

首届全国学习者语料库专题研讨会组委会

主任：梁茂成

成员：卫乃兴、文秋芳、许家金、何安平、李文中、洪华清、梁茂成、濮建忠（以姓氏笔划为序）

秘书：段长城、郑丹妮、陈功

主办：北京外国语大学中国外语教育研究中心

承办：北京外国语大学中国外语教育研究中心

协办：外语教学与研究出版社

目 录

| | |
|-------------|-----|
| 会议日程..... | iii |
| 小组发言安排..... | 1 |
| 主旨发言摘要..... | 1 |
| 小组发言摘要..... | 6 |
| 参会代表索引..... | 60 |

会议日程

2009年12月25日（星期五）

| 时间 | 议程 | 地点 |
|-------------|----|-------------------------|
| 08:00-18:00 | 报到 | 北外宾馆一层大厅（北京外国语大学西院东门北侧） |

2009年12月26日 上午（星期六）

| 时间 | 议程 | | 发言人 | | 地点 | |
|-------------|-----------------|------|---|-----|---------------------|-----|
| 08:30-09:00 | 致词 | | 文秋芳 | | 北外 中文学院 学术报告厅 | |
| 09:00-09:20 | 合影 | | | | | |
| 09:20-10:10 | 编号 | 议程 | 题目 | 发言人 | | 主持人 |
| | 1-1 | 主旨发言 | Phraseology in Contrast: A Multi-dimensional Contrastive Analysis of Learner Phraseologies across Corpora | 卫乃兴 | | 梁茂成 |
| 10:10-11:00 | 1-2 | 主旨发言 | “过多使用”，“过少使用”，“错误使用”，so what? | 何安平 | 文秋芳 | |
| 11:20-12:00 | 2-1-1~ 2-4-2 | 分组发言 | | | 第一至第四分会场（中文学院） | |

2009年12月26日 下午（星期六）

| 时间 | 编号 | 议程 | 题目 | 发言人 | 主持人 | 地点 |
|-------------|-----------------|------|---|-----|----------------|---------------------|
| 14:00-14:50 | 1-3 | 主旨发言 | Constructing a Learner English Portfolio Corpus on the Open Corpus Platform | 李文中 | 何安平 | 北外 中文学院 学术报告厅 |
| 15:00-15:50 | 1-4 | 主旨发言 | 中国英语专业大学生英语口语口头叙事中的自我形象：词语搭配研究视角 | 许家金 | 洪华清 | |
| 16:00-18:00 | 2-5-1~ 2-8-6 | 分组发言 | | | 第一至第四分会场（中文学院） | |

2009年12月27日 上午(星期日)

| 时间 | 编号 | 议程 | 题目 | 发言人 | 主持人 | 地点 |
|-------------|------------------|------|---|-----|-----|------------------------|
| 08:00-08:50 | 1-5 | 主旨发言 | 从词块的使用看学习者语言能力的 发展 | 濮建忠 | 李文中 | 北外 中文学院 学术报告厅 |
| 09:00-09:50 | 1-6 | 主旨发言 | Multimodal Learner Corpus Construction: Challenges and Directions | 洪华清 | 许家金 | |
| 10:00-12:00 | 2-9-1~ 2-12-6 | 分组发言 | | | | 第一至第四 分会场(中文 学院) |

2009年12月27日 下午(星期日)

| 时间 | 编号 | 议程 | | | | 地点 |
|-------------|-------------------|------|----------------------------------|-----|-----|------------------------|
| 14:00-14:40 | 2-13-1~ 2-16-2 | 分组发言 | | | | 第一至第四 分会场 (中文学院) |
| | 编号 | 议程 | 题目 | 发言人 | 主持人 | 地点 |
| 14:50-15:40 | 1-7 | 主旨发言 | Keywords Plus: 主题词 研究方法扩展及其实现 | 梁茂成 | 卫乃兴 | 北外 中文学院 学术报告厅 |
| 15:40-16:00 | 闭幕式 | | | | | |

小组发言安排

第一组

主持人：王立非

时 间：12月26日 上午

| 时 间 | 编 号 | 发 言 人 | 发 言 题 目 | 地 点 |
|-------------|-------|-------|---------------------------------|-------|
| 11:20—11:40 | 2-1-1 | 王立非 | 我国学生英语演讲中的话语语块特点 ——基于语料库的考察 | 第一分会场 |
| 11:40—12:00 | 2-1-2 | 柴 静 | 基于语料库对中国大学生英语口语中 个人立场词块的调查研究 | |

