

# 语料库在英语教学中的应用研究综述<sup>①</sup>

洪敏芳 王昕玥 李 平



**摘 要:**语料库已逐渐成为课堂中的重要教学工具。本研究综述语料库在英语教学中的应用研究,以中国知网上 2011—2020 年 320 篇核心期刊论文为研究对象,主要从语料库间接应用于教学、语料库直接应用于教学和语料库建设三个方面对论文进行细致的研究和分析。结果表明,该领域在过去 10 年取得了丰硕的研究成果,如研究领域更加广泛,人们对专门用途英语语料库的研究兴趣持续不断逐渐兴起等,很大程度上填补了前人研究空白。但同时也存在着一些不可忽视的问题,如理论研究成果在外语课堂实际应用研究不足,研究内容广泛而不均衡以及国内语料库的资源建设和共享不足等。

**关键词:**语料库;英语教学;文献综述

## A Review of Research on Corpora in English Teaching

HONG Minfang WANG Xinyue LI Ping

**Abstract:** Corpora has gradually become an important teaching tool in the classrooms. Based on previous studies, this study delivered a review on the research of corpora in English teaching and learning, taking 320 core journal papers on CNKI from 2011 to 2020 as the research object. The papers are meticulously studied and analyzed from the following three aspects: corpora indirectly applied to teaching, corpora directly applied to teaching and corpora construction. The results show that fruitful achievements have been made in this field in the past decade, including a wider range of research areas and the gradual rise of research on corpora for teaching English for special purposes, which has largely filled the gaps in previous research. However, there also existed some problems that cannot be ignored, such as insufficient application of theoretical research results into foreign language classrooms, extensive and uneven research, as well as insufficient construction and sharing of the corpus resources in

<sup>①</sup> 本文为江苏省研究生科研与实践创新计划项目(项目编号:SJCX21\_0217)的研究成果。

China.

**Keywords:** corpus; English teaching; literature review

## 1. 引言

Sinclair(1999)<sup>4</sup>把语料库定义为根据明确的语言标准选择和排序的语言片段集合,可作为某种语言的样本。语料库语言学则是以语料库提供的真实语言数据为研究对象,对大量的语言事实进行分析后总结出语言使用规律,这为语言研究提供了一种全新的研究思路。(杨惠中,2002)。语料库的优势之一在于其实证性质,它将大量真实的语言数据汇集在一起,让语言分析更加客观科学。随着信息技术的发展,语料库已被广泛运用于语言学研究的各个领域,如语用学、语义学、翻译与对比研究,二语习得和语言教学等,取得了不少的进展。数据驱动学习是一种基于语料库的外语学习方法,该方法一经提出,便受到广大研究者的关注,并被赞为“外语教学的一次变革”(Hunston,2002)<sup>1</sup>。

许多学者也曾对我国语料库和外语教学方面的研究进行了总结,如闫超亚(2010)和方秀才(2012)分别就2000—2009年和1981—2010年间发表在国内外语类核心期刊上有关语料库的英语教学与研究的论文做了统计分析,总结了这两个阶段研究的成就与不足。因此,本研究文献选取范围始于2011年初,止于2020年底。在两位学者的研究基础上,为了进一步了解国内语料库在英语教学领域的最新动态和发展趋势,笔者对2011—2020年发表在中文核心期刊的320篇研究论文做了统计分析,并针对研究现状所存在的问题提出建议,以期为今后国内语料库英语教学领域的发展提出建议和参考。

## 2. 研究目的与问题

通过对所选论文进行细致的分析归纳,以期了解我国语料库应用于英语教学研究的最新动态,本文回答以下研究问题:(1)在2011—2020年10年间语料库应用于英语教学研究的基本情况是怎样的?(2)语料库应用于英语教学研究有哪些主要内容及特征?(3)还存在哪些问题和不足?

