

潘之欣 著

交际性 听力理解考试的开发

—— 全国大学英语四、六级考试听力理解部分改革探索

DEVELOPING A COMMUNICATIVE LISTENING
COMPREHENSION TEST:
AN EXPLORATION INTO REVISING THE
LISTENING COMPREHENSION PART OF THE
CET-4 AND -6

上海交通大学出版社

交际性听力理解考试的开发

——全国大学英语四、六级考试听力理解部分改革探索

DEVELOPING A COMMUNICATIVE
LISTENING COMPREHENSION TEST:
AN EXPLORATION INTO REVISING THE LISTENING
COMPREHENSION PART OF THE CET-4 AND -6

潘之欣 著

上海交通大学出版社

图书在版编目(CIP)数据

交际性听力理解考试的开发: 全国大学英语四、六级
考试听力理解部分改革探索/潘之欣著. —上海: 上海
交通大学出版社, 2007

ISBN 978-7-313-04550-8

I. 交... II. 潘... III. 英语—听说教学—高等学
校—水平考试—教学改革—研究 IV. H319.9

中国版本图书馆 CIP 数据核字(2006)第 096687 号

交际性听力理解考试的开发

——全国大学英语四、六级考试听力理解部分改革探索

潘之欣 著

上海交通大学出版社出版发行

(上海市番禺路 877 号 邮政编码 200030)

电话: 64071208 出版人: 张天蔚

常熟市华通印刷有限公司印刷 全国新华书店经销

开本: 787mm×960mm 1/16 印张: 17.5 字数: 330 千字

2007 年 1 月第 1 版 2007 年 1 月第 1 次印刷

印数: 1—3050

ISBN 978-7-313-04550-8/H·633 定价: 25.00 元

版权所有 侵权必究

前 言

语言考试开发与应用语言学理论密不可分。作为应用语言学理论在实践中的具体运用,语言测试必然包含着一定的语言观和语言学习观,考试开发者在设计考试时必须考虑采纳当前被广泛接受的关于语言、语言能力和语言使用本质的观点(Alderson 和 Clapham 1992:149)。语言学学科在不断发展,语言考试设计不可能一劳永逸。随着新理论的出现,原先考试所依托的语言理论渐显不足,语言考试总是需要不断地修订、改革,以寻求更合理的理论基础。近期一个典型的例子就是国际上好几个英语考试都根据扩展了的语言能力观进行了改革。

20 世纪 80 年代,Dell Hymes(1979),Henry Widdowson(1978;1983),Sandra Savignon(1983),Michael Canale 和 Merrill Swain(1980)等语言学家拓展了语言能力的内涵,建立了交际能力(communicative competence)概念。语言使用是“在情景语境中对意义的协商”,语言能力是由多个成分构成的,不仅包括语法能力,还包括篇章能力、功能能力和社会语言能力(Bachman 2000:3)^①。不久,有关交际语言测试法(communicative language testing)的讨论也同时展开,并一直延续至今。概括地讲,交际测试有两大根本特点。其一,交际测试的目的是测试考生的交际语言能力,即有效应用语言完成特定真实交际任务的能力。其二,交际测试运用直接的(direct)或称真实的(authentic)测试方法,要求让学生完成的考试任务应尽可能地接近目标任务。尽管目前对交际测试法中的直接测试、测试真实性的本质还有很大争议(Bachman 2000:3),许多国际重要语言能力考试(如 IELTS, TOEFL, UETESOL 等)都在 20 世纪 90 年代进行了改革,在考试设计中增加了对语言使用中篇章特点、社会语言特点、情景语境等方面的考虑。进入 21 世纪,我国的全国大学英语四、六级考试(简称 CET 考试)也开始了对交际性听力理解考试开发的研究。

CET 考试是由教育部高等教育司主办的大规模标准化全国性统一考试。它根据《大学英语教学大纲》的要求,对我国大学生的实际使用英语的能力进行客观、公正、准确的评价(杨惠中和 Weir,1998)。由于其命题的科学性、评分的一致

