

基于语料库的口译研究：回顾与展望^①

李 婧 李德超 香港理工大学

摘要：过去15年来，基于语料库的翻译研究影响日增。但由于口译语料库的缺乏，基于语料库的口译研究仍不多见。本文在评价当今较具代表性的国内外三个口译语料库的基础上，归纳口译语料库的特点和基本的建立方法，指出基于语料库口译研究的主要问题和前景。

关键词：语料库；语料库翻译研究；口译研究

中图分类号：H315.9

文献标识码：A

文章编号：1672-9382 (2010) 05-0100-06

1 引言

作为一种实证性研究方法，语料库翻译研究因其可以更系统、更科学地描写和归纳可见的翻译现象而在过去15年间得到长足的进展（参见王克非，黄立波，2008）。但迄今为止，大多数这方面的研究都限于笔译，尤其是文学翻译的研究，把语料库运用于口译的研究仍不多见。

最先提出把语料库运用于口译研究的是以色列学者Miriam Shlesinger（1998），她认为语料库口译研究可以参照语料库笔译研究的一些课题，如研究口译中的词汇密度、类符一形符比、词汇频率、语法结构、文本样式、共生现象等等，亦可以与笔译语料库相对比，来探索口译文本的独特性。Shlesinger提出上述观点后的12年间，语料库口译研究取得了哪些进展？现存的口译语料库主要有哪些？语料库口译研究的前景如何？本文拟对上述问题作一探讨，以促进学界对这个新领域的了解。

2 口译语料库的类型

与笔译语料库相类，口译语料库亦可分为类比语料库和对应语料库两种（参见Shlesinger，1998）。

2.1 口译类比语料库

在语料库翻译研究中，类比语料库一般指针对某一领域的同一语言中翻译文本与非翻译文本组成的语料库。若是双语类比语料库则包括两种语言的相似文本。按Shlesinger（1998）的观点，最理想的口译类比语料库

至少包括以下三个子库：口译译文文本库、在相似场合发表的与源语同类的口语文本库及与以上源语口语文本库相对应的笔译译文库。这样设计的目的主要有以下用途：第一，用于研究口语文本。口译译文作为一种特殊的口语文本，研究者可以把收集的口译译文切分成片段，与相同语言的其他自发性口语文本（spontaneous discourse）相比较。这样即使不考虑源语，亦能研究口译文本的特殊结构；第二，用于研究口译译文与笔译文本的区别。口译译文也可以与相对应的笔译译文比较，便于找出口译文本结构的其他特征。换言之，口译类比语料库可以从不同角度分别探索口译文本的独特性，亦能跨越类型、语言和口译个案特征，从而研究口译文本作为整体的共通性。这些特性和共性的研究，对某些还没有被证实的口译处理过程，如处理能力的限制等必然大有裨益（Gile，1991：15-27），同时亦更能帮助更系统地描述口译现象，如口译中的转移（shift）等（Shlesinger，1989）。

2.2 口译对应语料库

基于语料库的笔译研究，其传统研究对象是语言对的特定因素（language-pair-specific factors）和个人变量，如：性别、经历、语言背景等对翻译的影响（Laviosa，1998：474-479）。但口译语料库除了上述这些领域外，还可以拓展至研究译文的结构与形态间的相互关系，而这就可以借助于口译对应语料库。Shlesinger认为，口译对应语料库也要包括三个子库：源语文本库、相对应的口译译文文本库和相对应的笔译译文文本库。通过对相同源语和两种译文的比较（口语形式和

作者简介：李婧，香港理工大学中文及双语学系研究生，研究方向：口译，E-mail：rebeccajinglee@gmail.com。李德超，博士，香港理工大学中文及双语学系助理教授，研究方向：翻译，E-mail：ctdechao@polyu.edu.hk。

书面语形式), 研究者可以测试两种语言形式取得功能等效(functional juxtaposition)时, 依赖形态(modality-dependent)的程度有多大。例如, 源语“可以”一词, 笔译可能为“may”, 而口译目标语可能除了“may”还有“can”、“shall”, 也即形态有了变化。而口译对应语料库的其中一个研究目的就可以用来发现和比较口译译文中形态变化的规律。这样设计的好处是不仅可以研究口译目标语的特点, 还能够进一步研究非言语因素对口译的方向性(directionality)和语言特征的影响。

