
汉语作为第二语言 学习者词汇习得研究

The Acquisition of Chinese Lexicon by CFL Learners

王 瑞◎著



汕頭大學出版社

汉语作为第二语言 学习者词汇习得研究

The Acquisition of Chinese Lexicon by CFL Learners

王 瑞◎著



汕頭大學出版社

图书在版编目(CIP)数据

汉语作为第二语言学习者词汇习得研究 / 王瑞著

— 汕头 : 汕头大学出版社, 2018.8

ISBN 978-7-5658-2922-2

I. ①汉… II. ①王… III. ①汉语—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(2018)第 202561 号

汉语作为第二语言学习者词汇习得研究

HANYU ZUOWEI DI-ER YUYAN XUEXIZHE CIHUI XIDE YANJIU

著 者: 王 瑞

责任编辑: 宋倩倩

责任技编: 黄东生

封面设计: 黑眼圈工作室

出版发行: 汕头大学出版社

广东省汕头市大学路 243 号汕头大学校园内 邮政编码: 515063

电 话: 0754-82904613

印 刷: 北京市金星印务有限公司

开 本: 710mm×1000mm 1/16

印 张: 11.25

字 数: 193 千字

版 次: 2018 年 8 月第 1 版

印 次: 2018 年 10 月第 1 次印刷

定 价: 42.00 元

ISBN 978-7-5658-2922-2

版权所有, 翻版必究

如发现印装质量问题, 请与承印厂联系退换

序

王瑞博士的《汉语作为第二语言学习者词汇习得研究》一书拟将出版，嘱我写一序。作为她的导师，我欣然承诺。该书有博士论文的底子，但作者又增加了一些新近的研究文献和内容，所以，我仍然从头到尾重新读了一遍。与此同时，我刚毕业的研究生的论文题目也是关于汉语学习者“生造词”的研究。（高飞 2018）虽然两项研究相隔近十年，但总体而言，汉语词汇习得领域关于“生造词”的产出和加工机制的研究却没有太多的新进展。这一领域研究相对滞后的原因有多种，但归结起来，无外两种：一是理论视野，二是研究方法。

汉语词汇习得研究，最初始于汉外词汇对比研究。受对比分析理论的影响，汉语教学界一度希望通过汉外语音、词汇、语法甚至文化的系统对比来探索汉语学习者习得规律。但事实证明，对比分析对探讨汉语学习者词汇习得过程和机制的贡献非常有限。之后，受偏误分析理论的影响，汉语教学界出现了大量汉语词汇、语法偏误分析的研究。然而，由于偏误分析本身在理论和方法上的局限，汉语词汇习得的偏误分析流于偏误表层结构和表层策略的描写，对学习者的词汇偏误产生的原因的阐释仁者见仁，智者见智。这种研究方法对研究学习者偏误的产出和加工机制显得无能为力。尽管如此，在众多词汇偏误分析研究中，邢红兵（2003）通过大量汉语学习者中介语语料库的统计分析，提出了词汇偏误中“生造词”这一具有重要研究价值的偏误类型。由此，汉语学习者“生造词”研究成为汉语词汇习得研究一个良好的起点和重要研究领域。

近些年来，围绕汉语学习者“生造词”的偏误分析仍然占有很大的比重。但是，

随着理论视野的拓宽,汉语词汇习得研究的心理语言学视角为“生造词”研究带来新的生机。特别是 Jiang Nan (2000, 2002, 2004) 关于词汇习得与表征的研究,引起汉语教学界学者对汉语词汇习得表征研究的极大关注。王瑞博士正是在这一理论框架下探讨汉语学习者“生造词”的习得与表征研究。这一研究也是汉语教学界从心理表征的角度较早探讨汉语学习者“生造词”产出和加工机制的行为实验研究。通过这一研究,我们对汉语学习者“生造词”心理表征和加工机制有了更为深入的认识。

