

绪 论

本章首先介绍本书的研究背景；其次简要描述本书的主要内容与研究问题；然后指出其研究意义，包括理论意义与实践意义；最后描述本书的结构。

1. 研究背景

自我国中学新的英语课程标准和教学改革于2002年在全国实施以来，英语教育研究界对如何提高中学英语教学质量展开了讨论。所争论的问题很多，涉及我国英语课程标准、教材、教法、师资培养等方方面面。然而，纵观争论的正反两方面的观点，我们发现它们均缺乏实证研究证据。我们认为，要想真正提高我国中学英语教育水平，对我国中学生英语学习和教学进行系统的科学实证研究至关重要。而系统的科学实证研究的起点应该是：搞清楚我国以中学生为主体的英语初学者¹中介语的现状如何？其所学的英语具有哪些发展规律性特征？这些发展特征的前提影响因素是什么？只有回答了这些问题，才有助于真正解决学好或教好英语的问题。

通过查阅我国公开发表的有关研究文献，我们发现，近几年来我国英语界对中国学习者的英语及其学习过程的研究主要集中在：（1）采用普遍

¹ 在本书中，英语初学者指的是中国初中生与高中生。

语法视角对大学生英语作文的偏误分析；(2) 运用语料库方法对大学英语一些词类（如冠词、形容词、动词、名词）的使用频率、分布情况、搭配模式、出现错误种类等的统计分析；(3) 对中国大学生和英语本族语者在一些语法现象和数量上的差异研究。

这些研究在不同程度上对我国英语教学水平的提高做出了贡献，但是它们很少涉及作为我国英语学习者主体的中学生。另外，这些研究缺乏基于我国英语学习者特点的理论建构，在收集数据时所采用的工具比较单一，多使用调查问卷或大学四六级命题作文语料。因此，很多学者指出，我国英语教与学研究的现状与目前英语教与学的国情不符，没有形成自己系统的研究成果，很少参与国际二语学术研究互动并发出自己研究的声音。

由此可见，对我国中学生英语系统的形成和发展进行系统的实证研究意义重大。它不仅验证西方的二语习得理论是否适合我国教学环境下的英语学习实践，揭示我国中学生英语系统形成和发展的系统性和差异性特征，而且也能够对我国的基础英语教学、英语课程及教材编排提供有用的实证数据。

基于上述认识，我们设计了“基于使用的英语初学者口笔语发展现状及规律研究”项目，并得到了国家哲学社会科学基金青年项目的支持（编号14CYY017）。本课题研究的“英语学习者的语言”在二语习得领域一般称为“过渡语”或“中介语”，是个非常复杂的概念（Selinker 1979）。根据学界普遍接受的看法，本研究将“英语学习者的语言”定义为：具有某种学习动机的学习者通过一定的努力和学习策略，在目的语（英语）输入的基础上所形成的一种既不同于其第一语言（汉语），也不同于目的语（英语），而是随着学习的进展向目的语逐渐过渡的动态的语言系统。

2. 主要内容和研究问题

本研究首先对“基于使用的”（usage-based）语言观以及“基于使用的”二语习得实证研究进行系统的文献研究；通过对文献的梳理与评价，指出“基于使用的”是与“基于规则的”（rule-based）截然不同的二语研究视角，它为传统中介语研究开辟了新的视角与方法论。其次，本研究

在“基于使用的”语言观指导下,论述了“语言接触量假说”的理论内涵、外延。再次,在充分论证的基础上,结合我国汉语背景下英语学习的特点,提出了本研究的理论框架——“汉语背景的英语学习语言接触量框架”(language exposure framework of English learning in Chinese background)。之后,通过采集大量笔语与口语语料,建成了两个中国英语初学者语料库——“中国英语初学者笔语语料库”与“中国英语初学者口语语料库”。最后,基于所建构的理论框架,采用基于语料库的研究方法,本研究进行了较为系统的实证研究,拟回答如下具体研究问题:

- 1) 中国英语初学者各阶段的二语使用特征有何共性?
- 2) 中国英语初学者各阶段的二语使用特征有何差异性?
- 3) 中国英语初学者的二语使用呈现出何种发展规律性特征?该发展规律性特征的前提影响因素是什么?