^① Bachman 2000:3 表示作者是 Bachman,出版年份是 2000,引文内容在第 3 页。这个意思以后不一一说明。

性、组织的严密性和成绩的可比性,CET 考试对推动大学英语教学改革、实施大学英语教学大纲和提高教学质量做出了很大贡献,并已得到社会承认,成为社会人才选拔中英语水平的重要衡量标准,目前每年考生上千万。CET 考试 1986 年设计,1987 年使用。那时在中国,结构主义测试已经开始衰落,综合测试正处于发展巅峰,而交际测试则刚刚起步。因此 CET 考试融合了这三种测试方法,但以综合测试法为主,个别部分采用结构主义测试法。考试包括四个以意义为纲的部分:听力理解、阅读理解、完形填空(四级)或综合改错(六级)及短文写作,另外再辅之以语言系统为纲的词汇和语法部分(具体考试结构见表 1)。

表 1 大学英语四、六级考试结构(《大学英语四、六级考试大纲及样卷》,1994)

试卷号	序号	题号	部分名称	回答类型	题目数	计分	考试时间
一	I	1~20	听力理解	标准多项选择	20	20	20
	II	20~40	阅读理解	标准多项选择	20	40	35
	III	41~70	词汇和语法	标准多项选择	30	15	20
	IV (四) (六)	71~90	完形填空	标准多项选择	20	10	15
		71~80	综合改错	短表达	10		
二	V	91/81	短文写作	长表达	1	15	30
合 计					91/81	100	120

其中听力理解部分是根据 1985 版《教学大纲》的相关规定设计的,目的是“测试学生获取口头信息的能力”。考试时间 20 分钟,共 20 题,总分 20 分,包括两节: A 节有十个短对话(20~30 字),对话为两个说话人之间的一次邻接应对(adjacency pair),每个对话后有一个问句;B 节有三篇短文(约 200 字),每篇后有三或四道问题。所有录音材料只播放一遍。每个问句后有短暂的间隙,要求考生从试卷所给出的四个选择项中选出一个最佳答案。对话部分为日常生活中的关于一般话题的对话,短文部分为关于日常生活、社会文化、科普知识的故事、讲话、叙述等。四级平均语速为每分钟 120 词,六级为每分钟 130 词。整个考试以信息为纲,主要测试考生理解重要或特定细节、推断隐含意义、获取篇章大意的能力。1999 年听力短文部分又增加了复合式听写新题型,与原有的回答问题题型交替使用。复合式听写含一篇短文,其录音播放三遍,要求学生收听录音时在该短文的文稿中填空。其中的前七个空,要求考生用所听到的原词填空,每个空一个词,考核正确掌握词汇包括拼写、词义和用法的能力;后三个较长的空,要求考生用听到的原文或用自己的语言填入有关的信息,每个空一个句子,考核在正确理解的基础上表达所要求信息的能力。

现在,交际教学和测试发展迅猛,对交际语言能力和语言使用的理解更加深入,系统引入交际测试法的时机已经成熟。更重要的是,一个交际性的 CET 考试更加符合当前的语言观,能够更有效地考察、预测学生在现实生活中运用语言的能力,从而更好地服务高校教学和社会人才选拔。此外,交际性考试被应用语言学家公认为具有最佳的反拨效应(washback effect)(Alderson 和 Wall 1993: 116),Alderson(1991:16)还指出在某些情况下可以利用考试对教学的反拨效应引发教学改革。基于 CET 考试对大学英语教学的巨大影响力,一个交际性的 CET 考试必将加快高校课堂教学内容和方法的更新,促进交际语言教学的推广应用,加强对学生的语言交际能力的培养。

