

语言测试学

引论

王英杰 编著

山西出版集团
山西人民出版社

ISBN 978-7-203-05831-1



9 787203 058311 >

定价:25.00 元

语言测试学

引论

王英杰 编著

山西出版集团
山西人民出版社

图书在版编目(CIP)数据

语言测试学引论 / 王英杰编著. —太原: 山西人民出版社, 2007.6

ISBN 978 - 7 - 203 - 05831 - 1

I. 语… II. 王… III. 英语 - 测试 - 研究 IV. H319

中国版本图书馆 CIP 数据核字(2007)第 076233 号

语言测试学引论

编 著: 王英杰

责任编辑: 杨 文 高美然

出 版 者: 山西出版集团·山西人民出版社

地 址: 太原市建设南路 21 号

邮 编: 030012

电 话: 0351 - 4922220(发行中心)

0351 - 4922208(综合办)

E-mail: fxzx@sxskcb.com

web@sxskcb.com

Renmshb@sxskcb.com

网 址: www.sxskcb.com

经 销 者: 山西出版集团·山西人民出版社

承 印 者: 山西文博印业有限公司

开 本: 850mm × 1168mm 1/32

印 张: 9.75

字 数: 260 千字

印 数: 1 - 1000 册

版 次: 2007 年 11 月 第 1 版

印 次: 2007 年 11 月 第 1 次印刷

书 号: ISBN 978 - 7 - 203 - 05831 - 1

定 价: 25.00 元

序

测试和测试分析是外语教学的重要组成部分,也是不可缺少的环节。通过不同类型的测试,教师可以获得多方面的信息,了解学生学习外语的素质、潜能,考查学生经过一个阶段课程学习后所取得的进步,从而掌握学生的实际外语水平。根据测试结果所提供的反馈信息,教师还可以检验教学大纲、教学方法和教学内容的可行性。学生参加测试可以达到复习、巩固和提高的目的,测试后能针对自己的薄弱环节,有的放矢地进行补课。

本书在简要叙述测试的发展后,介绍英、美有关将英语作为外语测试的理论,分析了在这些理论指导下设计的各种试卷的特点,以便广大读者参考。全书结构完整,主要内容包括以下几个方面:

- 第一,语言测试的发展和基本理论;
- 第二,语言测试的类别;
- 第三,语言测试的题型和任务;
- 第四,语言知识和技能的测试;
- 第五,命题的流程和质量。

本书基于作者本人讲授语言测试概论课程的讲义,本着这样的原则来进行这门课程的设计:以历史概述为出发点,以理论探讨为开端,到命题的流程和质量评估为终点,力争提供一个语言测试的全景图。为了实现这一目标,广泛吸取国内外的研究成果,结合语言测试的实践,通过大量的范例,来阐述测试各语言要素和各种语言能力的手段与方法,并且辩证分析各种测试方法与手段的长

处与不足、适用性与局限性。

本书集中论述英语试题。但是,不是谈怎么做题,怎么猜题,而是谈怎么去命题。因此,本书的对象首先是英语教师。现阶段,各学科的发展和社会的需求导致测试形式和内容发生了变化,英语教师如果不懂得命题的基本理论和实践,不仅命题不符合实际要求,而且势必大大削弱教学效果。本书的对象其次是英语专业的学生。他们要为今后的工作做好准备,日后一旦登上讲台,时常要命题。此外,其他专业学英语的学生如果具备一些测试方面的知识,也可以从中悟出一些道理,促进英语学习。

本书承蒙有关各方的关怀与支持,得以顺利完成。尤其是运城学院外语系的秦建华教授为本书的出版作出了辛勤的努力,在此谨致谢忱。

编著者

2007 年暑假

目 录

Contents

第一章	语言测试的定义和目的	[1]
第一节	测试的起源	[1]
第二节	世界各国的考试制度	[2]
第三节	语言测试的历史发展	[6]
第四节	语言测试的性质	[9]
第五节	语言测试的基本特点	[11]
第二章	语言测试的几个理论问题	[15]
第一节	测量与教育测量	[15]
第二节	交际	[18]
第三节	语言行为、语言能力、 交际行为和交际能力	[22]
第四节	记忆和思维	[28]
第五节	错误	[30]
第三章	语言测试的基本类别	[34]
第一节	按用途分类	[34]
第二节	按测试方法分类	[40]
第三节	按评分方法分类	[42]
第四节	其他分类方法	[48]
第四章	题型	[52]
第一节	题型	[52]
第二节	对部分题型的讨论	[56]