第二组

主持人：洪华清

时 间：12月26日 上午

| 时 间 | 编 号 | 发 言 人 | 发 言 题 目 | 地 点 |
|-------------|-------|-------------|-------------------------|-------|
| 11:20—11:40 | 2-2-1 | 杨瑞英、洪 华清 | 学习者英语学术论文语料库建设及应 用 | 第二分会场 |
| 11:40—12:00 | 2-2-2 | 张金福 | 中国英语专业学士学位论文语料库建设 构想 | |

第三组

主持人：张红燕

时 间：12月26日 上午

| 时 间 | 编 号 | 发 言 人 | 发 言 题 目 | 地 点 |
|-------------|-------|-------------|------------------------|-------|
| 11:20—11:40 | 2-3-1 | 张红燕 | 认知视角下的小学生语文作文话语分 析 | 第三分会场 |
| 11:40—12:00 | 2-3-2 | 刘泽权、刘 艳红 | 基于语料库的通用大学英语教材对比研 究 | |

第四组

主持人：罗卫华

时 间：12月26日 上午

| 时 间 | 编 号 | 发 言 人 | 发 言 题 目 | 地 点 |
|-------------|-------|-------------|-------------------------------------|-------|
| 11:20—11:40 | 2-4-1 | 邓耀臣、罗 卫华 | 基于 SECOPETS 语料库的学习者英语 口语失误分析 | 第四分会场 |
| 11:40—12:00 | 2-4-2 | 陈 佳 | 基于语料库的中国英语专业学生口语笔 语there 存在句使用研究 | |

第五组

主持人：陆军

时 间：12月26日 下午

| 时 间 | 编 号 | 发 言 人 | 发 言 题 目 | 地 点 |
|-------------|-------|------------|--------------------------------|-------|
| 16:00—16:20 | 2-5-1 | 陆 军 | 基于学习者英语语料库的扩展语义单位研究 | 第一分会场 |
| 16:20—16:40 | 2-5-2 | 何晶、张萍 | “CAUSE”——基于语料库的中国学习者学术写作词汇能力研究 | |
| 16:40—17:00 | 2-5-3 | 何 影 | 中国英语学习者写作中副词与形容词搭配的使用特点 | |
| 17:00—17:20 | 2-5-4 | 胡玉迎 | 专业物流英语词块特色研究 | |
| 17:20—17:40 | 2-5-5 | 黄瑾、崔艳 嫣 | 语料库在同义词搭配习得过程中的应用研究 | |
| 17:40—18:00 | 2-5-6 | 林维燕 | 论机切口语词块的元话语性 | |

第六组

主持人：黄大网

时 间：12月26日 下午

| 时 间 | 编 号 | 发 言 人 | 发 言 题 目 | 地 点 |
|-------------|-------|----------------|------------------------------|-------|
| 16:00—16:20 | 2-6-1 | 黄大网、徐 亮、杨廷君 | 一个社群、三类语境：中国学者期刊论文摘要的多维度体裁分析 | 第二分会场 |
| 16:20—16:40 | 2-6-2 | 邓海龙 | 基于语料库的中国英语变体研究 | |
| 16:40—17:00 | 2-6-3 | 江淑娟 | 语体研究与研究生科技论文英文摘要写作教学 | |
| 17:00—17:20 | 2-6-4 | 李学宁 | 多模态信息检索工具——MCA 及其在外语教学中的应用 | |
| 17:20—17:40 | 2-6-5 | 刘 芹、潘 鸣威 | 理工科大学生英语口语多模态学习者语料库的构建研究 | |
| 17:40—18:00 | 2-6-6 | 潘鸣威 | 英语专业写作教学语料库的建设与应用研究 | |

第七组

主持人：赵勇

时间：12月26日 下午

| 时间 | 编号 | 发言人 | 发言题目 | 地点 |
|-------------|-------|---------|--------------------------------|-------|
| 16:00—16:20 | 2-7-1 | 赵勇 | 作为中介语的“翻译腔”问题初探——一项基于学习者语料库的研究 | 第三分会场 |
| 16:20—16:40 | 2-7-2 | 龙绍赟 | 情态动词在教学大纲中的语法排序对比——以中外语料库为根据 | |
| 16:40—17:00 | 2-7-3 | 邵莉 | 促进学习者自主性：基于学习者翻译语料库的翻译评估 | |
| 17:00—17:20 | 2-7-4 | 汪兴富、刘世平 | 运用美国当代英语语料库(COCA)提升学习者真实语感 | |
| 17:20—17:40 | 2-7-5 | 王亚军 | 中国英语学习者语料库在双语学习词典编纂中的应用 | |
| 17:40—18:00 | 2-7-6 | 张莎 | 基于语料库的中美大学生议论文高频动词及物性研究 | |