基于上述考虑,运用最新交际测试方法改革 CET 考试已成为当务之急。作为改革的重要组成部分,2000 年 12 月全国大学英语四、六级考试委员会批准立项了“交际性听力理解考试的开发——全国大学英语四、六级考试听力理解部分改革探索”的重要科研项目,目的就是为 CET 考试设计一个交际性的听力理解考试,测试高校大学生的交际听力能力,即大学生在他们于在读期间及毕业后可能遭遇的交际任务中从口头语篇中获取信息的能力,从而在保证该部分的高信度和可行性的基础上,进一步提高其测试的直接性和效度。在 2002、2003 两年中,本人有幸作为主要研究人员参与了整个项目的实施,本书正是根据项目研究成果撰写而成,详实记载了项目依据交际测试原理,全面、系统地采用 Bachman 和 Palmer 在 *Language Testing in Practice* (1996) 中提出的考试开发方法,设计一个新的交际性 CET 听力理解考试(CLCT)的全过程。CLCT 考试的设计开发遵循 Bachman 和 Palmer 提出的两大根本原则:一是通过运用任务特点描写系统和语言能力描写系统实现考试中的语言使用与非考试中的语言使用的相互一致;二是通过六大考试特性——效度、信度、真实性、交互性、后效作用和可行性——之间的合理平衡实现测试整体实用性的最大化。研究需要解决两个关键问题:一个是理论问题,一个是实践问题。

理论问题指建立交际性听力理解测试理论系统,这是本项目的理论基础。Swain(1985:36)曾指出交际语言测试的一条普遍原则就是考试开发应该建立在一定的理论框架之上。Alderson(1991:14)认为语言测试发展缓慢的主要原因之一就是缺乏一个统一、连贯的体系,如果语言测试界使用共同的理论体系,各种想法就可以互相借鉴,使得系统化研究体系中某一项目及其与体系中其他项目的关系成为可能。他指出 Bachman (1990)提出的理论模型在一定程度上解决了这个问题,为广大测试研究人员提供一个有益的起点,今后应该努力将其中的各部分运用到实际中去,以探索它们之间的关系。Chalhoub-Deville (1997:16)也认为 Bachman 提供的只是一个普遍的理论系统,代表适用于不同考试情况下的理论基

基础系统。我们还需要一个可操作的系统来描述适用于特定考试情况下的理论基础系统。例如, Weir (1993) 提出了一个与 Bachman 理论相类似的, 由三部分组成的系统, 包括活动(技能)、实施条件和期望实施水平。他不同于 Bachman 的贡献在于他将这一系统运用到听、说、读、写四个技能上去, 建立了四个不同的, 分别适用于这四个技能的更加具体的体系。本研究的目标之一就是建立一个比 Bachman 的一般系统和 Weir 的听力系统更加全面具体的交际性听力理解测试理论框架。

实践问题则有关控制相互矛盾的考试特性, 寻求它们之间的合理平衡, 以实现考试整体实用性的最大化。众所周知, 不同考试特性, 尤其是信度和效度之间确实存在着矛盾, Davies (1988: 7) 坦言, “信度的限制对交际性考试提出的挑战远远大于对语法考试的挑战, 因为后者可以做成完全客观的考试”。Morrow (1991) 在解释 CCSE 考试设计时也谈到交际测试中信度和效度之间的矛盾:

可以采用许多办法来提高 CCSE 考试的信度, 但是它们都和考试真实性原则有直接冲突。所做出的抉择反映出所持的基本教育理念。一种观点认为, 信度是最关键的, 应该尽可能扩大, 而真实性终归是第二位的。另一种观点则认为, 真实性是最根本的特性, 尽管已经意识到考试任务的完全真实性是不可能达到的, 但是所有的努力还是应该集中在真实性上。信度不能忽略, 必须努力保证信度得到足够的重视, 但是在考试设计中, 最终它并不是首要决定因素。

最后他承认“原则上这种矛盾似乎是无法解决的”。CET 考试是全国性统一考试, 目前每年全国考生上千万, 因此它受到的信度和可行性限制比其他的公共考试大得多。因此本研究的另一个目标就是在效度、反拨效应和信度、可行性之间找到最佳平衡, 开发出具有更高整体测试实用性的 CET 听力理解考试。

