



CLIC2013

第二届中国语料库语言学大会

中国语料库语言学研究会主办

东华大学外语学院承办

2013年10月19-20日

目 录

CLIC2013 第二届中国语料库语言学大会宗旨	2
CLIC2013 第二届中国语料库语言学大会议程	3
CLIC2013 第二届中国语料库语言学大会论文摘要	6
English for Specific Purposes	6
Corpus Construction and Software Development	18
Corpus and Translation	23
Learner Interlanguage Analysis	33
Other Issues on Studies in Corpus Linguistics	48
CLIC2013 第二届中国语料库语言学大会参会指南	65
CLIC2013 第二届中国语料库语言学大会会议通讯录错误!	

未定义书签。

CLIC2013 第二届中国语料库语言学大会宗旨

语料库语言学作为外语教学与研究领域中重要的支撑学科，理所当然地成为外语学科建设中重中之重的内容。本次大会是在充分总结了第一届大会成果的基础上，希望在语料库语言学的理论研究、技术运用和学科拓展方面有所提升。本次大会重点讨论专门用途英语、学习者中介语、语料库的建设与软件开发和语料库与翻译等四个方面的内容。就以上的内容和其所涉及的相关问题，与会的学者和语言教师进行广泛而深入的讨论，同时，就语料库语言学的建设内容和发展方向进行多方面的交流。

第二届中国语料库语言学大会由中国语料库语言学会研究会主办，东华大学外语学院承办。本次盛会学界鸿儒缜密筹划，年轻后学积极响应，齐聚东华者来自中国大陆和香港地区，以及英国的近七十所大学和研究机构的学界同仁。

本次大会将为中国语料库语言学学界同仁提供加深学术研究意识，提升学术研究能力的学术交流平台。

金秋十月，丹桂飘香。“学界盛事”当不负这美好的时节。

CLIC2013 第二届中国语料库语言学大会日程

日期	时 间	日程安排	
10 月 18 日	9: 00-21: 00	报 到: 登记入住 地点: 新沙珑国际宾馆/维也纳酒店	
	18: 30	主持人会议 (地点另行通知)	
10 月 19 日		早 餐 地点: 新沙珑国际宾馆/维也纳酒店	
	7: 30	乘车赴会场 地点: 维也纳酒店门口	
	7: 45	地点: 新沙珑国际宾馆门口	
	8: 30-9: 10	开幕式 地点: 东华大学图文中心第二报告厅	
		俞建勇副校长致辞 主持人: 外语学院院长 赵晓临	
		合影留念 (图文中心外)	
	9: 15-10: 05	主旨发言 (50 分钟) 地点: 东华大学图文中心第二报告厅 主题: <i>On Corpus Linguistic Approach to Unit of Meaning—Its Development and Pedagogic Implication</i> 发言人: 何安平 主持人: 李文中	
	10: 15-11: 30	分组研讨会	
		主题 I: English for Specific Purposes	地点: 图文中心第二会议室
		主题 II: Corpus Construction and Software Development	地点: 第三会议室
主题 III: Corpus and Translation		地点: 第四会议室	
主题 IV: Learner Interlanguage Analysis		地点: 第五会议室	
	主题 V: Other Issues on Studies in Corpus Linguistics (I)	地点: 第七会议室	
11: 45-12: 45	午 餐 地点: 东华大学二食堂三楼招待厅		
13: 00-13: 50	主旨发言 地点: 东华大学图文中心第二报告厅 主题: <i>An Empirical Account of Coordination Based on a Corpus of Spoken and Written British English</i> 发言人: Alex Chengyu Fang 主持人: 梁茂成		
14: 00-17: 00	分组研讨会		
	主题 I: English for Specific Purposes	地点: 第二会议室	
	主题 II: Corpus Construction and Software Development	地点: 第三会议室	
	主题 III: Corpus and Translation	地点: 第四会议室	
	主题 IV: Learner Interlanguage Analysis	地点: 第五会议室	
	主题 V: Other Issues on Studies in Corpus Linguistics (I)	地点: 第七会议室	
17: 45-19: 30	欢迎晚宴 地点: 新沙珑国际宾馆 东华大学外语学院党总支书记林嵘致辞		

日期	时间	日程安排
10 月 20 日		早餐 地点：新沙珑国际宾馆
	7: 45	乘车赴会场 地点：新沙珑国际宾馆门口
	8: 30-9: 20	主旨发言 地点：东华大学图文中心第二报告厅 主题： <i>Corpus Linguistics and English for Specific Purposes</i> 发言人： Dr Lynne Flowerdew 主持人：濮建忠
	9: 30-10: 50	专家论坛 主题： <i>The Changing Corpus Linguistics</i> 专家：何安平、李文中、梁茂成、濮建忠、卫乃兴等
	11: 00-11: 20	闭幕式 主持人：东华大学外语学院副院长、东华大学莎士比亚研究所所长 杨林贵
	11: 35-12: 15	午餐 地点：东华大学二食堂三楼招待厅
	13: 45-17: 00	松江半日游：泰晤士小镇、佘山风景区
	17: 30-19: 00	佘山度假村欢送晚宴
	19: 00	返回宾馆

分会场安排

第一分会场

上午 主持人：甄凤超、杨海燕
发言人：张继东、甄凤超、贺霞、翟萌
下午 主持人：张继东、吴蕾、周红莉
发言人：杨海燕、吴蕾、阮周林、张绪华、胡玉迎、于涛、张立英、朱玉山、张晓明、严齐、杨媛媛、崔珣丽

第二分会场

上午 主持人：许家金、熊文新
发言人：邓耀臣、金罡、金菊花、毛浩然
下午 主持人：邓耀臣、张立飞、许霄羽
发言人：张立飞、许霄羽、陈功、邓飞、贾云龙、赵志敏、房印杰、郑青、倪修璟、陈杰、吴雪颖、焦俊峰、巩海曦

第三分会场

上午 主持人：刘泽权、刘克强
发言人：刘泽权、李潇辰、刘克强、刘磊
下午 主持人：徐风华、林作帅、兰凤利
发言人：刘鼎甲、种法娟、谭业升、徐风华、陈昕、贾艳、兰凤利、林作帅、刘君、刘燕、于璐、彭杰、王余余、田建国

第四分会场

上午 主持人：刘国兵、陆军
发言人：刘国兵、刘霞、陆军、赵晓临
下午 主持人：杨素香、任再新、张莉
发言人：权立宏、杨素香、刘晓燕、曹竞、曹蔓、杨晓东、张金福、张莉、方秀才、郭放、黄青

第五分会场

上午 主持人：王俭、孙海燕
发言人：濮建忠、孙海燕、王俭、李艳娇
下午 主持人：王均松、潘璠、方秀才
发言人：熊文新、潘璠、张碧萱、王均松、李华勇、李威、刘宇婷、毛越浩、吴羊、杨元媛、印辉、刘瑛

CLIC2013 第二届中国语料库语言学大会论文摘要

English for Specific Purposes

Corpus Linguistics and English for Specific Purposes

Lynne Flowerdew Hong Kong University of Science and Technology

In this presentation I aim to showcase recent trends in corpus investigations of two key academic genres, namely the research article and postgraduate dissertation writing (see Boulton et al. 2012; Flowerdew in press, 2013, 2014a, b; Gotti & Giannoni in press, 2013). The implications of the corpus findings for pedagogy will also be discussed.

The research article has been the focus of many corpus studies, which is not surprising given that this genre is easily accessible. While disciplinary variation has been a key topic for investigation, much work has also been carried out from an intercultural rhetoric perspective. Postgraduate dissertation writing of both L1 and L2 English writers across a variety of disciplines has also been investigated by corpus researchers, albeit to a lesser extent, as this is somewhat of an occluded genre (Thompson 2013). However, the research article and the dissertation are constantly-evolving genres. Corpus studies have recently begun to take account of the increasingly interdisciplinary nature of academia, the impact of digital technologies and the English as a lingua franca factor. New software, new algorithms and new ways of data representation are also being employed.

Corpus-based Contrastive Study of Reporting in Academic Papers of English Applied Linguistics

He Xia Beijing University of China

Reporting plays a crucial role in the academic paper, as the use of reporting could introduce the previous research findings or results, and the point of views of the predecessors. Thus lay a solid foundation to support the author's arguments, or to overthrow the opponents' opinions. Based on the self-established corpora, the author selected randomly 30 papers respectively from the English applied linguistics papers in the last 3 years by the English native authors and Chinese authors, and tagged the reporting sentences and reporting verbs. This study intended to investigate reporting from four categories, that is, the distribution and different use of reporting patterns, reporting verbs, tenses and voices of reporting verbs in both corpora. This study found that there are certain similarities and differences in the use of reporting. First, among the four reporting patterns, the integral reporting pattern and the non-integral non-reporting pattern are the most frequently employed patterns, while the integral reporting pattern is more preferred in the Native Authors Corpus(NAC), the nonintegral nonreporting more often used in Chinese Author Corpus(CAC), and there

are significant differences between these two patterns in both corpora; second, English native authors use more various reporting verbs to introduce the previous research findings or opinions, and research verbs dominated the reporting verbs, however, Chinese authors use more discourse verbs and there is significant difference of discourse verbs in both corpora; third, simple present tense and present perfect tense are more often occurred in the English corpus whereas Chinese authors use more simple past tense than the native authors, and there is significant difference in simple present tense in both corpora; last, active voices are the most commonly used reporting voice in both corpora, while native authors use more passive voice than the counterpart and Chinese authors employ more active voice, there is significant difference of active voice of reporting verbs between English native authors and Chinese authors. The findings reflected that different thinking modes, cultural diversities and academic custom affect the reporting use of English applied linguistic papers by native authors and Chinese authors.

Semantic Prosodies of Near Synonyms in Academic Texts

Meng Zhai Beihang University

Being a recurrent theme in corpus linguistic research, the phenomenon of semantic prosody (SP), uncovered in the 90s by way of examining the semantic characteristics of the collocates accompanying certain lexical items, has been stirring up diverse arguments about its nature, significance and applicability in language use, as can be testified by the amounts of study that scrutinize the concept from different perspectives and under distinctive linguistic circumstances. The undetermined definition of semantic prosody, though to a certain extent may cause a misunderstanding or misuse of it (often blended with the concept of connotation and complimentary/derogatory meaning), motivate scholars in and out of corpus community to dig deeper on this line of research, which have shown interesting findings both theoretically and practically. Over the twenty years since the term semantic prosody has been coined, the concept of semantic prosody has undergone changes from the earlier semantic meaning attributed to it to the more recent pragmatic meaning linked to it. The previous research in the field of semantic prosody, though rich in general contexts, are rather limited in specialized especially academic texts. It may be claimed that semantic prosody, which is associated with the authors' attitudinal meaning, rarely occurs in the more objective academic context, but Wei (2002) examined semantic prosody in academic texts, and pointed that polarized semantic prosody exists in specialized text with "cause" attaining stronger negative prosody in specialized texts, "career" showing positive semantic tendency, and "probability of" demonstrating a mixed semantic prosody but with a negative tendency.

To begin with, the current research exhibits the three main stages where the contagious, connotational, and functional view has been held respectively toward the concept of semantic prosody. By providing a synthesis of the process of semantic

prosody study conducted, the present study endeavors to investigate the employment of semantic prosody in academic texts via inspecting three phrasal verbs in CAUSE group, namely “bring about”; “lead to”; and “result in”. In light of Sinclair’s (1996) notions of co-selection and units of meaning, the study followed Sinclair’s analysis procedure of lexical items, starting from examining the collocational patterning of a certain lexical item to abstracting its semantic prosody. Adopting JDEST corpus (1200 texts of academic English with 4 million running tokens) as the data resource, and through the WordSmith 4.0 concordance program, the research demonstrates that academic texts do show semantic prosody but in a more implicit manner, and lexical items with similar denotational meanings or functions have different semantic prosodic distributions on the semantic continuum in academic texts.

The study, though limited in its scale, indicates that explorations of academic texts from the perspective of semantic prosody may uncover interesting facts, and the phenomenon of semantic prosody in academic texts is significant for EFL language teaching and scientific translation. As an initial step for further study on semantic prosody in the academic context, the present study still has some limitations. In further research, large amounts of data would be needed to explain the phenomenon of semantic prosody in academic prose, and larger units of lexical items (endocentric or exocentric structures) could be examined. Furthermore, it could be rewarding if the unique characteristics of semantic prosody in academic texts be explored deeply in comparison with general texts, and comparative studies be carried out between academic texts written by Chinese scholars and those by Western scholars in terms of semantic prosody.

Phraseological Features of Collocational Framework a * of in Medical Research Article Introductions

Jidong Zhang Donghua University

Denoting a framework composed of two discontinuous words (mainly grammatical ones), “collocational framework” investigates the discontinuous lexical combinations which are often neglected in traditional lexical study. Studies on collocational frameworks can serve as a supplement to those on recurrent lexical sequences, and are therefore of huge research significance in Phraseology. This paper follows a corpus linguistics approach combined with genre analysis theories and explores the collocational framework a * of together with their collocates in the genre of English medical research article introductions. The framework and its collocates are detailedly analyzed through their concordance lines, MI value, semantic categories and functional characteristics in the paper. It is intended to make an inquiry into the role of collocational frameworks in the construction of the whole discourse and in the realization of each move’s communicative purposes with a special aim to offer some reference and help for composing this genre.

The Rhetorical Functions of Sub-Technical Vocabulary: An Analysis of Evidence, Research and Results in Linguistic Journal Articles

Fengchao Zhen Shanghai Jiao Tong University

This paper discusses the rhetorical functions of sub-technical vocabulary in different sections of linguistic journal articles by means of corpus-based analysis of KWIC, collocation, and grammatical pattern. We choose three sub-technical items as research objects, namely evidence, research and results. The evidences show that in different sections of the articles, there is an apparent difference in the usages and meanings of sub-technical vocabulary. And these different usages and meanings can help the writer achieve certain rhetorical functions in writing. For example, the word research is employed frequently in the section of introduction and mainly manifests itself in the patterns of “in + proper noun + research” and “research + in/into + proper noun”, by means of which the essay writer aims to establish the research area and subject. In the section of literature review, evidence and research both occur frequently and share the identical pattern of “there is + no/little + evidence/research + preposition” by which the writer intends to point out the problems in the previous researches and implies that his research can fill the gap. In the section of results and analysis, the word results occurs frequently with its typical pattern of “the results + of + this/our + study/analysis + indicate/show + that/noun” because in this part the writer mainly presents the data and major findings. In the part of discussion, the three sub-technical words all occur frequently but with different patterns. However, they are all employed to realize the same pragmatic purpose, that is, to interpret the research findings from the theoretical perspective and to reinforce the positive evaluation of the research. In the part of conclusion, the word research is used frequently with the typical pattern of “further/more research + is needed/needs to be done”, by means of which the writer aims to point out the limitations of his research.

This research has the pedagogic implication, namely that learners should have a good command of the sub-technical vocabulary in order to improve their skill of reading and writing academic papers.

The strategies of writer/reader relationship management in Chinese apprentice business letters

Hu Yuying Guangdong University of Foreign Studies

The research will investigate how apprentice Chinese business letter writers construct relationships with a range of readers within and outside organizations. It will be carried out on the basis of a self-built mini personal professional correspondence corpus of 1,0187 tokens. Following the correspondence text mapping framework (Mike Scott and Christopher Tribble, 1996:109-129), firstly, the researcher will draw a brief text mapping picture of the investigated business letters in terms of four cardinal points: Internal or External, Close or Distant. Secondly, the KW analysis

method will be mainly adopted to identify words that are important in backchannel move critical to the management of writer/reader relationships in business correspondence. The investigation focus of KW analysis will be laid on the following aspects: collocation, colligation and semantic prosodies of the investigated keywords. Thirdly, language features evaluation criteria such as ellipsis, vague language and lexical density will also be surveyed in the corpus data so as to support what have been illustrated in the previous 2 steps. It is predicted that the letters to different receivers within or outside organizations would locate on different text map positions; hence the language features would also be differently expected with formal or informal features and it would reveal the different strategies used by the letter writers. To sum up, the forward going research would draw a very brief picture illustrating how Chinese apprentice business letter writers construct writer/reader relationship using different writing strategies. Teachers can use these pedagogic resources to help students polish their successful writing strategies and improve the unsuccessful ones, too. Moreover, the findings would also be a helpful instruction both to business letters textbook compiling and syllabus design.

基于语料库的医学学术语篇结论部分的人际意义研究

马静，金城，颜海璐 东华大学

近年来，医学学术语篇作为学术语篇的一部分，因其在人们日常生活中的重要作用日益受到学者们的关注。在学术语篇受到愈来愈多学者重视之时，有许多学者开始就医学语篇的语类结构，语言特征等从系统功能语言学的概念功能和语篇功能的角度对其展开了多方研究，取得了不少研究成果。但借助系统功能语言学三大功能之一的人际功能就医学学术语篇的结论部分展开的探讨却不尽充分。本文借助功能语法的主要观点，从人际功能角度分析了论文作者是如何通过结论部分的情态、语态、人称等语言手段的合理运用来建立与读者之间的互动并最终实现其人际功能的。

人际功能是系统功能语言学三大元功能之一。本文采用语料库研究方法，从人际功能的语气、情态、人称等方面对 100 篇医学论文的结论部分进行了分析。研究发现医学论学术语篇结论部分常使用被动语态体现医学语篇客观性和简洁性的特征。在情态系统下，研究得出作者多使用中低级情态操作词而非高级情态词来表达可能性。在情态取向上，显性客观取向表达了更多的信息。另外，还发现使用最多的是动词性模糊语，其次依此为情态助动词，副词，形容词和名词。本研究旨在探讨和剖析作者是如何通过结论部分来建立与读者之间的联系并影响读者的。希望本研究的发现能对语篇人际功能的教学，毕业论文的撰写及论文的发表有所启示。

A Corpus-based SFG Study of Nominal Groups in Written Business Discourse: the Case of “Business”

Zhoulin Ruan Xi'an Jiaotong-Liverpool University

This paper investigates the internal structure of complex nominal groups (NGs) that occur in written business discourse, following the SFG (Halliday & Matthiessen, 2004) approach to interpreting their ideational meaning. The paper addresses two research questions: 1) Are there any patterns of the experiential and logical structures of complex NGs in written business discourse? 2) What are the features of the functional elements of NGs in written business discourse? To answer these research questions, a one-million-word corpus of financial corporate annual reports was built, from which complex NGs with “business” as the head noun were extracted through WordSmiths Tools (Version 5). The internal structure of each complex nominal group was then analysed in terms of the experiential structure, from deictic to classifier, and the logical structure, either univariate or multivariate. The analysis shows that non-specific deictics tend to co-occur, proportionally, more frequently, with interpersonal epithets, but less frequently with classifiers, whereas specific deictics behave in an opposite pattern in relation to the other two functional elements. It was also found that classifiers, which tend to rely on internal bracketing for specifying and classifying the Head noun, are the most frequently occurring functional element, indicating the pragmatic purpose and subject matter of nominal groups in the informational writing of business English. Finally the multivariate structure was found to be the preferred logical structure of NGs. The paper concludes with some pedagogical implications of the findings for the teaching of business English to ESL learners.

A Corpus based Contrastive Study of English Guidebooks: A case study of Shanghai

Haiyan Yang Shanghai University of International Business and Economics

Tourism and language are closely related. However, up to now, tourism discourse in China still lacks a systematic and comprehensive study due to the complexity of tourism phenomenon and large amounts of the tourism discourse are designed from the public. This study attempts to investigate how English is used in a typical tourism discourse-guidebook for westerners to tour in Shanghai, P.R. China. A contrastive study is carried out to explore the distinct linguistic features of English guidebooks of Shanghai published by Chinese and overseas publishers. This study addresses the following three questions:

1. What are the similarities and differences of English guidebooks published by Chinese and overseas publishers with regard to the linguistic features at the macro level?
2. What are the similarities and differences of English guidebooks published by Chinese and overseas publishers with regard to the linguistic features at the micro level?
3. What accounts for these similarities and differences between the two groups of English guidebooks?

The contrastive study of English guidebooks adopts the mixed methodology of quantitative and qualitative approach. The qualitative approach dominates the macro level analysis of the guidebook features. The similarities and differences are examined in terms of format, content, register and genre based on register analysis and generic analysis. At the micro level, the corpus approach is adopted to examine the lexical and grammatical features. The linguistic context, professional settings, social, cultural and cognitive context are taken into consideration at both two levels and the whole study is based on the self-built corpus of English guidebooks of Shanghai.

The result shows that at the macro and the micro linguistic level, the similarities and differences between guidebooks by Chinese publishers and overseas publishers do exist. The similarities arise from the fact that they share similar communicative purposes: informational, promotional and persuasive. The significant differences demonstrate that the two groups of guidebooks can be considered as sub genres since the two groups of guidebooks are written and edited by two groups of distinct authors or editors. The obvious differences between the authors or editors can be grouped into the variables such as English proficiency, cognitive style, intimacy with tourism destination and attitude toward tourism destination and its culture. These variables account for the significant differences between the two groups of guidebooks.

The main contribution of the present study lies in the fact that contrastive analysis of English guidebooks is conducted between two distinct groups of authors or editors. From the viewpoint of language proficiency, the Chinese authors and editors can be regarded as non-native English speakers whereas the overseas authors can be regarded as native English speakers. From the viewpoint of tourism destination, Chinese authors are locals whereas the overseas authors are non-locals, although they are familiar with Shanghai and China. Therefore the crossing features of the two groups of authors or editors results in noticeable differences between the groups of guidebooks with relation to the same tourism destination- Shanghai.