第八组

主持人：权立宏

时间：12月26日 下午

| 时间 | 编号 | 发言人 | 发言题目 | 地点 |
|-------------|-------|-------------|---|-------|
| 16:00—16:20 | 2-8-1 | 权立宏 | 中国大学生英语会话中的回指修补研究 | 第四分会场 |
| 16:20—16:40 | 2-8-2 | 邓飞 | 连接副词的语用功能与使用情况调查分析 | |
| 16:40—17:00 | 2-8-3 | 付玉萍 | A Corpus-based Study on Written Vocabulary by First-year English Majors | |
| 17:00—17:20 | 2-8-4 | 何华清 | 大学英语写作中的词形错误分析——一项基于语料库的研究 | |
| 17:20—17:40 | 2-8-5 | 黄斌兰、刘士祥、朱兵艳 | 从英语专业限时议论文看被动语态习得——基于 SWECCCL 的对比研究 | |
| 17:40—18:00 | 2-8-6 | 刘小燕 | 中国学习者语域意识的多维度分析 | |

第九组

主持人：杨 华

时 间：12月27日 上午

| 时 间 | 编 号 | 发 言 人 | 发 言 题 目 | 地 点 |
|-------------|-------|-------------|--|-------|
| 10:00—10:20 | 2-9-1 | 杨 华 | 中国研究生英语学术论文摘要中的词块研究 | 第一分会场 |
| 10:20—10:40 | 2-9-2 | 刘 浩 | 基于语料库的同义词语义韵辨析及对外语教学的启示——以 effect 为例 | |
| 10:40—11:00 | 2-9-3 | 刘士祥、韦保丞、朱兵艳 | 英语专业大一、大三学生英语议论文中后置定语“语块输出”特点——基于SWECCCL语料库的对比研究 | |
| 11:00—11:20 | 2-9-4 | 卢美华 | 组块在英汉平行语料库中的对齐和提取 | |
| 11:20—11:40 | 2-9-5 | 唐洁仪 | 基于语料库的英语专业教材词块分析 | |
| 11:40—12:00 | 2-9-6 | 赵秋荣、张红燕 | 基于语料库的增强语程度副词语义韵认知研究 | |

第十组

主持人：蒋凤霞

时 间：12月27日 上午

| 时 间 | 编 号 | 发 言 人 | 发 言 题 目 | 地 点 |
|-------------|--------|----------------|---|-------|
| 10:00—10:20 | 2-10-1 | 蒋凤霞 | A Corpus-based Study on Vowel Pronunciation of Chinese EFL Learners | 第二分会场 |
| 10:20—10:40 | 2-10-2 | 王晓雯 | 售书广告对书籍序言的语类侵殖——一项基于语料库的研究 | |
| 10:40—11:00 | 2-10-3 | 叶拉美 | 教学资源优化与微型文本创建 | |
| 11:00—11:20 | 2-10-4 | 曾小兵、杨尔弘、邱丽娜、张普 | 针对特定领域对外汉语教学的语料库建设 | |
| 11:20—11:40 | 2-10-5 | 程世禄 | IT 英语语料库统计数据对大学英语教学改革的启示 | |
| 11:40—12:00 | 2-10-6 | 张媛媛 | 中美大学生英语定题演讲句式对比研究——一项基于语料库的研究 | |

第十一组

主持人：朱红强

时 间：12月27日 上午

| 时 间 | 编 号 | 发 言 人 | 发 言 题 目 | 地 点 |
|-------------|--------|--------|---|-------|
| 10:00—10:20 | 2-11-1 | 朱红强 | 中国英语学习者与本族英语学习者情态动词 CAN 的对比研究——一项基于语料库的调查 | 第三分会场 |
| 10:20—10:40 | 2-11-2 | 王 勇、杨晶 | 语料库在同义词辨析中的作用 | |
| 10:40—11:00 | 2-11-3 | 张利杰 | 基于语料库的连贯关系比较研究 | |
| 11:00—11:20 | 2-11-4 | 张 跃 | 基于语料库的外语教研平台 Direct Data | |
| 11:20—11:40 | 2-11-5 | 赵 晴 | 基于语料库的《红楼梦》两个英译本的译者风格研究 | |
| 11:40—12:00 | 2-11-6 | 郑洪波 | 基于语料库的中国学习者写作中单词“TRUE”的使用研究 | |

