

基于语料库的英语语法化研究

向朋友

(北京航空航天大学 外国语学院, 北京 100191)

摘要: 基于语料库的语法化研究在西方现已盛行, 根据 20 世纪 90 年代中期以来出版的相关专论, 概述语料库对语法化研究的贡献以及两者的相互促进和影响, 以唤起中国学者对此方面研究的关注。

关键词: 语料库; 语法化; 语言学

中图分类号: H314 文献标识码: B 文章编号: 1008-2204(2010)01-0087-06

A Corpus-based Approach to English Grammaticalization

XIANG Ming-you

(School of Foreign languages, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

Abstract: Corpus-based researches on English Grammaticalization have been proved successful in the west. Based on the books and papers on English Grammaticalization published recently, this paper introduces the explanatory power of corpus in grammaticalization studies with intention to arouse Chinese scholars' interests in corpus-based approach to grammaticalization.

Key words: corpus-based; grammaticalization; linguistics

Mair^{[1]121} 描述过这样一个现象: 20 世纪 70 年代末、80 年代初, 沉寂一时的语料库语言学和语法化理论不约而同地东山再起, 双双成为现代语言学研究的热门话题。遗憾的是它们平行发展十几年, 直到 20 世纪 90 年代中期才有交叉。语料库在更灵活生动地描述某些语法现象的语法化过程, 验证及进一步探究已有语法化理论方面彰显出巨大的解释力和权威性。笔者拟根据 20 世纪 90 年代中期以来出版的 4 本有关英语语法化研究的专著^[2-5]及部分相关论文, 在了解相关语料库的基础上, 从语法化现象描述, 语法化理论验证、拓展等方面简述语料库在英语语法化研究中的巨大作为。

一、与语法化研究相关的语料库

欧、美诸国当前广泛应用于英语语法化研究的语料库有赫尔辛基英语文本语料库 (Helsinki Corpus of English Texts)、布朗语料库 (Brown Corpus)、兰卡斯特 - 奥斯陆\卑尔根语料库 (Lancaster\Oslo-Bergen Corpus, 简称 LOB 语料库)、ARCHER 语料库 (A Representative Corpus of Historical English Regis-

ters)、约克英语语料库 (York Corpus)、牛津英语词典语料库 (OED Corpus)、多伦多古英语语料库 (Toronto Corpus of Old English)、世纪语料库 (Century Corpus)、不列颠国家语料库 (The British National Corpus)、Gutenberg 语料库 (Project Gutenberg)、COLT 语料库 (Bergen Corpus of London Teenage Language)、MICASE 语料库 (Michigan Corpus of Academic Spoken English)、Frown 语料库 (Freiburg-Brown Corpus of American English) 和 FLOB 语料库 (Freiburg-LOB Corpus of British English) 等。在论述语料库对语法化研究的贡献之前, 有必要简单了解几个与语法化研究密切相关的知名语料库: 布朗语料库, LOB 语料库, 赫尔辛基英语文本语料库及 ARCHER 语料库。

1961 年至 1964 年间, 美国布朗 (Brown) 大学的 Nelson Francis 和 Henry Kučera 建立了可机读的布朗语料库, 词库容量为 100 万词次, 包括 15 种文体的 500 个书面英语文本, 这些文本分为标注文本和非标注文本。这是世界上第一个根据系统性原则采集样本的平衡结构语料库。该语料库能充分反映语言的共时性。布朗语料库是一个代表当代美国英语

的语料库。由英国兰卡斯特 (Lancaster) 大学的 Geoffrey Leech 教授倡议, 挪威奥斯陆 (Oslo) 大学的 Stig Johansson 教授主持完成, 最后在挪威卑尔根 (Bergen) 大学的挪威人文科学计算中心联合建立的 LOB 语料库, 旨在与布朗语料库进行平行对比研究, 其规模与布朗语料库相当。该语料库始建于 1970 年, 于 1978 年完成, 这是一个代表当代英国英语的语料库, 采用 TAGIT 标注系统, 以统计的方式建立换算机率矩阵, 大大提高了标注正确率。与布朗语料库不同的是它能对英语的不同变体进行分析。