为回答这三个研究问题,本研究以“汉语背景的英语学习语言接触量框架”为理论依据,基于中国英语初学者口笔语语料库,以英语学习的几个方面(写作词汇丰富性、写作句法复杂度、话语联系语、话语标记语、口语词块、口语动词搭配、口语模糊限制语)为切入点,研究中国英语初学者各个学习阶段之间的共性与差异性,并探讨这些共性与差异性特征所呈现出的二语发展规律性特征。同时基于本理论框架,结合访谈法,讨论上述发展规律性特征的前提影响因素。这几个方面分别代表了英语学习的不同层面——词汇层(词汇丰富性、词块、动词搭配)、句法层(句法复杂度)、语篇层(话语联系语、话语标记语)与语用层(模糊限制语)。在研究初学者的每一个具体学习方面时,均尝试探讨三个研究问题。第一个是关于各个学习阶段之间的共性特征。第二个是关于各个学习阶段之间的差异性特征。而通过研究各学习阶段之间的共性与差异性,可发现初学者二语的总体使用情况及发展趋势,这在一定程度上可反映出二语习得的某种发展规律性特征。因此,第三个研究问题则是挖掘这些共性与差异性特征所呈现出的发展规律性特征,同时,探讨呈现出这些发展规律性特征的前提影响因素。换言之,各实证研究章节均从不同方面回答各学习阶段之间的共性与差异性,挖掘其发展规律性特征,并探讨呈现出某种发展规律性特征的前提影响因素。

3. 研究意义

1) 理论意义

首先，“基于使用的”是与“基于规则的”截然不同的二语研究视角 (Tomasello 2003)，它为传统中介语研究开辟了新的视角与方法论 (Ortega 2013)。我国英语初学者主要通过课堂学习英语，缺乏英语学习的语境，所以研究如何“使用”英语、“使用”结果如何就变得至关重要。因此，基于“使用”视角下的二语习得研究意义重大。

其次，“语言接触量假说” (language exposure hypothesis) 是“基于使用的”语言观视角下的一个重要假说，已引起学界的广泛关注，成为国际二语习得领域的理论前沿。在深入理解“语言接触量假说”内涵与外延的基础上，本研究结合我国英语初学者的学习环境，参考“语言接触量假说”，建构适合我国英语初学者特点的理论框架——“汉语背景的英语学习语言接触量框架”，并将其作为本研究的数据解读理论依据，系统研究不同年龄段英语初学者的习得现状和发展趋势，并探讨出现某些发展规律性特征的前提影响因素。

最后，本研究以我国英语初学者为研究对象，其认知能力发展具有特殊性，其英语学习特征也具有独特性。对其英语学习规律性特征的探索有助于深入理解二语习得的普遍规律。同时，本研究拟推导我国英语初学者英语学习的某些普遍规律，进一步验证支持“语言接触量假说”。

2) 现实意义

本研究的数据源于中国东北地区中学生的限时课堂写作与即时对话语篇，具有完全的真实性与一定的代表性。在系统研究英语初学者二语习得现状及发展趋势、挖掘其发展规律性特征并探寻其前提影响因素的过程之中，本研究的数据结果分析与目前的中学英语课堂教学、中学英语课程标准以及中学英语教材紧密结合，所得出的结论具有较强的现实性与一定的指导性。

首先，本研究可为中学一线英语教师课堂教学方案的进一步调整提供现实参考依据。同时，本研究还可为当前中学英语课程标准的进一步修订、中学英语教材的编写提供可靠的实证依据。

其次，我国英语初学者可通过了解自身语言使用现状及发展趋势，更

有效地扩展显性知识,并逐步内化英语词汇、句法、语篇、语用概念,提高学习效率,加速习得进程。

最后,目前国内尚无话题一致的中学生英语口语语料库,尤其是中学生口语语料库极度缺乏。而进行英语初学者的口笔语比较研究,或研究其口笔语的发展规律性特征具有重要的理论及现实意义。因此,本研究新建的英语初学者限定话题口笔语语料库在国内尚属首个,这可为二语研究及英语课堂教学提供重要资源,这也是本研究对学界二语研究和中学英语教学的一大贡献。