整个项目分三阶段进行。第一阶段是理论基础研究阶段, 目标是为交际性听力理解考试构建理论框架系统。研究首先对交际性语言测试, 尤其是 Bachman 和 Palmer (1996) 提出的考试开发方法, 进行了评论性综述, 从而明确了交际性语言测试的特点、原则、理论框架、考试开发过程等研究的一般理论基础。然后, 分别从社会语言学和心理语言学两方面探究实际生活中的听力行为的特点。最后综合认知心理学对交际语言能力的解释以及心理语言学中关于听力理解心理过程的理论, 建立了交际性听力能力模型; 同时依据社会语言学中的交际事件构成框架和应用语言学相关文献, 建立了交际性听力任务特点的描写框架。

第二阶段目标是利用所建理论框架系统确立 CLCT 考试规范。为此研究人

员进行了两项前期研究工作。一是通过对大学生英语听力需求的问卷调查,掌握大学生在在读期间及毕业后实际可能遭遇的交际听力任务,从而确定 CLCT 目标语言使用领域和任务。二是通过对影响广泛的国内外听力教材和大规模考试中听力任务的分析,掌握当前听力教学和测试中的普遍做法以及不同交际事件类型和不同听力活动类型的特点。根据研究结果,确定 CLCT 听力任务的特点,并最终形成 CLCT 考试规范,力保四、六级考试间有显著难度差异,并具有尽可能大的测试整体实用性。

第三阶段目标是对所设计的 CLCT-4 和 CLCT-6 进行难度和实用性检验。鉴于人力物力限制,项目仅对新增设的笔记填空(简称 CLCTNF)题型部分实施检验。研究人员制作了两个 CLCTNF 样卷各一套,并实施了一系列后期验证实验,如 CLCTNF-4 和 CLCTNF-6 试测、CLCTNF-4 和 CLCTNF-6 对比研究、CLCTNF 与现行 CET 对比研究以及 CLCTNF-4 和 CLCTNF-6 考生与监考教师反馈问卷调查。研究结果表明 CLCTNF-4 和 CLCTNF-6 新题型具有相当高的效度、信度和可行性,并表现出显著的难度阶梯,一个实用的交际性听力理解考试初步开发成功。当然,在 CLCT 正式为 CET 考试所采用前,还需要进一步完善样题,实施更大规模的试测,进行更全面的后期效度检验。

对应于项目的三个阶段,本书可分为三大部分。第一部分包括头三章,是交际性听力理解测试基础理论部分。其中,第一章介绍交际测试法的一般特点和原则以及 Bachman 和 Palmer (1996)的考试开发方法;第二章介绍实际生活中听力行为的心理和社会特点;第三章是以此为基础所建立的交际性听力理解考试理论框架系统。第二部分叙述了在 CLCT 考试规范建立阶段实施的几项实证研究。第四章记述了大学生英语听力需求问卷调查;第五章记述了国内外听力教材和大规模考试中听力任务分析;第六章描述了通过综合这些研究结果所建立的 CLCT-4 和 CLCT-6 考试规范。第三部分为第七章,汇报了对所设计的 CLCT-4 和 CLCT-6 中新增设的笔记填空题型的难度和实用性的检验方法和结果。最后第八章是全书的总结。

语言测试,包括课堂测验、学期考试、大规模标准化考试等等,是语言教学中必不可少的环节、人才选拔中广泛使用的手段。语言测试不是试题的随意拼凑,而是一门科学、一门艺术,科学严谨的考试开发和试卷编撰是极其复杂的过程,必须遵循一定的原则、一定的程序。本书将语言测试理论与语言考试开发实践相结合,是考试设计的一个成功范例,而且全书结构清晰,条理性强,用词准确,数据详实,可为各类语言考试的设计开发者和研究者开拓思路,也可作为广大语言教师的案头参考书。书中所建立的交际性听力理解考试理论框架系统,理论基础扎实、广博,对交际性听力考试的开发和交际听力教学的研究均有参考和借鉴价值,

可广泛用于指导听力课堂教学、教材编写和考试设计。书中记述的大学生英语听力需求分析结果、国内外听力教材和大规模考试听力任务分析结果也为大学英语听力考试开发、听力课堂教学、听力教材编写提供了有价值的信息。书中记录的努力提高听力考试整体测试实用性的过程对其他大规模考试的听力考试开发也有参考价值。

