

Experimental Research Method
in Applied Linguistics

应用语言学

实验研究方法



Z分数	-2	-1	0	1	2
T分数	30	40	50	60	70
CEEB分数	300	400	500	600	700
TOEFL分数	400	450	500	550	600
标准差	1	3	5	7	9
正态曲线每部分	1.37	28.94	50	71.06	92.12
百分位	2.28	15.87	50	84.13	97.73

林连书 著

中山大学出版社

应用语言学实验研究方法

林连书 著

中山大学出版社

·广州·

版权所有 翻印必究

图书在版编目(CIP)数据

应用语言学实验研究方法/林连书著. —广州:中山大学出版社,2001.8

ISBN 7-306-01806-X

I. 应…

II. 林…

III. 应用语言学—实验方法

IV. H08-33

中国版本图书馆 CIP 数据核字(2001)第 13372 号

中山大学出版社出版发行

(地址:广州市新港西路 135 号 邮编:510275

电话:020-84111998、84037215)

广东新华发行集团股份有限公司经销

广东从化市印刷厂印刷

(地址:广东从化市街口新村路 15 号 邮编:510900 电话:020-87916882)

787 毫米×1092 毫米 16 开本 15 印张 368 千字

2001 年 8 月第 1 版 2001 年 8 月第 1 次印刷

印数:1-1000 册 定价:28.50 元

如发现因印装质量问题影响阅读,请与承印厂联系调换

序

林连书老师的著作《应用语言学实验研究方法》出版，我感到高兴。这有三层理由：

首先，本书所讲的是一个重要然而人们不大熟悉的题目——应用语言学（主要是英语教学）的实验方法。大家知道，英语教学研究有两条途径：一条注意总体效果，一条注意细部分析；前者多用定性方法，后者则用定量方法。我们有些教师英语教得很好，可是问他们教学目标是什么，所用手段是什么，什么地方最成功，什么地方不满意，往往答不出或说不全，更提不出什么数据。原因是，他们向来只着重状态描写，不搞概率统计。他们的劳绩应当受到重视，但是在定量分析方面最好下些工夫，弥补自己的不足。

其次，本书的问世，正是我国大搞教学改革，但英语应用语言学实验工作还缺乏必要的指引的时候。把数理方法运用于语言研究和英语教学，目前在国内还处于萌芽时期。心理测验，有人出版了专著；外语测试，有人已写了若干篇文章，若干本专著。可是在应用语言学实验方面，应该如何设计，如何做具体工作，如何凭统计数字检验自己的假设，我这孤陋寡闻的人看到的参考资料不多。林老师这本书的出版，令人欣慰。

最后，本书著者具备了做他所做的工作所需要的学历、能力和经验。林老师曾在大学本科专攻数学五年，一般英语教师很少具有这样的条件。他在中山大学英语培训中心教英语已经教了20年，并曾两度赴美国深造，最后一次是专门研究应用语言学实验方法的。由于情况熟，资料新，数据多，他讨论问题时左右逢源。这里还不妨提一下：本书初稿已在中山大学外国语学院对硕士生、博士生和澳门研究生教过多次，反映良好。

也许有人说，这本书沉甸甸的，共有200多页，我怎么知道它的内容对我适合不适合。关于这个问题，我看是有办法的。古语云：尝鼎一脔，便知其味。这本书有多少用处，请试读7-8页（那是讨论如何选定研究题目的）就知道。

王宗炎
于中山大学回春楼
2001年1月

前 言

当今英语应用语言学研究的特点是范围大、方法多、内容丰富。在应用语言学发展之初，它研究的内容主要是英语作为第二语言教学的问题。但是，在短短几十年的时间里，它已经逐步发展到对所有有关语言和语言学的各种实际问题的研究，形成了各种各样的语言理论体系和语言运用模式。应用语言学从解决实际问题的目标出发，广泛地用到社会学、心理学、人类学、信息科学、概率统计等学科的知识，因而开拓了很多前所未闻的新学科和边缘学科。实际上，应用语言学的研究范围已扩大到包括普通语言学、语音学、语法学、语义学、语用学、词汇学、语段分析、文体学、社会语言学、心理语言学、专用外语、第一语言和第二语言习得、方言研究、移民语言研究、语病学、神经语言学、语言教学法、语言教学大纲设计、语言规划、语言测试、数理语言学等各个方面。而对某一个学科的问题的研究，往往触发对另一个问题或一个新学科的研究。例如，当语法研究深入到深层结构、语义功能和语言运用的心理过程时，便触发了对转换生成语法、生成语义学和格语法等的研究。又如，当语言学研究涉及语言习得和运用与社会和心理过程的关系时，就产生了社会语言学和心理语言学。由于应用语言学的研究涉及这么多的学科，因而研究方法也有了很大的变化。以往的语言学研究，往往偏重于理论的建立和文献资料的分类、综合和整理，因而常用的方法是综合、总结、分析、对比等等。这种依靠图书资料进行的研究，国外有些学者称之为 secondary research。叫它 secondary 是因为它所用的资料是第二手的。这些方法到目前为止还有其使用的价值。但是在应用语言学的研究中，涉及的问题大部分是实际的问题，有的是前所未有的，有的是前人曾经提及的经验、方法或是未经验证的理论和假设，光用 secondary research 便很不够。因而有些语言学家把原来属于自然科学的研究方法移植到语言学研究中来，这就产生了所谓科学研究方法（也称心理测量法）。科学研究方法强调实验设计、数据处理和假设检验。一个好的实验设计可在某个具体的研究中得到无可辩驳的结果，而且其结果可以经得起时间、空间和重复实验的考验。由于科学实验方法有这些优点，因而在应用语言学的研究中得到越来越广泛的应用。在美国，相当大一部分应用语言学期刊发表的论文是实验性的。越来越多的语言学博士、硕士研究生在写学位论文时，采用实验性研究方法。除了科学研究方法外，目前国外广泛采用的方法还有民族志研究（ethnographic research）、个案研究（case study）、内省法（introspective methods）、相互作用分析（interactional analysis）等。这些方法各适用于某些特殊问题的研究，各有其特点和优缺点，因而在不同的场合中得到了应用。上述的这些方法国外叫做 primary research，因为研究人员得到的资料是第一手的。研究范围的扩大和研究方法的多样化应用为应用语言学研究注入了很多新的内容。例如，在言语习得研究中，个性、资质、理解能力、学习动力等难以直接观察的心理因素与言语习得的关系，用图书资料研究法是很难进行研究的，用心理测量方法就可以顺利进行。