The study can shed some lights on the further tourism discourse studies, specialized discourse studies and discourse studies. The study can also provide useful references for the improvement of readability of English guidebooks by Chinese publishers. In addition, the study provides valuable resources for teachers and learners in their tourism English teaching and learning, Shanghai's tour guide English teaching and learning in particular. Furthermore, the study also provides effective perspective and suggestions for tourism administrative bureaus and tourism enterprises on how to promote tourism destination in an effective and efficient way so that more and more potential tourists can be converted into actual tourists.

基于语料库的 IF 条件句句法语义特征的语体差异研究

于涛 北京外国语大学

条件句是句法、语义和语用等因素相互交织的复杂语言现象，通常揭示命题间逻辑关系，而这种关系主要体现在语义和语用层面。长久以来，哲学、心理学及语言学研究从真值条件、推理过程及逻辑和意义等不同视角探讨这个颇有意义

的话题。然而，当前国内外条件句的研究多集中在理论探讨，实证研究则仅涉及其句法结构研究，忽略了条件句中最为重要的语义关系。本研究采用 Crown 和 CLOB 赋码语料库，考察 IF 条件句在句法和语义层面在四种语体间的差异；其中句法结构主要涉及主从句中动词的时体，语义层面主要考察主从句之间的语义关系（开放条件句、假设条件句和修辞条件句）；为了研究目的，IF 条件句须有以下特征：1)条件句分句以 IF 引导；2)主、分句完整；3)IF 前无 ONLY 或 EVEN 等副词；4)剔除引导补充成分的 IF 小句。本研究拟回答三个研究问题：1) IF 条件句频数在各语体的总体分布特征体现在哪些方面？2) IF 条件句主从句中主动词的时体在语体之间是否有差异？体现在哪些方面？3) IF 条件句的语义关系在语体之间是否有差异？体现在哪些方面？为了回答以上问题，首先，使用 PowerGrep 软件借助正则表达式提取符合以上特征的 IF 条件句及其分句中主动词的时体；其次，IF 条件句中三种语义关系的人工分类、识别和统计主要使用 Excel。研究结果显示：1) IF 条件句占各语料库中句子总数的比例依次为新闻>通用>小说>学术，新闻语体往往借助条件句引发读者的注意和推理，因而频数较高，而学术语体强调信息表达准确性，因而少用条件句；2) 新闻语体为了追求时效性在主从句中多采用现在时，小说和学术语体中大都描述过去发生的事情，因而在主句和从句中多使用过去时，而通用语体没有明显的倾向性；3) 新闻语体为了信息传达的准确性，假设条件句的比例明显少于开放条件句；小说语体中的条件句多出现在对话中，说话人为了降低语势多使用假设条件句；学术语体中假设条件句和开放条件句数量相当，使用假设条件句主要目的在于提出研究假设，而开放条件句主要目的在于体现主句命题的真实性；修辞条件句在以上三种语体中极少出现；通用语体与以上三种语体有所不同，偏向于三种条件句混合使用。

新闻报道中完成时态的使用及语篇功能

张立英 解放军外国语学院/上海交通大学

对于完成时态的研究，传统语言学一般局限于句子层面，时态的语篇或叙事功能探讨多局限在口头或文学叙事。本研究探讨现在完成时在新闻报道中的语篇功能，并尝试将语篇功能与时态的意义、动词特征联系起来，从而衔接微观和宏观层面的研究。研究对象为自建并已完成时态标注的《纽约时报》语料库，在探讨用于完成时态的动词特征时，借助 Wordsmith 中的 MI 和 LL，计算出与完成时态有关联的动词，并对计算出的动词，按照 Vendler 的分类进行归类分析，从而总结出新闻报道中完成时态动词使用的趋势特征。为进一步探讨完成时态的语篇功能，我们统计了完成时与其他时态连用的情况，提取出高频的时态群集 (cluster) 及其分布特征。初步发现完成时主要用于：(i) 连续使用 (PP+PP+PP...); (ii) 引导一系列连续使用的一般过去时 (PP+P+P+P...); (iii) 穿插在一般过去时或一般现在时中使用 (SP/Pr+SP/Pr+...+PrP+ SP/Pr+...+SP/Pr) (iv) 用于一系列一般过去时的结尾处 (SP+SP+...+SP+PrP)。对这些时态群集的分析表明，连续使用的完成时态可以：1、增强语气表达一种情感；2、与一般过去时构成对比，标志两个不同时间的故事主线。单个完成时态，在与其他时态连用时，可以：1、引入一个话题或者标志一个新话题的开始；2、突出记者认为有趣或重要的部分；3、标志报道或者一部分 episode 的结束。完成时的这些功能可以从完成时表达的意义与动词的特征得到解释，从而从微观层面解释完成时的语篇功能。

The Generic Structure of PhD Concluding Chapters in Applied Linguistics

Xiaoming Zhang, Zaixin Ren, Rongming Wang Donghua University

Genre analysis of academic discourse has attracted much scholastic interest ever since 1980s. Most of the attention has been devoted to the research article (RA). Almost all structural organization of different sections of RA have been examined by researchers. Even research paper titles (Haggan, 2004) have been examined. *Conclusions* have usually been regarded as part of the *Discussion* section of a research article (RA) (Swales, 1990), so up to now only a few studies have investigated conclusions as a separate part in APs. Not to mention the study on *Conclusions* in PhD dissertations.

This essay aims to investigate, based on the move-step framework proposed by Bunton, a comparative study of the generic structure of PhD dissertation conclusion chapters by native speakers of English and Chinese.

Previous research on conclusion chapter has shown that *Conclusions* are less prominent in RAs than *Discussions* and has attracted little attentions. *Conclusions* in longer theses esp. PhD dissertations have by and large been neglected in the literature, with reference to Bunton (1998, 2005).

Bunton (1998) found that *Conclusion(s)* was almost the compulsory chapter in PhD dissertations, while sometimes this chapter was called *Summary*. In 2005, Bunton revised his move-structure framework as the following:

Moves found thesis-oriented PhD *conclusions*:

- 1) Introductory restatement
- 2) Consolidation of research space
- 3) Practical applications/ implications/recommendations (Embedded elsewhere)
- 4) Recommendations for future research (Embedded elsewhere)
- 5) Concluding restatement (Bunton,2005,p217)

Up to now, few books of advice to students on how to write a Phd dissertation could be got at home or abroad. Bunton's research (1998, 2005) and the move-structure framework that he has proposed for PhD dissertation concluding chapters are of great value for doctoral students and their supervisors.

This study is based on 40 PhD dissertations in Applied Linguistics by native speakers of English and Chinese from CNKI and ProQuest with the focus on structural move analysis, which is chiefly because generic structure is of great pedagogical significance. Either sub-corpus (ns &nns) consisted of 20 texts from about 15 degree-award universities both in the United States and in China. All the 40 dissertations were finished and submitted between the year of 2005 and 2010.

After the required corpus was obtained, a structural move analysis was performed. Before the move analysis, the concluding chapters of the dissertations were identified.

Findings will be discussed in great details, first on the titles of the concluding chapters, then on their lengths, number of references, section headings, and finally the

generic moves and steps identified according to Bunton's model (2005). Divergences have been found, mainly in the frequencies of certain moves and steps. This practice will make PhD students move-sensitive and will foster in them to write good concluding chapters.

I'm going to...: Lexical Bundles in Academic Spoken Disciplines

Xuhua Zhang Fudan University

This paper investigates the use of lexical bundles in four academic spoken disciplinary areas: arts and humanities, life and medical sciences, physical sciences, social studies and sciences. Lexical bundles are the most frequently recurring sequences of words in a given register (Biber & Barbieri, 2007: 264). Following Altenberg (1998), Biber et al. (1999), Biber, Conrad, and Cortes (2004), Biber & Barbieri (2007) and Hyland (2008), we took a frequency-driven approach to the identification of lexical bundles in four academic spoken disciplines and applied a functional taxonomy, including stance expressions, discourse organizers, and referential expressions. The use of the lexical bundles in each of these four disciplinary areas was analyzed and compared to published results in research on written and spoken academic prose. The texts and data used in this study come from the British Academic Spoken English (BASE) corpus (see Thompson & Nasi, 2001). The findings show that lexical bundles are prevalent in the above mentioned academic spoken disciplines. The most surprising finding to emerge from the present study is the great number of stance bundles in all these academic spoken disciplines. Speakers use more stance bundles to convey attitudes or assessments of certainty than they do with discourse organizers to signal relationships between prior and coming discourse, or referential bundles to identify entities or specific parts of entities. This study also reveals the extent of disciplinary non-specificity in all four disciplines. In fact, more than half of all items in the top 100 bundles in each discipline occur in the top 100 of any other discipline.

基于语料库的《供应家具的房间》小说文体特征

周红莉 东华大学

欧·亨利是美国最杰出的短篇小说家之一，其作品风趣幽默，构思精巧，生动描写了美国人民的生活尤其是下层小人物的生活状态，被誉为“美国生活的幽默百科全书”。他的大多数作品体现了作家对人性之善和美好社会的强烈向往与追求。尽管不乏对黑暗冷漠社会的抨击和讽刺，但却立足于人性的光辉，歌颂小人物在贫困生存中患难与共、相濡以沫的崇高品格，在夸张、嘲讽、风趣、诙谐的幽默之中，饱含凄楚的情绪，充满对美国社会千百万小人物的极大同情，形成其独特的“含泪的微笑”的艺术风格。

然而，小说《供应家具出租的房间》无论从主题思想还是写作手法上，都堪称欧·亨利风格变异的代表作。作者用细腻、娴熟的笔触描写了纽约下西区一个

供应家具的房间，用神秘的气氛渲染了一对爱人先后在同一房间里自杀的悲剧。小说一扫往日作品的调侃与幽默风格，笔调更加酸楚凄凉，有的只是血与泪的奔流，痛苦地呐喊与控诉，表现了欧·亨利小说鲜有的悲怆的风格。这种变异，是“欧·亨利手法”同中见异的变奏，使欧·亨利的创作风格更加奇特，更加典型。正因如此，这部小说才具有很高的思想价值和艺术价值，被列为欧·亨利最佳短篇小说之一。

以往的研究基于传统的直觉分析，主要集中于对小说作品进行定性研究。随着语料库语言学在国内的迅速发展，越来越多的研究者尝试着将语料库语言学与文学相结合，利用语料库收集的语料对文本进行量化分析，从而开拓全新的文学研究视角。本研究拟借助语料库研究方法，运用语料库检索工具 AntConc 软件提取数据，探讨小说的文体特征，以期对小说有更深一步的理解与认识。

Structural Distribution of Metadiscourse in English Major's Argumentative Essays

Yushan Zhu Shanghai International Studies University

By analyzing the sample essays from the Corpus of English Writing constructed by Shanghai International Studies University, this research takes a corpus-based comparative approach to explore the structural patterns and variations of metadiscourse use in argumentative essays written by English majors at some Chinese universities, following the Interpersonal Model of Metadiscourse proposed by Hyland (2005). Previous studies on metadiscourse use in student writers' texts in the ESL/EFL context focus primarily on published research articles, whereas less effort has been put into the student learners' assignment essays in language classrooms (Crismore et al. 1993; Bloch & Chi 1995; Moreno 1997; Moreno 2004; Milne 2003; Hyland 2004; Gao, 2009, etc.). Moreover, the types of student assignment essays investigated are restricted to metadiscourse use in the essays as a whole (Intaraprawat and Steffensen 1995; Xu 2001; Adel 2006; Cao, 2009, etc.), while very few studies focus on the structural distribution of metadiscourse resource in the separate parts of essays, namely, the introduction, the body and the conclusion. This essay probes into the structural patterns and variations of metadiscourse distribution in argumentative essays.

Research results unveil that metadiscourse resources as a whole are used consistently most often in the conclusions and least in the body paragraphs. The distributions of attitude markers and engagement markers follow this pattern, while other subcategories do not. Hedges occur with higher frequencies across the three parts of essays but their structural distributions do not display any significant differences. The frequencies of other subcategories present significant differences across the three parts, with endophoric markers, evidentials, code glosses and self mentions occurring most often in the introductions and other five subcategories in the conclusions. Moreover, more structure-dependent features are also found with regard to the distribution of certain types of the subcategories. It is argued that the structural distribution patterns and variations of metadiscourse may be attributed to the discourse functions of each part of the essay.

中外运动医学学术论文英文摘要动词使用的对比研究

崔珣丽, 陈杰 北京体育大学

动词是构成句子和篇章的最重要元素之一 (Quirk et al., 1995)。除了时态和语态功能, 动词的含义和其在文本中的作用也是学术写作中需要注意的关键问题 (Hinkel, 2004)。在体育领域, 中国人的运动成绩已令世界瞩目, 然而在运动科学方面的研究成果——至少在国际期刊发表论文方面——仍与发达国家水平差距甚殊。目的: 本文旨在通过自建语料库, 从动词使用的角度研究国内运动医学核心期刊论文英文摘要和国际相应领域学术论文摘要的差异, 并加以分析。方法: 选取国内外运动医学领域具有代表性的期刊的英文摘要组建语料库。国内期刊选取了《运动医学杂志》(2011—2012 年间摘要 248 篇, 64,732 词), 国际期刊选取了《美国运动医学杂志》(2010—2011 年, 558 篇, 179,261 词) 和《英国运动医学杂志》(2010—2011 年, 253 篇, 62,079 词)。清理后的语料用 TreeTagger 标注, 之后用 PowerGREP 统计三种期刊摘要中动词词频, 对位于前 30 的动词进行对比分析, 并通过 Antconc 和 BFSU Collocator 软件观察这些动词的使用情况。运用卡方检验差异的显著性。结果: 中外期刊中动词使用情况差异较大, 除了像 compared, showed, measured 等少数几个动词使用频率无显著差异, 与国外学术论文相比, 国内论文英语摘要动词存在显著过度使用或使用不足。对于像 increased, increasing, decreased, divided 这样的动词使用过度体现出国内作者词汇量不够丰富、倾向使用简单的熟词; 而像 appeared 一词的误用显示中国作者对于学术研究结果没有表现出应有的信心。Study 做动词时, 国内论文摘要中都是作为原形动词出现, 用于阐述研究目的; 而在国外论文摘要中都是以被动语态出现, 用来说明研究对象。Have 和 has 在国际论文摘要中的大量使用和在国内论文摘要中的缺失, 揭示了国内论文英文摘要没有“背景”这一语步。结论: 中外运动医学类学术论文常用的高频动词中, 仅有少数动词使用情况相似, 而更多地显示出惊人的差异。除了语步缺失的原因, 中国作者词汇量不够丰富、对词语的语义搭配掌握不足以及在学术写作中缺乏训练从而信心不足也是产生现状的原因。

基于 MICUSP 的学术英语写作教学模式研究

吴雪颖 上海大学

MICUSP (密歇根学术英语写作语料库) 是目前世界上最方便教学的写作语料库之一, 所搜集的语料涉及 16 门学科, 涵盖人文艺术、社会科学、生物学及物理学四大学科领域, 是进行学术英语写作教学的重要媒介。本文在研究该语料库教学实用特色的基础上, 构建基于 MICUSP 的学术英语写作教学模式, 探讨如何利用该语料库提供的模型进行基于内容的跨学科写作教学, 同时重点针对中国学生写作水平发展的认知规律和语言知识习得的难点进行训练。

A comparative study of abstracts in medical journals from China and the USA/UK

Junfeng Jiao Zhengjiang university of sciences

English-language abstracts are required not only by medical journals in majority English-speaking countries such as the USA and the UK, but also in journals published in China. By analysing a corpus of 100 abstracts in English from seven leading Chinese medical journals and 100 abstracts from five leading American and British medical journals, this study examines cross-cultural differences and linguistic variation between the two groups of abstracts. An approach that combines genre and corpus analysis is used to establish salient differences. The results suggest that Chinese authors are more aware of the standard macro structure of abstracts, and that the two groups demonstrate variation in their use of voice, tenses, finite verb forms, lexical hedging and certain collocations. The study also considers the implications of these differences.

Corpus Construction and Software Development

Incorporating Morphological Features for the Automatic Multi-Word Term Extraction

Deng Yaochen Dalian Maritime University

From the linguistic dimension of terminology, the present paper proposes a new strategy for automatic multi-word term extraction by incorporating the morphological features into the current approaches. The experiments indicate that the incorporation of morphological features into the rule-based approach is robust in improving the efficiency of automatic term extraction. The score of morphological features is effective in distinguishing terms from non-terms. The study not only reveals the significance of linguistic knowledge in automatic term extraction but also provides strong support for the linguistics-underlain paradigm of natural language engineering.

Implementation of Parallel Corpus Involving 'Phags-pa and Uyghur Mongolian Scripts

Jin Gang Inner Mongolia Normal University

The Mongolian Nationality has used different scripts, such as Uyghur Mongolian script, 'Phags-pa script, Ali-gali letters, Clear script and Soyombo alphabet, for writing and recording. Documents written in these scripts have offered valuable information for studying Mongolian, with all kinds of issues regarding to Mongolian linguistics being able to be explained through comparison among them. In classic literature research, documents are usually interpreted in such a way that the unknown

sections are interpreted according to the unscrambled parts, which is the predecessor of now parallel-corpus based research. This approach has achieved great results in such studies as language comparison, disambiguation and translation.

In this research, with the aim to provide assistance to studies on documents written in scripts above, a parallel corpus involving 'Phags-pa and Uyghur Mongolian scripts is built with XML format and the corresponding retrieval platform is designed, with the rules for using TEI being followed. The connection between the corpora of the two scripts is established by annotating 'Phags-pa and Uyghur Mongolian scripts with the equivalent Mongolian lemmas now in use. Besides, complex retrieval through lemma, morphological information and word position, as well as regular expression based retrieving pattern, has also been implemented, with notes helpful to Mongolian literature study being added.

The application results shows that XML formatted linguistic data of Mongolian documents in different scripts and Mongolian lemma annotations have played a beneficial role in the retrieval of different kinds of scripts within the same operation platform. The approach, which is greatly favorable for comparison study, can also be extendedly used in the establishment of diachronic corpora involving different scripts and the corresponding application platform.

韩汉平行语料库的设计及其在翻译研究中的应用

——以韩汉信息不对称成语翻译为例

金菊花 山东大学外国语学院

平行语料库作为语料库的一个分支,越来越受到人们的重视,它不仅用于语言学研究,在语言工程领域也得到广泛的应用,推动了语言学理论研究和实用语言工程领域研究的结合。本文以信息不对称韩汉成语翻译研究为例,介绍用于韩汉翻译研究的平行语料库的收集、标注、对齐和加工,并以成语翻译个案为例介绍具体的研究步骤,帮助了解平行语料库的设计与从中提取研究对象的相关过程及相关数据的统计分析。

作为抽样研究的研究对象,我们选择了成语。成语是各种语言中最具民族特色而又普遍存在的“文化化石”。由于不同民族文化背景之间的“信息不对称”,使得不同语言之间成语的翻译成为一大瓶颈。韩国历史上曾经长期使用汉字,汉字词至今仍占韩语词汇的60%以上,其中包括汉字成语。但是韩国语中仍然存在大量自造的汉字成语和固有的成语,因此韩汉成语的翻译也存在信息不对称的困扰。可见,本研究有一定的针对性和研究意义。

论文由三部分构成。第一部分,主要介绍自建平行语料库的设计,内容涉及语料库的创建目的、语料构成、加工对齐等。第二部分,以韩汉信息不对称成语为例,统计从自建平行语料库中提取的数据,并对此进行分类。第三部分,基于数据统计和分析,结合以往研究,探讨这一方法对韩汉信息不对称成语翻译策略研究的意义。

本研究采用平行语料库方法,抽样建立规范的韩汉双语平行语料库,以权威翻译家的作品为样本,对成语等进行人工标注,形成信息不对称成语翻译的样板库,从中挖掘信息不对称成语的翻译类型,以便对其翻译原则和方法进行进一步

分析。为确保语料的代表性、权威性、口语性，韩汉平行语料库由口语性较强的非政论文语料构成，主要包括《高丽亚娜》杂志、大众报道、励志小说、韩语教材等四大类。其中，汉语语料 105.3 万字，韩语语料 379,550 词。全部加工标注成平行附码语料库，在此基础上，区分双语对应、不对称、半对称等不同类型，进行定量分析，探讨信息不对称成语的翻译类型。

应对突发公共事件舆情的官方话语语料库建设

毛浩然 福建师范大学

应对突发公共事件舆情的官方话语研究语料库子类别主要有：酒店类（消防、刑事案件、食品安全）、交通安全类（公交、动车、火车、飞机）、地震类、考试类（高考、考研、四六级等）、公共卫生类（疾控与食品医药安全）、安全生产类（矿难、森林火灾、旅游类）、工程质量类（建筑倒塌、桥梁断裂）、环保类等。官方话语指官方在突发公共事件中回应社会舆情的公关话语。语料来源包括电视、官方网站、官方微博等媒体上的官员表态、新闻发言人话语、案情通报、事故调查，原因披露等语料。课题主要研究问题是：在记者招待会、政府网站、电视台或政务微博等传播平台，政府、官员或新闻发言人的官方话语应如何回应舆情才能有效化解突发公共事件中的舆情危机，以及如何对官方话语效果进行及时客观的评价。本语料库结合舆情危机处理 5S 原则（承担责任 *Shoulder the matter*，真诚沟通 *Sincerity*，速度第一 *Speed*，系统运行 *System*，权威证实 *Standard*），参照舆情漏斗模型和危机管理模型，从热点事件追踪、社会心理分析、舆情主题深化、正负情感评价，意见领袖分析、舆情危机预警、话语修辞艺术等要素入手，采集包括 2002~2016 年 15 年间突发公共事件中舆情应对的国内官方话语案例语料（占 80%）和部分国际官方话语成功案例语料（占 20%），以新闻发言人话语和政务微博话语为语料来源重点，借助语音识别软件和语料库标注软件，语料转写、标注和校对采用软件与人工相结合，以传播者、受众、内容、效果为主要标注元素，对官方话语语料进行自动转写标注和人工校对标注，建成国内首个舆情应对的官方话语小型语料库；借助百度搜索平台和网络舆情监测系统（如：军犬、利盾等），实时收集官方话语，创建官方话语专题网站等多元网络互动平台，借助网络平台和监测软件进行云数据处理、智能提取和语料分析，并借助语料库检索与标注功能对语料进行词性、语义和修辞分析。