在项目进行和本书撰写过程中,原全国大学英语四、六级考试委员会主任委员、项目负责人、我的导师杨惠中教授给予了全程悉心指导;现任委员会主任委员金艳博士提供了大量具体建议和帮助;委员会中负责听力测试的董亚芬教授和夏国佐教授帮助审订了 CLCT 考试规范和样卷;全国大学英语四、六级考试委员会办公室工作人员朱正才博士、罗鹏博士、杨浩然工程师在数据整理、统计方面提供了帮助;上海交通大学的博士生王跃武、辜向东等参加了问卷调查、阅卷,在此向他们表示衷心的感谢。最后,还要感谢大力支持项目中问卷调查和 CLCT 试测的上海交通大学、复旦大学、上海师范大学、同济大学、华东理工大学、华东师范大学、东华大学、上海海运学院、上海外贸学院、上海第二医科大学的广大师生。

潘之欣

2006 年 1 月

目 录

第一章 交际语言测试——评论性综述	1
1.1 语言测试方法	1
1.1.1 结构主义测试法	1
1.1.2 综合测试法	2
1.1.3 交际测试法	2
1.2 交际测试合成法	5
1.2.1 实际生活法和交互能力法	5
1.2.2 考试的实用性	6
1.2.3 考试语言使用与非考试语言使用之间的一致性	11
1.2.4 考试开发流程	23
1.3 小结	25
第二章 实际生活中的听力行为——跨学科综述	27
2.1 导言	27
2.2 听力行为是一种社交活动	31
2.2.1 语言交际	31
2.2.2 口头交际	36
2.3 听力行为是一种认知活动	42
2.3.1 语言理解	42
2.3.2 听力理解	49
2.4 小结	51
第三章 交际性听力测试——理论系统	52
3.1 交际性听力能力	52
3.2 交际性听力任务	57
3.2.1 考试规则	58
3.2.2 情景语境的特点	59
3.2.3 输入语篇的特点	64

3.2.4 听力活动的特点	69
3.2.5 期望回答的特点	77
3.3 小结	81
第四章 大学生英语听力需求分析	84
4.1 研究目的	84
4.2 研究方法	85
4.2.1 调查问卷	85
4.2.2 调查对象	87
4.3 研究结果	88
4.3.1 英语听力的相对重要性	88
4.3.2 重要的英语听力使用领域	89
4.3.3 重要的英语听力使用情景和任务	91
4.3.4 对 CET 听力理解考试的建议	95
4.4 启示	96
第五章 英语听力教材和考试任务分析	100
5.1 研究目的	100
5.2 研究方法	101
5.3 研究结果	101
5.3.1 考试规则	101
5.3.2 情景语境	102
5.3.3 馈入语篇	104
5.3.4 听力高等技能	106
5.3.5 听力活动	108
5.3.6 期望回答	111
5.3.7 交际事件类型研究	112
5.3.8 听力活动类型研究	115
5.4 启示	118
第六章 CLCT 考试规范	122
6.1 设计说明	122
6.1.1 考试目的	122
6.1.2 受试人群	123

6.1.3 目标听力领域和任务	123
6.1.4 目标能力构成	124
6.1.5 考试实用性特点	125
6.2 CLCT 考试任务的总体特点	128
6.2.1 考试规则	128
6.2.2 情景语境	129
6.2.3 馈入语篇	130
6.2.4 听力活动	131
6.2.5 期望回答	133
6.3 考试规范	134
6.3.1 标准型 CLCT 考试规范	134
6.3.2 过渡型 CLCT-4 和-6 考试规范	136
6.4 小结	143
第七章 CLCT 考试实用性检验	145
7.1 试卷编制	145
7.1.1 听力语篇选材	145
7.1.2 听力活动设计	147
7.1.3 构卷	148
7.1.4 小规模试测	148
7.1.5 修改	149
7.2 考试实施	150
7.2.1 施考	150
7.2.2 阅卷	152
7.2.3 CLCTNF 试题分析	155
7.2.4 CLCTNF-4 与-6 对比研究	157
7.2.5 CLCTNF 与现行 CET 对比研究	159
7.2.6 考生与监考教师反馈问卷调查	163
7.3 小结	169
7.3.1 考试难度和考试实用性	169
7.3.2 对今后试卷编制的建议	171
第八章 结论	173
8.1 交际性听力理解测试理论系统	173

