

# 外国学习者 汉语词汇 深度习得研究

---

毛嘉宾 等 著



济南大学出版社

# 外国学习者汉语词汇 深度习得研究

毛嘉宾 等 著



西南交通大学出版社

· 成 都 ·

图书在版编目 (CIP) 数据

外国学习者汉语词汇深度习得研究 / 毛嘉宾等著.  
—成都: 西南交通大学出版社, 2017.12  
ISBN 978-7-5643-5924-9

I. ①外… II. ①毛… III. ①汉语—词汇—对外汉语  
教学—自学参考资料 IV. ①H195.4

中国版本图书馆 CIP 数据核字 (2017) 第 292609 号

外国学习者汉语词汇  
深度习得研究

毛嘉宾 等 著

责任编辑 祁素玲  
封面设计 原谋书装

印张 10.25 字数 164千

成品尺寸 165 mm × 230 mm

版次 2017年12月第1版

印次 2017年12月第1次

印刷 四川煤田地质制图印刷厂

书号 ISBN 978-7-5643-5924-9

出版发行 西南交通大学出版社

网址 <http://www.xnjdcbs.com>

地址 四川省成都市二环路北一段111号  
西南交通大学创新大厦21楼

邮政编码 610031

发行部电话 028-87600564 028-87600533

定价 59.00元

图书如有印装质量问题 本社负责退换  
版权所有 盗版必究 举报电话: 028-87600562

## 前言

二语词汇学习与教学越来越受到学界的重视。近十几年来,国内外二语词汇研究的新走向是二语词汇的理论研究和实证研究,并且词汇研究的关注点也从词汇的广度向词汇的深度转移。另外,语料库语言学研究盛行,基于大规模真实文本的语料库在语言研究中发挥了无可替代的作用。本书就是在这种语言背景下,基于中介语语料库,专门针对国别化学生汉语词汇深度习得而做的研究。

关于词汇的习得,可以从发展的角度出发,把认识和运用词汇的数量分出不同的层级;也可以从词汇的多个维度出发,对每个词的各种属性进行分析和定性。正如 Qian (1999)所指出的:词汇知识应该包括两个方面,即词汇的广度和深度。掌握词汇的广度重在词汇的“量”,而掌握词汇的深度重在词汇的“质”。研究词汇,提高词汇教学和习得的水平,就必须兼顾词汇的“质”和“量”。

习得一个词,除了掌握它基本的音、义,还应掌握许多其他方面的知识,包括词的各种不同的意义,它同其他词会发生哪些联系(如聚合关系和组合关系),在言语使用中还会发生哪些意义上的变异等。一个词有很多属性,要完全“懂得一个词”,就必然涉及各方面的知识。

本书在研究西方词汇习得理论的基础上,针对外国学生汉语词汇习得中出现的问题,结合汉语中介语语料库中的偏误实例,力图探寻汉语词汇深度习得的知识框架,探究国别化汉语学习者词汇深度习得框架结构的差异性,研究成果将用于指导国别化学生汉语词汇深度习得,有利于监控其汉语词汇习得过程,修正其汉语词汇习得方式,指导教师减少教学误区,提高教学实效。

本书以外国汉语学习者为研究对象,基于汉语词汇等级大纲和汉语中介语语料库选取词汇研究的范畴。选取的词汇为外国汉语学习者较为常用的,且是中介语语料库中偏误率较高的词语。在所确定的汉语词汇

知识框架下，在中介语语料库中检索汉语使用情况主要是检索偏误情况，在此数据基础上全方位多角度地展现外国学生汉语词汇深度习得出现的偏误面貌。对这些偏误，以语料库语言学、对比语言学、认知语言学为支撑，深度探讨外国学生词汇习得的偏误类型以及习得中偏误产生的原因。

基于上述词汇习得类型和偏误类型的分析及产生的原因，我们探索汉语教学策略。就如何高效习得汉语词汇、深度习得汉语词汇及预治或消除偏误进行实证调查，系统地梳拢出外国学生汉语词汇深度习得状态与趋向，在实证研究和理论分析的基础上提出规律性教学策略。