汉语学习者“生造词”研究作为观察学习者词汇表征的建立过程,以及“生造词”产生的原因和心理机制提供了一个特定的窗口。通过这个窗口,王瑞的研究有两个重要的发现,一是汉语学习者词汇表征的建立不是一蹴而就的,而是经历了“形式联结”“形式/词目联结”“L1 词目中介”和“重组”四个阶段。在心理表征建立之前,学习者根据已有的构词法知识进行构词,必然会产生“生造词”;二是整词表征和词素表征混淆。研究表明,对低水平学习者而言,更多地依赖语素表征单位进行构词,而高水平学习者更多依赖整词表征单位构词。这就是说,表征单位是汉语学习者产生“生造词”的主要原因。但是,由于“生造词”在学习者的心理词典中没有对应的心理表征,因此,“生造词”具有临时性和不稳定性。这些结论说明,“生造词”这一特定窗口为观察其生成机制提供了基础和前提,但是,特定的理论视野和研究范式则是获得上述结论的必要条件。这也是理论的价值所在,当然也是作者该项研究的价值所在。

顺便说一下汉语语素教学的问题。在汉语教学界,一直存在着“字本位”和“词本位”的争论。一派认为,“字本位”反映了汉语的本质特征,符合汉语的特点和规律,因而主张“字本位”教学。一派认为,“词”是交际的基本单位,“词本位”符合汉语教学的特点和规律。与前一种观点相似的是语素教学观点。在我们看来,教不教语素,怎么教语素,不能一概而论。从王瑞博士的研究以及其他相关研究的结论看来,汉语教学初级阶段似乎不适合语素教学。如果在初级阶段过于强调语素教学,但缺少构词法知识的教学,会使学习者产生更多的“生造词”,也不利于学习者整词表征的建立。教学顺序应该先整词,后语素。通过大量的含有相同语素的整词学习,学习者的语素意识会自然涌现。另外,从“生造词”的结构类型的分析可以看出,汉语学习者的构词法意识出现的比较早,这符合成人二语学习者的认知特点。但是,

成人二语学习者这一优势往往是产生“生造词”的一个重要原因。

总之，王瑞博士的研究为汉语词汇习得研究提供了一种新视角，研究的结论对汉语词汇教学教什么、怎么教提供了实验研究的依据。在未来的研究中，应进一步拓宽理论视野，采取新的研究范式，进一步揭示汉语学习者词汇习得心理机制和认知神经机制。

是为序。

北京语言大学对外汉语研究中心

王建勤

2018年7月3日于北京

摘 要

Levelt (1989) 指出, 表征研究和习得研究密不可分。但是, 与习得研究相比, 表征研究一直没有受到足够的重视 (Jiang 2000)。汉语作为第二语言 (L2) 的词汇研究尤其如此。表征作为习得的心理机制, 必然会体现在学习者的语言加工之中。从汉语词汇习得的角度看, “造词偏误”是最常见的一类词汇加工偏误。这类偏误是学习者根据已知或已有规则和词素“创造”出的中介词汇形式, 反映了学习者词汇心理知识表征的发展状况。把词汇心理表征研究与词汇习得研究相结合, 能帮助我们深入了解词汇习得过程、发现词汇习得的内部机制, 对词汇教学更有着直接的指导意义。

本书采用实验研究的方法, 首先验证 Jiang (2000) 提出的 L2 词汇心理表征发展模型假说对汉语词汇习得的适用程度。然后根据对学习者的造词偏误的分析, 考察与偏误相关的主要词汇知识类别的发展状况, 探讨母语为英语的汉语第二语言学习者造词偏误的心理机制。

研究一和研究二通过考察不同习得阶段 L1 和 L2 词目信息在汉语词汇认知中的作用, 在验证 Jiang (2000) 提出的 L2 词汇心理表征发展模型的基础上, 提出汉语词汇心理表征发展要经历的四个阶段: “形式联结阶段”“形式/词目联结阶段”“L1 词目中介阶段”和“重组阶段”。

造词偏误的产生是由于词条表征内容信息不足或不完善。偏误分析表明, 与造词偏误相关的词汇知识主要有三类: 语义、词边界和构词知识。因此, 研究三、研究四、研究五分别考察这三类词汇知识表征的发展过程, 从认知的角度, 对学习者的