4. 本书的结构

本书是依据国家哲学社会科学基金青年项目“基于使用的英语初学者口笔语发展现状及规律研究”形成的一部专著。本书既含有二语习得理论研究,又包含基于某种理论的实证研究;既有语料库的建构,也有基于语料库的系统实证研究。本书尽量做到见木见林,从而使读者据此择其感兴趣的问题,做更进一步的理论或实证研究,为我国的外语教学与研究做出贡献。

本书共分十一章。

第一章对近30年国内外权威核心期刊百余篇关于“基于使用的”二语习得研究进行文献回顾与梳理,发现国内外研究基本可分为五大类:(1)“基于使用的”二语习得观的建构;(2)“基于使用的”二语输入频率与习得效果关系研究;(3)“基于使用的”二语输出频率与习得特征关系研究;(4)“基于使用的”教学与二语习得关系研究;(5)“基于使用的”二语习得观述评。文献梳理发现,国内外研究在研究对象、语料样本规模、研究语体、研究维度、数据分析深度、该视角在学界的受重视程度这六个方面仍有待进一步拓展,国内较国外在研究广度、深度及数量上尚显不足。本章从这六个方面展望“基于使用的”二语习得研究发展前景,同时为本书的实证研究提供理论视角。

第二章首先从三个方面阐述“语言接触量假说”的内涵、外延。其次,基于文献阅读,本研究认为语言接触量是影响二语习得的前提因素,二语学习者不仅要接触二语语言,还要接触二语所在的社会。然而,就我国英语学习环境而论,学习者基本是在课堂中学习英语,学习的时间长短、

课堂的话语质量等往往决定了学习者的二语质量。最后,结合我国英语初学者的实际,参考“语言接触量假说”,本研究建构了“汉语背景的英语学习语言接触量框架”。“语言接触量假说”是研究二语偏误问题的新视角,基于该假说建构的理论框架为以下各章的系列实证研究提供了理论根据与数据分析路径。

第三章首先讨论语料库在二语习得研究中的重要性,并概括语料库在二语习得研究中所能回答的研究问题。然后对本研究所建设的语料库——“中国英语初学者笔语语料库”与“中国英语初学者口语语料库”进行描述。在描述中主要对语料库的抽样、语料来源、标注工具、库容大小等进行必要的说明。之后,对本研究基于语料库的研究方法中所使用的检索方法和统计方法进行简要图示解读。最后,简要介绍本研究的辅助研究方法——访谈法。

第四章采用本研究建构的理论框架,基于“中国英语初学者笔语语料库”,以Read(2000)的“词汇丰富性多维度评价模型”为数据分类框架,从词汇复杂度、词汇变化度、词汇密度和词汇偏误率四个维度,对比英语初学者各学习阶段之间写作词汇丰富性的共性及差异性,分析其词汇丰富性的总体使用现状及发展规律性特征。

第五章采用本研究建构的理论框架,基于“中国英语初学者笔语语料库”,从产出单位长度、从属结构数量、并列结构数量以及短语复杂度四个维度,统计英语初学者写作句法复杂度各指标的使用情况,对比各学习阶段之间句法复杂度的共性及差异性,探究其句法复杂度的总体使用现状及发展规律性特征。

第六章采用关联理论和本研究建构的理论框架,基于“中国英语初学者笔语语料库”,对比英语初学者三个阶段议论文写作中话语联系语使用特征的共性和差异性,分析初学者话语联系语的总体使用现状及发展规律性特征。

第七章采用本研究建构的理论框架,依据Maschler(1994)的话语标记语分类框架,基于“中国英语初学者口语语料库”,对比英语初学者三个阶段口语中指称类话语标记语使用特征的共性和差异性,探讨其口语话语标记语的总体使用现状及发展规律性特征。

第八章采用本研究建构的理论框架,依据Nattinger & DeCarrico (1992)的词块分类框架,基于“中国英语初学者口语语料库”,对比英语初学者三个阶段口语中词块使用的共性与差异性,探索其词块的总体使用现状及发展规律性特征。

