

第四届中国语料库语言学大会

会议手册

主办：中国英汉语比较研究会
语料库语言学专业委员会

承办：华中科技大学外国语学院

赞助：上海外语教育出版社

2018年6月1-3日

中国 武汉

The 4th Corpus Linguistics in China Conference

Conference Program

Organizer: Corpus Linguistics Society of China (CLSC)

Host: School of Foreign Languages

Huazhong University of Science and Technology

Sponsor: Shanghai Foreign Language Education Press

June 1-3, 2018

Wuhan, China

Scientific Committee

Deng Yaochen, Dalian University of Foreign Languages

He Anping, South China Normal University

Li Wenzhong, Zhejiang Gongshang University

Liang Maocheng, Beijing Foreign Studies University

Liu Guobing, Henan Normal University

Liu Zequan, Henan University

Pan Fan, Huazhong University of Science and Technology

Pu Jianzhong, Zhejiang Gongshang University

Qin Hongwu, Qufu Normal University

Wei Naixing, Beihang University

Xu Jiajin, Beijing Foreign Studies University

大会组委会

Conference Organizing Committee

President

Professor Xu Mingwu, School of Foreign Languages, HUST

Co-presidents

Professor Liang Maocheng, Beijing Foreign Studies University

Professor Lei Lei, School of Foreign Languages, HUST

Secretaries General

Feng Xuefang

Yang Hongyan

Secretaries

Ren Yu

Sun Yunmei

Wang Ruowei

Huang Chen

Gao Yu

Conference Affairs Supervisors

Wang Ruowei

Yang Hongyan

Huang Chen

Gao Yu

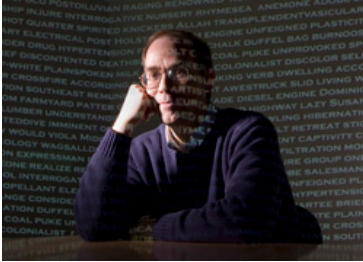
目 录 Contents

主旨发言人 Keynote Speakers	1
会议日程 Conference Agenda	4
分组讨论 Concurrent Session	6
论文摘要 Abstracts	20
主旨发言摘要 Abstracts of the Plenary Sessions	20
分组讨论摘要 Abstracts of Concurrent Sessions	26
通讯录 Contact List.....	89
校园地图 Campus Map.....	96
华中科技大学简介	97
Introduction to Huazhong University of Science and Technology	99
华中科技大学外国语学院简介	101
Introduction to School of Foreign Languages, HUST	102
华中科技大学外国语学院教师招聘启事.....	103
华中科技大学外国语学院博士后招聘启事.....	105
《语料库语言学》期刊征稿启事.....	106

主旨发言人 Keynote Speakers

Mark Davies

Brigham Young University



Mark Davies, Ph.D., is professor of Linguistics at Brigham Young University in Provo, Utah, USA. His primary fields of research are corpus linguistics, and language variation and change. He is the author of more than seventy articles and six books, and he has received several large grants from the US National Endowment for the Humanities (NEH) and the US National Science Foundation (NSF), to create and analyze corpora. He has created several large corpora that are available from corpus.byu.edu, most of which he has created by himself. These corpora are used by more than 130,000 distinct researchers, teachers, and students from throughout the world each month, which probably makes them the most widely-used corpora in existence. Data from these corpora are used for several hundred articles and books each year, and the data has been used by hundreds of companies to develop a wide range of language applications. More information at: davies-linguistics.byu.edu.

Jiang Feng

Jilin University



Feng (Kevin) Jiang is Kuang Yaming Distinguished Professor in applied linguistics in the School of Foreign Language Education at Jilin University, China and gained his PhD under the supervision of Professor Ken Hyland at the University of Hong Kong. His research interests include disciplinary discourse, corpus studies and academic writing, and his papers have appeared in most applied linguistics journals, including *Applied Linguistics*, *Journal of English for Academic Purposes*, *English for Specific Purposes*, *International Journal of Corpus Linguistics* and *Discourse Studies*. He is the winner of the *John Sinclair Bursary* 2016 and the *Fawzia Braine Memorial Award* by Hong Kong Association of Applied Linguistics for young scholarly excellence in 2016.

Liang Maocheng

Beijing Foreign Studies University



Liang Maocheng, Ph.D., is a professor of Applied Linguistics at the National Research Center for Foreign Language Education, Beijing Foreign Studies University. Dr. Liang's background is in Corpus Linguistics, Second Language Acquisition (SLA) and Natural Language Processing (NLP). He is currently the chairperson of the Corpus Linguistics Society of China. Recently, his research has focused on computational approaches to second language texts, with an emphasis on Automated Essay Scoring (AES) and Automated Writing Evaluation (AWE). He is the systems architect of iWrite®, an online writing tutoring and automated assessment system. He has published two monographs in this area, namely, *Constructing a Model for the Computer-Assisted Scoring of Chinese EFL Learners' Argumentative Essays* (2011) and *Toward an Automated Essay Scoring System for Large-Scale English Language Tests* (2012). He is also the author of *What is Corpus Linguistics* (2016), a co-author of *Using Corpora: A Practical Coursebook* (2010), the author of a number of journal articles, and the systems architect and developer of several computer systems for corpus analysis.

Liu Zequan

Henan University



Liu Zequan is currently professor and Ph.D. advisor at Henan University, with the title of “The Yellow River Scholar”. He obtained his doctoral degree in Translation Studies from the National University of Singapore and his first master degree in English for Specific Purpose from Huazhong University of Science and Technology. He has keen research interests in corpus-based translation studies, discourse analysis, and academic writing. Up to now, he has chaired three national and provincial research projects respectively, including the on-going national key project “A Critical Survey of Women Translators in Taiwan, Hong Kong and Mainland China (2015-present)”. Publication-wise, he has got articles carried in a narrow range of prestigious language- and translation-related international and domestic journals, including *Meta*, *Journal of Chinese Linguistics*, *Chinese Translators Journal*, etc.

Michaela Mahlberg

University of Birmingham



Michaela Mahlberg is Professor of corpus linguistics at the University of Birmingham, UK, where she is also the Director of the Centre for Corpus Research and the Director of Research and Knowledge Transfer for the College of Arts and Law. Michaela is the editor of the *International Journal of Corpus Linguistics* (John Benjamins) and together with Wolfgang Teubert she edits the book series *Corpus and Discourse* (Bloomsbury). One of her main areas of research is Dickens's fiction and the socio-cultural context of the 19th century. Her publications include *Corpus Stylistics and Dickens's Fiction* (Routledge, 2013), *English General Nouns: a Corpus Theoretical Approach* (John Benjamins, 2005) and *Text, Discourse and Corpora. Theory and Analysis* (Continuum, 2007, co-authored with Michael Hoey, Michael Stubbs and Wolfgang Teubert). Michaela was the Principal Investigator on the AHRC-funded project *CLiC Dickens: Characterisation in the representation of speech and body language from a corpus linguistic perspective* which led to the development of the CLiC web app.

Wei Naixing

Beihang University



Wei Naixing is Professor of English and Director of the Research Institute of Linguistics Science and Engineering at Beihang University, China. He holds a MA in English degree of Birmingham University and a Ph.D in Linguistics degree of Shanghai Jiao Tong University. He is on the Editorial Board for the *Journal of Foreign Languages and Their Teaching*, the *Journal of Contemporary Foreign Language Studies*, the *Journal of Corpus Linguistics* and the *Journal of PLA Foreign Language University*. Wei Naixing has served as Professor of English at Shanghai Jiao Tong University, and Adjunct Professor at Huazhong University of Science and Technology, Chongqing University, Graduate University of Chinese Academy of Sciences, and Beijing University of Technology. He is a founding member of the Corpus Linguistics Society of China (CLSC) and served as its Chairman from April 2009 to April 2017. In 2010 he transferred to Beihang University in the capacity of Outstanding Professor in Humanities. Wei Naixing's major research interests cover corpus linguistics, phraseology, English for academic purposes, discourse analysis and contrastive studies of English and Chinese. He has published 6 monographs and 50 research papers.

会议日程

Conference Schedule

◆ 2018年6月1日(星期五 Friday)

时间 Time	内容 Item	地点 Venue
8:00-21:30	报到 Registration	华中科技大学学术交流中心 8 号楼 Building No. 8, Academic Exchange Center, HUST
9:00-11:30	工作坊 Workshop	科技楼北楼 602 Room 602, North Building of Keji Lou
14:00-16:30	工作坊 Workshop	科技楼北楼 602 Room 602, North Building of Keji Lou
18:00-20:00	晚餐 Dinner	接待中心一号楼瑜园餐厅 Yuyuan Dining Hall

◆ 2018年6月2日(星期六 Saturday)

时间 Time	内容 Item	主持人 Host/Hostess	地点 Venue
8:20-8:30	Opening Ceremony 开幕式 ● 外国语学院领导致辞(5 mins) Dean Welcoming Speech ● 中国语料库语言学专业委员会 致辞 (5 mins) Speech from CLSC	Lei Lei	科技楼南楼 104 Room 104, South Building of Keji Lou
8:30-9:20	主旨发言 1 Keynote Speech 1 ● 发言人: Michaela Mahlberg ● 题目: Pushing the Boundaries of Corpus Linguistics by Studying Fiction	Wei Naixing	科技楼南楼 104 Room 104, South Building of Keji Lou
9:20-10:10	主旨发言 2 Keynote Speech 2 ● 发言人: Liu Zequn ● 题目: A Corpus-based Comparison of Six Chinese Translations of the Japanese Novel <i>The Tale of Genji</i>	Michaela Mahlberg	科技楼南楼 104 Room 104, South Building of Keji Lou

The 4th Corpus Linguistics in China Conference

10:10-10:50		集体照 Group Photo 茶歇 Tea Break	
10:50-11:40	主旨发言 3 Keynote Speech 3 ● 发言人: Wei Naixing ● 题目: A Local Grammar of Apology: A Corpus-based Investigation	Liu Zequan	科技楼南楼 104 Room 104, South Building of Keji Lou
11:50-13:30	午餐 Lunch 接待中心一号楼瑜园餐厅 Yuyuan Dining Hall		
13:30-15:10	分组讨论 Concurrent Sessions (科技楼北楼 North Building of Keji Lou)		
15:10-15:40		茶歇 Tea Break	
15:40-17:45	分组讨论 Concurrent Sessions (科技楼北楼 North Building of Keji Lou)		
18:00-20:00	晚餐 Dinner 接待中心一号楼瑜园餐厅 Yuyuan Dining Hall		

◆ 2018年6月3日(星期日 Sunday)

时间 Time	内容 Item	主持人 Host/Hostess	地点 Venue
8:20-9:10	主旨发言 4 Keynote Speech 4 ● 发言人: Mark Davies ● 题目: iWeb: The 14 Billion Word, Learner-oriented Web Corpus	Liang Maocheng	科技楼南楼 104 Room 104, South Building of Keji Lou
9:10-10:00	主旨发言 5 Keynote Speech 5 ● 发言人: Jiang Feng ● 题目: Nouns and Academic Interactions: Structure and Function in Disciplinary Writing	Mark Davies	科技楼南楼 104 Room 104, South Building of Keji Lou
10:00-10:50	主旨发言 6 Keynote Speech 6 ● 发言人: Liang Maocheng ● 题目: Distributional Hypothesis Calibrated	Jiang Feng	科技楼南楼 104 Room 104, South Building of Keji Lou
10:50-11:10	Recapping the Conference 会议总结 Closing Speech 闭幕致词		科技楼南楼 104 Room 104, South Building of Keji Lou
11:30-13:30	Lunch 午餐 接待中心一号楼瑜园餐厅 Yuyuan Dining Hall		

分组讨论 (6月2日下午)

Concurrent Sessions (Afternoon, 2nd June)

Group 1

Room: Room 500, North Building of Keji Lou

Time: 13:30-17:20

Chair: Li Jingjie

Time	Presenters	Abstract No.	Titles
13:30-13:55	Ding Yingying	1-1	A Corpus-based Approach in Translation Training: A Case of Company Websites
13:55-14:20	Li Huayong	1-2	基于语料库的翻译汉语语义韵研究——以 cause 的汉译对应词“导致”、“引起”为个案
14:20-14:45	Zhang Dong & Xu Jiajin	1-3	英汉与格交替现象的多因素对比研究
14:45-15:10	Wang Xueqin & Fang Kai	1-4	A Comparative Corpus-based Study on Style of Translators of <i>The Classic of Mountains and Seas</i>
15:10-15:40 Tea Break			
15:40-16:05	Kang Jiaping	1-5	A Corpus-based Study on Linguistic Features in the English Translation of <i>Report on the Work of the Government</i>
16:05-16:30	Liu Miao	1-6	A Corpus - based Study on the Language Features of Russian - Chinese Translation Novels — Taking Deputy Verbs as the Example
16:30-16:55	He Qiong	1-7	Using Corpora to Study L2 Acquisition of English Tense of Chinese non-English-major university students —A Study Based on Chinese Learner English Corpus
16:55-17:20	Li Jingjie, Hou Huili	1-8	硬科学论文中 we 的学科分布及话语功能研究

Group 2

Room: Room 505, North Building of Keji Lou

Time: 13:30-17:20

Chair: Liu Dilin

Time	Presenters	Abstract No.	Titles
13:30-13:55	Liu Dilin	2-1	Teaching Appropriate Use of Synonyms: Corpus-based Discovery Learning and Reinforcement Practices
13:55-14:20	Hsien-Chin Liou	2-2	Synergy of Corpus Consultation, Corrective Feedback, and Motivation in a Year-Long College Freshman Writing Study
14:20-14:45	Nanthini Arivuselvan & Seetha Lakshmi	2-3	Discourse Analysis in Tamil Classrooms Using Corpus Data: A Study on the Student Agency Affecting the Level of Classroom Interaction
14:45-15:10	Seetha Lakshmi & Letchmi Devi d/o Ponnusamy & Nanthini Arivuselvan	2-4	Tamil Corpus Bank and Students' Spoken Language —An overview
15:10-15:40 Tea Break			
15:40-16:05	Chen Lidan & He Anping	2-5	Corpus Keyword-aided Business English Instruction with Dual Aim
16:05-16:30	Li Yanjiao & Alex C. Fang	2-6	A Corpus-based Study of Feedbacks in Daily Conversation for Pedagogical Applications
16:30-16:55	Long Shaoyun & Liu Dilin	2-7	Grade Features of Epistemic Modality's Keyness in English Major's Argumentative Essays
16:55-17:20	Yuan Fei & Sun Youzhi & Tian Zuoyu	2-8	MOOC Instructors' Linguistic Features — A Corpus-Based Research

Group 3

Room: Room 509, North Building of Keji Lou

Time: 13:30-17:20

Chair: Zhang Xuhua

Time	Presenters	Abstract No.	Titles
13:30-13:55	Lei Lei & Liu Dilin	3-1	Most Common Academic English Collocations A Corpus-driven Study
13:55-14:20	Zhang Xuhua & Zhang Daqun	3-2	线性单位语法框架下的学术英语口语词块研究
14:20-14:45	Dong Min & Qiu Xixin	3-3	Cohesive Verbs: Patterns and Collocations in Research Articles
14:45-15:10	Li Xiaohong	3-4	A Contrastive Analysis of Discourse Strategies in Ph. D Dissertation Abstracts in Linguistics
15:10-15:40 Tea Break			
15:40-16:05	He Mengyu	3-5	Comparing Engagement Markers in Economics Research Articles and Popular Science Writing: A Corpus-based Study
16:05-16:30	ANG Leng Hong	3-6	Exploring Phraseological Variation in Academic Writing: A Corpus Analysis
16:30-16:55	Yang Yuesen & Wei Naixing	3-7	Patterns and Meanings of Verbal Hedges in English Research Articles by Chinese and Western Scholars: A Corpus-based Contrastive Study
16:55-17:20	Guan Xinchao	3-8	专业语篇 N-grams 提取及其 MTI 教学和论文写作应用

Group 4

Room: Room 510, North Building of Keji Lou

Time: 13:30-17:30

Chair: Chen Jiansheng

Time	Presenters	Abstract No.	Titles
13:30-13:55	Li Bin & Song Li & Kuang Haibo & Dai Rubing	4-1	英汉双语认知属性库的构建及计量分析
13:55-14:20	Chen Jiansheng	4-2	论体育英语参照语料库建设的三维模式
14:20-14:45	Bai Yu	4-3	“一带一路”核心语汇多模态汉俄平行语料库建设（2013—2021）
14:45-15:10	Li Xiaqing	4-4	基于知识图谱理论的国家“语言战略”研究热点可视化分析
15:10-15:40 Tea Break			
15:40-16:05	Xiao Cheng	4-5	武汉地区孤独症儿童语料库研究——以名词为例
16:05-16:30	Tian Xiaoyu & Hu Meifen & Zhang Weiwei	4-6	形式动词“加以”和“予以”近义辨析：基于语料库的构式搭配分析
16:30-17:05	Zhang Baolin	4-7	“HSK 动态作文语料库”（V2.0）的设计理念与功能
17:05-17:30	Hou Zhide	4-8	基于语料库的西方媒体之“深圳形象”研究

Group 5

Room: Room 513, North Building of Keji Lou

Time: 13:30-17:20

Chair: Xing Fukun

Time	Presenters	Abstract No.	Titles
13:30-13:55	Lu Jun	5-1	Integrating Corpus-based Analysis with Experimental Data: A Corpus-based Experimental Study of Implicit and Explicit L2 Collocational Knowledge
13:55-14:20	Xu Jianjun	5-2	An Empirical Study on the Cognitive Level of Two-syllable Lexicalization of Chinese Dimensionative Adjectives
14:20-14:45	Zhang Le	5-3	Phraseological Features of Application Letters Written by Chinese College Students
14:45-15:10	Hong Yang & Pan Hong	5-4	Comparative Analysis of the Polysemy of Preposition "in" by Native English Speakers and Chinese English Learners — A Corpus-based Study
15:10-15:40 Tea Break			
15:40-16:05	Liu Yunfeng	5-5	Meaning Shift Units in the Perspective of Local Grammars: A Study of <i>UPON</i> in Business English Correspondence Texts
16:05-16:30	Jin Yingying	5-6	A Corpus-Based Analysis of Synonyms —Take <i>absolutely</i> and <i>thoroughly</i> for example
16:30-16:55	Lei Shujie & Xing Fukun	5-7	Research on the Structure of the English Names of Military Equipment for NLP
16:55-17:20	Deng Hailong	5-8	基于词向量模型的中国学习者英语词汇语义特征研究

Group 6

Room: Room 600, North Building of Keji Lou

Time: 13:30-17:20

Chair: Jiang Zhanhao

Time	Presenters	Abstract No.	Titles
13:30-13:55	Jiang Zhanhao	6-1	Pragmatic Annotations in Corpus Construction
13:55-14:20	Zhao Yonggang	6-2	并购话语中的中国企业形象研究：基于语料库的话语—历史观方法
14:20-14:45	Zhao Chaoyong & Hu Jing	6-3	基于语料库的中-美政府工作报告语域特征多维分析
14:45-15:10	Zou Hong & He Huaqing	6-4	A Corpus-based Study on Contrastive Discourse Markers in Chinese English Writing
15:10-15:40 Tea Break			
15:40-16:05	Xiao Xianming	6-5	Genre Theory and Corpus-driving Introduction Writing Models in Scientific Academic Articles
16:05-16:30	Su Hang	6-6	Changing Patterns of Apology in British English: A Local Grammar Based Diachronic Investigation
16:30-16:55	Yang Zitong & Liu Aijun	6-7	基于语料库的美国“一带一路”新闻报道的批评话语分析
16:55-17:20	Wang Zhong & Fang Chengyu Alex & Fan Weiwei	6-8	The Influence of Topic Location on Information Distribution in <i>it</i> -extraposition and Non-extraposition: An Investigation Based on ICE-GB

Group 7

Room: Room 602, North Building of Keji Lou

Time: 13:30-17:20

Chair: Hu Chunyu

Time	Presenters	Abstract No.	Titles
13:30-13:55	Deng Sirui & Zhou Jie	7-1	基于语料库的“一带一路”英译在东盟国家的认知研究
13:55-14:20	Hu Chunyu & Lu Chunmei	7-2	Organic and Mechanical Metaphors in <i>the Wealth of Nations</i>
14:20-14:45	Fang Yinjie	7-3	构式的操作界定:认知固化与条件概率
14:45-15:10	Yang Linlin	7-4	认知视域下记者招待会中隐喻口译策略的实证研究
15:10-15:40 Tea Break			
15:40-16:05	Hu Chunyu & Wang Jinxiao	7-5	基于营销学语料库的旅途隐喻研究
16:05-16:30	Zhang Xuemei & Liu Ping	7-6	英语学习词典中外壳名词的释义模式初探——以《牛津高阶英语词典》(第8版)为例
16:30 -16:55	Zhao Jiamei & Chen Jiansheng	7-7	A Study of Phrase Frames in Argumentative Essays by Chinese English-majors
16:55-17:20	Chen Jiansheng & Huang Hui	7-8	A Corpus-based Study of Explication of Personal Pronouns in Chinese-English Consecutive Interpreting at Press Conference

Group 8

Room: Room 605, North Building of Keji Lou

Time: 13:30-17:20

Chair: Shao Bin

Time	Presenters	Abstract No.	Titles
13:30-13:55	Kang Ning	8-1	基于语料库的译者风格研究：以《骆驼祥子》葛浩文译本与施晓菁译本为例
13:55-14:20	Chen Cheng	8-2	A Semiotic Interpretation of Indian Media's Emotions to China —A Corpus-based Study
14:20-14:45	Zhang Jialing	8-3	基于语料库研究方法分析戏剧中感叹句的使用情况——以俄罗斯喜剧《聪明误》为例
14:45-15:10	Hou Yu	8-4	基于语料库的《红楼梦》人称指示视点翻译转移比较研究
15:10-15:40 Tea Break			
15:40-16:05	Shao Bin & Cai Yingying & Chen Lingwei	8-5	A Corpus-based Study on the Representation of "The Belt and Road Initiative" in the Western Media
16:05-16:30	Stephen Jeaco	8-6	Exploring Language and Literature with <i>The Prime Machine</i>
16:30-16:55	Xuan Jin	8-7	当代中国大众读者对精英文学《浮士德》的接受研究——以 2005-2018 年《浮士德》豆瓣评论为例
16:55-17:20	Zhang Ruidan	8-8	基于语料库的模糊限制语使用的性别差异研究

Group 9

Room: Room 606, North Building of Keji Lou

Time: 13:30-17:20

Chair: Deng Yaochen

Time	Presenters	Abstract No.	Titles
13:30-13:55	Deng Yaochen	9-1	A Multifactorial Analysis of the Choice of Necessity Modals <i>Must</i> and <i>Have to</i> in Chinese Learner English
13:55-14:20	Zhang Zihan & Li Ye & Huang Yi	9-2	A Corpus-based Study of Epistemic Modality in Chinese University Students' English Writing
14:20-14:45	Shi Xinyu	9-3	翻译与中国自我形象建构——基于语料库的《毛泽东选集》英译研究
14:45-15:10	Wu Liangping	9-4	多因素分析的标注瓶颈与分离式标注
15: 10-15:40 Tea Break			
15:40-16:05	Zou Yucheng & Hu Suhua	9-5	基于语料库的彝语诺苏话话题标记研究
16:05-16:30	Lei Xiaodan	9-6	A Corpus-based Comparative Study on Cohesive Devices in Geological English Abstracts of Chinese and International Journals
16:30-16:55	Sun Wan	9-7	基于虚词词典与语料库的“的”字短语释义研究
16:55-17:20	Qian Zhen	9-8	语料库语言学在古汉语研究中的应用——以“最”、“至”、“极”的语体正式度差异为例

Group 10

Room: Room 607, North Building of Keji Lou

Time: 13:30-17:45

Chair: Zhang Liying

Time	Presenters	Abstract No.	Titles
13:30-13:55	Yang Linwei	10-1	基于 RNN 的主谓一致错误自动识别研究
13:55-14:20	Pan Hong	10-2	对构式“什么 A (什么) B”认知解读
14:20-14:45	Liu Yang	10-3	中国学习者英语语法形位习得顺序——来自语料库和实验的证据
14:45-15:10	Yang Chenjiao	10-4	中西学者学术论文中作者声音构建的对比研究：以“名词+that”结构为例
15:10-15:40 Tea Break			
15:40-16:05	Yuan Xinhua	10-5	局部语法研究源流与发展述评
16:05-16:30	Yao Xinyue	10-6	Developments in Australian English Grammar from 1931 to 2006: An Aggregate Approach to Dialectal Variation and Change
16:30-16:55	Zhang Mengmeng	10-7	对英语迂回式使役结构中使役动词的多维度研究
16:55-17:20	Zhang Liying	10-8	Properties of Verb and the Present Perfect Construction
17:20-17:45	Chen Fu & Zhang Xuemei	10-9	基于义项标注法的构式搭配分析新路径：以英语被动构式为例