概括起来, 口译语料库可以帮助我们探寻口译译文的特点, 检验现存的口译理论, 考察各种口译策略对口译效果的直接影响, 还可以帮助我们撇开偏见, 研究语际交流和口译的特殊性。这种基于语料库语言学的方法和技术将最终帮助口译研究实现从规定性研究到描写性研究的飞跃(Baker, 1993: 248)。

3 现有口译语料库评介

迄今为止, 国内外建成并完成研究的口译语料库并不多见。^②在国外, 大型的口译语料库有两个, 即日本名古屋大学建立的CIAIR口译语料库和意大利博洛尼亚大学建立的EPIC欧洲议会口译语料库。国内则有中国大学生英汉汉英口笔译语料库。

3.1 名古屋大学综合语音信息研究中心同声传译语料库

日本名古屋大学的综合语音信息研究中心从1999年开始至2003年完成此语料库(简称CIAIR)的初步建设。CIAIR语料库收录了总时长达182小时的录音资料, 并全部将其转写成为文字, 转写后的文字总数约有100万词, 这让CIAIR语料库成为目前世界最大的同声传译语料库。语料库创建的目的在于研究语言信息处理技术, 提高语言翻译技巧以及完善口译理论。数据库语料也对外出售, 所得费用用于语料库维护和大学学术研究(参见Tohyama, 2006)。

从CIAIR语料库录音的内容来看, 所用语言组合均为英语和日语, 言语种类(speech type)包括对话和单独演讲。文本单位(discourse unit)同时具备时间标注(time tag)和文本识别标注(discourse ID)。单独演讲由5兆到150兆不等; 对话由2.5兆到110兆不等。每一段文本都配备了讲者声音同译者声音结合在一起的波段, 且均经过声道处理。当戴上耳机听时, 左右耳会分别传出讲者和译者

的声音。

到目前为止, CIAIR语料库共收录了30位英语讲者、15位日语讲者和31位口译员的录音及这些录音所转写的文本。这些录音均来自拟真场景, 而非真正的口译现场。口译员在同传厢中通过麦克风发声, 可以从耳机清晰地听到讲者并透过玻璃看到讲者的一举一动。讲者无法听到口译员的声音, 所以可以按照自己自然的语速讲演。从录音内容来看, 1999年单独演讲录音题材主要涉及日常话题及信息科技领域, 而对话录音则涉及旅游环境话题。

名古屋大学的研究者据此语料库作出的研究包括机器同声传译、同声传译中口译员的语速和翻译单位、同传译员和讲者的语言输出时间差等。

CIAIR语料库的最大优势在于其较大的规模, 且具有相对完善的建库技术, 非常有利于机器翻译, 包括机器口译的研制与发展。机器口译(亦称为机助口译或自助口译)是自然语音识别技术与自动翻译系统相结合而形成的一种新的翻译模式。而大规模的CIAIR口译语料库丰富了机器口译研究依赖的语料以及语言实践操作的基础, 有助于挖掘口译中的一些自动对应的结构和规则。CIAIR语料库的不足之处在于其收集语料的主题有限, 离满足不同交际情景的机器口译需求还有相当大的差距。另外, 由于口译语料库收集语料的环境为口译训练教室, 而非真实的口译工作环境, 并不能客观反映实际口译操作时的情景, 这在口译研究的“生态效力”(ecological validity)上亦打了折扣(Lindquist & Giambruno, 2006; Jakobsen et al., 2007: 228)。

3.2 欧洲议会口译语料库

欧洲议会口译语料库(European Parliament Interpreting Corpus, 简称EPIC)由意大利博洛尼亚大学的翻译语言和文化研究小组于2004所建。^③此语料库为电子对应语料库, 涉及的语种包括意大利语、英语和西班牙语。EPIC语料库收集的录音均来自欧盟议会全体会议的口译。语料库总规模为140个4小时的录影带, 包括5场2004年2月到7月的分会的口译录音(Bendazzoli & Sandrelli, 2005)。这些音像资料包括全体会议中的源语演讲(标识为“Org”)以及英语, 意大利语和西班牙语声道的同声传译(标识为“Int”), 亦有欧洲议会的新闻发布会的传译内容。