第五章	语言测试的任务和题型	[71]
第一节	测试语言的理解能力	[71]
第二节	测试语言的表达能力	[95]
第三节	测试综合的交际能力	[107]
第六章	语言知识测试	[125]
第一节	语法测试	[126]
第二节	词汇测试	[141]
第七章	听力测试	[155]
第一节	听力题的命题	[155]
第二节	听力多项选择题举例	[160]
第三节	听写	[168]
第八章	阅读的测试	[172]
第一节	对阅读的理解	[172]
第二节	对阅读的测试的理解	[181]
第三节	题型:两种常见阅读测试 方法的分析	[188]
第四节	结论	[206]
第九章	口语测试	[210]
第一节	概述	[210]
第二节	测试内容	[212]
第三节	测试形式	[213]
第四节	评分	[224]
第十章	写作测试	[235]
第一节	概述	[235]
第二节	命题	[237]
第三节	评分	[256]

第十一章 命题流程	[271]
第一节 确定目标、类别、内容和方法	[271]
第二节 写出命题大纲、制定内容细目表	[273]
第三节 编题、校对和拼题成卷	[277]
第四节 审卷	[278]
第五节 预测	[280]
第六节 定稿和印卷	[281]
第七节 数据的采集和分析	[281]
第十二章 命题的质量	[283]
第一节 效度评估	[283]
第二节 信度评估	[291]
第三节 区分度	[294]
参考文献	[299]

第一章 语言测试的定义和目的

测试是社会发展的结果,是社会为了选拔人才、检查教育效果和学习效果而采取的可以用来推测候选人能力的一种手段。作为测试的一个分类,语言测试也不例外。语言测试可以为社会选拔人才,也可以作为有效的手段来检验教和学的质量;并为教和学提供反馈。

第一节 测试的起源

测试是人类社会发展的结果,是人们为了适应社会发展的需求而设计,并逐步发展起来的。

人类在劳动的过程中,感到有必要确定物体的属性,例如数量、大小、重量之类的物理属性,于是就产生了测量,并研究用测量的方法来了解这些属性。随着人类文明的发展,测量的范围越来越广,测量的工具越来越多,测量的方法也越来越多。测量的对象逐渐从物扩大到人,从对人体形体的测量扩大到对人其他方面的测量,例如对人的知识和能力的测量。

对人的知识和能力的测量,始于原始社会后期。当时,生产水平日益提高,氏族、部落之间的交往增多,冲突也逐渐加剧。为了维持氏族、部落的生存和发展,要向年轻的一代传授生产技能、军事技术、宗教知识,于是产生了教育。为了检验教育的成果,开始用骑马、射箭、搏斗等演示活动来考核技艺。青少年一般要经过严格的考核之后才能被承认是氏族、部落的正式成员。

后来,随着社会生产水平提高,社会出现了阶级,社会分工更明确了,教育逐渐从生产、劳动中分化出来,出现了专门从事教育的人和专门从事教育的机构,也就是说,出现了教师和学校。随着教育的发展,测量受教育者的方法日趋完善。人们在教学的过程中不仅把测量作为一种手段,而且开始研究教育测量(educational measurement),因此出现了教育测量学。

第二节 世界各国的考试制度

一、中国是测试故乡

中国是考试的故乡。就出现大规模的、严格管理的考试而言,中国早于西方各国。我国的考试制度在奴隶社会末期就初步形成了。《礼记》比较详细地介绍了秦汉以前贵族的教育制度、教学内容和教学方法;同时显示出,在西周的时候就有定期考查学生作业的考试制度,而且当时的“国学”中的大学,每隔一年必须考一次。