Group 11

Room: Room 609, North Building of Keji Lou

Time: 13:30-17:45

Chair: Gao Xia

Time	Presenters	Abstract No.	Titles
13:30-13:55	Gao Xia	11-1	'This means...': A Contrastive Analysis of Authorial Practice of Attended/Unattended <i>this</i> in Physics and Linguistics Research Articles
13:55-14:20	Li Jinfeng & Xiao Wei	11-2	A Comparative Analysis on Stance Markers in Research Article Abstracts of Theoretical Linguistics and Quantitative Linguistics: A Corpus-Based Study
14:20-14:45	Yu Xuan & Xiao Wei	11-3	A Comparative Analysis on Reporting Verbs in Research Article Introductions Across Disciplines
14:45-15:10	Malila Carvalho de Almeida Prado	11-4	Contrasting Corpus Analysis Data to a Proficiency Scale: The Case of Aviation English
15:10-15:40 Tea Break			
15:40-16:05	Tao Guoxiao & Pan Fan & Fang Alex Chengyu	11-5	A Developmental Study of Prepositional Phrases in English L1 Students' Academic Writing
16:05-16:30	Zhang Jinfu	11-6	中国高校英语专业本科毕业论文结语研究
16:30-16:55	Yan Pengfei	11-7	Research on Modal Sequences in L2 Academic Writing
16:55-17:20	Zhao Lishu	11-8	A Corpus Driven Analysis of Representations Around the Term 'THAAD' in the Chinese, South Korean and American Press 2017
17:20-17:45	Cao Jing & Fang Alex C.	11-9	Identifying the Linguistic Features of Research Paper Titles in Biomedicine: A Corpus-Based Study

Group 12

Room: Room 610, North Building of Keji Lou

Time: 13:30-17:20

Chair: Zhao Xiaolin

Time	Presenters	Abstract No.	Titles
13:30-13:55	Zhao Xiaolin	12-1	Construction of Authorial Identity in HK MA Students' Academic Writing: Phraseological Sequences of SUGGEST as a Case
13:55-14:20	Liu Fan	12-2	A Corpus-based Study of Defamation Law in Mainland China, Hong Kong and UK: A Gricean Perspective
14:20-14:45	Dong Jihua & Chigan Menghua	12-3	Discipline-Specific Stance Construing in Academic Discourse
14:45-15:10	Wang Mimi	12-4	基于语料库的中外学者学术英语词块使用对比研究
15:10-15:40 Tea Break			
15:40-16:05	Wang Zaiya	12-5	中外学者学术写作中立场形容词对比研究
16:05-16:30	Wang Minli & Li Lixia	12-6	农业学术语篇中主语 it 外置结构的形式与功能——基于动物科学及园艺科学语料库的研究
16:30-16:55	Fu Ying	12-7	扩展意义单位下中外学者学术写作的对比研究——以最高程度增强语为例
16:55-17:20	Ma Haoran & Li Jinman	12-8	The Newer, the better? —Testing four wordlists with two corpora

Group 13

Room: Room 612, North Building of Keji Lou

Time: 13:30-17:20

Chair: Ma Xiaolei

Time	Presenters	Abstract No.	Titles
13:30-13:55	Ma Xiaolei & Chen Yingfang	13-1	英语学术语篇中外壳名词的动词搭配研究
13:55-14:20	Peng Lei	13-2	A Corpus-based Genre Study on the Essential Moves of Household Manuals
14:20-14:45	Li Longxing & Huang Juren	13-3	基于词汇素描系统的“reform”与“改革”对比研究
14:45-15:10	Sun Haiyan	13-4	基于语料库的动词配价型式特点对比研究
15:10-15:40 Tea Break			
15:40-16:05	Song Yajun	13-5	The Analysis of Structural and Functional Features of Lexical Bundles in English Major Academic Dissertations
16:05-16:30	Gao Lin	13-6	基于平行语料库的汉英心理形容词语义对比研究——以《萧萧》原本和英译本为例
16:30-16:55	Li Yuantian	13-7	基于语料库的视觉行为动词词义基元结构研究
16:55-17:20	Miao Lixia	13-8	中国学习者英语多义词语义知识的使用和发展

Group 14

Room: Room 613, North Building of Keji Lou

Time: 13:30-17:45

Chair: Li Wenzhong

Time	Presenters	Abstract No.	Titles
13:30-13:55	Li Wenzhong	14-1	信任文本与局部语法——以“N1 +死了+ N2”为例
13:55-14:20	Xiao Jiayan	14-2	Beyond Translation: HONGLOUMENG and China Image
14:20-14:45	Zheng Jianning	14-3	汉学家中国典籍英译语料库建设：理论与实践的思考
14:45-15:10	Li Tao	14-4	Reshaping China's Image: A Corpus-based Discourse Analysis of the English Translation of Chinese Political Discourse
15:10-15:40 Tea Break			
15:40-16:05	Wen Mengyuan	14-5	The English Translation of Traffic Aids Based on Parallel Corpus
16:05-16:30	Zheng Hongbo	14-6	基于多译本语料库的儿童文学翻译译文对比研究——对《卖火柴的小女孩》五译本的对比分析
16:30-16:55	Yang Liu	14-7	On Narrative Progression of English Fan Translation of Chinese Web-Novel <i>Coiling Dragon</i> —with Corpus Narratology Method
16:55-17:20	Huang Baorong	14-8	美国计算机协会通讯中译语料库建设及翻译研究
17:20-17:45	Zhou Zhiyi	14-9	Corpus-based Study of Semantic Co-selection among Words in English for Logistical Purpose

论文摘要 Abstracts

主旨发言摘要 Abstracts of the Plenary Sessions

iWeb: The 14 Billion Word, Learner-oriented Web Corpus

Mark Davies

Brigham Young University

The iWeb corpus was released in May 2018. It has many features that make it unique, even compared to the other BYU corpora such as COCA, COHA, NOW, GloWbE, and BYU-BNC.

At 14 billion words, iWeb is more than 25 times as large as the 560 million word COCA corpus. iWeb also has a much wider range of web-based materials than does COCA, since it is based on 22 million web pages in nearly 100,000 carefully selected websites.

Unlike any of the other BYU corpora, iWeb allows you to browse through the top 60,000 words in the corpus, and to search this list by word form, part of speech, rank (#1-60,000), and even pronunciation.

Most importantly, you can then see detailed information on each of the top 60,000 words in the corpus – definition, frequency information, synonyms and other related words (from WordNet), collocates (in a much improved format), related “topics” (perhaps much more useful than collocates), “clusters” (new in iWeb), relevant websites, and sample concordance/KWIC lines.

In addition, for each of these 60,000 words, there are “quick links” to related data from other websites – pronunciation, additional definitions, images, videos, and translations (for more than 100 languages).

Finally, in terms of “standard” corpus searches, iWeb is faster than any of the other BYU corpora, and much, much faster than other large, 10-20 billion word online corpora. iWeb also allows you to quickly and easily create “virtual corpora” on nearly any topic, and these virtual corpora can then be searched as their own “stand-alone” corpora, or compared to other virtual corpora that you have created.

In this presentation, I will present many searches that show how the corpus can be used to find a wealth of information on English, which is not available from any other corpus – and especially how iWeb can be used for teaching and learning English.

Nouns and Academic Interactions: Structure and Function in Disciplinary Writing

Feng (Kevin) Jiang

Jilin University

The rhetorical role nouns play in shaping persuasive academic discourse has received relatively less attention than other linguistic resources such as adjectives and verbs. In this presentation, I look at a particular type of nouns which I call *metadiscursive nouns* (as *advantage in the reflector has the advantage that it is resistant to changes in temperature*), and by drawing on Hyland's (2005) model of metadiscourse, I see metadiscursive nouns as functioning in both interactive and interactional dimensions. Here I examine "determiner + noun" and "noun + post-nominal clause" structure in a 1.7-million word corpus of 160 research articles from soft and hard knowledge fields, exploring how the nouns are used to mediate writer-reader interactions in academic writing.

Results show that the soft disciplines use metadiscursive nouns far more frequently than hard domains, which reflects the discursive nature of academic arguments and the need to reach a less homogeneous audience. Additionally, writers also vary in their use of metadiscursive nouns when developing arguments across different divisions of research articles, making the most of metadiscursive nouns in the literature review and result sections. Thus, we can see the choice of metadiscursive nouns does not only structure discourse or specify aspects of the topic, but also reflects the knowledge-making practices and discursive norms in the disciplines.

Distributional Hypothesis Calibrated

Liang Maocheng

Beijing Foreign Studies University

Distributional Hypothesis, originally proposed by Zellig Harris (1954), claims that words that occur in the same contexts tend to have similar meanings, and that words which have similar meanings tend to occur in the same contexts. On the operational level, while the distributional theory of semantics inevitably involves sophisticated computation, it is based on the simple fact that semantically related words tend to co-occur in order to make up logical and sensible texts. Over the years, the implementation of the hypothesis has brought about the success of a number of applications, most notably the well recognized theory of Latent Semantic Analysis and the relatively new Word2vec technology. In almost all such applications, one important factor has gone largely unnoticed – the operational definition of context, in which semantically related words are supposed to co-occur.

In this paper, we argue that to maintain the progression of the topic, semantically related words do not co-occur evenly over different sections of a text. More often, they are densely populated in only one part of the text. With the progression of the topic, a different set of semantically related words will become common, with the result that the topic of the text gradually changes to something slightly different. It is therefore proposed that in the implementation of the Distributional Hypothesis, the strength of word co-occurrence be measured within spans of variable lengths. Experiments with the proposed implementation have consistently produced better word embedding.

A Corpus-based Comparison of Six Chinese Translations of the Japanese Novel *The Tale of Genji*

Liu Zequan

Henan University

The Japanese novel *The Tale of Genji* (or *Genji Monogatari*) written at the beginning of the eleventh century, has been likened to *Hong Lou Meng* (or *Dream of the Red Chamber*) of the Tang Dynasty (618-907) of China, and, owing to its supreme achievements in the monogatari genre, has been deemed as the world's first modern novel ever coming out of the hand of a literatus. *The tale*, which comes in 54 chapters in “concertina” or “orihon” with nearly one million words in *kana* (the archaic Japanese phonetic script) and *tanka* (classic Japanese poetry), is a unique depiction of the lifestyles of high courtiers during the Heian Period, focusing on the love affairs of the hero Hikaru Genji, or “Shining Genji”, the demoted son of an emperor.

Ever since Taiwanese female translator Lin Wenye (1978) and Mainland China translator Feng Zikai (1980) published their respective complete Chinese translations of *The Tale*, nearly a dozen new versions or “retranslations” have sprung up, including Yin (1996), Xia (2002), Liang (2003), Yao (2006), Zheng (2006), Tang (2008), Wang (2010), Qiao (2012) and Ye (2014). Understandably, the craze for retranslation gives rise to an enthusiasm for studies of the translations, a survey of which seems to boil down to two conspicuous conclusions. While most studies devote themselves to a comparison of and conclude with compliments to Lin's and Feng's versions, others employ an elusive approach to the appraisal of sporadic excerpts of translations. While these observations point to a lack of systematic investigation of more translations on an objective or quantitative basis, some studies do come up with claims or even open outcries over the poor quality of the later translations, and even the “over-reference” to or “plagiarizing” of their predecessors, especially Yin (1996), Liang (2003), and Yao (2006).

Against this backdrop, the current study sets out to make a corpus-based comparison of six Chinese translations of *The Tale*, i.e., those of Feng, Lin, Yao, Ye, Yin and Zheng. The parameters to be examined are threefold: lexical, including TTR, density, as well as the occurrences of some remarkable types of lexis (e.g., reporting verbs, conjunctions, modal particles, and “love-hatred” expressions); average sentence length; and techniques whereby by the *tanka* verses are dealt with.

The study intends to answer two major questions, i.e., what linguistic and textual features (or translator style) the six translations register, and, to what extent these translations are similar to each other.

Pushing the boundaries of corpus linguistics by studying fiction

Michaela Mahlberg

University of Birmingham

In corpus linguistics, literary texts have often been neglected. Examples from literary texts may be seen as exceptions or outliers, or as Sinclair (2007: 3) puts it “a distinctive literary text is just not worth including in a general corpus because it will disappear below the waves”. Alternatively, fiction is treated as a register that can be compared against other registers and hence literary qualities of individual texts are levelled out in the resulting linguistic description. The umbrella term ‘corpus stylistics’ (cf. e.g. Semino and Short 2004) is increasingly used to describe approaches that employ corpus linguistic method to study specific literary qualities and individual literary texts. However, in order to be able to account as fully as possible for features of literary texts we need to create new tools and develop methodologies that are tailored to the task at hand. In this paper, I will illustrate key functionalities of the web application CLiC (<http://clic.bham.ac.uk/>) that has been specifically designed for the corpus linguistic study of narrative fiction. The case study that I will present looks at textual patterns that contribute to the creation of fictional characters. The examples will be drawn from the CLiC corpora. The most recent version of the CLiC corpora now comprise more than 130 books across four subcorpora: the corpus of Dickens’s Novels, the 19th Century Reference Corpus (19C), the Corpus of 19th Century Children’s Literature (ChiLit) and the Corpus of Additional Requested Texts (ArTs). For all CLiC texts, direct speech and specific places around speech have been marked up (Mahlberg et al. 2016) so that CLiC can run searches across defined textual subsets and thus support the analysis of literary features of texts. This approach illustrates how corpus methods are not only applied to the study of literary texts, but also need to be developed in response to research questions that arise from literary concerns.

A Local Grammar of Apology: A Corpus -based Investigation

Wei Naixing, Su Hang

Beihang University

Corpus linguistics has witnessed the emerging of a series of local grammar research in the past decade (Gross 1993, Hunston & Sinclair 2000, Barnbrook 2002, Allen 2005, Bednarek 2008, Cheng & Ching 2016, etc.). Local grammar is a new approach to the description and theorising of language. Previous local grammar studies indicate that they can offer more specific, more transparent, and more precise information about language in actual use, and they signify a new research orientation in corpus linguistics. This paper extends the concept of local grammar to speech act studies, focusing specifically on apologising. Key terms (e.g. *sorry*, *apologize*) that are frequently involved in apology expressions are used to automatically retrieve speech act instances in a corpus of scripted TV conversations. The retrieved instances are further manually examined in order to make sure that all remaining instances have illocutionary forces and thus qualify for further analysis. These instances are then analysed from a local grammar perspective, which facilitates the establishment of a local grammar of apology, comprising 14 local grammar patterns. The analyses show that it is promising to develop a set of local grammars to account more adequately for speech acts in general. The relationship between local grammars, functional grammars, and general grammars are further discussed. It is argued that local grammars can contribute to a more-functional account of language in use and that local grammars are of general applicability in linguistic description and explanation. Finally, directions for future research are suggested, and implications are briefly discussed.

分组讨论摘要 Abstracts of Concurrent Sessions

Group 1

1-1

A Corpus-based Approach in Translation Training: A Case of Company Websites

Ding Yingying, University of Bologna

In this paper, we propose to investigate problems of translated company websites from Italian into Chinese and provide some strategies in order to be integrated in translation training, as well as to improve the quality of translation.

Using a corpus-based approach, we will explore how a comparable corpus can be applied in Italian-Chinese translation to assist translators to make their decisions and discuss some findings with respect of their didactic utility for translation. The corpus contains 50 Chinese company websites in the field of food, and another 50 company websites which are translated from Italian into Chinese. The comparison between the texts suggests that the translated texts contain a higher level of frequency in the use of passive voice, as well as structural features of the Italian language. The detailed results of this study are expected to shed new lights on translation teaching and cultural implications for translation from Italian into Chinese.

1-2

基于语料库的翻译汉语语义韵研究——以 cause 的汉译对应词“导致”、“引起”为个案

李华勇，四川文理学院

翻译汉语作为一种客观存在的汉语变体，除词汇、句法层面存在不同于原创汉语的差异外，词项构筑的语义韵与其在原创汉语中构筑的语义韵也存在显著差异。研究显示，cause的汉译对应词“导致”和“引起”在翻译汉语语料库TED-cn中构筑的语义韵与其在原创汉语语料库BJKY中构筑的语义韵存在显著性差异。造成这种差异的原因一方面是英语源语渗透效应、共选机制在译语中的词汇搭配偏差产生的影响，另一方面是译者在翻译时积极调和的结果。这一研究结果说明：语义韵具有语类依赖性，同一词项在不同语类构筑的语义韵具有差异；在构建语义韵理论时，除要考虑语言类型、语境等因素外，还要考虑同一语言中的语言变体这一因素。

1-3

英汉与格交替现象的多因素对比研究

张 懂，许家金，北京外国语大学

与格交替 (dative alternation) 指的是介词与格构式 (如 John sent the book to Mary) 和双宾构式 (如 John sent Mary the book) 的交替使用。此现象在英汉语中皆有体现。本文运用多因素分析法 (multifactorial analysis)，对英汉与格交替制约因素之异同加以对比。研究发现：英语与格交替的制约因素按程度大小依次是：受事可及性 > 成分重量 >

与事代词性 > 与事可及性 > 与事具体性 > 受事有定性；汉语与格交替的制约因素排序依次是：受事有定性 > 受事具体性 > 与事有生性。英汉与格交替制约因素的共性和差异表现为：受事为有定成分时，英汉均显著倾向使用介词与格构式；二者在可及性和成分重量方面表现出显著差异。上述异同可从语言类型差异和认知机制得到解释。

1-4

A Comparative Corpus-based Study on Style of Translators of *The Classic of Mountains and Seas*

Wang Xueqin, Fang Kai, Wuhan University of Technology

Based on the English versions of *The Classic of Mountains and Seas*, this paper investigates the style of translators from lexical level, grammatical level and text level with the aid of corpus. The findings are as following: (1) the vocabulary of Birrell's version is more extensive and the text information is larger, but the average sentence length is shorter than Wang Hong's version, which indicates that the sentence structure is more simple. (2) Wang Hong's version adopts more subordinate clauses with longer sentences and more complicated structure. (3) Conjunction words and personal pronouns also show varying degrees of difference. (4) Besides, the two translator adopts different strategies to translate the proper nouns and culture-loaded words. As the cultural backgrounds and translation purposes of the two translators are quite different, the English versions of *The Classic of Mountains and Seas* presents different styles.

1-5

A Corpus-based Study on Linguistic Features in the English Translation of *Report on the Work of the Government*

Kang Jiaping, Xi'an International Studies University

Based on the corpus techniques, with an integration of corpus and SPSS methods, this paper conducts a contrastive analysis of the linguistic features of the two corpora: English translation of *Report on the Work of the Government and State of the Union Address* from 2009 to 2018 respectively. The result shows that there are differences in linguistic features between the two political discourse texts which are similar in function. This paper analyzes the differences specifically from three aspects: lexical features, syntactic features and textual features. The reasons of differences are discussed and some suggestions are given on the translation of *Report on the Work of the Government*.

1-6

A Corpus - based Study on the Language Features of Russian - Chinese Translation Novels —Taking deputy verbs as the example

Liu Miao, Beijing International Studies University

The deputy verb is a typical linguistic phenomenon in Russian literary texts and a "language vacuum item" in the process of mutual translation between Russian and Chinese.

Taking the source text of Chekhov's short stories and its three Chinese versions as data, based on corpus-driven approach, dissertation systematically describes the standing linguistic features of the Russian-Chinese literary translations from deputy verb. The research shows that the deputy verbs are more densely distributed in the literary texts and the translation tends to be integrated in the sentence structure. The translation lacks the grammatical means of lexicon and grammar, and the semantic relationship between the main verb and the deputy verb is obscured in the translation.

1-7

Using Corpora to Study L2 Acquisition of English Tense of Chinese non-English-major University Students —A Study Based on Chinese Learner English Corpus

He Qiong, Wuhan University

Numerous studies on L2 acquisition of English tense have been conducted home and abroad over the past 30 years, however, since corpus-based study on the acquisition of tense of Chinese college English students is still lacking, the hypotheses drawn upon previous studies haven't been testified with large quantities of written texts of Chinese college English students. The present paper seeks to testify the hypotheses on the basis of Chinese Learner English Corpus (CLEC), and then sketch out the implication of the study for College English Education.

1-8

硬科学论文中 *we* 的学科分布及话语功能研究

李晶洁, 侯绘丽, 东华大学

人称介入是一种人际策略, 服务于学术文本的交际目的。在所有人称介入表达中, *we* 的使用最为频繁, 语义指向也最复杂。本文基于语料库数据, 概括与描述硬科学论文中 *we* 的频数分布、语义指向及典型话语行为。研究发现, 近年来硬科学作者越来越多地使用 *we*, 或突显个人身份, 或与读者互动。一些传统的自然科学, 如数学和物理, 对 *we* 的使用频率甚至超过了部分软科学。硬科学论文中的 *we* 通常指向 4 类群体: 作者及其研究团队、作者与读者、整个学科和泛指一般人。每类语义指向下的 *we* 可搭配不同动词, 构成丰富多样的共选型式, 以实施多重学术话语行为。本研究可为新手作者提供便捷资源, 提高其语言使用的准确和适切, 因此具有学术英语教学价值。

Group 2

2-1

Teaching Appropriate Use of Synonyms: Corpus-based Discovery Learning and Reinforcement Practices

Liu Dilin, Alabama University

Synonyms or near-synonyms (as all synonyms are truly near-synonyms) express essentially the same meaning, but they often carry different attitudes, connotations, and/or connotations. Therefore, synonyms are generally not completely interchangeable. As such, they are vital for precise and effective communication, but very difficult for language users, especially L2 learners. An added difficulty for L2 learners in dealing with synonyms is the fact that words that are synonymous in their L1 may not be so in the target language or vice versa. This presentation discusses how the use of corpora may help us learn/teach synonyms more effectively. Drawing on recent research including his own published studies (2010, 2012, 2013, 2018), the speaker will first show that synonyms are best distinguished by their typical collocates (i.e., the words they usually co-occur) and that learning the appropriate use of synonyms is largely the learning of their collocates. Then, the speaker will present corpus-driven discovery learning activities for learning/teaching three sets of difficult synonyms (*circumstance/case/situation*, *demand/require*, and *important/significant*). These three sets of synonyms are especially challenging for Chinese learners of English because the words in each set, while sharing some semantic aspects, differ substantially in some other aspects that are, however, are not highly distinctive in the Chinese language as evidenced by the fact that the synonyms in each set are frequently translated into one same Chinese word (情况, 要求, 重要 respectively). As a result, Chinese EFL learners have often been found to misuse these synonyms, sometimes resulting in serious communication problems (author, 2018).

After presenting the corpus-based discovery learning activities, the speaker will introduce some corpus-based practice exercises that are designed and used to reinforce the learning of these difficult-to-distinguish lexical items. Principles and specific procedures for developing and using these corpus-based activities and exercises will be described and explained during the presentation.

2-2

Synergy of Corpus Consultation, Corrective Feedback, and Motivation in a Year-long College Freshman Writing Study

Hsien-Chin Liou, Feng Chia University

Research on L2 learners' use of online concordance programs has rarely been addressed to scaffold students' correction and uptake of written corrective feedback (WCF) for revision by taking into account important considerations of their processing of the feedback. As web-based concordancers become readily available with Internet connection, corpus consultation has confirmed its complementary role to help L2 writing process. This study attempted to achieve the synergies among insights of DDL, WCF, and motivation by

recruiting fifteen EFL college freshman students in an Asian country from a required yearly writing course. They wrote six multi-draft essay assignments plus diagnostic essays as part of their regular course work in various genre types: narration, description, comparison/contrast or exemplification. They used two Chinese-English concordancers and COCA (all new tools to the learners) to help their revisions after several prominent errors were marked on essay-drafts using an agreed-upon list of error codes. Multiple data were analyzed and triangulated: background/evaluation questionnaires, a writing motivation questionnaire, retrospective interviews, students' marked and revised essays, and video recording of learner-corpus interactions. Frequency and linguistic features of corpus consultation, error types and reduction rates were documented across six assignments over the year. Incidents of language awareness, correct revisions, and correct use in the new pieces of writing were also identified as evidence of linguistic development of these participants' writing accuracy in the year. In spite of the fluctuating trend, generally the students made progress as their writing skills became mature with the help of the in-class instruction and corpus-assisted error correction. Although most participants appreciated the assistance the new e-reference tools offered to their revision process, negative effects and no-effects were also noted in the process. Two focal cases indicate close links among motivation, task engagement, and feedback processing. Research implications on linking motivation to WCF and DDL are addressed.