专业科技英语语料库的建设与应用

吴蕾，赵晓临，张继东 东华大学

我国英文版科技期刊的整体质量有待提高，有必要建立起专门的相关语料库对专业英语论文写作展开研究。本文以“东华科技英语语料库”中“专业科技英语语料库”的建设为例，从语料来源、命名方式、相关软件工具、文本整理等几个方面描述了建立一个中小型语料库的具体步骤及相关软件的操作说明。文中一张表格的数据展示了本语料库中不同语类的采样篇数和所占比例，另一张表格的数据展示了各学科语篇的篇数及标准化类符形符比。最后，文章通过其中“环境语料子库”里的具体单词“insect”与“insects”的搭配差异说明建立该语料库对

科研人员撰写规范的学术英语论文具有重要的参考价值。

型式语法的形式化探索

陈功 对外经济贸易大学

型式语法 (Pattern Grammar) 是 Hunston & Francis (2000) 在 COBUILD 语法研究成果的基础上所作的深入理论探讨, 它不仅是 COBUILD 项目语法成果的全面总结, 也是当前描写语法理论研究的典型代表, “开创了语法描写的新途径” (Owen 1993)。作为第一部语料库驱动的英语描写语法, 型式语法可谓是开创性的, 它颠覆了以往语言学家看待语言、获取语言知识的传统, 给语言研究者展现出了一个全新的视角。

近年来, 随着自然语言处理界对语料库的不断重视, 型式语法也开始受到愈来愈多的关注。型式语法的核心思想是“语法-词汇不可分”、“结构-意义相联系”, 而形式化体系在发展过程中一直都忽略掉了这个问题 (Gamallo Otero & González López 2011)。词汇-语法分离、结构-意义分离的后果就是计算机无法对自然语言进行更加准确的分析和处理, 无法对搭配、习语等占据了语言相当大部分的固定或半固定短语的结构和意义做出积极的反馈。

冯志伟 (2010: 2) 指出, 计算机对自然语言的研究和处理, 一般应经过四个方面的过程: 形式化、算法化、程序化和实用化。为了更好地将型式语法应用于自然语言处理, 本研究将在形式化方面进行努力, 在已有形式化体系“链语法” (Link Grammar) (Sleator & Temperley 1991) 的基础上, 对型式语法的形式化进行探索, 旨在找到将型式语法的语言学思想转化为计算机规则的方法。

面向教学的报刊英语语料库的创建及其应用研究

邓飞, 李元科 华南师范大学

本研究运用语料库语言学的“教学加工”理念及其研究方法, 从教学资源 and 教学手段两方面入手对报刊英语教学进行改革, 目的是改变传统的报刊英语教与学的手段和模式, 提供基于语料库的受频数驱动的、以学习为中心的教学模式和自主性的、互动式的、开放型的、分析归类型的学习模式, 从而帮助学生快速掌握报刊英语的语言特征, 提高外报外刊阅读能力, 提高跨文化交际能力。具体内容包括:

1) 创建 300 万词次的报刊英语语料库, 选材包括美国和英国发行量较大的英语报纸, 它们分别是《纽约时报》、《华盛顿邮报》、《波士顿环球报》、《芝加哥论坛报》、《独立报》、《卫报》、《泰晤士报》、《经济学人》。本课题综合 Biber (1988) 和张健 (2007) 对新闻域的划分标准, 把报刊新闻分为新闻评论和纯新闻报道两大类, 纯新闻报道又分为财经新闻, 体育新闻, 科技新闻, 法律新闻, 政治新闻、社会新闻和文化新闻等七类。报刊英语语料库分设美国和英国两个分库, 两个分库又按照新闻域分层为不同的子库。

2) 在课堂上把语料库手段转化为师生能够接受的教学形式, 如“建立微本 (2003, SINCLAIR) ----编辑加工----教师给出指令和问题----学生按指令观察、分小组讨论问题----学生总结归纳----教师总结”, 引导学生借助语料库检索工具 AntConc

的语境共现行，多字词表，词频表，关键词表，搭配词表等，观察、分析和归纳报刊英语甚至报刊英语不同新闻域的词汇特点，语篇特点，帮助学生快速掌握报刊英语的常用词，提高外报外刊的阅读能力。此外，引导他们了解报刊英语的写作风格，熟悉报刊文章的文体特征，了解英语国家的社会、历史、生活、文化等背景知识，培养学生的跨文化交际意识。

基于 R-gram 的语料库分析软件 PowerConc 的设计与开发

贾云龙 新东方教育科技集团 许家金 北京外国语大学

在继承以往语料库分析软件优点的基础上，本研究开发了具有独立知识产权的 PowerConc 语料库分析工具。PowerConc 对传统的词汇索引、词表生成、主题词计算等功能进行了重构、扩展和优化。整个软件以基于正则表达式 (regular expressions) 的 N 元组 (N-gram) 为基础。二者的有机结合即本文所提出的 R-gram。R-gram 这一概念大大增强了检索和匹配的灵活性。同时我们设计了兼容正则表达式的简易输入语法——Smart Input，降低了用户使用的难度，提高了软件的易用性。PowerConc 软件基于面向对象的思想开发，核心功能被封装在不同的类中，与界面分离，具有很好的扩展性和可维护性。PowerConc 的开发将有效促进语料库语言学研究的开展。

以地质英语语料库为抓手，促进地学领域专门用途英语的学科建设

张翼翼，苗永胜 中国地质大学

语料库是为专门目的按照明确的设计标准收集的文章的集合 (Granger, 1998)，不仅对应用语言学在词汇学、翻译、语言教学等方面的研究有促进作用，还对语言比较和机器翻译起到了重大的推动作用 (王建新, 2005)。

计算机技术空前迅速的发展，使包含广泛自然语料的各种语料库得以建立。然而在地学领域，目前尚缺乏较为系统、完整的地质英语语料库，用于在教学中指导学生练习相关文献的翻译和学术论文的写作。地质英语语料库是以地学领域的英语教学和教材编写为目的收集的文章集合，取样的文本在最大程度上应代表地学环境下的英语语言及变体。作为一种专用型语料库，地质英语语料库含有大量的地学信息 (张翼翼, 2008)，可以辅助教师提高地质英语的课堂教学效率，加强学生在专业背景下使用专门用途英语的能力，能够促进新兴的语料库语言学为地质英语的学科建设服务。

完善现有的语料库，一方面应不断补充、更新相关语料，另一方面还要利用标记软件改进词汇和句子的标注方式，以此加强语料库的“兼容性”，使其能够识别更为复杂的句式，才能从根本上提高基于语料库的机器翻译质量，并为地质英语的课堂教学搭建良好的平台。

基于语料库的《军事外宣翻译词典》编纂研究

赵志敏 解放军外国语学院

《军事外宣翻译词典》是基于“军事外宣平行语料库”生成的词典，依照 Sinclair 有关“一个搭配即一个意义单位”的理念和“语料库驱动”的原则，从军事外宣平行语料库中提取对应单位，经过人工干预后成为词典中的词条及其对应的翻译。此外还构建了军事外宣可比语料库与监控语料库，作为平行语料库的有益补充。前者搜集英美国家的军事宣传语料作参考，用来评估对应单位的翻译信度与接受效果；后者依托网络技术及时捕捉新的对应单位，以便及时更新词典力求与时俱进。最终产品以光盘和网络为载体，将词典编纂所依赖的语料库也内嵌其中并提供“索引”功能。该词典的编纂旨在为军事外宣的翻译研究与翻译实践提供一定的参考和帮助。

Corpus and Translation

Translation Universals in the Translation of Motion Events—A Corpus-Based Investigation

Xiaochen Li Beihan University

The stylistic features of translated texts have long been discussed in corpus-based translation studies. Yet few consensuses have been reached regarding the nature and category of the unique stylistic features, i.e. translation universals (Baker 1993). A group of such features have been identified in translated texts of different languages, such as explicitness and implicitness. Yet previous research mainly works on text level, while this research goes a step further and analyzes the manifestation of translation universals in the translation of motion events, i.e. a group of phrases which describe the motion of an object. These phrases are chosen because they have very similar patterns of form-meaning association in Chinese and English, and the translated texts are highly comparable to the original ones. This research takes Event semantics (Talmy 2000) as its theoretical framework and adopts a corpus-informed method. By analyzing the data retrieved from TED corpus, it is discovered that explicitness and implicitness are stable features of translated texts and can be identified even in the translation of phrases with almost identical form-meaning associations in the two languages. Moreover, explicitness and implicitness, though apparently in contrast to each other, have similar discursal functions. Both of them are used to bridge the gaps between two languages and enhance the acceptability of the translated text. It is hoped that this paper will offer a more detailed account of the translation universals by combining the insights from translation studies, corpus linguistics, and cognitive semantics.

基于《水浒传》一本四译汉英平行语料库对晋词翻译的研究

刘克强 云南红河学院

本文以赛珍珠、杰克逊、沙博理及登特-杨译《水浒传》与原文建立的一对四句对齐汉英平行语料库的基础上，对詈词“厮”及“鸟”进行对比研究，结果发现：在詈词“厮”的翻译中，赛译基本上呈中性；杰译也基本呈中性，部分未译，因而赛译与杰译有“净化”的趋向；与赛译及杰译相反的是，沙译与登译往往使用含有感情色彩较浓、有明显谩骂的词汇来表达“厮”的内涵意义。相比于沙译，登译的数量更加丰富，词汇的贬义或蔑视的程度更加强烈。“鸟”的翻译对比研究发现赛译单一，杰译大多略去不译，沙译和登译丰富多样，在大多数场合中主要用来表示说话者辱骂、愤怒、怨恨和鄙视等感情或是表示说话者诅咒、厌烦、惊讶和强调等语气。需要指出的是，在中国软件实力不断增强及中华文化走向世界的背景下，探讨蕴藏于典籍中的中国文化的翻译具有重要的现实意义。

英语关系从句及其汉译的语法结构不对应研究

刘磊 北京外国语大学

关系从句(简称 RC)由中心名词短语(简称 NPmat)和限定小句(简称 Srel)两部分构成，后者的作用是界定前者的所指范围，并指定关系化名词短语(简称 NPrel)的语法功能(Andrews 2007: 206; 许余龙 2011: 645)。英语 RC 中的 Srel 置于 NPmat 之后，由关系代词引导；而汉语 RC 中的 Srel 置于 NPmat 之前，两者之间通常插入关系化标记“的”(Li & Thompson 1981: 579)。

本研究旨在考察英语 RC 译成汉语时，两类英语名词短语 NPmat 和 NPrel 的语法结构(如主语、宾语、谓语和旁语等)是如何在汉语译文中产生变化的。研究基于英-汉平行语料库 Babel，首先提取英语源语中的 RC 及其汉译文本，获得平行句对；其次制定标注体系手工标注源语和目标语中 NPmat 和 NPrel 的语法结构。本研究发现大约 50% 的英语 RC 译成汉语时，会产生语法结构的不对应现象，具体的不对应类别及其产生的语境如下：(一) 英语 RC 译为汉语中的“名词短语+的+名词短语”结构，如将“the wounds that comes with life”译为“生活中的伤痛”；(二) 英语 RC 译为汉语 RC，但汉语中 NPrel 的语法关系发生变化，如将“the challenges that face the continent”译为“大陆所面临的挑战”。其中 NPrel 由主语变为宾语；(三) 英语 RC 译为汉语中的兼语，如将“an Africa that is integrated in the world on and equal basis”译为“使非洲平等地融入世界”；(四) 英语 RC 译为汉语中的一个小句，如将“Ours is [a rebirth that must deal with problems] that...”译为“我们的复兴必须解决一系列问题”。其中 NPmat 由谓语变为主语；(五) 英语 RC 译为汉语中的两个小句，如将“listening to [songs of longing and loss that seemed to speak directly to us]”译为“聆听一些关于渴盼与失落的歌曲，那些歌曲好像正是唱给我们听的。”其中 NPmat 在目标语中重复使用，成为第一个小句的宾语，第二个小句的主语。进一步分析语料发现，产生不对应现象的一个重要原因是英语 NPmat 结构的内部复杂度：当 NPmat 由“N+of+N+RC”或“ADJ+N+RC”等复杂结构构成时，汉语译文更倾向于产生语法结构的变化。与之前的英语关系从句的汉译研究(李映霞 1990)相比，本文细化了英语关系从句汉译过程中语法结构不对应的种类及其频率分布，并对产生不对应的原因进行了探讨。

意识形态操控与动态读者取位：基于语料库与评价理论的切尼演讲二汉语译文分析

刘泽权，刘鼎甲 燕山大学

本文以2004年4月15日美国前副总统切尼在复旦大学发表的演讲及其美国和中国官方中文译本为语料，探讨译者与读者的主体间角色及其在翻译过程中所使用的人际策略、动态读者取位方法与意识形态的操控。切尼的演讲及其随后的答听众问，涉及到中国的政治经济改革、台海关系、朝鲜核问题等，并多次影射我国的政治现状，反复强调美国固守“台湾关系法”的立场。毫无疑问，这些言论富于敏感性，其中有些涉及到我国的主权和中美关系的基础。因此，我国官方报纸《人民日报》的门户网站《人民网》在刊载切尼演讲和答听众问的中译“全文”时，采用了“改编的手法”，这与美国国务院国家翻译局所发布的中文译文有着显著的差异。

本文综合运用系统功能语言学的评价资源分析工具，在批判性话语分析的框架下，分析比较了演讲原文、美国官方译文和中国官方译文中显式与隐式评价资源的运用，并结合相关翻译技巧的探讨，试图揭示作者、译者如何通过对意识形态的操控与评价模式的选择实现翻译过程中作者人际策略的选择与动态读者取位的支配。除此之外，本文建立了切尼演讲的英文转写文本及其两个汉语译本的平行语料库，分别使用单语与平行语料库检索工具检索分析文本中的高频词、词汇搭配、词丛、核心情感词汇的语义韵与翻译对应关系，试图讨论译者在话语与翻译过程中采用的人际策略。为验证评价理论分析结果的有效性，本文对比了评价分析和语料库分析的结果，发现评价分析的结果与基于语料库分析的结果不但在一定程度上互相吻合，而且互为补充，二者的结合使用可以更加全面地揭示翻译过程中作者、译者的主体间角色（“这里的主体间角色指作者或译者在发话或者翻译的过程中所采取的跟读者交际的策略以及期望达到的预设的人际功能所采用的方法”）、动态读者取位策略和意识形态操控行为。本文据此也讨论了语料库和评价理论的结合作为一种新的研究范式运用于翻译研究的可行性。

“X/XY+感”构式及英语对等研究

种法娟 青岛农业大学

本文主要研究现代汉语名词性“X/XY+感”构式中“感”的语义类别及其常用英语对等形式；本文首先通过北京大学现代汉语语料库数据统计和分析，将“X/XY+感”构式中“感”的语义细化为①“感1”：表感觉，本语义可以再细化为“感1-1”：自有的感觉；“感1-2”：别人给自己的印象；“感1-3”：别人使自己具有的感受；②“感2”：语义为感知能力，例如“方向感”；③“感3”：语义为意识。其次，通过Ant Conc英汉平行语料库，对“X/XY+感1”、“X/XY+感2”和“X/XY+感3”的具体例式进行检索、统计和对等研究，统计分析“感1”、“感2”和“感3”常用的英语对等形式。最后本文研究了本构式的句法功能、否定形式；构式内部成分“X”、“XY”的性质特征，包括音节特征、词类分布、“X”、“XY”与“感”的结构关系、“X”、“XY”进入本构式的结构限制和类型限制；本构式的语用价值、能产性表现及其动因。

基于语料库方法的《红楼梦》“社会脸”翻译研究

谭业升 上海外国语大学

本文基于语料库方法对《红楼梦》中具有社会义的“脸面”表达（即“社会脸”表达）的分布模式以及在两个英译文中的对应实现情况进行了分析和考察。并以此为基础基于意象图式、隐喻这两个认知语言学理论概念对“脸”意象图式在《红楼梦》及其两个译本中的例示情况进行了探讨。之后，基于社会义的“face”表达在两个英语本族语语料库 BNC 和 COCA 的分布情况的对比，对两个译本中的 face 表达的认知来源进行了探讨和推测。文章的研究表明：图式-例示的级阶可以很好地揭示和解释语言对比和翻译所涉及的语言文化之间的“同中有异，异中有同”关系，以及以此为基础的翻译认知策略。一方面，基于共性涉身体验的意象图式在翻译中可以在不同的意象维度上做隐喻性或转喻性拓展例示，并呈现语言偏好的社会关系概念化方式的差异；另一方面，基于共享的意象图式表征和可通约的涉身体验，两种语言偏好的概念化方式可以在翻译中进行合成，而新的译语概念组织方式和表达形式是合成概念化的结果。译者在此创造性的翻译过程中也表现出个性化的特征。这一认识，对于理解翻译的认知创造性具有重要的启示和意义。“社会脸”翻译的例证说明，翻译中的会通不是简单的激活，而更是两种语言表征空间互动和合成的过程。

本文先后通过翻译文本的对比和本族语语料库的对比验证，获得了有关语言特有概念化方式和译者创造性翻译过程的新认识。此种研究表明，以翻译语料库所揭示的问题（主要是对系统差异性的期望或假设）为引子，充分利用其它种类的语料库或其它实证研究方法加以综合对比和验证，开展认知研究的方法，有望给我们带来更多有关语言认知共性和个性关系的更具体、更深入的认识。

语料库翻译研究：基于语料库还是语料库驱动

陈昕 解放军外国语学院

Tognini-Bonelli (2001) 区分了“基于语料库”和“语料库驱动”两种方法论途径，后者又发展为“语料库驱动语言学”。Sinclair 认同这种区分 (2003)。两者的本质区别在于前者使用语料库数据验证已存在的理论假设，后者则以语料库数据为证据提出新的理论假设或挑战已有的理论假设。作为 Sinclair 的学生和第一个倡导把语料库应用于翻译研究的学者，Mona Baker (1993) 并没有明确区分翻译研究是“基于语料库”还是“语料库驱动”。Baker 提出使用语料库对翻译共性进行研究实质上是“基于语料库”的，因其目的是为了验证翻译文本是具有普遍性特征的、不同于源语文本与译语原创文本的“第三类语码”。但 Baker 也指出，某些翻译规范只在特定社会文化语境中发生作用，而语料库能够为研究不同社会文化、历史背景下的各种“操作规范”提供有力的数据支撑。从这一点来看，Baker 的语料库研究思想又含有“语料库驱动”的成分。

本文通过统计国内外语料库翻译研究成果（期刊论文、论文集、专著、科研项目等）题名中的关键字“基于语料库”或“语料库驱动”并梳理其研究内容和结果，得出结论：目前的语料库翻译研究既有“基于语料库”来验证已有假设的情形，也有“语料库驱动”来揭示新的翻译现象和规律的情形，但翻译研究学界对于这两种

方法论的概念仍存在混淆和误用,绝大多数翻译研究者在研究中并不区分这两种方法论,“基于语料库”的使用居多,“语料库驱动”有时被等同于“基于语料库”,还有“语料库辅助”的提法。诚然,任何研究都可能含有基于语料库的成分,也可能有语料库驱动的成分,只是程度不同而已。或许对这两种方法论进行区分的必要性在翻译研究学界尚未凸显。但语料库研究的主流方法是数据驱动,“语料库驱动”的翻译研究是有望发现新的翻译事实和机制的有效手段,现有假设要尽量少用,否则语料库翻译研究可能陷入 Snell-Hornby(2006)所言的“U 形转向”——“重新发明轮子”的窘境之中。因此,有必要对此正名,以使翻译研究者厘清两种途径的根本区别并在未来的研究中正确使用。

基于语料库的中国古诗英译研究路线图探索

黄培希 东华大学

本文选取国内曾经入选南大和北大外语类核心期刊自创刊以来历年刊登的古诗英译研究文章建立语料库,分别按照期刊种类、年份、篇名、主题、关键词、作者、摘要、全文、参考文献、文献来源、以及基金支持等进行设置,对数据进行统计和分析,以此展现建国以来中国古诗英译研究的详细情况。通过研究,我们将明晰中国古诗英译文章在外语类核心期刊(重要期刊)中的发表情况;文章在各种期刊中的数量分布、年份变化和作者分布;几十年来的研究主题及其变化;参考文献的出现频率和引用情况,以及各类文章所获得基金支持等。文章在对数据进行对比、描述各种变化的同时,运用翻译研究理论并对变化的原因进行剖析。