因为偏误词汇数量庞大，故不能进行完全的研究，本书主要采取例证方式，研讨外国学生偏误率较高的易混词、形容词、副词、介词、汉字词、与汉语相关的文化类词以及流行语等词语的深度习得情况。

本书为山东省社科规划项目“中介语语料库视角下的国别化汉语词汇深度习得研究”（项目编号：11CWZX02）的成果。在写作的过程中，项目组成员通力合作，胡晓清撰写了第二章，苏向丽撰写了第一、第三章，焉德才撰写了第四章，毛嘉宾撰写了第五至第九章并进行了统稿。周晓波、许小星等老师也提供了大力支持与帮助。本书还得到了不少外国朋友的协助与建议。在此一并致以衷心的感谢！

作 者

2017年4月于烟台

# 目录

|                                             |     |
|---------------------------------------------|-----|
| 第一章 汉语词汇知识基本框架及词汇深度习得方式的确立 .....            | 001 |
| 第一节 汉语词汇知识框架的建立 .....                       | 001 |
| 第二节 词汇深度习得过程与方式 .....                       | 002 |
| 第二章 汉语中介语语料库的建立及应用情况 .....                  | 006 |
| 第一节 语料库概况 .....                             | 007 |
| 第二节 语料库建设的必要性 .....                         | 009 |
| 第三节 语料库的特点 .....                            | 011 |
| 第四节 语料库建设中遇到的问题与解决方案 .....                  | 013 |
| 第五节 语料库的发展 .....                            | 015 |
| 第三章 外国学生汉语易混词深度习得分析<br>——以“大”和时间量度词为例 ..... | 023 |
| 第一节 外国汉语学习者“大”的习得分析 .....                   | 024 |
| 第二节 外国汉语学习者时间量度形容词的习得分析 .....               | 034 |
| 第四章 外国学生汉语副词深度习得分析<br>——以“很”和“就”为例 .....    | 041 |
| 第一节 外国汉语学习者程度副词“很”的习得分析 .....               | 041 |
| 第二节 外国汉语学习者“就”的习得分析 .....                   | 056 |
| 第五章 外国学生汉语介词深度习得分析<br>——以“比”为例 .....        | 062 |
| 第一节 外国留学生“比”字句习得状况分析 .....                  | 063 |
| 第二节 外国留学生“比”字句偏误类型及偏误频率分析 .....             | 064 |

|                                |     |
|--------------------------------|-----|
| 第三节 外国留学生“比”字句偏误成因分析 .....     | 069 |
| 第六章 外国学生汉外同形词深度习得分析            |     |
| ——以初级阶段形容词为例 .....             | 074 |
| 第一节 汉外同形词的对比分析 .....           | 074 |
| 第二节 汉韩同形词使用偏误分析 .....          | 090 |
| 第七章 外国学生汉语固定短语深度习得分析           |     |
| ——以奇位嵌数成语为例 .....              | 094 |
| 第一节 汉语奇位嵌数成语的结构特征 .....        | 095 |
| 第二节 汉外奇位嵌数成语对比 .....           | 104 |
| 第三节 汉外奇位嵌数成语异同的原因分析及偏误分析 ..... | 108 |
| 第八章 外国学生汉语外来词深度习得分析 .....      | 110 |
| 第一节 汉外借用外来语的历史 .....           | 111 |
| 第二节 汉外外来语的基本形式 .....           | 113 |
| 第三节 汉外外来语对比分析 .....            | 116 |
| 第四节 汉外外来语折射的不同文化心理 .....       | 123 |
| 第九章 外国学生汉语流行语汇深度习得分析           |     |
| ——以近十年来流行语为例 .....             | 125 |
| 第一节 汉外流行语的基本分类对比 .....         | 125 |
| 第二节 汉外流行语来源对比 .....            | 130 |
| 第三节 汉外流行语结构形式对比 .....          | 137 |
| 第四节 外国学习者汉语流行语学习建议 .....       | 147 |
| 参考文献 .....                     | 149 |
| 后 记 .....                      | 158 |



