

# 二语习得视角下“被”搭配的跨语料库对比研究\*

方雪至<sup>1,2,3</sup>, 谢永慧<sup>1,2,3</sup>, 崔悦<sup>1,2,3</sup>, 杨尔弘<sup>2,3</sup>

(1. 北京语言大学 信息科学学院, 北京 100083;

2. 北京语言大学 国家语言资源监测与研究平面媒体中心, 北京 100083;

3. 北京语言大学 语言资源高精尖创新中心, 北京 100083)

([Jasonfang3900@gmail.com](mailto:Jasonfang3900@gmail.com); [1098751719@qq.com](mailto:1098751719@qq.com); [1132174749@qq.com](mailto:1132174749@qq.com); [yerhong@blcu.edu.cn](mailto:yerhong@blcu.edu.cn))

**摘要:** 词语搭配是语言学领域的一个重要概念,也是二语学习者掌握一门语言的关键因素。本文对汉语母语语料库和汉语学习者语料库中“被”的搭配进行对比,从类联接、语义倾向和语义韵角度进行考察,揭示学习者语料库在“被”的搭配上与母语语料库之间存在的差异,并分析原因,以便更好的进行“被”的研究与学习。

**关键词:** “被”, 语料库, 搭配, 二语习得

## A Comparative Study on the Collocation of "Bei" in Cross-corpus from the Perspective of Second Language Acquisition

**Abstract:** Collocation is an important concept of linguistics and applied linguistics. It is also a key factor for second language learners to master a language. Therefore, based on a systematic comparison of the collocation of bei (被) in native corpus and learner corpus, this paper studies bei(被) in three aspects, i.e. colligations, semantic preferences and semantic prosody and explores the changes of the collocation of bei(被) in the two corpus. Finally, the thesis analyzes the causes, aiming to be helpful to the learning and studying of bei(被).

**Keywords:** “bei”, corpus, collocation; second language acquisition

### 1 前言

随着定量研究的不断增多,语料库语言学已经成为语言研究的主流。在语料库语言学中,就如何对检索行所提供的大量的语言示例进行统计分析,1996年Sinclair提出了扩展意义单位模型:搭配、类联接、语义倾向和语义韵。这一模型,将词汇、语法、语义和语用四个层面巧妙地结合起来,为语料库语言学的研究提供了新的方法手段。

#### 1.1 扩展意义单位模型

目前学界对扩展意义单位模型的界定已基本达成共识,分别为1)搭配:两个或两个以上的词在文本中相距不远处共同出现;2)类联接:词语与语法范畴之间的结伴关系;3)语义倾向:搭配词的共同语义特征;4)语义韵:两个或两个以上词的习惯性共现所呈现的语义氛围,分为积极、中性与消极三类语义韵。

采用扩展意义单位模型展开的研究主要有两个方面:一是词语个案的研究,多为同义词、近义词的辨析,同类词的对比研究,如范一凡(2019)对“交际、交往、来往”等交际类动词进行了研究;二是应用的研究,应用于词汇教学、词典编纂等领域的研究,如张会平(2013)通过对常用动词的考察,探索概念迁移规律,对英语词汇教与学具有启示意义。

\* 语言资源高精尖创新中心项目(No. TYZ19005)

介词“被”及其构成的“被”字句一直是二语教学的重难点，故本文选取“被”为研究对象，属于第二种应用类的研究。

## 1.2 二语视角下的“被”

在汉语作为第二语言学习领域，“被”的研究成果颇多，学界普遍基于多样的语料进行研究。李婕、陈晨（2012）结合北京语言大学 HSK 动态作文语料库，对不同汉语水平的留学生“被”字句习得情况进行了调查。周华文、肖溪强（2009）结合对汉语母语者语料库和中介语语料库考察“被”的分类；彭淑莉（2008）结合中山大学中介语语料库和 200 多名初中高级留学生在规定时间内随堂创造的语料，对“被”字句习得难度和习得顺序进行考察研究；张信爱（2013）设计问卷调查从而分析韩国留学生“被”字句中的偏误。

