



DUIWAI HANYU
JIAOXUE LUNCONG

对外汉语

教学论丛

(第三辑)

主编 刘 荣



四川大学出版社



DUIWAI HANYU
JIAOXUE LUNCONG

对外汉语

教学论丛

(第三辑)

主 编 刘 荣

副主编 何 婉 胡 晓



四川大学出版社

责任编辑:周 洁
责任校对:余 芳
封面设计:墨创文化
责任印制:王 炜

图书在版编目(CIP)数据

对外汉语教学论丛. 第三辑 / 刘荣主编. —成都:
四川大学出版社, 2016. 4
ISBN 978-7-5614-9380-9

I. ①对… II. ①刘… III. ①汉语—对外汉语教学—
教学研究—文集 IV. ①H195.3-53

中国版本图书馆 CIP 数据核字 (2016) 第 076471 号

书名 对外汉语教学论丛(第三辑)

主 编	刘 荣
出 版	四川大学出版社
地 址	成都市一环路南一段 24 号 (610065)
发 行	四川大学出版社
书 号	ISBN 978-7-5614-9380-9
印 刷	郫县犀浦印刷厂
成品尺寸	165 mm×235 mm
印 张	14.5
字 数	259 千字
版 次	2016 年 4 月第 1 版
印 次	2016 年 4 月第 1 次印刷
定 价	36.00 元

版权所有◆侵权必究

◆读者邮购本书,请与本社发行科联系。

电话:(028)85408408/(028)85401670/
(028)85408023 邮政编码:610065

◆本社图书如有印装质量问题,请
寄回出版社调换。

◆网址:<http://www.scupress.net>

序

中华文明五千年的历史，通过汉语的书面形式薪火相传。时至今日，随着中华民族的崛起，汉语逐渐走向世界。学习汉语的热潮在全球范围内蔚为壮观。正因如此，中国的对外汉语教学研究具有了“世界性”的意义。

在西方《圣经·创世纪》中，上帝为阻止人类建造通天塔，而变乱原本统一的语言，人类自此各散东西，发展出了不同的语言和种族。虽然“五方之民，言语不通，嗜欲不同”，但“东海西海，心理攸同；南学北学，道术未裂”。即使文化和语言千差万别，四海之内的“人心”也是相通的。而对外汉语教学的目的正是要架起一座沟通不同文化和语言的桥梁。为推进汉语事业的发展，我国对外汉语教学事业的先驱周祖谟等人一路披荆斩棘，筚路蓝缕，为此领域的理论建设铺下了第一块坚固的基石。如今放眼世界，这项“国家和民族的事业”正在以更快的速度和更旺盛的生命力迅猛发展。

《对外汉语教学论丛（第三辑）》的出版正是为国家文化事业建设添砖加瓦。该丛书不仅涉及对外汉语教学法的讨论、汉语学习的研究以及汉语本体的研究等诸多方面，而且学理性强，有较强的现实指导意义。它向人们立体而生动地呈现了一幅关于对外汉语教学理论种种课题实践的图景。虽然研究内容丰富多样，但都与对外汉语教学息息相关。所收录的论文多是在教与学的过程中诞生的成果，教学相长，可谓贵乎！

教学相长，是此书的初衷，也是最终目的。《学记》有言：“虽有佳肴，弗食不知其旨也；虽有至道，弗学不知其善也。是故学然后知不足，教然后知困。知不足，然后能反省也；知困，然后能自强也。故曰：教学相长也。《兑命》曰：学学半，其此之谓乎！”虽然教学有法，教无定法，但本书的诸位作者，在学海中孜孜不倦，不断探求规律，顺应时代发展，挖掘理论热



点，将思想的精华付诸笔端，结出了累累硕果。所谓学者，即是在悠闲的时光中忙碌自己的思想和智慧。而此书中的思想正是作者在教学之余刻苦钻研的结晶，定会给其他学者一定的启发，也必然会引起更多人对这一领域的关注。

“百川学海，而至于海；丘陵学山，不至于山”，为师者和治学者应当包举宇内，境界高远。正如“学其上，仅得其中；学其中，斯为下矣”，故“入门须正，立志须高”，当“通大义，识大体，立根本”，更应学贯中西，博学强知，“淹博，识断，精审”，方能高屋建瓴。然而“生也有涯，无涯唯智”，望诸位勤学苦思，切磋琢磨，潜心著书，放宽诸心，道理自会在潜移默化中显现。