## 第一章

# 汉语词汇知识基本框架及词汇深度习得方式的确立

发端于国外语言学界的词汇深度习得理论不断发展并取得一些研究进展。词汇研究从宽度向深度的转变,显示了人们不仅关注词汇的“量”,而且愈发关注词汇的“质”。关于词汇深度的界定,国外也有不同的观点,由 Cornbach (1942) 开始, Richards (1976)、Nation (1990)、Schmitt (1990)、Laufer (1997) 等为主要代表人物。他们认为如果“要懂得一个词”,那么词的一系列特征,比如形式、结构、搭配、句法、情感义、语用义、能产性运用等诸多方面也都要兼及,因此在词汇学习时应依照词汇深度知识框架来进行。国内很多学者也将此理论应用到英语教学中。近年来,国际汉语教学学界也越来越关注词汇的深度习得问题。

## 第一节 汉语词汇知识框架的建立

根据语言的共性和汉语的特性,我们认为,要了解汉语词的价值表现,可以从以下这些属性特征出发进行考察:

(1) 形式属性(体现形式价值): 语音形式(拼音); 发音; 书写形式(汉字)。

(2) 认知属性(体现认知价值): 是否基本范畴词汇; 是否反映了最重要的基本概念。

(3) 结构属性(体现结构价值): 是单纯或复合结构; 能否单用和自由运用; 构词能力大小。

(4) 语法属性(体现语法价值): 所属词类; 有无兼类和词类活用。

(5) 语用属性(体现语用价值): 常用度(词频); 释义能力。

(6) 语义属性(体现语义价值): 义项多寡; 词汇意义和语法意义。

微观语义属性: ① 基义, 包括普通义位、学科义位; ② 陪义, 包括各种陪义, 如语体、语域、褒贬、时代、方言、外来陪义等; ③ 义域, 包括大小域、多少域、伙伴域、适用域。

宏观语义属性(词汇联想模式): ① 聚合关系, 包括同义、上下义、反义、类义、多义; ② 组合关系, 包括语素的组合能力、义位的搭配能力。

这样的框架可以为汉语词汇习得研究提供一个量化考察的多元模板。通过对每个词的种种属性和价值的考察, 可以获得词的深度理解。以“老”为例:

“老”不论在古今、中外还是方言、普通话中都是一个基本词, 具有较高的认知价值, 熟知度高, 有很高的使用频率。这个古老的独体字, 从历史上说, 一向都能单用, 自由度高, 并具有较强的构词能力; 从语法属性上看, 具有较强的灵活性, 是兼类词, 可以做形容词(显得很老)、名词(男女老少)、副词(老早就来了), 还可以做前缀(老张); 从语义上看, 其基本义位和常用义位呈中性语体, 其派生义位中有的具有口语陪义(老好人)或方言陪义(吴语“老好”: 很好); 从词汇联想模式上看, 是聚合关系中的主导词, 具有多个同义词(上年纪、旧、陈旧、原来、非常), 反义词多(少、幼、新、嫩), 义项也多, 《现代汉语词典》(第5版)中就列了16个义项, 是形容词中义项最丰富的词。此外, “老”的组合能力很强, 既能构词(老人、老板、老师), 也能构语(老马识途、老谋深算、老少皆宜)。“老”在汉语中具有重要的位置, 它的词汇知识也较为复杂, 要深度习得和掌握“老”这个词是不容易的, 却又是十分重要的。

## 第二节 词汇深度习得过程与方式

事实上, 即便是自己的母语, 人们也不可能掌握每个词的全部知识和用法, 对词汇知识的全面描述是一种理想要求, 而不是对语言使用者的必备要求。词汇知识的研究描述了学习者应掌握的知识类型, 却缺乏