## 3. 研究方法

### 3.1 数据搜集工具

本文主要以CiteSpace为主要工具,将近10年来语料库英语教学领域的研究文献导入软件进行图谱分析,并根据所得图像进行深入的精细化的分类统计。再使用教育文献法对文献进行深入的解读,了解该领域的发展态势。CiteSpace以知识域为对象,

以图像的形式显示科学知识的发展进程与结构关系,从而揭示科学结构的发展现状和趋势(陈悦,2015)。CiteSpace 的分析结果具有直观立体快速科学的特点,软件自开发以来就受到了众多学者的青睐,在国内也得到了广泛使用。

### 3.2 研究对象

笔者采用高级检索方式,以“语料库”“教学”以及“语料库语言学”“数据驱动”“课堂教学”“网络教学”等关键词在 CNKI 全文数据库进行检索,设置检索时间为 2011—2020 年,共 10 年,数据源设置为北大核心和 CSSCI。经过人工剔除会议信息、期刊简介、书评,会议论文,对外汉语教学和小语种相关研究,以及只涵盖语料库或者只涵盖教学的文献之后,共得到有效文献 320 条,本文将以此 320 条文献作为研究对象。

学术论文的发表数量是评价某个领域研究态势的重要指标,通过对发表文献进行统计,可以得出该领域的发展趋势。所选文献年度发文数量绘制的统计如图 1 所示,总体来看,国内核心期刊在此领域上的发文量并不高,年均发文量 32.8 篇,发文量最高的是 2011 年和 2015 年。除此之外,笔者在检索时发现这些文章部分来自于外语类的核心期刊,如外语电化教学和外语界等,但还有相当一部分来自于非外语类核心期刊,如中国成人教育、电化教育研究和中国电化教育等,这一现象体现了语料库教学的跨学科的本质,它以语料库为基本知识源,研究自然语言规律,从而应用到语言教学中去。

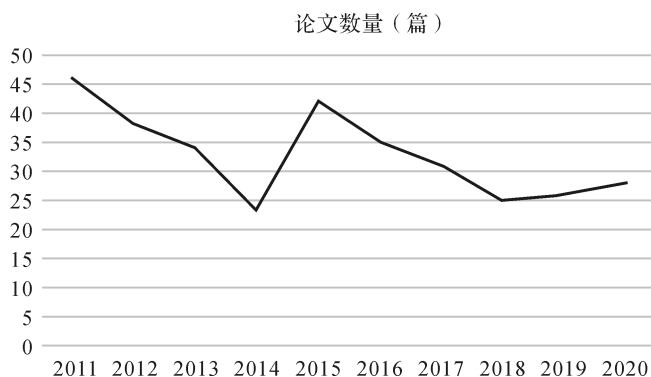


图 1 年度发文数量统计

### 3.3 具体分类情况

Leech(1997)将语料库在英语教育教学中的应用分为以下三个方面:第一部分是间接应用,主要是指教学资源 and 语言测试等;第二部分是直接应用,主要是指如何用语料库来进行课堂教学以及教学生如何使用语料库等;第三部分是面向教学的语料库建设,如英语学习者语料库和专门用途语言语料库等。

语料库间接用于教学主要指把语料库和教学理论相结合研究,从而探讨出更好的教学资源 and 模式,用以指导教学实践。语料库在外语教学的直接应用首先是外语教学研究,偏向实际应用,所以这部分大都为实证研究,用以检测语料库在英语教学各个方

面的辅助是否有效,此分类下涵盖英语教学的听说读写等各个方面。第三类则是前两类的基础,语料库在前两个重点中起着强大的辅助作用,只有把语料库建设好,才能更好地促进实际课堂语言教学。通过以上分析,本文主要参照 Leech 的分类情况,结合所得论文实际情况,将统计内容进一步分为四大类:第一类是语料库间接应用于教学,主要是理论研究;第二类是语料库直接应用于教学,主要为实证研究;第三类是面向教学语料库的建设和利用;最后一类主要是一些会议评述等纯理论。每个大类下面都有次类,情况详如表 1 所示。

表 1 论文分类统计情况

分类	分类范围		篇数		比例(%)	
1	语料库间接应用于教学(理论)	教学大纲、教学资源、教师	133	18	41.60	5.60
		教学模式研究		39		12.19
		词汇教学		19		5.94
		语法教学		4		1.25
		写作教学		11		3.43
		阅读教学		8		2.50
		听力教学		2		0.63
		口语教学		1		0.31
		翻译教学		31		9.69
2	语料库直接应用于教学(实践)	课堂话语研究	93	7	29.70	2.19
		教学模式研究		9		2.50
		写作教学		19		5.93
		口语教学		12		3.75
		语法教学		7		2.19
		阅读教学		1		0.31
		词汇教学		32		10.31
		听力教学		1		0.31
		翻译教学		8		2.50
3	面向教学的语料库建设	英语语料库的建设和利用	57	42	17.80	13.13
		ESP 语料库		14		4.38
4	其他	理论和意义分析	35	15	10.90	4.69
		综述		20		6.25

