

基于语料库的存在句主语确指性及其习得研究

赵秀莲^{1,2} 徐芝苹¹

(1 宁波大学外语学院 浙江宁波 315211; 2 虞城高中 河南商丘 476000)

【摘要】关于英语存在句“实义主语”是否具有“确指性限制”，争论已久。为了解决这一理论问题并探究确指性限制的二语习得问题，本研究利用语料库实证方法，对 2 个英美经典语料库和 3 个中国二语学习者语料库进行检索。根据前 2 个语料库，发现“确指性限制”总体看并不存在，但人称代词表现出绝对的限制性，且存在句中使用确指性主语的现象并不常见。而根据后 3 个语料库，一方面，中国学生使用存在句的频率稍微偏高，但与 L1 使用者并无显著差异。另一方面，中国学生使用确指性主语的频率大大低于 L1 使用者， P 值 = 0.00 表现出显著的差异性。这些现象是由于语言迁移和回避策略造成的。

【关键词】语料库；存在句主语；确指性限制；习得

【中图分类号】H 319 **【文献标识码】**A **【文章编号】**1672-4860(2011)01-0054-06

A Corpus-based Study of Definite Subjects of Existential Sentences and Its L2 Acquisition

ZHAO Xiu-lian^{1,2}, XU Zhi-ping¹

(1. Faculty of Foreign Languages, Ningbo University, Ningbo 315211, Zhejiang, China

2. Yucheng Senior Middle School, Shangqiu 476000, Henan, China)

Abstract There have been long arguments about whether there is definiteness restriction to the subject of existential sentences. To solve this theoretical problem and related problems in L2 acquisition, an empirical and systematic study is made of 2 L1 speakers' corpora and 3 Chinese L2 learners' corpora. According to the first two corpora, findings show that "definiteness restriction" does not exist as a whole but there is absolute restriction to personal pronouns and relatively rare use of definite subjects. According to the last 3 corpora, findings show that Chinese L2 learners tend to use a little more existential sentences though they use far less definite subjects than L1 speakers, which may be caused by linguistic transfer and evasion strategy of L2 learners.

Key words Corpus; Subject of existential; Definiteness restriction; Acquisition

一、引言

“确指性限制”(definiteness restriction)是指存在句的“实义主语”不能为确指性 NP，即不能是专有名词、指示代词、人称代词、名词性物主代词及由定冠词 the 指示代词、物主代词、专有名词所有格所限定的名词。^[1]在语言学理论领域，关于英语存在句“实义主语”的“确指性限制”国内外学者的观点有二：一种认为存在“确指性限制”，即存在句“实义主语”应为非确指性的 NP^{[2-7][19-22]}。另一种观点则认为“实义主语”可以由确指性的 NP 充当，即充当存在句“实义主语”的 NP 没有所谓的“确指性限制”^{[8-13][23-27]}。以往学界对

收稿日期：2010-07-08

作者简介：赵秀莲(1980-)，女，河南商丘人，硕士研究生，研究方向：翻译理论与实践，英语教育。

徐芝苹(1984-)，女，浙江温州人，硕士研究生，研究方向：英语教育。

于英语存在句主语的“确指性限制”的研究都是在理论层面上进行的, 其论据主要来自文献理论或者具体的例证。本研究第一部分拟借用语料库实证手段, 对英美经典语料库 (LOB, BROWN) 进行系统地检索, 以检验双方的争论焦点——确指性限制是否存在, 如果不存在限制, 确指性 NP 作主语的存在句所占比例是多少, 即确指性主语的用法是否常见。