8.2 CLCT 考试及其整体实用性	177
附录	184
附录 1 听力促成技能分类	184
附录 2 大学生英语听力需求分析(在读生问卷)	192
附录 3 大学生英语听力需求分析(毕业生问卷)	194
附录 4 听力任务分析英语听力教材一览表	196
附录 5 听力任务分析英语听力考试一览表	197
附录 6 英语听力教材任务分析结果	198
附录 7 英语听力考试任务分析结果	205
附录 8 标准型 CLCT 样卷	214
附录 9 CLCTNF-4 与-6 小规模试测	230
附录 10 CLCTNF-4 样卷	231
附录 11 CLCTNF-6 样卷	237
附录 12 CLCTNF-4 样卷评分标准	242
附录 13 CLCTNF-6 样卷评分标准	245
附录 14 CLCTNF-4 试测	248
附录 15 CLCTNF-6 试测	250
附录 16 CLCTNF-4 与-6 对比研究	252
附录 17 CLCTNF 表面效度研究(考生问卷)	254
附录 18 CLCTNF 表面效度研究(监考教师问卷)	256
参考书目	258

第一章

交际语言测试——评论性综述

1.1 语言测试方法

考试设计的灵魂在于该考试所体现的对语言和语言使用的认识。随着语言学与应用语言学理论的不断进步,在语言测试发展过程中出现了众多建立在不同语言理论基础之上的测试方法和模式。它们对于语言知识是如何构成的、语言是如何使用的持有不同观点。这些基本观点不仅决定考试的目标语言能力(test construct),即考试希望考察受试者的什么知识或技能,还决定考试使用的题型(test method),即考试采用什么方法达到上述目的。

1.1.1 结构主义测试法

结构主义测试法是建立在以美国语言学家 Bloomfield 等人为代表的结构主义语言学的基础上的(Baker 1989)。结构主义语言学着重对语言的形式系统进行客观、科学的描写。通过将语言系统中音的序列逐级分解成句子、词组、词、语素、音位等从大到小不同单位的成分,并对它们进行分析和归类,结构主义语言学认为语言是由音系、词汇、句法三个层次构成的系统,每个层次上都有不同类别的成分,如音系层次上的音位、重音、语调,词汇层次上的词汇,以及句法层次上的语法结构。另外,由于所有这些语言成分都在四种语言使用模式中活动,由此可区分听、说、读、写四种语言技能。

Lado(1961)指出在结构主义测试法中,语言是一个需要掌握的形式系统,语言能力就是对语言形式系统的掌握。语言测试的目的就是考察考生关于这个语言形式系统中的成分的知识,而不涉及使用这些知识进行交际的能力,因而这种方法可称作是以语言系统为纲的语言测试。在考试方法上,结构主义测试法使用离散考点(discrete point)题型,将所测试的语言成分从上下文中分离出来,每题只通过一种技能、一个孤立的句子测试一个语言成分的知识,以排除其他成分的干扰。因此,词汇和语法的多项选择测试使用广泛,对考生语言使用的评价主要在于语言的准确性。考试设计强调信度和客观性。