布朗语料库和 LOB 语料库收录的分别是 1961 年出版的 500 个美国书面英语文本和 500 个英国书面英语文本。1991 年, 在 Christian Mair 教授的倡导垂范下, 德国弗莱堡 (Freiburg) 大学的一帮学者以能反映 20 世纪 90 年代早期英、美英语特征的各类英语书面文本为语料, 比照布朗语料库和 LOB 语料库原有模式, 按 15 种不同文体, 分别精心挑选了每个长度在 2 000 字左右的文本 500 个, 便形成了前面提到的 Frown 语料库和 FLOB 语料库, 其库容量同样都在 100 万词次。这两个语料库不仅能够满足学者们关于现代英、美英语变化的研究, 更能帮助学者们进行英、美英语与印度、澳大利亚和新西兰英语的对比研究。这两个语料库在英语语法化研究中也发挥着重要作用。

赫尔辛基英语文本语料库由 Matti Rissanen 教授和 Ossi Ihlainen 教授等一批语言学家牵头, 从 1984 年着手在芬兰赫尔辛基大学英语系建成。该库涵盖公元 750 年至 1700 年间各类英语语篇, 语料按古英语、中古英语和现代英语区分, 并以每 100 年分段, 词库容量达 1 600 千万词次。作为第一个大型英语历时语料库, 其在历史语言学及语法化研究中发挥巨大作用。涵盖公元 1650 年 ~ 1990 年间英、美各类文体的 ARCHER 语料库由 Douglas Biber 教授和 Edward Finegan 教授在 20 世纪 90 年代发起创建。作为开放的历时语料库, 其语料还在不断扩容, 并由美国、芬兰、瑞典、瑞士、德国、英国和西班牙等 7 个国家的 13 所大学共同协调管理。该库的 3.1 版本语料已收至 2006 年的各类文本, 3.2 版本已扩充, 加注到 2008 年 11 月的语料。该语料库对研究英语各类变化意义重大, 在英语语法化研究中发挥着不可替代的作用。

二、语法化现象的描述

没有语料库支撑的语法化研究多以零星个例为

依据, 以定性分析为主, 其结论以假设居多。语料库的引入, 使语法化研究不仅可以定性和定量相结合, 探究现有语法现象的形成过程, 还能探究在语法体系形成过程中消亡的语法现象。基于语料库的英语语法化现象描述主要集中在对英语代词、副词、情态动词等语法化过程的研究上。Rissanen 等^[2] 著作的第一部基于语料库的英语语法化研究专论, 着重描述相关代词和副词的语法化过程。Palander-Collin^[3] 可以看成是 Rissanen 等部分研究的延续。Krug^[4] 集中探究英语情态动词的语法化问题。下面笔者将按专题简要介绍这 3 部专论的相关内容。