不同类型造词偏误的心理机制进行说明。第一类偏误是同 / 近义替代类偏误，形成的原因是表征语义知识发展不完善。汉语词语义的提取需要通过母语即英语对译词完成，当两个汉语词对应同一个英语对译词，就形成了 L2-L1 之间多对一的联结关系，造成混淆。当多个汉语词对应多个英语对译词，就形成了 L2-L1 之间多对多的联结关系，使得词汇的认知更加困难；第二类偏误是整词与词素混用类偏误，此类偏误的形成是由于词边界信息不明，学习者无法区分词与词素边界；第三类偏误是无汉语对应词的生造词，主要是由于学习者利用已经有的汉语构词知识，通过提取已知词素创造出来的。

根据以上实验研究结论，本书进一步将词汇心理表征研究与造词偏误类词汇知识发展研究相结合，建立汉语生造词这一特殊类别的汉语词汇心理表征模型，并探讨了表征的发展模式。最后，从教学的角度思考词汇心理表征相关问题，讨论汉语词汇教学的任务和策略。

关键词： 词汇心理表征，词汇偏误，生造词，心理机制

Abstract: Chinese Lexical Acquisition of CSL Learners

Any theory of second language acquisition is incomplete without a representation component, because, as pointed out by Levelt (1989), representation and acquisition cannot be studied independently of each other. However, compared to acquisition, representation has received little attention from second language researchers (Jiang 2000). This is especially true with Chinese vocabulary acquisition as a second language (L2). As the inner mechanism of acquisition, representation unavoidably affects the lexical processing. From the perspective of Chinese vocabulary acquisition, word-generating errors are among the most commonly-seen lexical processing errors. They are interlanguage forms *created* by the learners with morphemes and regulations available in their brains, reflecting the developing status of their mental lexical representation. Combining the study of Chinese lexical representation and that of vocabulary acquisition would promote our in-depth understanding of the process and mechanism of Chinese vocabulary acquisition, and thereby help to improve the Chinese L2 vocabulary teaching techniques.

The current study applies empirical approaches to test the applicability of the L2 lexical representation development model proposed by Jiang (2000) in Chinese vocabulary learning, and then explores the development of the main types of lexical knowledge based on word-generating error analysis. The mechanism of the errors is therefore discussed.

Study I and study II explore respectively the mediation of L1 and L1 lemma in Chinese vocabulary cognition. Based on Jiang's model, four stages of Chinese mental lexical

development are proposed: *formal association stage*, *formal/lemma association stage*, *L1 lemma mediation stage* and *reconstructing stage*.

Word-generating errors are results of the inadequate representation of lexical knowledge which can be categorized into three types based on the error analysis: word meaning, word boundary and morphology. Study III, study IV and study V are thereby designed to probe into the development of these three types of lexical knowledge representation, and to explain the mechanism of the word-generating errors from a cognitive perspective. Error type I, the synonym/related words substitution error, appears when two or more Chinese synonyms or related words share the same English translation because the cognizing of Chinese words need to activate the links between Chinese words and their L1 translations. Error type II, the words and morphemes substitution error, is due to the little or inadequate word boundary information presented in the lexical entries. Error type III, the typical self-generated word error, is caused when the learners create new words with morphology and morpheme information they have achieved.

Based on the above empirical research results, a specific representation model for learners' self-generated words is set up, and the developing mode of this model is further discussed from a theoretical perspective. The implications for Chinese L2 vocabulary teaching are also discussed.

Keywords: lexical representation, lexical error, self-generated word, mechanism

目 录

第一章 绪 论	001
1.1 研究缘起	001
1.2 理论背景	002
1.3 本书研究的问题	007
1.4 本书的结构	009
第二章 文献综述	010
2.1 第二语言习得领域的词汇研究	010
2.2 心理语言学的第二语言词汇研究	018
2.3 汉语作为第二语言的词汇研究	025
第三章 汉语词汇心理表征发展过程研究	031
3.1 研究一 L1 在汉语词汇认知中的作用	031
3.2 研究二 L1 词目信息在汉语词汇认知中的作用	039
3.3 汉语词汇心理表征的发展过程	048
第四章 造词偏误的心理机制考察	056
4.1 汉语学习者生造词的偏误分析	056
4.2 研究三 两种多重映射条件下的汉—英联结发展考察	059