对知识习得过程的解释, 尽量详细地描述词的属性特征作为接受性知识(receptive knowledge)是可能的, 但是把它作为产出性知识(productive knowledge)要求习得者掌握则是不现实的。词汇知识框架的研究可以为词汇深度习得提供充分的知识储备, 为深度知识的选择提供依据, 但词汇深度习得并不只是掌握知识的过程, 更重要的是言语过程中的实践积累。语言教学不仅要给定一个有效的词汇知识框架, 便于学生掌握词汇知识, 还要针对学生的需求和条件设计一套合理的逐渐走向深度习得的教学方法。

在词汇教学的过程中, 我们既不能忽视学习者的主体因素, 又要寻求恰当的方法以实现词汇知识的深度习得。词汇的习得, 是一个认知的过程。词汇知识的教学必须充分地关注习得的心理过程, 明确学习者在习得过程中的主体地位, 努力激发其主动积累心理词汇的自觉性。学习者的词汇深度习得过程是从最初遇到这一个词到与该词多次相遇和反复使用, 到最后对这个词的多方面属性特征的多维掌握。只有调动学习者的主体性, 才能获得词汇的深度习得。因此, 对词汇知识的研究, 不应该仅从客观上对词汇所涵盖的各类知识进行描述, 还应该从二语词汇习得的认知主体出发, 考查学习者的认知过程, 使词汇教学切合他们的知识结构和认知能力。

词汇深度习得如何进行, 以及如何提高这种习得效率是本研究需要解决的问题。词汇知识的深度习得不同于词汇广度(词汇量)的习得, 词汇量可随着学习进程的推进自然增加。词汇深度习得则不同, 一般的学习者即使到了中高级阶段, 虽然能够辨认、理解大量的词汇, 却只能自如运用其中一小部分, 他们对多数词汇的了解往往还停留在拼写、发音、基本语义、基本句法等较浅的层次, 而对深层次的释义能力、词的语体和语域分布、词义的引申能力、聚合、组合能力等的了解并不深入。词汇的深度习得必须在汉语词汇知识框架的基础上, 通过深化训练才能获得。同时, 在词汇教学中要调动学习者的积极性, 引导他们主动参与到交际实践之中, 使之成为信息的加工者、创造者, 而不仅仅是被动接受者, 从而把知识转化为能力。

词汇深度习得是层层递进的。根据原型理论, 词汇的习得是从原型向边缘扩散, 按照词汇知识之间联系的远近隐显, 由易而难习得。而从词价标准看, 一个词的价值越高, 则与其他词建立联系的可能性越多, 联系的多少, 决定了词汇习得的难度和深度。

在各种词汇系统中,词价最高的词是本系统中具有原型意义的词,如基本词汇中的核心词,就是习得者最容易掌握的词。以基本词的基本义为核心意义,通过认知联想机制,可以与相关的词汇知识相连接,进而通过激活已知知识来拓展目标词知识。这种联想可以根据语形、语义、语法等各方面获得,使新旧知识相联系,并通过采用各种深化训练方法掌握目标词,提高认知能力和使用能力。词汇深度习得的层层递进的过程可以描述为如下步骤:

- (1) 确立与目标词相关的核心词、原型词;
- (2) 根据联想认知机制建立原型词与目标词之间的联系;
- (3) 围绕原型构建各种语义、语用网络,形成相关词汇知识集;
- (3) 根据已知原型词推测未知的目标词;
- (4) 目标词汇语境化,经过识别语境,拓展目标词知识;
- (5) 学习者在多语境中反复运用目标词,形成词汇应用能力;
- (6) 发挥学习者的主体性,积极参与新旧知识的连接,总结梳理。

通过以上系列操作和训练,帮助学习者重新整理或改变词汇的储存方式,逐步建立二语词汇语义网络,为最终综合二语词汇知识奠定基础。这样的词汇深度习得进程可以用图 1-1 来描述。

