

# 基于CLEC语料库词频对中国学生过去式习得的影响

杨秋萍

(广西民族师范学院 外语系,广西 崇左 532200)

**【摘要】**过去式是英语学习中很重要的一个语法项目。本文采用语料库这一新的研究方法,通过统计分析处于不同水平组的中国学生的作文中出现的过去时态的错误特征,研究词频对动词过去式习得的影响。

**【关键词】**过去式;习得;语料库

**【中图分类号】**H314 **【文献标识码】**A **【文章编号】**1673-1883(2011)01-0140-04

## 一 引言

对于作为形态学研究重点之一的动词过去式习得,Pinker(1988)等人提出了双机制理论:儿童在掌握英语过去时态时,有两种不同的机制在起作用,而Rumelhart与McClelland提出的单机制英语过去时态模型强调儿童学习过去式时只有一种联接主义机制在起作用。词频对过去时习得的影响是两种理论的分歧点之一。到目前为止,针对中国英语学习者过去式习得的研究并不多。基于此,本文采用语料库的研究方法,通过统计分析处于不同水平组的中国学生的作文中出现的过去时态的错误特征,研究词频对动词过去式习得的影响,旨在测试中国学生的过去式习得符合上述哪种机制,从而验证过去式习得理论的实用性及科学性。

## 二 相关研究回顾

英语中动词数目多于10,000个,包括规则与不规则动词,其中不规则占160多个。研究者发现在过去式习得中出现了所谓的“U—形学习效应”,即儿童习得过去式的三个阶段:在第一阶段,只会使用为数不多的动词的过去式,且主要是不规则动词;在第二阶段,开始使用大量动词的过去式,其中大部分是规则动词,而将规则泛化到不规则动词的过去式中,出现很多错误;在第三阶段,恢复使用不规则动词过去式的能力,并保留了规则动词变形(conjugation)的能力。

为了解释该效应,Rumelhart与McClelland

(1986)提出了英语过去时态模型,他们的模型由三部分组成:一个固定的编码网络,一个带有可修正联接的模式连接器(a pattern-associator with modifiable connections)网络和一个竞争的解码约束网络(a competitive decoding-binding network)。他们使用此网络来模拟儿童的学习过程,发现该网络能产生“U—形学习效应”。1988年,Pinker(1988;Pinker 1991,1999;)等人提出了双机制理论:认为儿童在掌握英语过去时态时,有两种不同的机制在起作用:一种是学习一般性的形态规则,由此产生泛化的结果。另一种是“联想学习”(associative learning),将不规则动词的形态与其基本形式逐个对应起来。前者负责一般规律,后者负责单个例外。就是因为学习者初期见到的不规则动词很少,所以正确率高,过了一段时间,遇到的不规则动词数目增多,所以正确率下降,然而继续学习之后,学习者自己总结了规则动词的规律,将不规则动词联想学习,过去式产出正确率增加。

在国内也有研究者研究过去式习得这一问题,不过没有采取语料库的研究方法。本文采用中国学习者语料库语料,就过去式习得中词频的影响进行研究。

## 三 研究目的和方法

本文选取中国学习者语料库CLEC中六类不同水平组中的四类学生作为研究受试,包括高中生学生ST2,大学1-2年级非英语专业学生ST3,大学3-4年级非英语专业学生ST5以及英语专业学生ST6。

表1 标注说明

码	分类	类别	说明
fm1	word	Spelling(拼写)	spelling, coinage, abbreviation, apostrophe
fm2	word	word building(构词)	derivation, inflection, compounding, plurality(noun), irregularity(verb), 3rd person singular form(verb), syllabification, hyphenation, word division or fusion

首先使用AntConc语料库提取软件,按照CLEC语料中的错误标注符[fm1],[fm2](表1)(桂诗春,杨惠,

2003)检索出各组学生作文中的错误过去式,再对所得错误过去式进行分析归类。然后选取了ST2组中高频

收稿日期:2010-11-16

作者简介:杨秋萍(1982-),女,甘肃秦安人,硕士,研究方向为外国语言学及应用语言学。

率的规则不规则动词各10个,低频率的规则和不规则动词各10个(表2),利用SPSS数据分析软件,研究选取的20个动词出现的错误,进而验证理论的科学性。

#### 四 研究问题

1.中国学生在英语过去式习得中是否普遍存在规则化的现象?