第十二组

主持人：张雪梅

时 间：12月27日 上午

| 时 间 | 编 号 | 发 言 人 | 发 言 题 目 | 地 点 |
|-------------|--------|------------|---|-------|
| 10:00—10:20 | 2-12-1 | 张雪梅、杨滢滢、冯辉 | 中国英语学习者的时态习得现状分析——基于 CEM 和 CLEC 的分析 | 第四分会场 |
| 10:20—10:40 | 2-12-2 | 马艳姿、谢晓莺 | 中国学生习得虚拟语气特征分析——基于语料库研究 | |
| 10:40—11:00 | 2-12-3 | 潘不寒 | A Study of the Use of the Conjunction BECAUSE in L2 Hong Kong University Students' Writings | |
| 11:00—11:20 | 2-12-4 | 王岩、吴丽萍 | 中国学习者英语口语中人称代词与互动的研究 | |
| 11:20—11:40 | 2-12-5 | 张 祺 | 学生会话话轮转换处的会话填充语——一项基于语料库的研究 | |
| 11:40—12:00 | 2-12-6 | 王 颖 | 中国英语学习者英语介词 at 的使用研究——基于语料库的分析 | |

第十三组

主持人：陈浩然

时 间：12月27日 下午

| 时 间 | 编 号 | 发 言 人 | 发 言 题 目 | 地 点 |
|-------------|--------|-------|---|-------|
| 14:00—14:20 | 2-13-1 | 陈浩然 | Semi-automatic Retrieval of Collocation Errors based on Taiwanese Learner Corpora | 第一分会场 |
| 14:20—14:40 | 2-13-2 | 陈 颖 | 英语增强词的语料库对比研究——以 very much 为例 | |

第十四组

主持人：钱敏汝

时 间：12月27日 下午

| 时 间 | 编 号 | 发 言 人 | 发 言 题 目 | 地 点 |
|-------------|--------|--------|---------------------|-------|
| 14:00—14:20 | 2-14-1 | 钱敏汝、李婧 | 中国大学生德语中介语语料库的创建及应用 | 第二分会场 |
| 14:20—14:40 | 2-14-2 | 常 哲 | 中医药实验性学位论文的英文摘要探析 | |

第十五组

主持人：刘泽权

时 间：12月27日 下午

| 时 间 | 编 号 | 发 言 人 | 发 言 题 目 | 地 点 |
|-------------|--------|---------|--------------------------|-------|
| 14:00—14:20 | 2-15-1 | 刘泽权、刘鼎甲 | 基于语料库的翻译教学与学习者译本评析 | 第三分会场 |
| 14:20—14:40 | 2-15-2 | 李 渝 | 四级机考形势下大学英语写作教学新模式的研究与探索 | |

第十六组

主持人：郭孝田

时 间：12月27日 下午

| 时 间 | 编 号 | 发 言 人 | 发 言 题 目 | 地 点 |
|-------------|--------|-------|--|-------|
| 14:00—14:20 | 2-16-1 | 郭孝田 | Examining the Parameters of Comparison in Learner English Research—Beyond “overuse” and “underuse” | 第四分会场 |
| 14:20—14:40 | 2-16-2 | 陈 功 | 中国学生英语写作中语法化连接词的研究 | |

主旨发言摘要

1-1

12月26日，星期六，09:30-10:20
北外中文学院学术报告厅

卫乃兴

上海交通大学

Phraseology in Contrast: A Multi-dimensional Contrastive Analysis of Learner Phraseologies across Corpora

Abstract: This paper sets out to address a multi-dimensional model for analyzing EFL learners' phraseologies in production. It argues that the unit of analysis in learner corpora studies should be a phrase rather than a word in isolation. On the one hand, it has been a consensus view that a word on its own does not carry much meaning whilst it is phrases of various kinds which realize meaning in native speaker communication. On the other hand, long-established paradigms in learner corpora study have not given adequate attention to learner phraseologies. The paper proposes to adopt learner phraseologies as the unit of analysis, which include, among others, collocations, semi-fixed expressions and lexico-grammatical sequences. The analysis is multi-dimensional in the sense that the phraseology is to be examined in terms of its formal, semantic and pragmatic features. It attempts to construct a practical mode for examining the features and patterning of the learner phrase, by contrasting it with data from other types of corpora, comparable corpora and parallel corpora included. Several frequently occurring learner phrases are selected for study and related data across Native Speaker English Corpora, Chinese Corpora and Parallel Corpora are extracted for comparison. Feature differences are traced to a variety of sources. And implication for learner interlanguage studies is addressed.

Keywords: unit of analysis, learner phraseology, comparable corpora, parallel corpora, attitudinal meaning

1-2

12月26日，星期六，10:30-11:20
北外中文学院报告厅

何安平

华南师范大学

“过多使用”，“过少使用”，“错误使用”，so what?