在研究内容上主要包括句法、语义、搭配等方面。在句法方面，刘月华（2001）、周文华、肖溪强（2009）、李婕、陈晨（2012）先后提出了“被”字句句式的不同分类，以便更好地服务于二语教学；在语义方面，传统上“被”字句一直表示“不如意”“不愉快”的感情色彩，但随着研究的逐渐深入，萧斧（1952）提出“被”字句的另一种意义，即“无所谓愉不愉快”，马纯武（1981）认为“被”字句既可以表示贬义和中性意义，也可以表示褒义，祖人植（1997）提出“被”字句有三种语义，分别是：绝对不如意、相对不如意和非不如意；在搭配方面，根据知网的检索，仅有汪路遥（2019）《汉语中介语中“挨”“受”“被”的搭配研究——基于 HSK 动态作文语料库》一篇从二语习得视角下对“被”的搭配展开研究，该文先在汉语词典与语料库中考察“挨”“受”“被”的释义与一般搭配情况，之后基于 HSK 动态作文语料库对正确搭配从词语的词性、等级、语义类、感情色彩等方面分析，对错误搭配进行偏误分析。

## 1.3 研究设计

本文在以上研究的基础上，尝试将类联接、语义倾向、语义韵的方法应用于对比汉语母语语料库与汉语学习者语料库（以下简称为母语语料库与学习者语料库）中介词“被”的搭配研究，从而对语言学角度的“被”进行更全面和细致的考察。介词“被”的搭配形式包括“被”字句和“被”字结构。“被”字句是指“被”将施事置于句子主谓之间或者单纯用“被”附在句中动词之前表示被动的句子。“被”字结构是指“被”与其他动词构成短语在句中充当修饰成分。无论在“被”字句中还是“被”字结构中，节点词“被”后搭配的动词都至关重要。动词前可加修饰成分构成状中结构，研究表明，汉语中状语的构成成分非常丰富，各类实词大都可以作状语，但其中最常见构成成分为副词和形容词。因此本文以“被”后搭配的动词为中心，研究节点词“被”的动词搭配及修饰动词的副词、形容词搭配在母语语料库和学习者语料库中的差异，并分析差异产生的原因，以期对“被”在二语学习及语言深入研究提供一些参考。

# 2 在语料库中提取“被”字句及“被”字结构

## 2.1 语料筛选及数据处理

本研究收集了约 433 万字的汉语母语语料，及约 426 万字汉语学习者语料。其中，汉语母语者语料来自 DCC 动态流通语料库，汉语学习者语料来自 HSK 动态作文语料库。语料选取方面我们考虑到了语言的时间特征，将两类语料筛选条件都限制在 1992-2005 年的范围内。另外，由于学习者语料同期的数据量远远小于母语者语料，因此我们对母语者语料以篇为单位进行了随机抽样，在保证时间分布均匀的基础上使得两类语料规模没有显著差异。

我们对语料进行了清洗、分词、词性标注（词性标签采用中文宾州树库标签集），并对本文的研究对象“被”的分词错误进行了人工纠正。纠错后，母语语料及学习者语料的分词结果分别为 265 万词和 293 万词。

## 2.2 类联接抽取

我们使用 python 语言编写了索引行提取、词频统计、类联接抽取等程序。

图 1 部分索引行

结业/NN 考核/NN 时/LC , /PU 他/PN <被> 评为/VV 23/CD 名/M 指挥组/NN 学员/NN  
 , /PU 被/SB 勒令/VV 退学/VV 或者/CC <被> 开除/VV 都/AD 将/AD 是/VC 生命/NN 该  
 /DT 成果/NN <被> 国家/NN 专利局/NN 授予/VV “/PU 实用/NN  
 的/DEC 低矮/JJ 房屋/NN 必然/AD 逐渐/AD <被> 冷落/VV 。/PU  
 就/AD 是/VC 拿/VV 第二/OD , /PU <被> 大伙/NN 誉为/VV 艇/NN 上/LC 的/DEG