试析语料库对翻译理论和翻译实践的促进作用

贾艳 西北工业大学

如今,语料库与语料库语言学已经不再是新奇的学术名词。近 30 年来,基于语料库的各种研究发展迅猛,比如:二语习得、对比语言学、词典编纂等。语料库的应用极大地拓宽了应用语言学的视野,为外语教学研究提供了一种新的哲学思维模式。语料库与翻译研究的结合比其它领域来得晚些,语料库翻译研究是 20 世纪 90 年代兴起的全新的翻译研究范式。宏观意义上的语料库翻译研究代表了翻译研究中的一种新兴模式,它在研究手段、研究理念及研究目的上均有一定的特色,并在一定程度上推动了翻译研究的进程。语料库不仅为纯翻译研究提供了有力的工具,而且在应用翻译研究中也大有作为。语料库的开发和利用,为进一步开展翻译研究提供了新的方法论基础,即以真实语料为研究对象,以概率统计为研究手段,通过考察语言间的相互作用及文本内外的各种相关因素,对翻译过程及其结果进行解释和评价。翻译语料库的建立和使用不仅可以摆脱传统翻译研究的局限,更新翻译研究的思路和方法,同时亦为翻译实践提供了便利。但迄今为止,基于语料库的翻译研究多以相关事实的量化分析和描写为主,而对其背后内在动因和机制解释得较少,所以只有把语料库技术手段与文本生成过程联系起来,从社会、文化、历史等多视角对量化研究结果进行探究,才能完整地勾勒出翻译现象的概貌。本文探讨了相关语料库对翻译理论和实践(应用)研究所起的促进作用。

中医学概念隐喻词源、文化与翻译——基于语料库的探索和方法论反思

兰凤利 上海中医药大学

无论在中国还是在西方，隐喻研究均早已超越了诗学、修辞学、语言学的局限，成为语言哲学、科学哲学、认知哲学所关注的重要命题。隐喻是沟通东西方文化与科学、促进彼此相互理解的重要桥梁。

取象比类是中医学的核心方法论。“观物-取象-比类-体道”贯穿于中医学基本概念的形成、医理阐述和临床发展等各个方面，体现着“象-比-喻”的思维过程，并最终形成中医学中的隐喻、指导着临床实践、促进着中医学的发展。究竟如何翻译中医学中的概念隐喻——是保留概念的原始隐喻内涵，还是仅传达其医学价值？这是所有关注中医对外译介、传播与交流的学者，需要共同思考解决的问题，也是中医学跨文化传播的关键和难点之一。利用语料库的方法和技术，基于隐喻认知探索、研究中医学概念隐喻的词源、文化内涵及其翻译原则、策略与方法，具有现实的迫切性和独特的语言学、文化学与医学价值。

语料库语言学是语言学研究科学化的产物，旨在基于客观、量化的数据，分析、总结出一定的语言学规律。语料库语言学、基于语料库的翻译研究，在国内外语文学界日趋流行，为语言学研究带来重大而深远的影响。基于语料库的翻译研究，是按照尊重自然语言与自然翻译，尊重语言事实与翻译事实进行的，即：“好”的翻译会被不断地重复。何为“好”的翻译？当某一隐喻的翻译，基于隐喻认知的翻译结论和基于语料库频次的翻译结论发生矛盾时，到底应该采用、推广哪一个？基于语料库和基于隐喻认知的翻译研究，就像东、西方两种不同的科学途径——一个注重哲学思辨，一个注重量化分析，从不同视角了解同一事物，丰富了对这一事物的认识——语料库中出现频次最高的翻译就是当下最流行的，基于隐喻认知的翻译代表着将来的发展趋势，二者相互补充，与时代和文化背景相结合，共同阐释着多元文化背景中隐喻丰富的外延和内涵。

On Factor Balancing and Strategy Adopting in Translating

Zuoshuai Lin Weiyan Shen Donghua University

This study purports to discuss the question of factor balancing and strategy adopting in translating, through a contrastive analysis of the two translated versions of “Liu Lao Lao Yi Jin Rong Guo Fu”, the 6th episode of Hong Lou Meng by Xueqin Cao.

Xueqin Cao (about 1715-1763) is a very prestigious Chinese novelist, and his work Hong Lou Meng is regarded as an encyclopedia of ancient Chinese culture. Since its publication, this novel has aroused sensations for several times among Chinese readers, scholars and western sinologists.

As regards the translation of this novel, as yet, there are only two full versions: One, entitled A Dream of Red Mansions, is produced by Xianyi Yang (1915-2009) and his wife Gladys Yang (1919-1999); and the other, entitled The Story of the Stone, is done by David Hawkes (1923-2009) and his son-in-law John Minford (1946-). These two versions showcase quite striking features, and both of them are greatly welcomed by readers and scholars both in and out of China.

This study tries to conduct, with some notions from corporal linguistics, a contrastive analysis of the above-mentioned two versions, in order to find out if there are some significant differences in these two versions, if there are some differences in strategy adopting of the two groups of translators, and what are the underlying causes that lead to their different strategies.

In this research, first, with the assistance of the software Concordance, the length of each version (namely tokens), the total number of words used in each version, the common words shared by the two versions, the unique words adopted by each version, and the degree of difference between the two versions, etc. are to be examined and some assumptions will be made on the basis of the data obtained from the result of the computer-aided analysis. Then, the researcher will trace the location, in a version, of the unique words and closely read either version to find the difference between the two versions. Next, the researcher will judge, according to the result of the close reading of each version, whether the preliminary assumptions can be verified. And last, analysis and discussion will be made by the researcher to infer the factors that might have influences upon the translators' strategy adopting.

On one hand, Hawkes' version is about 1/5 longer than that of Yang, and on the other hand, a few sentences of the original text even have not been treated by Hawkes at all.

Upon close reading, the researcher finds out that Hawkes' version makes some paraphrases, supplies some supplemental information and adopts more cohesive devices. And also the two groups of translators adopt different strategies in treating terms concerned with Chinese culture: Yang adopts the strategy of foreignisation, whereas Hawkes adopts the strategy of domestication.

What are the underlying causes that might lead to the different performances of the two groups of translators? Through the review of related literature, the researcher comes to know that on one hand, the two groups of translators' activities are activated by different mechanisms; and on the other hand, the would-to-be final products of the two groups of readers are intended to meet readers in different socio-cultural contexts.

A study on translation selection of Japanese verb —Taking Japanese-Chinese homograph verb "dahasuru" as an example

Jun Liu Guangxi University

The ambiguity resolution and translation of natural language is still one of difficult problems in machine translation. Because of the ambiguity of natural language, during the translation, a source language word often corresponds to many target language words. Therefore, translation must be selected. Verbs are the core of the sentence, which occupy a dominant position in the sentence. Translation selection of verbs has an important influence on the quality of translation of the entire sentence, which is urgently needed to be resolved in machine translation. There are still some translation errors in the current online free translation software. For the sake of the improvement of machine translation quality, taking Japanese-Chinese homograph

verb "dahasuru" as an example, a translation method based on the collocation statistics of target language was used in a artificial simulated translation test to verify its effectiveness. First, analysis the semantic feature of homographs based on the Japanese-Chinese bilingual dictionary; second, calculate MI-value and T-value with retrieval tool and statistical software to analysis the feature of homographs' collocation; then, sentences extracted from the Japanese corpus will be translated into Chinese to establish a mini parallel corpus, then extract the optional Chinese translation of Japanese-Chinese homograph verb "dahasuru"; last, calculate the co-occurrence probability of the optional Chinese translation of homographs' collocations in the target language Chinese corpus, and selected a reasonable translation based on the maximum probability.

基于语料库方法的英汉人体隐喻对比研究

刘燕 北京外国语大学

本研究以概念隐喻为理论框架，采用语料库分析方法，对比英汉语篇中的人体隐喻。隐喻不仅是一种语言现象，更是人们认识和感知世界的思维方式，其本质是借助熟悉的具体事物去体验周围陌生的抽象事物(Lakoff & Johnson 1980)。隐喻现象的产生主要是基于人们的涉身体验，其中，人体及其部位是人们认知体验的基础和出发点之一。

本研究以人体隐喻（人体词的隐喻化现象）为切入点，研究其在英汉语篇里的分布及语言特征。研究主要围绕三个方面展开：1）人体隐喻在英汉语篇中的总体分布；2）人体词的隐喻意义在英汉语篇中的共性与差异；3）英汉人体隐喻在构词、短语搭配、语义韵等方面的异同。研究中使用的语料是布朗家族语料库的英语单语语料库 CLOB 以及汉语单语语料库 Torch2009(Texts of Recent Chinese 2009)，以上两类语料的取样及语料规模相同，因此可作为类比语料库，从总体上反应英汉两种语言人体词的隐喻化差异。本研究检索的人体词主要有头、眼、唇、耳、鼻、手、腿等，以及相对应的英语形式。对含有人体词的索引行，我们采用 MIPVU 的隐喻识别方法 (Pragglejaz group 2007)，选出包含隐喻义的短语结构，在此基础上归纳其语言特征。

本研究发现英汉两种语言中，一些人体词具有大致相同的隐喻用法，但是其在隐喻意义及搭配结构上表现出细微差异；人体隐喻在小说类、新闻类等语篇题材中出现频率较高；一些共同的隐喻并不仅仅是基于人们的涉身体验，有些是因不同文化理念而产生。

本研究结果证明，概念隐喻理论可作为人体隐喻类型的解释框架，本研究结果同时对中西方思维对比有一定的贡献。

拜伦诗歌《唐璜》汉译本中显化现象的研究

于璐，王均松 西北工业大学

20 世纪 90 年代，语料库语言学快速发展起来。它的发展为翻译学科寻找普遍规则和规律提供了有利的手段。显化的概念首先由 Vinay 和 Darbelnet 提出，

他们将其定义为一种文体翻译策略。随着语料库语言学的兴起, Mona Baker 首先意识到了语料库对于翻译学科的价值和意义, 提出了“翻译普遍特征”的假说。其中, 作为翻译普遍特征假说之一的“显化”受到了十分广泛的关注。

显化又称明晰化或明朗化, 是指译者在目的语文本中将源语文本中隐含的但可以根据具体语境条件推导的信息加以明示。显化可以分为语际显化和语内显化, 前者是源语与目标语文本比较分析的结果, 而后者是基于目标语中翻译文本与原创文本对比研究的结果。

翻译显化现象虽然受到了学者的普遍关注, 但基于汉英语言的语际显化研究在文体类型的选择上仍存在局限。国内的大多数研究多数以小说、杂志以及新闻等为研究语料, 而对诗歌这一题材就行研究的则为数不多。因此本文采用描述性的研究方法, 选取查良铮和朱维基所翻译的《唐璜》汉译文本作为语料, 自建语料库, 对其英译中翻译过程中的人称代词主语的显化特征做定量和定性研究。本文分别对第一人称 (I, we)、第二人称 (you)、第三人称 (he, she, they, it) 在英语原文和对应汉语译文中出现的数量与频次进行统计, 观察译文相较于原文在人称代词主语的显化是否有增加, 从而探索诗歌这一翻译文本类型在人称代词主语层面上是否存在语际显化特征, 并试图探究这一显化特征的动因。

本文希望通过对诗歌这类文本类型的研究扩大显化的验证范围; 通过考察平行语料库中人称代词主语的显化策略, 对翻译中的显化有更加全面更加深刻的认识, 从而为实现等效翻译提供可能。

From “Aggression” to “Green-Growth”

Qing Zheng Honghe University Ning He Nanjing University

George Steiner puts forward the concept of the hermeneutic motion consisting of trust, aggression, incorporation and compensation in his book *After Babel: Aspects of Language and Translation* published in 1975. The author holds that in a mutable shadow of blind “trust”, subjective “aggression” may undermine the meaning out of the source text that leads to an incomplete and off-balance “incorporation,” thus “compensation” is required to the target language text. Zhuang Chuaiyi translated most parts of Steiner’s book in his *Babel* in 1987. In Liao Qiyi’s *Contemporary Translation Studies in the U.K.* in 2001, Chapter Two is of George Steiner. Xu Jun thoroughly analysed Steiner’s four-step hermeneutic motion in his article *On Live Experience and Theoretical Exploration of Translation Process* in 2003. Yang Jing suggests in her *A Complement to Steiner’s Hermeneutic Motion* in 2007 that it is to be complimented in a mutual communication perspective. Xia Tian points out the problems in applying Steiner’s theory as well as its positive effects in his paper *Problems and Methods in Applying Steiner’s Hermeneutic Motion* in 2009.

To the flaw of George Steiner’s four-step hermeneutic motion, the author puts forward four steps of fiction translation: “selection”, “green-growth”, “deverbalization” and “appetization.” “Selection” means to choose good source language text (SLT), because the ultimate value of translation with translator’s time, energy and efforts is directly determined by it. Choice should be made in accordance with SLT’s content and value that valued by various social demands. On the other hand, translator’s

judgement, insight and inclination should have a say in this choice. “Green-growth” refers to the method and process the translator marking the points he/ she suspended or polished during comprehending the SLT and expressing target language text (TLT). This method suggests that the current comprehension or some expression is not perfect, and it should be marked out. This process implies that translator’s thorough comprehension of the SLT and perfect deliverance in the TLT, under the limits of his bi-lingual competence, knowledge, artistic appetite and spiritual quality, are not within his reach, but will be closer and closer through his relentless efforts and continuous improvement. “Deverbalization” refers to the process in which translator overcoming obstacles of source language as delivering a text in target language. In English-Chinese fiction translation, the language obstacles to idiomatic Chinese are mainly Americanized or Britainized Chinese. “Appetization” refers to the process in which translator making relentless efforts to pursue the art of expression so that target language readers will be attracted to his translation. To translate an English fiction into Chinese, translator should revise his translation again and again in order to choose even create better expressions. Meanwhile, translator should be in possession of the courage and persistence to retranslate so that his translation will get a new birth as a phoenix. To peptuate a translation is translator’s permanent pursuit.

A Review of the Trinity of Culture, Language and Translating Concerning the Choice between Domestication and Foreignization

Jie Peng Honghe University

Translating is a complex process that involves much more than a simple switching of languages. Many scholars and translators alike consider language a carrier of culture and that translating is a cross-cultural communication while cultural discrepancies and cultural voids always pose problems to such communicating process. In addressing these issues, translators frequently turn to two major approaches for a solution - domestication and foreignization. However, no matter which approach of these two the translators choose, the trinity of culture, language and translating shall be regarded as the basis for the choice between domestication and foreignization. Starting from a brief review of the history of translation theoretical studies and real practices both in China and western countries, the discussion below will mainly focus on the trinity of culture, language and translating for a better understanding of the choice between domestication and foreignization.

试析语料库的存在价值——从翻译教学谈起

王余余，田建国 西北工业大学

近年来，中国翻译教学取得了一些进步，但这种进步与社会的需求和期望相比，还是有很大差距的。随着中国对外开放与交流的步伐的加快，社会各个领域对翻译人才的需求也大大增加，要求也更高。但是国内翻译教学的方式和内容不能

适合现代社会的需求,培养出来的人才在实际工作、生活中翻译能力并不强。这就容易导致翻译教学与社会需求的供需脱节。随着科学技术的发展,语料库技术已经成为现代译者必须掌握的技能之一。语料库在翻译领域中的应用不仅体现在翻译研究理论方面,而且极大的促进了翻译教学的发展。基于语料库的翻译教学方法是一种行之有效的方法,符合以学生为中心的翻译教学理念,可以弥补传统翻译教学法的不足。

我们可以利用语料库中丰富、真实并且可进行比较分析的翻译语料来作为教学资料,教授翻译模式和知识,从而使翻译学习者受到更专业的训练。获得并证实译文模式和翻译知识,使其更具专业性。

翻译语料库拥有大量真实语料与丰富的分析工具,它强调系统性、科学性及宏观性,使小规模、局限于单一文本的个案研究转向大规模、综合性的研究成为可能。在翻译教学中运用语料库能够提供丰富的完整的译文实例以及连贯的语境,为教师教学及学生学习提供很大的帮助。采用基于语料库的方法进行翻译教学也符合以人为本、注重学生的现代教育理念,学生成为翻译教学的主动参与者和研究者,能调动其学习兴趣和积极性,从而能更好地提高学生们的翻译能力,达到更好的翻译教学效果。

但是,目前语料库在翻译教学中的运用还很有限,大多数师生对其认识不够,不能很好地借助语料库来辅助翻译教学或学习。这需要转变观念。学习现代技术并利用其带来的便利,才能更好地推动翻译教学的改进和提高。

本文简要回顾了基于语料库的翻译研究的发展,并在应用翻译研究的范畴内对语料库在翻译教学中的实用价值进行了有益探索,提出了合理建议。旨在打破传统的翻译教学观念,将教师从翻译知识传授者的身份转换为翻译教学活动的组织者、设计者和指导者;与此同时,学生将变为研究者和参与者。最终从质量和数量上同时满足各个领域对翻译人才的需求。

Learner Interlanguage Analysis

基于衔接特征的英语学习者书面语语篇连贯自动评价研究

刘国兵 河南师范大学

能够有效反映语篇质量的语言特征主要包括语篇连贯性、词汇复杂性、句法复杂性以及语法准确性等,语篇连贯是高质量作文的核心特征之一。本研究综合国内外英语作文自动评分与语篇连贯研究两个领域的知识,利用语料库语言学的方法,通过提取学生作文中与语篇连贯相关的文本特征,尝试构建中国英语学习者书面语语篇连贯评价模型。本研究重点回答两个问题:(1)提取的变量中,哪些变量能够有效预测中国英语学习者书面语的连贯性?预测力有多大?(2)构建的连贯性评价模型是否可以用于学习者书面语连贯性的自动检测?准确性如何?

本研究使用的语料选自中国学生英语口语语料库。语料选取采用随机抽样的方式进行。研究发现包括三个方面:(1)提取的45个变量中,通过相关分析,共发现23个语篇整体连贯预测因子和18个语篇局部连贯预测因子。(2)语篇主

题相关连结数、中心句百分比、相邻连系数、二元单位复现及连结密度等 5 个预测因子作为有效自变量进入了整体连贯方程模型。方程的复相关系数为.782，决定系数为.612。从语篇局部连贯评价模型的构建结果来看，非中心句百分比、语篇主题连系数、句子数、词元复现以及二元单位复现等 5 个预测因子作为有效自变量进入了局部连贯方程模型。该模型的复相关系数为.604，决定系数为.365。

(3) 本研究从机器评分与人工评分的相关性以及二者的相对一致性两个方面对模型进行验证。结果表明，语篇整体连贯与局部连贯两个评价模型完全能够应用于学习者书面语语篇连贯性的自动评价，机器评分与人工评分在相对一致性方面能够达到令人满意的效果。

中国英语学习者口头叙事话语结构研究

刘霞 北京外国语大学/中国外语教育研究中心

叙事话语是人类言语交流的一种基本类型。自人类开始组词造句，便产生了叙事话语。但国内叙事话语的研究与其他体裁的话语研究相比（如学术话语），未得到应有关注。

本研究依据 Biber 等(2007)提出的基于语料库的自上而下的话步分析(move analysis)方法，旨在发现中外同题口头叙事话语的宏观结构及不同话步的语言特征的异同。研究步骤为：1) 32 位中国英语学习者和 32 位英语本族语者分别观看主题为“梨子故事”的无对白视频短片；2) 本族语者用英语复述视频内容，学习者先用英语复述再用汉语复述视频内容；3) 研究者对复述内容进行转写，进而得到基于同一录像的英语口语叙事研究语料；4) 根据场景、人物和情节变化，将 96 篇文本切分成五个情节话步；5) 依据 Stein & Glenn (1979) 的叙事结构模型，以命题 (proposition) 为最小分析单位，利用 BFSU Qualitative Coder 对文本话语进行手工标注；6) 利用软件进行检索并分析。

经过此分析步骤后，每一篇口头叙事话语都成为了一副故事结构的树图，经过对比后发现，学习者的叙事结构过于简化。首先，学习者的叙事话语缺乏对场景 (setting) 的描述，开篇往往直接引入人物或事件；然后，学习者大多直接描述结果，缺乏在达到该结果前，对人物的动作及心理的具体描述，即经常省略叙事结构中的努力 (attempt) 和内在心理反应 (internal response) 成分，由此可能导致叙事话语不够连贯；最后，学习者的叙事结构与本族语者相比，存在顺序不一致的现象。比如有些学习者会先说结果 (direct consequence)，再描述为达到该结果主要人物所做的努力，最后才引出新出现的与该结果相关的次要人物或事件 (initiating event)。

Lexical Priming, a Cross-linguistic Perspective

Jun Lu Yangzhou University

The theory of lexical priming argues that lexis, being primed for collocational use (in a broad sense), is complexly and systematically structured and that grammar is an outcome of this lexical structure. The present study is intended to re-examine lexical priming from a cross-linguistic perspective, namely through a corpus-based

contrastive analysis of English and Chinese passives within the model of extended units of meaning (EUM model). The study involves multiple comparisons between Chinese and English (both L1 English and L2 English) in such aspects as collocation, colligation, semantic preference and semantic prosody. Both comparable corpora and parallel corpora were employed to provide evidence for the contrastive analysis.

The results show that i) the passives of English and Chinese equivalent verbs may exhibit marked difference in frequency, collocation, colligation, semantic preference and semantic prosody. It varies considerably among different node verbs; ii) for the given node words under investigation, English words are more frequently primed for passive use than their Chinese equivalents for different pragmatic functions; iii) In English-Chinese parallel corpora, English passives and their Chinese translation exhibit considerable difference in colligation, but tend to be highly congruent in semantic prosody; iv) the passives of given node words in the English writings by Chinese EFL learners exhibit marked difference from native English, but share much congruence with their Chinese translation equivalents, especially in semantic prosody.