本书主要讨论 primary research，而科学实验方法是其重点。很多语言学研究人员对

实验研究方法缺乏了解，不知道对什么问题用什么实验设计、采用什么数据处理的方法、如何用假设检验法来验证自己原来的假设，所以，有时尽管有良好的愿望，仍然得不到令人信服的结果。本书对上述问题都有详细的讨论。所以，读完这本书，读者就可根据实际情况，对要进行研究的问题选用适当的研究方法。

在实验性研究方法中，概率统计推理是极有用的工具，所以本书用较大的篇幅介绍了概率统计的基本原理和方法。概率统计对有些语言学研究人员和师生来说是比较困难的，但是，不懂得这种方法，就失去了一个重要的语言学研究工具，不但难以使用实验研究方法，而且难以看懂国外最新发表的、涉及实验研究方法的研究文章。所以，本书试图以深入浅出、逐步引入的方法，使读者了解概率统计的基本理论和方法。在描述中，尽量避免繁复的数学推导和令人目眩的符号，着重介绍原理和应用方法。数理基础较好的读者可适当了解某些公式的来源；数理基础比较薄弱的读者可避开这些内容，着重于公式表示的意义和如何应用。书中对每种实验研究都附有实例，而且采取先举例、后说明的方法，这样读者便比较容易看懂，并可举一反三，对类似的研究问题采用类似的研究方法。

概率统计方法与实验设计是分不开的。在提出原假设的阶段，研究人员就必须明确其实验设计和将要进行的数据处理和假设检验。否则，研究就是盲目的。研究必须有系统性、逻辑性、实践性、可行性和可重复性，实验结果必须能经受可靠性、有效性的检验。本书对这些关系到研究的科学性的问题都有所讨论。这样，读完了这本书，读者在进行实验性研究时，考虑问题就会更加周密，就能防止研究中容易产生的漏洞。

本书第一章介绍了应用语言学研究方法的分类和实验性研究方法的特点。第二章至第六章介绍了实验设计和假设检验所需的预备知识；第二章讨论了随机事件及其概率，第三章讨论了总体与样本的关系和样本的频率分布，第四章、第五章分别介绍了离散型和连续型的各类概率分布，第六章讨论参数估计。接着，第七章介绍了多种实验设计和假设检验法，学习了这些设计和检验法，就能掌握数据处理的方法。第八章讨论了另一种研究方法：相关系数法以及一元和多元线性回归。在此之后的第九、第十章讨论了实验性研究必须具有的两个特性：信度和效度。第十一章讨论语言测试中的题目分析。第十二章讨论多变量研究方法，其中包括主成分分析和因素分析；第八章中多元线性回归也是多变量研究的一部分。在本书的最后一章，我们介绍实验性论文的写法，并从国外语言学期刊中选出一些研究实例，对它们的设计、结果和存在的问题进行讨论。

当今世界，电脑得到越来越普遍的应用，本书所介绍的统计方法，都可以通过电脑统计软件（如 SPSS）从电脑中得到结果。所以，在有电脑的地方，使用实验研究方法就更加方便。但是，应该指出的是，电脑的工作效率虽然很高，但却不能完全代替人的工作。研究人员必须作好研究设计，告诉电脑要做什么，电脑才能按照你的指令去工作。从电脑中得到数据分析的结果后，研究人员还必须对这些结果进行讨论，得出结论，指出实验中的问题，对以后可能进行的下一步的研究进行展望，这些都是电脑所做不到的。所以，研究人员在有可能的情况下，要多学一点实验研究方法和与它有关的统计学。这样不但能大大提高自己的科研能力，而且使电脑得到更充分的应用。