理论和实践研究主要包括对教学资源、总体教学模式和英语教学各个层面的教学模式,如词汇、口语、语法、写作、阅读、听力和翻译等方面。数据显示,从研究内容上看,近十年来语料库外语教学研究涉及内容广泛,不同方面文章数量差别很大。有关理论

研究的有 133 篇,占论文总量的 41.6%,其中单独对课堂整体教学模式的研究就有 40 篇,占总论文的 12.5%。实践研究的有 95 篇,占 29.7%,其中词汇教学占的比例最大,其次是写作和口语教学。在语料库建设方面,有 43 篇论文探讨通用英语语料库的建设和利用,以口译和笔译语料库为主。除此之外,更多的学者开始关注专门用途英语(ESP)语料库的建设,呈现出多元化的特点,ESP 语料库的建设主要有旅游英语,航海英语,海事英语,学术英语,商务英语,信息技术英语,新闻英语和财经英语等方面。在具体的英语模块教学理论中,被研究最多的是翻译教学,其次是词汇和写作教学。在具体的教学应用中,被研究最多的是词汇教学,其次是写作教学和口语教学。由此可见,教学模式研究、翻译教学、写作教学、词汇教学和口语教学是 2011—2020 年的教学研究热点。以上数据更细致地反映了这 10 年间语料库英语教学研究的现状与重点。

## 4. 研究结果与讨论

在选取的 320 篇论文中,研究对象呈现出广泛且多元化的特点,涉及成人教育学生、高职院校学生、英语专业的本科生和硕士生、非英语专业的本科生和硕士生以及外语教师等,涵盖了我国高等教育大部分办学层次。与前人对过去 30 年语料库外语教学研究的成果相对比,近 10 年的进展主要体现在研究领域更加广泛。在课堂教学方面,翻译教学依旧是热点,近 10 年来词汇口语和写作教学也开始受到重视。英语各方面的学习是一个有机整体,相互联系,拓宽研究范围可以更好地提高学生的英语综合素质。在语料库建设方面,除了通用的英语大型语料库建设外,专门用途英语教学语料库的研究逐渐兴起,以满足各特色专业院校人才培养的需要。

### 4.1 教学模式研究

在理论研究方面,针对教学模式的研究占了很大一部分,共有 40 篇,占了 12.5%。这些教学模式中有针对英语教学的通用模式,如王晓雯、葛诗利(2015)以广东外语外贸大学《影视英语》精品课程改革为例设计了数据驱动教学模式,并探讨其在教学过程中的具体实施方案。除此以外,学者们也在探索专门用途英语的教学模式,如高书娜(2016)从我国汽车专业英语教学模式的弊端出发,试图把语料库融入英语课堂教学中,提出了“数据驱动”的汽车专业英语教学模式。这些教学模式虽各有其特色,但都是以语料库为基础进行探索,强调激发学生个性化和自主化学习,让学生通过语料库这种真实的语料工具来发现语言的使用规律和特点,与传统的教师传授模式形成鲜明对比。值得关注的是,关于教学模式的实证研究却仅有 9 篇,如戈玲玲等(2014)提出了基于语料库的大学英语分析教学模式,并进行了实证研究。在表格中同样可以看出语料库在教学资源应用方面的研究依然不够。

## 4.2 听说教学

听力在语言四大技能中列于首位,是其他技能不断提升的基础。若将语言能力进行量化,我们听到的是说的2倍,是阅读的4倍,是写作的5倍(Morley,1991)。此外,听力理解是其他语言技能不断发展的基础。从表24.7中可以看出,无论是在理论还是实践层次,针对听力教学的研究数量还不够,只有2篇,口语教学的实证研究较多,达到了12篇。杨林伟、李婷(2015)采用行动研究方法开展多媒体新闻语料库自主学习研究,用问卷收集学习者对于这种教学平台的看法,分析结果表明,语料库对听力学习能起到很好的辅助的作用。

郑李卉和肖忠华(2015)以“中国学生英语口语语料库”为基础,研究中国英语学习者口语中常见的搭配错误,总结出了6种不同搭配错误的分布,并建议将语料库方法运用到大学英语课堂词汇教学中。李振营(2013)运用CIA(中介语对比分析法)对COLSEC和BNC这两个异国口语语料库进行词汇和词项特征的对比研究,结果表明与英语母语者相比,中国学生的英语口语词汇的使用不足和过度使用现象。