图 1 展示了从语料库中提取到的“被”的左右侧各五个形符跨距内的索引行形式。每个形符（词或标点）之间以空格为分隔符，右侧“/”后为词性标记，中心词“被”两侧分别加“<”“>”以区别索引行中其他的“被”。母语语料中“被”字句及“被”字结构共计 2713 条，学习者语料中“被”字句及“被”字结构共计 2483 条。

基于提取的索引行，我们利用正则表达式分别提取了两类语料中待分析的两种类联接：“被”+VV（核心动词），“被”+AD/VA（修饰核心动词的成分）。之后，对提取结果进行了人工检查，以保证后续分析的准确性和可信度。在母语语料库及学习者语料库中，提取到的“被”+VV（核心动词）的数目分别是 2600、3036；“被”+AD/VA（修饰核心动词的成分）的数目分别为 1292、2319。

## 3 “被”的搭配在母语语料库和学习者语料库中的比较

### 3.1 “被”的类联接在母语语料库与学习者语料库中的比较

#### 3.1.1 “被”+VV

由于语料规模巨大，所以取搭配频数在 10 次以上的动词进行分析。

表 1 母语语料库中“被”+VV 共现频数统计

VV 分类	及物动词	不及物动词	总计
频数	846	0	846
比例	100%	0	100%
词数	36	0	36

由表 1 可知：在母语语料中，“被”的搭配动词全部是及物动词。在被字句中，“被”的搭配动词通常是一个具有较强动作性的动词，具有处置性，可以单独和受事主语构成动宾关系。

表 2 学习者语料库中“被”+VV 共现频数统计

VV 分类	及物动词	不及物动词	总计
-------	------	-------	----

频数	1067	468	1535
比例	69.51%	30.49%	100%
词数	33	3	36

由表 2 可知：在学习者语料中，“被”的搭配词不仅有及物动词，也有不及物动词。其中，不及物动词的频数占有 30.49%，主要有：“罚款”“吸烟”“流行”，如：

例 1 在公共场所边走边抽烟的将被罚款。

例 2 因为据说吸烟时，被吸烟的人更不好。

例 3 有不少源于恋爱而被流行的。

### 3.1.2 “被”+AD/VA

在语料中，修饰动词的副词与形容词有两种位置分布，一是附在“被”字后动词前，二是在“被”字前，其具体分布情况如表 3 所示。

表 3 母语语料库与学习者语料库中“被”+AD/VA 共现频数比较

总频数	母语语料库		学习者语料库	
	“被”前	“被”后	“被”前	“被”后
1204	1061	143	1599	68
比例	88.12%	11.88%	95.92%	4.08%

由表 3 可知，学习者语料中作状语的副词、形容词在数量上明显多于母语语料，但学习者语料中“被”后的副词、形容词使用量较少。

## 3.2 “被”的词语搭配在母语语料库与学习者语料库中的比较

### 3.2.1 “被”的动词搭配比较

表 4 共现频数在 20 次以上的动词

母语语料库		学习者语料库	
搭配动词	频数	搭配动词	频数
评为	96	罚款	370
称为	93	污染	151
誉为	80	吸烟	76
授予	48	影响	44
列为	37	发现	35
判处	26	接受	29
任命	25	打	29
淘汰	24	骂	27
发现	24	认为	25
视为	24	称为	25
认为	24	杀	25
占	21	规定	25
列入	20	流行	22
		淘汰	21
		吸引	21
		批评	20

由表 4 可知,学习者语料中高频搭配的动词类符数超过了母语语料,且学习者语料的高频搭配的频数也明显高于母语语料。母语语料中“被”的不同搭配动词的频数分布较均匀,最高频数为 96 次,而在学习者语料中出现了极端值,最高频数达 370 次。这是由于学习者语料来自 HSK 动态作文语料库,其考试作文具有一定的主题性,所以个别搭配词的频数出现极端值。

