

基于语料库的 外国人汉语句式习得研究

张宝林 等著



中国书籍出版社
China Book Press

教

基金资助

项目批准号：08JA740006

基于语料库的 外国人汉语句式习得研究

张宝林 等著



中国书籍出版社
China Book Press

图书在版编目(CIP)数据

基于语料库的外国人汉语句式习得研究 / 张宝林等

著. —北京: 中国书籍出版社, 2014. 12

ISBN 978-7-5068-4581-6

I. ①基… II. ①张… III. ①汉语—语法—对外汉语
教学—教学研究 IV. ①H195.3

中国版本图书馆CIP数据核字(2014)第273706号

基于语料库的外国人汉语句式习得研究

张宝林 等著

策划编辑 安玉霞

责任编辑 安玉霞

责任印制 孙马飞 马 芝

版式设计 中尚图

出版发行 中国书籍出版社

地 址 北京市丰台区三路居路97号(邮编: 100073)

电 话 (010) 52257143(总编室) (010) 52257153(发行部)

电子邮箱 yywhbjb@126.com

经 销 全国新华书店

印 刷 北京潮河印刷有限公司

开 本 710毫米×1000毫米 1/16

字 数 490千字

印 张 27.75

版 次 2014年12月第1版 2016年7月第2次印刷

书 号 ISBN 978-7-5068-4581-6

定 价 72.00元

版权所有 翻印必究

前 言

以 20 世纪 80 年代中期鲁健骥先生的汉语中介语研究为标志^①，国外第二语言习得理论开始进入我国，“它在国内的发展速度是空前的，吸引了一大批研究者，也涌现出不少优秀成果，在某种程度上甚至成了一个热门领域”。（施家炜，2006）30 年来，对外汉语教学领域的习得研究涉及汉语字、词、句、篇等各个层面。其中句式习得研究一向受到学界重视，研究主要集中于某些句式的偏误分析和习得顺序方面，取得了多方面的研究成果，极大地提高了对外汉语教学研究的水平。另一方面，研究中也还存在一些问题，例如研究范围狭窄，主要集中于“把”字句、“被”字句、趋向补语句等；考察的样本及语料规模小，难以保证研究结论的客观性、普遍性和稳定性；研究不够具体，不够深入；研究方法多是内省思辨，缺乏实证研究；研究结果对改进教学缺乏实际的参考价值；等等。

针对研究中存在的种种问题，我们在建成“HSK 动态作文语料库”之后，即开始申报课题“基于语料库的外国人汉语句式习得研究”，并得到教育部人文社会科学研究项目基金资助（项目批准号：08JA740006）。我们试图在一个较大规模的汉语中介语语料库的支持下，对外国汉语学习者的句式习得情况进行深入研究，全面考察外国学习者的汉语句式习得情况，根据语言事实，归纳偏误类型，分析产生偏误的原因，提出改进教学的对策，以期提高汉语句式教学的效率与水平，进而提高学习者的汉语表达能力，同时探索一种新的习得研究范式。书中留学生的例句均来源于语料库，为呈现语料的真实面貌，对例句的错别字、病句、标点、繁简体均不进行处理。

① 吕必松（1990）指出：鲁健骥的《中介语理论与外国人学习汉语的语音偏误分析》和《外国人学习汉语的词语偏误分析》，首次引进了中介语理论，标志着我国中介语理论研究的开端，在对外汉语教学中开辟了一个新的研究领域。

二

提高句式教学的效率与水平,探索新的习得研究范式,这是我们的愿景,无疑具有十分重要的实践意义和理论价值。然而其难度也是巨大的。在研究前期的探索过程中,我们并未能跳出以往研究的窠臼,前人研究中存在的一些不足在我们的研究中同样存在。随着研究的持续深入,认识的不断提高,我们的研究工作逐渐形成了自己的特点。