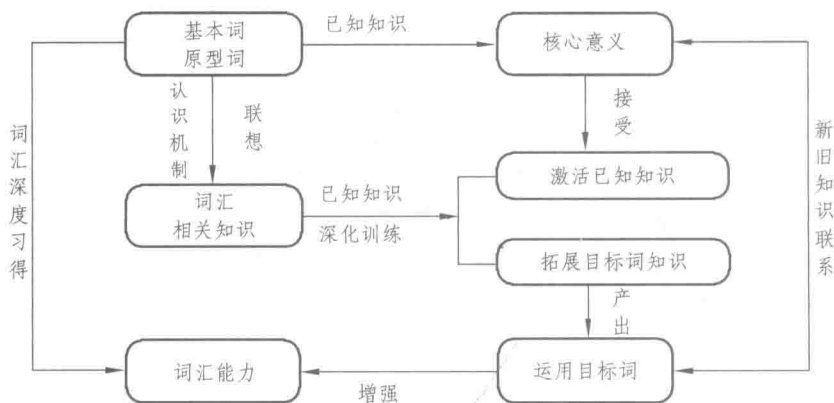


图 1-1 汉语词汇深度习得进程

词汇深度习得的过程是学习者深化已有的词汇知识,用已知词促进未知词的习得过程。熟知的词一般处于心理词汇的中心位置,由此向外扩散,分别涵盖不同习得程度的词。一个词被习得的程度是由该词与心

理词汇间建立的联系强弱决定的，联系越强，习得程度越深。在初始阶段，目标词与已知词之间的联系较明显；到了中高级阶段，联系往往转入隐性，特别是深层语义之间的联系。联系的隐性强弱程度不仅体现了习得是一个层层递进、循序渐进的过程，也明确了不同阶段的习得任务，这对词汇习得和教学有一定的指导意义。在深度习得过程中，我们必须在已知词和未知词之间建立联系，充分发挥学习者的主体性，根据已知原型词激活未知目标词，将新旧知识结合，让学习者在多语境中反复运用，使词汇语境化，最终形成实际的词汇能力。

对于外国学习者而言，不同国别由于母语的影响，会产生词汇习得中不同重难点。对于韩国汉语学习者来说，其文化背景与中国文化息息相关，较之其他国别汉语学习者会出现更多的同中存异、同异交错的现象，更容易因学习者负迁移而产生词汇的混淆、错用，因此我们更应该关注这方面负载信息量较多的汉字词、同形词以及文化类词汇。



## 第二章

### 汉语中介语语料库的建立及应用情况

本章我们把视角投到语料库，以探讨外国学习者词汇深度习得情况。利用外国学习者的中介语语料库，我们可以比较详尽地获得外国学习者的词汇习得状况，了解其词汇习得中最容易发生偏误的所在，据此，可以分析产生偏误的原因，并进一步分析对策，以提高其词汇习得能力和效率。

外国学习者的汉语中介语语料库的建设始于上个世纪 90 年代，第一个语料库即北京语言学院于 1995 年建成的汉语中介语语料库（储诚志、陈小荷，1993）；之后，出现了南京师范大学外国学生汉语中介语偏误信息语料库、中山大学汉字偏误标注连续性中介语语料库、暨南大学留学生汉语中介语语料库、南京大学留学生汉语中介语纵向语料库、香港中文大学外国留学生汉语口语语料库、苏州大学留学生口语语料库、台湾师范大学汉语学习者汉语偏误数据资料库，等等，中介语语料库建设如雨后春笋，一派蓬勃发展之势。北京语言大学 HSK 动态作文语料库更是一枝独秀，从 HSK 留学生作文试卷中选取 11 569 篇作文，约 400 万字，从字、词、句、篇、标点等多角度进行全面标注，是汉语中介语语料库中规模最大的语料库，且在线上全面开放使用（张宝林，2003）。而正在建设中的全球汉语学习者语料库更是规模空前，语料来源包括在中国的外国汉语学习者和在学习者所属国或其他国家学习汉语的学习者的语料，覆盖书面语和口语语料，规模空前，用途广泛（崔希亮、张宝林，2011）。