在二语习得实践领域, 关于存在句的研究相当薄弱。仅有的研究有的是利用问卷调查的方式进行习得水平分析^[28]; 有的是对 there be 结构中的动词误用和影响习得的因素进行分析^[29]; 有的是对该结构各种形式的分布频率进行检索^[30]。而关于存在句确指性主语习得的研究目前国内外仅有 L. White 教授涉及过。她在实验中发现, 二语学习者虽然在习得冠词时出现各种错误 (冠词缺失、冠词冗余和冠词混淆), 但在习得存在句主语的冠词时却正确率很高, 一般不会出现定冠词 (the) 或限定意味很强的限定词 (all, most, every, demonstratives, possessives, personal pronouns) 来修饰存在句主语, 她把这种现象称为“限定效应” (Definiteness Effect), 简称 DE^[14]。本研究的第二部分拟利用中国学生英语口笔语语料库 (SW ECCL 2.0) 和中国学习者英语语料库 (CLEC), 检验 DE 是否存在, 是否有反例出现, 反例出现的频率与 L1 语料库中的频率是否有显著性差异。

二、研究设计

本研究使用的工具为检索软件 Antconc3.2.1、标注软件 GO TAGGER、数据统计分析软件 SPSS16.0 参照语料库 LOB、BROWN 和观察语料库 SW ECCL、CLEC。

首先, 使用标注软件 Go Tagger 对 LOB、BROWN、SW ECCL 和 CLEC 四个语料库进行词性标注, 再使用 Antconc 对各个语料库中存在句进行提取, 列出各个语料库中存在句总数, 以了解各个语料库中存在句的使用频率。然后, 用 concordance 功能对其中以确指性名词短语作主语的存在句进行检索, 并对检索结果进行人工筛选, 列出每个语料库中各种确指性主语的使用次数及其总数, 并计算出各个库中使用确指性主语的存在句之于存在句总数的比率, 以了解确指性主语是否常用。最后, 用 SPSS 对观察语料库进行单样本 t 检验, 对比观察语料库与参照语料库中存在句使用频率的差异和确指性主语使用频率的差异。这里根据 C. Lyons 的分类标准把作存在句主语的确指性名词结构分为专有名词、人称代词、指示代词及带有定冠词、指示代词和物主代词所修饰的名词结构^[15]。

三、结果及讨论

(一) 英美经典语料库 LOB 和 BROWN 的数据及讨论

经过反复检索和人工把关, 得出以下数据: 在 LOB 语料库 (共 1123380 形符) 中共有 2788 个存在句。表存在的 there (there_EX) 在语料库中的词频密度^①约为 0.25%。剔除似是而非的结构如 “there was the unexplainable”, 其中使用确指性主语的有 180 个, 约占所有存在句的 6.46%, 表 1 是各类确指性主语的分布情况:

表 1 LOB 语料库的存在句使用频率

专有名词及用它作 定语的名词结构	人称代词	指示代词	定冠词修饰 的名词结构	指示代词修饰 的名词结构	物主代词修饰 的名词结构
34 (1.22%)	0 (0.00%)	5 (0.18%)	131 (4.70%)	8 (0.29%)	2 (0.07%)

现就各种确指性主语各摘取一例:

- There was Edinburgh and...
- There is those that...
- There is the horse-racing...
- Again there was that awful silence...
- There was his garage hand Dai...

在 BROWN 语料库 (1014300形符)中共有 2167个存在句,表存在的 there在语料库中的词频密度约为 0 21%。其中使用确指性主语的有 129个,约占有存在句的 5 95%,详细数据见表 2。

表 2 BROWN 语料库的存在句使用频率

专有名词及用它作 定语的名词结构	人称代词	指示代词	定冠词修饰 的名词结构	指示代词修饰 的名词结构	物主代词修饰 的名词结构
35(1 61%)	0(0 00%)	4(0 18%)	83(3 83%)	5(0 23%)	2(0 09%)

例如:

There was the Polish Revolution which...

There is St. Louis County...

There were those who...

There was the further complication that...

There was this poet in Italy...

There lies your answer...