2-3

Discourse Analysis in Tamil Classrooms Using Corpus Data: A Study on the Student Agency Affecting the Level of Classroom Interaction

Nanthini Arivuselvan, Seetha Lakshmi, Nanyang Technological University

This paper aims to discuss how student agency affects the level of classroom interaction through discourse analysis in Tamil classrooms. This study is part of a larger research that aims to analyse the language acquisition in Singapore Tamil classrooms. Before learning can take place, it is important for the students to feel that they are the key agents in learning. They need to have a sense of comradeship in classroom where they feel a sense of connectedness and feel empowered to develop their capacity in learning. This is only possible when teachers create a suitable learning environment where they are responsive to students' voices and it subsequently leads to actions (Williams, 2017). Student agency is critical in language learning and discourse analysis will enable one to study the interaction that takes place in a language classroom. We can then understand how the learning environment is structured and to what extent there is a capacity for student agency.

This paper focuses on the data collected from the Tamil classrooms under the research project (DEV03/SL15) (Seetha Lakshmi et al., 2015). Data has been audio and video recorded and transcribed into Tamil and translated into English. The preliminary finding of two Primary classroom data reveals that the level of classroom interaction is increased by student agency and teachers' facilitation on sustained speaking.

2-4

Tamil Corpus Bank and Students' Spoken Language - An overview

**Seetha Lakshmi, Nanyang, Letchmi Devi d/o Ponnusamy, Nanthini Arivuselvan, Nanyang
Technological University**

This paper highlights the use of Standard Spoken Tamil (SST) in Tamil classrooms in Singapore. Tamil language is one of the official languages in Singapore and is taught as a Mother Tongue Language subject at second language level in schools. Due to changing demographic patterns, majority of young students come from homes where English is the dominant language and are acquiring Standard Spoken Tamil (MOE, 2005) in their Tamil language classes. The implementation of Standard Spoken Tamil into curriculum took place in 2005. The research project team (DEV13/15 SL) has been working on creating a corpus bank of the classroom spoken language. Out of the 27 classrooms selected, data was collected from 14 classrooms through audio and video recordings which have been transcribed into Tamil and translated into English to create a corpus. The corpus has the potential materials to conduct courses for teachers to develop their teaching (Chambers, Farr and O'Riordan, 2011; Cubero and Ignacio, 2011; Honan, 2010, Zaki, 2017) approaches and strategies based on the students' use of language in the Tamil classrooms. Data from six Primary and eight Secondary classroom lessons were analysed to examine the teachers' and students' use of SST during their lessons. Preliminary findings show that there is a significant improvement in the use SST between teachers and students due to their mutual understanding, sustained engagement, encouragement and motivation in the class.

2-5

Corpus Keyword-aided Business English Instruction with dual aim

**CHEN Li-dan, He An-ping , Guangdong University of Foreign Studies, Guangdong University of
Foreign Studies/ South China Normal University**

The present study explores the application of corpus linguistics keyword approach in Business English instruction. In light of CLIL(Content and Language Integrated Learning) theory and based on self-constructed Business English thematic corpora, the authors developed a series of instructional designs including subject matter clarification by keywords, thematic knowledge framing by keyword clustering, and knowledge refinement by keyword collocations and theme-specific vocabulary learning by keyword concordances. A one-semester implementation on undergraduates' Comprehensive Business English course suggested that all these designs can improve learners' ability in framing theme-based business knowledge and expressing the subject matter in English compositions, thus contributing to the dual aim of disciplinary knowledge construction and English competence development.

2-6

A Corpus-Based Study of Feedbacks in Daily Conversation for Pedagogical Applications

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

As a necessary part of classroom interaction, teacher feedbacks are of crucial importance for effective classroom teaching and learning, even more so for learners of English in China, where English classes are still a primary source for their formal English education. While the research topic is largely understudied at home (Zhang & Dai, 2001), previous studies of teacher feedback have largely focused on the typology of various feedbacks (Cullen, 2002; Long, 1996; Lyster & Ranta, 1997; Vigil & Oller, 1976) based on data collected from the classroom. The current study is motivated by the rationale that a fresh insight can be gained from authentic daily conversations for more effective and natural teacher feedback and aims at an investigation of the feedbacks observed in the Switchboard Dialogue Act Corpus, a native English spoken corpus with 1.5 million words. With a much larger size of data than previous studies, the corpus is also chosen for its annotation of dialogue act information (Fang, Cao, Bunt & Liu 2012) pertaining to the communicative function and semantic content of each utterance. The current study will examine the different types of feedbacks in this corpus and, more importantly, produce an account of the various types of linguistic expressions of feedback showing specific context of use. Finally, pedagogical implications will be discussed based on findings arising from the study.

2-7

Grade Features of Epistemic Modality's Keyness in English Major's Argumentative Essays

Long Shaoyun, Liu Dilin, Jiangxi Normal University, The University of Alabama

With BaRENLOB(including BAWE+RA+ENS+LOCNESS+BNC essays) as the reference corpus, this paper makes a comparison of the grade features of the cognitive modality sequences used in the argumentative essays in Spoken and Written English Corpus of Chinese Learners 1.0(SWECCL 1.0).Research findings reveal:1) All BIS, LOG_L and LOG_R basically present significant grade differences except for the analogue cognitive sequences whether it is through variance analysis or robustness test.2) Means plots of variance analysis suggest that, as for cognitive sequence, both BIS and LOG_L rise with grade and reach the peak in the third grade , and fall to the lowest in the forth grade ,while LOG_R goes up slowly with grade, reaching its peak value in the forth grade.With regards to the obligation sequence, BIS and LOG_L are gradually decreasing with the rise of grade and there are reversed and high values in the second grade while LOG_R goes down in riptide and slack tide;concerning the analogue cognitive sequences,there exist no significant grade difference although BIS,the LOG_L and LOG_R all remain basically flat and consistent as the grade rises, sometimes fluctuated slightly, only to attain the best in the fourth grade .3) From the above conclusions, it can be seen that: A. The more the cognitive sequence appears, the less the obligation sequence does.B. The results of parametric test and non-parametric test

are fundamentally different, and some conclusions are completely contrary. These phenomena highlight the systematicness, incompleteness and instability of interlanguage language. In the end, the paper explores the specific causes of these phenomena from such theories as the phenomenon of language tide, interlanguage features and the practice of modal language acquisition.

2-8

MOOC Instructors' Linguistic Features – A Corpus-Based Research

Yuan Fei, Sun Youzhi, Tian Zuoyu, Tsinghua University, China University of Mining and Technology, Indiana University

The current study focuses on instructors' linguistic features in Massive Open online Course (hereafter MOOC). The researchers built a corpus from seven popular MOOCs provided by Tsinghua University, and intended to find the relationship between instructors' linguistic features and learners' involvement in the courses. It was found that the rate of participation of courses varied significantly, and the instructors of the more popular ones showed a faster pace of speaking and a higher likelihood of using adverbs and prepositions.

Group 3

3-1

Most Common Academic English Collocations: A Corpus-Driven Study

Lei Lei, Liu Dilin, Huazhong University of Science and Technology, The University of Alabama

Word collocations are very important albeit challenging lexical usages for ESL/EFL learners. Hence, educators and researchers have long tried to identify the typical English collocations and the challenges involved in learning them. Based on research on the types of collocations most useful to ESL/EFL learners and by employing a large balanced academic English corpus and applying a rigorous five-step extraction/selection method that draws on new research-based procedures, this study has developed a comprehensive and type-balanced academic English collocation list (AECL). A close comparison of the collocations on AECL with those found in well-known collocation dictionaries of general English and on the three existing academic English collocation lists indicates that AECL indeed contains mainly academic English-specific collocations rather than general English collocations. In addition, it offers substantially more academic collocations, especially the types learners may need to grasp, than the two existing lists do. Pedagogical and research implications are also discussed.

3-2

线性单位语法框架下的学术英语口语词块研究

张绪华, 张大群, 复旦大学, 南昌大学

目前常见的词块研究方法, 包括基于解读索引行 (Sinclair, 1991, 2004; Stubbs, 2001; 卫乃兴, 2011; 甄凤超&王华, 2012), 基于短语复现统计 (Biber et al., 1999; Biber et al., 2004; Hyland, 2008), 以及基于语言认知 (Divjak & Gries: 2008) 等方法, 从不同的角度对多词单位的形式和意义进行了解读。然而, 除了在多词单位中占比较小的习语外, 其他词块的存在形式和确切边界仍未得到充分证实。

线性单位语法 (Sinclair & Mauranen, 2006) 针对传统语法没有给予语言的线性特征以足够重视, 不能解释会话者如何共同构建会话, 不关注语言的线性特征和词块研究等问题, 提出从语言自然的线性组合关系出发, 采用数据驱动方法进行文本切分。线性单位语法尚处于初始阶段, 包括文本切块在内的语料分析都依赖于人工处理, 其机器自动实现还有待摸索。

以语言流利性为探究目标的学者认为, 口语中的停顿和重音等语音连贯性信息, 可以作为词块的边界标识 (Pawley, 1986; Rigggenbach, 1991; Hickey, 1993; Dahlmann & Adolphs, 2007)。有鉴于此, 本文尝试在线性单位语法框架下, 以语音连贯性作为词块整体性的判断标准, 即以停顿信息作为词块的边界标识, 将学术口语语料切分为词块, 并就词块的复现情况、长度信息, 以及词块自动切分方法进行讨论。

3-3

Cohesive Verbs: Patterns and Collocations in Research Articles

Dong Min, Qiu Xixin, Beihang University

Cohesion realized by demonstrative pronouns (Martin, 1992; Swales, 2005; Wulff, Roemer and Swales, 2012; Gray and Cortes, 2011), shell nouns (Aktas and Cortes, 2008; Hunston & Francis, 2000; Flowerdew and Forest, 2015; Schmid, 2000) and cohesive nouns (Hinkel, 2004) is found to be associated with a stretch of discourse and has a close tie with verbs such as *allow*, *involve* and *suggest*. However, little attention has been drawn on the corpus-based empirical study of cohesion realized by such verbs. This paper investigates the frequent grammar patterns and collocations and discusses their applicability to EAP pedagogy.

Two sub-corpora extracted from JDEST are analyzed(see below for overall description) and then POS tagged using AUTASYS (Fang, 2007)

	Disciplines	Texts	Type/Token Ratio
Natural	25	2,447	1.90
Social	15	683	2.08

The present study attempts to retrieve and identify such cohesive verbs by searching with two patterns (see below) co-occurring with the following node nominals: 1) self-mentioned research nouns (paper, section, study); 2) research-activity-related nouns (analysis, approach, method, model, process, theory); and 3) shell nouns (case, problem, example).

- a) This/this + (Node) + Verb + Noun/Noun to-inf/wh-cl/that-cl
- b) Node + Verb + Noun/Noun to-inf/wh-cl/that-cl

The results indicate that journal scholars' use of cohesive verbs shares the same yet different in order top 5 words in frequency(indicate, mean, provide, suggest, show), and both express cohesion with This/this/Node + Verb + Noun/that-cl patterns more often than other grammar patterns. And interestingly, some node nominals(e.g. allow and mean almost exclusively collocate with clause-initial this) are found to have collocational preference with certain cohesive verbs and some verbs(e.g. show and reveal are overwhelmingly used in past or perfect tense while suggest and involve prefer present tense.) to have preference with certain tense and modality.

3-4

A contrastive Analysis of Discourse Strategies in Ph. D Dissertation Abstracts in Linguistics

Li Xiaohong, Shanghai Ocean University

This paper takes as a starting point highly frequent sequences in the abstracts of doctoral dissertations by Chinese ESL learners and English native speakers. The typical phraseological patterns taking these sequences as essential constitutes are examined with their discourse functions recognised in rhetorical structures of the abstracts. Contrastive analyses are made to explore discourse strategies used by Chinese ESL writers. The study shows that the discourse

strategies employed by Chinese Ph.D students differ from those by native speakers, which include establishing an approving link with the literature, highlighting the theoretical basis, maintaining an impersonal attitude and focusing on comparison and description. The study could provide theoretical and linguistic evidence for EAP studies, studies of phraseology and teaching of English academic writing.

3-5

Comparing Engagement Markers in Economics Research Articles and Popular Science Writing: A Corpus-based Study He Mengyu, Universiti Sains Malaysia

Engagement markers serve to signal the presence of readers and help writers craft a credentialed text. As a key rhetorical feature, the use of engagement markers in academic writing has been explored in recent years. Despite this, few studies have investigated the use of engagement markers in research articles and popular science writing. In addition, studies on engagement markers in academic genres have not drawn extensively on corpus methodology. Following Hyland's (2005) framework of engagement markers, this study explores the use of engagement markers in economics research articles (henceforth RAs) and popular science writing (henceforth POPs). The analysis was based on two self-compiled corpora of RAs and POPs. The corpus analysis tool *LanCSBox* (Brezina, McEnery, & Wattam, 2015) was used in identifying and analysing the engagement markers. The results of this study show that RAs writers used more engagement markers than POPs writers. Among all the different subtypes of engagement markers, directives were predominantly employed by the writers in both genres. The use of engagement markers by writers in both genres are discussed in relation to the generic targeted audiences and different communicative purposes of the two genres. These findings indicate important implications for EAP writing.

3-6

Exploring Phraseological Variation in Academic Writing: A Corpus Analysis Ang Leng Hong, Universiti Sains Malaysia

Phraseology has been explored widely in recent years. Nevertheless, the phenomenon of variation within the phraseological sequences has not received considerable attention in the literature. This study aims at filling in this research gap by looking at both the continuous and discontinuous phraseological sequences in journal articles published in the field of International Business Management (IBM). The study is concerned with the analyses of lexical bundles and lexical frames. The software programs *Collocate*, *Antconc* and *kfNgram* are used in this study for research data extraction and analysis purposes. The findings of this study indicate that 3-word lexical bundles and the structure "noun phrase + *of*-phrase fragment" are most widely used by expert writers in IBM. With regard to the functional categories, referential expressions are the most frequent lexical bundles, followed by stance expressions and discourse organisers. 3-word lexical frames are found to be dominant in IBM. There are more fixed lexical frames in the category of 3-word lexical frame. The category of

4-word lexical frame contains more highly predictable lexical frames than the category of 3-word lexical frames. The structure “function word frame” is the most common frame structure in IBM. Also, the variable slots are mostly filled up by content words. This study contributes to the development of EAP course as the lexical bundles and lexical frames investigated could be used to develop teaching materials in the course. This study also has implication on the understanding of phraseology in which phraseology can in fact be variable. This contributes to the understanding of actual phraseological tendency in academic writing.

3-7

Patterns and Meanings of Verbal Hedges in English Research Articles by Chinese and Western Scholars: A Corpus-based Contrastive Study

Yang Yuesen, Wei Naixing, Beihang University

The traditional view of academic discourse held that it is an objective genre featuring exactness and precision in expression and neutral stance conveyed by the writer. However, various studies (e.g. Hyland, 1996, 1998; Markkanen & Schröder, 1997) demonstrate that deliberate imprecision or hedging expression is characteristic of academic discourse, and the appropriate use of this kind of device forms an indispensable part of successful academic communication. The present research conducts a corpus-based contrastive analysis of patterns and meanings of verbal hedges in academic texts by Chinese and Western scholars. The corpora analyzed are the Chinese Scholars Academic English Corpus (CSAEC) and the Western Scholars Academic English Corpus (WSAEC). As a crucial notion involved in the present study, a working definition for hedging pattern in academic discourse is proposed. On the basis of the definition, this article goes on to discuss the possible forms of the patterns of verbal hedges in English research articles written by Chinese and Western scholars, explore the discourse functions and strategic meanings realized by recurrent hedging lexical sequences, and compare the hedging features of academic texts written by Chinese and Western scholars. The findings of this study have thrown important insights for understanding the feature and nature of hedging patterns in association with academic evaluation and strategy. They are of potential value for improving academic writing pedagogy.

3-8

专业语篇N-grams提取及其MTI教学和论文写作应用

管新潮，上海交通大学

本文描述了如何运用Python编程提取N-grams的算法设计过程。NLTK虽然提供了各种提取N-grams的可能性，但如何将不同代码有效地关联在一起，则直接影响到N-grams的提取效率和有效性。通过对法律语篇的N-grams提取操作发现，需要为不同学科语篇的提取工作设置不同的停用词，这一点甚至适用于不同的法律语篇。其提取结果可显著提升法律语篇的学习效率以及明显改进法律汉译英实践。N-grams提取的论文写作应用可参见雷蕾的论文“Research Trends in Applied Linguistics from 2005 to 2016: A Bibliometric Analysis and Its Implications”。该文实现了将Python和AntConc两种不同工具的最佳功能进行组合应用，并进而通过N-grams确定应用语言学的研究趋势。

Group 4

4-1

英汉双语认知属性库的构建及计量分析

李斌, 宋丽, 匡海波, 戴茹冰, 南京师范大学

词语的认知属性是指语言使用者对词语代表的概念或实体的认知体验凝结到词义中的各种属性。传统的词义研究方法往往依赖于内省法和问卷调查来进行词语认知意义的个案研究。本文则使用大数据方法, 借助中英文搜索引擎, 采集形成了20多万条的中文认知属性和2万多条英文属性。进而使用统计方法对比了双语的认知属性异同, 并使用《知网》的双语词典进行可视化, 最后进行了双语认知属性相似度计算, 得出了更符合认知的词语计算方法。

4-2

论体育英语参照语料库建设的三维模式

陈建生, 广州体育学院

本文旨在为建设体育英语参照语料库提供一个理论框架和操作指引。本研究首先从体育的社会性, 主要包括体育的民族性、文化性和经济性三个方面, 厘清主要体育项目; 提出保证不同体育项目的语料均衡性的量化标准。其次, 以语域类型为参考, 对体育英语话语进行分类, 为体育英语语料库子库的组成提供语言学依据。最后, 从语料取样的量化准则出发, 强调了体育英语语料库需要尊重建库的量化标准。这个“社会-语言-量化”的三维模式, 从体育项目分类, 到语言功能分类, 到语料比例的确定为体育英语参照语料库建设提出了整体构想。

4-3

“一带一路”核心语汇多模态汉俄平行语料库建设(2013—2021)

白禹, 新疆大学

“一带一路”大背景、新形势、新理念、新方法为汉语传播和语言服务的发展带来了前所未有的机遇。我们应该以全新的国际传播与共享视角让世界了解中国的发展、中国的文化。

多模态语料库已成为语料库语言学的新起之秀, 被称为“语料库的4.0版”。笔者以“一带一路”背景下各领域的汉俄互译核心语汇和文字、语音、动态图像等多模态平行语料库建设两个部分为研究对象, 将按照沿线国家、地区和城市; 政策法规、文件、活动; 各领域主要代表人物等三大分类标准搜集、筛选出2000-3000个核心语汇, 目标是建设一个新型的、兼具语料数据多模态集成功能及较为强大的检索功能的《“一带一路”核心语汇多模态汉俄平行语料库》电脑软件和手机App, 预期成果形式如下:



本《语料库》创新性地将文本符号与视觉符号和听觉符号相关联，将符号和“一带一路”这一社会政治文化大背景相联系。紧跟多模态语言学研究的前沿问题，是一种简单、高效获得汉俄双语语言服务的新模式。其成果将应用于教学研究和翻译实践，不仅可以有效发挥语言服务在推进“一带一路”建设中的重要支撑保障作用，还可为其他人文社科领域研究提供参考和借鉴。

4-4

基于知识图谱理论的国家“语言战略”研究热点可视化分析

李夏青，中央民族大学

With the acceleration of the globalization, people's language view has undergone major changes. Some linguists are no longer limited to the study of linguistic microcosmic phenomena such as phonetics, vocabulary, and grammar, but are instead turning their attention to macro-language phenomena that are closely related to national interests and national security. Especially with the vigorous advancement of the construction of “the Belt and Road Initiatives” in recent years, language strategies have also become an important topic for linguists. Based on the theory of knowledge maps, this study takes 162 journal papers related to language strategy in China national knowledge internet(CNKI) as the research object, and analyzes the hotspots and frontiers of language strategy research in the past 20 years. It provides help not only for researchers to understand the language strategy aspects but also provides some inspiration for the country to find its own deficiencies in language strategy and formulate a suitable language strategy.

4-5

武汉地区孤独症儿童语料库研究——以名词为例

肖诚

孤独症谱系障碍(ASD)是全球性儿童发育障碍中较为普遍的一种病型，主要的病理表现包括语言发育水平偏低、兴趣范围窄且单一、缺乏基本的社交和行为刻板。

文章结合国内外对孤独症儿童语言障碍与康复方面的相关研究，阐明孤独症的定义、

类型及病理表现。通过在武汉市精神卫生中心六个月的实地观察与调研，参照普通儿童的语言发育水平，以名词为切入点，设计出符合3~8岁孤独症儿童语言规律的调查问卷，选取生活在武汉地区的孤独症儿童为样本，建立孤独症儿童语言习得情况的小型语料库。

文章以名词为例，基于孤独症儿童的真实语料库，着重从横向、纵向两方面对武汉地区孤独症儿童的名词习得情况进行分析和整理。在横向对比上，梳理了孤独症儿童的称谓、自然、颜色、交通工具、时间、形状、家居生活、动物、地理、方位、食物、服装、抽象物体、其他等15类名词习得障碍；在纵向对比上，归纳了对孤独症儿童的常用名词、抽象名词、复合名词等名词习得规律，试图对武汉地区的孤独症儿童的语言康复训练提出合理的语言学建议。

4-6

形式动词“加以”和“予以”近义辨析：基于语料库的构式搭配分析

田笑语，胡梅芬，张炜炜，上海外国语大学

形式动词（袁杰、夏允贻1984，朱德熙1985）是汉语动词中一个特殊的小类，它具备普通动词的语法功能，但没有具体语义，只起到“形式”上的作用，加在另一个动词的前面构成动宾结构，如“进行”、“加以”、“予以”、“给予”、“作”等。目前对形式动词的研究多是定性的描写和分类（如，周刚1987；胡裕树、范晓1995；王永娜2010），但是要辨析各个形式动词之间的语义差异，离不开基于自然语料的定量研究。构式搭配分析法（Stefanowitsch & Gries 2003）有效地结合了认知语言学的理论成果与语料库语言学的研究方法，以量化的手段揭示构式之间的意义差异。本研究基于北京大学现代汉语语料库，在构式语法理论（Goldberg 1995, 2006）框架下，采用构式搭配分析法考察[加以+动词]和[予以+动词]两个近义构式的动词搭配偏好，从而达到对“加以”和“予以”这两个形式动词进行近义辨析的目的。研究发现，二者的搭配偏好有较大差异：“加以”搭配的动词语义范围更为广泛，一般为主体有意识主动进行的动作行为，带有强烈的主观能动性，涉及表示采取积极措施意义的抽象动词等；而“予以”搭配的动词集中于表示奖惩意义的法律用语。文章最后总结了本次研究的不足之处，以期为将来关于形式动词的定量研究提供启发。

4-7

“HSK动态作文语料库”（V2.0）的设计理念与功能

张宝林，北京语言大学

“HSK 动态作文语料库”是母语非汉语的外国人参加高等汉语水平考试(HSK高等)作文考试的答卷语料库，自2006年12月建成伊始，即向海内外各界用户免费开放，极大地促进了作为第二语言的汉语教学与相关研究。然而，该库软件系统技术陈旧，难以保证正常运转，严重影响广大用户的教学与研究使用。为此，我们采用目前主流的计算机语言，重新开发了该库的软件系统。

我们的基本设计理念，一是要确保语料库安全运转；二是增强功能，为用户提供更多的使用便利。相比于1.1版，增加了下列功能：

1.查询，在“一般检索”之外，增加了“特定条件检索”，极大地提高了检索的功能与效率。

2.修改, 用户可以对语料中和标注中的错误进行修改, 经管理员确认, 可以更新语料, 提高语料及其标注的质量。

3.下载, 提供了自动下载功能, 500条以下全部下载, 500条以上随机下载, 可以保证相关研究的科学性。

4.数据统计增加了可视化图形设计。

5.强化用户与语料库建设者的沟通, 增加了留言及其反馈功能。

6.增加了用户的“个人工作区”, 用户可以在此对自己的相关信息进行维护。今后还将使其成为用户分析语料与进行研究的工作平台。

目前语料库还存在一些不足之处, 例如标注内容还不够全面, 按词性检索词语搭配尚未实现等。我们将在下一步的工作中对其进行改进。

4-8

基于语料库的西方媒体之“深圳形象”研究

侯智德, 暨南大学深圳旅游学院

迄今关于“深圳形象”的研究主要是从城市文化、城市发展和新闻传播视角对“深圳形象”的塑造、宣传和建构展开分析。然而, 就媒体有关“涉深报道”隐性建构的“深圳形象”实证研究还不够系统和全面。本文旨在研究西方媒体“涉深报道”中“Shenzhen”一词在新闻报道所表征的“深圳形象”。通过利用语料库语言学工具Concgram 1.0 (Greaves 2009), 对创建的“深圳形象媒体话语语料库”做批量处理: 从确定“Shenzhen”和“Shenzhen + 同现词”为核心 (core) 开始, 在索引行中展开由搭配 (collocation) 到类联接 (colligation), 到语义趋向 (semantic preference), 再到语义韵 (semantic prosody) 的分析。即从具体化的搭配到范畴化的类联接, 再到更为抽象的语义趋向和语义韵, 形成“Shenzhen”完整的具有累积效果的语义轮廓 (Sinclair 2004; Cheng 2006; Cheng et al. 2009; Hou 2018)。本文试图揭示体现“深圳形象”的西方媒体话语建构的特征及其组成元素, 为提升“深圳形象”话语权建设制定相应的对外策略和提供客观的语言学依据。

Group 5

5-1

Integrating Corpus-based Analysis with Experimental Data: A Corpus-based Experimental Study of Implicit and Explicit L2 Collocational Knowledge

Lu Jun, Yangzhou University

Compared with psycholinguists, who often exploit the benefit of combining corpus and experimental data, corpus linguists do so far less frequently. Corpora, as a supreme tool for linguistic data analysis, are well-known for such advantages as natural language context and large data range while experimental approaches have such well-known advantages: allowing the study of low-frequency phenomena, especially those that are too infrequent in corpora; making it possible to systematically control certain variables; and permitting the study of online process. These advantages of corpora and experiments are largely complementary. Thus, if aided by experimental data, corpus linguistic research might have greater potential in addressing relevant linguistic issues. The present study begins with addressing the typicality issue in the study of L2 knowledge and proposes a corpus-based psycholinguistic experimental paradigm. Within the paradigm, the developmental pattern of and the interface relationship between implicit and explicit knowledge of L2 collocations are examined. The data shows that L2 transparent and semi-fixed collocations differ considerably in the construct and in the development of implicit and explicit knowledge. The developmental pattern indicates a weak interface between implicit and explicit knowledge of L2 collocations. Further analysis suggests that the above features are closely associated with L1 influence. L1 collocational knowledge facilitates the hypothesis formation of L2 transparent collocations, yielding Type I implicit collocational knowledge. And hypothesis testing produces explicit collocational knowledge. As for semi-fixed collocations, the mere congruency with their L1 equivalents at phrasal level facilitates holistic processing of such collocations, yielding explicit L2 collocational knowledge. Both hypothesis testing and holistic processing contribute to repeated priming of the co-occurrence of collocates, yielding Type II implicit knowledge of transparent and semi-fixed collocations respectively. Finally, both theoretical and pedagogical implications are discussed.