考试制度到了封建社会得到进一步发展、运用。据记载,早在公元前165年的时候,汉文帝就使用正式的考试来挑选官职候选人。到了隋朝,这一做法发展成为系统的科举制度。唐代继续实行科举制度。那时,科举大体可分为两类:一类是常科,有秀才、明经、进士、明书、明法、明算等基本科目,每年定期举行;另一类是制科,由皇帝根据需要临时下令举行。宋代科举的办法,大体上用的是唐代的老办法。明代做了一些变更,其中最为突出的是宪宗成化年间(1465—1487)开始盛行所谓的“八股文”。这种特殊的文体,对于明清两朝的教育和文风都有重大影响。

科举考试的方法,同现代的考试有相似之处。例如,题型之一是贴经,这种题型多少有点像今天的填空题。所谓贴经,就是从经书中任揭一页,留出一行,其余的文字都盖起来。留出的一行,要

用纸贴盖三个字。考生得把这三个字说出来,或者写出来。又如,考试常常用策问。策问就当时的政治、经济、军事、生产等方面的问题提问,要考生书面答题,阐述自己的意见。这与今天政治课的问题颇有类似之处。

中国的科举考试,不仅有笔试,还有口试。例如,唐代有一种考试方法,称为口义,要应考者口述经义。《新唐书·选举志上》有这样的记载:“凡明经,先贴文,然后口试经文大义十条……元和二年……明经停口义,复试墨义十条。”“口义”是口试经义;“墨义”是将所答内容,录于纸上,“直书其义,不假文言”。

封建帝王用科举考试来选拔官员,具有一定的积极意义。首先,科举制度有利于加强中央集权制,把选择官员的权利从世家大家族手里夺过来。其次,科举制度可以确保封建帝王根据考试成绩比较广泛地从下层吸收、选拔人才。这种做法,不仅可以提高官员的文化素质,还调动了知识阶层,特别是中小地主阶层子弟的学习积极性。历代以科举入仕的人,有才识的人不少。可以说,我们的祖先很早就认识到了考试的重大作用,为我们积累了丰富的经验。

但由于前人不大注意考试方法的改革,考试内容陈腐、形式僵化,常常把读书人引入歧途,使他们轻视自然科学知识,轻视能解决问题的社会科学知识,把大量的时间和精力浪费在死记硬背古书的文句之中。特别是由于封建统治阶级为维护其既得利益,尽管有科举考试,重门第、走后门的现象始终未能杜绝。到了明、清时期,舞弊之风更盛,科举考试的弊端更明显。中国的科举制度陪伴着封建制度生存了一千三百多年,由于不能适应社会发展的需要,到清光绪三十一年(1905年)推行学校教育制度时而被废除。但考试作为发现、选拔人才的一种方法,并没有因此而完全被废除。

当代中国的考试,继承了中国古代考试的一些优良传统,也从国外吸取了大量科学的理论和有用的经验。以学校考试为例,教育的发展,大大地促进了学校考试的改革和发展。教育工作者逐步懂得,考试要为社会服务,要遵循教学的客观规律;考试的目标、内容、方法、标准等的确定,都要有利于受教育者的身心发展,有利于社会的发展。

现代的学校考试,与旧社会的那一套已完全不同了。首先,内容发生了重大变化。不再局限于古代圣贤教诲和文学、政治经典,而是现代社会科学、自然科学的各个方面;不再局限于记忆和背诵,比较注重理解和应用。其次,教育方法有重大改进。

20 世纪 70 年代以来,我国英语考试的改革走在其他学科的考试改革之前。我国英语界最早引进多项选择题技巧 (multiple choice technique),后来又率先采用完形填空技巧 (cloze technique),为题型的改革作了有益的尝试。1981 年,在当时高教部的领导下,我国第一个标准化考试(简称 EPT,即 English Proficiency Test)诞生。在此基础上,又推出了 MET (Matriculation English Test)、CET (College English Test) 等颇有影响的英语标准化考试。另外,英语界最早使用光电阅读器阅卷,最早在像高考这样的大型考试中应用教育统计学的理论指导实践。全国英语界认真地学习和宣传国内外的先进经验,为中国的考试改革做出了突出的贡献。