以上两个语料库的检索数据显示,在 L1使用者实际交际中,确指性限制总体上看并不存在,因为在六种确指性主语中,除人称代词外,其余五种都作过存在句主语,尤其是定冠词修饰的名词短语所占的比率相当显著(分别为 4 7%和 3 83%)。而且确指性主语总和在两库中的比率分别为 6 46%和 5 95%,即 100个存在句中,有 6个左右使用确指性主语,具有显著意义。可见,确指性的 NP可以充当存在句的“实义主语”,研究结果基本支持第二种观点。另一方面,确指性主语的使用频率虽然达到了显性程度,但相对于非确指性主语 (100% - 6 21% = 93 79%)的比率,确指性主语并不十分常用。而且各种确指性主语的分布并不均匀,其中人称代词在两个语料库的检索结果均为 0表现出绝对的限制性;物主代词修饰的名词结构的用法每库仅出现两次。两次相对于两千多个存在句来说,几乎没有统计意义;同样,指示代词及其所修饰的名词结构的比率总和在两库中分别为 0 47%和 0 41%,尚不足 0 5%,这与物主代词修饰的名词结构一样,表现出较强的限制性。可见,在五种确指性主语中有四种表现出绝对限制性或较强限制性,这也解释了为什么第一组学者坚持认为存在句“实义主语”存在“确指性限制”。

(二)中国二语学习者语料库 SWECCL及 CLEC的数据及讨论

中国学生英语口语笔语语料库 SWECCL2 0版包括书面语语料库 WECCL和口语语料库 SECCL,其中前者容量是 1248476个形符,包括 3417个存在句, there_EX的词频密度约为 0 27%。其中使用确指性主语的有 55个,约占有存在句的 1 61%。详细数据见表 3。

表 3 WECCL 语料库的存在句使用频率

专有名词及用它作 定语的名词结构	人称代词	指示代词	定冠词修饰 的名词结构	指示代词修饰 的名词结构	物主代词修饰 的名词结构
18(0 53%)	0(0 00%)	2(0 06%)	29(0 85%)	2(0 06)	4(0 12%)

SECCL容量为 779731个形符,包括 2621个存在句, there_EX的词频密度约为 0 37%。其中使用确指性主语的存在句有 45个,约占有存在句的 1 72%,各种确指性主语的具体分布如下:

表 4 SECCL 语料库的存在句使用频率

专有名词及用它作 定语的名词结构	人称代词	指示代词	定冠词修饰 的名词结构	指示代词修饰 的名词结构	物主代词修饰 的名词结构
11(0 42%)	0(0 00%)	0 00%	24(0 92%)	0(0%)	10(0 38%)

表 5 CLEC 语料库的存在句使用频率

专有名词及用它作 定语的名词结构	人称代词	指示代词	定冠词修饰 的名词结构	指示代词修饰 的名词结构	物主代词修饰 的名词结构
28(0 74%)	0(0 00%)	0 00%	31(0 82%)	3(0 08%)	10(0 27%)

CLEC 语料库的容量为 1070602个形符 包括 3800个存在句 表存在的 there词频密度约为 0 35% 其

中使用确指性主语的有 72 个, 约占存在句总数的 1.89%。具体数据见表 5。

为了直观起见, 下面把各个语料库的存在句 *there* 词频密度列成图表。

图 1 中柱 1 到柱 5 依次代表 LOB, BROWN, WECCL, SECCL, CLEC 语料库中存在句 *there* 的词频密度。图表显示, 与英美经典语料库 LOB, BROWN 语料库相比, 3 个中国二语学习者语料库中存在句的使用密度均有偏高的趋势。为了使数据分析更加精确, 下面使用 SPSS 软件对观察语料库存在句使用频率进行单样本检验, 已知参考语料库的存在句频率的均值是 0.23%, 目的在于分析两者之间是否存在显著差异。表 6 是检验结果。