建库时, 首先是将演讲者的源语录影带数字化, 形成影音文件。在此过程中, 将原来以意大利、英语和西班牙语的源语录音及相对

应的同传译文分离，分别存储为独立的音频和视频片段。视频信息亦保留，以利日后进一步分析。对于源语录音的转写，研究者主要参考欧盟在会议后发布的详尽的官方稿件，完成文字初稿，再经审阅而得。而口译录音的转写则较为复杂。具体而言，研究者采用同传培训中经常使用的影子跟读方法，一边听口译员的录音，一边大声重复他们的译文，同时利用语言识别软件将复述的言语自动输出文字稿。译文中的副语言特征 (paralinguistic features) 则由研究者补充完成。其他与译员的口译输出相关的信息，如言语的长度 (长、短或中等)、发言模式 (即兴、带稿或两者兼有) 和平均速度 (快、中等、慢) 以及有关讲者的名字、国籍、性别和政治背景等，都记录在文字稿一个有着特殊设计的标头里，并可以用作检索的参数之一。在EPIC的网页上，如键入“发言时间”或者“译员、发言者的国籍”等，就可以搜索到整个语料库中具有所指定言语或讲者特征的语料。

EPIC语料库采用POS标注，意大利语和英语的文本用的是Treetagger软件，西班牙语文本采用Freeling软件。建成之后的语料库共涵盖三个源语文本 (分属意大利语、英语和西班牙语) 的子语料库和六个译语文本 (分属意大利语、英语和西班牙语) 的子语料库。换言之，每一种源语都有其他两个语言的译本，三种源语就有六个译本语料库。

EPIC语料库最大的特点就是其语料的同质性。从译员角度看，欧盟译员全部是经过严格筛选的专业人士，专业水平相当，且口译方向均是由译者的被动语言 (passive language) 译为主动语言 (active language)，即母语。其他影响译员表现的外界因素变量，如会前的准备、工作的设备等对于所有欧盟译员而言都相似 (如大家会前获得资料的途径和资料内容等都是一致的)。EPIC语料的这些同质特点尤其适合用于研究特殊体裁同传，例如欧洲议会辩论口译的文体研究等。

由于EPIC语料库下共有九个不同语言构成的子库，这种结构上的复杂性为多角度的研究口译提供了可能。例如，研究者可以比较自然英文口语和口译英文在句法、修辞、词汇等方面的各种差别，进一步检验Laviosa (2002) 提出的非翻译文本 (non-translational) 和翻译文本 (translational) 两者本质不同的假设。研究者亦可研究同声传译译文的方向问题，以及不同语言组合在口译中策略和结构上的特点比较等。

依笔者看来，除了研究作用外，EPIC语料库还可以广泛应用于教学目的，包括普通外语

教学和口译教学。在外语教学上，库中的源语视频片段是听力练习的好材料，而相对应的文字稿亦能帮助学生认识未知的词语和结构，以便更有效地更正错误和吸收新知识。而听力练习也可帮助学生提高外语的发音技巧。在口译教学上，博洛尼亚大学已经把语料库运用于学院的翻译和口译教学培训课程之中。其中的一些做法包括：把EPIC的视频片段和文稿作为口译练习材料，或者把议会译员的输出作为学生考评或者学生自我评价的标准，以增强学生对于自己优缺点的认识。另外，口译语料库还可以成为选取真实场景作业的题库，教师可以直接运用其中的某些录音或是通过软件切分加工语料，让练习语料符合学生的实际水平。

除了EPIC语料库，Bendazzoli (2009) 等人还建立了DIRSI语料库和FOOTIE语料库。DIRSI语料库 (Directionality in Simultaneous Interpreting) 主要用于研究同传的方向。在欧盟，同传通常要求译者只能从被动语言译至主动语言，而反方向则通常不被鼓励，但这却与市场的实际情况 (如东欧一些国家) 背道而驰。DIRSI语料库就是为专门研究同传方向对口译质量的影响而建。它收集了三年 (2005—2007) 在意大利召开的国际会议中所有的同传录音，包括会议的开场白、陈述、辩论问答环节。其中辩论问答环节因为其较强的互动性而与其他独白体部分区别开来。以上录音都经过如EPIC语料库般的转写、对齐和标注等处理。