2.词频对中国学生的过去时习得有何影响?

#### 五 研究设计和步骤

1.统计并且分类四个不同水平组学生的作文中出现的过去式错误特征(表2)。不规则动词泛化为规则动词的现象不明显,只在ST2中出现一例(1)。其中将“辅音+y”结尾与“元音+y结尾”的动词以及“元音+辅音结尾”的动词应如何变换过去式混淆的特征陈上升趋势。(3)(4)除此之外,还出现了将元音“a”变为“o”,双重过去式化,如eat-ated,以及缺失辅音“d”,“t”等现象。

表2 各组学生作文中的过去式错误分布表

Groups	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ST2	0.7%	7.3%	11.7%	5.8%	11.7%	2.9%	4.4%	55.5%
ST3	0	12.9%	14.3%	1.4%	5.7%	0%	2.9%	62.9%
ST5	0	2.9%	26.5%	5.9%	2.9%	2.9%	0	58.8%
ST6	0	0	31.6%	7.9%	0	0	2.6%	57.9%

备注:(1)不规则化,如relay-relaid;(2)缺失或添加末尾辅音t,d,如feel-felt;(3)双写或未双写末尾辅音t,p,n,l,如stop-stoped;(4)将末尾y变为ied或者未变,如delay-delaid;(5)元音变化,如catch-caught;(6)过去分词,如grow-grown;(7)过度时态,如

eat-eated;(8)规则化,如sit-sitted.

2.本研究将所选动词(表3)在BNC中的子语料库S-classroom(包含433646词)检索,总体上高低频率分类相似,保证了数据的科学性(表4)。数据导入SPSS利用单因变量多因素方差分析对数据进行分析。

表3 所选动词

词	规 则		不 规 则				
	词频	词	词频	词			
flow	1	jump	49	burst	0	sit	79
occur	1	carry	73	spread	5	stand	94
obey	6	happen	75	beat	16	grow	143
regret	6	stop	94	rise	19	hear	215
reply	10	decide	130	hit	19	buy	226
quarrel	10	finish	180	draw	35	write	262
increase	16	hope	204	drink	42	find	340
drop	21	try	249	show	59	know	417
step	30	play	254	fall	59	think	537
laugh	31	study	471	sleep	73	get	555
	低		高		低		高

表4 所选动词在BNC子语料库S-classroom中的频率

词	规 则		不 规 则				
	词频	词	词频	词			
flow	12	jump	19	burst	0	sit	121
occur	12	carry	95	spread	12	stand	91
obey	0	happen	459	beat	11	grow	135
regret	2	stop	134	rise	5	hear	147
reply	8	decide	63	hit	29	buy	152
quarrel	0	finish	180	draw	269	write	554
increase	32	hope	15	drink	12	find	537
drop	34	try	638	show	195	know	1843
step	2	play	152	fall	38	think	1969
laugh	13	study	5	sleep	11	get	4432

对于高频动词,平均值 $ST6=.1000 < ST5=.4000 < ST3=.5000 < ST2=.7500$ 。习得英语时间越长的学生所犯过去式类错误越少;对于低频动词,平均值 $ST6=.3500 < ST5=.6000 < ST3=1.8500 > ST2=.7500$ ,非英语专业大一大二年级学生比高中学生的过去式错

误数更多,而后也是按照习得英语时间越长的学生所犯过去式类错误越少的规律变化。这一结果可以帮助大学英语教师调整授课计划,即涵盖过去式这个语法点,而不是理所当然地认为学生在中学阶段已学,无需讲解及强化。

表5 各组学生过去式错误平均值

因变量:错误数

组	词频	平均值	Std.Deviation	N
ST2	低频	.7500	1.16416	20
	高频	1.9500	2.21181	20
	总计	1.3500	1.84738	40
ST3	低频	1.8500	2.99605	20
	高频	.5000	.68825	20
	总计	1.1750	2.25192	40
ST5	低频	.6000	.94032	20
	高频	.4000	.94032	20
	总计	.5000	.93370	40
ST6	低频	.3500	.93330	20
	高频	.1000	.30779	20
	总计	.2250	.69752	40

表6说明词频和组的交互作用对学生产出错误过去式的多少具有显著影响, $p=.001 < .05$ ,具有统计

学意义。将表5与表6的数据相结合,问题出现了:词频的高低对动词过去式的错误数有何影响?