5-2

An Empirical Study on the Cognitive Level of Two-syllable Lexicalization of Chinese Dimensionative Adjectives

Xu Jianjun, Ningbo University

Based on the authoritative corpus, from the diachronic and synchronic perspectives, this paper examines the possibility of Chinese monosyllabic dimension adjectives into DTM adjectives and explores the cognitive level and cognitive interpretation of Chinese monophonic dimension adjectives. The study found that, firstly, from the diachronic and synchronic perspectives, the mono-scale adjectives "high" and "deep" always occupy the primary position in people's cognitive order; secondly, from the diachronic and synchronic

perspectives, monosyllabic dimension adjectives "thick", "long", "big" in people's perceptions vary with time, and in the synchronic level of written language and spoken language, cognitive order is also different. Thirdly, from the investigation of the diachronic corpus, the cognitive order of "thick" and "long" is always in the same level.

5-3

Phraseological Features of Application Letters Written by Chinese College Students

Zhang Le, University of Shanghai for Science and Technology

Application letters is an important promotional business-writing genre of English for Occupational Purposes (EOP). This paper adopts a corpus-based phraseological perspective, and extracts a number of multiword sequences from Chinese Science and Engineering Majors Written English Corpus by using node-word approach, phrase-frame approach, and phrase-list approach. The purpose is to examine formal and functional features of multi-word sequences in application letters written by Chinese college students. It is found that Chinese students are generally capable of producing various semi-fixed phrases with specific semantic, functional, and move features so as to achieve promotional communicative purposes. Differences are also found between the learner corpus and the reference corpus with regard to collocation, degree of informativeness and the use of politeness strategy. The results are expected to be helpful for research on English for Specific Purposes and for EOP writing courses.

5-4

Comparative Analysis of the Polysemy of Preposition "in" by Native English Speakers and Chinese English Learners—A Corpus-based Study

Hong Yang, Pan Hong, Wuhan University of Technology

This paper aims to investigate the usage of prepositions in by native English speakers and Chinese English learners, which including differences in frequency and differences in the usage of its polysemy. First, with the help of AntConc3.2.1, the index lines of preposition "in" both in the Brown corpus and CLEC corpus are retrieved. (There were 21,422 cases in the Brown Corpus and 22,500 in the CLEC.) Then, using the Conc Sampler to randomly get 5000 observation lines both in the two corpora for further artificial annotation of semantic information: (1) prototypical meaning - "YX"; (2) independent meaning - "DL"; (3) metaphorical meaning - "YY". Finally, compare the difference between Chinese English learners and English native speakers in the usage of prepositions "in" in its polysemy. The results show there are significant differences between Chinese students and native English speakers when using the proposition "in". The reasons include: the acquisition of prototype meaning, the negative transfer of mother tongue and the student's personal study habits and teacher's teaching influence.

5-5

**Meaning Shift Units in the Perspective of Local Grammars:
A Study of *UPON* in Business English Correspondence Texts**
Liu Yunfeng, Beijing Foreign Studies University/ Anhui Polytechnic University

Local grammars have become a new approach to language description in corpus linguistics in recent more than 10 years. However, the studies of local grammars based on specific word or word form center mostly on lexical words and rarely on grammatical words; furthermore they involve few functional categories in research procedures. In the perspective of local grammars, this paper describes the local grammar patterns of meaning shift units around *UPON* in Business English Correspondence Corpus with functional categories. It has been found that in business English correspondence texts all sentences in which *UPON* occurs explicitly show the function of service; meaning shift units around *UPON* present three canonical forms, which demonstrate the phenomenon of meaning shift under the restriction of co-selection, and respectively express positive commitment to anticipated service, regret for the occurrence of negative events, and objective description of transaction feasibility. The paper concludes that in the professional text, based on specific grammatical word and by establishing functional categories, the description of local grammars can present dynamic co-selection of meaning shift units in terms of lexis, grammar, meaning and function, and can reveal the fact that grammatical words are as important as lexical words in constructing phrase structure & its overall meaning, which has new implications for the combination of phraseology and text analysis.

5-6

A Corpus-Based Analysis of Synonyms—Take *absolutely* and *thoroughly* for Example
Jin Yingying, University of Shanghai for Science and Technology

Based on the BNC corpus, this paper makes an analysis of the word *absolutely* and *thoroughly* from these aspects: word frequency, frequent register, collocation, colligation and semantic prosody. Therefore, the results of the analysis can be listed as following. Firstly, the word frequency of *absolutely* is higher than *thoroughly* in the BNC corpus. Secondly, *absolutely* is mainly used in spoken English and fiction while *thoroughly* in fiction and magazine. Besides, the frequency of *absolutely* is higher than *thoroughly* in fiction in BNC. Thirdly, the obvious verb collocates of *absolutely* mainly are negative words while the obvious verb collocates of *thoroughly* mainly are neutral words and negative words. The obvious verb collocates of *absolutely* mainly are negative words while the obvious verb collocates of *thoroughly* mainly are neutral words and negative words. Fourthly, both *thoroughly* and *absolutely* can be used in the following colligations: V+Adv. Adv.+V Adv.+V+N Adv.+Adj. Adv.+Adj.+N Adv.+Adj.+prep.+N V+Adv.+Adj. BE+Adv.+Adj. Art.+Adv.+Adj.+N. However, *absolutely* can be used in the pattern “Int.+Adv.” while *thoroughly* cannot; *thoroughly* can be used in the following pattern “V+Pron.+Adv.” while *absolutely* cannot. However, there exist some differences in the following sentence pattern, that is, [It is *absolutely/thoroughly* Adj. that]. It is clearly that the degree of collocation of

absolutely is higher than *thoroughly* in the sentence pattern. In this pattern, the obvious adjectives of *absolutely* can be classified into three types: positive words, negative words and neutral words. While the obvious adjective of *thoroughly* is only related to one neutral word. Fifthly, different part of speech of collocates makes the semantic prosody of *absolutely* and *thoroughly* different. If the obvious collocate of *absolutely* is adjective, the semantic prosody of *absolutely* is positive prosody; If verb, then negative prosody. If the obvious collocate of *thoroughly* is adjective, the semantic prosody of *thoroughly* is negative prosody; If verb, then negative or neutral prosody. This paper aims to make a comprehensive analysis of the synonyms, thus, contributing to foreign language learning and teaching.

5-7

Research on the Structure of the English Names of Military Equipment for NLP

Lei Shujie, Xing Fukun, Luoyang Foreign Language University

Military equipment name is a sub-class of named entity and is an important expression in military field. It is of great value to study the constructive models of military equipment name and the distribution of its internal composition for linguistic ontology research and automatic processing of information in specific fields. This paper takes English names of military equipment as its subject and constructs a basic database covering the names of many kinds of military equipment. Through manual annotations, this paper conducts a comprehensive survey on the constructive models of English names of military equipment and the distribution of their internal composition, summarizes the main constructive models, constructive elements and their distribution characteristics. Besides, English names of military equipment are also analyzed compared with the common named entities such as personal names, place names and so on. In this paper, the proposed constructive models are tested with the untagged datasets of English names of military equipment. The test results show that the proposed constructive models have good descriptive ability and can identify the English names of military equipment with high accuracy. The research methods and results of this paper can not only effectively support the information processing and related teaching work in the military field, but also have important reference value for the related studies in other languages.

5-8

基于词向量模型的中国学习者英语词汇语义特征研究

邓海龙, 北京外国语大学/赣南师范大学

近年来, 自然语言处理领域中基于分布式语义表示的词向量模型取得了重要进展, 它能够高效计算和准确表示自然语言中的词汇语义关系。词向量模型在学习者语言研究领域具有很大应用潜力。从英语学习者词汇运用研究来看, 限于传统语言分析技术, 迄今为止大部分基于语料库的研究都仅仅停留在词形(包括词组或者赋码后的词形及词组)频数统计及对比分析之上, 很难进一步深入到语义层面; 虽然采用精细标注语料并

结合定性研究的方法可以做到深度分析，但工作量巨大，能够考察的学习者语料数量极其有限。本研究从外语教学与研究出版社的iWrite平台抽样十万多篇英语作文（约1500多万形符）建成学习者语料库，并运用word2vec工具分别对所建学习语料库和英语国家语料库（BNC）构建词向量，然后将所构建的两个词向量进行对比分析，考察中国学习者英语词汇语义特征。结果显示，学习者英语词汇语义精细度不如本族语，词义重心也往往与本族语存在较大差异。研究表明，基于词向量模型的方法能够对大量语料进行快速词汇语义特征分析，有利于全面系统地把握学习者词汇语义特征。

Group 6

6-1

语料库中的语用标注 姜占好, 西安外国语大学

Both researchers and teachers pay more and more attention to corpus due to its data richness and authenticity. Pragmatic annotation, as an approach to keeping language authenticity, has been one of the heated topics in corpus linguistics. This paper, on the basis of literature review, has explored its necessity, its annotated objects, and its challenges as well as its potential research foci.

6-2

并购话语中的中国企业形象研究：基于语料库的话语—历史观方法 赵永刚, 北京外国语大学

本研究以话语—历史观 (Reisigl & Wodak 2009) 为研究框架, 结合及物性理论和评价理论, 基于自建的“中国企业海外并购新闻语料库”, 对比分析了中美英文媒体 *China Daily* (CD) 和 *The New York Times* (NYT) 对中国企业形象的构建及形象构建的话语策略。先导研究发现:

- 1) CD、NYT 都将中国企业描述为最突出、最活跃的参与者, 但 NYT 更感兴趣的是中国企业的政治背景, 而 CD 局限于对事件本身的报道。
- 2) NYT 和 CD 都更多地将间接转述融入报道者的立场, 但 CD 尤其偏爱具体的专家意见; NYT 倾向于平衡的消息来源, 但在来源的确定性上稍显不足。
- 3) CD 和 NYT 在报道中都有偏向本国企业的倾向, 但 NYT 更多地操控指称策略和述谓策略来引导和影响读者的立场。
- 4) CD 最多使用的论辩策略是权威和有利, 即诉诸权威, 论证中国企业“走出去”对中国企业的利益, 而忽视了“对方利益”; NYT 则使用了威胁、公正等论辩策略来暗示中国企业的海外并购是国家的“侵略”行为。

6-3

基于语料库的中-美政府工作报告语域特征多维分析 赵朝永, 胡婧, 华东师范大学

本研究运用多维分析法 (MDA), 对比考察 2000-2018 年间中国《政府工作报告 (英文版)》与美国《国情咨文》的语域特征, 探析两者语域特征的拟合度。研究表明, 中美两国的政府工作报告在 5 个维度 (交互性/信息性表达、叙述性/非叙述性关切、指称明晰性/情境依赖型指称、信息抽象与具体程度、即席信息组织精细度) 上均存在显著差异, 其中维度 3 (指称明晰性/情境依赖型指称) 差异最大, 维度 1 (交互性/信息性表达) 和维度 4 (信息抽象与具体程度) 依序次之。线性回归结果显示, 造成不同维度差异的影响因素不尽相同, 这一发现对我国政治文献的英译具有一定启示。

6-4

A Corpus-based Study on Contrastive Discourse Markers in Chinese English Writing

Zou Hong, He Huaqing, China West Normal University

This study investigates the use of Contrastive Discourse Markers (CDMs) by Chinese middle school students in English writing via adopting contrastive interlanguage analysis and makes a comparison with native English speakers so as to discover the differences between them. Additionally, this paper would point out the errors of CDMs appeared in students' compositions in order to find deficiencies.

The paper would discuss three research questions: 1. Is there a significant difference on the overall frequency of CDMs between Chinese middle school students and native English speakers? 2. In terms of syntactic positions, is there a significant difference between them? 3. What error types of CDMs misused in students' compositions?

With the aid of the on-line retrieval platform with multilingual corpora at Beijing Foreign Studies University (BFSU), the study selects TECCL as a learner corpus and Crown as a reference corpus. The research procedures: Firstly, confirm the items of CDMs. Secondly, select the needed texts. Then, retrieve the search words one by one by means of CQPweb and cull the indexes that are not in the scope of the study so that count the frequencies accurately. Finally, employ Excel and Log-likelihood and Chi-square Calculator in data processing. The findings: 1. There is a significant difference on the overall frequency of CDMs between Chinese middle school students and native English speakers. 2. In the aspect of syntactic positions, the uses of CDMs by middle school students are not as flexible as native English speakers and the difference is significant. 3. The error types of CDMs misused in students' compositions are primarily on word form, preposition and conjunction.

The results of the study are analyzed and some suggestions are put forward in the hope of providing enlightenments for English teaching and learning and parallel studies.

6-5

Genre Theory and Corpus-driving Introduction Writing Models in Scientific Academic Articles

Xiao Xianming, Wuhan University of Technology

Introduction is one of the major constituents for an academic article. The discourse model stipulated from the theory of genre is hypothesized to be an important method for introduction writing and learning. This paper is intended to explore first the discourse models of the introduction in genre theory and the linguistic characteristics of moves in the introduction discourses. Then, it is to test the hypothesis that the models of introduction writing pedagogically from the genre theory can be an effective way of teaching introduction writing in EAP discourse by analyzing the self-established corpus of introductions down-loaded from Nature and Science. Therefore, it is of great value that the models of corpus-driving introduction writing shed light on how to write high-quality introductions of academic papers in English.

6-6

Changing Patterns of Apology in British English: A Local Grammar Based Diachronic Investigation

Su Hang, Beihang University

This paper extends our recent research on local grammars (Hunston & Su 2017; Su 2017, 2018, in press 2019; Su & Wei in press 2018) to diachronic speech act analysis, aiming to present a methodological innovation in corpus-based speech act studies in general. A brief review suggests that previous studies have moved directly from realisation forms to pragmatic speech functions with little consideration of the semantics of those forms and as such they only capture regularities in formal realisations but not the underlying semantic regularities. As a consequence, speech act realisations may have not been reliably quantified and, consequently, observations drawn by analysing formal realisations might be potentially questionable. This indicates the necessity to explore alternative method(s) which can be used to quantify more reliably speech act realisations across time. This paper argues that local grammars can be one candidate, which is subsequently demonstrated by tracing one particular speech act, viz. apology, in BNC1994 and BNC2014. The analyses show a less frequent use of performative verbs and a tendency of using more routinised, detached expressions to apologise. A possible explanation, as Leech (2014) notes, might be that social distance between individuals in British culture has been reduced, everyone being equal and claims solidarity with others, which in turn supports Taavitsainen and Jucker's (2008: 8) argument that "the diachronic development of speech acts can reveal the formulation of cultural scripts".

6-7

基于语料库的美国“一带一路”新闻报道的批评话语分析

杨紫彤, 刘爱军, 北京邮电大学

国家主席习近平在2013年首次提出“新丝绸之路经济带”和“21世纪海上丝绸之路”的合作倡议,二者简称为“一带一路”倡议,至今已将近五年。该倡议涉及经济、政治、科技、教育等诸多方面,并在海内外赢得了高度关注。西方媒体对此的报道十分广泛;作为在西方国家占主导地位的美国主流媒体,其报道的内容具有一定的研究价值。作者利用搜索主题词的方法筛选出《华盛顿邮报》中2017年的相关新闻报道,一共77篇,并利用AntConc3.4.4自建语料库。本文基于话语分析理论,将从语言的元功能:概念功能、人际功能和语篇功能三个方面进行分析讨论,总结美国“一带一路”新闻报道中的语言使用特点、媒体态度倾向以及社会意识形态影响。本文对读者提高政治经济类文献阅读提供了新思路新方法,有助于提高读者的批评性思维阅读能力及思辨能力。

6-8

The Influence of Topic Location on Information Distribution in *It*-extraposition and Non-extraposition: An Investigation Based on ICE-GB

Wang Zhong, Fang Chengyu Alex, Fan Weiwei, Huazhong University of Science and Technology/City University of Hong Kong, City University of Hong Kong, Huazhong University of Science and Technology

It is argued that the process of postposing sentence elements is related to given and new information distribution in a sentence (e.g. Birner & Ward, 1998). *It*-extraposition, with subject clause moved rightward to its noncanonical position and a dummy *it* occupying the subject position, is related to writers' changing of canonical word-order to locate new information to the right hemisphere of a sentence. However, not all *it*-extrapositions express new information (16.3%; Kaltenböck, 2004). This deviance is assumed to result from the location of the source topic for the ensuing sentence (Miller, 2001) and different text genres (Kaltenböck, 2005). This study attempts to bridge the relationship between (a) information status, composed of two dimensions including given- v.s. new- information and degrees of givenness/ ungivenness, (b) topic location, and (c) the choices of (non-)extraposition (primarily *it*-extraposition) in spoken and written texts. The primary data of this research are extracted from the British subcorpus of the International Corpus of English (ICE-GB). Over 1700 instances of *it*-extraposition were identified in both spoken and written texts by searching for constructions involving anticipatory *it* (marked as ANTIT in ICE-GB) as a provisional subject; a much smaller number of non-extraposed *it*-extraposition (a clause as logical and provisional subject of matrix clause) were also identified by semi-automatic methods. Information status (including given v.s. new information and degrees of givenness or ungivenness) and source of topic for the following sentence are then manually marked on matrix predicate and subject clause of (non-)extraposition. The relationships between the above-mentioned (a), (b), (c) variables are analyzed through statistical calculation and qualitative discussion of these identified cases of *it*-extraposition and non-extraposition in ICE-GB.

Group 7

7-1

基于语料库的“一带一路”英译在东盟国家的认知研究

邓斯芮, 周杰, 贵州大学

东盟国家处于“一带一路”的陆海交汇地带, 是中国推进“一带一路”建设的优先方向和重要伙伴。研究“一带一路”英译在东盟国家的使用情况有利于了解东盟国家对我国“一带一路”倡议的接受和认知情况, 从而为发展中国东盟关系提供一定的参考。本文通过自建东盟十国主流英文媒体“一带一路”新闻报道语料库, 利用AntConc3.4.4软件, 从共时和历时角度分析东盟国家对“一带一路”英语表达方式的接受度和使用情况。此外, 通过关键词表, 揭示东盟国家媒体对“一带一路”的倡议认知和主要关注点。

7-2

Organic and Mechanical Metaphors in *the Wealth of Nations*

Hu Chunyu, Lu Chunmei, Guangdong University of Foreign Studies

Cognitive linguists contend that metaphors shape our understanding of the world and our action in it, and that our ordinary conceptual system is fundamentally metaphorical in nature. In a discipline as abstract as economics, it is expected that metaphorical conceptualization is not an uncommon phenomenon. This paper aims to explore mechanical and organic metaphors, two significant root metaphors in economics, used by Adam Smith in his seminal work *An Inquiry into the Nature and Causes of the Wealth of Nations* (WN). The results show that 1) organic metaphors are more frequently employed, with a total of 226 occurrences, than mechanical metaphors (n=148); 2) three functions of metaphors have been achieved among the 31 metaphorical examples in WN; 3) the construction of metaphors and contextual factors are two prominent factors influencing the choice of preference to organic metaphor.

7-3

构式的操作界定: 认知固化与条件概率

房印杰, 北京邮电大学

构式语法内部先后提出过两种构式定义: 狭义定义(Goldberg, 1995)和广义定义(Goldberg, 2006)。狭义定义下的构式具有不可预测性(Unpredictable Property), 广义定义将该特征剔除, 代之以‘足够频次’(Goldberg, 2006: 5)作为构式的界定标准。这使得构式的涵盖范围扩大, 但在具体研究中对构式的认定出现困难。本文提出学界对构式范畴的争议源于没有厘清其理论/操作层面, 基于频次的认知固化可以作为构式的操作界定。认知固化的渐变性、层级性、动态性对应构式的原型性、抽象度、动态性。通过对英语本族语者的关系代词取舍开展基于条件概率的共现词分析发现: 1) 隐性/显性关系代词从句呈现基于条件概率的认知固化, 可以被界定为构式; 2) 隐性关系代词从句比显性关系代词从句更加靠近具体词语层面, 认知加工负载更小。

7-4

认知视域下记者招待会中隐喻口译策略的实证研究

杨璘璘, 广东行政职业学院

基于概念隐喻理论,对政府记者招待会平行语料库中的隐喻进行识别和归类。同时,基于认知翻译学,对语料中的隐喻口译策略进行归纳,并统计每一种口译策略的使用频率,揭示隐喻口译的特点以及译者口译策略选择的原因。研究发现语料中概念隐喻使用丰富,直译隐喻法是使用频率最高的策略。译者通常根据隐喻的语篇功能、中英文政治文本隐喻类别的相似程度、隐喻的语境和搭配原则等进行隐喻口译策略的选择。

7-5

基于营销学语料库的旅途隐喻研究

胡春雨,王进晓,广东外语外贸大学

隐喻是传递营销理念、形成有效营销策略的强大认知工具,其中概念隐喻MARKETING IS JOURNEY被视作一个非常重要的营销根隐喻。本文以五本权威的英文营销学教科书和高影响因子的三本英文营销学期刊中的论文为研究语料,建成三百万词规模的营销学英语语料库。借助已有旅途隐喻研究,制作源域词语检索清单,并结合“隐喻识别程序”(MIP),识别两个子库中旅途隐喻的使用状况。研究发现:1)不论是营销教科书还是期刊,对于旅途终点的重视性都大于旅途起点;2)营销教科书和期刊中有关旅途的主题有障碍和辅助、旅途方向、上下坡,以及快与慢的旅途等,其中期刊论文中所用隐喻比教科书更注重旅途的方向;3)由于教科书和期刊的对象读者不同,教科书针对学科新手,而期刊面向专家学者,导致两类营销学话语在旅途隐喻的选择上有所不同。

7-6

英语学习词典中外壳名词的释义模式初探

——以《牛津高阶英语词典》(第8版)为例

张雪梅,刘萍,上海外国语大学

英语学习词典是外语学习不可或缺的工具,从释义局部语法视角研究英语学习词典的释义模式具有理论和现实意义。Barnbrook(1995,2002)对《柯林斯学习词典》中整句释义的局部语法进行了研究,但是没有考虑不同词类释义模式的差异,也没有对释义的右半部分进行详细分析。《牛津高阶英语词典》中的名词释义多为短语释义,相当于整句释义的右半部分。本研究基于整句释义的局部语法研究成果,以Schmid(2000)对外壳名词的研究框架为依据,从《牛津高阶英语词典》(第8版)中抽取656个外壳名词的释义,使用“上义词”、“区别成分”、“触发语”及“详述”等功能范畴构建局部语法,探究释义模式。研究表明,外壳名词主要有四种释义模式:“名词+从句”、“名词+介词短语”、“名词+分词短语”、“名词+不定式短语”;每种模式均有不同的局部语法实现型式;局部语法型式在不同类型外壳名词释义中的分布存在差异。本研究从局部语法视角揭示外壳名词的释义模式,不仅为英语学习词典编纂的理论与实践研究提供有益参考,也将为对外汉语学习词典的编纂提供借鉴。

7-7

A Study of Phrase Frames in Argumentative Essays by Chinese English-majors

Zhao Jiamei, Chen Jiansheng, Tianjin University of Science & Technology

Based on corpus linguistics, this paper studies the characteristics of phrase frames in the argumentative essays by Chinese English-majors at two levels. The results show that Chinese English learners are more dependent on phrase frames than native speakers of English, and most of them are not used properly or accurately, namely, there exist underuse, overuse and misuse of phrase frames; only a limited range of words are used to fill in the variable slots; most of the phrase frames are topic-related; learners tend to be colloquial when writing, displaying vague stylistic awareness. All of the above characteristics are closely related with negative transfer of mother tongue and strategies of second-language learning.