(一) 代词的语法化

Rissanen 等中提及与英语代词语法化研究相关的论文有 2 篇: Brunberg 和 Tarkka^[6], Rissanen^[7]。

Brunberg 和 Tarkka 探讨英语单数不定代词的语法化问题。他们的研究主要基于芬兰赫尔辛基英语文本语料库, 并辅以多伦多古英语语料库、莎士比亚全集电子版、Milic^[8] 世纪语料库以及圣经现代英语译本。为方便计算机识别, 该论文把研究对象集中在从古英语阶段延续到早期现代英语阶段, 包含 some, any, every, each 和 no 等不定成分的单数不定代词上, 排除诸如 whoever (古英语的 SWA HWA SWA), 与之对应的古英语 SE DE 结构和之后的 he that/who 及 one another 等相互代词。HWA 或 HWILC 这类疑问代词虽被用作不定代词, 但他们在古英语的较早时期就已消亡, 故也未被列入考查对象。论文从不定代词的性质, 形式和相应的分布时段三个视角把研究对象进行列表归类。在上表归类的基础上, Brunberg 和 Tarkka 把不定代词的历史分布时段进一步细分为古英语 (~1150), 中古英语 1 ~ 2 期 (1150 ~ 1350), 中古英语 3 ~ 4 期 (1350 ~ 1500), 早期现代英语 1 ~ 2 期 (1500 ~ 1640) 和早期现代英语 3 期 (1640 ~ 710) 等 5 个阶段。从语料库中取出与上述 5 个不同时段对应的各 10 万字容量的语料, 计算出各断定 (assertive) 代词、非断定 (non-assertive) 代词、否定 (negative) 代词及全称 (universal) 代词的出现频率^{[6]31}, 再以断定代词、非断定代词、否定代词和全称代词为主线, 以库内机读语料为依托细致描述 4 类不定代词各表现形式出现的具体语境、频率、语义、句法特征及相关变体的演化过程。除却以上系统分析外, 该论文还借助语料库的便利, 对上述各不定代词在不同方言或体裁文本中出现的频率和语境作了翔实描述。在翔实语料支持下, 研究者不断指证现代英语中简单式不定代词 EVERY 和 EACH 与古英语的 GEHWA AEGH-

WILC 及早期中古英语的 AEFRE 和 AELC 等 5 个复合词一脉相承,由它们语法化而成,而且还生动地描绘了由 MAN, BODY, ONE 构成的复合不定代词的语法化过程。该研究不仅通过对不定代词语法化过程及其相关特征的探究充分支持了把英语划分为古英语、中古英语和现代英语 3 个阶段的合理性,还证实在包含 BODY, ONE 的复合不定代词的语法化过程中起关键作用的是对称性而不是人们所接受的经济性。

Rissanen et al 中与代词语法化相关的第 2 篇论文 Rissanen 主要研究英语数词 ONE 转化为代词的语法化过程。此项研究主要依托芬兰赫尔辛基英语文本语料库。研究的跨度从古英语一直延伸到现代英语,但研究重点集中在中古英语及早期现代英语阶段。古英语与现代英语 ONE 对应的是数词 AN。受说话人主观态度影响,数词 AN (ONE) 与其它更大数字相比,更易于体现整体性或一体性 (oneness)。该论文基于语料库充分的语料,印证 AN (ONE) 在古英语时代,受主观化因素影响,逐步形成个指 (individualizing)、类指或泛指 (generic\generalizing) 功能。随着其数词意义逐渐虚空 (bleaching),经过漫长语法化过程,到古英语后期,AN 便获取代词的句法功能。到中古英语及现代英语时期,代词 AN 的形式被 ONE 替代,并从意义上粗略地朝着 6 个方向不断延续着其语法化进程,形成 6 类代词 ONE: 人称特指词 (personal-specific)、非人称特指词 (personal-non-specific)、人称类指词 (personal-generic)、替代词 (substitutive)、与不定代词并用代词 (with indefinite pronouns) 以及代替词 (prop-word)。^{[7]88-89} 在语料库支持下,该论文深入探究了中古英语及现代英语时期上述 6 类代词 ONE 语法化的过程后,得出以下结论: 数词 ONE 的个指及类指功能的形成是典型的渐进性语法化,在中古英语及早期现代英语语篇里可以得到充分印证; 在 ONE 的语法化过程中,主观化发挥关键作用; ONE 的代词化与古英语曲折词缀的拉平 (levelling) 或消失,继而导致其各系统崩溃密切相关; 现代英语语料证实代词 ONE 仍保留类指词、替代词、代替词和后代词 (post-pronominal) 等用法,而人称特指词和非人称特指词用法分别被 someone 和 anyone 所替代。