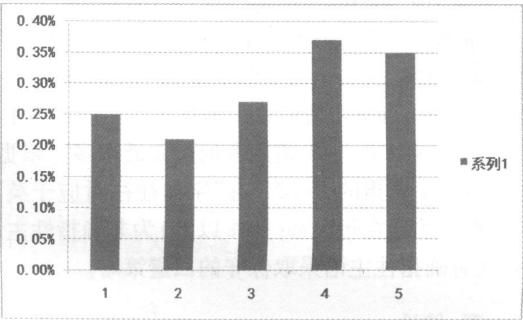


图 1 各个语料库中存在句 *there* 的词频密度

表 6 观察与参照语料库存在句使用频率差异

One-Sample Test						
Test Value = 0.0023						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower	Upper
观察与参照语料库存在句使用频率差异	3.273	2	0.082	0.0010000	-0.000314	0.002314

双尾 *t* 检验的显著性概率为: 0.082, 大于 0.05, 说明不存在显著差异^②。平均值差的置信区间含有零 (从下限 -0.000314 到上限 0.002314 之间有零存在), 也说明观察语料库平均数与参考语料库平均数之间不存在显著差异。

简言之, 相对于 L1 使用者, 中国二语学习者使用存在句虽有偏高的趋势, 但差异并不显著, 即不能简单的说中国学习者过度使用存在句。这一结果不支持 Yip 和 Chou 所说的过度使用现象^[16]。至于这种稍微偏高的趋势存在的原因, 我们认为 Chou 的分析是有道理的: 汉语类似句型“有”字句的转换作用; 汉语的“主题突出性”导致学习者倾向使用 *There BE* 句型来引出语篇中的新主题^[17]。

再把各个语料库中使用确指性主语的存在比率制成图表, 以便比较:

图 2 中柱 1、2、3、4、5 依次代表 LOB, BROWN, WECCL, SECCL, CLEC 语料库中确指性主语出现的频率。由图示看出, 英美国语料库中确指性主语的出现频率为 6% 左右, 中国二语学习者语料库中确指性主语的频率均不足 2%, 前者是后者的 3 倍左右, 呈现极大的差异性。

为了更加精确地验证两者差异的显著性, 表 7 使用 SPSS 对观察语料库的确指主语使用频率进行单样本检验, 已知参考语料库的确指主语使用频率均值是 6.20%, 以比较两者是否存在显著差异:

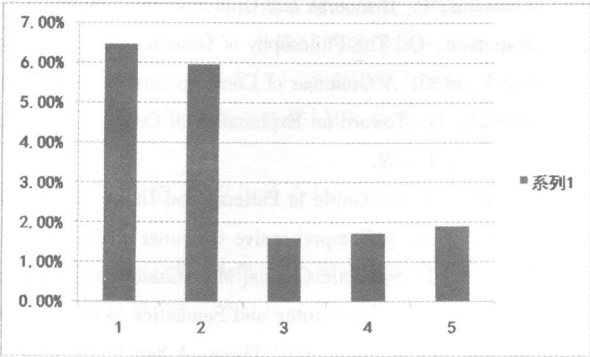


图 2 各个语料库确指主语的使用频率

表 7 观察与参照语料库确指主语使用频率差异

One-Sample Test						
Test Value = 0.0620						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower	Upper
观察与参照语料库存在句使用频率差异	-54.761	2	0.000	-0.0446000	-0.048104	-0.041096

双尾 t 检验的显著性概率为: 0.00 远远小于 0.05, 呈现绝对显著差异; 平均值差的置信区间不含有零 (从下限 - 0.048104 到上限 - 0.041096 之间没有零), 也说明观察语料库与参考语料库之间存在显著差异。