随着探索的逐步深入，学者们开始对该领域的研究现状展开反思、提出质疑，如对汉语中介语语料库的建设和加工提出建设性改进意见（任

海波, 2010); 张宝林 (2010); 崔希亮 (2011) 从更宏观的角度出发, 站在已有的研究基础之上, 提出了建设全球汉语学习者语料库的设想。

有学者也再次提出了建设汉语学习者口语中介语语料库的设想 (杨翼等, 2006), 我们期待在此基础上, 本书能为学界全面、深入、系统地开展汉语中介语研究提供语料支持。

本书所选用的语料来源主要为鲁东大学国别化汉语中介语动态语料库, 该库目前是国内规模最大的单国别中介语语料库, 已收录 400 万字生语料, 标注完成 300 万字熟语料, 下面对该语料库有关情况做一介绍。

## 第一节 语料库概况

鲁东大学国别化汉语中介语动态语料库始建于 2008 年。开始名称为“韩国留学生中介语语料库”。其语料主要取自鲁东大学国际交流学院 2006—2010 年间在校就读的韩国留学生的作业和试卷, 形式上分为造句语料和作文语料两种, 层级兼顾初、中、高级不同程度。

该语料库收集录入了 102 万字原始语料, 从标点符号、字、词、句、篇章五个层面对语料进行了偏误标注。语料库分为语料处理系统、数据库和用户检索系统三部分。语料处理系统包括录入的原始语料、已登录的语料信息、已加工的熟语料三部分。录入的原始语料存有扫描版和手录版两个版本。手录时对错字、不规范字和无法识别的字直接标记, 若检索可通过标记链接扫描语料进行对照、查询。语料信息登录了写作者来源、姓名、性别、出生年月、国籍、写作学年、学期、日期、所在班级、写作类型等 10 项信息。在原始语料中上述信息也一并提供。已加工的熟语料主要进行了机器自动分词、人工干预处理, 然后从标点、字、词、句、篇章五个层面进行偏误标注。数据库中存有标点信息库、字信息库、词信息库、句信息库、篇章信息库 5 个子信息库。用户检索系统为自开发的语料库检索软件。该软件向用户开放, 用户可根据自己的需要自主选择检索并输出语料库中的字、词、句、篇章、标点使用信息及偏误信息。该语料库具体偏误标注类型见图 2-1。