The above findings suggest that i) both English and Chinese lexis are primed for collocational use, which is characterized by specific grammatical structures/features, semantic preference and pragmatic functions; ii) English and Chinese verbs are primed differently for passive use, especially in terms of attitudinal meaning or communicative purpose; iii) Receptive priming and productive priming work differently and tend to have different impact on language acquisition; iv) There is considerable difference between the priming of L1 and L2 words in that, in addition to the cracking in Hoey's term, L2 priming also involves the competition between bilingual equivalent words (mother language and target language).

The study, from a cross-linguistic perspective, comes up with a competition model of L2 priming, which further develops the theory of lexical priming and lends some support to the argument that semantic prosody has a leading role to play in the integration of a lexical item and its environment, and illustrates the relationship between local grammar and local function. The implications of these findings for translation practice and language teaching research are also discussed.

A study of self-repair operations by Chinese English learners

Lihong Quan Guangdong University of Foreign Studies

Self-repair is the process by which speakers revise their prior talk themselves, at their own initiation. Among the different self-repair operations, recycling and replacement are the most commonly used devices to carry out the repair mechanism. Recycling (or repetition) refers to the repeating, either with no apparent changes or with some additions or deletions, of the repaired segment. The other commonly used self-repair device is replacement, which is the substitution of a quasi-lexical or lexical item or items for another quasi-lexical item/ items when both the repaired and the repairing segment belong to the same syntactic class.

Previous works focusing on the relationship between grammar and self-repair

maintained that some typological features, such as basic word order, syntactic constituent types, and morphological structure, influence the self-repair patterns across languages (Rieger, 2003; Wouk, 2005; Fox, et al, 2010, Nemeth, 2012). The studies found that, for languages with function words preceding content words, there seemed to be a universal pattern in which the speakers tended to use function words as destinations of recycling more often than content words, while replacing content words at a disproportionately high rate.

This study attempts to examine the similarities and differences between Chinese English learners and native speakers (henceforth NSs) in the practice of two self-repair operations, namely, simple recycling and simple replacement in English conversations. Rationale for the selection of Chinese English learners in this study stems from the fact that most of the prior studies dealt with the practice of recycling and replacement operations by native speakers. And comparative studies between L2 learners and native speakers are even fewer. It appears to be conducive to the field to look into the matter with a focus on a comparison between Chinese English learners and NSs.

This study aims to seek answers for the following questions:

- (1) In general, what are the similarities and differences between L2 learners and NSs in the use of recycling and replacement repairs?
- (2) What are the similarities and differences between L2 learners and NSs in the use of destinations of recycling and replaced items? What are the possible reasons for the differences?
- (3) What types of syntactic / lexical items are most repeated by L2 learners and NSs?

The study used database from the Chinese component of LINDSEI (Louvain International Database of Spoken English Interlanguage). A comparison was made with this study and data from the LOCNEC (Louvain Corpus of Native English Conversation) corpus compiled by Sylvie De Cock (Gilquin, et al. 2010), which is the native speaker counterpart of LINDSEI corpus. The data have been analyzed with a special focus on two aspects: the syntactic class of words both Chinese learners and native speakers tend to initiate recycling and replacement in, and what types of syntactic /lexical elements are most repeated.

The study has the following bullet points: (1) Learners use more content words as destinations of recycling, compared with natives. (2) There is an over-representation of verbs in L2 recycling and replacement repairs. (3) The study proves that the salient constituents in L1 may not be salient in L2. (4) L1 transfer, automaticity and attention in L2 are possible explanatory factors. (4) Learners and natives employ more word-level recycling than group-level recycling.

Finally, the findings were discussed with reference to morpho-syntactic patterns of English language, combined with theories of attention and automaticity in L2.

Stance Adverbs in a Chinese Learner Corpus
Xiaolin Zhao Jingjie Li Donghua University

A corpus-based contrastive approach is adopted to characterize the unique patterns and meanings of stance adverbs in a written English corpus of Chinese Learners of English (CLEs). The British Student (BS) written corpus is used as a reference corpus, and the Bank of English (BoE) from the Birmingham University, is supplemented as the backup for contextual information in Native Speaker (NS) texts.

A combined approach is adopted to explore the expression of stance in three domains, and accordingly three major categories of stance adverbs are dealt with: epistemic, attitudinal and style. The study of stance adverbs focuses on the most common controlling words that have occurred frequently in university registers (Biber 2006: 92). This study indicates that while making epistemic stance, the epistemic stance adverbs in CLEs texts show different semantic preferences and semantic prosodies from those in NSs texts that yield unique pragmatic functions and discoursal effects. Epistemic stance is polarized with a weak re-adjustment between “yes” and “no”, but constructed by thematization, which contributes to explicitness of the expression of epistemic stance. Regarding two subcategories of epistemic adverbs, certainty epistemic adverbs are apt to ensure authorial certainty about the events concerning the writer, while likelihood epistemic adverbs tend to vaguely comment on the events on part of the reader.

The expression of attitude stance is prominent due to the influence of Chinese culture. CLEs tend to use paired attitude stance adverbs for expressing the two sides of an event, particularly expounding the positive effect of a consequence. Moreover, CLEs tend to use the stance adjectives in the *it*-extraposition structure in addition to the attitude stance adverbs.

The expression of style stance is characterized primarily by thematization, subjectivity, explicitness and a consensus of authorial style stance with the imagined reader by using the *we*-clause, which is similar to the expression of epistemic and attitudinal stance. Unlike NSs, CLEs not only project their attitudes to the text, but shape the reader’s point of view. The writer positioning is characterized by attitudinal polarization, high authorial visibility, attitudinal thematization and empathy in stance making; the reader positioning is featured by attitudinal involvement and attitudinal inclination. CLEs’ stance making reveals the extralinguistic influence, including not only socio-cultural transfers such as L1 cultural transfer and transfer of Chinese social values, but interpersonal functions in EFL teaching in China such as the lack of systematic summary of stance expressions in the textbooks and the lack of overt instruction in classes.

A Corpus-based Investigation of Dialogue Acts in Conversations by Chinese EFL Learners

Jing Cao Zhongnan University of Economics and Law

Chengyu Fang City University of Hong Kong

The current research intends to examine the communicative behavior of Chinese EFL learners from the perspective of dialogue acts, with an ultimate goal of offering

valuable insight into the English education. The on-going research is initially motivated by the much less attention that spoken English of Chinese EFL learners has received when compared with their written English.

The comparatively less interest in spoken English is mainly due to the practical difficulty in the construction of spoken corpora (Wei et al., 2007) and the annotations (Xu, 2007). Encouragingly, the availability of officially released spoken corpora of Chinese EFL learners, such as the College Learners Spoken English Corpus (COLSEC, Yang and Wei, 2005) and the Spoken English Corpus of Chinese Learners (SECCL, Wen et al., 2005), provides naturally occurred materials for fundamental and scientific research. Research on the spoken English of Chinese EFL learners has been looking into the various research areas, such as

- Lexical features (e.g. Wen, 2006; Wu and Zou, 2009)
- Pronunciation (e.g. Zheng, 2013; Bi and Chen, 2013)
- Lexical chunks (e.g. Wei, 2004; 2007; Yang, 2004; Zheng and Zhao, 2011)
- Discourse markers (e.g. He and Xu, 2003; Wang and Zhu, 2005; Xu, 2013)
- Conversational moves (e.g. Yang, 2004; Wei, 2004)
- Communicative management (e.g. Wei, 2004; Xu and Xu, 2007; Chen and Pu, 2007)

Moreover, many of the past studies have also employed reference corpora for comparative analysis, including the spoken section of well recognized corpora, such as BNC, ICE-GB and COCA, as well as the spoken corpora, such as the London-Lund Corpus of Spoken English (LLC), the Louvain International Database of Spoken English Interlanguage (LINDSEI), the Bergen Corpus of London Teenage Language (COLT), and the Michigan Corpus of Academic Spoken English (MICASE).

To further research on the spoken English of Chinese EFL learners, the current study proposes to investigate conversations from the perspective of dialogue acts (DAs). ISO 24617-2 (2012) defines dialogue acts as the ‘communicative activity of a participant in dialogue interpreted as having a certain communicative function and semantic content, and possibly also having certain functional dependence relations, rhetorical relations and feedback dependence relations’ (p. 3). According to Bunt et al. (2010), dialogue acts play a key role in the interpretation of the communicative behavior of dialogue participants. Thus, the study intends to examine the communicative behavior of Chinese EFL learners.

In this paper we report a case study of Positive Feedback, ‘a communicative function of a dialogue act performed by the sender, S, in order to inform the addressee, A that S believes that S’s processing of the previous utterance(s) was successful’ (ISO 24617-2, 2012: 71). To observe the dialogue acts used by Chinese EFL learners, the section of Task 3 in the SECCL 1.0 (revised version, Wen et al., 2005) is chosen as the corpus resource. The Switchboard Dialog Act (SWBD-DA) Corpus is employed as the reference corpus. Different from other spoken corpus of native speakers, the SWBD-DA corpus is annotated with 59 different dialogue acts according to the SWBD-DAMSL scheme. In our study, the most frequently used 10 expressions of Positive Feedback are identified and extracted from the SECCL Task 3, and then

compared with the top 10 expressions in the SWBD-DA Corpus. The finding shows that Chinese EFL learners display quite different communicative behavior from native English speakers in terms of Positive Feedback. The results show that the most commonly used expression 'Uh-huh' by the native English speakers never occurred in the conversation by Chinese EFL learners. The data also reveal the errors made by Chinese EFL learners in the use of expressions such as 'Oh'. After discussing the idiomatic expressions functioned as Positive Feedback with corpus evidence, the paper will conclude with some pedagogical applications of dialogue act analysis.

Syntactic Complexity Analysis of Oral Writing Based on Self-built Corpus of Non-English major Graduate Students

Man Cao Wuhan University of Technology

Syntactic complexity has been a focus of second language writing (SLW) research (Ortega, 2003), but measuring syntactic complexity is difficult. Wolfe-Quintero, Inagaki, and Kim (1998) summarized more than a hundred indices to measure syntactic complexity. As the indices of measure vary across different studies, inconsistent results were presented. For example, there is no definite conclusion on the relationship between language proficiency and syntactic complexity. SLW researchers often apply T-unit (including the main and all clauses) and subordination in analyzing grammatical complexity, concluding that more number of clauses indicates higher writing level. Lu (2011) compiled L2 Syntactic Complexity Analyzer to measure 14 indices in five aspects: length of production, sentence complexity, subordination, coordination and particular structures. Lu applied the 14 indices to Written English Corpus of Chinese Learners and identified seven indices more appropriate to measure writing abilities of learners at different language level: complex nominals per clause (CN/C), mean length of clause (MLC), complex nominals per T-unit (CN/T), mean length of sentence (MLS), mean length of T-unit (MLT), coordinate phrases per clause (CP/C), and coordinate phrases per T-unit (CP/T). However, Biber, Gray and Poonpon (2011) point out that T-units and clausal subordination may not be effective to measure complex noun phrase constituents since subordination is used far more frequent in conversation than in academic writing.

Based on the suggestion of Biber, Gray and Poonpon (2011) and Lu (2011), the current study applied 4 indices, namely CN/C, CN/T, CP/C, CP/T, to measure the syntactic complexity of untimed oral writing taken from a self-built EFL graduate students' writing corpus. The research tool was L2 Syntactic Complexity Analyzer, and the data include 50 writing samples from students of two humanistic majors and 50 from students of two scientific majors. Then SPSS 19.0 was used to determine whether students' English level (their scores in College English Test) and writing performance were respectively relevant with the above four indices, and that whether there were differences of the 4 indices between students from different majors. The results showed that students' English level and writing performance were relevant

with the above 4 indices, and that there were no significant differences of the four indices between students from different majors. The findings of the study point out the direction to teaching academic writing to EFL graduate students. Targeted exercises can be designed to raise students' awareness of syntactic features of academic writing and to enhance the students' use of complex nominals and coordinate phrases in writing production.

“四位一体”的学习者英语分析框架——以“-ed 小句”为例

方秀才 上海外国语大学

本文通过综述国内外学习者 EFL 语言研究发现：基于语料库的对比中介语分析研究多是以个别词或短语为对象，很少对某一类现象进行探讨；大多数研究未能很好兼顾研究的深度和广度，也未能将理论指导、方法论框架、解释维度和应用探讨综合到同一框架内彼此关联。基于此，本研究构建词汇语法整体观、对比中介语分析方法、学习者特征解释维度以及研究应用价值探讨“四位一体”的研究框架，并以“-ed 小句”为研究焦点，探讨中国英语学习者的使用特征、从多维视角解释特征背后的原因并探讨研究发现的 application 价值。

“四位一体”的学习者英语分析框架包括以下四个部分。首先，文章整合现代语言学从整体观来讨论词汇和语法关系的主要语言学理论论证“词汇和语法的整体性”，它们包括：Firth (1957)的语境论、Sinclair (2000)的词汇语法理论、Hunston 和 Francis (2000)的型式语法理论、Hoey (2005)的词汇触发理论和认知语法相关理论 (Langacker 1987, 1991)。这一节旨在论证在“词汇语法整体观”指导下运用基于语料库的方法来分析词汇意义和语法功能时的合理性和可操作性；其次，“对比中介语分析方法” (Granger, 1996; 2002; 2004) 勾画了整个研究的方法论框架：即 NS/NNS 母语者与非母语者比较、NNS/NNS 非母语者之间的比较。前者强调非母语者语言的“非母语性特征”，后者有助于区分学习者特征是中介语发展特征还是特定母语相关特征。再次，“中介语产出的解释框架”为学习者语言使用特征的解释提供综合的多维解释框架，包括外在因素：主题语境、L1 迁移、L2 输入、L2 产出环境；内在因素：L2 水平、学习者特征、语域意识以及石化现象；最后，“学习者特征研究的应用价值”讨论学习者特征研究对于教学、教材编写及测试等方面的启示，构建应用价值讨论框架。

在“四位一体”的框架下研究以“-ed 小句”，提出以下研究问题：

第一、中国英语学习者和英语母语者在“-ed 小句”的使用上是否有差异？呈现什么特征？这些特征背后的原因有哪些？

第二、不同水平的中国英语学习者之间在“-ed 小句”的使用上是否有差异？呈现什么特征？这些特征背后的原因有哪些？

第三、中国英语学习者在“-ed 小句”的使用特征主要体现“中介语发展特征”还是体现“母语相关特征”？

第四、中国英语学习者“-ed 小句”的特征研究和原因解释能够给英语教学等领域带来哪些启示？

后续研究通过对比中介语分析比较中国英语学习者和英语母语者及法德两国英语学习者以及不同水平的中国英语学习者之间的“-ed 小句”的使用差异；根据自动词性标注和人工语法语义标注总结了“中国英语学习者”的“中介语特征”；在

“多维解释框架”下分析“中介语产出原因”；在“应用价值框架”下讨论学习者特征研究的应用价值。

中国学生学者学术论文引言的短语学分析

房印杰 北京邮电大学

随着全球范围内针对学术英语的需求不断增长,非英语本族语者在学术英语使用中的特征研究成为一个热点问题。本次研究以通信工程类学术论文的引言部分为切入点,分别建立了学习者学术论文引言语料库(CUPT)与专家学术论文引言语料库(Telecom)。并采用 Swales(1990)中提出的“建立学术研究空间”模式,自上而下的将两个语料库分别切分为三个部分,以对应模式中的三个语步:确定研究范围、建立研究空间、占领该空间。而后,采用自下而上的方法,从词语表、短语序列、短语框架等层面对两个语料库进行分析,并结合 Halliday 的元语篇功能来进行阐释。研究表明,学习者学术论文引言语料库与专家学术论文引言语料库在三个语步上均存在一定程度的差异,其中,在第一语步层面,参照语料库表现出更加侧重学科术语;在第二语步上,参照语料库更加频繁的使用转折连词;在第三语步上,虽然两个语料库均使用 paper 和 section 的相关短语序列,但是参照语料库的相关短语序列有着更高的多样性,更高的频率。两个语料库的语步对比说明:中国学者已经开始意识到了学术论文引言部分的不同交际功能/语步,但是,他们还未能更深入的掌握对应这些交际功能/语步的短语序列。

学习者英语议论文中主语生命度研究

吉洁 中国外语教育研究中心

生命度(animacy)是名词的固有属性之一,在书面语中可以体现作者在认知上的主体融入程度。本研究对英语本族语者、中国英语学习者议论文中的主语生命度进行对比分析,探究学习者议论文中主语的生命度特征,同时通过对比不同得分的学习者语料,分析主语生命度对议论文写作质量的影响。

本研究使用的本族语者语料来自 LOCNESS 中所有议论文,学习者的语料来自 WECCL2.0 中所有英语专业学生的议论文和 CLEC 中的非英语专业学生的作文(四、六级子库)。本文考察的是三个语料库中定式小句(finite clause,即不包括不定式非时态小句)的主语。所有定式小句必须包含一个主语和一个谓语,即在主语和谓语共现的情况下,谓语前的名词性成分大多就是主语。研究的操作步骤如下:将所有可担当主语和谓语的成分列举出来、编写各种不同主谓结构的正则表达式、在 PowerGREP 中依次检索、人工排查与区分、统计。

研究发现:1) 本族语者在议论文写作中倾向于使用无灵主语,而学习者倾向于使用有灵主语,学习者议论文的主语生命度呈现整体偏高的现象;在有灵主语中,学习者过多使用第一、二人称代词,具有较高的作者/读者显现度和作文主观性;在无灵主语中,学习者使用的无灵名词缺乏丰富多变性,指示代词和 it 做主语的用法较少,动名词和从句作主语时存在使用泛化。学习者主语生命度偏高可能与汉语思维迁移、产出性词汇匮乏、偏爱使用主动语态、和惯于以人举例的论证方

式有关。2) 主语生命度是影响议论文质量的因素之一。使用大量有灵主语, 在提供证据时惯以个人事例进行假设举例, 甚至以举例代替论证、缺乏理性思辨和逻辑推理, 极大地削弱了议论的说服力和深度, 影响了议论文的整体质量。

中国英语学习者虚化动词搭配发展特征

刘晓燕 河南师范大学

目前学术界关于学习者虚化动词使用的研究已经为数不少, 如邓耀臣(2005), 赵晓临(2010), 刘国兵(2011), 杨纳枝(2011), 张琳(2012)等。通过对比分析, 该类研究发现了中国英语学习者在虚化动词使用上存在的诸多特点。但问题在于, 由于缺乏可比性较高的本族语者语料库, 已有研究通常把 LOCNESS 或其他通用语料库(如 BNC)的笔语子库与学习者语料库进行对比, 之后得出学习者与本族语者在虚化动词的使用上存在显著性差异的结论。我们认为, 这些本族语语料库与学习者语料库相比, 在题材、体裁、内容以及文本长度等多个方面都有很大不同。因此, 此类研究得出的结果也有待进一步验证。本研究选取 WECCL 作为学习者语料库, NESSIE 作为本族语者参照语料库。通过对这两个语料库中 GIVE 使用情况的比较, 试图从多个方面重新揭示中国英语学习者虚化动词的使用情况。

研究发现, 1) 从整体来看, 与英美本族语者相比, 中国英语学习者过多使用此类虚化动词, 搭配模式也较为单一; 2) 学习者作文中使用的虚化动词, 表现出明显的口语化倾向, 与书面语语体存在不一致的情况; 3) 在使用虚化动词准确性与地道性方面, 高年级学生与低年级学生之间差异较为显著; 4) 在虚化动词的使用方面, 高年级学生更接近于母语水平, 二者之间不存在显著性差异; 低年级学生与本族语者之间差异则较为显著。结果表明, 对于虚化动词的掌握和使用, 学习者会随着语言整体水平的提高而提高; 高级英语学习者可以无限接近本族语者水平。

Units of Meaning in Chinese University Students' Argumentative Writings: the Case of Collocational Framework "the * of"

Suxiang Yang School of Foreign Studies Henan Polytechnic University

This paper is grounded in Wittgenstein's philosophical perspective on language meaning and Firth's contextual theory of meaning. It employs Sinclair's lexical grammar model to systematically investigate the units of meanings based on the collocational framework in Written English Corpus of Chinese Learners (WECCL) and Louvain Corpus of Native English Essays (LOCNESS). The purpose is to find the collocational features of collocational framework "the * of" in the native speaker and Chinese university students' argumentative writings.

There are two reasons for choosing "the * of". Firstly, as previous corpus researches on units of meaning were mostly based on continuous words or phrases, this research tries to extend the units of meaning study to discontinuous words or phrases. The second reason is that "the * of" is the most frequently occurred

collocational framework in English. Therefore, the research questions of this study are as follows: 1) What are the lexical grammatical profiles of “the * of” in Native speaker’s argumentative writings? 2) What are the lexical grammatical profiles of “the * of” in Chinese university students’ argumentative writings? 3) How do the features of units of meaning in Chinese university students’ argumentative writings differ from those in native speakers’ argumentative writings?

Enlightened by, but not strictly following, Sinclair’s (1996/2004a) lexical grammar model and Hunston’s (2008) semantic sequence approach, the author regards “the * of” as a core of units of meaning and expands it as a seven-word string with the words on the positions of Left 2 and Left 1 (hereinafter: L2C and L1C respectively) before the core, the collocated words in the middle position of the core (hereinafter: MidC), and the words in Right 1 and Right 2 positions (hereinafter R1C and R2C respectively) after the core, hence a string of units of meaning based on collocational framework covering “the * of” is “L2C + L1C + the + MidC + of + R1C + R2C”.