摘要: 与目标语本族语者语料库相比较，学习者语料库的研究往往首先发现的是某些语言形式（尤其是词汇形式）的过多使用、过少使用或者是错误使用。至此如何将研究深入下去并取得更有理论或应用价值的结果？这个问题常常困扰着许多刚起步的语料库研究者。本文力

图通过文献考察和个案分析来展示：1) 如何应用语料库语言学的短语理念将表层形式的频数统计延伸至语法结构、语义和语用功能的深层调查。2) 如何继续使用语料库手段对学习中介语特征的成因进行推导和论证。

关键词： 过多使用、过少使用、错误使用、语料库研究、短语理念、中介语研究

1-3

12月26日，星期六，14:00-14:50

北外中文学院学术报告厅

李文中

河南师范大学

Constructing a Learner English Portfolio Corpus on the Open Corpus Platform —A research proposal for diachronic study of the learner language

Abstract: A learner English portfolio corpus is dynamic collection of each learner's English writing production in a diachronic manner and aims at tracing the learner paths towards the goal of English learning. Learner language is on a continuum of development with great variations as language learning itself is much contextualized and individualized. The learner corpora available designed to capture the synchronic features of the learner language fail to present data on the learner language development across time. Therefore it is proposed that a learner English portfolio corpus be constructed on the Open Corpus Platform with the following objectives: 1) to keep a tracing record of the learner language production for dynamic observation of its development; 2) to place each learner on an individual and contextualized evaluative continuum based on the learner progression; 3) to seek individualized solutions to the specific difficulties and problems encountering the learners. Such a portfolio corpus differs from the previous learner corpora by the fact that it is: 1) designed for a specified group of learners in a particular instruction setting by the field teachers; 2) learner specific diachronic text collection; 3) combination of corpus database and management system. The corpus construction is based on the Open Corpus Platform, which we have developed integrating a series of innovative technology in corpus construction and processing. It is believed that the research opens up a new dimension for learner language analysis and enables the researcher to have a diachronic view of learner language progression with individualized focus.

Keywords: learner English portfolio corpus, Open Corpus Platform, diachronic study of learner language

1-4

12月26日，星期六，15:00-15:50

北外中文学院学术报告厅

许家金

北京外国语大学

中国英语专业大学生英语口语叙事中的自我形象 ——词语搭配研究视角

摘要: 本研究通过提取和分析中国学生英语口语叙事话语语料中的框合结构 (congram), 试图挖掘文本背后中国英语专业学生的自我形象。本文首先简述有关“身份认同”和“话语身份”的研究, 特别是其中有关“自我”的相关研究。其次, 本文介绍借助词语搭配开展的话语及话语身份研究。最后, 本文提出了利用框合结构进行话语身份研究的理论框架和分析方法, 并以中国学生的英语口语叙事话语为例探讨中国学生在自述个人经历中所共同构建的自我形象。

关键词: 自我形象、框合结构、英语口语叙事话语、学习者语料库

1-5

12月27日, 星期日, 08:00-08:50

北外中文学院学术报告厅

濮建忠

解放军外国语学院

从词块的使用看学习者语言能力的发展

摘要: 词块在语言研究、语言教学与学习中的地位日益凸显。研究显示, 英语写作能力与词块的使用能力密切相关。本文以词块的使用为切入点, 从多文本和单文本两个视角, 观察我国英语专业大学生英语写作能力的发展情况。从观察的几类词的使用情况看, 这一学习者群体已经具备较好的词块意识, 即他们在写作时使用了大量的词块, 而且许多是地道的。但本研究也发现, 不论是从整体还是从个体来看, 通过将近两年时间的学习, 学习者词块使用能力的发展却不够明显。造成这一现象的原因可能是词块教学还没有有效地在英语教学中展开。

关键词: 词块使用、写作能力、学习者语料库

1-6

12月27日, 星期日, 09:00-09:50

北外中文学院学术报告厅

洪华清

新加坡南洋理工大学

Multimodal Learner Corpus Construction: Challenges and Directions

Abstract: This paper attempts to discuss the challenges and directions of learner corpus research, particularly focused on the construction and application of a multimodal and multiple-layer-annotated corpus of learner language. The construction of a corpus is a significant

step in the field of corpus linguistics, and the development of multimodal corpus database of dynamic texts can provide much more informative resources for evidence-based research on teaching and learning. Unfortunately, there are not many learner corpora of such a kind available in the region and world. On the one hand, the access to a multimodal or multi-layered annotated corpus of learners' oral products is not an easy task, and thus the research interests and foci of most of learner corpus projects have been mostly centered on text analysis of learners' written artefacts. On the other, the application of computer corpus methodology to the study of multimodal texts, for instance, video/audio taped classroom interactions, is comparatively a very new research area (Seedhouse 2004 & 2006; Al-Garawi 2005; Walsh 2006; Baldry 2005 & 2007). To bridge this gap, a list of linguistic and technological challenges has to be answered. Linguistically, we need to identify a set of evaluation criteria that suits best for the adoption of multimodal and multidimensional analyses of learner data. Technically the lack of sophisticated annotation, indexing and querying tools also makes the attempt an expensive and labour-intensive task. In this respect, we need a comprehensive methodology, which can effectively adapt the computer technology for this purpose.