第一,课题依据目前规模最大、标注内容最多、背景信息最全的汉语中介语语料库——“HSK 动态作文语料库”进行研究,该语料库 1.0 版包括外国考生参加高等汉语水平考试的作文答卷 10 740 篇,语料共约 400 万字;其 1.1 版含作文答卷 11 569 篇,语料 424 万字。本研究所考察的学习者样本数量之多、语料规模之大,在对外汉语教学的习得研究中是前所未有的。且语料来自标准化水平考试,能够最大限度地真实反映考生的汉语水平。如此广泛而坚实的研究基础使我们的研究得到一些新的发现,得出的研究结论具有较大的普遍意义。例如,学界普遍认为“把”字句是外国人学习汉语的最大难点,“把”字句教学的最大问题是学习者采取了回避策略。然而研究发现:学习者在 87.48% 的概率上是能够正确使用“把”字句的;学习者的“把”字句使用率高于《人民日报》的使用率;回避现象固然存在,但还有大量与之相反的泛化现象,二者的比例几乎完全相等。这些发现与学界以往的认识大相径庭,且数据翔实,具有很强的说服力。

第二,在偏误分析的基础上,进一步考察学习者正确的语言表现,并与母语者的同类语言表现进行对比研究,从而把偏误分析提升为表现分析。偏误分析无疑是很意义的,这也是其在对外汉语研究领域 30 年来长盛不衰的原因所在。但是,只看偏误而置学习者正确的语言表现于不顾,难免只见树木不见森林,无法对学习者的习得情况做出全面而准确的评价。而表现分析则弥补了此种缺陷,可以使我们对学习者的习得情况的认识更清楚,更全面,更准确,因而也就更可靠。

第三,对偏误原因的研究较为具体、深入,特别是部分章节通过测验对偏误成因的推测进行了实证研究,在一定程度上使主观性的研究结论得到了客观的证明。例如《“是……的”句(一)的偏误研究》一文,通过对北京语言大学三大对外汉语学院 5 种教材的详细考察,印证了偏误原因在于教学环节的偏差的推论。而“是……的”句(二)、“使”字句、“给”字句、“被”字句的习得考察则以测试的方式,并通过推论统计证实了研究者对偏误原因的一些推测。“比”字句的习得考察虽然实证研究的结果并不支持研究者的推测,不是证实,而是证

伪，但至少排除了一种并不恰当的主观认识，还是很有意义的。

第四，对教学对策进行验证。以往的研究在归纳偏误现象、分析偏误产生的原因之后，常会用比较少的篇幅，非常简略地，甚至是“象征式”地提出一些教学建议、对策或设想。这些建议、对策与设想是在对相关的偏误现象与成因进行研究的基础上提出的，因而有其意义与价值。但并未经过教学实践的检验，也没有实证研究的支持，因而缺乏客观性，其价值也就很难被证实。实事求是地说，本书在这方面并无根本性突破，但《趋向补语“起来”的习得考察》还是通过试验班与对照班的对比教学，对所提出的教学对策的有效性进行了实验，证实了教学对策的可行性与优越性，是一次十分有益的探索与尝试。

第五，在偏误原因的研究中试图有所创新。以往的研究在探讨偏误原因时，非常注重学习者母语的个性特征，常常归因于学习者母语的负迁移。这种解释当然是有其理论依据和参考价值的，但其考虑的角度显得比较单一，很难再得出进一步的结论。本书除从这一角度进行研究之外，还试图从语言共性的角度来观察学习者产生偏误的原因，探讨不同母语背景的学习者存在的相同的汉语句式偏误现象。例如《“把”字句偏误分析》，通过研究得出了“产生所谓‘回避’的根本原因是汉语和其他语言在句子类型上的不匹配，不对应；而直接原因则是二语者对‘把’字句的使用规则（主要是语义语用特征）没有充分掌握，即目的语知识不足”的结论。

第六，通过研究，我们形成了一种“大样本、重实证”的研究理念，即从大规模的语料样本中归纳语言事实，通过主观思辨分析偏误产生的原因，采用实证方法加以证明，特别注重定量分析，追求结论的客观性。这种研究理念与实践最终形成了一种较为新颖的习得研究范式，即：基于大规模真实语料的、定量分析与定性分析相结合的实证性研究。“HSK 动态作文语料库”中的 424 万字真实语料给我们奠定了定量分析的坚实基础，翔实的数据、测试与实验方法的运用为我们内省的偏误原因推测加以验证。研究范式的探索是本书最为重要的研究目标之一，对汉语习得研究具有十分重要的意义。