除此之外,母语语料中的“被”的 13 种搭配动词中,只有 3 种是贬义类动词,分别是:判处、淘汰、占;有 6 种是中性的动词,分别是:称为、列为、发现、视为、认为、列入;有 4 种是褒义类动词,分别是:评为、誉为、授予、任命。而在学习者语料中的“被”的 16 种搭配动词里,有 8 种是贬义类动词,分别是:罚款、污染、吸烟、打、骂、杀、淘汰、批评;有 6 种是中性的动词,分别是:影响、发现、认为、称为、规定、吸引;有 2 种是褒义类动词,分别是:接受、流行。

由此可知,母语语料中“被”的搭配动词的种类排序是:中性类>褒义类>贬义类;学习者语料中“被”的搭配动词的种类排序是:贬义类>中性类>褒义类。

### 3.2.2 “被”的副词、形容词搭配比较

表 5 母语语料库中共现频数大于 3 的 AD/VA 统计表

“被”的左搭配	已、就、不、也、将、都、又、还、多、并、曾、但、却、已经、全、先后、未、均、便、才、快、分别、即、仍、几乎、刚、再、终于、一直、早、没、一度、立即、正、不幸、唯一、逐渐、长期、共、往往、还是、容易、再次、曾经、竟、像
“被”的右搭配	依法、及时、广泛、严重、正式、强行、非法、有效、合法、坚决、陆续

从表 5 可观察到汉语中“被”搭配副词和形容词的用法特点:

1) “被”的左搭配副词和形容词明显丰富于“被”的右搭配。

2) 从音节角度看,“被”的右搭配副词和形容词都是双音节词,在共现频次小于 3 的搭配词中仅发现 1 例为四音节词,其余均为双音节词。这说明“被”的右搭配副词、形容词主要为双音节词。“被”的左搭配副词与形容词既有单音节词,又有双音节词,其占比分别是 62.5%和 37.5%,单音节词更多。

表 6 学习者语料库中共现频数大于 3 的 AD/VA 统计表

“被”的左搭配	将、不、也、就、都、没有、很、还、已经、容易、已、但、才、却、常常、多、经常、没、一直、可是、便、又、自然、一定、最、曾、只、有时、渐渐、常、还是、往往、一般、总是、非常、快、当时、完全、慢慢、自、也许、再、深深、一样、逐渐、好像、太多、所谓、曾经、比较、总
“被”的右搭配	严重、非常、严厉、无情、深深

比较表 5 和表 6 可观察到母语者和学习者习得“被”搭配的同异:

1) 学习者语料中“被”的左搭配副词和形容词明显丰富于“被”的右搭配,这符合汉语中的用法。

2) 从音节角度看,学习语料中“被”的左搭配副词、形容词也以单、双音节为主,但双音节词比单音节词多 3.85%,这与母语语料有所区别。“被”的右搭配副词、形容词均为双音节词,符合母语使用规律。

3) 学习者语料中“被”的右搭配副词、形容词明显少于母语语料中的搭配。观察可知,

学习者语料中右搭配多为形容词，缺少副词。

### 3.3 “被”的语义倾向在母语语料库与学习者语料库中的比较

#### 3.3.1 “被”搭配动词的语义倾向比较

我们统计了与“被”共现频数在10次以上的所有动词，并根据语义将其归类，详见表7、表8。

表7 母语语料库中“被”搭配动词的语义倾向

主要语义倾向		搭配词
1	遭受义	判处、淘汰、占、关押、盗、破坏、查处、剥夺、占领、撤销、打死、绑架、辞退、逮捕、围困、拘留、感染、开除
2	承赐义 <sup>1</sup>	评为、称为、誉为、列为、视为、认为、看作、选为、称作、定为
3	弱及物义 <sup>2</sup>	授予、任命、列入、用、命名、确定、取消
4	感知心理义	发现