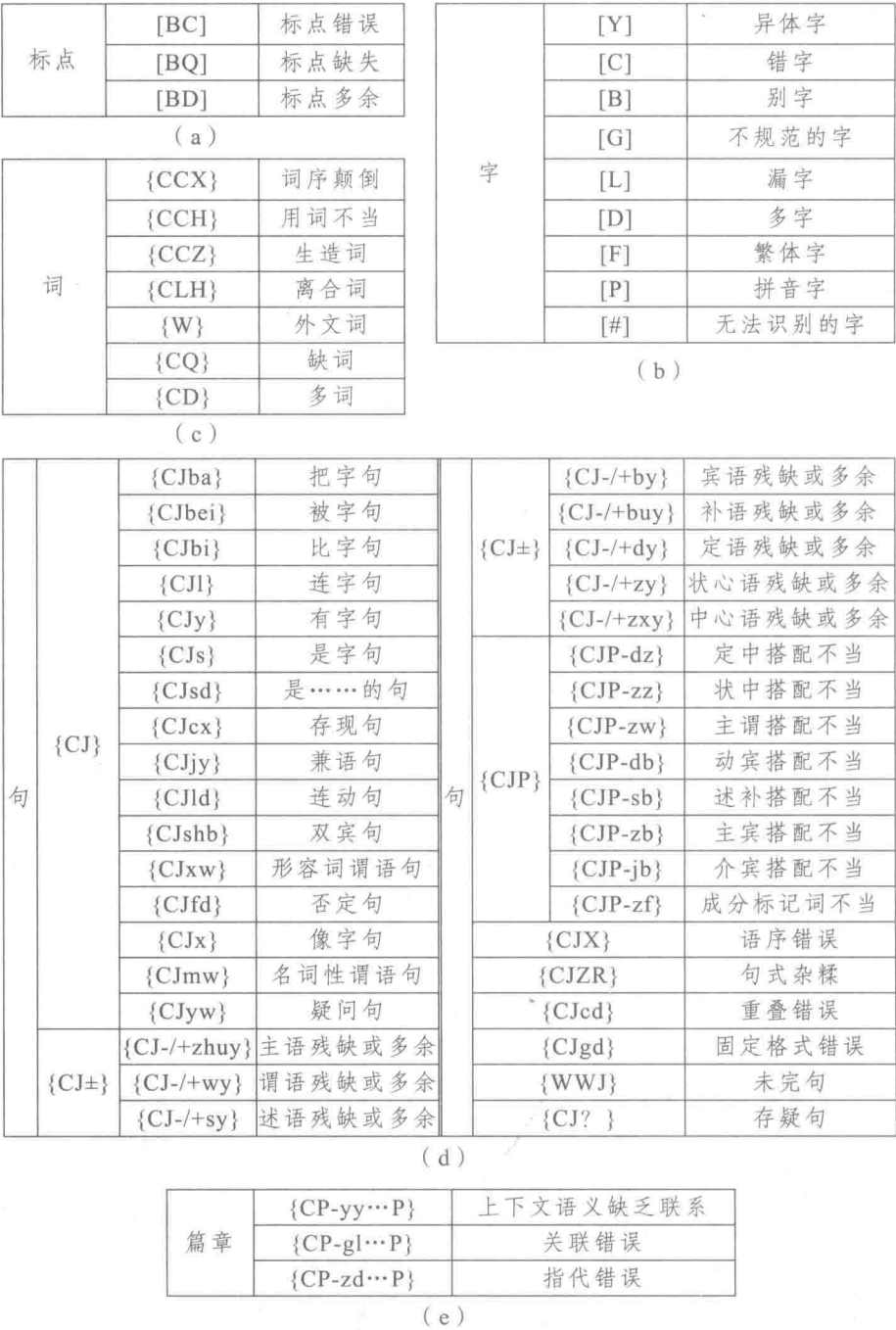


图 2-1 语料库偏误标注类型

数据库中存在标点信息库、字信息库、词信息库、句信息库、篇章信息库 5 个子信息库。具体内容见图 2-2。

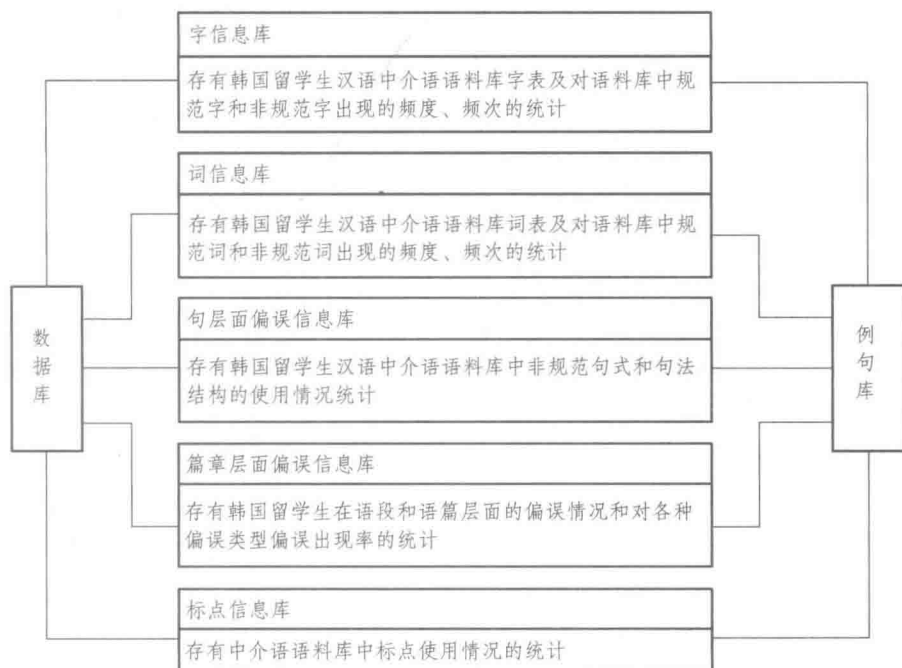


图 2-2 国别化语料库的内容构成

用户检索系统为自开发的语料库检索软件。该软件向用户开放，用户可根据自己的需要自主选择检索并输出语料库中的字、词、句、篇章、标点使用信息及偏误信息。随着语料库建设的进展，上述部分，尤其是语料加工中的偏误标注可能会有相应调整或细化。

## 第二节 语料库建设的必要性

语料库语言学的研究不但要基于本体语料库，也需要大规模中介语语料库的支撑。目前国内中介语语料库的数量尚满足不了语料库语言学发展的需求。而第二语言习得与汉语教学研究更迫切要求建立规模大、种类全、功能细的汉语中介语语料库。

截至目前,国内已开发的中介语语料库主要有汉语中介语语料库、HSK 动态作文语料库,以及南京师范大学汉语中介语偏误信息库(简称南师大语料库)、中山大学中介语语料库(简称中山大学语料库)、暨南大学留学生语料库(简称暨南大学语料库)等。其中 HSK 动态作文语料库全开放使用,其他语料库部分开放或封闭使用。很多人包括部分学者曾对单国别汉语中介语语料库的建设有所质疑,认为多国别中介语语料库中已包含的国别中介语情况可在多国别语料库中检索、提取,与其建设单国别中介语语料库,不如扩大多国别语料库的规模。对此,我们持有不同意见。