FOOTIE语料库的录音均来自2008年欧锦赛意大利队总共16场比赛前后的新闻发布会口译内容，涉及意大利语、英语、法语和西班牙语。为了保证语料的同质性，文字转写的部分均为源语为意大利语和目标语为英语的录音，源语材料超过两个半小时。所有的译文都由同一位口译员提供，均从意大利语译成英语。FOOTIE语料库为同传语料库，新闻发布会的特征体现为高互动性的对话体，因此从这个角度看，FOOTIE的语料与通常会议中的问答环节有相似的特点，均采用一人对一人或一人对多人的交流形式。但研究时需注意的是口译听众的特殊性，即除了直接观众 (如在场的记者和工作人员) 外，还应当包括不在现场但关注新闻发布会的人士，如世界各地的球迷以及撰写赛况和文章的媒体从业人员等。这些都是利用FOOTIE语料库作研究时需要注意的。

除上述两个大型的、由机构所建立的口译语料库之外，国外亦有少数为专门研究某个问题由研究者自行建立的小型口译语料库。如Meyer (2008) 创建的用于研究交际和同传中对人名的处理和“医疗口译”的K6语料

库; Petite (2005) 所建的专门用于研究同声传译中的纠正机制 (Repair Mechanism) 的语料库; Cencini (2000) 建立的用于研究电视口译特点的口译语料库 (Television Interpreting Corpus); Fumagalli (1999—2000) 建立的研究英一意时事交传特点的类比和对应语料库等等。

以上口译语料库均属于专门用途口译语料库, 针对性强, 通常为了解决某些专门问题而设, 且库容较小, 影响和应用虽不及前两者大型口语语料库, 但其运用语料和实验, 而不是传统口译研究中常用的思辨、内省来研究口译的方式, 无疑开辟了口译研究的新天地, 值得借鉴。

3.3 中国大学生英汉汉英口笔译语料库

中国大学生英汉汉英口笔译语料库 (简称 PACCEL) 是由我国文秋芳教授等建立的包含中国大学生口译和笔译语料的大型学习者语料库。根据文秋芳等 (2008) 介绍, PACCEL 语料库收集了2003年到2007年全国18所高校英语专业的学生在大三大四时进行英汉汉英互译测试的语料, 总规模为210万字。PACCEL 主要分为两个子语料库: 口译对应语料库 (简称: PACCEL-S) 和笔译对应语料库 (简称 PACCEL-W), 规模分别为50万字和160万字。光盘里, PACCEL 语料库的结构包含三个部分: 口译语料库, 笔译语料库和检索工具。口译语料库又包括音频和文字转写, 文字转写在句子层面实现了源语和译文的对齐, 还进一步具备生文本和词性标注的文本。笔译语料库则包括文字部分, 并且所有文字部分也如口译语料库的文字转写部分一样, 实现了句级对齐并具备生文本和词性标注两类文本。

该语料库使用了高质量的音频设备保证较高质量的数字化转录, 即把磁带与电脑连接一起, 在播放磁带的同时采集语音。所有录音文件也都使用广泛兼容的MP3格式保存以保证文件大小适中, 然后用年份、组别、序列号组成每一份音频文件的名称, 为检索打下基础。

PACCEL语料库的口译语料主要收录的是TEM8测试中的口试部分, 由于所选语料范围针对性强, 所选群体为英语专业学生, 故其同质性较高, 再加上所选语料为TEM8全国统一考试, 其测试标准一致, 所以该语料库在代表性方面具有突出的优势, 它的建成可以帮助研究者了解中国学生学习英语的过程, 对教学、研究、测试、培训以及教材编写、网络远程教育等都具有重要意义 (文秋芳等, 2008)。