表8 学习者语料库中“被”搭配动词的语义倾向

主要语义倾向		搭配词
1	遭受义	罚款、污染、吸烟、影响、打、骂、杀、淘汰、破坏、抛弃、伤害、害、判、罚
2	承赐义	认为、称为、视为
3	弱及物义	规定、流行、吸引、录取、受到、给、知道、采访、使用、说、派到
4	态度义	批评、允许、禁止、尊重、接受、重视
5	感知心理义	看、发现

对比表7和表8可知，学习者语料中“被”的搭配动词的语义类别多于母语语料，“态度义”在母语语料中没有出现。因为在汉语中，表示被动意义的词除“被”之外，还有“受”“挨”等，如：“受重视”“挨批评”，还可以用“得到”“获得”等词，如：“得到允许”“获得尊重”，但区分这些词的用法是一大难点。对于学习者而言，用目的语产出句子的过程实则是进行翻译的过程，以英语为例：英语中被动句中的“be 动词”相当于汉语中的介词“被”，所以学习者在任何表被动的情况下都会首选“被”。除此之外，两个语料库搭配动词的类符总数相同，都是36个，但母语语料中，“遭受义”和“承赐义”的动词种类最为丰富，而在学习者语料中，“遭受义”和“弱及物义”的动词种类最多。这是由于如今的介词“被”是由表遭受义的动词“被”虚化而来，“遭受义”是“被”字句最基本的语义倾向，所以两个语料库中都是“遭受义”的动词最多。

由此可见，“被”的语义倾向在母语语料和学习者语料中的主要差异是：

- 1) 母语语料中没有“态度义”动词，而学习者语料中有“态度义”动词。
- 2) 母语语料中“承赐义”的动词种类远比学习者语料中的丰富。
- 3) 母语语料中“弱及物义”的动词远没有学习者语料中的多。

#### 3.3.2 “被”搭配副词、形容词语义倾向比较

<sup>1</sup> 承赐义“被”字句是由邢福义先生提出的。

<sup>2</sup> 弱及物义是由朱一凡、胡开宝先生提出。

结合黄伯荣、廖序东版《现代汉语》对状语的语义类别的分类方法，我们对母语语料库和学习者语料库中与介词“被”的共现频数在3次以上的副词与形容词进行汇总，将其语义倾向做如下统计。

表9 母语语料库中“被”搭配AD/VA的语义倾向

		搭配词
限制性	时间	已、将、曾、还、已经、才、仍、刚、一直、早、正、往往、曾经、还是、立即、终于、逐渐、再次、一度、及时
	程度	几乎、竟、像
	否定	不、未、没
	方式、手段	先后、分别、长期、依法、陆续、强行
	范围	都、但、全、均、共
	数量	多
	语气	就、并
	关联	也、又、却、便、再
描写性	性质	快、容易、不幸、唯一、严重、正式、非法、有效、合法、广泛、坚决

从表9可观察出汉语中“被”字的语义倾向的特点：

- 1) 汉语中“被”字的语义倾向主要分为九类：表时间、表程度、表否定、表方式手段、表范围、表数量、表语气、表关联、表性质。
- 2) 其中表时间、表性质和表方式手段的搭配用词较为丰富。

表10 学习者语料库中“被”搭配AD/VA的语义倾向

语义倾向		搭配词
限制类	时间	已、将、曾、还、已经、才、一直、曾经、还是、常、常常、经常、渐渐、总是、逐渐、自、当时、有时、再、往往
	程度	很、非常、好像、比较、最
	否定	没有、没
	范围	都、但、只、总、完全
	数量	大多、多
	语气	就、也许
	关联	也、又、却、便
描写性	性质	快、容易、严重、深深、慢慢、一样、严厉、无情、自然、一定、一般、所谓

