

朗声英语教师进修丛书



英语语言测试 基础

曾用强 著

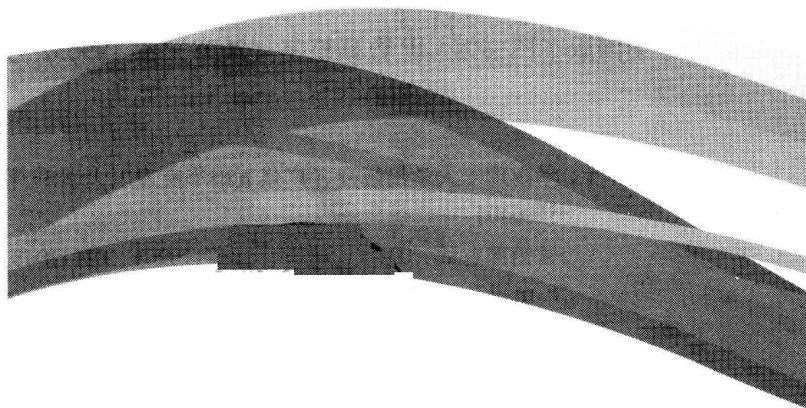
工大学出版社

朗声英语教师进修丛书

英语语言测试

基础

曾用强 著



华南理工大学出版社

·广州·

图书在版编目 (CIP) 数据

英语语言测试基础/曾用强著. —广州: 华南理工大学出版社,
2009.1

(朗声英语教师进修丛书)

ISBN 978 - 7 - 5623 - 2982 - 4

I. 英… II. 曾… III. 英语 - 测试 - 中学 - 教学参考资料
IV. G633.413

中国版本图书馆 CIP 数据核字 (2008) 第 199632 号

总发行: 华南理工大学出版社 (广州五山华南理工大学 17 号楼, 邮编
510640)

营销部电话: 020 - 87113487 87110964 87111048 (传真)

E-mail: z2cb@scut.edu.cn <http://www.scutpress.com.cn>

策 划: 何 娴

责任编辑: 谢茉莉

印 刷 者: 湛江南华印务有限公司

开 本: 880mm × 1230mm 1/32 印张: 7.25 字数: 202 千

版 次: 2009 年 1 月第 1 版 2009 年 1 月第 1 次印刷

定 价: 20.00 元

版权所有 盗版必究

前 言

10 多年来,由于研究兴趣的缘故,我一直与中学英语教师有来往。在往来的过程中,我深感第一线老师的辛苦和压力。这些辛苦和压力主要来自课堂的教学和考试的备考,尤其是每一次的教学和考试改革都会给我们的老师添加几分危机感,总是担心自己的知识更新赶不上改革的速度。

很多老师希望能有一套适合中学英语教师使用的进修丛书,在紧张的工作之余能提高自己的专业素质。这个愿望今天总算可以实现了。应广州市朗声图书有限公司约稿,我负责编写了这本《英语语言测试基础》,目的是为英语教师解决在教学过程中遇到的有关考试方面的实际问题。本书不仅是简单地提供问题的答案,更是提供思想和方法。阅读这本书的最终目的不是简单地知道一些技巧,而是要培养一种富有创造性的思维,达到有效利用考试服务教学的目标。

本书的主要特点可以归纳为以下几个方面:

(1) 注重理论与实践相结合。本书既不是纯理论的研究报告,也不是纯应用的技巧传授;而是在现代语言测试理论框架下,分析并解决英语教师实践中遇到的有关考试的实际问题。

(2) 突出问题的现实性。本书所谈及的考试问题都是作者多年教学或调研中发现的问题,这些问题都是英语教师普遍关心的问题。

(3) 体现内容编排的新颖性。每个章节的内容编排:首先通过问题讨论引出一个知识图表,它是知识的概括,更是读者思维的导引;然后详解本章节的语言测试知识;最后进行知识归纳,并引导读者应用本章节知识解释或分析教学中遇到的一些考试现象和问题。