概率统计方法广泛应用于自然科学（如应用数学、医学、生物学、化学、气象学等）和社会科学（如经济学、社会学、心理学等）。因而，掌握了这种方法，语言学研

究人员甚至可以参予经济、管理、社会学、心理学等与统计学有关的学科的研究工作，对学科的相互借用、渗透和促进有很大的作用。这就有利于建立新学科、边缘学科和新的研究领域，使语言学研究从其他学科的研究中得到营养。

本书适用于任何对英语研究有兴趣的读者。语言学研究人员、外语院校和大学外语系师生、中学英语教师、在外国语言学专业学习的留学生，都是它的对象。其他对实验性研究和标准化测试有兴趣的读者，也可以从本书中学到基本的原理和方法，扩大眼界，增长知识。

近十多年来，应用语言学实验性研究在我国有较大的发展。例如，广东外语外贸大学的桂诗春、李筱菊等教授在英语测试理论方面进行了比较全面、深入的研究，在全国推广了英语标准测试，并培养在这方面有专长的研究生。桂诗春、宁春岩教授的《语言学方法论》于1997年出版。严辰松教授也于2000年出版了《定量型社会科学研究方法》。中山大学广州英语培训中心长期在开设的硕士研究生应用语言学专业中讲授统计学和英语测试的理论和方法，因而硕士研究生的学位论文几乎全部都是实验性的。在我国的语言学期刊和学术论文汇编中，近年来出现了一些实验性研究论文。老一代的语言学家，如中山大学外语学院的王宗炎教授，对新一代的语言学者总是寄以厚望，鼓励他们努力进取，掌握新理论、新方法，促进语言学的发展。时代对应用语言学的发展是很有利的。我们希望有更多的中青年英语教师和英语研究爱好者，投身到应用语言学的实验研究中来，真正学会如何作实验性研究，而不是空谈别人的研究。

我要感谢美国 San Diego State University 语言学与东方语言系的 Charlotte Webb, Ann Johns 和 Soonja Choi 三位教授，他们为我提供了大量有关言语习得、心理语言学、英语阅读写作实验性研究的论文和书籍，使我从中了解到国外语言学界在实验性研究方面的发展。系主任 Charlotte Webb 曾让我为系里部分研究生开有关实验性研究的讲座，帮助他们搞实验设计和统计，使我在实践中学到不少东西，同时也了解了学生在实验研究中存在的问题。在我写这本书的过程中，他们不断给予我鼓励，并提供多方面的支持。没有他们的支持，这本书的完成是不可能的。

同样，我也要感谢中山大学外语学院、广州英语培训中心的领导和同事们给予我的时间上、精神上的支持。没有他们的支持，这本书的完成也是不可能的。

为此，我把这本书献给所有帮助过这本书的完成的人们。

林连书

2001年1月

目 录

第一章 应用语言学实验性研究	(1)
第一节 应用语言学研究的类型	(1)
第二节 实验性研究的特点	(4)
第三节 实研性研究的过程	(5)
第二章 随机事件、概率和随机变量	(11)
第一节 随机现象和随机事件	(11)
第二节 随机事件的概率	(13)
第三节 变量及其类型	(18)
第四节 变量的量表	(20)
第三章 总体与样本	(22)
第一节 总体与样本的定义	(22)
第二节 随机抽样	(23)
第三节 样本数据的初步处理	(24)
第四节 样本的数字特征——平均值、标准差和方差	(30)
第四章 离散型随机变量的概率分布	(36)
第五章 连续型随机变量的概率分布	(44)
第一节 连续型随机变量	(44)
第二节 正态分布	(44)
第三节 t 分布	(55)
第四节 χ^2 分布	(56)
第五节 F 分布	(58)
第六章 参数估计	(60)
第一节 参数的点估计	(60)
第二节 参数的区间估计	(62)
第七章 假设检验	(69)
第一节 U 检验法	(69)
第二节 t 检验法	(72)
第三节 χ^2 检验法	(79)
第四节 方差分析和 F 检验法	(86)
第八章 相关系数和线性回归	(120)
第一节 相关系数	(120)
第二节 一元线性回归	(129)
第三节 多元线性回归和复相关简介	(136)
第九章 信 度	(143)

第一节	信度的基本定义	(143)
第二节	信度的种类和计算方法	(145)
第三节	影响信度的种种因素	(150)
第十章	效度	(153)
第一节	效度的种类	(153)
第二节	影响效度的种种因素	(157)
第三节	效度与信度的关系	(158)
第十一章	题目分析	(160)
第一节	题目的难度	(160)
第二节	题目区分度	(162)
第三节	答题差异性	(164)
第十二章	主成分分析和因素分析简介	(166)
第一节	主成分分析	(166)
第二节	因素分析	(170)
第十三章	实验设计与效度的关系及实验性研究论文的写法	(173)
第一节	内部效度和外部效度	(173)
第二节	实验性研究设计的主要类型及其与效度的关系	(175)
第三节	实验性研究论文的写法	(179)
结束语		(185)
附录 A	统计表	(186)
附录 B	语言学研究参考资料来源	(207)
附录 C	应用语言学实验性研究常用术语 (英汉对照)	(213)
附录 D	实验性研究实例	(223)
参考书目		(230)