第九章采用母语负迁移理论和本研究建构的理论框架,基于“中国英语初学者口语语料库”,分析英语初学者常用动词、动词搭配的使用特征,并对比各阶段使用特征的共性与差异性,探索其动词搭配能力的总现状况及发展规律性特征。

第十章采用本研究建构的理论框架,以Prince (1982)的分类标准为分类框架,基于“中国英语初学者口语语料库”,对比各阶段英语初学者模糊限制语的使用共性与差异性特征,探索模糊限制语的总体使用现状及发展规律性特征。

本书的每个实证章节(包括第四章至第十章),在研究初学者各种语言能力的总现状况及发展规律性特征之后,均尝试探讨这些发展规律性特征的前提影响因素。

第十一章是本书的总结与启示。本章首先对实证研究部分的研究发现进行总结,归纳英语初学者能力发展的规律性特征;然后根据研究发现,提出中学英语课堂教学、中学英语课程标准的修订以及中学英语教材编写的几点启示;最后说明本研究的局限性,并展望未来发展方向。

我们希望本书:(1)能够为了解我国中学生英语学习现状及发展规律提供一些实证研究数据;(2)能够为我国学界对我国中学生英语的系统性和差异性特征做进一步研究提供理论与方法上的启示;(3)能够为进一步提高我国中学英语教学质量提供有证据的建议。

然而,由于我们的英语初学者语料库尚在扩建之中,在语料库的建设、语料的开发和利用等方面还有很多问题需要解决,所以本书的研究成果存在一些缺点和局限性在所难免,真诚希望同行专家学者和广大读者批评指正。

第一章 “基于使用的”二语习得研究回顾与展望

1.1 引言

二语习得研究受语言学本体研究影响，概括起来主要有三种研究范式，即两种语言结构的对比分析、心灵主义普遍语法驱动“基于规则的”二语研究、“基于使用的”二语研究 (Langacker 2000)。“基于使用的”是与“基于规则的”截然不同的二语研究视角，它为传统中介语研究开辟了新的视角与方法论。虽然受认知语言学与系统功能语言学等多个流派影响，但“基于使用的”二语研究认为：二语与母语一样，是社会共享的符号系统，该系统同样受社会文化和认知方式影响，在使用中学习和建构 (Tyler 2010)。简言之，二语知识源于语言使用体验，二语在交际使用过程中习得或学得 (Ellis *et al.* 2013)。因此，一方面，二语的使用影响习得；另一方面，研究二语习得不可与二语使用分离 (McGloin & Konishi 2010)。在“基于使用的”二语研究中，频率是重要参考依据 (Bybee 2006)：真实话语的输入频率对习得有重要影响；语言输出频率可反映习得的诸多规律。而语料库是“语言系统真实使用的样本”，具有“以频率发现规律”的特性，因此，语料库为“基于使用的”二语研究提供了必要途径。“基于使用的”语言观作为新的二语研究视角，在二语习得和语料库教学研究中获得了广泛应用。基于此，本书首先阐释“基于使用的”语言观的核心观点与核心理念，然后对国内外近30年“基于使用的”二语研究文献进行回顾和梳理。

1.2 “基于使用的”语言观核心观点及核心理念

Langacker (1987) 提出的“基于使用的”语言观以体验哲学为基础,认为语言知识源自语言实际使用的认知体验,即受人的生理条件、认知机制限制,在实际语言体验中建构而成。语言习得过程受到普遍的认知机制 (domain-general mechanism) 的驱使,这些认知机制包括感知和概念系统的结构 (structure of the perceptual and conceptual system)、输入处理能力 (input processing capacities)、输入属性 (properties of input), 等等 (O'Grady 2008)。该语言观还强调,输入属性决定词汇和形态句法的习得速度,这些输入属性包括形符、类符、搭配频率、凸显性和语义复杂度等 (Bybee 2008; Lieven & Tomasello 2008; Tomasello 2003)。“基于使用的”语言观并非一个孤立的理论,而是用来概括一组具有认知功能主义倾向的语言理论,包括认知语言学、构式语法、动态系统理论、复杂系统理论、涌现论、社会语言学和语料库语言学。其中,频率是语言实际使用的一个重要方面,包括形符频率 (token frequency) 和类符频率 (type frequency)。前者是指一个语言表达式在实际语言使用中出现的次数,它决定了其固化程度和被提取使用的流利程度;而后者是指一个语言型式在实际语言使用中的具体表现个数,它决定了一个语言型式使用的可生成性 (王初明 2011)。