## 4.3 读写教学

英语教学中把阅读视为输入,写作视为输出,两端处于平衡状态才能达到最佳的学习效果。写作教学无论在理论研究还是实践研究中都是一个热点,但阅读教学研究受到的关注度却比较低,只有2篇,可见关于读写的研究处在一个失衡的状态。徐曼菲、何安平(2016)把语料库的关键词探究法和英语教学中的深度阅读相结合,在英语专业的精读课程中探讨了语料库辅助下的阅读教学模式,结果表明数据驱动的英语阅读教学有利于培养学生多种阅读微技能,提升学生的深度阅读能力。

董洪学、初胜华(2010)在分析我国大学英语写作教学现状后,以输入假设理论为基础,将多媒体技术和网络语料库相结合,提出了基于计算机网络语料库驱动的英语写作教学模式,并进行了实证研究,研究结果表明该架构能够显著提高学生的英语写作运用能力。王华(2020)提出将语料库驱动学习的理念及语料库技术应用到学术英语写作教学中,探索一种具有先进教学理念、又切实可行的学术英语写作教学模式,并应用于学术论文摘要写作教学。

## 4.4 语法词汇和翻译教学

翻译教学是一大热点,占到了12%,这在一定程度上也是机器和神经翻译快速发展所带来的结果。曾祥宏(2019)从MTI专业学位翻译教学的培养目标出发,针对传统翻译教学不足的分析,从教材编写、教学模式和翻译作业语料库建立这三个角度探讨如何将语料库有效应用于MTI翻译教学中。刘晓风、李德凤(2020)尝试将COCA单语语料库应用到英汉商务翻译教学中,提出了“一个中心、两个基本点和三种能力”的教学方法。词汇学习是我国英语学习者的一大拦路虎,主要体现在单词记不住以及记住了却

无法正确运用,这一现状也引起了学界的关注。词汇教学不论是在理论还是实践研究上都是一个热点,在所有的研究中占比 21%。闫长红(2018)在分析传统英语词汇教学课堂现状的基础上提出语料库的运用优势,并进一步分析了语料库在词汇教学中的应用原则和策略。孟超和马庆林(2019)针对英语专业学生法律词汇学学习情况构建了基于在线语料库的法律英语词汇教学模型,实证研究表明该模型能有效促进学生法律英语词汇的认知与习得。相比之下,语法教学则不太受重视,理论和实践研究加起来一共 11 篇。吴军(2010)提出可以把语料库运用到英语语法教学中,来消除我国传统语法教学的弊端,并指明了未来研究改革和努力的方向。

#### 4.5 语料库建设和利用

在语料库建设方面,除了通用英语语料库的建设,还有很多学者提出了 ESP 语料库的建设,主要包括旅游英语,航海英语,海事教学,学术英语,公安英语,商务英语,信息技术英语等。冯正斌、王峰(2016)从语料规模、语料采集、语料赋码方案等方面探讨了财经英语新闻语料库的初步建库理念,并探索语料库在专业英语词汇教学方面的应用。邓军涛、刘梦莲(2020)以欧盟委员会“虚拟现实口译”项目的视频语料资源库为例,从素材选取、语料切分、信息标注与场景模拟四个方面阐述如何深度开发面向口译教学的视频语料库。这说明国内语料库建设正在走向多元化,语料库作为工具提供的服务涵盖的范围越来越广泛,针对的对象也越来越精准。

### 5. 问题与启示

本文选取的北大核心和 CSSCI 期刊代表了学科发展的最新动态,在 2011—2020 年间,语料库在外语教学中的研究取得了一定的成就。黄萍、闫超亚(2010)在梳理 2000—2009 年的相关文献时曾指出学界对教学大纲设计、教材编写、教师、口语教学、翻译教学、学习模式的探讨关注得很少,文体学教学、阅读分析课教学等小众领域没有涉及。这一现象在过去 10 年已经得到改进,王伟(2013),刘雯(2015)和余剑虹(2017)等陆续将语料库引进诗歌教学和英美文学教学课堂,填补前人研究空白。通过对 2011—2020 年间的文献梳理,本文也发现了一些研究不足之处。