第一章 应用语言学实验性研究

第一节 应用语言学研究的类型

进行应用语言学研究，首先要弄清什么是研究。Nunan (1992) 曾叫修习研究方法课的研究生回答下列两个问题：

- 1) 研究是什么？
- 2) 研究的目的是什么？

对第一个问题，学生有如下的答案：

1) 研究是对问题的探讨。研究包括过程和结果两部分。研究过程与探讨的领域和方法有关。结果是在研究中得到的知识。

2) 研究是一个包括提出问题、说明研究目的和作出假设的过程。它通过对资料的收集、分类、分析和解释实现研究目的。

3) 进行系统的考察，更好地了解自己从事的领域，并使人们了解它。

4) 研究是一个分析和评估某个问题的活动。

5) 收集和分析某个具体领域的资料，从而证明你的理论。

6) 根据某些预定的方法在某个领域中提出问题，进行评估、考察和分析，证实假设，综合事实，收集和分析资料。

对第二个问题，也有各种各样的答案：

1) 以科学的方法客观地，而不是主观地取得结果。

2) 解决问题，证明理论的应用，更好地洞察事物。

3) 启迪研究人员和有兴趣的读者。

4) 证明或否定新的或已有的概念，找出某些现象的特征，达到个人或集体的目标。

5) 通过证明或反证、澄清事物、有计划的探索来支持某个观点，揭示未知事物的原因，找出问题的解决方法，等等。

上面的回答有一定的代表性，也有一定的道理，都从某个角度上勾划出了应用语言学研究中的某些特征。要找出一个包罗万象的定义是很困难的。Nunan 从这些回答中归纳出一个最低限度的定义：

Research is a systematic process of inquiry consisting of three elements or components:

① a question, problem, or hypothesis, ② data, ③ analysis and interpretation of data.

这个定义指出，研究是一个系统的探索过程，它至少包括问题或假设、资料和对资料的分析和解释三部分，缺少其中任何一部分的活动都不能算是研究。

研究有各种各样的类型，但是如果根据收集资料和分析资料的方法来看，研究可分成定性研究和定量研究两种。定性研究假定知识是相对的，任何研究都带有主观的成

分。研究者可通过文字资料或是逻辑推理来探索语言现象的规律、性质和特征，建立理论和模式。它强调总体印象，而不寻求事物的数量关系；它只揭示现象，却不对现象作类推。这种研究方法广泛地用在对付图书资料的综合、分类、分析、对比、比较和理论或假设的建立上。定量研究特别注重研究的客观性，强调用事实和数字来说明问题。不少人认为，应用语言学研究提出来的理论和假设，必须用科学的方法进行验证才能接受。所以这种研究方法特别注重实验设计，对实验条件进行人为的控制，以期得到预想的结果，并通过统计推理得出结论，作出结论。它常常采取缩小范围、孤立变量、逐个击破的方法来进行，所以解决的问题常常是一个大问题中的一个个具体问题，与定性研究强调总体印象形成对照。从应用语言学研究的历史和现状看，两种方法各有特点，都有存在的价值。因此，从实用的观点来看，应该避免采取极端的立场，使两种方法互相对立，而应该充分利用它们各自的长处，互相推动，互相补充。实际上，在不少应用语言学著作中，既用到定性分析，又用到定量分析。在讨论大的问题时，用定性分析多些；在讨论一个个具体的问题，又牵涉到科学性时，用定量分析多些，这不能不说是一种好的方法。

Chaudron (1988) 试图突破定性和定量的这种区分，提出应用语言学研究的四个传统方法：心理测量法、相互作用分析法、语段分析法、民族志研究法。心理测量法即本书将要详细讨论的科学实验研究法；相互作用研究法是通过有系统、有计划、有计划的观察来研究师生的相互作用对学习行为影响的程度；语段分析法是将在课堂上收集到的资料进行语言学上的分类和分析；民族志研究法是通过自然的、无控制的观察和描述，将课堂看成一个文化系统来进行研究。Chaudron 的分类法比较具体；但是他把相互作用法和语段分析法当作两种独立的研究方法列了进去，但又不包括其他一些重要的研究方法。