后生可畏，焉知来者之不如今也？

是为序！

曹顺庆

目 录

在语言能力测验中建立构想效度的可能性	陈泓明 (1)
汉语中高级学习者习得汉语词汇方法思考	程 文 (18)
四川省对外汉语教学的地域文化优势及运用	郭 宏 (24)
对外汉语课程与小学语文课程的性质、目标及内容之对比研究	郭莉莎 (33)
高级汉语水平韩国留学生声调声学实验与分析	何 婉 (46)
以大学为依托建立对外汉语教学微信公众平台的必要性	胡 晓 (56)
《國語》單音詞同義詞“征”“伐”“討”考辨	雷 莉 (66)
交互主观性研究与对外汉语词汇教学	李 韵 (74)
利用新技术开展专门用途汉语的自主学习 ——以导游汉语为例	刘 荣 彭 越 (82)
留学生汉语称谓语掌握情况调查与分析	罗艺雪 (96)
析“字本位”论的认识错误	任珣璁 (110)
介词“从”的语块分析	王 川 (120)
“A 有 B (这么/那么) C”结构的语义认知分析	王 丹 (140)
汉语和 21 世纪	王文虎 (150)
浅谈语言测验的取向 ——语言能力	王 燕 (158)
字母 A, B, C, D 和汉语水平考试简称 HSK 的念法	王永场 (164)



语法化理论在对外汉语教学中的适当运用

——以“把”字句为例 徐 蔚 (168)

对外汉语高级口语课中的会话修正研究 杨 恬 (181)

中高级综合汉语课堂话题讨论方法初探 尹绍华 (196)

从对外汉语教学的视角看《现代汉语词典》编纂中的释义问题

..... 游 黎 (201)

中级汉语听力课堂教学思考 张晓洪 (211)

分层次教学模式在对外汉语教学中的应用及启示

——以美国犹他大学三年级中文课为例 周 丹 (216)

在语言能力测验中建立构想效度的可能性

陈泓明

摘 要：本文讨论的是在语言能力测验中建立构想效度的可能性问题。笔者认为，语言能力测验中构想效度的有效性同时指向理论、构想、测量以及对测量结果的解释，要验证语言能力测试中的构想效度的有效性，就必须对以上四个方面一一加以考察，但是目前要进行这样的考察还面临着相当大的困难。因此本文建议考察被试完成特定目的语任务的能力而非笼统地测试其语言能力。另外，构想效度的概念也有必要进行一定的修正。

关键词：构想；构想效度；语言能力测试

The Possibility of Establishing Construct Validity in Language Proficiency Test

Chen Hongming

Abstract: This paper examines the possibility of establishing construct validity in language proficiency test. It is argued that the validation of construct validity in language proficiency test presupposes, simultaneously, the validation of the underlying theory, the construct, the test, and the interpretation of the test scores. Further, this paper presents the difficulties of validating the underlying theory, the construct, the test, as well as the



interpretation based on test scores. It concludes by suggesting the practicability of assessing, in language proficiency test, test takers' ability of fulfilling specific target language tasks rather than their language competence. It also indicates the necessity of altering our view of construct validity.

Key words: construct; construct validity; language proficiency test

1 引言

语言测验质量的核心是效度问题,其中最本质、最核心的是结构(构想)效度问题^①(陈宏,1999)。语言测验的开发者需要在测验中建立良好的构想效度,然而在现阶段我们是否能够做到这一点?本文将首先讨论构想效度有效性的指向问题。笔者认为,语言能力测验中构想效度的有效性同时指向理论、构想、测量以及对测量的解释,进而本文将从上述四个方面讨论在语言能力测验中建立构想效度的困难所在。

首先需要声明,本文所讨论的测量(文中指语言能力测验)指的是探索性测量,是作为应用性测量的对立面出现的。张凯(2004b)指出,在语言测量中讨论构想效度有必要区分探索性测量和应用性测量。探索性测量指的是建立理论(尽管不是所有理论)、证明假说所必需的测量;应用性测量指的是测量的实际应用。我们同意张凯(2004b)的观点,即探索性测量是理论的组成部分,应用性测量是理论的产物;构想效度仅仅是理论探索阶段的问题,不是应用阶段的问题。这也是笔者选择在探索性测量中讨论构想效度问题的一个原因。另一个原因是应用性测量的有效性最终还是要靠相关的探索性测量的有效性来保证。