“基于使用的”语言观的核心观点是,使用者的语言是在处理大量语言使用事件 (usage event) 中涌现出来的 (Tyler 2010)。使用事件是指在特定的交际情境下,语言使用者为了表达特定的意义而理解或产出语言的语境实例。说话人在交流时,人们往往认为他试图通过对语言策略的有意选择而实现与话语社团成员进行交际的特定目标。将使用事件与特定话语社团相联系反映出这一理念:语言的实际使用是蕴藏在特定的文化与语境中的。

“基于使用的”语言观坚持以下五个核心理念 (Tyler 2010)。

1) 语言存在的根本目的是为了交际,交际使用塑造了语言本身。正如 Tomasello (2003: 5) 针对一语习得所说:“语言的结构源于语言的使用。”

2) 任何自然语言总是在语境中使用,对话语建构的选择受到一系列语境因素的影响。语境本身即为一个复杂的建构,并非是将各个维度组合在一起而不考虑多维度语境中细微的、相互作用的语言映射。比如,所有的“基于使用的”语言观都认识到交际中受众或参与者的作用,将

其看作语境的主要方面 (Chafe 1994; Langacker 1987, 1991; Martin & Rose 2008)。说话人与受众之间关系的细微变化影响着说话人对语言的选择。

3) 语言是后天学得的。语言的使用形式, 包括搭配 (词汇层与语法层) 和频率, 对于学习语言至关重要。这一理念引发了对存在单独语言模块的反驳, 否定了接触主要语言数据能够激发预制语法配置这一说法。

4) 意义不仅来自词汇项, 语法形式本身也有意义, 并能表达意义 (如 Givón 2001; Langacker 2008)。

5) 不必通过划分语言层次来解析语言, 单层模式也能够精确并完全地解析语言, 即不必将句法划分为诸如深层结构与表层结构两个 (或更多) 层次。可相互替换的句法结构是有意义的独立结构, 比如, 主动句与被动句各自独立, 并非相互转换而来 (Zyzik 2009)。

1.3 “基于使用的”二语习得研究回顾

“基于使用的”语言观最初由 Tomasello (2003) 引入儿童母语习得领域, 认为语言习得归根结底是对构式的习得。儿童母语学习者通过自身认知能力, 能够从具体的语言构式中归纳、抽取、概括构式范例, 最终形成抽象的语言构式, 这一过程揭示了语言习得的本质特征。此后, “基于使用的”语言观也逐步受到二语学界的关注。“基于使用的”二语习得研究时代已经来临 (Ortega 2015)。

“基于使用的”二语习得研究在国内外备受关注。学者们多从认知视角出发, 采用实验或语料库的方法进行研究, 取得了令人瞩目的成果, 但国内研究较国外研究略显不足。其研究内容主要包括以下五个方面。

1.3.1 “基于使用的”二语习得观建构研究

语言学习是在语言使用中学习结构的过程, 即学习语言最基本的表征“形式-功能配对体”的过程。为了实现语言的最优加工, 习得机制对语言学习而言是必不可少的证据 (Ellis & Robinson 2008)。语言习得的过程即构式习得的过程, 二语习得依赖学习者语言使用的经历, 其中的决定因素包括: (1) 输入频率 (形次比、Zipfian 分布、语言近期性); (2) 形式 (凸

显和视角);(3) 功能(意义的典型性、信息理解的形式、冗余度);(4) 上述因素的互动(形式与功能关联度)(Ellis *et al.* 2015)。“基于使用的”习得观认为,句法结构、词汇和语义是无法分割的,这些构式在使用的过程中互相联系,互相影响。