Grotjahn (1987) 认为把研究分成定性和定量两种是把问题简单化。他认为要将研究分类，必须考虑三个因素：一是收集资料的方法（实验性的还是非实验性的）；二是资料的类型（定量的还是定性的）；三是对资料的分析（统计性的还是解释性的）。将这三个因素互相配搭，可以得到很多不同的研究类型，其中最基本的两种是“分析—法理性”（analytical - nomological）模式和“探索—解释性”（exploratory - interpretive）模式。前者是用实验性方法收集资料，产生定量资料，并对数据进行统计分析；后者是用非实验性方法收集资料，得到定性资料，并对资料进行解释性的分析。除了这两种最纯粹的方法外，Grotjahn 提出另外六种方法组合：①实验—定性—解释性（experimental - qualitative - interpretive）；②实验—定性—统计性（experimental - qualitative - statistical）；③探索—定性—统计性（exploratory - qualitative - statistical）；④探索—定量—统计性（exploratory - quantitative - statistical）；⑤探索—定性—解释性（exploratory - qualitative - interpretive）；⑥实验—定量—解释性（experimental - quantitative - interpretive）。

Grotjahn 的分类方法在一定程度上反映了应用语言学研究的实际。很多专著和学术论文都在实验—非实验、定量—非定量、统计—非统计这些方法中作出选择，来适应各种不同的研究目的、环境和内容。在不同的情况中又会有所侧重，互相渗透。但是，也不能不承认，定量分析和定性分析还是代表了研究方法的主流。

Van Lier (1988; 1990) 则认为应用语言学研究可以根据两个参数进行分类：一是

干预者 (interventionist), 一是选择性 (selectivity)。干预者参数指的是实验人员对研究环境干预的程度。例如, 有些研究是在实验条件下进行的, 实验人员对研究环境干预很大; 有些是自然研究 (naturalistic study), 对研究环境不作干预。选择性参数指的是研究人员事先指定研究对象的程度。例如, 有些研究要事先指定变量, 有些则是民族志方法研究, 等观察到结果再说。由这两个参数也可对不同的研究方法归类, 即分成干预性高的或低的, 选择性高的或低的。本人认为, 这样的分类法揭示了研究的一些特点, 但不够全面, 似乎不能包括所有的研究。

值得一提的是 Brown (1988) 的分类法。Brown 于 1980~1982 年在广州中山大学广州英语培训中心工作期间, 就曾在应用语言学实验研究方法的讨论班中提出过这种分类法。他根据研究资料的来源将研究分成第一手研究 (primary research) 和第二手研究 (secondary research)。第一手研究指的是对第一手材料的研究, 如实验性研究、调查 (survey)、个案研究 (case study) 等等。如果按照这个定义, 所有进入实地获取第一手材料的研究, 如民族志研究、相互作用分析、自然观察、内省法 (introspective methods) 等, 都可以列入这个范围。第一手研究又分成统计性和非统计性的。例如, 实验性研究和调查就是统计性的, 个案研究、民族志研究就是非统计性的。与第一手研究相反, 第二手研究就是对第二手材料的研究, 这主要是图书资料的研究, 如对材料的综合、分析、分类、比较、对比、理论和假设的建立等等。这种分类法可用树形图表示如下:

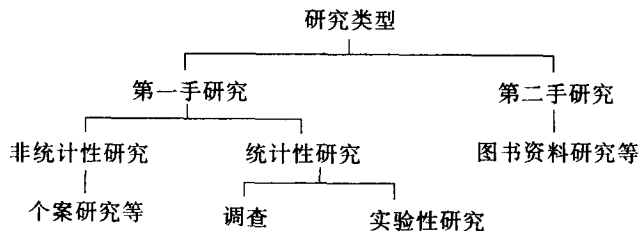


图 1.1 应用语言学研究的类型

Brown 的这个分类法简单明了, 突出了应用语言学研究的特点。按照这种分类法, 我国以往的研究以第二手研究居多。这种研究在我国有悠久而良好的传统; 各种字典的编写、语言学理论著作、语法书、教科书的编写, 大部分基于这种传统。因而这种研究还会继续下去, 在应用语言学研究中继续发挥它的作用。但是, 应当指出, 第一手研究, 特别是实验性研究, 目前在国外已经比较普及。不少研究人员都乐于使用它。大学语言学系的本科生和研究生都需要阅读大量第一手研究文章, 以加强学习和研究的能力。本科生和硕士、博士研究生在写论文时, 常常要进行独立的实验研究, 并作统计分析。不少应用语言学研究论文, 都是实验性论文。在其结果部分, 常常有大量的图表和统计数字。它们所用的研究工具, 牵涉到实验设计和大量统计推理的理论和方法。所以, 对不熟悉第一手研究的、特别是实验性研究的人来说, 有些论文是很难看懂的。既然应用语言学研究存在另一些重要的研究方法, 既然这些方法在我国比较生疏, 我们就应对其进行研究。近年来, 我国语言学界与国外接触明显增多, 新的语言理论、教学方法、图书资料大量输入我国, 我国语言学界在实验性研究方面有一些进展。但是总的来

说，这种方法不够普及，了解和应用它的人不多，深入研究它的人更少。因此，除保持我们在应用语言学研究的优点和特点之外，还应该学习、研究和推广新的研究方法。为了这个目的，这本书将重点讨论第一手研究，特别是实验性研究，而将第二手研究留给在这方面富有经验的学者们去讨论。