三

本课题组成员大多是研究生导师，我们认为研究生的主要培养方式是参加实际的科学研究，在科研实践中学习新知，经受训练，提高研究能力与水平。因此我们把课题研究与研究生培养相结合，带领研究生参加课题研究工作。从 2008 年 11 月课题立项到 2013 年 8 月，共有 30 名研究生参加本课题的研究工作，完

成硕士学位论文 30 篇。

在此段时间内,课题组师生参加国际国内学术会议 18 人次,在国内学术期刊和多种论文集中发表学术论文 16 篇。

在此基础上,以课题组的教师成员为主对上述研究成果进行整理,包括改编与改写,有些章节经历了多次反复修改,最后由张宝林统稿,形成本书。

成书阶段的分工如下:

第一章、第二章、第九章、第十六章、第十七章、第二十二章,由张宝林写成或整理。

第三章、第十一章,由林君峰、王小玲整理。

第四章由王静写成。

第五章、第十三章、第十五章、第十九章,由刘雪春指导王婷婷、吴慧芹、马驰佳、田萍、商玥整理。

第十二章、第十八章,由栾育青整理。

第十四章,由田旭红、金海月写成并整理。

第七章、第十章,由贾钰写成或整理。

第六章、第八章,由金海月整理。

第二十章、第二十一章,由施家炜整理。

各个章节由哪篇或哪几篇论文加工整理而成,在各章标题上加了题注,这里不一一说明。

关于本书各章的排列顺序,第一章为全书的研究思路与理念,第二十二章为各章的研究结论对汉语句式教学的启示,均为宏观性论述。其他各章主要按标题首字的音序排列,只有两处例外:一是为阅读方便,把三章关于补语习得的研究放在了一起;二是把研究否定句和正反问句的两章放在了其他句式之后,因为这两章是对句类习得情况的研究,虽然十分重要,但严格说来,并不属于句式的范畴。

四

我们期待本书的研究理念、研究范式与研究结论能为汉语习得研究带来积极的影响,能为汉语中介语研究、汉语作为第二语言的教学理论研究、课堂教学研究、教材编写与研究、汉语水平考试研究等提供重要参考,发挥其应有的作用。同时,我们也期待学界的专家学者不吝赐教,批评指正,以使我们的研究不断前进。

张宝林

2014 年 12 月 5 日

目 录

前 言	1
第一章 外国人汉语句式习得研究的方法论思考	1
一 引言	1
二 句式习得研究的症结	2
三 研究策略	4
第二章 “把”字句偏误分析	11
一 引言	11
二 “把”字句的偏误类型	13
三 关于回避与泛化问题的讨论	18
四 对汉语教学与测试的启示	27
第三章 “被”字句习得研究	29
一 引言	29
二 “被”字句偏误类型	29
三 “被”字句正确句与母语者使用情况对比	38
四 偏误成因探析及教学建议	41
五 余论	48
第四章 “被”字宾语隐现规律习得研究	49
一 引言	49
二 来自汉语本体研究的成果	50
三 “被”字宾语在语篇中的衔接力	51
四 小说中自由 B 的衔接力	53
五 HSK 中自由 B 的衔接力	55
六 自由 B 在小说与 HSK 中的语篇衔接力对比	56
第五章 “比”字句习得研究	60
一 引言	60
二 外国学习者“比”字句表现分析	61
三 “比”字句偏误原因分析	70

四	关于“比”字句的教学建议	75
五	结语	76
第六章	程度补语句习得研究	80
一	引言	80
二	程度补语句习得情况分析	80
三	偏误成因探讨	92
四	对程度补语句教学的启示	93
五	结语	95
第七章	情态补语习得研究	97
一	引言	97
二	情态补语句习得情况考察	100
三	对教材的相关考察	112
四	教学建议	116
第八章	趋向补语“起来”习得研究	118
一	引言	118
二	趋向补语“起来”习得情况分析	119
三	偏误成因探讨	127
四	对“起来”教学的启示	130
五	教学实验	132
六	结语	140
第九章	“得”标记重动句习得研究	143
一	引言	143
二	“得”标记重动句习得现状考察	144
三	偏误原因探析	156
四	教学建议	160
五	结语	162
第十章	动词重叠偏误研究	163
一	引言	163
二	偏误类型	166
三	偏误原因分析	171
四	研究与教学建议	175
五	结语	177
第十一章	“对于”句习得研究	178
一	引言	178