4.3	研究四 词边界意识发展考察	067
4.4	研究五 构词意识发展状况	070
4.5	研究小结	075
第五章	理论讨论	078
5.1	实验研究回顾	078
5.2	生造词的心理表征构成	079
5.3	生造词的心理表征发展模式	082
5.4	表征的相关理论问题	086
5.5	对本研究的思考	094
第六章	本研究对汉语词汇教学的启示	099
6.1	造词偏误心理机制研究对教学的启示	099
6.2	汉语词汇心理表征发展过程研究对教学的启示	109
6.3	小 结	114
第七章	余 论	117
7.1	本研究存在的问题	117
7.2	需要进一步研究的问题	118
参考资料		119
附 录		151

第一章 绪 论

1.1 研究缘起

近 20 年,词汇习得研究在第二语言习得研究中越来越受到关注。Meara (1996a) 指出:词汇能力是语言交际能力的核心。词汇问题就是语言问题的核心 (Verhallen et al. 1998)。对于学习者而言,学习和使用第二语言 (以下简称“L2”, 母语或第一语言简称“L1”) 的主要挑战就是词汇的掌握 (Singleton 1999)。目前,词汇习得研究已经成为第二语言习得研究的重要领域。

在对外汉语教学界,词汇教学和习得研究也越来越受到重视。杨惠元 (2003) 认为,在整个对外汉语教学中,词语教学自始至终都应该放在语言要素教学的中心位置。学者们从不同角度提出,学习者学习汉语的困难,在很大程度上是词汇问题 (何干俊 2002, 陈贤纯 1999), 即使进入中、高级阶段,词汇偏误仍是影响表达,提高汉语水平的主要问题 (张若莹 2000, 沈履伟 2002)。

综观国内外词汇习得研究,尽管近些年来研究成果斐然,但是正如 Meara (2002) 所指出的,研究模式和研究方法非常有限,大部分词汇研究还是沿着传统的词汇习得研究思路进行。“当前的心理语言学、计算机研究等领域的研究方法和研究成果似乎没有影响到这一领域。” (Meara 2002: 406) 从研究内容上看,还缺乏对词汇习得的发展过程、词汇知识多层面的研究 (Nation 2001), 从学习者的角度考察 L2 词汇发展过程的研究更少 (Lin 2004)。整体上看,第二语言词汇习得研究一直是欧洲

语言的天下 (Meara 1996b, Larsen-Freeman 2000), 针对其他语言的词汇习得发展研究非常有限, 更不用说像汉语这样与拼音文字语言正字法系统完全不同的语言。这种“一头沉”的现象会使整个研究领域陷于畸形发展 (Meara 1996b), 不利于学科理论建设。

Koda (1996) 指出, L2 词汇识别能力与两种语言的正字法系统相似程度相关。学习者在学习一种母语的亲属语言时, L2 词汇的表征常常不是大的问题。正像 Odlin (1989) 描述的: “掌握大多数欧洲语言的词汇就好像学着识别一群有着小小伪装的老朋友, 再花些工夫学习剩下的不认识的词。” (转引自 Lin 2004) 尽管这样的描述有些过于简单化, 但是研究表明, 当学习者的 L1 与 L2 正字法系统迥然相异时, 例如英语和汉语, L2 的词汇加工会比较困难缓慢 (Everson 1988, Mori 1998)。由于 L1 到 L2 的正字法距离效应 (Koda 1996) 和 L1 对 L2 的迁移效应 (Odlin 1989, Koda 1997), 母语为英语的汉语学习者建立 L2 词汇表征, 要比他们学习其他英语的亲属语言, 如西班牙语或法语, 要跨度更大难度更大, 汉语词汇习得的阶段性也许会更强烈。词汇的发展过程是 L2 词汇习得研究的核心, 也是词汇习得研究的前沿问题。因此, 我们有必要对汉语作为 L2 的词汇发展过程进行研究。

另外, 词汇发展过程涉及学习者词汇心理表征、词汇习得、词汇加工等各个方面。有关这些问题的研究对我们深入了解词汇习得的内部机制、发现词汇偏误的心理机制都至关重要, 对词汇教学更有着直接的指导意义。由于传统的第二语言习得理论模式多以语法项目的习得为主要对象, 无法直接应用在 L2 词汇习得研究中 (Lin 2004, Jarvis 2009), 这就需要在现有理论的基础上, 借鉴其他相关领域的理论与方法, 结合汉语的实际情况, 考察汉语作为第二语言的词汇发展模式, 并在此基础上探讨词汇习得的相关问题, 为第二语言的词汇习得的理论建设和教学实践做出贡献。