7-8

**A Corpus-based Study of Explicitation of Personal Pronouns in Chinese-English
Consecutive Interpreting at Press Confer**

Chen Jiansheng, Huang Hui, Tianjin University of Science & Technology

The existing studies of personal pronouns are mostly limited to the English-Chinese written translation of literature texts from the inter-lingual perspective, and there are relatively fewer studies of personal pronouns based on Chinese-English political interpreting texts from both inter-lingual and intra-lingual perspectives. Based on the self-built Parallel Corpus of Chinese-English Consecutive Interpreting at Press Conference as well as Comparable Corpus of Press Conference Interpreted English and Press Conference Original English, this research attempts to explore the characteristics in the use of personal pronouns in Chinese-English consecutive interpreting at press conference from an inter-lingual and intra-lingual perspective respectively. The study shows: 1) Compared with the personal pronouns in press conference original Chinese, those in press conference interpreted English show a tendency of explicitation, especially the first and the third personal pronouns; 2) Compared with press conference original English, however, there is an overall tendency of implicitation in the personal pronouns in press conference interpreted English. Finally, the paper discusses the above phenomenon on the basis of linguistic difference between English and Chinese, interpersonal and textual functions of personal pronouns, and Source Language Shinning Through.

Group 8

8-1

基于语料库的译者风格研究：以《骆驼祥子》葛浩文译本与施晓菁译本为例 康宁，青岛科技大学

基于语料库的译者风格研究是语料库翻译学的重要课题之一，该研究模式由“英语翻译语料库（TEC）”创始人 Mona Baker 教授于2000年提出。此研究不同于一般意义上以源文本、原作者文体或风格为关注焦点的翻译文体研究，而是以某一位或几位译者的文学翻译文本为语料，从语言形式特征，如类符-形符比、平均句长、词汇密度等方面对语料进行数据统计，关注文学翻译译者独有的翻译语言特征，旨在发现译者在译文中表现出的规律性的、反复出现的语言行为模式。本文以《骆驼祥子》英语译文为语料，对比分析葛浩文和施晓菁两译者的译者风格。一方面基于译文考察两译者的语言模式，另一方面基于原文对比分析两译者在处理思想表达（*thought presentation*）模式和文化负载词时所用的翻译策略。本研究基于自建的小型平行语料库，采取定性与定量相结合的方法。

研究发现，两译者在语言模式选择上存在差异。葛浩文译本的平均句长高于施晓菁译本，其中一个原因在于葛浩文译本大量使用分号连接独立分句。对比分析两译者的标准类符-型符比，葛浩文译本的比值小于施晓菁译本，说明葛浩文译文表现出用词更趋简化的特征。两译者在词汇密度方面存在较小差异，但施晓菁较多地使用形容词和副词，且其使用频率与葛浩文存在显著性差异。这一点通过对比两译本关键词表可以得到印证。语法缩略形式方面，葛浩文译本显著多于施晓菁译本。在思想表达模式处理上，葛浩文译本倾向于自由直接思想模式，而施晓菁译本则没有使用这种模式；施晓菁译本相比葛浩文译本更多使用自由间接思想模式。在文化负载词翻译策略上，葛浩文倾向于选择异化策略保留汉语文化特征，尤其是在歇后语的翻译上。施晓菁相对均衡地选用了归化和异化两种翻译策略，因此施晓菁译本可读性更高。

本研究在一定程度上反映出两译者在翻译同一源文本时表现出的不同风格，揭示了基于语料库的译者风格研究的可行性。但由于语料的局限性以及所选分析角度还不够全面，对翻译文本的译者风格研究还需进一步补充和完善。

8-2

A Semiotic Interpretation of Indian Media's Emotions to China

——A Corpus-based Study

Chen Cheng, Zhejiang University

The present study conducts a semiotic interpretation of Indian media's emotions to China through a semantic discourse analysis at three levels: conceptual sense, inter-textual contextual sense, and extra-textual contextual sense. The three levels are respectively realized by three corpus-based discourse mechanisms: annotating and classifying the emotional semantic tags, retrieving and counting the collocations of the emotional signs, analyzing the social meanings. The results are at two folds. First, "violent" and "worry" were Indian media's dominant emotions towards China through ten years, which were mainly caused by three main factors "Boundary issue", "Sino-Pak relation" and "separatist terrorism". Second,

“liking” is the stably highest positive emotion, which was caused by “Wen’s effort” and “high-tech products”. The further social semantic analysis deepens the emotional senses that the violent and worrying emotions were supported by Indians’ nationalism sentiment and derived from Sino-Indian boundary disputes. The economic exchange alleviated their negative emotions and increased their liking emotions, which mainly realized by government’s economic diplomatic promotion and high-tech product export. The methodology and structure of the present study offer a quantitized pattern for emotional analysis in discourses.

8-3

基于语料库研究方法分析戏剧中感叹句的使用情况 ——以俄罗斯喜剧《聪明误》为例 张家玲，北京第二外国语学院

Corpus-based research methods, which have been used for linguistic phenomena in recent years, are objective and reasonable research methods based on big data. As an important literary theme, drama is different from other literary themes such as novels, prose, poetry and so on in the aspects of language style, narrative angle, figure image creation and plot conflict. Drama has its own distinctive features, especially its vivid language style is very unique, dramatic language has a special action, stage performance and so on. One of the features of drama language is the exclamatory phrase, which is widely used in the drama. The use of corpus-based research methods to study exclamations in drama can help us better explore the dramatic language features.

8-4

基于语料库的《红楼梦》人称指示视点翻译转移比较研究 侯羽，燕山大学

视点指“说话人站在什么角度说话”（冉永平，2007：332），而人称指示指“通过公开或隐含的人称代词把话语中涉及的人或物与说话者、听话者或第三者联系起来，表明彼此间的关系”（陈辉、陈国华，2001：175）。陈辉和陈国华（2001）对《红楼梦》原文中人称指示视点的选择问题进行了研究，本研究旨在考察杨宪益夫妇和霍克思如何针对《红楼梦》原文人称指示视点进行翻译转移。

本研究的主要问题有三个：1）两位译者针对原文进行人称指示视点翻译转移的主要类型和方式以及所采取的主要翻译形式有何差异；2）与原文视点的选择相比，两个译本视点的选择对陈辉和陈国华（2001）提出的亲疏和地位原则遵循的程度有何变化；3）他们进行人称指示视点翻译转移的主要动机是什么。

本研究方法包括设计理论框架和制定具体研究步骤。理论框架包括核心概念的界定和人称代词的理论分类。本研究将人称指示视点界定为说话者在听话者面前使用人称指示语+称谓语（如“我父亲”、“你女儿”、“他妈妈”）指称第三者的视点，使用不同的人称代词代表不同人的视点。本研究将视点翻译转移界定为译文视点偏离原文中视点，但两者关系性质不变，视域相同。本研究从理论上将人称代词分为第一人称代词，第二人称代词、第三人称代词。

本研究主要步骤包括语料的检索、确定、分类、标注、统计和分析。首先，将人称代词作为检索项，借助ParaConc软件，依次对“《红楼梦》中英文语料库”原文及其杨宪益夫妇译本和霍克思译本进行检索，获取人称代词所在的检索行及其译文。其次，删除人称代词后未出现称谓语的检索结果，并将其按照是否进行了视点转移分成两类，并对进行了视点转移的例子按照转移类型、方式和翻译形式进一步分类和统计。此外，依据陈辉和陈国华（2001）界定的亲疏关系和地位原则，分析视点转移译例对它们遵循程度的变化。最后，解释译者进行人称视点转移的原因。

本研究的意义主要有三方面：首先，它将丰富我们对霍克思和杨氏夫妇英译《红楼梦》主体性的了解，即译者们在翻译中所表现出的主观能动性。此外，它从翻译视角拓宽了人称指示视点转移研究的范围，将有助于深化我们对人称指示视点转移选择和动因的了解。最后，它是将语料库翻译学研究和语用研究相结合的一次有益探索。

8-5

A Corpus-based Study on the Representation of “The Belt and Road Initiative” in the Western Media

Shao Bin, Cai Yingying, Chen Lingwei, Zhejiang University of Finance & Economics

Since its debut in 2013, “The Belt and Road Initiative” (hereafter BRI) has drawn great attention at home and abroad. Nevertheless, there have been few studies pertaining to the image of BRI in the eyes of western media. Based on the English news collected from LexisNexis news database, this paper builds the “The Belt and Road Initiative Corpus” (hereafter BRIC) and attempts to investigate how BRI is represented in the western media. Concordance analysis is used to reveal the overall attitudes of western countries towards BRI and cluster analysis scrutinizes each individual attitude towards BRI and clusters the similar attitudes in the same category. In addition, keyword analysis explores different focuses on BRI of British, American, German, Australian, and New Zealand’s media. It finds that 1) while a few western media intentionally misinterpret it and deem it as a manifestation of China’s ambition to dominate the world, most of them hold positive attitudes towards BRI; 2) media of America and countries of the British Commonwealth, such as Australia and Canada, sometimes have negative view towards BRI. On the contrary, those countries along the route, especially central and eastern European countries, are inclined to have positive view; 3) western countries vary in their focus on BRI considering their national and economic interest and ideology.

8-6

Exploring Language and Literature with *The Prime Machine*

Stephen Jeaco, Xi’an Jiaotong-Liverpool University

Corpus approaches provide insights into patterns of language use and they provide evidence of marked and unmarked usage, or of norms and deviations from those norms (Hanks, 2013; Hoey, 2005). Examples from corpora can help language learners in many ways, helping them notice common patterns associated with specific words and collocations, compare and choose between pairs of words with similar meanings and functions, and gain

insights into patterns of use across different text types (Thomas, 2015; Tsui, 2004). For the analysis of literary language, corpus tools can be used in at least two ways: drawing on corpora as a representative sample of a specific language variety, or focusing on a corpus containing all the texts of a specific author or group (Mahlberg, 2013). Corpus tools as a language reference help users find concrete evidence of how a specific use of language outside the corpus is similar or different to uses in other texts. If a literary text or collection of interest is loaded into the software, corpus tools can also be used to find and explore all instances of a specified language feature and corpus-driven methods can be adopted to identify features to explore in the text.

This paper presents a software demonstration of *The Prime Machine V3*, introducing some of the ways it supports the exploration of English words and phrases for the study of English language and English literature. Building on earlier versions of this tool, *The Prime Machine* provides user-friendly access to pre-processed corpora and has search screen support and results pages designed to help students notice and explore tendencies of words and phrases to occur in different kinds of environment (Jeaco, 2017). Software features will be introduced with pedagogical rationales, specifically considering the needs of Chinese university English language learners, and Chinese English Majors in particular.

The software hosted on a server in Suzhou, China and is available from www.theprimemachine.net.

8-7

当代中国大众读者对精英文学《浮士德》的接受研究 ——以2005-2018年《浮士德》豆瓣评论为例 宣瑾，华中科技大学

该研究以豆瓣平台上共计34个《浮士德》译本的短评和书评为考察对象，依托读者反映批评理论，运用LDA主题模型、社会网络分析和情感分析方法，探究豆瓣成立以来其用户对《浮士德》的接受状况。研究表明：其一，2005-2018年，豆瓣用户对《浮士德》的讨论热度呈上升趋势，情感分析结果显示，历年积极情感占比均在50%左右，正面评价为讨论中的主流观点；其二，《浮士德》的读者讨论网络呈现出一种高评价与低理解并存的矛盾。尽管对《浮士德》的溢美之词俯拾即是，但“难以理解”这一话题却处于核心地位，由于审美趣味和翻译等原因的存在，读者未能对作品内涵和译本话题发表更多见解，《浮士德》与大众读者之间仍然存在难以逾越的鸿沟；其三，豆瓣这一媒介为大众对《浮士德》的阅读反馈提供了平台，“沉默的大多数”的发声在一定程度上消解了《浮士德》作为精英文学的“光晕”，同时挑战了少数读者对于文学阐释的话语权，促成了读者审美新面貌的形成。

8-8

基于语料库的模糊限制语使用的性别差异研究
张瑞丹, 上海外国语大学

本研究以上海交通大学JDEST学术英语语料库为语料源, 以E. F. Prince、J. Frader和C. Bosk对模糊限制语的分类方法为标准, 即变动型模糊限制语(程度模糊限制语和范围模糊限制语)和缓和型模糊限制语(直接缓和语和间接缓和语), 采用定量与定性相结合的方法, 探究男女性在61个较为常用的四类模糊限制语使用上的差异, 并试图对这些差异作出解释。该研究是对模糊限制语研究的拓展, 对英语教学也有一定的启示意义。

Group 9

9-1

A Multifactorial Analysis of the Choice of Necessity Modals *Must* and *Have to* in Chinese Learner English

Deng Yaochen, Dalian University of Foreign Languages

In this study, we apply the MuPDAR approach to the comparison of alternating behaviors of two English necessity modals, *must* and *have to*, in native English and Chinese learner English, hoping to reveal the different patterns of choice as well as the grammatical contexts which constrain these two groups of speakers in the choice of this pair of modals of obligations. More than 2700 examples from the comparable corpora of LOCNESS and CLEC were extracted and annotated manually with more than 20 morphosyntactic and semantic features. A hierarchical cluster analysis and two logistic regression analyses provide inspiring results which are beyond previous SLA research and show exactly where learners' modal use still deviates from that of native speakers.

9-2

A Corpus-based Study of Epistemic Modality in Chinese University Students' English Writing

Zhang Zihan, Li Ye, Huang Yi, Qilu Normal University

Academic writing is full of epistemic modality expressions, which allows writers to show their "evaluation on possibilities", or "the degree of confidence" (Coates, 1983). The appropriate use of epistemic devices (EDs) is crucial to successful writings, for it not only helps writers express their certainty and doubt exactly, but contributes to their relationship with readers (Oh, 2007; Hyland & Milton, 1997). Ultimately, it leads learners to present statements in "acceptable and persuasive ways" (Hu & Li, 2015).

Despite its importance, epistemic modality is considered one of the most difficult and complex items for EFL learners to master (Palmer, 1979; Cheng & Qiu, 2007). This study investigates how epistemic modality is expressed in Chinese college students' English writings by comparing the essays from Chinese CLEC and native speakers' LOCNESS corpora. We assume that Chinese university students may overuse or underuse some epistemic devices compared with British and American university students. Moreover, the high-proficiency NNS group better approximates the NS group than the low-proficiency NNS group. Explanations from the perspectives of pedagogical reasons, cultural values and language transfer are discussed afterwards, which help to reveal some pedagogical implications.

9-3

翻译与中国自我形象建构
——基于语料库的《毛泽东选集》英译研究

石欣玉, 西安外国语大学

本文从形象学视角着手翻译研究, 探讨翻译与中国自我形象建构问题。本文将《毛泽东选集》(1-4卷, 1961-1965) 外文社版作为考察对象, 使用语料库研究方法, 从关键词类切入, 选取关键词, 借助对关键词的搭配特征研究, 探讨《毛选》外文社版译本所建构出的中国自我形象; 此外, 本文通过分析影响翻译活动的外部因素, 阐释译本如此建构中国自我形象的背后动因。

9-4

多因素分析的标注瓶颈与分离式标注

吴良平, 湖南商学院/北京外国语大学

语料库语言学中的多因素分析考察影响某一语言现象的所有潜在因素, 开展共时分析, 进而确定各个因素的权重及交互效应 (Gries 2003; 房印杰 2016), 以期对该语言现象进行全面解释。由于其统计方法上的严谨, 受到越来越多的语料库研究者的关注和重视, 在词汇、句法、语义诸多语言层面开展了大量研究, 并被广泛应用于二语习得、认知语言学等领域。然而, 由于这一方法常常涉及上十甚至数十变量的人工标注, 给研究者带来了前所未有的挑战。本文认为, 这一问题的解决方案在于采用分离式标注 (stand-off annotation)。通过运用MMAX2对影响中国英语学习者直宾型that从句引导词显隐的诸多因素进行分离式标注, 本文讨论该标注方法的优势, 并提出改进意见, 以期推动多因素分析方法的应用与发展。

9-5

基于语料库的彝语诺苏话话题标记研究

邹雨橙, 胡素华, 中央民族大学

彝语中有两个单音节的话题标记li33和nu33, 二者的功能有区别, 但极其细微, 不易捉摸。以往的研究从自然口语和民间故事的考察中发现, li33是句子话题标记, nu33是篇章话题标记, 并且二者与信息的新旧程度及指称关系相关 (胡素华, 赵镜, 2015)。本文旨在通过现有的彝语语料库 (彝语基础例句语料库、彝语民间故事语料库) 的量化统计, 进一步证实和精确化话题标记li33和nu33在篇章结构和句子结构层面上的分布和功能。通过考察li33和nu33标注的短语在句子和段落中的位置 (首、中、末)、所标注的短语属性 (名词性、动词性、形容词) 及其谓语属性 (属性谓语、事件谓语), 从而理清li33和nu33各自的功能, 以期了解彝语的信息结构及其实现的手段, 同时也探讨语料库语言学在实证研究中的作用。

9-6

**A Corpus-based Comparative Study on Cohesive Devices in Geological English
Abstracts of Chinese and International Journals**
Lei Xiaodan, China University of Geosciences

English abstracts are of great importance in new knowledge dissemination and scholarly communication in a disciplinary community. A well-written abstract will surely promote the text attached to it more effectively. To compare the similarities and differences of Chinese and English writers in the use of cohesive devices of English abstracts, two corpora are built, which are *Chinese Journal of Engineering and Geology* which is the school journal of CUG (China University of Geosciences) and *The Bulletin of Engineering Geology and Environment* which is the authoritative journal of IAEG (the International Association of Engineering Geology). By examining the concordance, concordance plot, cluster, collocates and word list of cohesive devices in the two different journal corpora, this study finds the similarities and differences in the cohesive devices of English abstracts and analyzes the reasons to shed light on the degree to which students are correct in the use of cohesive devices of English abstracts in geological research articles and to provide relevant materials of English for Specific Purpose (ESP) for both English teachers and geological writers.

9-7

基于虚词词典与语料库的“的”字短语释义研究
孙婉, 张延成, 武汉大学

本文通过全面考察目前已有的虚词词典中“的”字的释义, 使用台湾中研院平行语料库进行“的”字短语各种类使用频率抽样统计, 结合目前学界对“的”字短语的研究现状, 根据句法、语义对“的”字短语进行分类细化, 探究对外汉语虚词词典的编纂的释义问题, 认为:

现有虚词词典中, 对于“的”用法的分类往往较为简单, 尤其是面向对外汉语教学的虚词词典。“的”字短语使用情况复杂, 编写对外汉语虚词词典和教材时, 应该考虑习得的难易程度, 进行相应的分级。

作定语标志的“的”字短语(例如: 红的花、开车的技术)与其他语言中存在的定语小句较为相似, 习得难度较低。根据台湾中研院平行语料库抽样显示, “的”作定语标志的情况最多, 可以考虑让具有初级汉语水平的学习者掌握。

独立“的”字短语(例如: 父亲是镇上卖猪肉的。), 虽然具有名词性成分的句法地位, 却并不表示固定事物, 因而也没有固定的所指, 其所指取决于上下文或者具体的语境。这种类型的“的”字短语习得难度较高, 在外向型词典释义处理中应有标明。论文还专门探讨伪定语、事态助词前置等特殊“的”字的释义和句法例释问题。

9-8

语料库语言学在古汉语研究中的应用
——以“最”、“至”、“极”的语体正式度差异为例
钱珍, 广东财经大学

词义的理解,除了参考词典上注释出的义项之外,必须在一定的语境中才能明确意思。以至于很多语料库语言学者认为,孤立的词具有多义性,其词义一般是模糊而不确定的(Sinclair, J. 1996)。排除多义必须向上下文中探求。因此,意义单位(a unit of meaning)不仅是关键词(keyword),还应当加上那些使关键词的意义确定下来成分,即与其搭配使用的那些词(Teubert & Čermáková 2007/2009:17-22)。

本文认为,这样的思路对古汉语词义研究非常有帮助。众所周知,古汉语研究很难像现代汉语那样,藉助语感进行。而利用语料库的检索技术,以及语料库语言学的相关分析,可以在一定程度上降低这种局限性,尤其是在词义判定以及同义词的辨析方面。本文即是对此思路的一个实践。利用台湾中央研究院的“上古汉语标记语料库”及作者自建的《新五代史》语料库,本文分析了上古到中古时期,三个同义副词“最”、“至”、“极”的语体正式度差异,并在此基础上,探讨古汉语分语体语料库的构建设想。

Group 10

10-1

基于RNN的主谓一致错误自动识别研究

杨林伟, 北京外国语大学

近十年来, 二语学习者自动语法检查研究备受计算语言学界关注。其方法历经从基于词性赋码、依存关系等语法知识的规则匹配, 到基于N元概率统计的语言模型, 再到基于学习者语料库的统计建模与机器学习三个阶段。基于语法特征的传统机器学习与基于“平行”语料库的机器翻译法已成为自动语法检查领域的主流研究方法。随着神经网络与深度学习的出现, 语法自动检查研究在方法上得到进一步发展 (Wang et al. 2017)。但是, 深度学习在自动语法检查研究中的应用刚刚起步, 能否成为主流研究方法需要更多的深入研究加以验证。

本研究聚焦中国学习者英语动词语法错误中的主谓一致错误, 基于BNC语料库与iwrite中国学习者错误标注语料库, 尝试采用深度学习的方法, 利用循环神经网络(RNN)构建中国英语学习者英语主谓一致错误自动识别模型。实验结果显示, 基于BNC语料库训练的RNN模型对主谓一致错误的识别准确率优于传统模型, 值得深入研究。

10-2

基于构式语法对构式“什么A(什么)B”的认知解读

——以《红楼梦》文本为例

潘鸿, 武汉理工大学

构式一词由来已久, 语言学中研究的任何内容几乎都可以称作为“构式”。构式“什么A什么B”这种具体的表达形式, 也可以被看作是习语, 广泛的被运用在日常生活和文学作品中。本文结合北京语言大学语料库中心(BCC)的数据, 对该构式做一个简单的数据统计, 探讨该构式的意义以及其变体。利用红楼梦汉英平行语料库将“什么A什么B”构式全部检索出来, 自建一个“什么A什么B”的小型语料库。通过语料标记, 借助HyConc语料处理软件提取结果进行分析。本研究包括三部分: 一、分析“什么A什么B”构式的几种情况。二、探讨该构式是如何形成的。三、研究“什么A什么B”构式的语义韵及其翻译。