第二节 实验性研究的特点

实验性研究具有很强的探索性。它表现在：①不盲目相信人家提出来的理论和假设。理论和假设必须用科学的方法进行检验，才能知道它们的真伪。②实验性研究必须揭示规律性。例如，对一组中国留美学生作应用语言学实验，应该研究得出的结果在多大的范围和多大程度上反映了语言学上的某些规律。③对别人做过的实验性研究，可以对它们的原理、方法和结果进行再检验，也就是说，可以对同一个问题再做一次实验，并对结果进行比较。④可以改变实验内容和实验条件，扩充别人的研究成果，探索新的成果。⑤可以指出自己实验的局限性，让别人去补充，扩大自己的研究。Brown (1988) 提出实验性研究几个突出的特点，这就是系统性、逻辑性、触感性 (tangible)、可重复性 (replicable) 和可说服力 (reductive)。所谓系统性，是指研究必须遵循明确的步骤，有一整套实验设计、统计程序和假设检验方法，能保证研究的信度和效度。例如进行相关研究，可通过相关系数法和一元、多元线性回归去解决；进行因果关系的研究，可用诸如方差分析的方法去解决。所谓逻辑性，是指研究遵循科学的推理方法。假设检验就是人们从长期的实验性研究中总结出来的一套推理方法。例如，在以后将要详细讨论的 t 检验、 χ^2 检验和 F 检验中，可以对从不同组观察到的数值进行数理统计，用反证法来检验不同组间是否存在显著差别；换句话说，就是可以检验观察到的差别是随机因素引起的，还是某些研究中的因素引起的。第三个特点是触感性，它指的是资料来源于真实世界。实验性研究得到的资料应可以化为数据，进行定量分析。可重复性是指实验研究应经得起重复试验的考验。实验人员应对自己的实验对象、材料、方法和结果作详细的叙述，使别人能在同样的实验条件中重复自己的实验，以证实结果的正确性。最后一个特点是可说服力，指的是实验结果应回答提出来的问题，澄清事实，以事实和检验结果支持或否定某种理论或假设，并通过推论，揭示一定的规律。所有这些特点可归结为实验性研究的信度和效度。

信度也叫可靠性，指的是研究的准确性和稳定性。信度高的研究误差小，在同样的实验条件下做多次同样的实验，得出的结果应该十分相似。信度可分为内部信度和外部信度。内部信度指的是在同一次实验中，得到的数据误差小，变化小，也即具有统一性 (consistency)。例如，在英语测试中，两个评分员对同一组考生的某个技能进行评分，如果得出的结果类似，我们就说测试的评分员相关信度高。评分员相关信度是内部信度的一种。也可以将测试再做一次，如果得到的结果与第一次相似，我们就说测试具有重复测试信度。重复测试信度也是内部信度的一种。实验性研究有多种方法测定内部信度，我们将在第九章作详细的讨论。外部信度指的是研究由不同研究人员各自独立去做，得出的结果是否类似；也就是说，研究是否有可重复性 (replicability)。如果得出的结果类似，我们就说研究具有较高的外部信度。缺乏信度的研究是靠不住的研究，其

研究结果是不可信的。

效度也叫有效性，指的是研究结果是否反映了研究人员想要研究的东西。效度也可分为内部效度和外部效度。内部效度指的是研究的可解释性（interpretability）。在实验性研究中，常常要控制实验条件，使得出的结果反映你想要研究的东西的结果。例如，要比较交际教学法与传统教学法的教学效果，可以用分组对比的方法。但是影响教学效果的因素是很多的，如教师、教材、学习环境、学生的个人因素（如原来的水平、年龄、智力、学习动力、个性、性向等）、课外活动等。影响测验的结果也是很多的。分组的方法对实验结果也有影响。因此，研究人员必须从分组、教学、测试各方面保证两组的情况除教学方法之外一切都相同，得出的结果才能反映教学方法对教学效果的影响。又如，要测量学生的阅读能力，不能用考语法或词汇的试题来考，阅读试题的内容必须全面且有代表性，不能有偏差，也就是测试必须在内容上保证其效度。内部效度也有很多种，我们将在第十章作详细的介绍。外部效度指的是实验结果能否从少数的实验对象推论到大量的研究对象，也就是实验结果的可推广性（generalizability）。例如，要研究广州地区大学非英语专业学生在英语发音上的问题，在人力、物力、时间上都不可能把所有这些学生都拿来研究。必须从这些学生的全部（称为总体）中抽取一定数量的学生（称为样本）来进行研究。这些学生的代表性对研究结果的可推广性就起很大的作用。如果抽样是随机进行的，即每个学生有同等的机会被抽中，得出的结果就可推广到学生的总体中去。也就是说，抽出的学生有什么语音问题，整个广州非英语专业的学生就很可能有什么语音问题。可推广性在揭示规律性方面起很大的作用。如果一次研究得到的结果只限于被研究的实验对象本身，研究的意义就不是很大。但如果得到的结果可推广到较大的范围，研究成果就可以用来指导我们的工作，可以用它来进行比较、预测、控制等其他的研究。要保证研究有比较高的效度，通常采用随机抽样、随机分组、使用控制组、使用对等试题、使用初测和后测等方法。这些内容将分别在第三章第二节和其他有关章节进行讨论。