即相对于 L1 使用者, 中国二语学习者很少使用确指性主语, 即在存在句确指性主语使用方面表现出很强的限制性。这一结果不但支持 L. White 教授利用问卷等实验方式所得出的研究结果——限定效应 (DE Effect), 即中国学习者所使用的存在句主语基本上都具有不确指性特征, 而且还表明中国二语学习者确指性主语的使用频率比 L1 使用者的还要低得多。根据语言迁移理论, 这种现象可能因为汉语的“有”字句中存在主体是不确指的^[18], 而“有”字句往往对应于英语的 There Be 句。还有可能因为二语学习者所接受的输入语料中的存在句大都 (90% 以上) 为非确指性主语, 因而难以判断何种情况下英语存在句的主语是确指性的, 就对确指性主语采取保守的回避策略。

四、结论

从 L1 使用者语料库看, “确指性限制”总体看来并不存在, 但人称代词表现出绝对的限制性, 且存在句使用确指性主语的现象并不常见。从中国二语学习者语料库看, 中国学习者一方面使用存在句的频率稍微偏高, 但与 L1 使用者并无显著差异。另一方面, 中国二语学习者存在句确指性主语的使用频率是 L1 使用者的三分之一, P 值 = 0.00 表现出绝对的显著性差异。这不仅证明“限定性效应” (DE) 是存在的, 而且说明中国学习者使用确指性主语的频率比 L1 使用者低得多, 这种现象可能是由于汉语“有”字句的迁移和学习者的回避策略导致的。

注释

- ① 这里词频密度即类符频率, 它反映某个类符在特定语料库中出现的密度, 其计算公式为: 类符频率 = 该类符观察频数 ÷ 语料库形符总量。^[31]
- ② 双尾 t 检验的显著性概率即 P 值, 当 $P < 0.05$ 时, 说明观察数据与参照的数据之间存在显著性差异, 当 $P > 0.05$ 时, 说明不存在显著性差异。^[32]

参考文献

- [1] Erdmann, P. Discourse and Grammar[M]. Munchen: Tuduv, 1999.
- [2] Jespersen, O. The Philosophy of Grammar[M]. New York: Norton, 1924.
- [3] Quirk et al. A Grammar of Contemporary English[M]. London: Longman, 1972.
- [4] Mikami, G. Toward an Explanation of Certain Peculiarities of the Existential Construction in English[J]. Linguistic Analysis, 1977, 3: 1-29.
- [5] Hornby, A. S. Guide to Patterns and Usage in English[M]. London: Oxford University Press, 1978.
- [6] Quirk et al. A Comprehensive Grammar of the English Language[M]. London: Longman, 1985.
- [7] Safir, K. J. Syntactic Chains[M]. Cambridge: CUP, 1985.
- [8] Nathan, G. S. The Syntax and Semantics of the Existential Construction[M]. Hawaii: University of Hawaii, 1978.
- [9] Brevik, L. E. Existential There: A Synchronic and Diachronic Study[M]. Bergen: University of Bergen, 1983.
- [10] Hornback, H. An Interpretive Solution to the Definiteness Effect Problem[J]. Linguistic Analysis, 1984, 13: 195-215.
- [11] Hannay, M. English Existentials in Functional Grammar[M]. Dordrecht: Foris, 1985.
- [12] Abbot, B. Some Remarks on Specificity[J]. Linguistic Inquiry, 1995, 26: 341-347.
- [13] Ward, G., B. J. Bimer. Response to Abbot[J]. Language, 1997, 73: 109-112.
- [14] White, L. The Definiteness Effect in L2 Acquisition[R]. Ningbo: Ningbo University, 2010.
- [15] Lyons, C. Definiteness[M]. Cambridge: CUP, 1999.
- [16] Yin, V. Interlanguage and Learnability: From Chinese to English[M]. Amsterdam: John Benjamins Publishing Company, 1995.
- [17] Chou Minchih(周敏洁). Chinese Learners' Overgeneralization of English Existential Constructions[J]. Studies in Linguistics.