10-3

中国学习者英语语法形位习得顺序

——来自语料库和实验的证据

刘洋, 北京外国语大学

英语中的形位长期以来一直是非母语者习得的难点之一。Krashen (1977) 提出“自然顺序”习得假说, 声称无论学习者何种母语背景, 其形位习得遵循同样的顺序。本研究采用EFCAMDAT语料库的中国学习者部分, 试图探究各水平中国学习者的形位习得顺序如何, 是否与“自然顺序”一致? 同时, 使用SLOTE Test对120名不同水平学习者进行实验, 与语料库数据结合, 互相佐证。此外, 本研究还将中国学习者语料与来自母

语背景为德语、法语和日语的学习者数据进行比较,探究中国学习者在形位习得中是否受到母语的影响。这对我们加深语言发展机制的了解、“自然习得”顺序的验证和我国英语教材的编写和教学都具有重要价值。

10-4

中西学者学术论文中作者声音构建的对比研究:以“名词+that”结构为例

杨晨姣,北京航空航天大学

“作者声音”这一概念近来在写作研究中被屡屡提及。但是,一直以来,学术文本由于其特殊性,不易看出文章作者真实声音。基于Sinclair“自引(averral)”和“他引(attribution)”两个重要概念,本研究将通过“名词+that”结构,探讨中西方学者学术论文中作者声音构建的异同以及作者如何使用该结构与读者进行人际互动。研究选取了BEIJING CARE学术论文可比语料库中西方学者和中国学者材料学子库,共计391篇论文,旨在找出与that从句搭配的名词形式以及这些搭配如何表达不同类型、不同程度的作者声音。研究发现,两个作者群体“名词+that”结构使用的频次差异显著,西方学者共527条远高于中国学者的104条,但两方学者均通过使用“the observation that”以及“the evidence that”等固定搭配表现“自引”,突显自我声音、表达立场,倾向使用“the problem that”等固定搭配表现“他引”,表达对论点的质疑。然而,中国学者几乎不使用西方学者较多使用的“there be+that”搭配表达“隐性自引”。理解西方学者如何通过使用“名词+that”结构展示“真实的”作者声音,对学术英语教学及中国学者国际学术交流效益的提升有一定的促进作用。

10-5

局部语法研究源流与发展述评

袁新华,北京外国语大学

本文在梳理局部语法理论发展脉络的基础上,综述了国内外研究的现状及进展情况,着重阐述局部语法的核心观点、基本思路和研究范式,结合研究实例阐释了局部语法已经成为语料库语言学发展的新趋势。同时,本文也指出了目前该领域存在的问题与不足。针对目前国内的研究现状,初步展望局部语法在我国未来可能的发展方向 and 主要研究内容。本文认为局部语法的分析框架已经成为语料库语言学发展的新趋势。

10-6

Developments in Australian English Grammar from 1931 to 2006: An Aggregate Approach to Dialectal Variation and Change

Yao Xinyue, Peter Collins, Renmin University of China, University of New South Wales

Australian English (AusE) has never ceased to be a topic of scholarly and popular interest. Commonly defined as the English used by (non-aboriginal) native-born Australians, it serves as the dominant form of, and the national linguistic standard for communication in Australia. With a short history of no more than 250 years, AusE has undergone substantial changes over time. This study is the first to investigate overall developments of AusE grammar in the 20th century, making comparisons with British and American English (BrE and AmE) of the same

period. AusE data was drawn from balanced diachronic corpora designed after the Brown family – which we refer to as “AusBrown”. The corpus was sampled for four “time points”, the 1930s, 1960s, 1990s and 2006, spanning a total of roughly 80 years. Around 180,000 words of running text, equally divided among three written registers, were chosen for one time period. On the methodological side, this study adopts an aggregate approach to grammatical variability. It investigates changes in the text frequencies of 69 (lexico-)grammatical features which have been shown by previous studies to be sensitive to variation across regions and over time. With exploratory multivariate techniques including hierarchical cluster analysis and multidimensional scaling, we provide a more realistic and reliable picture of grammatical changes in written AusE in the past century and consider them in relation to those in BrE and AmE. We also explore potential areas of grammar where AusE significantly diverges from the two varieties, and seek to offer linguistic and socio-historical explanations for patterns identified.

10-7

对英语迂回式使役结构中使役动词的多维度研究

张蒙蒙 北京外国语大学

使役结构是自然语言中的一种常见结构。跨语研究发现使役结构可分为词汇使役结构 (lexical causatives) 和迂回式使役结构 (periphrastic causatives)。本研究主要对英语迂回式使役结构中的使役动词 (make, cause, have, get, let) 在不同英语变体和语体中的用法及变化进行近半个世纪的历时考察, 结果发现: 1) 迂回式使役动词在英国英语中的使用呈先下降后上升姿态, 而在美国英语中的使用则呈不断上升姿态; 2) 不同的迂回式使役动词在不同语体中分布不同, 这种不同的分布在不同的时期不同的变体有着惊人的相似性。3) 迂回式使役动词在小说语体中的使用频率远远高于其他语体, 并具有显著性差异。

10-8

Properties of Verb and the Present Perfect Construction

Zhang Liying, Information Engineering University

The present study investigates the three uses of the present perfect and the properties of verbs in the present perfect. The data consists of the fiction and newspaper sections of the British National Corpus. Collostructional analysis is employed to select verbs strongly attracted to and strongly repelled by the present perfect. The selected verbs are classified into six categories of achievement, activity, accomplishment, stage-level state, individual-level state and semelfactive. Results indicate that the Vendler or aspectual classification of verbs cannot explain why the present perfect is selective of some verbs not others. Further analysis of verbs strongly attracted to and repelled by the present perfect shows that the present perfect is sensitive to the manner/result distinction of dynamic verb, the stage/individual distinction of state verb and the causative/noncausative distinction of mental verb. The present perfect is a type-selecting construction and asks for verbs denoting result of a process. Analysis of the three uses of the present perfect shows that the three uses of the present perfect show distinct

patterns. The resultative perfect is predominantly used while the experiential and continuative perfect are marginally used. The three uses of the present perfect share the meaning of a time span.

10-9

基于义项标注法的构式搭配分析新路径：以英语被动构式为例

陈孚，张雪梅，上海师范大学，上海外国语大学

构式搭配分析现已形成了一组相互联系的方法体系，包括共现词素分析(Collexeme Analysis)、显著共现词素分析(Distinctive Collexeme Analysis)以及共变词素分析(Covarying Collexeme Analysis)。传统的构式搭配分析仅以构式某槽位上的词型(lexeme)为计算单位。然而，面对一词多义现象，这一做法无法揭示该多义词型的哪个(些)义项与构式的搭配关系最为紧密，因而该方法的语义区分度较低。本研究以英语被动构式为测试对象，对近年来出现的基于义项标注法(sense-based)的构式搭配分析的这一新路径进行了有效性检验。结果发现，一、传统方法某些未达到显著性的词型的某个(些)义项得以凸显，并成为了显著共现词素；二、经过义项标注后得到的显著共现词素在语义区分度上得到了明显提升，语义信息量加大，便于研究者做出更加清晰的判断和概括；三、义项标注法不仅有助于提升基于Fisher精确检验的双向性(bidirectional)搭配强度计算的准确性，同时也增加了单向性(unidirectional)构式搭配计算指标(如 ΔP)的可解释性。

Group 11

11-1

‘This means...’: A Contrastive Analysis of Authorial Practice of Attended/Unattended *this* in Physics and Linguistics Research Articles

Gao Xia, Beihang University

This paper aims at quantitatively and qualitatively analyzing the use and distribution of demonstrative *this* in a comparable corpus of research articles written by L1 Chinese scholars (CSs) and L1 English scholars (ESs) in Physics and Linguistics. The analysis of 8,004 instances of *this* obtained demonstrates that: (i) ESs employ both attended and unattended *this* significantly more frequently than CSs; (ii) native physicists are found to employ substantially more (un)attended *this* than native linguists; (iii) both ESs and CSs prefer to use *this* sentence initially; (iv) a group of verbs has been identified to be distinctive for unattended *this*; and (v) writers’ choice of (un)attended *this* hinges in part on its antecedent, the distance between *this* and its antecedent, and the verbs co-occurring with *this*. These findings may hopefully shed lights on EAP teaching and cross-cultural understanding among academics.

11-2

A Comparative Analysis on Stance Markers in Research Article Abstracts of Theoretical Linguistics and Quantitative Linguistics: A Corpus-based Study

Li Jinfeng, Xiao Wei, Chongqing University

Academic writing serves for persuading the peer researchers, and stance markers are an essential rhetoric means to express author’s opinions, attitudes and commitments for achieving author’s positions and interaction with readers. The frequently used stance markers are hedges, boosters, attitude markers and self-mentions. The previous studies have found that different disciplines often display different patterns of use of stance markers, but the differences between sub-disciplines are yet unclear. To this end, the present study aims to investigate the differences of distribution of stance markers in theoretical linguistics and quantitative linguistics, two substantially different sub-disciplines of linguistics. This study build up two corpora, each comprises 40 research article abstracts chosen from authoritative journals like *Language*, *Linguistics* and *Quantitative Linguistics* during 2010-2018, and double coding is used to improve accuracy. The findings indicate that researchers in theoretical linguistics and quantitative linguistics use stance markers frequently in order to interacting with readers. Both sub-disciplines use stance markers near 60 times per 1000 word. Despite similarity in general, they differ in the distribution of four types of stance markers ($\chi^2 = 13.528$, $p = 0.04$). Theoretical linguists use attitude markers (27.36%) more often than quantitative linguists do (20.10%). Besides, theoretical linguists prefer hedges with a proportion of 22.28% compared with quantitative linguists of 17.84%. However, self-mentions in quantitative linguistics occupy a larger percentage with 41.35% than that in theoretical linguistics (30.75%). As for hedges, they account for a similar percentage near 20% in both sub-disciplines. These results can be accounted for by characteristics and differences of the two sub-disciplines.

11-3

A Comparative Analysis on Reporting Verbs in Research Article Introductions across Disciplines

Yu Xuan, Xiao Wei, Chongqing University

Reporting verbs are a rhetorical device to express authors' opinions. They are generally classified into three categories: textual, mental and research ones. In the past decades, numerous studies related to reporting verbs have been done by scholars abroad and at home. However, little attention has been paid to the differences across disciplines, especially across the humanities, social sciences and natural sciences. To this end, this research article is designed to thoroughly depict reporting verbs in research article introductions across the humanities, social science, and natural science. In this paper, 60 research articles (about 60000 words) from three disciplines (mechanism, economics and philosophy) have been collected randomly from SpringerLink to build up three corpora for this research. The findings show that (1)Mechanism discipline obtains the largest percentage of reporting verbs compared with its total word tokens, which accounts for 2%, followed by economics and philosophy, whose reporting verbs account for 1.85% and 1.32% respectively. (2)In mechanism discipline, research reporting verbs account for the largest proportion; in economics, textual and research reporting verbs obtain similar proportion, far surpassing mental reporting verbs; in philosophy, textual reporting verbs obtain the largest percentage, far surpassing mental and research reporting verbs. (3)Reporting verbs in economics and philosophy discipline show a similar pattern in that the two disciplines use a wider range of reporting verbs compared to mechanism. These findings can be accounted for by discipline differences.

11-4

Contrasting Corpus Analysis Data to a Proficiency Scale: The Case of Aviation English

Malila Carvalho de Almeida Prado, University of São Paulo

ICAO (International Civil Aviation Organization) has required governments throughout the world to certify their pilots' and air traffic controllers' English proficiency to operate internationally. With this aim, ICAO has established a proficiency scale as guidelines to the testing of aviation English. The construct to be assessed is the scope that corresponds to the spontaneous language to be used on the radio in case Aeronautical Phraseology, a scripted and documented language, does not suffice. In an attempt to understand how this language is used in real contexts, we have compiled a corpus of radiocommunications between the abovementioned professionals in unusual circumstances. By means of Corpus Linguistics tools, we have extracted important pragmatic features from the corpus, especially in the analysis of fluencemes (Götz, 2013), turn-initiators (Tao, 2003) and deixis (Rühlemann & O'Donnell, 2015). We intend to contrast the pragmatic elements highlighted from the corpus to the proficiency scale proposed by ICAO, which focuses more on linguistic competence. Such an investigation may provide a better comprehension of the proficiency scale – and to what extent it suits the professionals' needs – as well as assist the teaching and testing of aviation English.

11-5

A Developmental Study of Prepositional Phrases in English L1 Students' Academic Writing

Tao Guoxiao, Pan Fan, Fang Alex Chengyu, Huazhong University of Science and Technology, City University of Hong Kong,

Modern academic writing is said to drift towards a structurally compressed discourse style by using various compressed phrasal devices (Biber et al., 2011, 2016). Following this trend, most studies have investigated the use of these compressed devices in academic writing. Among them, however, only a few have paid attention to prepositional phrases (PPs), with even fewer to PPs in English L1 students' writing, which is hypothesized to show a trend of development over school years (2011). Based on BAWE (i.e. British Academic Written English), this study will examine the use of PPs in English L1 students' writing across four school years to testify the hypothesized developmental sequences. More specifically, the internal structures, syntactic functions and semantics of PPs will be explored respectively by using various corpus analysis tools. Findings are expected to be: 1) the internal structure of PPs will become more complex, with the use of more various types of prepositions and prepositional complements; 2) PPs will play a variety of different structural roles, changing from adverbials to nominal post-modifiers; 3) PPs will be used to reveal more semantic information, extending from concrete to abstract meaning. In a word, this study will be of great significance for the English for Academic Purposes instruction in the EFL environments by revealing the trend of use of one of the phrasal features in English native students' writing.

11-6

中国高校英语专业本科毕业论文结语研究

张金福, 九江学院

结语(结论)通常是学术(学位)论文正文的最后部分,是作者总结研究的精华所在,是读者理解全文的关键部分,因此其重要性不容质疑。然而问卷调查和访谈发现,大多数英语专业本科毕业生并不清楚毕业论文结语的基本功能、语言特点和体裁特征。文献检索显示,国内外有关学术论文结语的研究并不多,有关学位论文的更少,且大多数是关于硕、博论文结语研究。本研究基于两个自建语料库(《中国高校英语专业本科毕业论文结语语料库》和《英语学术论文结语语料库》)对我国英语专业本科毕业论文结语从基本概貌、语言特点和体裁特征等方面进行检索分析,发现其结语确实存在抄袭行为,但占比并不高;用词偏口语化,句子偏短,在词汇丰富度等方面与学者学术论文有较明显差异;从体裁特征来看,二要素较多,且主要是研究背景概括和研究结果陈述,大多忽略研究不足与展望等语轮。

11-7

Research on Modal Sequences in L2 Academic Writing

Yan Pengfei, Beijing Foreign Studies University

This study explores the features of modal sequences in Chinese science and engineering learners' L2 academic writing through corpus-based contrastive interlanguage analysis. Results show that learners tend to overuse deontic modal sequences and underuse epistemic ones. Specifically, modal verbs are overly preceded by personal pronouns and followed by dynamic verbs which lack marked aspect and voice, and epistemic modal adverbs for modification. In addition, modal verbs are seldom used in attributive clauses and *THERE BE* sentence pattern to interpret or evaluate preceding propositions. Above tendencies would largely result in a certain degree of monotone and subjectivity in modal patterns and meaning, and relative lack of logic and precision in learners' L2 academic writing.

11-8

A Corpus Driven Analysis of Representations Around the Term ‘THAAD’ in the Chinese, South Korean and American Press 2017

Zhao Lishu, Changchun Normal University

This study merges corpus linguistics and critical discourse analysis to explore patterns of representation around the term “THAAD” in Chinese, South Korean and American Newspapers. The self-built corpus involves news articles of the THAAD Event in China Daily, The Korea Times and New York Times from January 2017 to December 2017. With the tool Sketch Engine, an analysis of noun collocates of THAAD was conducted, and several categories were found to be of particularly lexical richness in the three medias. Following this, an analysis to explore the characteristics of agenda setting and discursive constructions of the national images of China, South Korea and USA was also made. Finally, this study took the legitimation strategies that enabled editors to print more controversial representations into consideration and attempted to reveal the different ideologies behind those reports.

11-9

Identifying the Linguistic Features of Research Paper Titles in Biomedicine: A Corpus-Based Study

Cao Jing, Fang Alex C., Zhongnan University of Economics and Law, City University of Hong Kong

The importance of research paper (RP) titles has been widely acknowledged (e.g. Goodman et al. 2001; Haggan 2004; Soler 2007 & 2011; Aleixandre-Benavent et al. 2014; Méndez & Alcaraz 2015; Chen 2016), but title writing has received disproportionately little attention when compared with other RP components such as the abstract. Previous studies have often focused on revealing disciplinary variations between different subject domains. Anthony (2001) has suggested that the sub-disciplines of individual fields could also show further variation, which is increasingly true today when inter-disciplinary approaches are

applied. As a branch of medical science, biomedicine has attracted paramount attention in recent years and resulted in a vast number of publications. Though the RP titles in this area have been studied, data size has been rather limited. The current study intends to be a remedy of this shortfall and is based on a large corpus of 25,000 titles randomly sampled from PubMed between 2010-2014 under the topics of drug side effects and adverse reactions. In addition, the corpus has been grammatically annotated through the use of AUTASYS (Fang 1996), whose annotation scheme provides not only rich POS information but also specific subcategorization features revealing phrasal patterns. This study will investigate the titles from some of the conventional perspectives such as title lengths and the use of punctuation. More significantly, this study will focus on the grammatical aspects, including the syntactic features according to phrase structures as well as the grammatical composition of the titles according to different parts of speech. The paper will conclude with discussions on potential pedagogical applications as suggested by the empirical findings.

Group 12

12-1

Construction of Authorial Identity in HK MA Students' Academic Writing: Phraseological Sequences of SUGGEST as a Case

Zhao Xiaolin, Donghua University

Academic writing is not only to convey factual and propositional information, but to construct the authorial identity of being a credible insider. Based on an applied linguistics MA theses corpus data, this paper investigates three types of phraseological sequences of SUGGEST, i.e. SUGGEST with human subjects, SUGGEST with abstract subjects and SUGGEST with the concealed subject, aiming to examine their typical uses in HK MA students' academic writing. The investigation demonstrates that the first type of the sequences is eight times as much as the second, and 13 folds of the third. Among the first type, "authors' names + SUGGEST (*that*)", "I + SUGGEST (*that*)" and "I + Modal *suggest* (*that*)" are the most common means. Surprisingly, *we* is not found. It is revealed that HK students likely express authorial identity in an explicit way when making suggestion, and meanwhile construct a responsible, a credible and an approachable insider within applied linguistics studies.

12-2

A Corpus-based Study of Defamation Law in Mainland China, Hong Kong and UK: A Gricean Perspective

Liu Fan, Zhejiang University

Defamation law has been the focusing point of plenty of research. Previous studies center on interpreting intentions of defamation law as well as emphasizes the balance between freedom of speech and protection of reputation. By three self-built corpora of court judgments in Mainland China, Hong Kong and UK, and using *Antconc 3.5.6* as the analytical tool, this study is designed to compare the law of defamation among the three jurisdictions so as to figure out their distinctive priorities as well as the experiences Mainland China can learn from UK and Hong Kong. As we figure out from the corpus analysis, among the TOP 50 most frequent content words, we can find that the two most frequent content words occurring in defamation cases are defense and meaning except for those words related to procedural law: *plaintiff(s); defendant(s); court; case; law; evidence; judge; judgment; trial; action(s); claim; article; order; appeal; issue; matter; question*. Therefore, the study will probe into different types of defenses in these three jurisdictions and the approach they take to deal with meaning. Moreover, linguistic inspection provide insight to facilitate the trial of defamation cases. This study adopts Grice's meaning theory to distinguish natural and non-natural meaning since which is a crucial element in defamation cases. Through the comparative analysis and the introduction of Grice's meaning theory, this study is expected to push forward the improvement of defamation law in Mainland China and help juries discern defamatory language as well as explore the intended meaning of the utterer.

12-3

Discipline-Specific Stance Construing in Academic Discourse

Dong Jihua, Chigan Menghua, Li Lixia, Northwest A&F University

Stance expression plays an essential role in the effective communications between writers and readers in academic discourse. The present study is based on two self-built corpora, the first one consists of the papers published in SCI journals in the discipline of engineering; the other is composed of the course papers written by Chinese EFL postgraduates majoring in engineering. The stance expressions were concordanced and analyzed in terms of variety, frequency, and left- and right-hand context. The study found that Chinese students tended to use attitude stance more frequently, while the professional authors were more likely to use more hedges and reference stance expressions. The findings indicate that the students were lack of the awareness and strategies in communicating and interacting with readers; students seemed to hold more affirmative tone and have a higher affective presence than the professional writers; they were also found to lack the strategies of referring to other papers and organizing the discourse. The above differences were explained by the different readers targeted, students' lack of exposure to a range of stance resources and their insufficient knowledge about the academic writing. The findings could contribute to our understanding of the use of stance expressions in the students writing and that of the professional writers. Pedagogically, the finding could provide guidance for the practitioners in academic writing from the perspective of book compiling, curriculum designing and the class teaching, etc.

12-4

基于语料库的中外学者学术英语词块使用对比研究

王蜜蜜, 河南师范大学

词块作为语篇构建的基本单位, 在学术英语写作中承担着重要的语篇功能。学术语篇中的词块使用已引起语料库语言学和二语习得领域的关注, 以往研究已明确词块使用存在学科差异, 但是国内针对理科英语词块的研究较少。因此本文基于课题组自建语料库, 以国内理科期刊英语论文语料库为目标语料库, 国际理科期刊论文语料库为参照语料库, 以四词词块为研究对象, 从结构和功能两个角度探析中外学者在期刊论文中使用词块的异同。结果表明: 频数上, 中国学者词块使用的数量和种类均显著低于国际学者; 结构上, 中国学者过多使用动词词块, 而过少使用名词词块和It结构, 在介词词块的使用上无显著性差异; 语篇功能上, 中国学者过少使用指示词块和组篇词块, 在立场词块使用上无显著性差异。这表明中国学者在使用词块时仍存在困难, 需要提高词块使用意识, 和恰当使用词块来实现语篇功能的能力。本研究对学术英语写作和教学具有借鉴意义。

12-5

中外学者学术写作中立场形容词对比研究

王在雅, 河南师范大学

作为表达个人情感、态度和价值判断的方式,立场表达是一个极为重要的研究课题,它包含丰富的表达形式,其中立场形容词是重要形式之一。国内研究大多将立场形容词贯穿于立场标记语之中,专门针对立场形容词的研究相对较少。本研究基于Biber et al.(1999)的分类框架,利用课题组自建语料库——中国理科学者期刊论文和国际理科学者期刊论文语料库,对比分析中外学者学术写作中立场形容词的使用特征。结果表明:语义类别上,中外学者的立场形容词类别无显著性差异,立场形容词集中出现在认知立场和态度立场两大语义类中,但中国学者使用立场形容词的总频数不足国际学者四分之一,且词汇丰富度不及国际学者;语法标记型式上,中外学者不存在显著性差异,均多用隐性手段表达立场,且较多使用“形容词+不定式补足语”结构。这表明中国学者基本上能恰当使用立场形容词,在今后的写作中需强立场形容词的使用意识,丰富词汇多样性。本研究对理科学者学术英语写作和教学具有借鉴意义。

12-6

农业学术语篇中主语it外置结构的形式与功能

——基于动物科学及园艺科学语料库的研究

王闰利, 李丽霞, 西安交通大学, 西北农林科技大学

本研究基于农业学术文本语料库,旨在分析主语it外置结构的形式特征和功能特征。具体研究问题如下:1)在结构上,主语it外置结构呈现哪些特点?2)在功能上,它们实现了哪些人际功能?体现了农业学术语篇的哪些特点?3)主语it外置结构人际功能的实现涉及哪些语言特征的共选?研究发现:1)在结构上,主语it外置结构主要由形容词主导,且that-从句出现频数比to-不定式高;2)在功能上,该外置结构主要表达情态意义,表明农业学术文本侧重对事实的阐述与说明、更依赖实证研究;3)该结构中存在着词汇与语法层面的共选,语言的形式和意义不可分割性及相互关联的统一的特点。本研究结果有助于农业学术论文写作的教学研究。

12-7

扩展意义单位下中外学者学术写作的对比研究

——以最高程度增强语为例

浮莹, 河南师范大学

作为增强事物性质或状态的程度词汇,增强语是学术英语写作中必不可少的成分。由于其意义的相似性与表达形式的多样性,作者在运用增强语的过程中往往出现困惑。因此,有效分析增强语的使用特征,正确运用增强语对学术写作具有重大意义。本研究以Sinclair(2004)的“扩展意义单位”为研究视角,基于两个课题组自建语料库——中国理科学者期刊论文语料库和国际理科学者期刊论文语料库,运用语料库驱动的研究方法,对比分析中外学者最高程度增强语的使用特征。结果发现:中国学者的最高程度增强语使用频率和语义韵与国际学者大致相同,但在其搭配、类联接与语义趋向方面与国

际学者有明显差异。搭配方面，中国学者的词汇丰富度不及国际学者；类链接方面，国内期刊中fully与perfectly的类链接结构种类明显少于国际期刊；语义趋向方面，国际期刊语义趋向呈现丰富、多元化，而国内期刊语义趋向较为单一。这表明中国学者能够恰当运用最高程度增强语，在今后写作中需注意运用多种搭配词汇与语法结构，使最高程度增强语的使用更加地道。

12-8

The Newer, the Better?