比较表9和表10可观察出母语语料和学习者语料中“被”的语义倾向的异同：

- 1) 学习者语料中“被”字的语义倾向主要分为八类：表时间、表程度、表否定、表范围、表数量、表语气、表关联、表性质，这八类同时也是母语语料中所具有的。
- 2) 学习者语料中缺少表方式手段的语义倾向。通过观察母语语料，我们发现表方式手段的搭配词除了“长期”是形容词，其余都为副词，通过查阅新HSK词汇等级表得知，其中仅有“陆续”与“分别”在词汇大纲中且为5级词汇，而其他词汇都属于超纲词汇。故由于词汇储备的欠缺使得此类表方式手段的语义倾向缺失。
- 3) 学习者语料中表时间和表性质的搭配词较为丰富，符合汉语使用特点。

表 11 母语语料库与学习者语料库中“被”搭配 AD/VA 的语义倾向频数比较

语义倾向	汉语母语语料库		汉语学习者语料库	
	频数	比例	频数	比例
时间	343	33.50%	509	39.83%
程度	17	1.66%	68	5.32%
否定	108	10.55%	66	5.16%
方式手段	77	7.52%	0	0
范围	120	11.72%	175	13.69%
数量	35	3.42%	26	2.03%
语气	126	12.30%	139	10.88%
关联	138	13.48%	191	14.95%
性质	63	5.85%	104	8.14%

如表 11 所示，母语语料库中各语义倾向所占比例排序为：

表时间>表关联>表语气>表范围>表否定>表方式手段>表性质>表数量>表程度

学习者语料库中各语义倾向所占比例排序为：

表时间>表关联>表范围>表语气>表性质>表程度>表否定>表数量>表方式手段

通过对比得知，有三类语义倾向的排序浮动较大，其中表否定、表性质类语义倾向的排序浮动为两位；表程度类语义倾向的排序浮动为三位。学习者语料中表否定类词排序下降，说明学习者对否定词“不”和“未”的掌握不佳；学习者语料中表性质类词排序上升，且此类词语都为形容词，而形容词与动词搭配多需要加“地”，如：“被严厉地受罚”“被深深地影响”，说明学习者倾向于使用带有“地”字结构的状中短语；学习者语料中表程度类词排序上升幅度大，观察“很、非常、好像、比较、最”发现其都属于初中水平的词汇，学习者容易掌握，且这几个程度副词也是汉语作为二语教学中的重点，故学习者使用的频率大大提高。

### 3.4 “被”的语义韵在母语语料库与学习者语料库中的比较

朱一凡、胡开宝（2014）提出在进行语义韵判断时应遵循以下三个原则：

1) “被”字句的受事主语或“被”字结构的受事者如果是有生命的，其态度意义是针对受事主语或受事者而言的；

2) “被”字句的受事主语或“被”字结构的受事者如果是无生命的，如果有领属者和相关者则针对领属者或相关者而言，如无领属者和相关者则态度意义针对说话者而言；

3) “被”字句的受事主语和“被”字结构的受事者，或者其领属者和相关者如与说话人立场相对，则态度意义针对说话者而言<sup>3</sup>。

有别于传统的语义韵的三分法，本文在语义倾向的基础上结合三分法展开统计与分析。由于在“被”字句和“被”字结构中，其语义氛围主要由动词表达，副词和形容词只起修饰作用，故本文对于“被”语义韵的考察，主要通过动词来分析。

依据上述原则与方法，我们分别在两个语料库中随机挑选了 100 行索引行进行了语义韵的分析，分析结果如表 12 所示。

表 12 母语语料库与学习者语料库中“被”语义韵的对比

<sup>3</sup> 朱一凡，胡开宝.“被”字句的语义趋向与语义韵——基于翻译与原创新闻语料库的对比研究[J].外国语(上海外国语大学学报), 2014, 37(01):53-64.