Following the principle of frequency-driven in corpus linguistics, the author first retrieved cases of concordance lines based on the core “the * of” from the corpora, and set the minimum frequency as 2, therefore drew 366 (normalized frequency is 2,012) concordance lines in LOCNESS (181,879 words) and 734 (normalized frequency is 699) concordance lines in WECCL (1,049,680 words). After that, from the 366 concordance lines in LOCNESS, the author examined the top 200 collocates in the five positions: Left 2, Left 1, Middle, Right 1, and Right 2. Then categorizing the collocates at three levels — word class categories, semantic categories and functional categories, the co-occurrence of the top three categories at each position are analyzed so as to form typical patterns — colligational sequences, semantic sequences and functional sequences. For the 2,017 concordance lines in WECCL, the same methods are used.

The research results show that some salient features of collocational framework “the * of” in Chinese university students’ argumentative writings when compared with native English students’ argumentative writings. 1) the semantic sequence patterns in WECCL are more diversified instead of being typically patterned as in LOCNESS; 2) the nominalization rate of “the * of” in the middle position in WECCL is less than those in LOCNESS; and 3) the collocates in the Middle position have less abstract nouns expressing “Essence” in WECCL. All these show that Chinese university students’ writings, though in concordance with grammar rules, are stylistically oral orientated, less formal and less academic.

The significance of this study implies that function words with high frequency are better taught and learnt in terms of a meaningful collocational framework instead of in their isolated forms. They can assist students to efficiently identify the meaning of a word and disambiguating the meanings of words through the lexical grammar profile of the collocational framework, and finally to master the typical usage of collocational framework from form, meaning and function.

非英语母语者的英语使用能有多地道？

杨晓东 浙江工业大学

基于语料库的中国报刊新闻英语分析近年来也受到语言学者的关注，有的关注中国的英语新闻中词汇与主题表达之间的相互关系以及词汇的使用和语言学特征(高超, 2006)；有的关注中外报纸新闻标题语言及文化特点(周杰, 2008)；有的着眼于英语报纸新闻的批判性话语分析(钱毓芳, 2010)。但目前国内对于母语为非英语的记者和编辑所作英文新闻稿的母语性分析（analysis of nativeness）尚未见到。

作为非英语国家的英文刊物，许多稿件的撰写和编辑都是母语为非英语人员完成的，虽然有专业人员对将要刊载的稿件进行语言润色，我们还是希望知道这些稿件英语母语性的具体表征，即这些稿件英语表达的地道程度如何。由此可以进一步发现非母语人士在英语表达上跟母语人士存在什么差异，这对高层次英语教学与学习有非常重要的意义。

本研究选取中国较有影响力的两份英文报刊----《中国日报》和《上海日报》中中国作者所撰写的社评类文章共计 100 万英文单词组建小型语料库，对所搜集的语料做词频统计，并分析核心功能主题词，根据词频的统计数据 and 核心功能主题词的分布对语言事实进行定量描写和概括，并将其和英文同类型语料库（如 Frown）的语料进行对比，发现其词汇分布特征以及核心功能主题词的使用规律，从而推出中国英语报刊英语的母语性特征并分析其成因与意义。本研究将有助于我们加深语言习得过程中的母语趋近性的认识，也有助于我们管窥英语语言的细微之处，为英语学习者和教育工作者提供语法书上看不到的参考。

基于美国当代英语语料库对学生英语作文中词汇应用能力研究

张金福 九江学院

本研究以美国当代英语语料库（COCA）为教学平台，将过程教学法与数据驱动学习手段有机结合，引导学生在写作教学过程中的各个阶段（写作前、写作过程中、写作后）充分应用该库来提高词汇应用能力，从而达到提高写作质量的目的。研究发现：1）实验班学生在使用 COCA 辅助作文教学后，其词汇丰富度有一些变化。实验班后测与前测相比差异显著，实验班后测与对比班后测有差异，但无显著差异。原因可能有两个方面：一是实验班在培训 COCA 语料库检索技巧阶段占用了一些写作课时间；二是部分学生，特别是基础较差学生对语料库培训课程不重视，导致后来检索能力差而失去应用 COCA 信心。2）在“写作前”阶段，COCA 对学生“头脑风暴”能起到联想提示作用，同时对专业术语规范表达有促进作用，但主要只是在写部分话题的说明文时效果较为明显。在“写作中”阶段学生并不愿意应用，主要是认为时常检索容易打断写作思路，导致效率低。在“写作后”修改阶段应用可以提高学生词汇应用能力，使学生的用词更地道、表达更得体。3）在利用 COCA 辅助写作教学中，将其与其它工具如英语词典等配合使用效果更好。

东华大学本科生使用批改网评阅作文的反馈与思索

张莉 东华大学

批改网是基于语料库和计算机云技术研发的作文线上评阅综合系统。东华大学本科生从 2013 年上半年开始试用。使用者的反馈情况如何,给英语作文授课的教师带来什么样的思考,未来批改网的研发有什么样的方向和改进意见。这些是本论文努力探讨的研究内容。论文撰写者是试用批改网的一线教师,也是语料库语言学研究方向和多模态英语语言教学方向的研究型教师。

本研究的取样来自该教师三个本科生授课班级,93 份有效问卷中男女生比例六比四,文科生与理工科生所占比例分别为 7.3% 比 92.7%,样本具有理工科学学校学生的代表性。问卷的设计以三分和五分量表相结合的形式,揭示出使用者们对英语作文与英语学习的认知,以及对批改网使用的实际看法。开放式的问卷条目中,学生们反映的情况主要分为两部分。一是认为批改网能辅助英语作文教学,但教师的作用仍旧不可绝对替代。高达 80% 多的使用者仍呼吁教师对学生作文适当点评,结合范文进行分析。二是对批改网的设置和使用提供了学生们自己的一些使用体会和建议,主要集中在批改网提供的语句分析、语篇结构等方面的不足。有来自计算机专业的某位学生甚至提出了具体如设置学生提交修改上限的建议。

虽然这次研究的历时很短,样本量比较小,研究设计还很浅显,但以上总总使用反馈数据都可以为将来的英语作文教学和批改网的进一步开发完善提供有益的思索。

基于语料库的中国学习者英语程度副词的对比研究

巩海曦 田建国 西北工业大学

本文通过语料库的研究方法,在选定语料库 CLEC 和 LOB 中对比中国英语学习者和本族语者在一组易混程度副词 *fairly*, *rather*, *quite*, *pretty* 的使用分析,从其语用频率、类联接、语义韵等方面进行比较研究。本文利用 *wordsmith Tools* 提取索引行进行描述和分析,来研究选定程度副词在使用和搭配上有何异同。

本文选取一组易混程度副词作为主要研究对象,基于以下两点原因:1) 对许多易混的程度副词的教学,教师经常忽视词语在其语用、语义、语韵等层面的用法。2) 以往对于程度副词的研究都是倾向于总体的使用情况和分类,而没有从个别语义近似的程度副词中着手。

本文研究结果及创新之处总结为如下几点:

- 1) 从整体频率来看,本族语者在使用程度副词整体频率要高于中国英语学习者;在四个选定程度副词中,中国英语学习者和本族语者在使用 *rather* 和 *pretty* 两个词语的频率差异最大, *fairly* 次之, *quite* 则使用频率差距最小。相比其他程度副词而言,中国英语学习者已经基本掌握了 *quite* 的基本用法。相比本族语者而言,中国英语学习者使用程度副词的限制性与其背诵单一词义有关。
- 2) 在类联接方面,检索选定程度副词,得到基本的类联接形式有 *DA+adj*, *v+DA*, *DA+particle*, *DA+prep*, *DA+adv*。分析得出,中国学生和本族语者都倾向于使用 *DA+adj* 的类联接搭配模型。中国学生 *pretty* 和 *fairly* 的类联接与本族者类似,但 *rather* 和 *quite* 有显著差异。

3) 在语义韵方面, 中国大学生在英语程度副词搭配方面所呈现的语义韵特征也与本族语者的使用惯例之间有显著差异。中国学习者在使用 *fairly+adj* 和本族语者使用 *pretty+adj* 时, 都呈现积极的语义韵; 中国学习者程度副词 *quite* 和 *rather* 的语义韵特征与本族语者呈现较大差异, 分析体现了中国大学生更多地生搬硬套典型搭配, 搭配范围小, 搭配词项单一, 还会出现违反搭配规约的不符合语义韵和谐的大量搭配。

结果表明, 中国大学生在使用程度副词过程中, 出现使用单一, 类链接结构搭配模型僵化, 语义韵意识淡薄等问题。所以应该在外语教学中引入类链接、语义韵概念, 基于大量语料库证据, 引导学生分析、概括搭配词项的语义特点, 这将有助于他们二语习得的学习。

A Corpus-based Usage Contrast of Collocational Frameworks

Fang Guo Shandong University of Science and Technology

From the Birmingham Collection of English Text, Renouf and Sinclair (1991) have found out the following collocational frameworks that claim the highest frequency of occurrences: a * of, an * of, be * to, too * to, had * of, for * of, as * as, many * of. Collocational frameworks are also called frames. They pose a hot potato for EFL learners because they may cause difficulties in collocation, with the words on both sides fixed and the word in the middle variable. Longman Grammar of Spoken and Written English has also listed the following frames with high occurrences: as *as, in *to, with the * of. To further investigate Chinese EFL learners' use of these frames, the author makes a usage contrast of as * as, in* to, too* to between the writings of Chinese EFL Learners and Native Speakers.

The research questions are: (1) What are the distinctive features in the use of frames in the writings of Chinese EFL learners and native speakers? (2) To what extent do these frames account for the Chinese learners' lexical acquisition and written production? To answer these questions, the author makes a corpus-based study, in which two corpora are selected to form a contrast. One is the WECCL, which is a sub-corpus of the SWECCCL and serves as the learner corpus. The other is the LOCNESS, a sub-corpus of the ICLE and serve as the native corpus. Owing to the difference in their sizes, all the raw frequencies are computed into normalized frequencies. As to the software, AntConc is used as concordance software and SPSS (Statistical Package for the Social Science) is used to deal with statistics.

The result shows that: (1) Compared with native English speakers, Chinese learners tend to overuse those frames that have a direct Chinese equivalent and are emphasized by the teaches, such as as soon as, in order to and etc. Meanwhile, Chinese learners underuse, misuse or avoid using those frames that are not emphasized in teaching and that have different syntactic functions from their Chinese equivalents. (2) There is a lack of variation in the slot-filler used by Chinese learners. That is to say, Chinese learners can't select the appropriate slot-filler to match the two words at both sides. In contrast, the slot-filler used by native speakers is diversified and . (3) The slot-filler in Chinese learners' writing differs from that in native

speakers' writing in terms of semantic field and semantic prosody.

Finally, some pedagogical implications of the research are discussed, such as the emphasis on frames and lexical chunks in English teaching in order to improve learners' idiomaticity and fluency; the compilation of a corpus-based dictionary of lexical collocation and the utilization of corpora in various ways to facilitate EFL learning and teaching.

A Corpus-based Study of Syntactic Features in EFL Graduate Students' Writing

Qing Huang, Man Cao Wuhan University of Technology

Modern corpus linguistics is an emerging linguistic research science in the middle and late 20th century. In recent decade many trademark achievements have been achieved in terms of the construction and research of Chinese learners' second language acquisition corpus. The current learners' corpus (such as CLEC) has collected many compositions of middle school students and undergraduates while non-English postgraduates' corpus fails to be listed as samples. As a result the writing characteristics of these learners are little known. It's well known that syntactic complexity is one of the most important indexes of second language acquisition. The foreign relevant research begins early with comparatively abundant data which are mainly written language. In comparison, domestic relevant research is relatively fewer with the focus on the complexity and richness of vocabulary.

This study is directed at non-English majoring postgraduates who are attached little importance in previous research. The sentences in their compositions are collected to build WUT postgraduates' writing corpus by classifying data according to its grammatical type. The data of this corpus will be used to analyze the relationship of syntactic structure and semantic expression in their writing, reasons to form varied syntactic structure, the relationship of learners' reproductive sentences and their social environment, individual background, and other factors.

This study employs semi-experimental way to collect the compositions of the non-English major postgraduates, do sampling arrangements and build writing corpus. Based on the stratification and coding of the data some statistical software such as syntactic complexity analyze, Wordsmith and SPSS are employed to analyze and calculate the data.

The study is expected to find the syntactic features of theses learners' writing, which is helpful for teachers to guide students' writing and is helpful for the second language learners to know the relativity between selection of language form and writing purposes. As a result their academic writing abilities are effectively improved. At the same time the key points and difficult points of the writing are discovered and analyzed, which has significance for effective improvement of WUT postgraduates' writing qualities.

Other Issues on Studies in Corpus Linguistics

On corpus linguistic approach to unit of meaning --- its development and pedagogic implication

Anping He South China Normal University

One of the most important findings of corpus linguistics is the “phraseological tendency” in language use, where “unit of meaning are expected to be largely phrasal” (Sinclair, 2004: 29-30). It thus renewed our recognition of meaningful units in a language.

This presentation will first account for historical development of corpus study of unit of meaning. It will then focus on three subtopics. First, why phraseology can be a basic unit of meaning in corpus perspective and moreover of various types in nature? Second, what is the strength in studying various types of phraseology, with its pedagogical implication in particular? Third, what are further issues to be resolved in studying various types phraseology?

Key concept or issue will be addressed to each subtopic, accompanied by relevant studies at home and abroad, including the latest research findings from the corpus linguistics teams of South China Normal University.

A Corpus-based study of Repetition in Daily Conversations

Yanjiao Li, Alex C. Fang, Jing Cao City University of Hong Kong

The pervasiveness of repetition is fairly salient in daily conversations, which usually refers to say the same thing twice or more. According to Merrit (1994: 27), repetition is a universal phenomenon as well as a major resource in communication. Speakers can repeat not only their own words or utterances, but also what has been said by the interlocutor. Recent work on repetitions has been extensively investigated across different disciplines, from rhetoric to phonology, philosophy to psycholinguistics and literary studies, to discourse and conversation analysis (Francesco Straniero Sergio, 2012). However, only a few studies have focused on repetition in conversation or other non-formal texts (Norricks, 1987; Tannen, 1987, 1989; Bazzanella, 2011; Rababah, G. & Saileek A., 2012). It is therefore the main purpose of this paper to seek to examine repetition in daily conversations, varying in types and functions, by investigating data collected from Switchboard-Dialogue Act corpus. This corpus comprises 1,155 transcribed telephone conversations, totaling in 205,000 utterances or 1.4 million word tokens, which is fully annotated with dialogue act information (Jurafsky et al. 1997).

To achieve this goal, we first investigate a wide range of forms in repetition. In the corpus, auto-repetition (repeat self) has been distinguished from allo-repetition (mimic others) with different tags (^r and ^m respectively). Furthermore, both types have been divided into nonsense repetition (interjection), ritual repetition and

semantically charged repetition. The last aspect, the focus of this study, again has fell apart into exact repetition, part repetition and repetition with modification according to the linguistic form. Based on the data statistics and analysis, it has been found that auto-repetition is more frequently used than allo-repetition in daily telephone conversations. For auto-repetition, semantically charged repetition accounts for nearly half of the whole auto-repetition; by contrast, it takes almost ninety percent in allo-repetition, illustrating that speakers usually tend to mimic meaningful elements uttered by the interlocutor, rather than nonsense elements. By three subtypes of semantically charged repetition, exact repetition is rather significant in auto-repetition, compared to that in allo-repetition, while for part repetition, it is more frequently used in allo-repetition than that in auto-repetition. The contrast shows that it's more likely for speakers to repeat their own words exactly. Also they tend to mimic part of utterances uttered by the interlocutor rather than their own words. Still, there are various types within repetition with modification, which is worthwhile to further look into it.

Secondly, according to various forms of repetition in conversations, we try to identify their different functions. For auto-repetition, it can be defined with respect to the following functions:

An intensifying function, by which the speaker expresses high degree of agreement/disagreement with what has been said;

An emphasis function, by which the speaker intends to emphasize one point mentioned previously;

A response function, by which the speaker repeats the previous talk as the corresponding to the interlocutor's request.

For allo-repetition, it can be defined as follows:

An acknowledgement function, by which the speaker indicates that he or she is attending the conversation and following the interlocutor. What was said has been heard and interpreted via mimicking others' utterances;

A positive attitude function, by which the speaker show his or her agreement with what has been said by allo-repetition;

A request function, by which the speaker shows uncertainty and seeks for a confirmation of what has been said by the interlocutor. On this point, it's often a rising tone at the end of the repeated utterance.

In addition, repetition can expand the conversation or work as a cohesive device, especially for repetition with modification, a kind of paraphrasing, or reformulating. Repetition as participation, it helps develop the conversation.

Finally, we also examine the annotation of repetition in the existing coding schemes of dialogue act. A repetition has been handled quite differently in other existing schemes of dialogue acts (9 schemes mentioned in ISO) due to a wide range of functions. It has been assigned a particular label in 6 schemes (DAMSL, SWBD-DAMSL, ICSI-MRDA, MALTUS, Coconut and SPAAC). However, several schemes haven't employed a repetition as a dialogue act, such as Verbmobil-2, Traum's and AMI coding schemes. A repetition can be interpreted as having different communicative functions in various contexts, as the new ISO standard (ISO 24617-2,

2012) pointed out, “a repetition often has a feedback function, but it can also have other communicative functions” (P.9).

On Discovery Procedures of Units of Meaning Jianzhong Pu NLP

In 1996, in his paper “The search for the units of meaning”, Sinclair put forward a proposal that “a start can be made on an inventory of the units of meaning of English” (Sinclair 1996). About 17 years have passed since then, yet such an inventory is still not available. The reason for this, as we see it, is that the notion of unit of meaning is not clearly defined on the one hand, and on the other, there is still methodological difficulties in retrieving units of meaning in whatever language, be it English, Chinese, or any other language. The main purpose of the present paper is to discuss the discovery procedures of units of meaning in English, hoping that they can also be applied to other languages.

For this purpose, three case studies will be made, which focus on three word forms, to be exact, said, world and personal, respectively, all of which are common ones, but occupy different positions on the frequency list of word forms derived from the Brown corpus. Besides, three other corpora – LOB, Frown and FLOB – which are also available to us, are used in cases where we do not get enough instances from Brown alone. We principally follow the procedures set forward by Sinclair (1996), which identified five inherent components of units of meaning, i.e. core, collocation, colligation, semantic preference and semantic prosody. However, it is found out that, at the present level of linguistic theory and corpus technology, it is still hard to automatically discover the units of meaning centered around these word forms. There remains not only technical problems but more importantly theoretical and methodological ones. Theoretically, for example, we do not yet agree what are really units of meaning, or what corresponds to semantic prosody or semantic preference. Methodologically, for example, WordSmith 5.0 provides us with altogether 5 ways of calculating the collocabilities between two word forms, namely, MI, Z, MI3, Log L. and T Score, yet none of them proves to be ideal, and thus there is the question of what to choose as our primary statistical indicator. As a result, we encounter difficulty in separating significant collocations from casual ones. Besides, there are many other obstacles to arriving at the level of semantic prosody and thus identifying the units meaning around the above cores.

Nevertheless, as long as we get enough instances of a word form and its use from the corpus, its patterns of use are bound to emerge. In the case studies of this research, we have developed the strategy of finding predominant uses first, i.e. we only consider very significant collocations, and very typical colligations in the very beginning. Based on them, attempts are then made to further discover the semantic preferences and semantic prosody. It is found that collocations and colligations are relatively easy to identify, but such is not the case with semantic preferences and semantic prosody, the identification of which depends a lot on what corpus and

software tools we possess. By using the Brown corpus and additionally LOB, Frown and FLOB, with the help of Wordsmith 5.0, for example, we find that one predominant use of said is something like “somebody said something in a certain way”, that world is typically used in prepositional phrases (in the world, throughout the world, around the world) and compound nouns (world war, world peace, world citizen) or compound adjectives (world-wide), all of which have their further uses, that personal is often used in “(one’s) personal things” (with “things” realized by property, experience, etc.), which again have their further uses. As one may easily find, what we have arrived at are mostly phrases of some kinds, but not what Sinclair had in mind in discussing units of meaning or compound lexical items. To arrive there, we need bigger and better corpora, which is easy to understand, and more powerful software tools, by which we mean they should be able to deal with collocations between more than three word forms.

In conclusion, it is plausible to make a start on an inventory of the units of meaning in English, and in any language, of which we have a big enough corpus. Hopefully, the procedures of identifying the units of meaning in a language will become much more automatic through the combined efforts of both linguists and computer engineers.

A Corpus-based Functional Discourse Analysis — President Obama's Inaugural Speech in 2013

Yan Qi, WuShuang Donghua University

President Obama’s inaugural speech in 2013 was not simply a summary of his last term but also showed his promise for the second. This paper analyzed Obama's inaugural speech in 2013 through corpus-based and Halliday’s systemic-functional grammar approaches. With the aid of Wordsmith Tool 4.0 and POS tagging software Tree Tagger, the analysis was based on Halliday’s three meta-functions theories through the discussions of the transitivity system, modality system as well as theme system, aimed to find out Obama’s speech discourse strategies and his real intentions for this address. This study found that this speech was of moderate difficulty via the use of a number of long sentences to reserve the seriousness and formality as well as simple words to cater to the masses’ needs. The recurrences of such slogans like “equality” and “freedom” indicated his political stance and incentive strategy. The high frequency of “America”, “country”, and “people” showed Obama’s determination to construct America together with his people. Material process verbs occupied an overwhelming part in the transitivity system, which made the speech more convincing and practical. The adoption of many first person pronouns and high value model verbs effectively motivated the audience’s passion. Also, the main application of radialized thematic progression pattern constructed a great deal of parallel sentences so as to make the speech more invigorative. All the speech strategies he applied was to win Americans’ continued support and manifest his resolution to recover the economy and his majors concerns in the next term like

human rights, climate change, medical service, and global peace. Corpus research approach gave important data support to this functional discourse analysis, hence proved an effective approach in discourse analysis.