以上介绍的国内外三大口译语料库可谓各具特色。日本的CIAIR语料库的建库技术最为完善, 其基于口译语料库的机器口译研究走在世界前列, 但语料库在口译生态效力上却有所欠缺。意大利的EPIC语料库在研究、教学以及后期可持续扩充方面都比较完善, 且对语料的处理亦相当科学和严谨。我国PACCEL语料库检索功能齐全, 具备研究和教学价值, 但其本质上是一个学习者语料库, 研究领域受到局限。此外, 如要更好地发挥PACCEL语料库对学习进程的研究和比照作用, 宜在该语料库中增加与口译输出相比照的笔译文语料库。相对而言, 外国的大型口译语料库的建立目标明确, 研究对象清晰, 较易根据研究目的控制变量, 语料的后期处理亦更为科学, 且通常兼备常规和专门用途语料库的特点。我国口译语料库的发展还处于起步阶段, 建库通常用于教学目的, 若要进行口译研究的话, 语料涉及的变量很难控制, 这是我们以后建立大型口译语料库时需要注意的。

4 口译语料库的建库步骤

从以上对国内外主要口译语料库的回顾可以发现, 如要建立口译语料库, 一般要遵循以下步骤:

(1) 根据研究目的决定建立口译语料库的类型: 对应、类比, 还是两者兼具。

(2) 录音/影: 录音分为现场录音和实验室录音。两者均要获得作为研究用途的录制许可。尤其是现场录制, 应注意获得多方许可, 如讲者, 口译员以及会议组织方等。这是口译语料库建立的难点之一, 因为口译员往往有自我防护的意识, 再加上某些会议资料可能有版权保密要求, 所以要取得真实口译语料往往不易。

交传录音由于是单声道, 相对比较简单, 使用一般的录音设备或是软件录制即可。同传由于是双声道, 对于录音设备的技术要求相对较高。实验室录制时可以使用录音软件; 现场录音则要使用与同传厢的输入和输出设备相兼容的录音装置。录音过程也可以转为录影过程, 因为视频信息也可能是日后重要的分析材料。

(3) 录音/影资料数字化。有些通过设备获得的录音/影材料, 必须通过一定的软件处理, 成为音/视频文件, 才能被计算机读取 (Machine Readable) 或分析。可以用作数字化处理的软件有Pinnacle Studio (9.0), 它是一个影音捕捉和编辑软件, 影音文件的格式是

mpeg1。而口译数字化使用的一般是声音编辑软件,如: Cool Edit-Pro 2.0, 格式为.wav, 采样速率为32K; 声道为单声道; 采样分辨率达8 Bit (Bendazzoli & Sandrelli, 2005)。好音质为研究音韵特征, 如停顿, 犹豫的分布等提供了良好基础。

(4) 影音资料文字转写。文字转写应该是可机读的, 也应该是便于使用的, 同时还要决定采用何种转写规则。国际通用的转写规则是TEI (Text Encoding Initiative), 著名的英国国家语料库 (BNC) 使用的即是这个转写规则 (Burnard, 1994)。在技术层面上, 一般借助语言识别软件效率较高, 现在可用的语言识别软件有Dragon Naturally Speaking and IBM Via Voice等, 它们可以帮助获得初稿。但现阶段的汉语语音识别软件技术很不成熟, 出错率极高。所以文字转写的实质工作主要还是依靠人工完成, 这也是最耗时和耗力的部分。而言语特征, 例如未完成的句子, 错发音, 或者不合语法结构的表达, 在后期还需要通过精听, 于初稿基础上再进行补充。

(5) 文本标注和对齐: 在实现了文字转写之后, 需要对文字材料进行标注和对齐, 这可为日后的分析打基础。文本标注可以根据研究目的设置搜索参数; 文本的原文和译文的对齐则需要依赖对齐软件, 主要用于对应语料库。此外, 还可以进一步增加注解, 比如增加语言学或非语言学的特征: 句法、音律特征, 乃至讲者的肢体语言, 幽默使用等。

5 小结

语料库口译研究目前仍旧是语料库笔译研究的分支, 在许多方面都延续了语料库笔译研究的特点。如语料库的分类, 大体分为对应语料库和类比语料库; 开放型 (open-ended) 和非开放型语料库; 研究对象多是文字文本的对比分析。但随着时间和技术的发展, 语料库口译研究也开始显示出自身不同于笔译研究的特点, 如:

(1) 语料收集: 收集难度较大。录制、分类以及声音处理都极大地依赖各种以电脑技术和软件功能为主的科技手段, 收集程序比较繁复。此外, 语料来源相对较少, 真实场景的口译语料更是难得。

(2) 语料库的建立: 口译语料库建立的过程要比笔译语料库复杂, 主要在于数据搜集、文字转写、标注以及对齐等过程。其中文字转写要根据一定转写依据进行, 以及根据需要增加的笔译文本语料库。

(3) 语音特征: 口译语料库的基础在于口译资料, 与笔译语料库的文本资料最大的不同是能够凸显口语的特征。这对研究语音学, 音韵学相关问题也能提供重要依据。

(4) 教学价值: 口译语料库能为教学提供宝贵的材料。真实场景的语料可以为学生提供实用的口译练习环境, 职业译员的口译更能为学生提供参考和标准。对于外语学习者而言, 不同的口译语料库对于听力以及口音训练也能起到对比, 认知和改进作用。基于语料库进行的错误分析, 也可为教学分析和纠正学生问题找到更有效的途径。此外, 语料库还可以指导教材的编写, 从而“对用法更普遍的语言事实给予更多的关注” (黄昌宁、李涓子, 2002: 160)。

最后, 口译语料库建立时尤其需要注意以下几个问题: 第一, 语料的代表性。正如杨惠中所言, “语料库是否具有代表性直接关系到在语料库基础上所作的研究及其结论的可靠性和普遍性” (2002: 13)。但是牵制口译语料库代表性的因素繁复, 除却语料库的规模, 还有“语料加工层次、口译场景, 口译人员实践水平 (专业、学员) 等也受到其他语料库所没有的限制” (张威, 2009: 58), 所以如果单纯从代表性而言, 口译语料库很难做到绝对代表性, 解决办法之一可以从变量控制入手, 而变量的控制取决于研究目的。第二, 研究框架的科学性。口译语料库建立的目的之一要借助语料库从描写入手, 归纳演绎并结合分析和诠释, 以从不同角度加深对口译以及与相关现象的认识。但现在有些口译语料库设计理念不清, 研究目的不明确, 为建立而建立, 实属浪费资源。第三, 语料库的开放性。口译语料库的规模无论大小, 建成后应能继续充实语料和添加标注, 让先前收集的语料能“循环再用”, 形成语料库开放格局, 并应提供给同行借鉴, 以便共享和节约资源。□

注释

- ① 本研究得到香港理工大学项目资助, 编号为A-PC1M。黄立波博士曾对本文初稿提出宝贵意见, 特此鸣谢。
- ② 已完成的语料库口译研究案例较少。以下所介绍的口译语料库均是笔者在参考网络和大量书籍的基础上所得。参考角度包括中/外, 同传/交传, 语料库规模 (自建小型语料库/机构所建大型语料库), 语料库研究类型 (封闭型: 针对单个研究目的/开放型: 可以补充语料满足多次研究目的), 语料库口语场景: 真实语境/拟态 (实验室) 语境等。
- ③ 读者可参阅<http://dev.sslmit.unibo.it/corpora/>