1.2 理论背景

本研究的主要理论背景是 Jiang (2000) 从心理语言学角度提出的 L2 词汇发展模式假说。

1.2.1 词汇习得模型的提出

一般认为, 一个 L1 词条的心理表征包括语义、句法、词素和形式(语音和拼写)特征, 这些特征可以分为成两个层面: 词目层(lemma) 和词位层(lexeme), 如图 1-1。词目层包括一个词的语义和句法信息, 词位层包括语素和形式信息(Garrett 1975, Levelt 1989)。已有的心理学实验(Van Orden 1987, Perfetti et al. 1988 等)表明, 在 L1 词汇表征内部, 词目信息和词位信息具有高度整合性。也就是说, 一个心理词条一旦打开, 关于这个词的所有的音形义信息都同时自动启用。



图 1-1 词汇表征内部结构 (Levelt 1989)

Jiang (2000) 认为, 与学习 L1 相比, L2 学习者不可避免地要面临两个挑战: 一是语言输入上的质和量的不足; 二是在学习 L2 时, 学习者头脑中已经存在一套 L1 的概念和语义系统。这套概念语义系统是随着 L1 母语习得建立起来的, 必然要影响到 L2 的学习。从一方面看, 成年的 L2 学习者倾向于通过 L1 对译词通达已有的概念和意义系统学习新的 L2 词汇; 另一方面, 已有的概念和意义系统也可能会干扰或阻止新的 L2 词意义特征的生成。

这样看来, L1 词会和已有的语义系统一起作用, 干扰 L2 学习中新词汇语义特征的建立。而且, 由于语义系统是在 L1 词汇学习的过程中建立的, L1 对译词的激活, 或倾向于使用 L1 对译词, 就是 L2 词汇学习中不可避免的现象。

Jiang (2000) 认为, 正是由于 L2 学习的这种特殊性, 造成了 L2 词汇表征与 L1 词汇表征与发展模式的根本不同。根据这一假设, 他提出了 L2 词汇表征发展的三个阶段。

1.2.2 L2 词汇习得三个发展阶段

第一阶段, Jiang (2000) 称为“词汇发展的形式阶段”(the formal stage of lexical development), 是词汇表征发展的初级阶段。习得 L1 词汇时, 词汇的习得是

伴随着意义的建立完成的,即词的语义和词汇形式特征(即 lemma 层和 lexeme 层)同时创建。而在传统教学环境下,学习者的注意力最初主要集中在 L2 新词的形式特征上,比如拼写和发音,对意义的理解大都依靠 L1 对译词完成。所以,这一阶段几乎不会建立新的语义或句法信息。词汇表征的内容(见图 1-2a)仅仅包括新词的形式特征和一个纽带,纽带把 L2 新词和它的 L1 对译词联系在一起,学习者通过 L1 对译词通达语义。

如图 1-2a 所示,这一阶段词汇表征的特点是词目层空缺(De Bot et al. 1997)。但应指出的是,表征中缺少语义、句法和词素信息,并不是说这些信息对学习者是不可知的。学习者可以通过激活 L2-L1 之间的联结纽带,获得 L2 词的部分语法语义信息;另外,学习者也可以外显地学习语法规则。然而,这两种途径获得的语义和语法信息都不是 L2 词汇表征中的有机整合部分。相反,这部分信息是存储在表征之外,在自然交际中是无法自动提取的。它们只是学习者词汇知识的一部分,还没有变成词汇能力。

第一阶段使用 L2 词,需要激活 L2 词与 L1 对译词之间的联结纽带(见图 1-2b)。要表达的概念首先激活 L1 词典中词目信息与该概念相符的 L1 词,再通过有意识的回忆,找到与这个 L1 词联结的 L2 词。这一过程实际上是“词汇联结过程”,也就是双语词典研究者们(Potter et al. 1984)提出的“词汇联结假设”(Lexical Association Hypothesis)。识别词汇时,L2-L1 之间的联结首先被激活,由于 L1 词表征内部信息的高度整合性,L1 对译词的语义、句法和词素等信息也同时被激活,并帮助理解 L2 词。生成词汇时,要表达的概念先激活那些词义特征与概念相符的 L1 词汇,然后 L1 词汇通过联结纽带找到相应的 L2 词。