2004 30(2): 183- 214

- [18] Huang C- T. (黄正德). Existential Sentences in Chinese and (In) definiteness[A]. In E. Eeklund & A. Meulen(eds). The Representation of (In) definiteness[C]. Cambridge Mass MIT Press 1987 226- 253.
- [19] 章振邦. 新编英语语法 [M]. 上海: 上海译文出版社, 1989
- [20] 肖俊洪. 信息结构与 There- 存在句“实义主语”的确指性 [J]. 外语学刊, 1994(2): 22- 27
- [21] 谷化琳. 英语存在句的信息传递探索 [J]. 外国语, 1998(6): 32- 37
- [22] 何自然, 陈新仁. 英语语用语法 [M]. 北京: 外语教学与研究出版社, 2003
- [23] 肖俊洪. 存在结构 THERE + BE 刍议 [J]. 现代外语, 1985(2): 72- 73
- [24] 王玉环. 浅析英语 (there) 存在句的主要结构特征及句型转换 [J]. 浙江师范大学学报, 1996(2): 71- 75.
- [25] 戴曼纯崔刚. 英语存在句研究二题 [J]. 外语与外语教学, 2000(9): 24- 27
- [26] 余国良. 英汉存在句的比较研究 [J]. 四川外国语学院学报, 2000(2): 92- 94
- [27] 张绍杰, 于飞. 英语存在句“确指性限制”的语用解释 [J]. 外语学刊, 2005(2): 2- 7.
- [28] 李金玉. 英语 there be 存在句习得的中介语研究 [D]. 北京: 首都师范大学, 2005
- [29] 黄冰, 何安平. 中国大学生 There BE 句型习得初探 [J]. 现代外语 (1), 2007
- [30] 宋阳. 基于语料库的中国英语学习者英文写作中 There- be 结构习得的研究 [J]. 安徽广播电视大学学报, 2008(3).
- [31] 杨惠忠. 语料库语言学导论 [M]. 上海: 上海外语教育出版社, 2002 168
- [32] 秦晓晴. 外语教学研究中的定量数据分析 [M]. 武汉: 华中科技大学出版社, 2003 144- 155

(上接第 47 页)

独立地发挥其交际工具的作用, 而隐语行话的形成和使用, 则正好弥补了这一不足。此外, 两者的互补性还表现在: 隐语行话在民族共同语的基础上生成, 民族共同语为隐语行话的生成提供了音、形、义等物质材料; 而隐语行话呢? 它与共同语之间并无不可逾越的鸿沟, 一些隐语行话在某一秘密群体习相沿用之后, 久而久之, 会渗透到其它特定的社会群体, 或进入特定时代的文学戏曲语言, 并以此为过渡, 逐渐进入民族共同语或地域方言。隐语行话这种通用化传承, 在局部上促进了民族共同语的发展。

综上所述, 我们认为: 第一, 隐语行话生成于特定的社会群体, 运用于特定的社会群体。在秘密社会群体内部, 隐语行话具有交际工具的一切价值和功能, 反之, 离开了特定的社会群体, 便会失去其价值与功能。因此, 社会群体性是隐语行话赖以存在

的社会基础。第二, 隐语行话的构成元素全部取自于民族共同语, 隐语行话只有在民族共同语的基础之上才能有效运用。离开了民族共同语, 隐语行话的构成与运用便成了无本之木, 无源之水。因此, 民族共同语特有的民族性是隐语行话得以生成及有效运行的物质基础。

参考文献

- [1] 荀子· 正名 [M]. 上海古籍出版社, 1986
- [2] 潘家懿, 赵宏因. 一个特殊的隐语区 —— 夏县东 浒 “ 延话 ” 隐语) 调查纪实 [J]. 语文研究, 1986(3).
- [3] 郑杭生. 社会学概论新修 [M]. 中国人民大学出版社, 1997
- [4] AD 什维策尔. 现代社会语文学 [M]. 北京大学出版社, 1987.
- [5] 赵元任. 反切语八种 [A]. 国立中央研究院历史语言研究所集刊 (第二节) [C]. 1931.