从语料数量来看,目前语料库中韩国留学生中介语语料不够充足。如汉语中介语语料库 100 万字加工语料中朝鲜语占 15%(陈小荷,1996),即韩国学生汉语中介语加工语料为 15 万字。其他中介语语料库未见国别抽样具体数据,但只要是平衡语料,韩国留学生语料应不超过 100 万字(基于 HSK 动态作文语料库总规模 400 万字、南师大语料库 100 万字、中山大学语料库 100 余万字、暨南大学语料库 400 万字的初步数据信息)。如语料再进行程度分级,分布到每个层级的韩国学生中介语语料会更少,这样就无法为单国别中介语偏误研究和国别化汉语教学提供足量的语料。因此有必要建设较大规模的韩国留学生汉语中介语语料库。

从语料层级来看,目前规模最大的 HSK 动态作文语料库采自高级汉语水平考试作文语料,因此,语料均为高级学段作文。汉语中介语语料库中 15 万字韩国学生语料若分布到初、中、高三级,每一层级语料量会更少。其他类同。这就使基于中介语语料库进行汉语字、词、句、篇的难度序列研究受到分层级语料数量不足的制约。因此有必要对分层级中介语语料库予以关注。

从语料动态性来看,文中所涉中介语语料库均为动态语料库,但侧重点各有不同。HSK 动态作文语料库的动态性偏重于历时的可扩充性,即可随着 HSK 高级考试的逐年进行不断补充新的语料。然而,库中很难收录到同一学习者的历年动态语料。汉语中介语语料库本意也要对同一学习者不同学习阶段语料进行跟踪收集,以便开展跟踪性调查研究,然而在取样时为了“使核心语料中各种属性的语料分布比较均匀”(陈小荷,1996),不得不“损有余而补不足”(陈小荷,1996),规定同一作者的语料一般最多抽取 4 篇。如此便无法开展学习者个案跟踪研究。