2 语言能力测验中的构想效度有效性的指向

克龙巴赫和米尔(Cronbach & Meehl, 1955)指出:“构想效度的研究,本质上是证明测验背后的理论有效。”(转引自张凯,2004b)然而,关于构想效度是什么有效的问题一直难有定论。张凯(2004b)总结归纳了前

^① 在本文中结构效度和构想效度指的是同一个东西,都是译自英语的“construct validity”。

人对构想效度有效性指向的研究，比较明确的观点大致有以下四种：构想本身有效，测验有效，测验或分数的解释有效，以及理论有效。他同意“证明构想效度就是证明一个理论有效”，而且强调，“在整个证明过程中，构想、测量和理论哪一个也不能先于另外两个获得有效性”（张凯，2004b）。

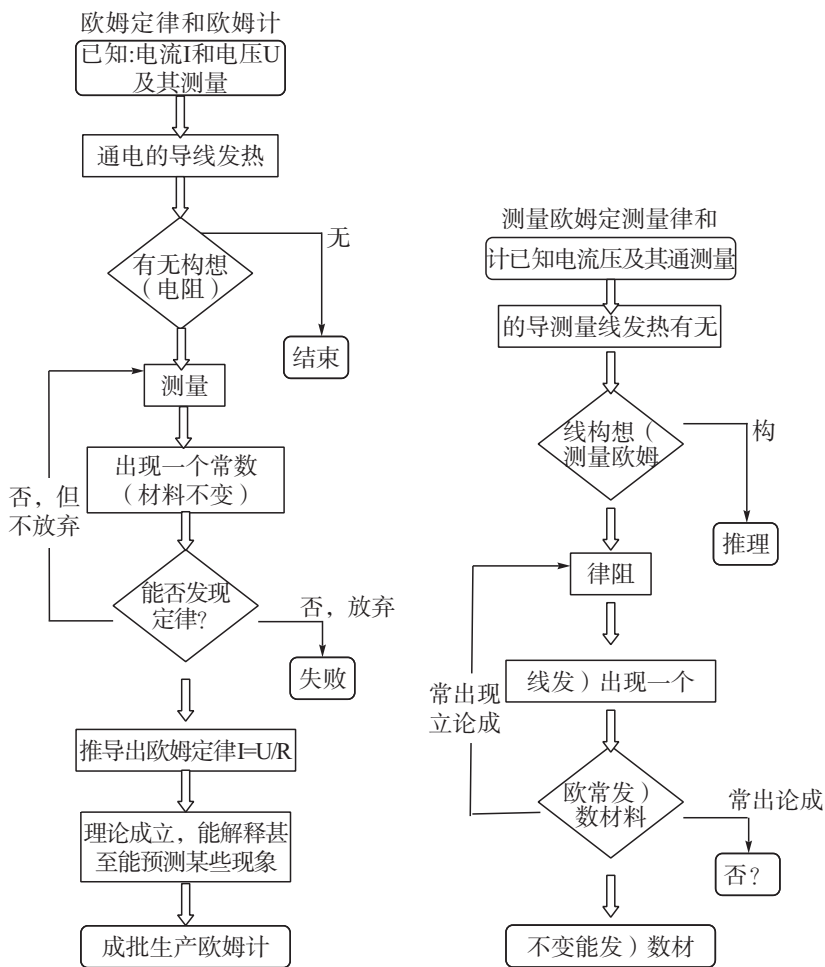


图1 欧姆定律的建立过程与语言能力测验的对比

笔者同意张凯所强调的内容。既然在构想效度的证明过程中构想、测量和理论哪一个也不能先于另外两个获得有效性，那么构想效度的证明在证明理论有效性的同时，也就包含了证明构想和测量的有效性。因而笔者认为，在语言能力测验中，构想效度的有效性是同时指向理论、构想、测验（测



量)、测验(测量)分数的解释有效,其中并不存在主次的问题,正如克龙巴赫和米尔(1955)所说:“我们不是先证明‘理论’而后证明测验有效,也不是反过来先证明测验再证明理论。当我们根据一类观察现象做出归纳推断时,我们考察的是整个理论网络与观察之间的关系。”(转引自张凯,2004b)。克龙巴赫和米尔的说法比较令人费解,而张凯(2004b)将欧姆定律的建立过程与语言能力测验对比则很能够说明问题。