(二) 英语副词的语法化研究

英语副词从构词角度看可分为复杂、复合和简单副词 3 类,复杂类指 easily, frankly 这类由形容词加副词词缀-ly 等构成的副词;复合类指 therefore, therein, thereabouts 这类由已有副词与相关介词组合

构成的副词;简单类则指 fast, early 这类不需添加任何成分的根词副词。从句法功能角度看英语副词可分为专门修饰动词的动词副词和修饰整个句子的句子副词。句子副词多承载说话人的主观评注。有关学者甚至把 I think, You know, methinks 这类话语标记也看成句子副词。现代英语以由形容词添加-ly 派生的复杂类副词和零派生简单类副词居多。Nevälainen^[9] 依托赫尔辛基英语语料库,以公元 1350 年~1700 年阶段(界于晚期中古英语阶段和早期现代英语阶段之间)的英语为基本语料,探究现代英语中构词能力极强的副词词缀 LY 的语法化过程以及构词能力极弱的零派生简单类副词的语法化过程,指出 LY 系古英语形容词后缀 LIC 和副词后缀 E 经由重新分析而构成的 LYCHE 的缩略形式。早期现代英语阶段之前英语副词多由形容词的零派生构成,历史进入早期现代英语阶段以后 LY 派生而来的副词占据主导地位。赫尔辛基语料库的中零派生简单类副词数量从晚期中古英语阶段的 21% 下降到早期现代英语阶段的 13%,到现代英语阶段几乎失去了派生能力,取而代之的是 LY 的崛起。

复合副词大都经历进一步语法化的过程。sterman^[10] 依托赫尔辛基英语语料库,布朗语料库和 LOB 语料库,探究了由副词 there 与 in, on, at, after, about 等介词构成的复合副词从古英语阶段直到现代英语阶段由兴起、巅峰到衰落整个语法化的过程,并详细阐述了其兴起或衰落的个中缘由,指出 there 复合副词的兴衰充分反映了英语语言类型学的演变,即由综合形式转向分析形式。I think, methinks 原是话语的语用标记,有些学者把它们看成句子副词。Palander-Collin^{[3][11]} 依托赫尔辛基英语语料库和早期英语书信语料库 (Corpus of Early English Correspondence),基于语用学、社会语言学及语法化的相关理念,通过对英语非人称动词 (impersonal verb) think 自中古英语以来不断语法化,逐渐演化为展示说话人的言据性 (evidentiality) 的人称动词 (personal verb),并与第一人称代词组成 methinks, I think 这类句子副词过程的系统分析,进一步验证了 Hopper^[12] 的语法化原则,明确指出语法化除受到语言系统内部因素的影响外,还受语域 (register)、文体、地域方言、言语交际者的性别、年龄、社会地位等社会因素的左右。Lorenz^[13] 依据语料库,从去词汇化视角入手,探究 really 这类英语强调词的语法化过程及其未来之前景。

(三) 英语情态动词的语法化研究

国外对英语情态助动词的研究可归结为一般特

性探究和类型学研究两类。具体到英语核心情态动词,相关研究主要集中在探究其现代语义特征和历史演变两个方面。Coates^[14]、Palmer^[15]及 Sweets-er^[16]主要探究英语核心情态动词的语义特征。集中探究英语核心情态动词历史演变的代表性著作有 Goossens^[17-18]、Lightfoot^[19-20]、Plank^[21]、Perkins^[22]、Kytö^[23-24]、Denison^[25]和 Fischer^[26]等。