全书分 6 章,第一章介绍语言测试的一些基本概念,包括语言测试的定义、考试的局限性、语言能力的基本结构及语言测试的发展趋势;第二章介绍考试的各种不同分类方法,重点介绍各类考试的特征及不同考试之间的异同点;第三章讨论考试的设计,主要涉及考试的评价标准——信度、效度和考试设计的基本原则,并通过实例介绍各类不同考试的设计;第四章介绍一些常用的考试题型,同时对语言知

识和语言技能的能力要求作了简要的描述；第五章是关于命题的基本原则，主要内容包括试题的评价标准，以及不同题型的试题特点和命题原则等；第六章讨论试题分析的意义、内容和应用。

由于时间仓促，我只能提纲式地把平时积累的经验和资料整理成册，很多内容都只能是“点到为止”，不能展开详细论述。希望待日后修订时，能够弥补这个遗憾。由于编者水平有限，书中存在错漏，恳请读者朋友批评指正。谢谢！

作 者

2008 年 12 月于托斯卡纳

目 录

第 1 章 语言测试的基本概念	1
1.1 语言测试的定义	4
1.2 考试的局限性	4
1.3 语言能力的基本结构	6
1.4 语言测试的发展趋势	9
1.5 英语教师关心的考试问题	12
1.5.1 考试的公平性	12
1.5.2 考试的反拨作用	12
1.5.3 考试的科学性	12
1.5.4 考试的稳定性	13
第 2 章 语言测试的分类	15
2.1 从考试的时期上区分	16
2.2 从考试的用途上区分	18
2.2.1 学业成就测试	19
2.2.2 进步测试	20
2.2.3 水平测试	20
2.2.4 学能测试	21
2.2.5 诊断性测试	22
2.2.6 分级测试	23
2.3 从反映分数的方法上区分	24
2.3.1 尺度参照性测试	24
2.3.2 常模参照性测试	25

2.4 从考试的内容上区分	27
2.4.1 综合性测试	27
2.4.2 离散考点测试	27
2.5 从考试的要求上区分	28
2.5.1 能力测试	28
2.5.2 速度测试	28
2.6 从考试的方式上区分	29
2.6.1 笔试	29
2.6.2 口试	29
2.6.3 做事测试	30
第3章 考试设计	33
3.1 考试的信度和效度问题	35
3.1.1 信度	36
3.1.2 效度	38
3.1.3 信度与效度的关系	40
3.2 考试设计的基本原则	41
3.3 考试设计举例1: 水平考试	45
3.3.1 基于理论定义的水平考试设计	45
3.3.2 基于目标情景的水平考试设计	48
3.3.3 2007年广东省高考英语的开发过程	48
3.4 考试设计举例2: 学业成就考试	54
3.5 考试设计举例3: 诊断考试	56
3.6 考试设计举例4: 竞赛英语	60
3.7 考试设计举例5: 口语考试	61
第4章 常用考试题型介绍	64
4.1 词汇考试	68

4.1.1	词汇能力的定义	68
4.1.2	词汇能力的测定	69
4.2	语法考试	74
4.2.1	语法能力的定义	74
4.2.2	语法能力的测定	75
4.3	听力考试	80
4.3.1	听力能力的定义	80
4.3.2	听力能力的测定	80
4.4	阅读考试	91
4.4.1	阅读能力的定义	91
4.4.2	阅读能力的测定	92
4.5	写作考试	101
4.5.1	写作能力的定义	101
4.5.2	写作能力的测定	102
4.6	口语考试	105
4.6.1	口语能力的定义	105
4.6.2	口语能力的测定	106
第5章	命题的基本原则	114
5.1	试题的评价标准	116
5.1.1	题干选材	116
5.1.2	考点设置	117
5.1.3	答案设定	118
5.1.4	干扰项设置	119
5.2	客观试题的命题	119
5.2.1	总原则	120
5.2.2	单项填空	120