二、西方的考试制度

西方国家建立国家考试制度,比中国晚得多。据史料记载,到 19 世纪,西方才有比较完备的考试制度(或称考绩制度),政府根据竞争性的考试以及对业绩的客观评价来选拔官员。最早建立这种制度的是普鲁士和英国。1883 年,美国正式用考试来挑选官员,用所谓的量才录用制取代原来的政党分摊制。

现在,除了欧洲、北美各国,很多国家实行国家公务员任职考

试制度。例如日本就有比较完善的公务员制度。据日本《国家公务员法》的规定,要做公务员就得参加考试。考试分初、中、高三三个等级。三个等级中,参加高、中级考试的人必须具备大学学历。参加高级考试的,名牌大学毕业生居多,当然要求也最高。以1987年的高级考试为例,34000人参加,最后仅仅录取了1562人,可见难度之大、竞争之烈。

很多国家非常重视中学统考和高等院校入学考试,把这一类考试看做是选拔人才的重要手段。

以美国为例,1959年,衣阿华州衣阿华市一个名为美国大专院校测验规划处(American College Testing Program,缩写为ACTP)的机构推出一种标准化考试系列,取名为美国大专院校测试(American College Tests,缩写为ACT)。测试分为课程测试(Academic Test),包括英语、数学、社会知识、自然科学四个认知领域;兴趣发明测试(Interest Invention Tests),包括六个方面、两大类。ACTP负责命题、制卷、阅卷等各项有关测试的工作,提供成绩报告,另外还提供学生简况(ACT Student Profile)。现在,很多中学把ACT作为高中结业或者毕业考试,很多高等学校把ACT分数作为录取新生的主要依据。

又如,美国的学习能力测试(SAT,即Scholastic Aptitude Test)也是一种被当做高等学校入学考试的著名的标准化考试。这一考试由举办托福的ETS(即Educational Testing Service)来举办,对象是十二年级的中学生,每年考生达120万。

考试这种测试形式不会因为朝代的变更或社会制度的变革而失去作用。随着社会的不断进步,考试越来越广泛地被人们接受,测试的方式方法也会越来越科学,而研究测试的重要性也会越来越明显。

第三节 语言测试的历史发展

语言测试,这里主要指外语测试,是随着外语教学而出现的。教学和测试有着密切的关系,要了解测试的历史,必须了解教学的历史。正规的外语教学进入学校至今大约几百年间,外语教学不断发展变化,衍生了名目繁多的外语教学方法与模式。外语测试的发展虽在外语教学之后,但也演绎出种种方法与模式。

教学理论决定着测试的理论、方法和实践,而测试的理论、方法和实践则反映着教学的理论。

纵观语言教学的历史,特别是欧美语言教学的历史,我们就会发现,在大约四百年前,欧洲的语言教学,尤其是外语教学,形成了一套完整的教学理论。当时,主导语言教学领域的是语法翻译法。这是同当时人们对语言的认识分不开的。当时主导欧洲的理论是“功能心理学”。人们认为拉丁语是最合乎逻辑的语言,它的语法可以用来训练人的思维。再加上当时的文艺复兴运动,鼓励人们去重新重视、挖掘古希腊、古罗马的文化财富,对古典文献给予了特别的关注。这种现象反映在外语教学领域,就是人们在学习语言的过程中把古典文献作为学习的材料。与此同时,人们认为各种语言之间并没有实质上的差异,不同语言的词汇表达的概念完全相同,因而外语学习也就等于不同语言之间的翻译而已。因此,外语学习也就是一个翻译古典材料、学习语法的过程。因此,我们可以说,语法翻译法强调的是知识的传授。