Krug 一改语法化研究中靠理论或凭直觉的传统,以系统详实的历史和当代英语语料为基础,定量、定性研究兼顾,归纳、演绎方法并用,对部分英语准情态动词的语法化过程展开了深入、系统的探究。其研究依托的历史语料来自赫尔辛基英语语料库和 ARCHER 语料库;其现代英语语料源于 LOB, FLOB, Brown 和 Frown 规模都在百万字以上的英、美书面语语料库,代表英国英语、兼收书面语和口语的不列颠国家语料库以及收录《卫报》等英文报纸和莎士比亚全集等文学名著的光盘。Krug 集中探究了 HAVE GOT TO/ GOTTA, HAVE TO/ HAFTA, WANT TO/WANNA, 以及边缘情态动词 DARE(TO), NEED TO 和 OUGHT TO 等语法化的动态过程。整个研究涉及句法重新分析、音位侵蚀、附着形式化(cliticization)、话语频率、相似性、时态中性、情态语义源起、情态动词从强制义到认识义或将来义的扩展、语义泛化、主观化及其他形式的语用推理等十多个方面,证实所研究的准情态动词正经历着语法化,并不断获得核心情态动词的特性。该研究凸现了话语频率和 Zipf 经济原则在语法化过程中的推波助澜作用,进一步验证了 Heine 关于词汇语用频率与语音侵蚀的历时序列关系的论述,即 Desemanticization → High frequency → Loss of information value → Erosion(去语义化 → 高频率 → 信息价值丢失 → 语音销蚀)

Krug 最后指出英语 got- 结构已取代核心情态动词,成为强制义情态动词的标记。

(四) 英语介词的语法化研究

英语介词语法化问题 Bybee 等^[27]都有涉及,但对复合介词结构语法化的研究并不多见。Hoffmann^[28]依托不列颠国家语料库, Gutenberg 语料库、牛津英语大词典和中古英语大词典,归纳出 126 个 PNP(preposition + noun + preposition)结构的复合介词,这些复合介词包括 in place of, by way of, on behalf of, in common with, in front of, in line with 等。在作好必要的理论铺垫和语料选取说明后,他从社会、功能角度对上述复合介词尤其是词频排在前 30 位的复合介词在现代英语中的特征及其语法化过程进

行了共时、历时并举,量化统计与定性分析结合的系统探究。之前 Seppänen 等坚称 PNP 结构中 PN 与 P 是分离的。Hoffmann 以理服人,有理有据地批驳了 Seppänen 等人的观点,证明了 PNP 结构作为复合介词的整体性。他指出在高词频的作用下,大量 PNP 结构不断从实义短语虚化为介词,逐步完成其语法化过程,但在类推的作用下,一些低频 PNP 结构也同样能够完成其自身的语法化进程。具有介词、副词和从属连词身份的 beside(s)在现代英语中使用频率较高。Rissanen^[29]借助赫尔辛基英语语料库和 ARCHER 语料库、辅以不列颠国家语料库、LOB、FLOB、Brown 和 Frown 语料库,早期现代英语语料库、英语散文语料库、古英语和中古英语语料库以及古英语书信样本语料库,系统考察了 beside(s)的缘起和整个语法化过程。他指出 beside(s)缘起于古英语的 be + sidan, 在中古英语时期开始语法化为介词,随后又很快语法化为连词。在其语法化过程中,语义空灵和主观化发挥巨大作用。

(五) 英语语用标记的语法化研究

语用标记,也有人称之为话语标记,最早为话语分析学者和语用学者所关注,近来又被看成语用法语化的典型案例,引来不少学者的关注。Brinton^[30]依据大量历史语篇,从语用、功能及语法化视角出发,在阐明其篇章、人际功能及句法环境的基础上,对中古英语语用标记 gan、中古和早期现代英语语用标记 anon、古英语语用标记 hwæt、古英语 gel-amp - 结构、中古英语 bifel - 结构、whan - 小句、语用标记 I gesse, God woot, Trusteth me wel 等的语法化过程和可能动因做了较详细的阐述,应证了语用法语化的假说,强调了主观化的作用。Thompson 和 Mulac^[31]对 I guess, I think 的语法化过程做过历时的量化研究,其结论与 Brinton 关于 I gesse 的探究结论出入较大。Fischer^[32]借助语料库,在比较了 Thompson, Mulac 和 Brinton 的相关结论,又对比了 I guess 和 I think 在荷兰语的对应式的语法化情况的基础上,运用 Hopper^[12]提出的语法化原则和 Lehmann^[33]提出的语法化参数分别测试了上述两结论,指出 I guess 和 I think 经历的不是语法化而是词汇化,它们在词汇化过程中失去了指称意义,仅保留认识和评价意义。Defour^[34]依托赫尔辛基英语语料库和早期英语书信语料库完成了对 well 等英语语用标记词语法化过程的探究,充分凸现了主观化在这些语用标记词的语义空灵、语用功能加强过程中所发挥的作用。Margerie^[35]依托 COLT 语料库、MICASE 语料库和 Santa Barbara 语料库对英国英语和