语义韵		母语语料库		学习者语料库			
		频数	百分比	频数	百分比		
积极	遭受义	1	1%	27%	0	0%	16%
	承赐义	13	13%		1	1%	
	弱及物义	13	13%		12	12%	
	态度义	0	0%		2	2%	
	感知心理义	0	0%		1	1%	
中性	遭受义	10	10%	38%	1	1%	28%
	承赐义	7	7%		1	1%	
	弱及物义	20	20%		19	19%	
	态度义	0	0%		5	5%	
	感知心理义	1	1%		2	2%	
消极	遭受义	27	27%	35%	37	37%	56%
	承赐义	0	0%		0	0%	
	弱及物义	8	8%		17	17%	
	态度义	0	0%		2	2%	
	感知心理义	0	0%		0	0%	

由表 12 可知，在两种语料中“被”的主导语义韵不同，母语语料中中性（38%）所占比重略高于消极（35%）；而学习者语料中消极（56%）所占比重远高于中性（28%）和积极（16%）。这主要是由于学习者语料“被”的中性语义韵中遭受义（1%）的比重较小，同时消极语义韵中遭受义（37%）和弱及物义（17%）的比重较大。这说明二语学习者很好地掌握了“被”的“遭受义”搭配，但同时也反映出随着语言的发展，“被”的“遭受义”不仅可用于消极语义韵，也可用于中性语义韵。那么如何在不同的语义韵下灵活搭配“遭受义”动词需要学习者进一步学习。

此外，“被”在学习者语料中的积极语义韵（16%）也低于母语语料（27%）。这主要是由于学习者语料中承赐义动词搭配少，在 100 个搭配词中，这类词只有 2 个。邢福义（2004）提出承赐义“被”字句在情绪倾向上表示称心的积极义。在母语语料中，在 100 个搭配词中，“承赐义”动词出现了 20 个，其中有 13 个表示积极的语义韵，所以母语语料中积极语义韵的比重高于学习者语料。

## 4 差异产生的原因分析

造成上述差异的原因可从三个角度来分析：一是母语的干扰；二是学习者的抽象思维造成的过度泛化；三是缺乏词语搭配储备。

### 4.1 母语的干扰

在二语学习的过程中，学习者会不自觉地把母语中与二语相似的部分进行迁移。以母语为英语的学习者为例，汉语中的“被”字句与英语中的被动句在结构上有极高的相似性，但汉语中“被”字句受自身语义的限制，其典型搭配为“遭受义”较强的动词，而在英语中没有这一限制，所以二语语料库中会出现较多的“态度义”动词，如“被允许（be

allowed) ” “被尊重 (be respected) ” “被接受 (be received) ”, 而这些都是常用于英语被动句中的动词。

## 4.2 学习者的抽象思维造成的过度泛化

类推能力是语言学习中比较重要的能力, 它可以使第二语言学习者把学到的一部分规则扩展到其他地方, 却是一把双刃剑。一方面, 有利于学习者快速掌握一门语言; 但另一方面, 也可能会造成过度泛化的偏误。如: “有不少源于恋爱而被流行的。” 该句的意思是“有不少歌因为题材是恋爱而流行起来”, 本不需要“被”, 但这里强行加入“被”造成了偏误, 这就是由于学习者把“被”的使用进行类推而造成的偏误。

## 4.3 缺乏词语搭配储备

由于本文中汉语学习者语料选自高等汉语水平考试 (HSK 高等) 的作文, 故 HSK 词汇大纲是学习者必备的词汇, 而词语搭配中或多或少会有超纲词汇。学习者在进行词语搭配时, 由于搭配词的语义类别比较丰富, 有些语义类别难以理解, 大脑中该语义类别的词语储备较少, 由此运用的搭配词有限, 种类较少。例如介词“被”可以搭配多种语义类别的动词。“遭受义”动词是“被”的典型搭配动词, 比较容易理解, 学习者对此类动词掌握良好; 但“承赐义”动词是汉语中特有的动词, 较难理解, 所以在学习者语料中只出现了“认为”“称为”“视为”这三类“承赐义”动词。介词“被”与表方式手段的副词、形容词的搭配在学习者语料库中缺失也是由于此类副词、形容词多为高级词汇或超纲词。