#### 5.1 外语课堂实际应用不足

在语料库间接应用于教学的理论研究方面的研究论文颇丰,占到了论文总量的 12.19%,相比之下,语料库辅助的课堂实际应用方面的研究较少,仅占 2.5%,理论成果的转化率还不够高。《课程要求》鼓励各高校充分运用现代化信息手段,改进过去以教师讲授为主的“填鸭式”教学,采用基于计算机的教学模式。学者们提出了众多教学模式,有广义上的多模态教学模式和翻转课堂模式等,也有如“一个中心、两个基本点和三种能力”这样的具体教学策略,但其中真正运用到实证研究中的教学模式却寥寥无

几。教育部在2007年颁布的《大学英语课程教学要求》中明确提出要努力开发以网络技术为支撑的新型教学模式,让英语的教与学朝着个性化和自主化的方向发展。从核心期刊文献数量来看,目前的实证研究成果远远不能满足我国各层次外语教学改革与发展的需要。从这个角度出发,充分利用语料库在外语课堂应用的理论成果,进行广泛的课堂实践应当成为今后的重点研究方向。只有经得住课堂实践考验的教学模式,才是真正适合中国英语实际教学的新时代教学方法。

## 5.2 研究内容广泛而不均衡

从表1可以看出,无论是在理论还是实践方面,针对语料库在外语教学中的研究领域很广,几乎涉及了外语学习者会接触到的方方面面,主要有学习模式的探讨、口语教学、写作教学、语法教学、词汇教学、翻译教学、阅读教学等。也对前人的研究领域进行拓宽,如教师课堂话语的研究,教学资源研究和语料库建设等。其中的热点主要是教学模式和翻译词汇教学研究,对中国学生最薄弱的口语和听力部分的研究占比依旧很低,分别只占到7.4%和1.79%,这反映出对英语实际运用研究的关注度还远远不够,口语是一个英语学习者综合能力的体现,“哑巴英语”依然是困扰中国英语学习者最大的桎梏。研究所涉及的领域大都为英语学习者的常见部分,但对于一些英语专业特有的课程如英语文体学和精读课程却没有涉及。

除此以外,对教师和教学资源(如课本和教学大纲),课堂话语等的理论较之前的研究有拓展,但关注度依旧不高。外语教学包括“教”与“学”,“教”的主体是老师,“学”的主体是学生,教师自身拥有过硬的语料库教学素养和语言意识对打造一堂好课至关重要。虽然针对语料库的课堂研究较多,但这些研究大都停留针对现有数据进行客观的描述和分析,没有深入分析受试者产生某方面外语学习障碍的根本原因。语料库语言学是外语教学的得力助手,在课堂上的教学能帮助学生暂时提高,只要找到学生产生这种现象的本质原因,辅以正确的学习方法和工具,才能真正达到授人以渔,让学生受益终身。

## 5.3 语料库利用率有待提升

语料库在外语课堂上的作用在于,训练学生通过语料分析主动去探索语言使用的型式和用法。虽然陆陆续续有不同高校的学者就自己所在学校的特色来进行ESP语料库建设,如财经英语新闻语料库,口译视频语料库和多媒体新闻语料库等,但是大都停留在建设构想层面,还未进行真正的建设;其次文献中涉及的语料库大都是国外的一些大型免费语料库,如BNC,COCA,BROWN等,国内的大型语料库,如CLEC(中国学习者语料库),CEME(中国英语专业语料库),PACCEL(汉英交替传译语料库)等十几个国内知名高校建立的大型语料库虽也有提到,但相比而言利用率并不高。还有相当一部分学者在进行研究时采用了自建语料库的形式,如自建课堂话语语料库,美英报刊英语标注语料库(NEC)和英语阅读教材语料库(PCRC)等。这些语料库的建造耗费了大



量的人力和物力,不加妥善利用,必然造成资源的浪费。虽然每个语料库是依照特定目的建立的,有着其自身独特的价值,但与国外相比,我国自建语料库的规模和内容无法满足学科发展所需。今后在资源建设和共享方面需要继续提高,只有充分吸收外来资源,加强国内语料库的建设与共享,才能更好地促进语料库教学研究的发展。

## 6. 结语

通过文献梳理分析,语料库在语言教学中应用研究在 2011—2020 年这个阶段趋于稳定,比过去 30 年的研究领域有所拓宽。所统计论文在总体上更侧重课堂教学理论创新,而忽视了外语教学实际。研究涉及内容广泛,但分布不均,教学模式研究、翻译教学、写作教学、词汇教学和口语教学是这一期间的研究热点。此外,国内语料库建设正在走向多元化,但现有语料库利用率有待提升。未来的研究方向应该朝英语能力各方面均衡发展,注重实证研究并加强 ESP 语料库的建设以及国内语料库资源的利用。本文对过去 10 年语料库外语教学的研究近况的概述和分析,为未来的研究方向提供了重要参考。

