生精品课程建设项目、南京师大优秀教学团队项目的资助。初稿曾在第十一届高校对外汉语教学国际研讨会(2013.7.3—7.5 英国·诺丁汉大学)上宣读。

① 肖奚强,男,语言学博士,教授,对外汉语专业博士生导师,南京师范大学国际汉语教育研究所所长。主要研究领域为现代汉语语法、语言教学与习得理论。颜明,女,对外汉语专业博士生,主要研究领域语言教学与习得理论。