二 偏误类型及统计分析	179
三 偏误原因及解决策略	189
四 结语	195
第十二章 “给”字句偏误研究	196
一 引言	196
二 S3 句式的主要偏误类型及统计分析	197
三 偏误原因实证研究	214
四 S3 句式的教学建议	220
第十三章 “连”字句习得探析	223
一 引言	223
二 “连”字句本体研究	223
三 母语为日语者“连”字句习得特点与偏误分析	230
四 “连”字句的教学策略	236
五 结语	243
第十四章 “使”字句习得研究	245
一 引言	245
二 “使”字句偏误类型及分析	246
三 偏误原因及实证研究	252
四 教学及教材编写策略	262
第十五章 “是”字句习得情况考察	264
一 引言	264
二 “是”字句偏误类型及统计分析	265
三 “是”字句偏误原因探析	276
四 正确使用“是”字句情况及对比	279
五 解决策略	284
六 结语	287
第十六章 “是……的”句（一）偏误研究	288
一 引言	288
二 “是……的”句（一）的偏误类型	289
三 偏误产生的原因	296
四 教材对“是……的”句（一）的处理	298
五 教学对策	304
六 结语	306
第十七章 “是……的”句（二）习得情况考察	307
一 引言	307

二 “是……的”句(二)的偏误类型	308
三 偏误原因探析	314
四 测试:留学生与母语者的对比分析	316
五 相关建议	325
第十八章 双宾结构偏误研究	329
一 引言	329
二 双宾结构偏误类型及统计分析	329
三 双宾结构泛化偏误原因分析	337
四 教学建议	343
第十九章 “有”字句偏误研究	346
一 引言	346
二 “有”字句偏误类型分析	347
三 偏误原因分析	357
四 教学策略与建议	362
五 结语	364
第二十章 “不”和“没(有)”否定结构习得研究	366
一 引言	366
二 汉语否定结构的习得情况	371
三 汉语否定结构习得的影响因素	380
四 结语	383
第二十一章 正反问句习得研究	386
一 引言	386
二 汉语正反问句的习得情况	389
三 汉语正反问句习得的影响因素	400
四 教学对策与建议	403
五 结语	405
第二十二章 从汉语句式习得研究看句式教学面临的任务	406
一 句式习得研究的现状	406
二 句式习得研究的新成果	409
三 句式教学面临的任务	411
参考文献	414
后 记	431

第一章 外国人汉语句式习得研究的方法论思考^①

一 引言

在对外汉语教学中，把字句一直被认为是外国人学习汉语语法的一个难点，学界普遍认为，外国人对把字句采取了回避策略。这种看法由来已久，已经成为学界的一项共识。

然而，笔者请留学生造把字句时，学生张口就能答出“把窗户打开”“把瓶子给你”之类的句子；和一些同事探讨，也有人表示把字句并不难教，所谓“把字句难”是一种误导；而学界在“共识”之外，也存在不同看法：把字句的难度在很大程度上被夸大了（刘颂浩，2003）。

我们不禁要问：

- （1）把字句究竟是不是难点？如果是，难到什么程度？
- （2）学生是否回避了把字句？如果是，究竟回避了什么？
- （3）除了回避，把字句教学还存在什么问题？

张宝林（2010）根据“HSK 动态作文语料库 1.0 版”中 10740 篇作文、共约 400 万字语料所做的研究表明：

（1）把字句的偏误率为 12.52%，正确率则达到 87.48%，并不像人们想象的那么难。

（2）把字句的问题并不限于回避，而是存在 3 种类型的偏误：回避（该用而未用）、泛化（不该用而误用）、内部偏误（该用也用了但存在错误）。其中“泛化”与“回避”的偏误率相差不足一个百分点，均在 35% 上下，而以往对这种偏误认识不足，重视不够。

（3）外国考生把字句的使用率为 0.092%，仅次于“是……的”句、“是”字句和“有”字句，而高于其他句式。据统计，《人民日报》“把”字句的使用

① 本章原载《华文教学与研究》2011 年第 2 期，作者：张宝林。

率在 0.0754% ~ 0.0767% 之间。从使用率的角度来看,不论是外国人和外国人的比,还是外国人和中国人比,外国汉语学习者使用把字句的情况都并不少。“回避说”显然根据不足。