### 参考文献:

- 陈悦,陈超美,刘刚渊,等,2015. CiteSpace 知识图谱的方法论功能[J]. 科学研究,33(2): 242-253.
- 邓军涛,刘梦莲,2020. 面向口译教学的视频语料资源库深度开发机制研究[J]. 外语教育研究前沿,3(1):37-43.
- 董洪学,初胜华,2010. 基于计算机网络语料库驱动的英语写作教学模式研究[J]. 外语电化教学(5):27-32.
- 方秀才,2012. 基于语料库的英语教学与研究综述:成就与不足——根据 22 种语言学类 CSSCI 来源期刊近 30 年的统计分析[J]. 外语电化教学(3):19-24.
- 冯正斌,王峰,2016. 财经英语新闻语料库的建设构想与教学应用[J]. 外语电化教学(2):54-58.
- 高书娜,2016. 语料库“数据驱动”的汽车专业英语教学模式改革探索[J]. 西南师范大学学报(自然科学版),41(10):168-172.
- 戈玲玲,李广伟,刘彬,2014. 基于语料库的大学英语分级教学模式实证研究[J]. 外国语文,30(3):150-152.
- 黄萍,闫超亚,2010. 国内语料库应用于英语教学研究现状分析[J]. 山东外语教学,31(5):44-50.
- 李振营,2013. 中国大学生英语口语词汇特征研究[J]. 辞书研究(1):35-41.
- 刘雯,2015. 语言学教学中语料库文体学与狄更斯小说述介[J]. 中国教育学刊(1):47-48.
- 刘晓东,李德凤,2020. COCA 英语语料库在英汉商务翻译教学中的应用[J]. 中国科技

- 翻译,33(1):29-32,61.
- 孟超,马庆林,2019.基于在线语料库的法律英语词汇教学模型实证研究[J].外语电化教学(2):82-89.
- 王华,2020.语料库驱动的学术英语写作教学模式探索:以摘要写作为例[J].外语学刊(1):49-55.
- 王伟,2013.读者反应教学法与网络资源结合在英语诗歌教学中的应用[J].外语教学理论与实践(2):79-82.
- 王晓雯,葛诗利,2015.大学英语的数据驱动教学模式研究:以广东外语外贸大学《影视英语》精品课程改革为例[J].外语研究(01):47-52.
- 吴军,2010.语料库与英语语法教学[J].外语学刊(01):104-107.
- 徐曼菲,何安平,2016.语料库技术促进英语深度阅读教学[J].中国电化教育(12):87-90.
- 闫长红,2018.英语课堂教学对基于语料库的词汇学习研究[J].中国教育学刊(S1):94-96.
- 杨惠中,2002.语料库语言导论[M].上海:上海外语教育出版社.
- 杨林伟,李婷,2015.网络多媒体新闻语料库辅助英语听力教学研究[J].中国成人教育(16):178-179.
- 余剑虹,2017.语料库在英美文学教学中的辅助应用——评《哈克贝利·费恩历险记》[J].中国教育学刊(1):112.
- 曾祥宏,2019.语料库辅助的MTI翻译教学探索[J].上海翻译(1):76-79.
- 郑李卉,肖忠华,2015.中国英语学习者的口语搭配行为:基于语料库的研究[J].外语教学理论与实践(4):29-36.
- HUNSTON S, 2002. Corpora in Applied Linguistics [M]. Cambridge: Cambridge University Press;1.
- LEECH G, 1997. Teaching and Language Corpora: A Convergence [C]//WICHMANN A, et al, Teaching and Language Corpora. London: Longman:1-23.
- MORLEY J, 1991. Listening comprehension in second foreign language instruction [C]//CELCE-MURCIA. Teaching English as a second language. New York: Newbury House:81-105.

作者简介:

洪敏芳,女,1999年2月生,江西上饶人,南京农业大学外国语学院硕士研究生,主要从事语言学、翻译学研究。

王昕玥,女,1998年7月生,江苏南通人,南京农业大学外国语学院硕士研究生,主要从事翻译研究。

李平,男,1971年9月生,湖北黄冈人,翻译学博士,南京农业大学外国语学院教授,硕士研究生导师,主要从事语言学、翻译学研究。