短语学视域下语法与词汇的界面研究

孙海燕 河南师范大学

组合关系和聚合关系是制约语言运作的两个重要关系。聚合轴强调语法关系，揭示了语言使用的可能性。组合轴则体现了意义的现实性，关注词语的横向组合。传统的语言学多强调聚合关系，将语法置于语言研究的核心位置。自语料库语言学迅速发展以来，语言研究的方向出现了重大的变化，现在的语言研究愈来愈重视组合关系的研究。Sinclair (1991: 65) 认为形式和意义之间从根本上讲是没有区别的，“意义和句法是相互联系的”。比如，want 后面接动词不定式，这一描述很难说是有关词语搭配的还是有关语法方面的(Hunston & Francis 2000: 251)，因此语法和词汇之间没有泾渭分明的界限。词汇和语法联系密切，在词汇和语法的连续统之间存在许多兼有词汇和语法特征的半固定短语，而言使用的过程就是将这些短语连接成话语的过程。

语料库驱动的短语学研究发现，意义体现于多个词及语法结构共选构成的短语之中。语料库驱动的研究方法挑战了以前的语言理论和范式，对传统的语法范畴提出质疑，认为词汇和语法、结构和意义之间没有截然界限，词汇和语法互相选择，结构和意义密切相关。Sinclair (2004)描述了意义单位和扩展的意义单位，认为这是一种具有一定抽象度的语义单位，是语法和词汇、词汇和词汇共选关系的体现。扩展意义单位的研究范畴包括节点词的搭配选择、类联接选择、语义趋向和语义韵。其中搭配选择属于组合关系，类联接属于聚合关系。这一体系将两种关系协调，兼顾横向的组合关系和纵向的聚合关系，从而改变了以往的研究重聚合、轻组合的作法。意义单位打破了词汇与语法之间的界限，将两者视为一体，既涵盖具体的词汇与句法特点，也包括说话人的意图及交际功能，是一个集形式、意义、功能于一体的概念。如果将句法和语义割裂开来，则不符合语言的本质和规律，有碍于外语研究和学习的进程。

Corpus-based literary interpretation of Hemingway novels—An exploratory study

Jian Wang Shanghai International Studies University

This dissertation intends to expand the application scope of corpus linguistics from language acquisition and translation to literary interpretation. This thesis aims to achieve three goals: 1) to demonstrate the usefulness of corpus linguistic techniques in a stylistic analysis of literary texts and corpora; 2) to gain stylistic and literary insights into the data collected from Hemingway novels with the help of corpus linguistic techniques; and 3) to explore the potential of the techniques used in the analyses so that they can be adapted to future analysis of literary texts and corpora.

Firstly, several corpora have been established and compared among each other

through WordSmith Tools, a corpus software, so as to extract keywords out of each of the four Hemingway major novels. Then, they are further analyzed so that Hemingway's personal preferences for different parts of speech are discovered, such as simple verbs, rare usage of adjectives and two successive prepositions.

Secondly, patterns of phrases in the four novels are extracted through another software kfNgram. Compared with the patterns in Hemingway's contemporary literary works, some characteristics conveyed by the phrases are found to be Hemingway's idiolect.

Finally, based on the themes, concordance lines of the keywords as node words are examined in details and the protagonists and the themes in the four novels are compared and discussed. Thus, traditional literary interpretations of these novels and Hemingway's style of iceberg principle have been partially proved.

The study has resulted in a few achievements. Practically, corpora of literary texts have been founded and the four Hemingway major novels have been diachronically studied and some new insights have been gained: 1) Hemingway has a tendency to use two prepositions together; 2) Catherine plays the role of the exemplary hero in the novel *A Farewell to Arms*; 3) The novel *The Old Man and the Sea* is very different from the other three novels.

双语平行语料库创建与加工的策略

熊文新 北京外国语大学中国外语教育研究中心

语料库是客观的、可靠的语言知识的来源，对语言研究具有重要意义（冯志伟，2011）。语料库建设既是一个科学问题（理论与架构），也是一个工程问题（组织与实施）。由于语料相对较难获得、对齐加工处理复杂，涉及不同语言的多（双）语平行语料库的创建尤其困难。本文以（熊文新，2013a, 2013b）为基础，提出利用 Web 资源，“分步骤按领域”创建“高大全”式英汉平行语料库的通用建设方法。

全文主要观点如下：

（1） 辨析 Web 和语料库的关系（Baroni & Bernardin, 2006），充分借助网上日益涌现的双语电子文本，解决传统方法由于采集困难、领域单一造成双语语料库数量、质量难以保证的弊病。

（2） 参照“子语言（*sublanguage*）”（Harris, 1989）的思想，选定领域专一、语言可靠、格式规范的文本，逐次建设特定领域的语料库，最后汇总成高质量、大规模、全领域的“高大全”式双语语料库。

（3） 借鉴“基于转换、错误驱动”（Brill, 1995）的处理策略，充分利用现有基于统计的自动句对齐工具，在已有较好对齐结果的基础上，根据当前处理语料的语言特点，有针对性地实施再对齐处理。

全文以一个实例贯穿始终，阐释如何在双语平行语料库的工程应用中贯彻实现以上原则，具体流程包括 Web 语料对象选择、网页采集、超文本页面的净化、纯文本的预处理、自动句对齐以及再对齐处理等。实验证明采用以上步骤能够较好地解决双语平行语料库的采集和质量瓶颈问题。

An Empirical Account of Coordination Based on a Corpus of Spoken and Written British English

Alex Chengyu Fang City University of Hong Kong

Coordination and subordination are pervasive in English syntax observable at all the possible junctures in the parse tree. Past studies in language variation have predominantly focused on subordination and shown that it exhibits different patterns of occurrence across different genres and registers. For example, adverbial clauses have been shown to be characteristic of formal discourse. This talk will address the issue of coordination and aim to investigate whether, like subordination, it also demonstrates a strong correlation with different degrees of formality. Methodologically, the ICE-GB corpus will be used for its detailed syntactic analyses as well as its extensive text categories. The primary data will consist of the complete corpus divided into speech and writing. The secondary data will consist of a subset of the corpus, divided into spontaneous speech, prepared speech, timed university essays and untimed university essays. Statistics will be reported regarding coordination of different syntactic categories such as the word, the phrase and the clause as well as different syntactic functions of the constituents. As will be shown through empirical evidence, like subordination, coordination demonstrates a positive correlation with text formality, suggesting that the interplay of the two constructions is a primary factor that contributes to syntactic complexities. The study will also show that coordination in speech and writing exhibits a preference for different syntactic categories: while clausal coordination is found to be preferred in spoken texts, the coordination of noun phrases is found to characterise written texts. This finding lends support to the past claim that writing is marked for a higher density of complex nominal groups. However, the higher proportion of clausal coordination found in the spoken texts cannot be taken to argue for a higher syntactic complexity in speech in general, an issue that warrants further research in the future.

A Corpus-based Contrastive Study on the Use of Signaling Nouns in NS and NNS Academic Writing

Fan Pan Hua Zhong University of Science & Technology

Signaling nouns are cohesive in nature to create textural cohesion in extended stretched of information in academic discourse. Since 1970s, researchers have become increasingly interested in this special type of abstract nouns. Various studies reveal that signaling nouns are highly prevalent in academic context and the acquisition of these nouns is one of the major challenges for ESL/EFL learners. Although considerable research has been done on the various aspects of signaling nouns used by NS academic writers, much less is known as to how non-native academic writers actually use these signaling nouns.

Using a corpus-based methodology, the present study aims to find out the

differences between the NS and NNS academic writing in terms of frequency distributions and lexico-grammatical realization patterns of signaling nouns. Two corpora concerning the field of Applied Linguistics are compiled for the study. The 1.1 million NS corpus consists of 120 research articles published on 6 leading international Linguistics journals during the year 2009 and 2010. The 0.6 million NNS corpus comprises 120 research articles published on one English journal issued by a Chinese institution during the same time period. By referring to Hinkel's (2004) list, 35 highly prevalent signaling nouns with a cohesive function in writing are identified from two corpora and then manually checked carefully to exclude the unqualified ones. The frequency distributions of the 35 signaling nouns in two corpora are compared by employing Chi-Square statistics to find out whether there is significant difference between NS and NNS corpora. By adopting Flowerdew's (2010) framework of lexico-grammatical realization patterns, the study also analyzes and compares the distribution of lexico-grammatical realization patterns of the six most frequent signaling nouns in NS and NNS corpora.

The study has the following findings. First, among 35 signaling nouns investigated in the present study, 10 signaling nouns have higher frequency in NS corpus than that in NNS corpus, but only 4 signaling nouns (effect, fact, factor, manner) reach the statistic significance. However, 11 signaling nouns (process, result, purpose, reason, aspect, problem, tendency, method, stage, topic, trend) have significantly higher frequency in NNS corpus than that in NS corpus; Second, in terms of lexical-grammatical realization patterns, some interesting differences and similarities are found. Though both NS and NNS writer use more anaphoric patterns than cataphoric patterns, NS writers seem to use across-clause patterns (anaphoric and cataphoric patterns) much more frequently than NNS writers. A detailed analysis into the in-clause patterns shows that, the most frequently used in-clause patterns of 6 top signaling nouns appear to be almost the same in both corpora, e.g. SN+of+NP is the dominant in-clause patterns of 5 signaling nouns (effect, type, form, process, factor) and SN+that clause is the dominant in-clause patterns of fact in both corpora. Nevertheless, NS writers use a larger variety of in-clause patterns than NNS writers. The findings indicate that, though Chinese linguistic writers (represented in the present study) use signaling nouns as frequently as NS writers, they tend to use the nouns in relatively limited types of in-clause patterns. Some pedagogical implications for EAP teaching of signaling nouns are also discussed in the study.

On Minimal Assumptions in Corpus Linguistic Studies

Naixing Wei Beihang University

There have been long-standing debates between the so-called corpus-based method and corpus-driven method in corpus linguistics, which have been held to be two sharply differing methods, even two fundamentally different approaches, which carve up current studies into two seemingly irreconcilable broad strands. Rigorous distinctions have been drawn (Tognini-Bonelli 2001) between the two methods and

strong criticisms leveled at a rival method from a different strand of research (McEnery & Hardie (2012). However, at the very heart of the Sinclairean corpus linguistics is the important idea of Minimal Assumptions (Mahlberg 2007a, 2007b), which has been neglected in on-going quarrels and arguments. This paper discusses the idea of Minimal Assumptions for corpus linguistics (1991, 1996, 2003, 2004, etc.) and addresses its usefulness in corpus linguistics in general. I will first look at Sinclair's original statements regarding the idea of minimal assumptions for an understanding of its essence and implications, before moving on to a more in-depth discussion of its central features and relationships with broad issues of corpus linguistics. More specifically, I focus on the following questions: 1) What are the central tenets of Minimal Assumptions and its cyclical principle in research? 2) In what relationships does the idea of Minimal Assumptions stand with the corpus-based and corpus-driven methods? 3) What are the relationships between theory and method in research? 4) What possible roles can the minimal assumptions method perform in future development of corpus linguistics? Throughout I will argue that the Sinclairean idea of Minimal Assumptions lies at the very heart of corpus-driven linguistics, which can be usefully applied to studies of differing strands, if they are intended to discover new meaningful facts and patterns and arrive at new interpretations.

Closed-class Keywords and Their Collocations: A Contrastive Analysis between Academic Texts of Experts and Chinese Learners

Bixuan Zhang Beihang University

The range of discourse analysis has been broadened by computer-assisted keyword analysis. Generally speaking, there are two frequently applied approaches: study focuses on words at the very top of the keyword list and disregarding words that fall below an arbitrarily set cut-off point; study focuses on certain categories of keywords and discard other categories, and categories selected are often open-class words such as nouns, verbs, adjectives and adverbs. There is considerable merit in these two approaches: words at the top of the keyword list have the highest keyness scores, thus can clearly reveal the features of certain discourse; open-class words contain more concrete meaning than grammatical words such as conjunctions, prepositions and pronouns and thus can be effective indicators of the "aboutness" of a text. However, a large quantity of corpus evidence has proved that meaning does not reside in a single word but in a sequence of words in certain context. To fill the gap, the current study conducts a corpus-based keyword analysis, focusing on closed-class keywords and their formulaic sequences in academic texts written by Chinese learners and native scholars.

The present study is a corpus-driven study in which no hypothesis is made before the study. Two specialized corpora are compiled. One contains all graduation theses of Chinese postgraduate students major in linguistics in colleges and universities of the 985 project, with a total number of running words for the whole corpus being approximately 1,005,000 words. The other corpus includes native scholars' academic

research articles published in internationally renowned journals with high impact factors in the field of linguistics, and the total number of running words for the whole corpus is approximately 1,001,000 words.

Both quantitative and qualitative analysis will be used in the present study, and corpus tool used is WordSmith Tool 4.0. The quantitative analysis will focus on surface closed-class keywords features of each corpus. The qualitative analysis reveals the different formulaic expressions of closed-class keywords as well as the hidden evaluative meaning and ideology in each corpus.

The present study starts with keywords analysis. The chi-square algorithm will be adopted in the keyness analysis. Two corpora analyzed in the current study will be mutual reference to each other in the keywords analysis. The current study focuses on the closed-class keywords, which are usually of high frequency in texts, thus the benchmark for minimum frequency of words analyzed is set to 50, higher than normal keyword analysis.

Closed-class keywords in each keyword list will be selected out for further collocational study, which adopts the method defined as “pattern flow” by Hunston and Francis (1998:68–69, 2000:207–215). “The observation of certain recurring expression is conducted both hierarchically, showing the keywords as a set of interlocking component units, and linearly, prioritizing the interdependence of lexical items and the grammar pattern with which they co-occur.” The analysis of formulaic expressions begins with a concordance analysis using certain closed-class keyword as node word, followed by a further concordance analysis using the pattern “collocates + closed-class keyword” as node words.

Several linguistic features as well as their hidden ideologies are discovered in the current study. For instance, closed-class keyword analysis shows that pronouns over-used in by Chinese learners are “them” “they” “their” “he” “it” and “its”, while pronouns under-used are “itself” “who” “she” “my” “her” and “I”. Comparison of these pronouns reveals the stronger group consciousness, and the less feminine consciousness and individual awareness of Chinese learners than native scholars.

Formulaic expressions analysis also indicates features of two corpora. For instance, analysis shows patterns with “into” are mainly used in the following contexts:

Segmentation: BE divided/ classified/ grouped/ categorized into ...(number) categories/ groups/ parts etc.

Transformation: BE translated into Chinese/ English, Condition A TURN into Condition B etc.

Utilization & Merging: Put into effect/ practice, Take/ bring/ integrate/ combine/ assimilate/ introduce information into existing units etc.

Investigation & research: Investigation into question, insight into question
These four types of formulaic expressions can be found in both corpora. However, a follow-up concordance analysis shows pattern with the meaning of segmentation covers the greatest proportion in academic texts of Chinese learners. In academic texts of native scholars, on the other hand, the most frequent meaning of “into” pattern is the meaning of investigation and research, followed by the meaning of utilization and

merging. This difference might be an indicator of differences of contents of academic texts written by Chinese learners and native scholars. Native scholars pay more attention on empirical study, thus use more patterns related with empirical study. Chinese learners, on the other hand, are more tend to conduct a theory-based analysis. Patterns related with theoretical analysis, therefore, are more often used, and a follow-up examination finds that patterns with the meaning of segmentation are often used by Chinese learners when introduce trends or sub-divisions of a certain theory.

拓展历史认知语言学研究新视野的工具

李华勇 四川文理学院

近 20 年来多年来, 200 多年前开创的历史语言学和现代认知语言学逐渐开始融合诞生了历史认知语言学。认知语言学有“基于用法的模式”来研究语言事实背后的理据性的传统, 这使历史认知语言学就有与语料库语言学结合的天然优势来观察相关语言现象和语言结构的历史变化过程、追寻其历史变化事实背后的内在规律。文章以具体的实例分别详细说明了 COHA 在对美语词汇、形态、句法、语义到文化方面的作用, 这些实例说明 COHA 极大地拓展了历史认知语言学研究的新视野, 也为历史认知语言学研究提供了强有力的工具保证。

基于语料库的新闻英语听力词汇习得

李威 东北师范大学

本文首先根据 Sinclair 的词汇语法理论、新闻英语词汇特征和语音及音位特征构建新闻英语听力的词汇习得框架。该框架包括: 1. 新闻英语特征词汇习得: 新闻小词、专有名词、以及这些词的搭配、类链接、词块; 2. 新闻英语通用词汇习得: 高频词、以及这些词的搭配、类链接、词块; 3. 新闻英语的语音和音位特征习得: 词汇的发音、重音、连读、语调。然后以自建多模态新闻英语地震语料库为例, 介绍语料库的建立, 听力词汇习得内容的提取, 并详述新闻英语听力词汇的习得方法。习得方法以习得策略为指导, 听入和读入相结合, 传统的词汇习得方法和 DDL 的词汇习得方法相结合, 计算机辅助学习与手机辅助学习相结合。目的是寻找一种适合英语专业学生学习实际的、高效的词汇习得方法, 增加新闻英语的听力词汇, 提高新闻英语的听力理解水平。

“适者生存”在《白牙》中的阐释——基于语料库的《白牙》文本检索分析

刘宇婷 河南师范大学

本文应用相关的语料库检索软件, 对美国著名作家杰克·伦敦的小说《白牙》进行分析, 旨在探究杰克·伦敦在《白牙》中对“适者生存”哲学思想做出的阐释。首先文章通过对小说主人公白牙发出的动作动词词表的分析, 发现白牙发出的动作动词中体现适应性的动词占得频率很高。然后本文再对频率较高的几个适应性动词索引行进行观察分析, 发现白牙一直在努力地学习、提高、改变来适应不同

的生存环境。

基于语料库的英语阅读资源评估初探

毛越浩, 曹竟中 南财经政法大学

英语一直是我国高等教育重要的一部分, 拥有适合自身英语水平的学习资料对学习者来说至关重要。然而, 寻找这样的资料却并不容易。网络上、市面上充斥着大量针对不同水平学习者的复习资料, 但它们内容的难易程度却往往与真实考试的要求有着或大或小的差异。使用这样的学习资料, 学习者虽然付出了许多的时间, 精力与金钱, 但却并不能达到理想的学习效果。因此, 怎样有效的发现适合不同水平学习者的学习资料成为一个值得探究的话题。不过, 目前国内对于这方面的研究却寥寥无几。本研究以英语专业四级考试的阅读为例探讨英语阅读材料的选取和评估。

《高等学校英语专业英语教学大纲》(2000年修订版)指出, 达到英语专业四级水平应能读懂难度相当于美国 Newsweek 的国际新闻报道; 能读懂难度相当于 Sons and Lovers 的文学原著。基于这一要求, 本研究使用加拿大 Cobb 教授研发创建的 Compleat Lexical Tutor 网站中词汇分析工具分别对历年专业四级考试阅读真题、Newsweek 的新闻报道以及小说 Sons and Lovers 等语料进行分析。该评估是基于最常用 2000 词汇 (Michael West, 1953) 学术词汇 (Coxhead, 2000) 以及专业术语的分布比例的考量, 同时也考察文本的类符/形符比和词汇密度。结果显示以上材料的词汇分布比例非常相近, 表明它们的难度是相似的, 同时也说明使用该词汇分析工具来对文本的难度做分析是可行的。由此可见, 教师和学习者可借助该工具对获取到的文本进行分析, 量身定做符合不同学习层次要求的阅读资料, 减少不良资料的干扰, 从而提高学习效率。同时, 各类学习资料的作者和出版商也可以此工具作为参考, 对其作品进行校验。

现代汉语量词“副”的语义韵研究

王均松, 田建国 西北工业大学

语义韵的概念和理论首先由西方学者提出, 研究的对象也是以英语为代表的印欧语系语言为主, 其理论和方法是否具有普适性是一个非常值得关注的课题。而量词是汉藏语系区别于印欧语系的一种特殊语言现象, 因此汉语量词的语义韵研究对于建构完备的语义韵理论体系具有重要价值和意义。本文以 Sinclair 提出的扩展意义单位模型为理论框架, 基于现代汉语语料库对量词“副”的语义韵进行了全面的调查和分析。笔者首先在国家语委现代汉语语料库中以“副”为关键词进行检索, 共提取 330 条索引行。经过观察发现, 其中 187 条索引行中的“副”主要用于具体名词前, 表示成对或配套的物品, 如“一副眼镜”、“一副手套”、“一副对联”等; 其余 143 条索引行中的“一副”则主要用于描述人物的外部形象或面部表情, 如“一副老态龙钟的样子”、“一副冷冰冰的脸色”、“一副轻松的神态”等。鉴于“一副”用于具体名词前时往往侧重对于事物的客观描述, 说话人的态度倾向或评价意义较弱, 因此本研究重点考察“一副”用于神态或面部表情时的语义韵特征。研究表明, 现代汉语量词“副”在描述人物的神态或面部表情时主要出现在

“V+一副+ADJ+N”和“一副+ADJ+N”两种类联接型式中，其典型的搭配结构为“一副 X 的样子”。通过对随机选取的 50 条索引行的逐一验证发现，节点词结构表达消极语义趋向的有 39 条，占总数的 78%；表达积极语义趋向的只有 7 条，占总数的 14%。从总体上来讲，量词“副”用于描述人物的神态或面部表情时具有较强的消极语义趋向，传递否定或批判的语义韵。

综上所述，首先，语义韵不是印欧语系所独有的语言现象，它在汉语中也是普遍存在的，现代汉语量词“副”即是一个典型的例子。其次，语义韵是一种隐藏的评价意义，必须通过大量的索引行证据，结合上下文语境才能够确定。最后，语义韵由整个扩展意义单位共同营造，因此语义韵应置于扩展意义单位模型中进行研究。

A Multi-dimensional Analysis on Corporate Blog Yang Wu, Zhoulin Ruan Xi'an Jiaotong Liverpool University

This study employed a multi-dimensional (MD) analysis to investigate the co-occurring patterns of linguistic features and the functional linguistic variation in corporate blogs (CB). The corpus came from 1048 posts in 41 top ranking corporate blogs (among top 50), which includes textual, audio, visual, or commentary content. Among these posts, 1020 texts were selected to represent a corporate blog language because they have met such criteria as being currently active, written in English language, text-based, and containing more than 100 words, etc.