由图1我们不难看出,在欧姆定律的建立过程中,理论、构想、测量以及对测量的解释的有效性是互为前提的:如果关于欧姆定律的理论是一个在现实世界中根本不存在的东西,那么关于电阻的构想、各种测量电压和电流的方法过程以及对测量结果的解释,都将只能是徒劳;然而,即使欧姆定律是客观存在的,如果我们关于电阻的构想是错误的话,科学家们无论如何测量电压电流,无论如何解释其测量结果,也不能得到欧姆定律;另外,即使欧姆定律(理论)事实上存在(但是有待于发现),(各种)电阻也存在(构想有效),如果测量电压和电流的方法不对,或者由于测量工具的影响,测量结果得不到一致性保障(测量无效),那么再怎么解释测量值,科学家们还是不能够证实现实中存在的欧姆定律(理论)或是电阻(构想);最后,如果科学家们对于所测量的电压和电流没有给予正确的解释的话,那么欧姆定律、电阻和测量电压电流的方法三者的有效性还是不能得到证明。而构想在这里所扮演的角色是提供一座连接理论与现实的桥梁,正如安娜斯塔西(Anastasi, 1982)所说:“……结构一方面来自某种理论,一方面来自对行为的各种测度之间已经确立的相互关系的推论……”(转引自陈宏,1997b)。无论构想先于理论还是理论先于构想,构想都是从属于理论的,正如在欧姆定律中,电阻(构想)仅仅是整个理论的一个组成部分;另一方面,构想的证明是基于对现实的观察基础上的,正如电阻存在的证明(构想的有效)是建立在对电流和电压的测量(观察的有效)上的。

上面的论述只是想说明理论、构想、测量以及对于测量的解释的有效性是一种“连坐”的关系。再来看构想效度,《心理学大词典》的定义是:“结构效度亦称构想效度,或概念效度,即测验在多大程度上正确地验证了编制测验的理论构想。”(朱智贤,1989)根据这个定义,可以肯定的一点是证明构想效度肯定要证明构想和理论有效,然而“正确”地验证则需要测量及对测量结果有效性的保证。再来看欧姆定律的例子。要证明欧姆定律建立过程中的构想效度,也就是要证明在多大程度上欧姆定律(理论)、电阻(构

想)、对于电压和电流的测量(测量)以及对于电压和电流的测量结果的解释(对于测量结果的解释)这四个方面是同时有效的,或者说是同时正确的。

由此笔者认为,在语言能力测验中构想效度的证实就是指语言能力理论、关于语言能力的构想、语言能力测验自身、对语言能力测验结果的解释这四个方面同时有效,缺一不可。构想、测量和测量的解释是建立或证明理论的必要环节,因而也是理论的组成部分。它们之间的关系可以由下图表示:

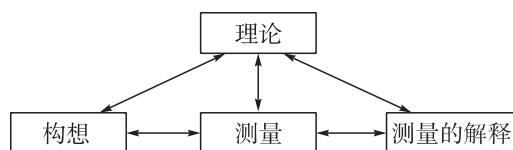


图2 理论与其组成部分的关系

现在面临的一个问题是,是否可能在语言能力测验中找到来自四个方面的足够的信息和证据,从而建立和证明构想效度?换句话说,如果要在语言能力测验中建立和验证构想效度,需要回答下述四个问题:语言能力的理论在多大程度上有效?关于语言能力的构想在多大程度上有效?测验本身在多大程度上有效?对测验的解释在多大程度上有效?事实上,正是在回答这些问题的時候,笔者遇到了困难。

3 语言能力的理论在多大程度上是有效的

张凯(1992)在讨论汉语水平考试(HSK)效度问题时指出:“我们对于理论建构的认识并不充分,因此没有建立起一个可以进行操作性定义的、可以被量化的关于语言能力结构的理论模型。”陈宏(1997b)指出,语言能力测验中结构效度所面临的理论上的挑战主要是语言能力的定义问题。事实上,围绕什么是语言能力这一问题的讨论几十年来一直没有定论。

拉多(Lado, 1961)从结构主义和行为主义的观点出发对语言能力测验进行解释,他认为语言测验必须包括两种变量:一种是语言本身的各种成分,另一种即是所谓的听、说、读、写四种技能。然而这种表述没有将语言能力与语言行为区分开来(陈宏, 1997b);巴赫曼(Bachman, 1990)指出