corporaproject.php?path=E.P.I.C.来获得关于此语料库的详细介绍。

参考文献

- [1] Aston, G. Corpus use and learning to translate [J]. *Textus*, 1999(12): 289-314. Also available at URL: <http://www.sslmit.unibo.it/~guy/textus.htm>.
- [2] Baker, M. Corpus linguistics and translation studies: Implications and applications [A]. In M. Baker G. Francis & E. Tognini-Bonelli(eds). *Text and Technology: In Honour of John Sinclair* [C]. Amsterdam: John Benjamins, 1993.
- [3] Bendazole, C. & A. Sandrelli. Corpus-based interpreting studies: Early work and future prospects [J]. *Tradumatica 7. L'aplicació dels corpus linguistics a la traducció*, 2009.
- [4] Bernard, Lou. The TEND: Towards an extensible standard for the encoding of texts [J]. Seamus Ross, Edward Higgs(eds). *Electronic Information Resources and Historians*. London: British Academy, 1994.
- [5] Cencini, M. On the importance of an encoding standard for corpus-based interpreting studies [J]. *inTRAlinea* (Special Issue). CULT2K, 2002.
- [6] Fumagalli, D. Alla ricerca dell'interprete. Uno studio sull'interpretazione consecutiva attraverso la corpus linguistics [D]. Unpublished dissertation, Advanced School for Translators and Interpreters(SSLMIT), University of Trieste, 1999-2000.
- [7] Gile, Daniel. The processing capacity issue in conference interpretation [J]. *Babel*, 1991, 37: 15-27.
- [8] Gile, Daniel. Methodological aspects of interpretation and translation research [J]. S. Lambert & B. Moser-Mercer(eds). *Bridging the Gap: Empirical Research in Simultaneous Interpretation*, Amsterdam & Philadelphia: John Benjamins, 1994.
- [9] Jakobsen, A. L., Jensen, K. T. H. & Mees, I. M. Comparing modalities: Idioms as a case in point [A]. In F. Pöchhacker, A. L. Jakobsen & I. Mees(eds). *Interpreting Studies and Beyond*(Copenhagen Studies in Language)[C]. Copenhagen: Samfundslitteratur, 2007: 217-249.
- [10] Laviosa, Sara. *Corpus-based Translation Studies: Theory, Findings, Applications* [M]. Amsterdam & New York: Rodopi, 2002.
- [11] Laviosa, Sara. The corpus-based approach: A new paradigm in translation studies [J]. *Meta*, vol. 43(4), 1998: 474-479.
- [12] Leech, G. The state of the art in Corpus Linguistics [A]. Aijmer, K. & Altenberg, B. (eds). *English Corpus Linguistics: Studies in Honor of Jan Svartvik*, London: Longman, 1991: 8-29.
- [13] Lindquist, P. P. & Giambruno, C. The MRC approach: Corpus-based techniques applied to interpreter performance analysis and instruction[J]. *Forum*, 2006, 4(1): 103-138.
- [14] Olohan, M. *Introducing Corpora in Translation Studies*. Oxfordshire [M]. Routledge, 2004.
- [15] Petite, C. Repair mechanisms in simultaneous interpreting: A corpus-based analysis of interpreters' deployment of processing resources(English/French/German) [D]. Unpublished PhD thesis, Heriot-Watt University, Edinburgh, 2004.
- [16] Shlesinger, Miriam. Simultaneous interpretation as a factor in effecting shifts in the position of texts on the oral-literate continuum [MA]. Thesis. Tel-Aviv University, 1989.
- [17] Shlesinger, Miriam. Corpus-based interpreting studies as an offshoot of corpus-based translation studies [J]. *Meta*, 1998, 43(4): 486-493.
- [18] S. Matsubara, Y. Aizawa, N. Kawaguchi, K. Toyama, Y. Inagaki. Construction and analysis of word-level time-aligned simultaneous interpretation corpus — A corpus-based analysis of simultaneous interpretation [Z]. 2006.
- [19] Vanderauwera, R. *Dutch Novels Translated into English: The Transformation of a "Minority" Literature* [M]. Amsterdam: Rodopi, 1985.
- [20] Van Besien, F. & Meuleman, C. Dealing with speakers' errors and speakers' repairs in simultaneous interpretation[J]. *The Translator* 10(1), 2004: 59-81.
- [21] Zanettin, F. Bilingual comparable corpora and the training of translators [J]. *Meta*, 1998, 43(4): 616-630.
- [22] 黄昌宁, 李涓子. 语料库语言学[M]. 北京: 商务印书馆, 2002.
- [23] 王克非. 语言与翻译研究并重的双语平行语料库 [A]. 北京外国语大学科研处. 学术论文集(上) [C]. 北京: 外语教学与研究出版社, 2001: 50-58.
- [24] 王克非, 黄立波. 语料库翻译学十五年[J]. 中国外语, 2008(6): 9-14.
- [25] 王克非, 胡显耀. 基于语料库的翻译汉语词汇特征研究 [J]. 中国翻译, 2008(6): 16-17.