1.1.2 综合测试法

不久,人们逐渐意识到结构主义测试法只注重语言系统知识自身,而忽视使用该知识进行交际的一面,于是将语言系统知识与具体语境相结合的综合测试法应运而生。该方法主要基于美国教育学家 John Oller(1979)提出的“语用期望语法”(pragmatic expectancy grammar)和“单一能力假说”(unitary competence hypothesis)之上(McNamara 1997; Baker 1989)。McNamara(1997:15)总结说,Oller 的理论为语言测试提供了一种全新的关于语言和语言使用的观点,不再注重语言知识自身,而是更多地关注语言使用中语言处理的心理过程。语言使用包括两个因素:一是对语言的实时处理(on-line processing),二是“语用映射”(pragmatic mapping),即利用关于语言系统特征的知识在语境中理解或表达意义。在 Oller 的理论中,语用期望语法是一种及时根据外界环境有序排列语言成分以建立或表达意义的心理机制,它就是在语言活动中综合语法、词汇、语境、语义和语用知识的能力。正是这种能力使得语言的产生和理解成为可能,因此它应该是语言测试的目标所在。Oller 进一步认为所有语言考试中的语言使用全都依赖于这种综合各种知识的内在能力,这种能力具有单一结构,是不可分的,这便是所谓的单一能力假说。为了更加经济高效地测试这种能力,Oller 进而设计出完形填空、成段篇章听写等综合测试题型。

尽管 Oller 的单一能力假说被证明是不可信的(Baker 1989:68-73),但是总体来说,综合测试法比结构主义测试法更接近真实语言使用。其目的是测量受试者综合利用语言系统知识进行交际的能力,着重于对在语境中的意义以及语篇的交际功能的理解和表达。但是它并不考虑采用可能与受试者未来语言使用相关的具体交际任务,因而可称作是以意义为纲的语言测试。主要题型有口试、论说文写作、较长的书面或口头篇章理解、完形填空、成段篇章听写等。评价标准立足于准确性和流利性。

1.1.3 交际测试法

交际测试法则是以 Hymes(1972)的交际能力理论为基础的。Hymes 认为掌握一种语言不仅仅需要掌握其语法规则,还必须掌握将语言与交际语境特点联系起来的该文化所特有的语言使用规则(McNamara 1997:16)。Bachman(1990:4)指出语境不仅包括句子所处的语篇,还包括它所处的社会语言情景(sociolinguistic situation),后者在很大程度上限制着语篇的功能和形式特点。简言之,交际能力理论进一步拓展了语言能力的内涵。相对于综合测试法,它针对前者提出的在语境中使用语言的能力这一概念又补充引入了社会语言学视角。McNamara(1997:17)指出,交际测试法代表了从视语言为内在现象的心理学语言

观向着重语言的外在社会功能的社会学语言观的转变。相应的,交际性语言使用也被看作是情景语境、语言使用人以及语篇三方面之间的动态的相互作用。语言使用被定义为“在特定情景中,两个或两个以上语言使用人之间主动地协商意向意义的过程”(Bachman 和 Palmer 1996:61)。

随着交际性语言能力和语言使用观的树立,关于交际性语言测试的讨论也在语言测试工作者中广泛展开。Morrow (1979) 认为交际性语言考试是以做事为纲的测试,意在探寻受试者能够用语言做些什么事情。这种测试方法必须首先通过发现在特定领域中受试者实际使用语言所做事物的类型来确定所要测量的交际类型。Morrow 还提出了一些交际语言使用的特点,他认为语言考试也应该具备这些特点(Skehan 1988:215)。真实的交际具有①交互性(发生在参与人之间,每个参与人都不断影响着对方语言行为的形式和内容);②一定程度的不可预测性(参与人必须处理新信息,被引向不可预知的方向);③语境化(发生在一定的语篇和社会文化情景中);④目的性(参与人总是试图利用语言达到例如劝说、欺骗之类的某种目的);⑤与语言运用的相关性(在记忆限制、时间压力、失误等真实的心理条件下进行);⑥真实性(所用材料不是简写材料或为考试特别编制的材料);⑦以行为为评价标准(考试语言使用是否合格取决于是否达到了交际目的)。