美国英语中的语用标记 kind of 和 kinda 的语法化过程做了较系统的探究,特别强调了主观化在这一语法化过程中的作用。Bogaert^[36]在语料库的基础上对英语语用标记 do you think, do you suppose 及 do you believe 等的语法化作了详尽的共时阐述。

三、语法化理论的验证和语料库语言学的拓展

通过上述介绍,不难看出语料库的巨大解释力主要体现在语法化个案研究上。然而,依托语料库等先进手段,语法化学者不仅能够较客观地描述和归纳现有语法现象形成的过程及规律,而且还能验证已有的研究结论,甚至拓展语法化的相关理论。同时,语法化研究也对语料库语言学提出新的挑战,对语料库语言学的发展不无裨益。Lindquist 和 Mair^[5]收录的相关论文不仅反映了语料库在语法化研究中的巨大解释力,而且着手探究两者相互促进及对语料库的客观认识问题。

笔者之前介绍过 Krug 依托赫尔辛基英语语料库、ARCHER 语料库、LOB 语料库、FLOB 语料库、Brown 语料库和 Frown 语料库以及不列颠国家语料库等提供的系统详实的历史和当代英语语料,对部分英语准情态动词的语法化过程展开了深入、系统的探究,并得出结论:英语 got-结构已取代核心情态动词,成为强制义情态动词的标记。但 Tagliamonte^[37]依托约克英语语料库,通过对英语“必须”和“强制”等不同表达法的研究,发现在保守的约克英语里,Krug 所指的取代变化根本就没发生。这就为日后研究 got to 和 gotta 何时、如何发生替代核心情态动词的变化提供了广阔、有趣的空间,同时也彰显了语料库用事实说话的分量及其对语法化理论的验证功能。

英语中进行体与 be 型被动态竞争、并存的历史并不长,但传统的、基于经验的语法化研究学者对此现象似乎已有定论。Hundt^[38]借助 ARCHER 语料库,对英语中的 be 型被动态(如 The house is being built)和表达被动态的主动进行体(如 The house is building)的竞争、并存过程进行系统的量化探究,在理清各自的语法化历程的同时,驳斥了前人就此语法现象得出的不符合事实的结论,雄辩地证明在两者竞争、并存的过程中,be 型被动态并没有将进行体排挤出局,两者不是处于过渡并存状态(transitional layering)而是处于稳定的并存状态(stable layering)。Hundt 还具体剖析了其中的缘由。Hundt

这项基于语料库的研究可视为语料库验证功能的又一次体现。

Hoffmann^[39]广泛依托不列颠国家语料库和 Gutenberg 语料库等提供的鲜活语料,以复合介词 a kind of 为个案,挑战语法化研究中关于频率与语法化关系的定论,指出低频复合介词同样可以借助与之相关的高频词,通过类推而实现语法化。其研究看似对传统语法化理论的挑战,实则是彰显语料库对语法化研究的贡献。