5.2.3	完型填空	126
5.2.4	阅读理解	129
5.3	主观试题的命题	135
第6章	考试分析及其应用	140
6.1	考试分析的意义	141
6.1.1	考试评价	142
6.1.2	试题评估	145
6.1.3	教学反拨	146
6.2	考试分析的内容	146
6.2.1	答卷情况描述	146
6.2.2	信效度分析	147
6.2.3	项目分析	148
6.2.4	干扰项评估	152
6.2.5	试题分析	154
6.2.6	作文分析	157
6.3	考试分析实例	159
6.3.1	成绩分布描述	159
6.3.2	各题型分析	162
6.3.3	客观试题项目分析	162
6.3.4	主观试题分析	164
6.3.5	试题分析	165
附录一	2008 年普通高等学校招生全国统一考试英语试题（广东卷）	177
附录二	2007 年普通高等学校招生全国统一考试广东省英语科 考试说明	192
附录三	2007 年广东省高考英语效度研究	205

第1章 语言测试的基本概念

问题思考

为什么英语教师需要掌握语言测试知识?

教师平时最关心的问题当然是如何教书,比较少关注考试的问题。对于很多老师而言,考试只是一个很简单过程,无非选用(或命制)一份试卷考学生,然后通过评卷给他们反馈一个成绩。确实,考试看似就这么简单,但是简单的过程却引发我们很多复杂的思考,比如:

- 考试与教学是指挥与被指挥的关系吗?
- 面对琳琅满目的中(高)考模拟试题,我们如何辨别它们的质量?
- 自己命题考试好,还是使用现成的试卷好?

这些都是老师在平时教学中经常要思考或感兴趣的问题。教学总是离不开考试,考试总是服务于教学。如果考试得到合理有效的使用,它必将对教学产生积极的促进作用。语言测试知识正是要引导英语教师合理有效地使用考试,充分发挥其对英语教学的积极反拨作用。

那么我们需要了解的语言测试是什么?首先让我们通过图 1.1 带领老师走进语言测试。

知识图解

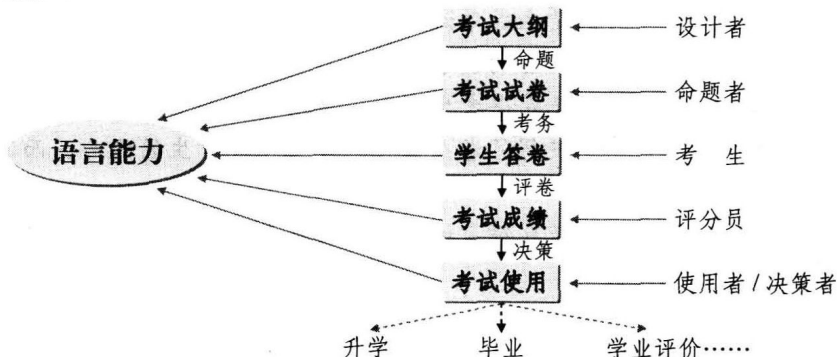


图 1.1 考试过程



知识详解

如图 1.1 所示, 考试过程包括了两大部分, 其中一个部分是可直接观察的, 即我们可以看得到的, 比如考试大纲、考试试卷和考试分数等, 它们都是每位老师熟悉的, 也是谈起考试时最受关注的内容; 另一部分则是不可直接观察的语言能力, 即我们不可以看得到的, 因为它是抽象的, 不能被直接测量。考试的过程就是建立这两部分之间的关系。

如何通过可直接观察的部分 (如考试大纲、考试试卷和考试分数) 反映不可直接观察的部分 (如语言能力), 从而为考试决策 (如升学和毕业等) 提供有效可信的依据, 这就是我们学习语言测试所需要回答的问题。只有当你知道了这个核心问题的正确答案时, 你也就找到了平时所遇到的关于考试问题的答案。

从图 1.1 中, 我们可以看出, 考试过程开始于考试大纲, 然后按照考试大纲的要求进行命题, 生成考试试卷, 并通过组织考试获得考生的答卷。考试之后组织评卷获得考试成绩, 即分数。这些分数最终被用作各种决策的依据, 比如决定升学、毕业或评价学业等。考试大纲就是基于考试决策的目标而制定的。由此可见, 考试过程中的每个环节都是互相联系又相互制约的。