## 5 结语

本文对母语语料库与汉语学习者语料库中“被”的词语搭配进行研究。在类联接方面, 确立了“被”+VV (核心动词), “被”+AD/VA (修饰核心动词的成分) 两类结构为研究对象。研究发现: 母语语料中“被”的动词搭配只有及物动词, 而在学习者语料中, “被”的动词搭配既有及物动词, 也有不及物动词; 对于副词、形容词, 学习者语料中“被”搭配的总频数多于母语语料, 但“被”的右搭配频数少于母语语料。

在搭配方面, 选取左右高频搭配词进行研究。研究发现: 学习者语料中高频搭配的动词类符数与频数皆多于母语语料; 学习者语料中“被”右搭配副词和形容词丰富度不及母语语料。

在语义倾向方面, 采用更贴切的语义分类法进行考察。研究发现: 对于动词, 母语语料中没有态度义, 而在学习者语料中出现了态度义; 对于副词、形容词, 学习者语料中缺少表方式手段的搭配。

在语义韵方面, 突破传统“积极/消极/中性”三分类描述方法, 结合语义类别进行语义韵的分类。研究发现: 母语语料中“被”的语义韵整体呈中性化趋势, 而学习者语料呈消极趋势。

经过分析, 造成上述差异的原因主要是母语的干扰、学习者抽象思维造成的过度泛化及学习者缺乏词语搭配储备。

本文存在的局限: 一是在做语义趋向和语义韵的分析时, 采取了随机抽取的方法, 未对整体数据进行研究; 二是在语义趋向和语义韵的判断仍是基于人的认知, 存在一定的主观性。

## 参 考 文 献

- [1] 梁茂成.什么是语料库语言学[M].上海外语教育出版社,2016.
- [2] 黄伯荣,廖旭东.现代汉语[M].高等教育出版社,2014.
- [3] 汪路遥.汉语中介语中“挨”“受”“被”的搭配研究[D].上海外国语大学,2019.
- [4] 孔信爱.韩国留学生被字句偏误分析与教学[D].辽宁师范大学,2013.
- [5] 范一凡.基于语料库的交际类动词研究[D].华侨大学,2019.
- [6] 顾鲜静.基于语义韵的同义词搭配个例研究[D].上海外国语大学,2012.
- [7] Stubbs,M.Two quantitative methods of studying phraseology in English [J] .International Journal of Corpus Linguistics,2002,7(2):215-244.
- [8] 李婕,陈晨.基于 HSK 动态作文语料库的“被”字句习得考察[J].国际汉语学报,2012,3(01):109-120.
- [9] 彭淑莉.汉语动词带宾语“被”字句习得研究[J].汉语学习,2008(02):91-99.
- [10] 张会平,刘永兵.英语词汇学习与概念迁移——以常用动词搭配、类联接与语义韵为例[J].外语与外语教学,2013(06):38-43.
- [11] 朱一凡,胡开宝.“被”字句的语义趋向与语义韵——基于翻译与原创新闻语料库的对比研究[J].外国语(上海外国语大学学报),2014,37(01):53-64.
- [12] 鲁健骥.偏误分析与对外汉语教学[J].语言文字应用,1992(01):69-73.
- [13] 鲁健骥.外国人学汉语的语法偏误分析[J].语言教学与研究,1994(01):49-64.
- [14] 周文华,肖奚强.基于语料库的外国学生“被”字句习得研究[J].暨南大学华文学报,2009(02):44-50+71.
- [15] 黄月圆,杨素英,高立群,张旺熹,崔希亮.汉语作为第二语言“被”字句习得的考察[J].世界汉语教学,2007(02):76-90+3.
- [16] 邢福义.承赐型“被”字句[J].语言研究,2004(01):1-11.
- [17] 刘月华.实用现代汉语语法[M].商务印书馆,2001.
- [18] 祖人植.“被”字句表义特性分析[J].汉语学习,1997(03):47-51.
- [19] 马纯武.也谈“被”字句的语义问题[J].汉语学习,1981(06):17-21.
- [20] 萧斧.“被动式”杂谈(上)[J].语文学习,1952(03):42-45.