现代英语反身代词有的用于强调,有的用于对比。Peitsara^[40]排除这两种用法的反身代词,集中探讨与动词紧密相连,作为动词使用策略的反身代词从中古英语到现代英语的语法化问题。该研究依托赫尔辛基英语语料库,从文本类型、文体需求、语言使用习惯、语境因素、个人言语特点及语言规范动机等方面入手,通过定量和定性相结合的方法对此类反身代词从中古英语到现代英语的演变过程作了较为详尽的分析,指出该语法化现象的实现机制是重新分析,信息明晰原则和经济原则一直左右该语法化的全过程。此项研究不仅局部验证了相关的语法化理论,而且开启了代词化与相关动词补充成分关系的研究。

当众学者为语料库对语法化研究的解释力欢呼雀跃的时候,西方学者中也有不同的声音,他们依托语料库来证实语料库在语法化研究中的局限性。对语料库持审慎态度的学者有 Nevalainen 和 Mair。Nevalainen^[41]依托 LOB 语料库、赫尔辛基英语文本语料库及 ARCHER 语料库,以英语 prett(il)y 和 fair (ly)两词语中副词化词缀 LY 为个案,探究了语法化研究的传统视角与语料库语言学和社会语言学视角的融合问题,强调语料库实证性方法论的价值和历史语料对区分语言的文体差异和历时变异的作用。Nevalainen 教授不仅阐述了语料库对语法化研究传统视角形成的冲击,明确指出语法化研究对语料库语言学构成的挑战,而且从社会语言学角度对语法化传统研究和基于语料库的语法化研究提出质疑。她最后提出将 3 个互补的视角融合在一起,形成语法化研究的新视角的倡议。这对语法化研究和语料库语言学研究视角的拓展都是很有启发意义的。

Mair^[1]从探讨语料库与语法化研究的关系入手,以牛津英语词典为历时语料库,对语法化研究中基于语料库的纯频率统计方法的研究价值提出质疑,强调语法化研究中关注语言历时变化的必要性和基于交际理论的定性研究的重要性。

四、结语

以上综述充分说明了语料库在语法化研究中所发挥的巨大解释力,而语法化研究对语料库标注所提出的新要求又反过来促进着语料库语言学的发展。以上提及的相关研究同时也提醒人们语料库这一有利工具并非万能,语料库对某些语法现象仍无法做出精准描述。人们在依托语料库做系统量化研究的同时,也不能偏废在历时语言学研究中业已证明行之有效的定性分析。

基于语料库的语法化研究在西方现已盛行,但在中国此类研究尚未起步。与国外近年来基于大型语料库的语法化研究相比,中国语法化研究在语料占有量方面略显羞涩,语料技术处理方面手段落后,从而使国内的相关研究从某种意义上讲缺乏强有力的量化数据支撑。因此,充分运用和建设语料库,尤其是英、汉共时、历时平行语料库,将是深入进行英、汉语法化研究必不可少的基础条件。

参考文献:

- [1] MAIR C. Corpus linguistics and grammaticalization theory: statistics, frequencies and beyond [C] // Hans Lindquist, Christian Mair. *Corpus Approaches to Grammaticalization in English*. Amsterdam: John Benjamins, 2004.
- [2] Rissanen M, Mekja Kyto, Kirsi Heikkonen. Grammaticalization at work: studies of long-term developments in English [M]. Berlin: Mouton de Gruyter, 1997.
- [3] Palander-collin M. Grammaticalization and social embedding: I think and methink in middle and early modern English [M]. Helsinki: Mémoires de la Société Néophilologique de Helsinki, Tome LV, 1999.
- [4] Krug M. Emerging English modals: a corpus-based study of grammaticalization [M]. Berlin: Mouton de Gruyter, 2000.
- [5] Lindquist, Mair. Corpus approaches to grammaticalization in English [C]. Amsterdam: John Benjamins, 2004.
- [6] Brunberg H R, Tarkka L K. In definite pronouns with singular human reference [C] // Rissanen M, Mekja Kyto, Kirsi Heikkonen. *Grammaticalization at Work: Studies of Long-term Developments in English*. Berlin: Mouton de Gruyter, 1997.
- [7] Rissanen Matti. The pronominalization of one [C] // Rissanen M, Mekja Kyto, Kirsi Heikkonen. *Grammaticalization at Work: Studies of Long-term Developments in English*. Berlin: Mouton de Gruyter, 1997.
- [8] Milic Louis T. A new historical corpus [J]. *ICAME Journal* 1990, 14: 26-39.
- [9] Nevalainen Terttu. The processes of adverb derivation in Late Middle and Early Modern English [C] // Rissanen M, Mekja Kyto, Kirsi Heikkonen. *Grammaticalization at Work: Studies of Long-term Developments in English*. Berlin: Mouton de Gruyter, 1997.
- [10] Österman Aune. The compounds in the history of English [C] // Rissanen M, Mekja Kyto, Kirsi Heikkonen. *Grammaticalization at Work: Studies of Long-term Developments in English*. Berlin: Mouton de Gruyter, 1997.
- [11] Palander-collin M. A medieval case of grammaticalization, methinks [C] // Rissanen M, Mekja Kyto, Kirsi Heikkonen. *Grammaticalization at Work: Studies of Long-term Developments in English*. Berlin: Mouton de Gruyter, 1997.
- [12] Hooper Paul J. On some principles of grammaticalization [C] // Traugott E C, Heine B. *Approaches to Grammaticalization*. Amsterdam: John Benjamins, 1991: 1-14.
- [13] Lorenz Gunter. Really worthwhile or not really significant? A corpus-based approach to the delexicalization and grammaticalization of intensifiers in Modern English [C] // Ilse Wischer, Gabriele Diewald. *New Reflections on Grammaticalization*. Amsterdam: John Benjamins, 2002.
- [14] Coates Jennifer. The semantics of the modal auxiliaries [M]. London: Croom Helm, 1983.
- [15] Palmer Frank. Modality and the English modals [M]. London: Longman, 1989.
- [16] Sweetser E. From etymology to pragmatics [M]. Cambridge: Cambridge University Press, 1990.
- [17] Goossens Louis. On the development of modals and of the epistemic function in English [C] // Ahlgren A. *Papers from the 5th International Conference on Historical Linguistics*. Amsterdam: John Benjamins, 1982, 21: 74-84.
- [18] Goossens Louis. The auxiliarization of English modals [C] // Morris M, Ramat P. *Historical Development of Auxiliaries*. Berlin: Mouton de Gruyter, 1987: 111-143.
- [19] Lightfoot David W. Principles of diachronic syntax [M]. Cambridge: Cambridge University Press, 1979.
- [20] Lightfoot David W. How to set parameters: arguments from language change [M]. Cambridge: MIT Press, 1991.
- [21] Plank Frans. The modals story retold [J]. *Studies in Language*, 1984, 8: 305-366.
- [22] Perkins Micheal. Modal expressions in English [M]. London: Francis Pinter, 1983.
- [23] Kyösti Merja. On the use of the modal auxiliaries indicating possibility in early American English [C] // Martin Harris, Paolo Ramat. *Historical Development of Auxiliaries*. Berlin: Mouton de Gruyter, 1987.
- [24] Kyösti Merja. Variation and diachrony, with early American English in focus [M]. Frankfurt: Peter Lang, 1991.
- [25] Denison David. English historical syntax: verbal constructions [M]. London: Longman, 1993.
- [26] Fischer Olga. The development of the modals in English: Radical versus gradual changes [C] // David Hart. *English Modality in Context: Diachronic Perspectives*. Bern: Peter Lang, 2004: 16-32.
- [27] Bybee Joan, Revere Perkins, William Paolucci. The evolution of grammar: tense, aspect, and modality in the languages of the world [M]. Chicago: The University of Chicago Press, 1994.