根据上面的讨论,可以得到两条启示:

(1) 在现代汉语中,把字句是最具特色的句式之一。与其他句式相比,在本体研究、对外汉语教学与第二语言习得等方面的研究中,把字句都得到了更多的关注与研究。但迄今为止,我们对外国人习得把字句的情况并不是很清楚,还存在许多问题。其他许多句式也存在类似情况。例如是字句的偏误率远远高于把字句(刘艳娇,2011);“是……的”句的偏误句数约为把字句的 4 倍,陆俭明先生(2000)曾高度关注这一句式的教学,认为该句式教学上的问题反映了本体研究的不足。但是,对汉语中这两种常用句式的习得情况的研究还很少,很不全面。因此,我们迫切需要对外国人汉语句式的习得情况进行全面的调查与深入的研究。

(2) 12 世纪法国杰出的思想家和哲学家阿伯拉尔指出,在学问上最好的解决问题的方法就是持续的和经常的怀疑:“由于怀疑,我们验证,由于验证,我们就获得真理。”(张芝联主编,1989:76) 当今社会普遍认同这样的观点:科学的发展与进步离不开怀疑,从某种意义上说,提出问题比解决问题更重要。从学术发展的过程来看,前人的观点无疑是非常重要的,前人的经验也是必须继承的。但如果看不到今天理论观念、研究方法、研究材料、教学对象等方面的变化,仍不加思索地将“把字句难”的看法奉为圭臬,其实是囿于成说,墨守陈规。因此需要开动脑筋,独立思考,在新形势下得出新的结论。

基于上述认识,本文拟对外国人汉语句式习得研究的方法论问题进行一些探讨,以期对汉语句式的习得与教学研究提供一些有价值的参考。

二 句式习得研究的症结

20 世纪 80 年代初期,国外第二语言习得理论开始进入我国,对外汉语教学界对外国人汉语习得情况的研究随之展开并不断深入。20 多年来,对外汉语教学领域的习得研究涉及汉语字、词、句、篇等各个方面。其中句式习得研究一向受到学界重视,研究主要集中于某些句式的偏误分析和习得顺序方面,例如把字句、被字句、某些补语句等,已经发表多种论著,取得了多方面的研究成果,对教学起到了积极的推动作用。(施家炜,2006)

另一方面,句式习得研究中也还存在一些问题。例如研究不够全面,不够平

衡。除对把字句、被字句、趋向补语句等句式研究较多之外,对是字句、有字句、存现句、连字句、重动句等句式则很少研究。研究不够具体,不够深入。例如很多研究认为学生对把字句采取了回避策略,但究竟是哪个阶段的学习者采取了这种策略,回避了哪些类型的把字句,怎么回避的,则缺乏探讨。对偏误原因的探讨一般都归结为母语的负迁移和目的语规则泛化,研究角度比较狭窄,而且分析笼统,缺乏针对性和个性化。教学对策只是用很少的篇幅非常笼统地提出一些意见、建议或设想,均为“感想式”或“应景式”的泛泛之论,缺乏实际的参考价值。

研究存在的主要问题有两个:

2.1 考察的范围小,结论的普遍性不强,对习得情况认识不清

以往的研究主要是基于个人的教学经验和个人收集的语料。然而,个人的经验有多大的普遍意义值得怀疑。一位一线教师终其一生,所能接触到的观察样本可能很有限,也许不会超过几百人;即使有意识地积累研究素材,大概也不过十几万字。而实际上,收集语料是一件存在多方面困难的艰巨工作,对于研究者个人来说,收集十几万字语料也是很难的。而且,在传统手工方式下,如何有效而充分地使用这些语料也是一个很大的问题。这就使得研究者难以观察到较多样本和较大规模的语料。那么在此基础上得出的研究结论的客观性、普遍性与稳定性也就很成问题了。

一个重要的客观原因是,过去没有大规模的汉语中介语语料库,不具备进行大规模样本考察的条件,导致研究者无法进行具有普遍意义的量化分析。

因此,考察的样本少,观察的语料不多,就成为以往研究中的一个普遍现象,有的研究甚至只是基于零星的样本。例如:在讨论把字句的回避现象时,有的研究只是根据个别学生的表现得出结论(参李宁、王小珊,2001)。绝大部分研究只有十几个考察对象,均属小规模样本的考察(参刘颂浩,2003;程乐乐,2006;黄月圆、杨素英,2004等)。