Taking the multimodal corpus of classroom interactions as example, it is well noted that the exercise of analysis of classroom interactional data as speech genres and discourses, we need both empiricist approaches to interaction or behaviour (thus linguistic or paralinguistic features) and functional-communicative approaches of classroom language as conversation (Rojo 2000). With the building of Singapore Corpus for Research in Education (SCoRE) (Hong 2005), we will present how corpus linguistic methodology can be turned into a synergy of these two approaches to pedagogical research as well as multimodal analysis of learner language. In doing so, first we discuss the feasibilities of applying a multimodal corpus approach to learners' language, and identify the significance of multiple-layered annotation in learner corpus research. Next, we demonstrate how to design, annotate, index and query such a corpus. Finally, a list of problems encountered and the answers to the challenges in various stages of corpus processing will also be reported.

Keywords: Multimodal learner corpus, corpus construction, multimodal annotation, SCoRE

References

- Al-Garawi, B. 2005. A Review of Two Approaches to L2 Classroom Interaction. *The Annual Review of Education, Communication, and Language Sciences*. Vol. 2.
- Baldry, A. 2005. A Multimodal Approach to Text Studies in English: the role of MCA in multimodal concordancing and multimodal corpus linguistics. Campobasso: Palladino.
- Baldry, A. 2007. The role of multimodal concordancers in multimodal corpus linguistics. In Terry Royce and Wendy Bowcher (eds.), *New Directions in the Analysis of Multimodal Discourse*. New Jersey: Erlbaum, 173-193.
- Biber, D., et al. 2004. *Representing language in the university: Analysis of the TOEFL 2000 spoken and written academic language corpus*. Princeton: ETS.
- Hong, Huaqing. 2005. SCORE: A Multimodal Corpus Database of Education Discourse. *Proceedings of International Conference of Corpus Linguistics, Birmingham, July 14-17, 2005*.
- Rojo, R. H. R. 2000. *A discursive approach to classroom interactions as speech genres: from heteroglossia and social languages to authoritative discourse*. Retrieved on July 05, 2007,

at: <http://www.fae.unicamp.br/br2000/trabs/1050.doc>.

Seedhouse, P. 2004. The interactional architecture of the language classroom: a conversation analysis perspective. Oxford: Blackwell.

Seedhouse, P. 2006. Classroom Interactions. *Revue Française de Linguistique Appliquée*. 11(2): 111-122.

Walsh, S. 2006. *Investigating Classroom Discourse*. London: Routledge.

1-7

12月27日, 星期日, 14:50-15:40

北外中文学院学术报告厅

梁茂成

北京外国语大学

Keywords Plus: 主题词研究方法扩展及其实现

摘要: 主题词方法是语料库对比中的最常见方法。在语料库对比研究中常常以卡方检验或对数似然率为测量工具, 考察某些词在两个语料库中的使用频率是否存在显著差异, 以发现某种性质的语言或某语言的某种变体具有何种特征。然而, 常见的语料库分析工具仅以单个的词(而非短语)为分析对象, 统计语料库容量时所考虑的也是形符数。随着当今语料库研究对短语和搭配越来越重视, “词”这一语言单位逐渐淡化, 因而以上分析方法所能解决的问题极为有限, 远不能通过对比的方法有效地发现文本间的差异。基于以上思想, 本研究设计并实现了一种新型研究工具 Keywords+, 突破“词”的界限, 将语料库对比中的分析单位扩展到词簇乃至词性赋码序列, 并将短语或赋码序列的对比、显著性计算和索引行析出结合起来, 一键输入, 完成全部分析, 对于有效分析文本特征具有重要意义。研究中借助这一工具, 对学习者的语言特征进行了尝试性的分析。