Wesche (1983:45-48) 也指出交际性语言测试的目的是考核“受试者在自然情景中用所学语言能够做些什么”,或者说“受试者能否有效运用所学语言达到特定交际目的”。这些目的是由受试者的语言需求决定的。关于交际性考试的特点,Wesche 罗列了六种特征:①考试应该是综合性、实用性的考试,让受试者在一定的语言语境和情景语境中使用自然的语言;②考试应该尽可能“直接”,尽量模拟目标语言使用情景;③考试应该涵盖反映考试目标的一系列情景;④考试应该采用尺度参照法(criterion referencing)评分,将每个受试者的考试语言运用与所定义的任务合格语言运用进行比较;⑤考试应该具有信度,和⑥可行性。

Swain (1985) 强调交际性语言测试的四条普遍原则:①始于一定的起点,考试开发应该以当前的交际能力理论为基础;②着重内容,交际性语言考试的内容应该具有激励性(使用生动合适的话题、引人入胜的形式,激发受试者完成真实的任务)、实际性(含有受试者不了解的新信息)、综合性(所有信息和活动都围绕着一个主题)和交互性(向受试者提供就话题交换意见的机会);③考试偏颇性应有利于受试者的最佳发挥,考试开发者应该尽可能引发出受试者的最佳表现;④反拨效应,考试开发者应该让教师参与到考试的开发、实施和评分中来,以改善课堂教学。

Morrow (1991) 通过开发新的 CCSE 考试系列总结了一个优秀考试的特点:①考试应该由几个独立的部分构成,每个部分测试听、说、读、写各个单项技能;②考试应该是对受试者语言运用——使用语言做什么事——的直接测量;③考试应

该采用语言任务,让参与人在理解或产生语言时完成任务,从而通过对意义的协商达成共识;④考试应该是尺度参照考试,评价受试者运用语言交流意义的能力;⑤考试应该反映和促进有益的课堂教学方法,对课堂教学产生良好的反拨效应。

总之,交际性测试最根本的特点在于这类考试的目的是考察受试者是否能够运用语言完成特定的交际任务。因此,这类考试可称作是以交际任务为纲的语言测试。而这一点又进一步决定了交际测试中确定受试者语言需求的必要性和所开发考试的特定性。交际测试法采用社会语言学视角,重视受试者在现实生活中可能承担的社会角色(McNamara 1997:17)。每个特定考试的目的就是测试受试者在与他们实际交际需求密切相关的语言使用情景中运用语言有效达到交际目的的能力(Wesche 1983:45, Clark 1983:433, Weir 1990:11)。正如 Bachman 和 Palmer 所言,“任何一个考试的开发都必须有明确的目标、特定的受试群体和特定的语言使用领域(target language use domain)(即受试者在语言考试以外使用目标语言的情景或语境)”(1996:18)。因此,需要采用不同的考试来评价具有不同目标语言使用领域需求的不同受试群体。

在考试方法上,交际测试法追求真实性或直接性,即考试应该尽可能地具有真实实际的特点,如交互性、信息的不可预测性、语境化等等。从根本上讲,考试应该让受试者实施足够长的接收性或产生性交际行为(McNamara 1997:16),并且为考试语言使用任务设置的情景应该尽可能接近真实生活中发生该类任务的条件(Clark 1983:433)。最后,对受试者语言运用的评定当然也应按照交际的有效性进行,不仅包括正确性、流利性,还包括恰当性。

这三种测试方法主要特点的比较(见表 1.1)显现出语言测试趋向测试直接性的发展趋势,即考试中的语言使用与实际情景中的语言使用越来越接近。Clark (1983)指出,既然语言教学的最终目的就是培养学生在真实交际场合恰当有效地使用目标语的能力,那么努力近似真实交际活动的交际测试法考试显然能够更直接,更有效地服务于教学目的,并且更能够激发学生的学习兴趣 and 教师的工作热情。

表 1.1 三大语言测试方法比较

语言测试方法	视 角	考什么	怎么考	评价标准
结构主义测试法	语言学	掌握语言形式系统的知识	离散考点	正确性
综合测试法	心理语言学	理解或表达语篇意义和交际功能	综合性	正确性 流利性
交际测试法	社会语言学	完成交际任务	综合性 真实性	交际有效性 { 正确性 流利性 恰当性