● **考试设计:** 目的是制定考试大纲, 要求考试大纲必须满足考试决策的目标。

● **考试命题:** 目的是生成考试试卷, 要求考试试卷必须符合考试大纲的试题结构和考试内容的要求。

● **考务组织:** 目的是获得考生答卷, 要求确保考生在公平的考试环境下充分发挥能力, 使得答卷在最大程度上体现考生的能力, 而不受其他因素干扰。

● **考试评卷:** 目的是反馈考试成绩, 要求考试成绩必须是考生答卷的真实可信的反映。

● **考试决策:** 目的是有效使用考试, 要求考试决策必须与考试设计的使用目的保持一致。

考试的目标是基于可信的考试分数做出有效的考试决策。因此,



考试至少要实现两个确保：考试分数要确保可信，考试决策要确保有效。前者涉及考试的信度问题，后者涉及考试的效度问题。关于考试中的信度和效度问题，第三章将作详细的讨论。

在这里我们需要讨论的是，这些分数凭什么可以用来作各种的决策？答案是：分数代表的是考生的语言能力，正是这种能力支撑考试的决策。“高分高能”才是符合考试决策的要求，反之，“高分低能”则说明考试的效度可能偏低。显而易见，考试的核心问题是，考试分数是否准确反映考生的能力水平。

综上所述，考试中的每一个过程都是围绕特定的语言能力而展开的。

● **考试设计**：是对目标语言能力的定义，即，考试设计者首先必须定义考试拟要测定的语言能力是什么。

● **考试命题**：是对目标语言能力的抽样，即，命题者根据考试设计提出的语言能力进行抽样，确保考试内容能够代表目标语言能力。

● **考务组织**：是对目标语言能力的测量，即，考务人员负责组织好考试的每一个环节，确保全体考生在一个平等的考试环境下测出自己的语言能力。

● **考试评卷**：是对目标语言能力的评估，即，评分员按照统一的标准，宽严一致地评估考生通过考试试卷所反映出的语言能力。

● **考试决策**：是对目标语言能力的使用，即，考试使用者使用考试分数进行决策，完成考试的过程。

图1.1详细剖析了语言测试的过程，回答了什么是语言测试，揭示了语言测试的本质，并引发我们对语言能力的深度思考。本章节将详细介绍以下内容：

- ▣ 语言测试的定义；
- ▣ 考试的局限性；
- ▣ 语言能力的基本结构；
- ▣ 语言测试的发展趋势；
- ▣ 英语教师关心的考试问题。



1.1 语言测试的定义

基于上述讨论，我们可以归纳：语言测试是对语言能力的一次抽样推理过程，即，通过一份试卷（抽样过程）的考试组织，使用其考试成绩推导考生的目标语言能力（推理过程）。比如，高考就是考生通过考试获得成绩，并以此推导他是否达到就读大学所应具备的能力水平；期末考试也是通过抽样，推导考生掌握教学内容的程度。

我们可以使用以下式子来解释语言测试的定义：

$$\text{考试分数} \geq \text{语言能力}$$

语言测试的本质就是使用考试分数推导语言能力，换言之，考试分数必须能够解释考生的语言能力。要实现这个式子必须满足两个最基本的条件：

- (1) 考试分数必须是考生语言能力的客观体现；
- (2) 考生体现的语言能力必须与考试者定义的语言能力相一致。

也就是说，有效的考试必须是考试试卷确实激活了考生的语言能力（而不是其他能力），而且这份试卷的内容确实代表了考试拟要测量的语言能力。比如，你要设计一份期末考试试卷，首先必须保证试卷的结构和内容能够代表你的教学内容，然后保证考生在一个平等的考试环境下发挥出自己语言能力。只有这样，考试分数才能准确代表目标语言能力。但是在现实中，要实现这个式子是不容易的，因为影响语言能力和考试分数的因素有很多，比如，不同的语言学习观可能影响语言能力的定义，不同的题型设计可能改变拟要测量的语言能力，评分员的宽严不一可能影响考试分数的公平性，考试环境的变化可能影响考生的语言能力发挥，等等。所以，语言测试要研究的问题就是，如何最大程度挖掘考生能力，把非能力因素的影响限制到最小程度。