(下转111页)

根本原因。

隐喻的翻译研究还激发了新的思考：比如传统的翻译策略“二分法”是否合理？如何恰当地表述隐喻翻译策略？如何建构科学的翻译评价体系？这些都需要我们进一步研究。□

参考文献

- [1] Lakoff, G. The contemporary theory of metaphor [A]. *Metaphor and Thought* [C]. Cambridge: Cambridge University Press, 1993.
- [2] Lakoff, G. & Johnson, M. *Philosophy in the Flesh—The Embodied Mind and Its Challenge to Western Thought* [M]. New York: Basic Books, 1999.
- [3] Lakoff G. & Johnson, M. *Metaphors We Live By* [M]. Chicago and London: The University of Chicago Press, 2003.
- [4] Mac Cormac, E. R. *A Cognitive Theory of Metaphor* [M]. Mass: MIT Press, 1985.
- [5] Nida, E. *Language and Culture* [M]. Shanghai: Shanghai Foreign Language Education Press, 2003.
- [6] Schäffner, C. Metaphor and translation: Some implications of a cognitive approach [J]. *Journal of Pragmatics*. www.elsevier.com/locate/pragma, 2004(36).
- [7] Toury, G. *Descriptive Translation Studies and Beyond* [M]. Amsterdam/Shanghai: John Benjamins Publishing Company/Shanghai Foreign Language Education Press, 1995/2002.
- [8] Zoltán Kövecses, *Metaphor(A Practical Introduction)* [M]. Oxford: Oxford University Press, 2002.
- [9] 方梦之(主编). 译学辞典 [Z]. 上海: 上海外语教育出版社, 2005.

- [10] 姜其煌. 欧美红学 [M]. 郑州: 大象出版社, 2005.
- [11] 宋德生. 认知的体验性对等值翻译的诠释 [J]. 中国翻译, 2005(5).
- [12] 王寅. 中西学者对体验哲学的论述对比初探 [J]. 外语与外语教学, 2004(5).
- [13] 王寅. 翻译的认知语言学模式 [J]. 中国翻译, 2005(5).
- [14] 许渊冲. 翻译的艺术 [M]. 北京: 五洲传播出版社, 2006.
- [15] 赵艳芳. 认知语言学概论 [M]. 上海: 上海外语教育出版社, 2004.
- [16] 郑海凌. 翻译标准新说: 和谐说 [J]. 中国翻译, 1999(4).
- [17] 郑海凌. 文学翻译的本质特征 [J]. 中国翻译, 1998年(6).

A Study of Metaphor Translation from the Perspective of Conceptual Metaphor

Abstract: Many inspirations are drawn from the turn of studies of metaphor which is from cognitive linguistics to rhetoric. Under the view of conceptual metaphor, some basic understandings of metaphor translation have been revealed. To take the metaphor translation in literary works for example, the metaphor translation is a cognitive process based on translators' cultural embodiment: cultural contexts, literary contexts and linguistic contexts are the inherent basis on which translators make their choices of translation strategy. The translation result, therefore, is always a kind of undertranslation, and the state of undertranslation can be vaguely evaluated in terms of the translation strategy.

Key Words: metaphor translation; conceptual metaphor; embodiment; context; undertranslation

(上接105页)

- [26] 文秋芳, 王金铨. 中国大学生英汉汉英口笔译语料库 [M]. 北京: 外语教学与研究出版社, 2008.
- [27] 薛沛文. 基于小型语料库的会议口译中显化现象研究 [MA]. 上海交通大学, 2007.
- [28] 杨惠中. 语料库语言学导论 [M]. 上海: 上海外语教育出版社, 2002.
- [29] 张威. 口译语料库的开发与建设: 理论与实践的若干问题 [J]. 中国翻译, 2009(3): 54-59.

Corpus-based Interpreting Studies:

The State of the Art

Abstract: Corpus-based translation studies as a

discipline has been gaining ground for the past 15 years. While there are abundant corpus-driven studies on written translation today, few similar studies are conducted in the field of interpreting studies largely due to the dearth of rigorously-built interpreting corpora. Based on the review of three existing interpreting corpora in the world, the paper summarizes the features and the necessary steps in building interpreting corpora. The paper concludes by pointing out the prospects and pitfalls of corpus-based interpreting studies.

Key Words: corpus; corpus-based translation studies; interpreting studies