关键词: 主题词分析、搭配、类联接

小组发言摘要

第一组

2-1-1

12月26日, 11:20—11:40

第一分会场

我国学生英语演讲中的话语语块特点 ——基于语料库的考察

摘要: 本文运用语料库考察中国英语学习者演讲话语语块使用的特点, 重点考察句法和话语立场二类语块。研究发现: 就句法语块而言, (1) 中国英语演讲选手使用4类语块: 多元词、短语架构、习语和句子构造框架, 4类语块的使用频数存在差异, 较少使用习语和多元词, 但后者的丰富度最大。(2) 中国演讲选手的语块使用与英国演讲者相比存在显著差异, 在语块总体丰富度上未见显著差异, 表现出过多使用习语语块的倾向。(3) 大学生与中学生无显著差异, 但大学生的语块丰富度明显高于中学生, 中学生的习语语块使用频数远远超过大学生, 丰富度却较低。就话语立场语块而言, (4) 中国学生采用词汇和句法二种手段表达话语立场, 其中更多使用的是句法手段。表达认知立场主要依靠句法语块; 表达态度立场主要依靠词汇语块; 表达风格立场只采用句法语块。演讲中认知立场和态度立场二类语块占较大比例, 风格立场语块极少。(5) 学生演讲中的立场语块总量超过英语本族语者, 二者未见显著差异。就立场语块的具体类型而言, 中国学生存在过度使用和使用不足并存的特点: 过度使用从句、情态动词和形容词, 过少使用动词和副词。但使用名词、介词短语、状语从句和修辞疑问句与本族语者没有显著差异, 本研究对认识中介语语块发展特征和英语演讲教学具有一定启示意义。

关键词: 句法语块、话语立场语块、英语演讲、语料库

A Corpus-based Study on Syntactic and Stance Chunks of Chinese EFL Public Speakers

Abstract: This paper describes the patterns of syntactic and stance chunks in unprepared public speeches by Chinese EFL learners quantitatively and qualitatively. The results show that syntactically Chinese EFL learners apply four types of chunks, among which sentence builders are most frequent and polywords have highest richness; Compared with native speakers, Chinese EFL learners show no significant difference in formulaic use, except for institutionalized expressions; University speakers are much better than middle school students; In terms of stance, Chinese EFL learners employ more syntactic than lexical stance constructions in their public speech. Chinese EFL speakers used more syntactic constructions of epistemic stance, while they choose a larger number of lexical constructions of attitudinal stance. They express style stance exclusively with syntactic constructions and with very few instances in their speaking. Moreover, Chinese EFL speakers use more stance constructions than native speakers, but show no significant differences in percentage. In terms of stance types, Chinese EFL speakers reveal a pattern of both overuse and

underuse, namely, the overuse of clause, modal verbs and adjectives, but the underuse of verbs and adverbs. There is no significant difference in the use of stance nouns, prepositional phrases, adverbial clauses and rhetorical questions. The findings of the study suggest some important implications for teaching EFL public speaking and understanding interlanguage development in L2 spoken discourse.

Keywords: syntactic chunks, stance chunks, English public speaking, corpus

王立非
对外经济贸易大学

2-1-2

12月26日, 11:40—12:00

第一分会场

基于语料库对中国大学生英语口语中个人立场词块的调查研究

摘要:近年来的许多研究表明,词块在本族语者的口语表达中起着举足轻重的作用,同时,它在二语习得中也起到重要的作用,它能够提高学习者在使用语言时接近本族语者的流利程度和选词的地道性。本文采用中介语对比分析的方法,基于中国大学生英语口语语料库和英国国家语料库的口语部分,比较分析中国大学生英语学习者和本族语者对个人立场词块的使用情况。在此文中个人立场词块指的是用来表达个人感情,想法或者愿望的词块。

研究发现是:(1)中国大学生英语学习者在口语中也使用大量的个人立场词块,而且有时比本族语者还多。学习者也有一定的词块能力。(2)但是,学习者与本族语者所具体使用的个人立场词块的类型和频率却不尽相同。(3)学习者少用一些能传达礼貌信息的词块,使得学习者在表达个人立场时趋于明显和直接。(4)学习者少用一些能起到人际交互作用的词块,致使学习者的对话缺乏应有的交互性。(5)学习者根据表达的需要创造出许多本族语者少用或不用的个人立场词块。

关键词:个人立场词块、语料库、中介语对比分析

柴静
东北林业大学

第二组

2-2-1

12月26日, 11:20—11:40

第二分会场

学习者英语学术论文语料库建设及应用

摘要:发表英语学术论文是当今世界传播、交流新科学知识的重要途径。英语学术论文的数

量也成为衡量一所大学乃至一个国家科学研究水平和实力的重要指标。然而,对于许多中国学者(也可以称为高级英语学习者)来说,用英文进行学术论文的写作仍然是一个极具挑战性的工作。怎样帮助英语为非母语的中国学者提高英文论文写作水平是摆在英语教师面前的一项重大研究课题。目前,国内虽然已有一些致力于帮助中国学者提高学术论文写作水平的教材和软件,但这些成果大多基于对专家语料的研究,普遍忽视了对学习者产生的学术论文手稿所存在问题的分析研究。众所周知,对学习者学习需要的了解和所存在问题的诊断是课程设计的重要环节。我们认为,只有通过对学生英语论文写作中出现的各种语言问题进行深入系统的研究,才有可能提出富有针对性的解决方案,提高学术英语写作教学的效率及中国学者的英文论文写作水平。