近年来,国内外学者主要从以下两个方面探讨“基于使用的”二语习得观的建构。

1) 二语习得观的建构研究

有学者结合“基于使用的”二语习得观,提出外语学习的语句习得假说(韩宝成 2018)。该假说认为,外语学习的基点是“语句”而非词汇和语法。学习者通过接触大量的语句实例,抽取有规律的语言形式,从而创造新的语句。也有学者提出“学伴用随”原则,强调外语学习中多变量的交互作用,由此产生丰富的语境来促学(王初明 2009)。还有学者探讨儿童早期句法发展到到底是受“基于使用”还是“基于规则”的影响,认为这两种多元的研究角度说明了儿童早期句法发展的复杂性(杨小璐 2012)。林正军、张宇(2020)遵循“基于使用的”语言观,从体认语言观视角出发,对外语教学步骤进行了探讨。首先,教会学习者获得语言知识,提高对实例构式的图示化能力;其次,教会学习者使用语言,加强学习者对概念内容的识解和表达能力。

2) “基于使用的”二语构式习得研究

例如, Ellis *et al.* (2016) 立足“基于使用的”语言观,采用语料库统计和心理实验研究法,调查母语和二语题元构式的存储、加工和习得规律,为构式习得和“基于使用的”语言习得理论提供了实证支持,证明了母语和二语题元构式中的动词频率效应的相似性。Bai (2014) 研究了 just me 构式,指出构式是形式和意义配对的语言系统。还有学者发现频率对图式构式的建构有显著影响(杨黎黎、汪国胜 2018)。王启(2019)以“基于使用的”语言理论为框架,提出了搭配构式优先,兼顾语法构式的语言使用模式。该模式认为,语法构式以搭配构式为核心,且仅在缺乏现成的搭配构式时才得以调用。该模式兼顾语义兼容原则,实现了规约性和能产性的统一。

1.3.2 “基于使用的”二语输入频率与习得效果关系研究

国内外学者主要从以下三个方面探讨“基于使用的”二语输入频率与习得效果的关系。

1) 二语输入频率与参数的习得

例如, Yang (2002, 2004, 2010) 构建了参数学习模型, 认为输入频率会影响习得的参数, 所以需要与习得机制结合起来; 该模型还认为, 儿童更容易习得输入中频率高而无歧义的参数。

2) 二语输入频率与词汇的习得

在词汇搭配研究中, 有学者探讨频率在成年人学习英语形容词-名词搭配中的积极作用 (Durrant & Schmitt 2008); 或研究频率对“can 搭配结构”习得的促进作用 (Eskildsen 2008); 或分析学习者搭配结构的习得因素, 发现输入频率对增加二语敏感度有重要作用 (Wolter & Gyllstad 2013)。在词块习得研究中, 有学者探讨频率对二语读写任务中语块使用的影响 (周丹丹、张萌 2016); 有学者探讨输入频率和输出频率对二语语块产出性知识习得的影响 (周正钟 2018); 还有学者研究英文电视剧对学习者的词汇学习的影响 (Kusyk & Sockett 2014)。

3) 二语输入频率与二语概念认知能力的提升

有学者发现, 学习者在语言使用过程中会对使用频率产生敏感性 (Ellis 2002)。Koster (2015) 从动态、“基于使用的”二语习得角度出发, 在教学过程中增加荷兰电影作为地道语言的输入, 发现德国的荷兰语学习者二语总体认知水平得到了显著提高。Stefanowitsch (2008) 则认为, 如果充分理解语言输入频率, 就可以从正面证据中推断出负面证据。他认为在“基于使用的”习得观中, 这种负面证据的概念化过程即负面固化的过程。也有学者发现, 基础迁移引起的语言认知能力的变化受到语言使用频率的显著影响 (Bao 2010)。还有学者研究频率作用对二语写作的影响, 发现频率对提高认知能力及写作水平有较大帮助 (周丹丹 2011)。

1.3.3 “基于使用的”二语输出频率与习得特征关系研究

二语习得研究认为, 二语输出频率是语言习得多维特征的重要反映。国内外学者主要从以下三个方面研究二语输出频率与习得特征的关系。