调查范围小,样本少,必然导致两个后果:一是观察到的现象不全面,所得结论的普遍性、稳定性不强;二是难以准确地判断与把握外国人学习汉语的实际情况。例如有学者认为形容词谓语句是教学难点(吕必松,1992:110;刘珣,2000:366),也有学者认为不是(赵金铭,2006)。研究者都是对外汉语教学领域的权威学者,所谈的教学对象都是英语背景的汉语学习者,依据的都是教学经验,结论却完全相反。情况究竟如何?似乎很难判断。前述把字句的研究情况也是非常典型的例证。

2.2 囿于主观判断，缺乏实证研究

国内的汉语习得研究在研究理念与研究方法上，普遍缺乏客观性、普遍性与稳定性观念，大多数研究者未能掌握统计分析和实证研究的方法。以往的研究大多停留在偏误分析阶段，而对偏误原因、教学对策的研究常常是基于研究者个人的内省思辨与主观判断，而不是在统计分析、实验基础上得出结论。江新（1999）对国内汉语学习和教学研究的论文进行的统计表明：理论介绍和经验总结文章所占比例高达 80.8%，实验研究仅占 3.6%；87.6% 的研究没有数据，1.3% 的研究虽有数据但无统计分析，有描述统计的研究占 7.8%，有推论统计分析的研究占 3.3%。王建勤（2006）指出：“汉语习得研究在研究方法上已经有所改观，学者们越来越认识到研究方法的改进对提高汉语习得研究质量的重要性。但目前的研究方法仍然问题较多：一是有的研究有数据、无统计，有的简单罗列百分比，缺少推理统计；二是被试选择无控制，要么被试太少，要么被试母语背景混杂；三是材料来源不清，统计方法选择不当。这在很大程度上影响了研究结论的可靠性和科学性，影响了研究质量的提高。”这些分析揭示了研究方法上的症结所在。

我们认为，基于主观思辨的研究固然有其意义与价值，但比较空泛，不是那么具体，也不是很深入，特别是缺乏客观性，因而其说服力不强。

三 研究策略

3.1 利用语料库，摸清外国人汉语句式习得的基本情况

习得研究对样本与语料的充分性和代表性非常重视，因为“语料是描写中介语的基础，如果语料不完整、不系统或不准确，那么在此基础上描写的中介语就不可靠”（孙德坤，1993）。在此基础上得到的研究结论也就无法反映学生真实的习得情况，进而会导致教学的针对性不强，难以切实提高教学水平与教学效率。

以外国人所写的汉语文本为语料来源建立起来的大规模汉语中介语语料库，为进行汉语句式习得研究提供了条件。例如目前收集样本最多、语料规模最大、标注内容最全、语料作者背景信息最完备的汉语中介语语料库——“HSK 动态作文语料库”（1.1 版），包括 11 569 篇外国人参加高等汉语水平考试主观性考

试的作文答卷,共计424万字。这些语料是学习汉语的外国人在标准化考试中不借助任何参考资料和工具书的情况下,运用汉语进行规定题目写作的即时表现,最真实地反映了他们运用汉语进行书面表达的实际水平。“HSK”动态作文语料库为针对外国人的汉语习得研究奠定了非常坚实的基础。可以说,该语料库所能考察的学习者样本数量之多、语料规模之大,在对外汉语教学的习得研究中是前所未有的。如此广泛的研究基础必将带来一些新的发现,使我们看到一些前人未曾看到的东西。我们可以考察外国人习得把字句、被字句、是字句、“是……的”句、形容词谓语句等各种句式的基本情况,了解这些句式的使用率、正确率和偏误率,研究各种句式具体的偏误类型。毫无疑问,这样得出的研究结论将具有更大的普遍意义。