In the multi-dimensional analysis, 32 linguistic features were tagged, counted and normalized. The normalized frequencies of these features were then entered into a factor analysis to find the co-occurring patterns. Findings indicate that there are six co-occurring patterns which can be named as follows: (1) personal involvement narration, (2) informational production, (3) interactive persuasion, (4) abstract/impersonal style, (5) evaluative stance, and (6) general information vs. technical information.

Findings also indicate that corporate blog registers cover a broad range of dimensions, characterized by a mixture of those reflected in earlier studies, with different sets of co-occurring features reflecting varied contextual functions. Another important finding of this study is that most dimensions show a weak polarization between spoken and written registers. These findings add to our understanding of the description of corporate blogs as a language variety in its own right, and provide some practical implications for effective corporate communications as well.

框架语义学与“好”字汉英翻译中的认知对比——一项基于语料库的研究 杨元媛 商丘师范学院 宋军 东华大学

框架语义学 (Frame Semantics) 理论是认知语言学的一个重要组成部分，要理解系统中的任何一个概念或词汇，都要基于人类经验、思维背景下形成的整体“框架”，以及在这个框架之下的具体语义。框架语义学理论可以被应用于英汉翻

译研究中。

汉字“好”经常被英译为 *good*，而基于汉语语料库 Chinese Gigaword Second Edition，英文语料库 British National Corpus，运用软件 Sketch Engine，可以探知两词词汇框架的不同之处。

分别以“好”和 *good* 作节点词，统计其搭配词的构式分布，发现“好”左一位为“更”、“较”、“很”等程度副词，而 *good* 左一位为冠词 *a* 和 *the*；“好”与 *good* 右二位以逗号、句号居多，除此以外，“好”右二位上使用名词或动词，而 *good* 右二位上使用介词 *of* 或 *to*。统计词项的 LogDice 值，发现与“好”搭配力强的词汇丰富，且较多为副词和实词；与 *good* 搭配力强的词汇略显简单单调，且虚词占很大比例。这种框架中填充项不同的特点是由两种语言的背景和概念结构决定的。

汉字“好”最常用的搭配框架为“好的效果”、“好的方向”等；其次为“好”字构成的名词化结构，如“好风气”、“好势头”；另外，“谁好”、“成绩好”这样的句法结构也占一定比例；“好”字之前还经常出现否定词“不”，形成如“干不好工作”、“治不好病”等模式。而 *good* 应用的结构和语境有所不同，最常见于 *good idea*, *good news* 类结构；其次为 *very good*, *perfectly good* 等句法形式；*good and gracious king*, *good old days* 等 *good* 与其他形容词连用的模式也很常见；*the weather is good*, *the snow was good* 形式等频繁出现，表现出西方人对天气状况的关注；*good over evil* 这一配位模式出现频度相当高，且 *good over* 之后只有 *evil* 一个固定论元，表现出西方文化中，善良最终能够战胜邪恶的生活信念。“好”与 *good* 近义词不同用法的分析也体现出中西方认知的差异。

在翻译中，应对比不同文化背景下的认知差异，选择与源语言相匹配的目的语词汇及框架，以使译文最接近于目的语使用者的语言习惯，达到最佳翻译效果。

Interactions between lexical items and Mandarin multi-verb constructions

Hui Yin Xi'an Jiaotong-liverpool University

In the literature, there has been little systematic exploration of Mandarin multi-verb constructions (MVCs) from corpora. This research is sought to fill this gap by exploring multi-verb constructions from the Lancaster Corpus of Mandarin Chinese—a publicly available balanced corpus containing one million written words. This study proposes that constructions may be characterized by particular patterns of lexical diversity and attractions of particular lexical items. One can track a certain MVC's lexical productivity (openness) or lexical restrictedness (fixed or preferred verbs in a particular slot). The corpus data show there is a strong interaction between lexical items and construction types. The findings reveal that generally there is a clear lexical attraction/affinity between certain verbs and the different slots of a multi-verb construction. Some verbs are easily attracted to a particular construction, while others are largely excluded from it (Stefanowitsch & Gries 2003). The attraction scores indicate that some verbs are more important than others for a given construction (Schmid 2000, Yin 2010). The corpus results suggest that high attraction of particular lexical items to V1 or V2 is symptomatic of being considered a somewhat specific construction. The corpus results also indicate that certain verbs are compatible with a given construction and may heavily color constructional meaning or bring the

particular functional value to the overall construction.

The Synchronic and Diachronic Study of the Use of Progressive Statives

Fenghua Xu, Jidong Zhang Donghua University

The acceptability of progressive statives has been supported by both the empirical study and corpora examples in the literature. The purpose of this research is to conduct a corpus-based synchronic and diachronic study of the use of progressive statives. According to the stative verbs used in Gennari & Poeppel (2003) and the tests for classifying verb categories, this research uses the Brown Family Corpus (1960), Frown Family Corpus (1990) and American English 2006 to study the use of progressive statives in American English. At the same time, this research uses the LOB Corpus (1960), FLOB Corpus (1990) and British English 2006 to study the use of progressive statives in British English. Furthermore, the similarities and differences between the American English and British English in the use of progressive statives are studied. The results reveal that the corpora representing the American English and British English respectively in the same year show that there are more stative verbs used in the progressive form in the British English than in the American English. The diachronic study reveals that, the use of progressive statives does not increase significantly with the change of time in number and frequency (except for certain stative verbs).

汉语指称性“X的”结构使用规律的认知阐释

张立飞 解放军外国语学院

“X的”是汉语中简化了的指称形式，可单独用来指称语境中的一个名词性成分，其初始动因是由经济性原则驱使的中心语省略。严辰松（2007）认为指称性“X的”功能在于从由中心语所指的事物中挑出特定的实体或类别，以区分于其他实体或类别，因而“X的”的各种类型，只要具有区分的功能，就潜在地具有指称功能。因此，从理论上讲，X本身的属性以及X与中心语之间的关系不会对“X的”单独用来指称构成障碍，只要语境适宜，“X的”就可用于指称。既然如此，那么“X的”在实际使用中呈现出何种特点呢？它在实际语料中的分布是否是均匀的呢？本文利用LCMC语料库对这个问题进行了探讨。

通过对索引行进行观察，我们选择从两个维度对“X的”用例进行手工标注，进而考察“X的”在实际使用中体现出的规律性。维度一：“X的”缺失的中心语是否在上下文语境中出现？维度二：X的语法属性，即是名词性的、动词性的、形容性的还是分句？经统计，我们发现维度一和维度二之间存在很强的关联，即不同类别的“X的”结构在有或无中心语的情况下的使用频率有很大起伏，如下表所示：

X的类别	合计				
名词性	形容词性	动词性	小句		
有中心语	71	67	131	12	281
	25.3%	23.8%	46.6%	4.3%	100.0%

无中心语	39	158	831	79	1107
	3.5%	14.3%	75.1%	7.1%	100.0%
合计	110	225	962	91	1388
	7.9%	16.2%	69.3%	6.6%	100.0%

从上表我们可以看出，相比中心语出现的语境，在中心语不出现的语境中，名词性和形容词性 X 加“的”用于单独指称的比例均有下滑，此消彼长，动词性和小句 X 加“的”用于指称的比例均有增长。其中，变化最明显要数动词性 X，从原本的 46.6% 增至 75.1%，其次是名词性 X，从 25.3% 减至 3.5%，而且它在无中心语的语境中出现的频率（71）竟然比有中心语的语境中的（39）还要低。可以说，从有中心语语境向无中心语语境的过渡中，名词性 X 和动词性 X 在构成单独指称的“X 的”结构时的表现截然相反。

“X 的”在实际语料中的上述分布规律从语法上体现了自主/依存的概念对立。概念的自主或依存是 Langacker 建构认知语法体系时赋予语言单位的语义属性。自主结构（*autonomous structure*）是指可以独立出现的、自身的语义描述不预设另一结构的音位结构或语义结构，而依存结构（*dependent structure*）是指其身的语义描述需预设另一结构的存在，或需另一结构作为支撑。按照 Langacker (1987; 2008) 的观点，名词表示实体，其语义描述无需预设其他成分，因而是概念自主的，而动词表示事物之间的关系，是概念依存的，其语义描述必定要预设事物。概念自主的名词做修饰语时，它和被修饰的成分之间的语义关系并不是与生俱来的、本质上的。在中心语不出现的情况下，名词性的 X 若用于指称性的“X 的”结构，会导致中心语的提取困难，从而使用受到限制。但概念依存的动词则没有上述限制。由于对动词的语义进行描述预设了的事物的存在，且该事物为相应动作的参与者。如此一来，当动词做修饰语时，其修饰的对象通常为动词的内部论元，是动作发生不可或缺的参与者。由于动词与中心语之间的语义关系密切，使得动词对于中心语有一定的预测性。这就导致即便是省去中心语，受话人也能对指称对象进行提取（动词缺失的论元即为指称中心语），因而使用频率较高。

A Contrastive Study of Nominalizations in Chinese Media English and British Media English

Alex Chengyu Fang, Liu Ying City University of Hong Kong

Variation is inherent in human language and linguistic variation is crucial for the study of language use. Influential frameworks such as Biber's multi-dimensional approach have greatly facilitated variational studies across registers, genres, even across different languages. With increasing numbers of people learning and using English all over the world, there is a growing recognition of "Englishes" in the plural sense such as world Englishes and international Englishes over the last three decades. The world English context has greatly extended the scope of variational studies, from register variation and genre variation to English variety variation. Among various varieties of English, China English began to draw more and more attention both at home and abroad due to the rise of China and its largest number of English learners in the world. Previous studies of English variety variation have concentrated on comparing two native varieties – British English and American English or comparing a

few non-native varieties like Indian English and Australian English with the two native varieties. However, the comparison is often made on a few lexical and grammatical features such as modal verbs, articles, or subjunctives and largely ignores variations at the syntactic level. Nominalization, as a syntactic phenomenon, is pervasive in language and has received much attention in linguistic enquiry. Previous studies show that nominalization is an important feature for linguistic variations. The distribution of nominalizations varies across registers and the more formal a text is, the more uses of nominalizations in it. But how the uses of nominalization vary across different English varieties is still an underexplored area. Seeing such a gap, the present study adopted a syntactico-structural approach to explore the uses of nominalizations in China English and British English by using data from two corpora – the Chinese Media English Corpus and the British Media English Corpus. This study has found that there are various differences in the uses of nominalizations in China English and British English in terms of structure, meaning and function. The study of nominalizations in two English varieties will not only deepen our understanding of linguistic variations across English varieties at the syntactic level but also provide a framework for identifying variational indicators for distinguishing English varieties.

What are the Things in the Source Domain of Marriage Metaphor?

Xiujing Ni University of Shanghai for Science and Technology

Metaphor is not just a rhetorical device, but a way of thinking, a tool for understanding and a matter of reason, too. Previous researches on metaphor were mainly based on introspection and a considerable amount of data was of intuitive in nature, rendering the conclusions unconvincing to some extent. This paper attempts to find the possible things in the source domain of marriage metaphor with the help of on-line corpora and other on-line resources.

美国大学校长毕业典礼演讲语料特征分析及启示

许霄羽 西北工业大学

作为演讲的一种形式，毕业典礼演讲主要由校长演讲与嘉宾演讲构成，演讲文本具有口语与书面语双重语言特征，由于毕业典礼受众的独特性与演讲目标的指向性，毕业典礼演讲在语言选择与价值观传递上都具有自身特点。论文以哈佛大学、普林斯顿大学、麻省理工学院等排名全美高校前 15 名的 10 篇大学校长毕业典礼演讲文本为研究对象，对美国大学校长毕业典礼演讲语言与价值观特点进行分析，进而为我国大学毕业典礼演讲提供借鉴。

CLIC2013 第二届中国语料库语言学大会参会指南

说明

18 日全天报到，时间（9：00-21：00），新沙珑国际宾馆（东华大学校园内无接待）。

新沙珑国际宾馆

宾馆位于上海市松江区梅家浜路 1501 号（梅家浜路和嘉松南路交叉口附近）。上海地铁 9 号线出站口西南（左前方）300 米处高楼下即是。照片如下：



上海各车站至松江大学城路线指南

松江南站（火车站）

乘坐地铁 9 号线（往“杨高中路”方向）至松江大学城站下（打车费约 37 元）。

虹桥机场或虹桥火车站（打车费 约 130 元）

线路 1

乘坐地铁 10 号线（往“新江湾城”方向）至“虹桥路”，换乘 3 号线或 4 号线至“宜山路”（1 站），换乘地铁九号线（往“松江南站”方向）至“松江大学城”站下。

线路 2

步行至“虹桥西交通中心”（约 5-10 分钟），乘坐虹桥枢纽 10 路公交车至“松江城东”站（票价 9 元），转乘松江 1 路公交车至“松江客运中心”（票价 1 元；松江大学城地铁站旁）下。

浦东机场（打车费约 270 元）

步行到地铁 2 号线浦东国际机场站，乘坐地铁 2 号线延伸段（“广兰路”方向）到“广兰路”下（注意收听提示地铁语言提示），换乘坐地铁 2 号线（“徐泾东”方向）至“世纪大道”下，换乘地铁 9 号线（“松江南站”方向）至“松江大学城”下。

上海火车站北广场（打车费约 105 元）

乘坐轨道交通 3 号线（“上海南站”方向）至“宜山路”下，换乘轨道交通 9 号线（“松江南站”方向）至“松江大学城”站下。

上海火车站南广场（打车费约 105 元）

乘坐地铁 1 号线（“莘庄”方向）至“徐家汇”站下车，换乘地铁 9 号线（“松江

南站”方向)至“松江大学城”站下。

上海火车南站 (打车费约 85 元)

乘坐地铁 3 号线(“江杨北路”方向)至“宜山路”站下,换乘地铁 9 号线(“松江南站”方向)至“松江大学城”站下。

注:

- ① 地铁换乘均为站内换乘,请勿出站;
- ② 虹桥火车站与虹桥机场 2 号航站楼相连;
- ③ 若您于 1 号航站楼到达,请先乘摆渡车至 2 号航站楼;
- ④ 虹桥枢纽 10 路运营时间:
 - 往“松江城东站”方向 6:00—23:00
 - 往“虹桥西交通中心”方向 5:00—22:00
- ⑤ “松江城东站”→“虹桥西交通中心”,中途无停靠站(返程可以考虑此线路)。

编号	姓名	单位
1	卫乃兴	北京航空航天大学
2	张碧萱	北京航空航天大学外国语学院
3	翟萌	北京航空航天大学外国语学院
4	李潇辰	北京航空航天大学外国语学院
5	陈杰	北京体育大学
6	崔珣丽	北京体育大学 外语系
7	贾云龙	北京外国语大学
8	刘磊	北京外国语大学 外语教育研究中心
9	于涛	北京外国语大学 中国外语教育 研究中心
10	许家金	北京外国语大学 中国外语教育 研究中心
11	熊文新	北京外国语大学 中国外语教育 研究中心
12	刘燕	北京外国语大学 中国外语教育 研究中心
13	刘霞	北京外国语大学 中国外语教育 研究中心
14	梁茂成	北京外国语大学 中国外语教育 研究中心
15	李文中	北京外国语大学 中国外语教育 研究中心
16	吉洁	北京外国语大学 中国外语教育 研究中心
17	房印杰	北京邮电大学 人文学院 北京外国语大学 中国外语教育 研究中心

18	吴进善	北外外研中心 河南师范大学
19	袁新华	北外外研中心 泰山学院
20	仇桂珍	北外外研中心 新乡医学院
21	田纯广	滨州学院外语系
22	邓耀臣	大连海事大学 外语学院
23	程昕	大连海事大学 外语学院
24	李威	东北师范大学 人文学院英语系
25	陈功	对外经济贸易大学英语学院
26	毛浩然	福建师范大学 外国语学院 英语教育系
27	张绪华	复旦大学外文学院
28	胡玉迎	广东外语外贸大学 英语教育学院
29	权立宏	广东外语外贸大学英语教育学院
30	刘君	广西大学 外国语学院日语系
31	冉志晗	合肥学院外语系
32	杨素香	河南理工大学 外国语学院
33	王胜利	河南农业大学 外语学院
34	张绵	河南师范大学
35	孙海燕	河南师范大学 外国语学院
36	刘宇婷	河南师范大学 外国语学院
37	刘晓燕	河南师范大学 外国语学院
38	刘国兵	河南师范大学 外国语学院
39	彭杰	红河学院 外国语学院
40	田臻	湖南大学外国语与国际教育
41	邓飞	华南师范大学

42	何安平	华南师范大学 外国语言文化学院
43	潘璠	华中科技大学 外国语学院
44	龙绍赟	江西师范大学
45	张立英	解放军外国语学院
46	张立飞	解放军外国语学院学报编辑部
47	赵志敏	解放军外国语学院英语系
48	陈昕	解放军外国语学院英语系
49	张金福	九江学院 外国语学院
50	林程	九江学院 外国语学院
51	濮建忠	洛阳外国语学院 语言工程系
52	杨震宇	内蒙古大学
53	金罡	内蒙古师范大学 蒙古学学院
54	叶章勇	宁波职业技术学院
55	种法娟	青岛大学 汉语言学院
56	金菊花	山东大学 外国语学院
57	郭放	山东科技大学 外国语学院
58	杨元媛	商丘师范学院 外语学院
59	吴雪颖	上海大学 外国语学院
60	吴羊	上海大学外语学院 西交利物浦大学
61	杨海燕	上海对外经贸大学外国语学院
62	杨媛媛	上海工程技术大学基础教学学院
63	刘国辉	上海海事大学
64	甄凤超	上海交通大学 外国语学院
65	倪修璟	上海理工大学 外语学院
66	谭业升	上海外国语大学《外国语》编辑部
67	朱玉山	上海外国语大学

		出国培训部
68	王俭	上海外国语大学 出国培训部
69	黄晓梅	西安理工大学
70	方秀才	上海外国语大学 研究生部
71	周丹	上海外国语大学 研究生院
72	兰凤利	上海中医药大学 外语教学中心
73	宋伟华	韶关学院外语学院
74	龙得银	四川文理学院
75	李华勇	四川文理学院
76	周福娟	苏州科技大学 外国语学院
77	Friedrich G. Wallner	维也纳大学哲学系
78	曹蔓	武汉理工大学
79	黄青	武汉理工大学 外国语学院
80	王余余	西北工业大学
81	王均松	西北工业大学
82	田建国	西北工业大学
83	贾艳	西北工业大学
84	巩海曦	西北工业大学 人文与经法学院 外国语言学及 应用语言学系
85	于璐	西北工业大学 人文与经法学院
86	陈鑫	西北工业大学 人文与经法学院 外语系
87	许霄羽	西北工业大学 人文与经法学院 外语教育系
88	阮周林	西交利物浦大学 英语与文化传播系
89	Hui Yin (印辉)	西交利物浦大学 英语与文化传播系
90	Dr. Alex Fang	香港城市大学

91	李艳娇	香港城市大学 人文社会科学学院中文语言翻译系
92	刘瑛	香港城市大学 中文语言翻译系
93	Dr. Lyn Flowerdew	香港科技大学
94	刘鼎甲	燕山大学
95	刘泽权	燕山大学 外国语学院
96	孙乐乐	扬州大学 外国语学院
97	陆军	扬州大学 外国语学院
98	刘克强	云南红河学院 外国语学院
99	郑青	云南红河学院 外国语学院英语系
100	杨晓东	浙江工业大学 外国语学院
101	焦俊峰	浙江理工大学 外语学院
102	张翼翼	中国地质大学（北京）外国语学院
103	毛越浩	中南财经政法大学
104	曹竞	中南财经政法大学外国语学院
105	贺霞	北京工业大学 外国语学院
106	赵晓临	东华大学外语学院
107	周红莉	东华大学外语学院
108	张晓明	东华大学外语学院
109	张莉	东华大学外语学院
110	张继东	东华大学外语学院
111	严齐	东华大学外语学院
112	徐风华	东华大学外语学院
113	吴蕾	东华大学外语学院
114	王荣明	东华大学外语学院
115	任再新	东华大学外语学院
116	马静	东华大学外语学院
117	林作帅	东华大学外语学院