表6 词频对过去式习得的影响

因变量:错误数

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
词频	.900	1	.900	.467	.495
词频*组	32.750	3	10.917	5.667	.001

a R Squared = .337(Adjusted R Squared=.268)

利用SPSS,得到表7和表8,其中规则动词类错误中, $p=.863 > .05$ ,说明词频和组的交互作用对动词的过去式错误数没有显著影响,而对于不规则动词类错误,词频和组的交互作用, $p=.001 < .05$ ,说明不规则类错误对动词的过去式错误数有极显著影响,具有统计学意义。结合表5与表6中的数据得出结论

为:在过去式的习得过程中,词频对习得是有影响的,对于不规则动词的习得,词频影响显著。如前所述单双机制对于词频的影响有分歧:双认为词频只对不规则动词的习得有影响,而单认为对规则和不规则都有影响。在此本文与双机制理论达成一致。

表7 词频对规则动词过去式习得的影响

因变量:规则动词类错误

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
组	2.700	3	.900	1.486	.226
词频	2.450	1	2.450	4.046	.048
词频*组	.450	3	.150	.248	.863

a R Squared=.114(Adjusted R Squared=.028)

表8 词频对不规则动词过去式习得的影响

因变量:不规则动词类错误

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
组	49.650	3	16.550	5.097	.003
词频	.050	1	.050	.015	.902
词频*组	58.450	3	19.483	6.000	.001

a R Squared=.316(Adjusted R Squared=.250)

## 六 结论与启示

研究结果表明:(1)随着英语水平的提高,各组学生表现出错误类型和错误特征类型都会减少。将规则动词泛化为不规则动词来处理在学习者也普遍存在。(2)通过进一步研究所选取的40个动

词,发现各组学生在不规则动词上所犯错误明显高于规则动词。词频对不规则动词得影响显著,对规则动词的习得影响不显著。本研究与双机制理论关于词频对动词过去式习得的影响的结论一致,验证了该理论的广泛性和有效性。

### 注释及参考文献:

- [1]Pinker, S. & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*[J]. 28, 73-193.
- [2]Pinker, S. 1991. Rules of language. *Science*[J]. 253, 530-535.
- [3]Rumelhart, D. E. & McClelland, J. L. (1986). On learning the past tense of English verbs[A]. In J.L.McClelland et al. (eds.). *Parallel Distributed Processing*[C]. 2, 216-271. Cambridge: MIT Press.
- [4]Timothy T. Rogers, David H. Rakison, & James L. McClelland (2004). U-shaped curves in development: A PDP approach. *Journal of Cognition and Development*[J]. 5, 137-145.
- [5]桂诗春, 杨惠. 中国学习者英语语料库[M]. 上海: 上海外语教育出版社, 2003.
- [6]李平. 语言习得的联结主义模式[J]. *当代语言学*, 2002, 3.
- [7]王克非. 双语对应语料库和应用[M]. 北京: 外语教学与研究出版社, 2004.
- [8]王宗炎. 谈谈桂诗春、杨惠中的科研项目“中国学习者语料库”[J]. *现代外语*, 2003, 2.
- [9]陆燧枫. 二语环境中动词过去式的实证调查与研究[D]. 南京: 南京师范大学, 2004.

## The Role of Frequency in Acquisition of Past Tense for EFL Learners

——A Corpus-based Study

YANG Qiu-ping

(Foreign Language Department, Guangxi Normal University for Nationalities, Chongzuo, Guangxi 532200)

**Abstract:** Past tense is one of the most important grammar points in English learning. The present study is a corpus-based study on Chinese EFL learner's past tense acquisition. By collecting and analyzing the past tense productions of Chinese EFL learners, this study reaches its conclusions in error features and the role of frequency in the acquisition of English past tense.

**Key words:** Past Tense; Acquisition; Corpus

(责任编辑:周锦鹤)