1.2 考试的局限性

从上一节的讨论，我们已经知道，考试是一个抽样推理的过程，而且影响这个过程的因素有很多。考试的抽样和推理属性也决定了考



试具有不确定性、间接性、不完整性、主观性和不精确性等特征。

不确定性 语言测试的测定对象是语言能力，它属于心理属性，具有不确定性的特征。不确定性一方面说明了语言能力没有一个确定的定义；另一方面也说明了考生本身对自己的语言能力也无法确定地知道。比如，关于阅读理解能力的定义有很多，不同的学者从不同的角度给出的定义都不一样；同样，考生也不可能确切知道自己的阅读能力有多强。

在考试中，一个准确的语言能力定义是很重要的，因为不同的定义可能导致不同的题型设计和考试内容。比如，把阅读能力定义为考生能够根据上下文猜测生词的意义，或把它定义为考生能够根据上下文推导作者的态度和写作意图，这两种不同的定义必然要求不同的考试方式，否则考试分数就不能准确解释所定义的阅读理解能力。不确定性特征决定了准确定义语言能力的难度，但是，考试设计的首要任务就是要对语言能力提出一个明确且具有可操作性的定义，因为它决定了考试题型和考试内容。所以，考试设计者首先必须基于考试的目的和目标考生群体的特征，从理论和操作两个层面对语言能力作出定义。

间接性 由于语言能力是一种心理属性，它不像物理属性（如身高和体重）一样，可以通过一定的工具（如尺子和秤）进行直接的测量。语言能力只能通过间接的手段（如考试试卷）进行测定，不可能有“能力尺”直接来测量考生的语言能力。考试试卷就是使不可观察（语言能力）变成可观察（考试分数）的一种间接测量工具。如同尺子是通过它的刻度表示人的身高一样，考试是通过分数表示考生的语言能力。与测量身高不同的是，考试是间接的。

不完整性 语言具有不可取尽的属性，它包含无限的语言形式和内容。所以语言能力的测定只是一个抽样的过程，从无限的语言材料中选择一定具有代表性的样本作为考试的内容。比如，要测定考生的语法水平，我们无法让他完成所有的语法项目，只能通过特定的题型让他完成一定数量的语法项目；同样，测定一个人的英语口语能力，我们无法让他说出所有话题的所有话语，只能是选择特定的话题和特定的形式，表达一定内容。



由于考试具有不完整性的特征，所以它对考试内容的抽样提出了很高的要求，否则就会影响考试的有效性。考试内容的抽样必须满足两个基本条件：代表性和覆盖面。代表性是指考试内容的抽样必须能够代表拟要测定的语言能力；覆盖面是指考试内容的抽样必须具有一定的数量，能够覆盖拟要测定的主要内容。

主观性 语言测试从语言能力的定义开始，到最终的考试决策，都有主观的成分。考试是间接的，它只能依靠人的努力来完成，包括考试开发者、命题者、考务人员、评分员以及考试使用者等。语言测试只能通过制定相关的原则和标准，保证最大限度地减少主观因素的影响作用。

不精确性 由于语言能力的测定具有主观性，误差是必然的。正是考试的这种不精确性，决定了语言测试只能努力降低考试的误差，而不可能消灭误差。

考试的不确定性、间接性、不完整性、主观性和不精确性特征，解释了考试的局限性。我们只有真正把握考试的本质，并清楚认识到考试的局限性，才能更好地使用考试，发挥考试的作用，否则就容易陷入考试万能的误区。

1.3 语言能力的基本结构

语言测试是以语言能力为核心的，因为它是语言测试的目标对象。只有当我们对目标对象有了很清楚的定义时，考试才能顺利开展。就像在教学中，要想给学生上好一堂阅读理解课，教师首先必须知道阅读理解指什么。对教学目标定义得越清楚，教学内容就越具体，思路也越清晰，教学效果当然就越理想。但是语言能力是一种不可直接观察的心理属性，它没有唯一的定义，考试开发者只能基于考试目的和目标考生群体的特征对语言能力提出假设。当然这种假设必须经过科学的验证，或效度验证。比如，高考所体现的是中学毕业生必须具备的语言能力，这种中学生必须具备的语言能力是不可直接观察和测量的，只是考试开发者根据高中课标和大学选拔人才的需求而定义的能力结构。