第三节 实验性研究的过程

实验性研究是用科学方法进行的研究，所以要遵循一定的研究程序。下面我们介绍实验性研究的主要步骤。

一、阅读文献、观察问题

进行实验性研究，必须首先选择研究的领域。一般说来，研究领域是你最有兴趣、比较了解而且比较实用的领域。例如，如果你是个英语教师，你可能会对第二语言/外语习得（second language/foreign language acquisition）这个领域产生兴趣，因为在你的工作中，你可能已经观察到不少有关外语习得的问题。作为英语教师，你可能已经学了不少语言学的理论和知识。在这个领域中的研究成果，对外语教学有帮助，可以说是比较实用的。但是，光凭兴趣和对某一领域的一知半解来进行研究还是不够的，这就要求对某一领域的理论和研究进行更加全面和深入的学习。这样的学习对自己的研究很有帮助，这是因为：①全面、深入地学习某个领域的著作，对该领域的理论和研究就会有比

较完全的了解，这样就可以避免考虑问题的片面性。②国外应用语言学研究所提出的理论、假设比较多，其中有些可能是正确的，有些是有争议、互相矛盾、甚至是不正确的。比较全面地了解人家的理论和假设，就可以将各家之说互相比较，根据自己的直觉、理性和经验对其真伪进行判断。有些听起来有理的理论和假设，可以用科学实验方法对其正确性进行检验，对一些你认为不正确的理论和假设，也可以用科学实验对其正确性进行检验，这样做就比较客观。③通过阅读别人的著作，你会了解国外做过了多少实验性研究，这样就能避免做别人早已做过的实验，浪费时间和精力。但这不是说，别人做过的实验就不能做。实验性研究的一个特点就是它的可重复性，但这只有在你了解别人实验的目的、方法和结果，而且明确了重复做同一实验的目的时去做才有意义。④阅读别人的著作，你会从中得到很多启发，发现很多可以研究的问题。每个实验研究能解决的问题都是很有限制的，有些研究已经得到一些结果，但可能部分问题还未得到解决。有时候，实验目的、方法、对象、条件可以改变，以适应新的情况，例如，适合外国情况的东西不一定适合中国的情况。这就给我们提供了许多可以研究的课题。因此，必须多读书，在阅读别人的著作时要做到有选择性和批判性。

在国外，有关某一领域所需的研究材料，可以从三方面去找：

1. 索引和摘要期刊

如英国的 *Language Teaching: The International Abstracting Journal for Language Teachers and Applied Linguists* 就提供了大量有关第二语言或外语教学的重要文章的摘要；美国密执安大学出版的 *Linguistics and Language Behavior Abstracts* 提供了 30 种语言、1000 多种出版物的文章的摘要。通过查索引和摘要，可以比较快地了解某一领域中已经进行过的研究。

2. 评述出版物和年鉴

例如，英国的 *Annual Review of Applied Linguistics* 集中了每年各种专题的评述文章，其内容包括交际教学法、第二语言习得、语段分析等等。每篇文章对某一领域的重要研究作综述，还附有参考书目。

3. 期刊

应用语言学的学术著作后面都附有参考文献，相当大一部分是在期刊上发表的论文。在国外，应用语言学的期刊很多。例如，英国的 *Applied Linguistics* 和 *Studies in Second Language Acquisition*，美国的 *TESOL Quarterly*，*English for Specific Purposes*，*Issues in Applied Linguistics* 等等。我国很多大学图书馆、资料室都订有不少国内外期刊，不难在这些期刊中找到一些有关的研究文章。

附录 B 列出国外一些主要语言学研究出版物的名称、出版者和内容简介，供读者参考。

在阅读别人的著作的时候，要联系实际，善于思考，才能找到研究课题。

二、提出问题和作出假设

找到研究领域并不等于找到研究课题。研究领域是很大的，必须缩小题目，才能用实验方法去解决。例如，通过阅读有关第二语言习得的文献，你可能会注意到它讨论过下列这些题目 (topics)：

- 1) 语言学习理论 (如行为说、天生说、相互作用说);
- 2) 语言习得模式 (如过渡语模式、文化适应模式、监控理论、语言普遍性假设、适应理论、语段理论、神经功能说等);
- 3) 语言相互作用对语言习得的影响 (如语段分析、foreigner talk、mother talk、teacher talk 等);
- 4) 语言输入对语言习得的影响 (如自然习得次序研究、输入与吸入、可理解的输入、语频的作用等);
- 5) 个人因素对语言习得的影响 (如年龄、智力、认知类型、个性、态度、动力对语言习得的影响);
- 6) 课堂教学对语言习得的作用 (如对习得次序、习得速度和习得成功率的作用)。

必须进一步缩小这些题目,直到你能够以现有的时间、人力、物力和你的能力进行解决为止。研究课题的选择有下面几个标准:

1. 有限、具体、明确

实验性研究大多不是综合性的研究,它研究的问题往往是一个个具体的问题。涉及的问题越多越复杂,实验条件就越难控制,设计就越复杂,所需的实验工具就越难掌握。所以,提出的问题应有限、具体、明确。例如,研究课堂教学对语言习得的影响,不宜提“课堂教学是否对语言习得有影响”这样笼统的问题,而应提什么样的课堂教学对哪方面的习得是否有影响的具体问题;同样,研究英语教学法,不宜提“交际教学法是否优于听说教学法”这样的问题,而应提在什么样的教学目标上、对什么样的学生、在什么样的教学环境下、在哪些技巧的习得上交际教学法是否优于听说教学法的问题。

2. 明确变量间的相互关系

实验性研究研究因果关系和相关关系。所以,必须把影响实验结果的种种因素(叫做变量)找出来,弄清它们与实验结果的关系以及它们之间的关系。例如,研究年龄对语音习得的影响,年龄就是一个变量(叫做自变量)。语音习得可通过可靠、有效的某种语音测试来进行,得到的分数就代表了习得的水平,叫做因变量。弄清了自变量与因变量和自变量之间的关系才有可能解决问题。

3. 问题应有可能、或假定有可能被解决

提出问题的目的是为了解决问题。所以,提出问题实际上已蕴涵如何解决问题。例如,研究年龄与语音习得的关系,就可以通过不同的设计用相关法或方差分析法进行解决。这就要求研究者熟悉实验设计和统计推理的方法。如果对实验研究方法不了解,应请教熟悉实验研究的人员或统计学家。

4. 问题应只有一个中心,不要和其他问题扯在一起

影响因变量的因素很多,所以,要把要对其进行研究的因素弄清楚。例如上面提到的课堂教学对语言习得的影响,你是要研究教师的教学风格对语言习得的影响,还是教学方法或教材对它的影响,还是某个学生的个人因素对它的影响,一定要弄清楚。几个问题扯在一起,它们对因变量的影响就会互相增强,或互相抵消,就难以找到规律。

研究课题常常可以归结为一个疑问句。例如,下面就是一些可以研究的问题:

- 1) 广州地区中学生对 14 种英语词素的习得次序是否与 R. Brown (1973) 发现的第二语言习得次序相同?

- 2) 在英语教学中使用真实可信的材料是否有助于英语教学?
- 3) 正式学习语法规则对英语的运用是否有帮助?
- 4) 背景知识对英语阅读理解是否有帮助?
- 5) 在英语学习中, 语音的正确性与学习效果有什么关系?
- 6) 学生的 I. Q. 与英语习得是否有关系?
- 7) 学生的年龄/兴趣/个性/认知类型与英语学习成绩有什么关系?
- 8) 中国英语教师与外籍英语教师在批改英语写作作业时是否有明显的不同?
- 9) 在大篇演讲中多用语段标记对听众的理解是否有帮助?
- 10) 进行短期的应试练习对 TOEFL 成绩的提高是否有帮助?
- 11) 中国学生对英语各种类型的否定句/疑问句/定语从句/状语从句是否有一定的习得次序?
- 12) 词频与词汇习得是否有关系?
- 13) 学生的记忆力与习得哪些技能(如听、说、写、读、语音、语法等)有关?
- 14) 进行英语 critical reading 的训练对提高阅读理解能力是否有帮助?
- 15) 用上下文记单词和孤立记单词效果是否不同?
- 16) 在同一时间内, 语言输入量与习得量是否有关系?
- 17) 教学方法 A 在培养学生英语交际能力上是否比教学方法 B 更加有效?
- 18) 在美国, 中国留学生入学的 TOEFL 成绩与其学业成绩是否有关系?
- 19) 英语课堂教学的次序是否能影响或改变自然习得次序?
- 20) Role-play 对提高学生口头交际能力是否有帮助?

仔细研究上述 20 个问题, 看它们是否符合研究课题的 4 个标准。

问题提出来以后, 应根据自己学过的理论、假设, 他人得到的实验结果, 自己的直觉和经验作出假设。假设具有 3 个特征: ①预测各个变量间的关系; ②明确地表达为一个陈述句; ③应具有可检验性。例如, 对上面提到的 20 个问题的前 5 个, 可作如下的假设:

问题 1: 广州地区中学生对 14 种英语词素的习得次序与 R. Brown (1973) 发现的第二语言习得次序相同。

问题 2: 在英语教学中使用真实可信的材料有助于英语教学。

问题 3: 正式学习语法规则对语言的运用有帮助。

问题 4: 背景知识对英语阅读理解有帮助。

问题 5: 在英语学习中, 语音的正确性对听说的学习有帮助, 对语法的学习没有帮助。

读者可对其他问题作出自己的假设(假设可以是肯定句, 也可以是否定句), 然后看自己的假设是否符合上述的 3 个特征。

在假设检验中, 通常先作出没有差别或不存在相关关系的假设, 这种假设称为零假设(null hypothesis), 然后找出支持或否定零假设的证据。有关零假设, 我们在讨论假设检验时再讲。