--- Testing Four Wordlists with Two Corpora

Ma Haoran, Li Jinman, Shanghai University of Finance and Economics

The General Service List (GSL) developed by West in 1953 and the Academic Word List (AWL) by Coxhead have been useful and helpful for English learning and teaching over the past decades. However, they have also been criticized for different reasons. Hence, a New General Service List (NGSL) was created by Dr. Charles Browne, Dr. Brent Culligan and Joseph Phillips in 2013 together with a New Academic Word List (NAWL). With 5 years having elapsed, the reliability and the validity of the two lists need a close scrutiny considering the rapid change of English language and the ongoing progress within academia in different disciplines or fields. Therefore, the present study aims to compare the two new word lists to their older counterparts with two corpora: the corpus of accounting (AC) and the corpus of finance (FC). Each of them comprises 150 journal articles in their own subjects from the year of 2012 – 2015 and contains about 310,000 tokens in total. The results unravel a clear difference in the coverage between the two pairs of word lists (GSL vs. NGSL and AWL vs. NAWL). In FC, the NGSL provides a better coverage of 85.85% than the GSL, which is 75.06%, while the NAWL gives a shockingly less coverage of only 3.67% than the AWL, which offers 13.45%, about 4 times more text coverage. Likewise, in AC, the NGSL with 84.26% coverage has outperformed the GSL with coverage of 72.4%, whereas the coverage AWL provides is 14.33% and that of NAWL is merely 3.33%. The significant asymmetry of coverage between the two pairs of word lists seems to suggest that the new general word list has been improved to work better but the new academic word list may not necessarily guarantee a better performance as claimed in academic texts from different disciplines.

Group 13

13-1

英语学术语篇中外壳名词的动词搭配研究

马晓雷, 陈颖芳, 国防科技大学

外壳名词的广泛使用是学术语篇的重要特征。以往文献主要依据Schmid (2000) 提出的构式框架开展研究, 却较少关注框架之外的动词与外壳名词之间的搭配规律。本研究利用语义相似度分析、主成分分析、聚类分析和多项特异共现词位分析等方法, 对497篇英语学术论文中的“实义动词+外壳名词”组合进行综合分析。结果表明, 实义动词体现出不同的外壳名词搭配倾向; 构成典型搭配关系的实义动词和外壳名词在语义上具有内在关联; 实义动词与外壳名词的搭配组合构成了科学研究事件域中的基本元素。外壳名词与实义动词的搭配规律可以为学术外语教学提供借鉴与启示。

13-2

A Corpus-based Genre Study on the Essential Moves of Household Manuals

Peng Lei, Chongqing University

Every time when we purchase a household appliance, there will certainly be a booklet for the user's reference. In any case, we call this kind of booklet a household manual. Usually, the reference booklet or household manual has its essential explanation parts, which contains the most important information of the product for the purchaser. However, purchasers seldom will read the manuals before they use the product and only come to its help when they meet problems. Thus, it is necessary to study the typical writing form of manuals to help manufacture writing good quality manuals as well as to help user understand the complex operating procedures. According to Swales, the household manuals contain certain genre moves, steps and sub-steps. Therefore, it can be regarded as a kind of manual genre, which worth studying. In this thesis, household manual is regarded as a certain genre and it will definitely have a prototypical structure. And in each move, step or sub-step, there must be a typical wording and diction to write the instructions. Thus, the thesis started from genre analyzing, using a corpus approach to study its typical writing form of a household manual. Corpus is used in this study to analyze the frequency of moves, steps, and sub-steps.

13-3

基于词汇素描系统的“reform”与“改革”对比研究

李龙兴^{1,2}, 黄居仁¹, 1 香港理工大学, 2 澳门大学

2018年是中国改革开放四十周年, 改革开放深刻地影响了中国的方方面面, 也产生了世界性影响。笔者通过知网对“改革开放”进行知识元检索, 结果显示与“改革开放”相关的文献大多分布在政治、经济学科, 这反映了该词的特性。在语言文化方面与“改革开放”相关的文献多是将其视为一个时代背景或时间界限, 从语言学角度对“改革开放”语篇进行的研究甚少。因此, 本文从语言学角度提出两个研究问题: 1. “改革”及其英文对应词“reform”在词汇语义方面有哪些相似点和不同点? 2. 中英文网络语料库

中对“改革开放”与“reform and opening-up”反映出的情感和态度有何不同？

为回答以上两个问题，本文将首先选取约17亿词的网络汉语语料库（zhTenTen11）以及200亿词的网络英语语料库（enTenTen13），通过词汇素描系统（Sketch Engine）中的词汇素描（word sketch），差异素描（sketch difference）等功能，对两个词进行跨语言对比，揭示其异同。然后通过系统自动生成的与“改革”搭配的“开放政策”以及与“reform”搭配的“and opening-up”词组，选取一定数量的索引行，分别对其进行极性分析和情感分析，以探究中英语料中显示出的对中国改革开放的态度和情感差异。

13-4

基于语料库的动词配价型式特点对比研究

孙海燕 河南师范大学

本文在配价语法视角下探讨中国学习者配价结构的使用特点，研究以动词 *provide* 为例，对比英语本族语者和中国学习者动词配价型式特点，考察节点词的语义和句法特征。研究语料选自学习者作文语料库 TECCL 和美国当代英语语料库 COCA，通过提取节点词的索引行，对比两个语料库中各个配价型式的类型、频数和准确性。节点词 *provide* 的配价型式主要有二价型式和三价型式，研究首先分析了 COCA 中各种配价型式的频数和比例，然后考察了 TECCL 中配价型式的使用比例和使用的准确性。数据显示两个语料库中动词配价型式在频数和类型上存在差异，学习者使用配价型式的类型比本族语者少，在配价型式的使用上存在误用现象，学习者主要使用 *provide* 的三价型式。差异原因可能与学习者语言输入不足以及母语概念迁移有关。本研究基于语料库提供的真实数据，描述了学习者高频动词配价型式使用特点，对外语教学有较好的启示作用。

13-5

The Analysis of Structural and Functional Features of Lexical Bundles in English Major Academic Dissertations

Song Yajun, Henan Institute of Technology

Based on the Hyland(2008) structural and functional framework of 4-word bundles, this paper explored the structural and functional features between English major students' dissertations and international journal papers. The findings showed that in general, the mean sentence length of students' dissertations was shorter than that of international journal papers; in structural aspect, students paid more attention to meaning correlation and explanation, however, international scholars were smarter on statement and abstract description; in functional aspect, students laid emphasis on environmental description, but not good at data mining and explanation and also personal standpoints. Based on real language material, the findings also have a pilot meaning for the improvement of EAP and academic communication ability.

13-6

中南财经政法大学-基于平行语料库的汉英心理形容词语义对比研究
—以《萧萧》原本和英译本为例
高林，中南财经政法大学

At present, the technology of Chinese-English translation contrastive study based on parallel corpus has been improved greatly. Shen Congwen's short story *Xiao Xiao* is famous at home and abroad for its unique style and customs in western Hunan. There are as many as five English versions of *Xiao Xiao*, but the study of its English versions is very limited. This paper summarizes the translation studies of *Xiao Xiao* as follows: at the micro level, there are semantic relations between Chinese and English characters; at the middle level, the study of culture-loaded words, translation strategies and translation activities, and at the macro level the study on the social and cultural context. However, the study of the semantic comparison between Chinese and English psychological adjectives in the original and English versions of *Xiao Xiao* is still vacant, especially from the perspective of micro study of word meaning with the method of semantic primitive structure analysis.

This paper uses parallel corpus as the research tool to study the psychological adjectives in the original and the three English versions of *Xiao Xiao* (Gladys Yang's, Robinson's and Ouyang Zhen's) from the perspective of semantic primitive structure with the method of Minimal pair theory of phonetics. Firstly, the lexical set of Chinese psychological adjectives and three lexical sets of English psychological adjectives are established; Secondly, the semantic primitive structure of Chinese psychological adjectives is analyzed, the semantic primitives are extracted and the primitive set is established. Thirdly, the semantic primitive structure of these English psychological adjectives is analyzed from the perspective of English semantics according to the context of the original context of *Xiao Xiao*, the semantic primitive is extracted and the primitive set is established. Finally, the contrastive study of the psychological adjectives in the original and three English versions of *Xiao Xiao* is carried out. This study not only broadens the perspective of contrastive translation studies, but also provides scientific and objective quantitative research method.

13-7

基于语料库的视觉行为动词词义基元结构研究
李远天，中南财经政法大学

人工智能时代，自然语言处理对词义精确度要求越来越高，词义微观结构的研究越来越受到重视。目前大量纸质词典存在大量循环释义的现象，词义研究存在语义颗粒度较粗，应用范围有限等问题，词义基元结构的研究是对这些进行问题解决的有效尝试，通过分析描写基本词位之间的关系，找出词的概念特征之间的差异，从而构建颗粒度更为精细的词义基元集。本文以北京大学CCL语料库、国家语委现代汉语语料库、中国传媒大学有声媒体语言文本语料库、《现代汉语词典》（第七版）等为基础，以视觉行为动词为研究对象，首先对视觉行为动词界定以便进行封闭、穷尽式搜集与分类，建立视觉行为动词词汇场；运用语音学最小对立体的研究方法研究词义；以一对意义相反

或相近的词为一个语义最小对立体，对视觉行为动词词义基元结构进行分析和基元提取，建立视觉行为动词词义基元集。视觉行为动词是感官类动词的次类范畴可作为个案来研究，初步用来探讨这类典型词的词义基元及基元集的建构，这不仅有助于提高词汇本身的处理技术，词汇语义处理技术的发展和提高为句处理带来突破性的意义，还进一步促进词典编纂，避免词典释义中的循环释义。

13-8

中国学习者英语多义词语义知识的使用和发展

苗丽霞，宁波大学

本研究采用中介语对比分析研究法，探究中国学习者英语多义词语义知识的使用和发展。研究发现：1)学习者使用的多义词义项数目较少，揭示了其大脑词库中狭窄的语义宽度。2)学习者在多义词的语义使用上具有异质性：其一，产出的最高频义项异于本族语者；其二，对多义词多个义项的使用频率失衡。3)学习者的多义词语义宽度知识发展缓慢，没有随着其二语水平的提高同步增长，且在中级阶段以后进入了高原期；4)学习者的多义词语义深度知识从初级到高级持续增长，且在中级阶段以后，表现出更为明显的增长趋势。

Group 14

14-1

信任文本与局部语法 ——以“N₁+死了+N₂”为例 李文中, 浙江工商大学

利用北京语言大学现代汉语语料库(BCC)“文学”子库(30亿字),检索词为“n死了n”(为表述方便,简称“王冕型式”),得到1595个结果,获许下载索引行,通过手工检查,删除不符合该型式以及重复的索引行,最后得到820行索引实例。使用Corpus Parser对索引文本分词,再使用正则表达式删除词性码(使用该软件分词的目的是在词语中间加上空格,以便进行索引分析,因此并不关心分词的准确率),利用AntConc进行索引分析,以描写出其局部语法。本研究从语言事实出发,采用短语学视角和数据驱动方法,并应用局部语法路径,观察和分析该词语型式的各种使用特征,并试图回答两个研究问题:1)“n死了n”作为词项,其强型式与变异型式有哪些?2)意义与型式共选呈现的关系有哪些?其意义移变的具体特征是什么?通过分析,本研究认为,动词“死了”对前后的名词或名词短语,以及后跟名词对前述名词之间都有很强的互选性和限制性,这种互选关系约束了“死了”前后名词的选择,如后跟的名词都是有灵的,一旦出现例外,就可能会发生为隐喻或形成习语。由此可见,汉语中很多所谓“不及物”动词都能放在其动作施行者前面,以减弱该名词在叙述中的权重,同时强调与动作施行者关系密切的前述名词,叙事的焦点不是呈现右侧名词与动作的事件本身,而是这件事对左侧的名词意味着什么。也就是说,在整个叙事进程中,右侧名词(人物或动物)发生的事件被边缘化了,重点强调该事件是发生在第一个名词(人物、处所、集体)身上的。因此,在“王冕死了父亲”这句话中,“父亲死了”不是重要信息,重要的是“父亲死了”这件事是王冕经历的,是他失去了父亲的怙恃。由于词语的共选形成了某种固定的型式,即N₁+死了+N₂。在分析中,该型式可进一步简化为N₁+V[[]]+N₂。本研究认为,该型式可以解释为是词语共选交互作用的结果,理据是:1)不是所有的不及物动词都出现在该结构中;2)一旦确定动词,则该动词对左右出现的词语都有共选限制;3)该型式是词语共选与语法共选适配的结果,但词语共选优先于语法共选,共选的结果是该型式成为一个意义移变单位。

14-2

Beyond Translation: HONGLOUMENG and China Image Xiao Jiayan, Hubei University of Technology

HONGLOUMENG as Chinese literary classics has been translated into more than 23 kinds of languages in the past years, and yet how much it has been received is left unfolded. Communication ensured widely and theoretically that translation of Chinese classics has facilitated the construction and development of China image in at least both of literature and culture, with this background, a case study is exemplified by probing into English translation of HONGLOUMENG both selected and completed by English native translators and Chinese translators, with comparative study of the translators' narrative account of Chinese culture, literature in translation on the specific data, to sort out what cognition of China they retained

and intended to consolidate, to evaluate if they adopted way of neo-liberal and culture-oriented, if the translation(s) convinced positive China image or not. The discussion covers the wording, narrative perspective, cultural awareness and modification, translator's personal interpretation and strategy selection, with the intention of classifying the factors of improving the intercultural communication and more positive response of Chinese classics, and simultaneously better context for China image to be developed. Since Chinese classics is inevitable component embedded with China image and Chinese cultural ecological construction, not only in English context and many others in relation by literature, culture and personality of Chinese individuals, the findings can bring something inspired to the inter-disciplined subject of both classics translation study and communication.

14-3

汉学家中国典籍英译语料库建设：理论与实践的思考

郑建宁，北京师范大学

中国典籍是中华文化的内核，集中体现中华民族的思想价值观念，是世界了解中国的桥梁，也是中国文化“走出去”的必要内容。国外汉学家凭借其英语母语优势及目标读者视角出版了众多的中国典籍英译本，在国外产生不小影响，是推动中国典籍对外传播的重要力量。然而，汉学家的中国典籍英译本尚没有统一的建库整理，已建的语料库也大多规模偏小、文本散缺、设计加工质量不高，并且共享程度有限，这阻碍了对汉学家中国典籍英译的深入研究，因此建设该语料库具有迫切性。在系统梳理现有典籍英译语料库建设及相关研究现状的基础上，本文详细阐述了汉学家中国典籍英译语料库建设对于翻译史研究、汉学家翻译规范研究、典籍英译策略研究乃至翻译教学研究等方面的意义，并重点介绍了语料库建设的总体规划设计与加工思路、原文及译文本的选择标准、文本采集、语料库加工等内容。文章最后对基于该语料库能够开展的研究课题进行了展望。

14-4

Reshaping China's Image: A Corpus-based Discourse Analysis of the English Translation of Chinese Political Discourse

Li Tao, Shanghai Ocean University

Drawing on a combined framework of Appraisal System (Martin and White 2005) and van Dijk's Ideological Square Model (1998; 2006), this paper conducts a corpus-based investigation of China's national image through the English translation of appraisal epithets in Chinese political discourse. The results show that (1) translation shifts do exist in the English translation of appraisal epithets in Chinese political discourse, even though equivalent translation strategy is a canonical principle for translators of Chinese political discourse; (2) translation patterns of appraisal epithets vary within the sub-categories of Appraisal System, with shifts found mostly in the translation of negative appraisal epithets in the subcategories of Engagement and Graduation; (3) China is more negatively presented in the translated English texts than in the source Chinese texts. It is argued that apart from other factors such as the differences between Chinese and English, ideological force is also a vital source to the translation shifts and the variation of China's image.

14-5

The English Translation of Traffic Aids Based on Parallel Corpus

Wen Mengyuan, Wuhan University of Technology

With the continuous deepening of the “One Belt and One Road,” the world’s economy has developed a lot, and the transportation sector is playing an increasingly important role. More and more cities have adopted bilingual translation on traffic aids. However, in the real life there are also many problems in English translation that require us to do further analysis. Therefore, this article uses Wuhan, an internationalized city as a research area, to establish the parallel corpora by collecting a large number of Chinese and English translations of Wuhan traffic pedigrees. To conducts qualitative analysis, and then combines SPSS software with quantitative analysis. The article aims to analyze the English translation features of traffic propaganda from the aspects of word use, sentence pattern, grammar, and culture. Finally to explore the differences between Chinese and English and find suitable English translation methods in the translation of traffic propaganda language. Therefore, it provides reference for regulating the English translation of traffic propaganda language, creates a good traffic language environment, helps foreign friends better understand the traffic propaganda language, further implements the “Belt and Road” policy, and increases the degree of Wuhan’s internationalization.

14-6

基于多译本语料库的儿童文学翻译译文对比研究

——对《卖火柴的小女孩》五译本的对比分析

郑洪波, 上海外国语大学语言研究院

文章抽取自建安徒生翻译多译本平行语料库中的《卖火柴的小女孩》林桦译本、石琴娥译本、任溶溶译本和叶君健译本为研究个案, 同时与小学语文课本收录的叶君健翻译《卖火柴的小女孩》的修改版进行对比, 采用语料库研究方法, 从句长、词类、标点等微观层面探讨译本间的差异。研究发现: 石译和任译句子偏长, 用词丰富; 叶译句子短, 叹号等标点符号频率高, 感情丰富; 叶译的类符/形符比低, 词汇密度低, 语言更简单, 易读性更强。四位译者的译文平均句长均高于课本收录文本, 研究证明了小学语文课本改编和收录叶君健译文的合理性。

14-7

On Narrative Progression of English Fan Translation of Chinese Web-Novel *Coiling Dragon*——with Corpus Narratology Method

Yang Liu, Shanghai International Studies University

Coiling Dragon is the one that ignite American readers’ enthusiasm for Chinese web-novels. A case study is made on the narrative progression of its fan translation on wuxiaworld website with the corpus stylistic method focused on four core parameters first put

forward by Michael Toolan and modified by the author for this research, namely, distribution of top-keyword name of a main character, VVD density, first sentence of each narrative paragraph and analysis of “fully lexical” frequent keywords and clusters. Another web-novel of a different subgenre Xianxia, *Stellar Transformations* and two classic wuxia novels written by Gu Long are taken as reference corpus respectively. The results show that translators left their marks on the English translations and that of “笑” (laugh) is a characteristic word to differentiate the two web-novels. However, they share common features on narrative progression that is overt and simple to predict when compared with Gu Long. The paper specifies the features based on the four parameters in order to offer reference for creative writing, rewriting and translation, and therefore to promote the communication of Chinese cultures since the simplistic progression tends to counteract the obstacles brought about by “cultural otherness”. On the other hand, when the positive implication is acknowledged, a further discussion is in need for the literary, cultural value and Chineseness of web-novels.

14-8

美国计算机协会通讯中译语料库建设及翻译研究

黄宝荣, 上海佑译信息科技有限公司

目前, 国内对计算机科学论文翻译方面的语料库建设和研究较为少见。本文探讨了基于中国计算机学会和美国计算机学会中国分会刊发的《计算机学会通讯》汉译论文英汉双语平行语料库兼中文译者核心期刊论文可比语料库的计算机科学论文汉译研究, 研究首先讨论了适用于论文翻译研究的语料库设计、构建及建库时的语料标注问题, 主要包括语料处理系统的建设、元数据标注、词性标注、翻译方法和技巧标注以及大规模标注时的质量检查和控制。通过与译者可比语料的句长、形次比等文本难度方面的分析, 作者试图揭示论文译文文本与译者自创文本之间的差异, 验证论文翻译中是否存在翻译共性, 并探讨了如何用可视化的方法清晰地展现差异和研究结果。

14-9

Corpus-based Study of Semantic Co-selection among Words in English for Logistical Purpose

Zhou Zhiyi, Guangdong University of Foreign Studies

Based on co-selection theory, this paper explores the basic meaning categories and semantic preference of English for logistical purpose from the perspective of semantic co-selectiveness. Focusing on the 100 keyword list generated by comparing the self-established ELP corpus and OANC (open American national corpus), this paper subdivided 100 keywords into 6 meaning categories and generalize the characteristics of semantic preference among words in English for logistical purpose, aiming to have some implications for logistical purpose teaching and dictionary compilation. (1) There are 6 basic meaning categories in ELP, namely “Logistics participants”, “Logistics entities”, “Logistics units”, “Logistics activities”, “Logistics region/orientation”, and “Logistics regulations”; (2) Semantic preference among representative keywords centers on such 7 meaning categories as “Logistics participants”, “Logistics entities”, “Logistics units”, “Logistics activities”,

“Logistics region/orientation”, “Logistics regulations” and “determiners”, which is in accordance with the basic meaning categories in ELP; (3) There are differences of semantic preference among representative keywords selected from different meaning categories; (4) There are also differences of semantic preference among representative keywords selected from the same meaning category; (5) Some of the selected representative keywords from different categories are co-selected.

(不发言作者论文摘要)

Probe into Gender Differences of Written English Syntactic Complexity by Vocational College Students

Song Renfu, Wuxi Vocational College of Science and Technology

In order to explore the written English syntactic complexity of higher vocational students, this paper investigated 9 discursal structures and 14 indices of syntactic complexity with the help of a self-built corpus of written English and L2 Syntactic Complexity Analyzer(L2SCA). The results showed that female students tended to get higher scores in aspects of 9 structures and 14 indices than males did, and that individual differences within-groups were rather obvious. The study offers empirical evidence and implications in the study of written English gender differences and English writing teaching in vocational colleges.

A Corpus-based Study of Noun Phrases to Express Stance in Chinese EFL Learners' MA Theses

Song Ruimei, Wuhan University

It has been well recognized that writers are expected to express their stance towards the proposition they are talking about in academic writing. This study aims to investigate the use of noun phrases to express stance by postgraduates of applied linguistics in their MA theses. Published journal articles in the field of applied linguistics will be collected to serve as a reference corpus. The overall complexity of noun phrases of stance is analyzed both quantitatively and qualitatively with the help of automated software such as Tree Tagger and AntConc 3.4.4. Then different aspects (including types, frequency and functions of pre-modifiers and post-modifiers) of noun phrase of stance will be investigated. Findings of this study may be of both theoretical and pedagogical implications for L2 academic writing research.

Sketch Engine Based Creation and Analysis of Parallel Corpora of New Year Address

Chen Tong, Meng Sheng, Chengdu University of Technology

Sketch Engine, the well-known online corpora software in the world, is a productive tool in corpora study and research for its powerful features, friendly UI and ease of use. This paper makes use of the "Create", "Compare" and other features of Sketch Engine, to build, analyze and search the C-E parallel corpora of New Year Address.

The Correlation among Title, Downloads and Citations of Linguistics Academic Papers

Liang, Shichen, Beijing normal university of china

Aiming at the different conclusions of the research on the academic thesis, this paper examines all titles of the CSSCI Linguistics Journals, then selects 2779 academic papers titles

to build a small-scale corpus. With Excel and SPSS 23.0, we run regression analysis and One-Way ANOVA. The results show that in terms of linguistics, the title length of the essay is positively related to the counts of essay downloads and citations, as the same time, essay citations is also positively correlated with downloads. Downloads plays a role as a mediation between titles and citations; structurally, compound titles (titles separated into two phrases or sentences by a punctuation) will generally have higher downloads and quotes, but in highly cited journals, it's counts of citations is below average.