什么是语言能力？随着语言学习观的不断发展，人们对这个问题的认识也不断向纵深发展。对这个问题的理解和回答决定了语言教学与考试的发展方向。回溯过去 30 年我国中学英语教材的发展和教学的改革，我们可以清楚地认识到，随着语言学理论的发展，语言学习观也随之变化，语言学习观的变化直接表现为语言教学方法和教材的变化。比如，在过去的 30 年间，我国英语教学经历了语法—翻译法、听说教学法、交际教学法、功能意念教学和任务型教学等教学模式。语言教学的发展也必然要求语言考试的改革。自从 20 世纪 80 年代中期我国推出标准化考试以来，英语考试经历了多次改革与发展，这些改革与发展都是语言学习观和教学理念发展的必然结果。中学自 2004 年开始实施新课程标准，也体现了现代学习理念的 latest 发展。2007 年的高考改革正是为了适应这种最新发展的需要。那么，我们应该如何理解最新发展的现代语言观呢？

语言学习的目的是为了使用语言。所以，语言学习过程中所培养的语言能力也应该是语言使用能力。使用语言首先需要—个情景，完成特定的语言任务，语言任务需要依靠语言技能（听、说、读、写）来完成，语言技能又必须以语言知识为其核心内容。因此，语言知识、语言技能和语言任务构成了语言能力的三大要素。语言（运用）能力就是具备扎实的语言知识，并能够有效地应用语言技能，完成特定情境下的语言任务。图 1.2 解释了语言能力的构成。

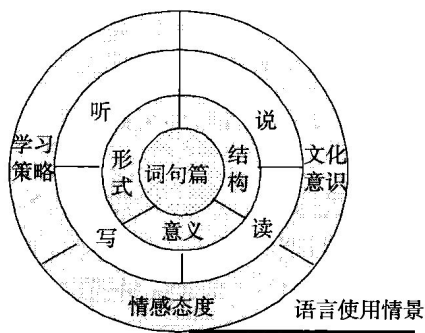


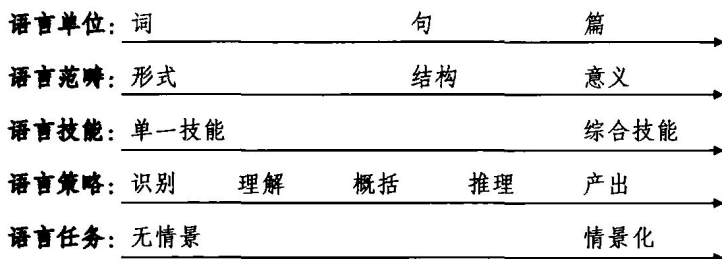
图 1.2 语言能力结构图



语言知识是材料，语言技能是手段，语言任务是目标。材料、手段和目标相互联系、相互依存、缺一不可。比如，你要听一次英国专家的讲座，然后参加他的讨论（语言情景或语言任务，目标），那么，你需要通过听来获取讲座的信息，通过写来记录你所听到的信息，并通过说来发表你的讨论（语言技能，手段）。要完成这项语言任务，你首先必须具有相关的语言知识（材料），比如特定的词汇和语法知识等。当然，语言应用能力除了这些语言因素外，还应包括一些非语言因素，比如文化意识、情感态度和学习策略等。

图 1.2 解释了语言能力的基本结构，其中每个部分又由若干成分组成，比如，图的中心部分是语言单位，它是由词、句和篇组成；图的中圈是语言范畴和语言技能，其中语言范畴包含了形式、结构和意义，语言技能包括听、说、读和写；图的外圈代表非语言因素，比如文化意识、情感态度和学习策略。

语言能力的基本结构不变，但是构成能力的组成成分总是在发展变化中。图 1.3 则解释了这种发展变化的趋势。



注：箭头代表语言能力的发展趋势

图 1.3 语言能力的发展趋势

从图 1.3 中，我们清楚地认识到，在现代的英语学习中我们应该培养怎样的语言能力。按照现代的语言学习观（如图 1.2 和图 1.3 所示），我们在语言学习过程中应该充分认识到以下问题：

（1）语言能力是一个整体概念，即，它的各个组成部分（语言知识、语言技能和语言任务）之间相互联系、互相依存、缺一不可。