了其理论的另外两个不足:知识和技能在这种理论中关系不明,而且这个理论没有考虑到语言环境的作用。

乔姆斯基(Chomsky, 1965)对语言能力(competence)和语言表现(performance)进行了区分。在乔姆斯基那里,语言能力被理想化了,被认为指的是理想化的听/说者关于语言的潜在知识;是同质的语言知识或语法知识,是心理状态、心理器官或认知结构,是天赋的、绝对的心理特征;语言能力不是怎样使用语言知识的能力,不是过程或实际运用(戴曼纯, 1997)。乔姆斯基试图在剥离语言习得过程中社会文化因素作用的高度纯净条件下,通过演绎的手段研究最抽象的语言能力。遗憾的是,这种理想化的假设至今没有能为建立语言教学和语言测试关于语言能力的理论模型提供任何实际的帮助(陈宏, 1996)。

海姆斯(Hymes, 1972)从社会语言学的角度出发,认为乔姆斯基没有考虑语言是如何应用的问题,他进而提出交际能力的概念。海姆斯认为语言能力包括潜在的知识(knowledge)和使用知识的能力(ability for use)。但是,海姆斯没能清楚地指出语言能力的本质是什么,他的语言能力的概念是非常宽泛的,由此带来了理论应用范围方面的模糊。例如,承认运用知识的能力也是语言能力的一部分,就相当于把动机等非认知因素也划归到语言能力的范围中了。海姆斯的语言能力理论对于语言测验来说,也仅仅是有效的语言交际行为的一部分特征和通过这种语言行为观察推断交际能力需要参照的基本范畴(陈宏, 1996)。

巴赫曼(1990, 1996)将语言能力放到语言使用的互动模型中加以讨论,并且提出了交际语言能力的理论框架。图3反映了交际语言能力的各个组成部分。

从上文的论述不难看出,对于什么是语言能力这一问题学者们还难有定论,而且就现阶段学者们提出的理论来看,距离“可进行操作性定义的、可以被量化的”语言能力结构还有相当一段距离。还有一个问题是,即使我们有了满足条件的理论模型,我们又能够在多大程度上说这种模型拟合了人类所具有的语言能力这种特质呢?正如爱因斯坦提出相对论之前,很多人认为牛顿的经典物理学理论能够解释自然界所有的现象,然而事实并非如此。这可能是一个更高层次上的问题:理论总是在不断地完善和演进的。但是另一方面,理论在某种程度上的被证伪是建立在那个理论能够为我们所证伪的前提下的:牛顿的经典物理学理论在微观领域能够被证伪是因为其本身具有这

种被证伪（或证实）的特质，或者说是一种科学理论的特质。但是目前的语言能力理论还缺少这种科学理论的特质——我们很难通过科学测量的手段对现有的语言能力理论进行证伪或者证实，这也造成了目前对什么是语言能力这一问题争论不休的局面。