光杆有定宾语的认知研究：情境植入视角

卢鑫莹，河南农业大学

本文旨在研究汉语光杆名词在宾语位置实现定指义的认知基础。本文在认知语法情境植入理论框架内，自建70万字的小型光杆名词语料库，在此基础上构建了光杆有定宾语的情境植入功能分析模型。本文发现，话语中光杆有定宾语凸显事物实例，并非事物类型。汉语光杆名词在实际话语中实现定指义并非受句法位置的影响，而是与不同情境植入方式有关，包括“型例合一的内在情境植入”、“领属连带型情境植入”、“时空锚定型情境植入”、“言语互动型情境植入”、“照应型情境植入”、和“量化型情境植入”六种。说话人通过使用不同的情境植入成分开辟了一条听话人与例示域中某一特定个体相联系的心理通道。

基于CiteSpace V的国内英汉平行语料库研究现状与热点分析

唐雪梅，北京师范大学

本文利用可视化分析软件CiteSpaceV,以2000年——2017年中国知网期刊全文数据库中英汉平行语料库研究的相关论文作为数据来源，根据CiteSpaceV提供的知识图谱，分析了相关论文的发表年代、核心作者群、发文机构、以及关键词聚类形成的研究热点。有助于了解英汉平行语料库的发展脉络和知识结构，以期为该领域未来的发展提供参考。本文利用CiteSpaceV对中国知网相关论文的作者、机构、关键词进行统计分析，得出三条结论（1）、英汉平行语料库的相关研究正处于蓬勃发展的阶段，已经形成了核心作者群，他们为英汉平行语料库的发展做出了贡献。（2）英汉平行语料库的研究机构多为国内高校的外语学院，鲜见有汉语相关的学院做关于英汉平行语料库的研究，即英汉平行语料库的研究在英语研究和汉语研究中出现不平衡的现象。（3）英汉平行语料库的研究方向主要分为两方面，构建和应用。构建主要包括语料库建设和创建方法两个研究热点。应用主要包括英汉翻译、英汉对比、词典编纂三个研究热点。

基于可比语料库的性别语言多维对比研究

唐美华，北京外国语大学中国外语与教育研究中心

国内外针对性别语言的实证研究存在以下局限：1）研究对象局限。多集中于口语和其他非正式书面语体，较少涉及正式书面语篇，致使研究的整体性和科学性存在不足；2）研究视角单一。大多聚焦于词汇层面的性别差异，对深层次的句法和语义层面未有涉及；3）对中国汉语语料的本土化研究将进一步加强。本文选取中国小说作家作品（汉语）和英美小说作家作品（英语）自建可比语料库，采用语内和语际对比双重视角，对

性别和母语背景对文本所产生的影响及其动因进行多维度（词汇、句法和语篇）探究。通过语内对比，归纳出单语内男、女性作家在词汇、句法和语篇各层面的独特特征；通过语际对比，探讨两种母语背景下的男（女）性作家在语言使用上的异同，探究背后的社会文化原因。

基于偏序结构理论的情态与时、体、态互动关系知识发现研究

耿延宏，燕山大学

本文基于一个120万字的英语语料库，采用偏序结构理论和方法，研究了情态动词语义与时、体、态句法特征的互动关系。以英语情态动词 *can* 为目标词，划分了 *can* 的 3 个词义：能力(ability)、允许(permission)和可能性(possibility)。以 *can* 与相邻词的互信息作为语义特征，显性和隐性时态、体、语态特征为句法特征，构建了反映情态动词 *can* 的 3 个词义与时、体、态句法特征共现关系的形式背景，并生成了可视化展示情态动词语义与其属性特征之间的偏序结构关系和概念层次关系的属性偏序图。基于此图，发现了情态动词 *can* 与时、体、态特征的互动关系等知识。本文研究结果为自然语言处理中的特征选择和规则提取研究提供了重要的理论和实践依据。

基于语料库对“报复”的词类标注研究

吴琪睿，四川外国语大学

我们通过对基本主流词典的调查，可以发现第五版和第六版现代汉语词典把“报复”标注为动词，释义为“打击批评自己或损害自己利益的人”，第三版现代汉语规范词典也将其归为动词，解释为“对批评自己或损害自己利益的人进行打击”。然而在国家语委语料库中对“报复”的生语料进行检索时，却能明显发现好几个名词的用法：

- ①其二，罪犯违法目的明显，或为财或为色或为报复，病态人格犯罪动机模糊
- ②这明明是报复，软招子。
- ③报复总归是自私的东西，犯罪的动机是永远不应当被支持，被放纵的！

分析发现，词典对“报复”的词类划分有所偏颇，“报复”有不仅可以用作动词，而且还有名词的用法，但是其名词用法却没有被归为词典当中，那么现代汉语词典应当如何对“报复”标注词类，进行词类判断的标准是什么，有什么理论依据，我们从中应该得到哪些启示呢？

本文基于国家语委语料库，在现代汉语语料库中对“报复”一词的使用进行统计研究。将词的表述功能作为词类判断标准，以双层词类范畴为指导，即只要一个词出现了固定的表述模式就可以单独设立一个义项。本文试着给汉语词典中的“报复”的词类标注提出建议，并且得到词典编纂的启示。

基于知网的维吾尔语语义知识库研究

姚洲，赵小兵，中央民族大学

知网是一个描述概念与概念之间的关系以及概念的属性与属性之间的关系的常识知识库，它的知识词典包含汉语词语5万条和对应的概念6万多条，以及与之对应的英语词语5.5万条和概念7万多条。维吾尔语目前有基于框架语义学的维吾尔语框架语义知识库（UyghurFrameNet），但尚未出现以知网为参照的语义知识库。知网是跨语言的，本文

以知网所提出的七个高层分类——万物、时间、空间、属性、属性值、事件和部件为框架，以知网所提取的两千多个义原为基础，以维汉双语电子词典为主要材料，构建维吾尔语的语义知识库，并扩大知网标注语言的范围，增强该库跨语言的适用性。

Corpus-based Study of Simplification Features of Nonnative English from Chinese Perspective

Ling Zhenghua¹, Ma Bosen², Cai Qiang¹, 1 Jiangxi University of Science and Technology, 2 Zhejiang University

Translation Universals (TUs) have often been investigated in translation studies, among which simplification features are relatively more controversial and problematic. Previous corpus-based translation studies have explored simplification mostly by using Laviosa's (1998b) measures, which "revolve[s] around shallow linguistic features" (Wang and Qin 2010, 165). So the present study applies a comparatively fuller range of measures to explore simplification features of nonnative English or ELF by using Laviosa's (1998b) "core patterns of lexical use" and Lu's (2010) L2 syntactic complexity measures based on a self-built corpus, CONTEC. Lexical data show that nonnative English texts use more lexical words and have a wider range of vocabulary in fiction texts, while the nonnative news ones exhibit a lower lexical density and STTR than the native English. The results of syntactic structures also reflect hybridity features, with a blend of both English and Chinese cultures.

通讯录 Contact List

序号	Name	Affiliation	E-mail
1	Malila Carvalho de Almeida Prado	University of São Paulo	
2	Ang Leng Hong	Universiti Sains Malaysia	
3	Nanthini Arivuselvan	Nanyang Technological University	
4	Mark Davies	Brigham Young University	
5	Ding Yingying	University of Bologna	
6	He Mengyu	Universiti Sains Malaysia	
7	Stephen Jeaco	Xi'an Jiaotong-Liverpool University	
8	Jiang Feng	Jilin University	
9	Seetha Lakshmi	Nanyang Technological University	
10	Liang Maocheng	Beijing Foreign Studies University	
11	Hsien-Chin Liou	Feng Chia University	
12	Liu Dilin	The University of Alabama	
13	Liu Zequan	Henan University	
14	Michaela Mahlberg	University of Birmingham	
15	Wei Naixing	Beihang University	
16	白禹	新疆大学	
17	蔡强	浙江大学	
18	蔡颖莹	浙江财经大学	
19	曹竞	中南财经政法大学	
20	陈程	浙江大学	
21	陈孚	上海师范大学, 上海外国语大学	
22	陈建生	广州体育学院	

The 4th Corpus Linguistics in China Conference

23	陈建生	天津科技大学	
24	陈丽丹	广东外语外贸大学	
25	陈凌微	浙江财经大学	
26	陈彤	成都理工大学	
27	叱干梦华	西北农林科技大学	
28	邓海龙	北京外国语大学/赣南师范大学	
29	邓斯芮	贵州大学	
30	邓耀臣	大连外国语大学	
31	董敏	北京航空航天大学	
32	方凯	武汉理工大学	
33	房印杰	北京邮电大学	
34	浮莹	河南师范大学	
35	高林	中南财经政法大学	
36	高霞	北京航空航天大学	
37	耿延宏	燕山大学	
38	管新潮	上海交通大学	
39	何华清	西华师范大学	
40	何琼	武汉大学	
41	洪扬	武汉理工大学	
42	侯羽	燕山大学	
43	侯智德	暨南大学深圳旅游学院	
44	胡春雨	广东外语外贸大学	
45	胡婧	华东师范大学	
46	胡梅芬	上海外国语大学	

The 4th Corpus Linguistics in China Conference

47	黄宝荣	上海佑译信息科技有限公司	
48	黄慧	天津科技大学	
49	姜占好	西安外国语大学	
50	金莹莹	上海理工大学	
51	康佳萍	西安外国语大学	
52	康宁	青岛科技大学	
53	雷蕾	华中科技大学,	
54	雷树杰	洛阳外国语学院	
55	雷小丹	中国地质大学	
56	李斌	南京师范大学	
57	李华勇	四川文理学院	
58	李金凤	重庆大学	
59	李晶洁	东华大学	
60	李丽霞	西北农林科技大学	
61	李龙兴	澳门大学	
62	李涛	上海海洋大学	
63	李文中	浙江工商大学	
64	李夏青	中央民族大学	
65	李晓红	上海海洋大学	
66	李艳娇	山东大学	
67	李远天	中南财经政法大学	
68	梁诗尘	北京师范大学	
69	凌征华	江西理工大学	
70	刘樊	浙江大学	

The 4th Corpus Linguistics in China Conference

71	刘淼	北京第二外国语学院	
72	刘萍	上海外国语大学	
73	刘洋	北京外国语大学	
74	刘运锋	北京外国语大学/安徽工程大学	
75	龙绍赞	江西师范大学	
76	卢春梅	广东外语外贸大学	
77	卢鑫莹	河南农业大学	
78	陆军	扬州大学	
79	马浩然	上海财经大学	
80	马晓雷	国防科技大学	
81	苗丽霞	宁波大学	
82	潘鸿	武汉理工大学	
83	彭雷	重庆大学	
84	钱珍	广东财经大学	
85	邱锡鑫	北京航空航天大学	
86	邵斌	浙江财经大学	
87	石欣玉	西安外国语大学	
88	宋仁福	江苏省无锡科技职业学院	
89	宋瑞梅	山东商务职业学院/武汉大学	
90	宋亚军	河南工学院	
91	苏杭	北京航空航天大学	
92	孙海燕	河南师范大学	
93	孙婉	武汉大学	
94	孙由之	中国矿业大学（北京）	

The 4th Corpus Linguistics in China Conference

95	唐美华	北京外国语大学	
96	唐雪梅	北京师范大学	
97	陶国潇	华中科技大学	
98	田笑语	上海外国语大学	
99	王进晓	广东外语外贸大学	
100	王蜜蜜	河南师范大学	
101	王闵利	西安交通大学	
102	王雪勤	武汉理工大学	
103	王在雅	河南师范大学	
104	王重	华中科技大学/香港城市大学	
105	温梦媛	武汉理工大学	
106	吴良平	湖南商学院/北京外国语大学	
107	吴琪睿	四川外国语大学	
108	肖诚	——	
109	肖家燕	湖北工业大学	
110	肖巍	重庆大学	
111	肖先明	武汉理工大学	
112	邢富坤	洛阳外国语学院	
113	徐坚俊	宁波大学	
114	宣瑾	华中科技大学	
115	闫鹏飞	北京外国语大学/北京理工大学	
116	杨晨姣	北京航空航天大学	
117	杨林伟	烟台大学/北京外国语大学	
118	杨璘璘	广东行政职业学院	

The 4th Corpus Linguistics in China Conference

119	杨柳	上海外国语大学	
120	杨越森	北京航空航天大学	
121	杨紫彤	北京邮电大学	
122	姚洲	中央民族大学	
123	要新乐	中国人民大学	
124	喻璇	重庆大学	
125	袁新华	北京外国语大学	
126	张宝林	北京语言大学	
127	张大群	南昌大学	
128	张 懂	北京外国语大学	
129	张家玲	北京第二外国语学院	
130	张金福	九江学院	
131	张乐	上海理工大学	
132	张立英	信息工程大学洛阳校区	
133	张蒙蒙	北京外国语大学	
134	张瑞丹	上海外国语大学	
135	张绪华	复旦大学,	
136	张雪梅	上海外国语大学	
137	张子菡	齐鲁师范学院	
138	赵朝永	华东师范大学	
139	赵佳美	天津科技大学	
140	赵立书	长春师范大学	
141	赵晓临	东华大学	
142	赵永刚	北京外国语大学	

The 4th Corpus Linguistics in China Conference

143	郑洪波	上海外国语大学	
144	郑建宁	北京师范大学	
145	周杰	贵州大学	
146	周芷伊	广东外语外贸大学	
147	邹洪	西华师范大学	
148	邹雨橙	中央民族大学	

校园地图 Campus Map



华中科技大学简介

华中科技大学是国家教育部直属重点综合性大学，由原华中理工大学、同济医科大学、武汉城市建设学院于2000年5月26日合并成立，是国家“211工程”重点建设和“985工程”建设高校之一，是首批“双一流”建设高校。

学校校园占地7000余亩，园内树木葱茏，碧草如茵，环境优雅，景色秀丽，绿化覆盖率72%，被誉为“森林式大学”。学校教学科研支撑体系完备，各项公共服务设施齐全。

学校学科齐全、结构合理，基本构建起综合性、研究型大学的学科体系。拥有哲学、经济学、法学、教育学、文学、理学、工学、医学、管理学、艺术学等10大学科门类；设有98个本科专业，214个硕士学位授权点，183个博士学位授权点，39个博士后科研流动站；现有一级学科国家重点学科7个，二级学科国家重点学科15个（内科学、外科学按三级），国家重点（培育）学科7个。在教育部第四轮学科评估中，14个学科进入A类（其中机械工程、光学工程、生物医学工程、公共卫生与预防医学4个学科进入A+），28个学科进入B类（其中19个进入B+）。

学校实施“人才兴校”战略，师资力量雄厚。现有专任教师3000余人，其中教授1000余人，副教授1300余人；教师中有院士16人，“千人计划”入选者38人，“外专千人计划”入选者7人，“青年千人计划”入选者123人，长江学者特聘教授59人、讲座教授42人，长江学者青年项目15人，国家杰出青年科学基金获得者65人，“973计划”项目首席科学家15人，重大科学研究计划项目首席科学家2人，国家重点研发计划项目首席科学家24人，973计划（含重大科学研究计划）青年科学家3人，优秀青年科学基金获得者43人，国家级教学名师9人，“万人计划”领军人才29人、青年拔尖人才21人，教育部新世纪优秀人才支持计划入选者224人，国家百千万人才工程入选者39人，湖北省“百人计划”入选者52人，“楚天学者”207人。

学校贯彻建设“学生、学者与学术的大学”的教育思想，秉承“育人为本、创新是魂、责任以行”的办学理念，坚持“一流教学一流本科”的建设目标，采取多种举措，深化教育教学改革，全面推进素质教育，构建和完善充满活力的创新人才培养体系。几十年来，已为国家培养了近40万名高级人才。

按照“应用领先、基础突破、协调发展”的科技发展方略，构建起了覆盖基础研究层、高新技术研究层、技术开发层三个层次的科技创新体系。建设有武汉光电国家研究

中心以及国家脉冲强磁场科学中心（筹）、精密重力测量研究设施等国家重大科技基础设施，还拥有 4 个国家重点实验室、1 个国防科技重点实验室、6 个国家工程（技术）研究中心、1 个国家工程实验室、2 个国家专业实验室及一批省部级研究基地。

学校坚持“服务乃宗旨，贡献即发展”的办学思路，面向国民经济建设主战场，学研产相结合，密切与地方政府、大型企业的合作关系，通过设立驻外研究院、产业化基地，开展横向科技合作等方式为区域经济建设服务。

学校坚持开放式办学理念，积极开展全方位、多层次的国际交流与合作，目前已与世界上 35 个国家和地区的 100 多所大学建立了良好的校际交流关系，每年有约 1600 余人次的国（境）外专家学者来我校任教、合作科研和开展学术交流活动。

附属协和医院、同济医院是集医疗、教学、科研、培训为一体的大型现代化综合性医院，是湖北省乃至中南地区的医疗诊治中心。附属梨园医院突出老年病学的特色，是湖北省老年病防治研究中心。

华中科技大学正以创建世界一流大学为目标，秉持“明德厚学，求是创新”的校训，敢于竞争，善于转化，聚精会神，科学发展，全面提升办学水平，努力开创更加辉煌灿烂的明天。

Introduction to Huazhong University of Science and Technology

Huazhong University of Science and Technology (HUST) is located in Wuhan, the capital city of Hubei Province, on the central banks of the Yangtze River. The university is composed of one main campus situated near the East Lake scenic spot and a satellite medical campus located in the central business district of Hankou on the other side of the Yangtze River. HUST's picturesque environment covers an area of over 460 hectares, and the campus has 72% greenery coverage, which has earned it the honorary title "University in the Forest".

HUST was established in 1952, originally under the name of Huazhong Institute of Technology. In 2000, the university merged with Tongji Medical University (founded in 1907) and Wuhan Urban Construction Institute (founded in 1952), to form the new Huazhong University of Science and Technology as it is known today.

HUST is a comprehensive and multi-disciplinary research university, covering 10 major disciplines: philosophy, economics, law, education, literature, science, engineering, medicine, management and arts, offering a variety of degree programs, including 98 undergraduate programs, 214 master programs, and 183 PhD programs. There are 39 post-doctoral research centers, 7 national key first-level disciplines and 15 national key sub-disciplines. Three disciplines namely, Mechanical Engineering, Optical Engineering, and Public Health and Preventive Medicine ranks top 1 among its peers in China, while both Electrical Engineering and Biomedical Engineering rank 2nd. HUST also hosts various national advanced talent training bases for undergraduate education, including the "National Exemplary Software Faculty", the "National Base of Life Science and Technological Talents", the "National Base of IC Talents", the "National Education Base of Basic Science Courses", and the "School of Engineering Sciences".

HUST hosts around 3000 full-time faculty members, more than 1000 of which are professors and 1300 associate professors. A large number of eminent scientists and professors are gathering here, including 16 academicians of the Chinese Academy of Science and the Chinese Academy of Engineering, 38 *Thousand Talents Program* members, 7 *The One-Thousand Foreign Experts Program* members, 123 *The Thousand Talents Program for Distinguished Young Scholars* members, 59 *Changjiang Scholars Program for Distinguished Professors*, 42 *Changjiang Scholars Program for Visiting Professors*, 42 winners of *The National Science Fund for Distinguished Young Scholars*, 15 Chief Scientists for *973 Program*, 2 Chief Scientists for *Key Research Programs*, 3 *Young Chief Scientists for 973 program*, 43 winners of *Research Fund for Distinguished Young Scientists*, 224 winners of *Program for New Century Excellent Talents in University*, 39 *Millions of Leading Engineering Talents*, 9 *National Teaching Masters*.

A total of more than 55,000 full-time students are enrolled, including approximately 32,000 undergraduates, 16,000 master candidates and 6,000 doctoral candidates' students have won prizes in various national and international competitions. WANG Liran of the School of Management won the "Future Business Leader" award at the final of 2014 CIMA held in Bombay on August 6, 2014. The CIMA Global Business Challenge is known as "the business Olympics for college students". At the World Final of the 23rd World Memory Championship 2014 held in Hainan on December 13, SHU Yubiao, a HUST physics sophomore and MAO Hongwei of the School of Law were crowned as "World Memory Masters". In recent years, the growing social impact of HUST on social life has helped the university admit the best possible students. HUST graduates have been enjoying popularity among employers in the job market, and the employment rate has reached above 95%.

HUST is proud of her national laboratory - "Wuhan National Laboratory for Optoelectronics" (one of the six national labs in China), the National Science Center-- "Wuhan National High Magnetic Field Center", the Center for Gravitational Experiments and the biggest oxy-fuel combustion carbon capture pilot system in China—"3 MWth Carbon Capture Test Base". In addition, various research facilities have been established. In recent years, a number of high qualified academic papers cited by SCI, EI, ISTP and Medical Literature Analysis and Retrieval System (MEDLARS) have been published. The number of published articles is among the top ten universities in China. In 2014, 7 HUST research projects were awarded by the annual National Science and Technology Awards Ceremony.

HUST has been paying sustainable attention to the research-industry transformation since its founding in 1952. The university services local economy by developing cooperative programs and setting up research institutes and industrialization bases with other neighboring province governments and industries. Now the university has developed itself into one of the seven national technology transfer centers in China.

HUST has established partnerships with more than 100 renowned universities and research institutions worldwide. Over 1,600 scholars and experts visit and lecture here annually. Since 2013 the campus has welcomed several Nobel Laureates including Prof. Alan J. Heeger, the Nobel Laureate in Chemistry, and Prof. Klaus von Klitzing, German Nobel Prize winner in Physics. Currently around 3000 international students are studying at HUST. The students coming from different parts of the world gather here for degree and non-degree programs. Some courses are taught in Chinese while most of the degree programs are taught in English. Scholarships are accessible to all the applicants.

华中科技大学外国语学院简介

华中科技大学外国语学院设有英语语言文学、日语语言文学、德语语言文学、翻译和法语语言文学等五个本科专业，负责全校公共外语教学的大学外语系，以及负责研究生专业课程教学的外国语言文学研究所。学院设有外国语言文学一级学科硕士学位授权点，外国语言学及应用语言学、英语语言文学、日语语言文学和德语语言文学等四个二级学科硕士学位授权点以及翻译硕士专业学位授权点。设有“中外语言文化比较研究”二级学科博士学位授权点。“外国语言文学”为湖北省重点学科。

学院现有专任教师 153 人，其中教授 17 人、副教授 65 人。博士生导师 13 人，硕士生导师 43 人。“湖北名师”1 人，校“教学名师奖”4 人，“新世纪优秀人才”1 人、“楚天学子”1 人、“华中学者”6 人，有 2 位教授在教育部高等学校外语教学指导委员会（大学外语及英语专业）、教育部大学英语四、六级考试委员会任职。现有“大学英语”和“英汉互译”2 门国家级精品课程，这 2 门课程已经升级为国家级精品资源共享课。

学院设有外语教育研究中心、语言学研究中心、翻译研究中心、比较文化（文学）研究中心、区域（国别）研究中心等五个研究中心。建有同声传译实验室、计算机辅助翻译实验室、口译实训平台以及湖北省语言学习示范中心。

Introduction to School of Foreign Languages, HUST

School of Foreign Languages of Huazhong University of Science and Technology offers undergraduate degrees in 5 majors, namely English Language & Literature, Japanese Language & Literature, German Language & Literature, French Language & Literature and Translation. The College English Teaching Center oversees the foreign language teaching of the whole university. The school has one authorized primary master program (Master of Language & Literature), and four authorized secondary master programs (Master of Linguistics & Applied Linguistics, English Language & Literature, Japanese Language & Literature, German Language & Literature) as well as one MTI program (Master of Translation and Interpretation). Besides, there is also one authorized secondary doctoral program (Comparative Study on Chinese and Foreign Language Culture). Moreover, “Foreign Languages & Literature” is the key discipline of Hubei Province.

There are in total 153 full-time teachers in the school, including 17 professors and 65 associate professors, 13 of them are doctoral supervisors and 43 of them are master supervisors. We have one “Outstanding Teacher of Hubei Province” and 4 top teachers of our university. Moreover, 2 of the professors hold a post in Teaching Steering Committee of Foreign Languages, CET-4 and CET-6 Examination Committee of Ministry of Education as well as English Professional Steering Committee. There are 3 professors respectively serving as a member of the China Bilingual Lexicography Committee, China Association for Comparative Studies of English and Chinese as well as a permanent member of Hubei Foreign Literature Association. Among all the courses, there are two national excellent resources sharing courses, namely “College English” and “English-Chinese Translation”. “Foreign Language and Literature” is one of the provincial key disciplines.

The school houses a range of advanced teaching and research facilities. There are 5 research centers, namely Center for Foreign Language Education Research, Center for Comparative Cultural & Literary Studies, Center for Foreign Language Research, Center for Translation Studies and Center for Regional Studies. Besides, there are state-of-the-art professional laboratories and language laboratories, such as Simultaneous Interpretation Laboratory, Computer-aided Translation Laboratory, Interpretation Training Platform, Televisic Language Laboratory and Language Learning Demonstration Center of Hubei Province.