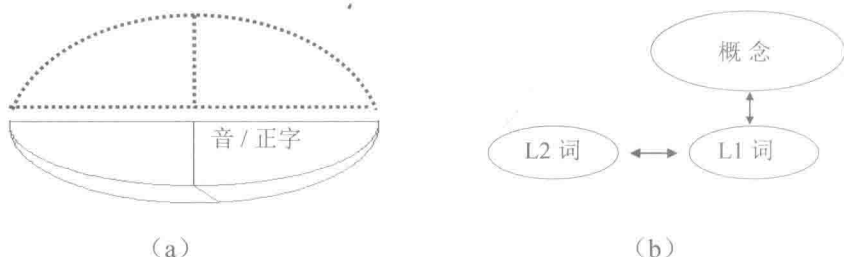


图 1-2 第一阶段词汇表征和 L2 词汇加工

第二阶段称为“L1 词目中介阶段”(the L1 lemma mediation stage),是词汇发

展的中级阶段。随着学习者 L2 语言经验的增加, L2 词形式特征与 L1 对译词之间的联结纽带不断被重复激活, 直到最后, L2 词与 L1 对译词的词目信息(语义和句法特征)的联结达到了自动激活的程度。这个时候, L1 词目信息实际上已经进入 L2 词汇表征, 或者说是被“复制”或“粘贴”到 L2 词汇表征的词目层上, 和 L2 词形一起组成了 L2 词汇表征(见图 1-3a)。

这一阶段的表征特点是 L1 词目“占领”了 L2 词条的词目空间, 但是词汇表征中词素信息依然空缺。因为语义和句法信息可以由两种语言共有, 而词素特征却是各个语言特有的, 不易发生迁移。另外一个特点是, L2 词与概念意义之间的联系不紧密, 原因是词目信息由 L1 复制而来, 不是在 L2 词汇学习过程中创建的, 所以还没有成为词条信息的有机组成部分。意义的提取依然要通过 L1 中介。

不同的是, 到了这个时候, 意义的提取有了两个路径(见图 1-3b): L2 词汇既可以通过表征中的 L1 词目信息通达概念和意义, 也可以通过第一阶段的词汇联结通达概念和意义。由于前者更为直接快捷, 就成为默认路径。这一提取过程与双语心理词典研究提出的“改进型层级模型”(Kroll & Stewart 1994)相一致。所以, 这一阶段 L2 词的使用特点是, L1 词形不再起主要作用。另外, 由于 L2 形式和 L1 词目信息建立了直接的联结, 在生成 L2 词汇时省略了有意识回忆的环节, 所以生成 L2 词的自动化程度大大提高。

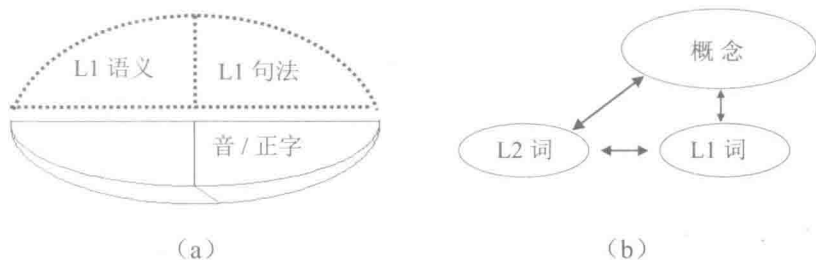


图 1-3 第二阶段的词汇表征和 L2 词汇加工

第三个阶段是 L2 词汇表征发展的“整合阶段”(the L2 integration stage), 也是词汇发展的最高阶段。学习者最终从语言的使用中提炼出 L2 词的语义、句法和词素特征信息, 并将这些信息整合进词汇表征中(见图 1-4)。在此阶段, 无论是词汇表征还是词汇加工过程, L2 词与 L1 词都非常接近。不但表征内容完整, 而且在 L2 使用上也具有较高的自动化程度。