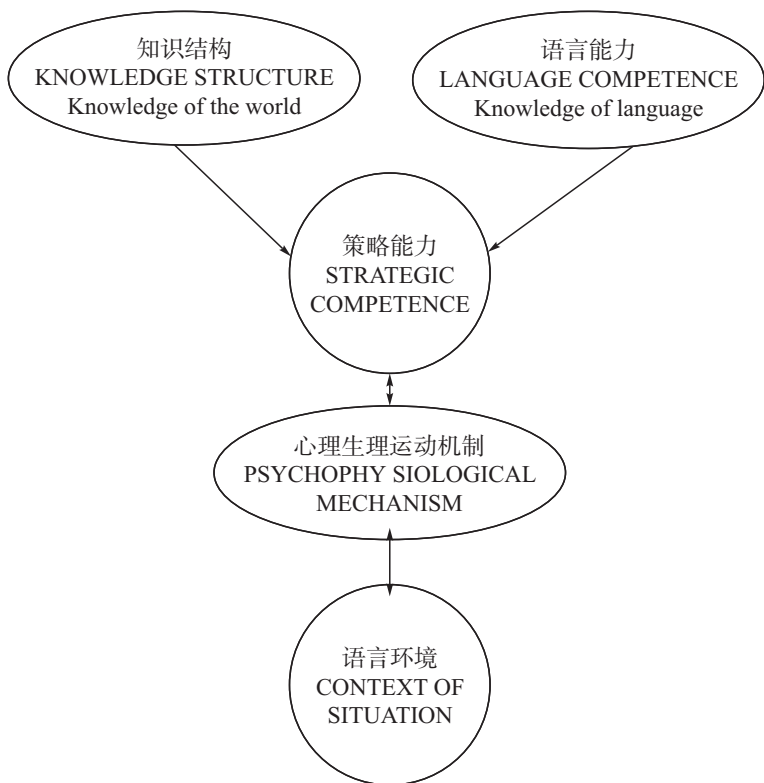


图3 交际语言能力的组成部分

然而，笔者认为，在语言能力测验中对语言能力理论的争议是可以暂时搁置的，因为任何理论在被证实之前都只能是假设成立的。正如张凯（2004b）所说：“在建立理论的过程中，测量不一定总是在提出定律之前进行……很多时候，测量一定要在提出定律之前进行，欧姆定律、量子论、语言测验、心理测验都是这种情况。”韩宁（Henning, 1987）说：“某个测验被说成是有效的，是指其在多大程度上测量了他应该测量的东西。因此我们可以推论‘有效’这个词，但用于描写一个测验时，常常应该配以介词‘对于’意义才会完整。因此任何测验对于某个目的都可以是有效的，而对于其



他目的又都是无效的。”事实上，如果我们承认语言能力测验中的构想效度是同时针对理论、构想、测量和对测量的解释的话，我们可以用其他三部分的有效性来证明理论的有效性。奥勒（Oller，1979：266）认为：“理论本身能够为与理论拟合得较好的测验程序提供效度的证据，因为理论本身在不同于语言测验数据和研究的效度检验中是与现实相连的。”^①正如在欧姆定律这样一种理论被证实之前，它只能作为一个假设成立的理论；恰恰是电阻（构想）、电压电流的测量（测量）和对所测到的电压和电流的正确解释（对于测量的解释），使得我们证明了欧姆定律的有效性；而欧姆定律的有效性是在随后的无数次测量中得到验证的。正如张凯（2004b）所说：“要使理论有效，构想、测量、定律^②三者必须同时有效”，“整个理论及其组成部分的有效性是有数学的有效性保证并有经验事实证明的”。

所以，在语言能力测验中，我们可以先根据某一个语言能力的构想（这个构想可以是基于某一种语言能力的理论，也可以是先有这个构想，而后再推导出相应的理论）来编制测验，通过对测验结果的解释来证实或者证伪某一种理论。正如我们可以先有电阻这个构想，设想它和电流、电压的关系（一个理论），然后通过测量和相应的解释来检验这种理论。如果理论成立，那么构想是有效的。这看起来似乎可行，然而一个重要前提是构想、测量以及对测量结果的解释必须同时有效，但是我们能够做到这一点吗？

4 语言能力的构想在多大程度上是有效的

梅西克（Messick，1975，转引自 Bachman，1990：255）说：“一个测量对人们拥有或显示的某些东西的量进行评估。一个（关于构想验证的^③）基本问题是：那个东西的本质是什么？”巴赫曼（1990：255）认为：“要回答这样一个问题，我们必须确认和定义我们想要测量的东西是什么，当我们对那是什么做出定义时，我们实际上是在定义一个构想。”

① 原文是“The theory can be used as an argument for the validity of proposed testing procedures that accord well with it because the theory itself is anchored to empirical reality through validity checks other than language test data and research”，参见 Oller（1979），⑩Language Tests at School⑩，p. 266。

② 我们把这里的定律理解成对于测量的解释。

③ 原文是“construct validation”。

事实上，在语言能力构想的有效性问题上，我们一直热衷于“定义”一个构想，而很少关注我们希望测量的语言能力的本质。这种对所测东西本质属性的忽视很可能造成我们长时间无谓地浪费大量时间和精力来测量一种根本不能为我们所测量的东西。以电阻为例，电阻是我们的一个构想，围绕电阻的所有测量及其解释是基于这样的一个假设，即电阻这种东西是可以为我们所测量的。同样，我们围绕语言能力所做的各种测量也是基于这样一个假设，即语言能力是能够为我们所测量的。但是事实上是这样的吗？我们又怎么能够证明语言能力这种人类的特质是可以为我们所测量的呢？