语法翻译法对语言教学的影响一直持续到 20 世纪 40 年代。当时,正值第二次世界大战,美国因为参战,需要把军队派往世界各地。为此,军方就邀请语言学家、心理学家等为其设计语言培训项目,这就是战后盛行于世界各地的“听说法”。当时,这个培训

项目发挥了特别积极的作用。参训人员经过数周的培训,就基本掌握了驻地的语言,可以被派往当地了。战后,经过各方面的协同发展,这个培训项目就演变成为流行于世界各地的“听说法”了。

“听说法”的语言学基础是结构主义语言学。结构主义语言学认为,语言是由无数的句子组成的,句子是由词组成的,词是由词素或语素组成的。因此,学语言就应该从语素开始,经过词的学习,最后才学习句子。尽管句子是无数的,但仍可以从中找出规律,这个规律就是句型。因此,掌握了句型,就可以掌握句子了。结构主义语言学还认为,语言是一套形式符号的习惯。因此,学习语言也就是学习一套习惯。形成一套习惯,就需要经过反复的操练,达到自动化(automation)。“听说法”的心理学基础是行为主义心理学。行为主义心理学可以追溯到巴甫洛夫(Pavlov)的条件反射心理学,经过斯金纳(Skinner)的发展用于语言教学领域。行为主义心理学典型的观点是刺激—反应理论(S-R)。

在具体的教学法中,“听说法”就是特别强调句型的反复训练。反映在中学的英语教学中,就是句型的替换练习。因此,我们可以说,“听说法”的中心是强调技能的训练。

20世纪50年代末,美国的语言学家乔姆斯基(Chomsky)对结构主义语言学、行为主义心理学发动了猛烈的攻击。他认为,人们语言的获得不是反复刺激的结果。我们能够经常发现,在获得语言的过程中,儿童总能够说出他们从来没有听到过的新句子。这就说明,人类的大脑中天生地存在一个帮助人们获得语言的机制,这就是语言习得机制(LAD)。语言习得机制是人们大脑中一套关于语言的知识,即语言能力(linguistic competence)。乔姆斯基认为,这个语言能力是一个典型的人关于语言的所有知识。

20世纪70年代,英国的语言学家海姆斯(Hymes)对乔姆斯基的观点提出了反动。他认为,在我们的现实世界中,我们找不到

这么一位理想的、抽象的人。每一个人都有其特点,都与他人存在着各种差异,都是社会的人,承担着社会的角色。因此,他提出了“交际能力”的概念。经过许多语言学家的发展,人们对“交际能力”达成以下的共识:

交际能力 { 语言能力:语言形象系统本身的操作能力
语篇能力:语言形式结构置于语篇中的能力
语用能力:语言形式结构置于情景中的能力

简单地说,语言就是用来做事的。把这一理论用于教学实践中,就是现在比较流行的“交际法”。反映在教学过程中,就是做事型练习。比如,写一封道歉信、下一个开会通知等。因此,我们可以说,“交际法”的中心是强调能力的培养。

我们可以发现,这三个不同的教学流派,反映了不同的语言学习理念,因而也就有了与之相适应的教学中心。不同的教学中心也就要求各自的测试理念和测试方式与之相适应。

既然“语法翻译法”的教学中心是知识的传授,它的测试理念也就是对所学知识记忆的检测。“语法翻译法”的测试方式就有“死”的和“活”的两大类。较“死”的题型有:标音、拼写、听写、配对、单词释义,给出同义词、反义词或派生词,朗读、背诵、词法变位、语法填空、完成句子、变换句型、词类辨义、句法分析、条件翻译、复述短文等。这些较“死”的题型大部分要求“记”的功夫,测试的目的是语言的形式、规则以及汉外对译等。如果学生能够较为用功去记这些知识,他们就能够获取比较高的分数,我们也就可以推测他们比较好地掌握了这些知识。较“活”的题型有:造句、改写、答问、翻译、作文、口头作文等。

“听说法”的教学中心是技能的培训,它的测试理念就是对所掌握技能的检测。“听说法”的主要测试方式有:语法填空、词汇填空、词语填空、完成句子、句型变换、句子辨错以及单句听义等。