以双宾语句为例。基于语料库的研究表明:双宾结构偏误中泛化偏误(指不该用双宾结构而用了的中介语现象)所占比例为67.99%,而其他类型的偏误很少;双宾动词中使用频率和偏误率最高的三个词是:“告诉”(42.98%/12.12%)、“教”(38%/13.24%)、“给”(29.96%/25.45%),可见这三个双宾动词是留学生双宾结构习得的重点和难点;留学生对于“给”的使用,用作介词的情况要远远多于用作动词,“给+NP1+V+NP2”“V+NP2+给+NP1”结构的使用要多于双宾结构的使用。(周岚钊,2009)对日本学生来说,连字句的偏误率高达49%,其中大部分偏误发生在语义层面,比例为52%。(张茜,2010)如果不利用汉语中介语语料库,我们是不可能得到这样的具体认识的。

虽然在国外“第二语言习得研究已从描述型转移到了解释型”(袁博平,1995),但就汉语的二语习得研究而言,描写的任务尚未完成,我们迫切需要对外国人汉语句式的习得情况进行全面的调查与深入的研究。

3.2 定量分析与定性分析相结合,尤其注重定量分析

传统的语法研究主要是一种定性分析,目前绝大多数的习得研究在探讨某种偏误现象的原因时也仍然主要采用定性分析的研究方法。这种研究无疑是有价值的,但我们很难断定其结论的正确性,特别是无法断定其普遍性有多大。因为它源于研究者的逻辑思辨与主观判断,难以得到客观而直接的证明。而“机控语库是开展大范围语言研究的极好语料源,因为它所提供的语料较之先前的材料更具有真实性,其层级结构更加明晰,因而更有助于对语言的不同层面进行描写研究,更有助于对不同语体的比较研究和开展量化与概率统计研究”。(丁信善,1998)在大规模真实语料基础上所进行的定量统计分析具有最大限度的客观性与普遍性,可以弥补定性分析的不足。

我们理解,定量分析的第一要义是“一切基于数字说话,不能拍脑袋”。(辛苑薇、侯继勇,2009)也就是说研究中得出的任何结论都要“言必有据”“无一字无来历”,这样才能保证结论的客观性。本文引言中关于“把”字句的相关数据就是证明。

定量分析的第二要义是数据的广泛性与代表性。即数据必须来自较多观察样本和较大规模语料,而不能仅仅来自少数的、甚至个别的样本或少量的观察材料。根据较少样本、较小规模语料得出的结论,如果扩大考察范围,增加样本数量与语料规模,数据可能就会发生变化,进而导致研究结论的改变。

例如同样是对把字句的量化研究,因考察的样本及语料规模不同,数据相差极大。罗青松(1999)根据日、韩留学生的考试作文所做的考察发现,把字句的使用率为0.022%。刘宏帆(2007)根据中山大学中介语语料库进行的研究结论是:把字句的偏误率为24%;其中“泛化”偏误句占有偏误句的54.29%。黄月圆、杨素英(2004)采用专题测试的方法考察发现,在不同类型的测试题中把字句泛化的错误率分别为77%、31%、52%。而笔者根据“HSK动态作文语料库”(1.0版)得到的结论是:把字句的使用率为0.092%,偏误率为12.52%，“泛化”句占偏误句总数的34.92%。

考察的句式相同,甚至使用的研究方法也相同,而结论竟有如此大的差异,其原因在于考察的样本与语料规模不同:黄月圆、杨素英的研究只有14个样本;罗青松的数据虽然来自45个样本,但只有2万多字语料;中山大学中介语语料库不超过75万字(张炯,2008);而“HSK动态作文语料库”(1.0版)则有10740个样本,约400万字语料。由于“比较大的样本可以减低样本统计量的变异”(戴维·S.穆尔,2003:162),依据较多样本和较大规模语料得到的数据无疑具有更大的普遍性,结论也就更可靠。

3.3 把偏误分析提升为表现分析,重视表现分析

在国内对外汉语习得研究中,从20世纪80年代开始,直至90年代,偏误分析都是研究最多的内容之一,并取得了很多研究成果,对教学起到了很好的促进作用。今天,偏误分析也还是一个受到较多关注的研究领域,仍然是一种常用方法。然而偏误分析是带有片面性的,因为它“只研究中介语的偏误部分,而且是横切面式的静态分析,并未研究中介语的正确部分”。(刘珣,2000:202)“也就是说,只看到了学生不能做的方面,而没有看到学生能够做的方面。”(袁博平,1995)这就难免只见树木、不见森林,看不到学习者全部的语言表现,特别是看不到正确的语言表现,进而夸大偏误的严重程度,