沃利斯（Wallis，2004）指出，在心理测量领域，对智商和性格结构的争论持续了一个多世纪却依然没有定论，一个主要原因就是研究者们“忽略了一个重要前提，即证明那些被测量的变量事实上是可以被测量的”。他引用米歇尔（Michell，1990）的研究，指出这些可以被测量的变量应该具有两个性质——顺序性（ordinality）和可加性（additivity）。所以在证明“那个东西”可以为我们所测量之前，任何有关结构的争论都没有太大的意义。只有当一种属性“可以作为一种可加的性质进行处理时，他才可以被描述成一种可测量的性质”（特拉斯特德，2002：70，转引自张凯，2004c）。沃利斯（2004）引用米歇尔的研究指出：要测量任何变量，第一步就是要证明这些变量具有和数字系统一样具有经验性的结构，而且这些变量所具有的经验性结构是独立于数字结构的。

正如心理测量中所面临的问题一样，语言能力测验首先要证明的是语言能力是可以为我们所测量的。张凯（2004b）认为，与构想效度有关的理论（语言能力理论、智力理论）是一种量化的理论，可以且必须用数学来表示。但是我们怎么才能判定语言能力理论是一种量化的理论呢？我们显然遇到了一个悖论：要证明语言能力能够被量化就需要通过量化的手段，然而我们在利用任何量化手段来证明语言能力这种特质（能够被量化）的时候都不能保证其有效性，因为我们不知道语言能力能否被量化。这是一个很令人头疼的问题，很可能语言能力这种特质的证明是与整个理论的证明同步的。

5 测验本身在多大程度上有效

格雷戈瑞（Gregory R. J.，2004：108）认为，所有心理构想具有的一个共同的特征就是“依据现有的关于一个构想的理论能够得到一个相互联系



的假想的网络”。

在编制语言能力测验时，我们所依据的就是基于关于构想的理论的相互联系的假想的网络。克龙巴赫和米尔（1955）认为，所谓构想，就是“人所具有的某个假设的特质（attribute），这个特质被假定为是能够在测验表现中有所反映的”，而且“拥有某一特质的人将会在 X 情境中以 Y 方式行事，而且两者之间有一种明确的概率（stated probability）关系”。这就是说，通过对语言行为的观察，我们可以推知语言能力，并且两者之间具有一种稳定的关系；而找出这种关系的手段便是语言测验。由于语言能力难以直接观察，我们为了得到可供推论的样本，必须对什么是语言能力进行操作性定义，根据操作性定义来设计测验选区相关的语言行为样本。这就是说，语言测验既是对语言能力操作性定义的结果，也是记录语言行为的工具。

语言测验的这种双重属性带来了一个问题，就是测验本身要同时满足解释的充分性和描写的充分性（陈宏，1997b）。

解释的充分性是针对语言能力的操作性定义的。自从奥勒（1983）承认一元特质假说的最强式是错误的以来，人们倾向于将语言能力看成是多个要素的综合体。相应地，我们倾向于把语言能力测验分成几个不同的分测验，分别对应语言能力的多个方面。理想的状态就是我们的分测验所测的仅仅是我们希望测到的那个语言能力的要素，而非其他要素。我们越是能够保证所测能力要素的单一性，基于测验结果的解释可信度也就越高。

但是在实际情况中这是很难实现的，而且也可能没有必要去实现。一方面，在实际操作中我们是通过听、说、读、写等不同语言行为来测量的，而其中的任何一方面都显然涉及整体的语言能力而非语言能力的某个要素。正如陈宏（1999）所说：“我们很难想象，在听、说、读、写测验中，我们又能观察到什么别的，而脱离了听、说、读、写，我们还能观察到什么。”另一方面，从真实性角度来说，语言能力测验不是为了测语言能力而测语言能力，而是“希望并且相信，我们可以从他在某个特定场合、某些数量有限的任务上的行为方式，对他在所有其他有关场合的行为方式做出概括”（Ingram, 1977, 转引自陈宏，1997b）。这就要求测验本身具有充分的真实性，因为真实性“为我们提供了一种工具，这种工具使我们能够考察对分数的解释在多大程度上能够超越测验表现而推广到目的语使用域（target language use domain），或是其他相同的非测试性质的语言使用域（nontest language use domain）”（Bachman, 1996: 24）。