本文旨在介绍还处在建设初期的中国学习者的学术英语语料库,包括其语料来源、设计结构及错误分类体系。该语料库的错误标注尝试包含汉式英语及语篇体裁层面的问题。汉式英语分为形式为主和意义为主两大类;语篇体裁层面的问题包括 *Unrealized purpose*, *Inadequate discussion*, *Irrelevant literature*, *No comments on results* 等。对学生汉式英语及语篇层面所存在问题的标注在已有的学习者语料库中还非常少见,因此该语料库的建立及相关研究的开展对系统地深入地研究中国学习者英语论文写作和相关课程的设计具有非常重要的意义。

关键词: 中国学习者、英语学术论文、语料库建设及应用

杨瑞英、洪华清
西安交通大学

2-2-2

12月26日,11:40—12:00

第二分会场

中国英语专业学士学位论文语料库建设构想

摘要: 最近的数十年见证了英语专业和英语学科的飞速发展。目前,全国的英语本科专业点多达 900 多个,属于国内大学扩招较多的专业之一。(戴炜栋、张雪梅 2007: 4) 2004 年教育部组织专家对国内 4 所院校进行试评估发现“毕业论文撰写的质量及其中教师的指导均不理想”,“毕业论文成了各高校的头等事……”(戴炜栋、张雪梅 2007: 5-8)

本课题组通过 CNKI 数据库文献检索发现:国内自 1979 年至今的 30 年间,有 48 篇公开发表的学术论文探讨了我国英语专业学位论文/毕业论文问题(主要在近 10 年)。其中有 27 篇是关于毕业论文现状调查分析并提出要加强管理和指导策略的,占 56.3%。有 5 篇谈到利用网络资源来指导学生毕业论文(谈网络资源的基本运用),占 10.4%,只有 3 篇是利用语料库方法分析毕业论文标题、内容摘要和连接副词等内容,仅占 6.25%。很显然已有研究对有 900 多个专业点、每年几十万毕业生且有 3 类生源的中国来说,其代表性值得怀疑。

随着计算机技术迅速发展,学习者语料库研究正在崛起,这种以概率和频率为基础的二语习得研究的全新方法,拓宽了应用语言学研究的视野,为外语教学研究提供了一种崭新的哲学思维方式。(文秋芳等 2005: 1) 英语专业学位论文属于学术论文,与基础写作有很大差异。论文写作的每个部分(摘要、关键词、引言、正文、结语、致谢、参考文献等)都是学位论文的有机组成部分,对学生来说都是十分重要的。在库容为上千万篇的 CNKI 数据库只能找到如此少量相关文章,这说明研究还远远不够。

建立英语专业学士学位论文语料库将填补我国英语学习者语料库的一项空白。此语料库一旦建成,研究人员就可以开展基于语料库的实证研究,多层面、多纬度地深入研究学习者学术语言特征和发展模式,丰富语言学理论和二语习得理论,提出有针对性的、有利于提高我国外语教学质量,特别是英语学位论文质量的建议。对今后的英语专业教学大纲修订和英语专业的学位论文/学术论文写作等教材的编写等都有参考价值;也为教育部正在进行的高等教育英语专业评估“毕业论文”部分提供更加翔实的数据;对英语专业学位论文指导教师也有参考价值。

本课题组基于以上分析拟(1)建立“中国英语学士学位论文语料库”(中型,历时性)。计划首先建立1500篇(约6百万词)学士学位论文语料库。其中从一类、二类和三类本科学校各选5所,每所学校分别随机抽300篇(其中2005年、2006年、2007年、2008年、2009年各抽60篇);为便于做对比研究,拟另外收集以英语为母语国家学生学士论文200篇。总计1700篇。(2)开展基于语料库的实证研究,多层面、多纬度地深入研究学习者学术语言特征和发展模式,丰富语言学理论和二语习得理论。难点是想办法处理好学位论文中的图片、图表、符号等非文字内容。因为目前语料库检索工具还无法处理这类文件格式。手工标记文本内容并不能保证其准确的还原,可以考虑采用国际主流用于数据交换的标记技术XML来存储信息,XML具有可扩展性、开放性、通用性和语义结构化而广受数据存储和交换界的青睐;此外也可以考虑结合关系型数据库建立语料库文本格式文件与原始学位论文文件的一一对应,辅助语料库来完成数据的检索和查询。

英语专业学士学位论文问题多,质量达不到要求。建立语料库便于研究的进一步深入,给目前英语专业学位论文一个比较准确的定位,找准学生学术论文写作的“最近发展区”,为本科专业评估提供基本依据,为提高英语专业本科教学质量服务。

关键词: 英语专业、学位论文、语料库、构想

张金福
九江学院

第三组

2-3-1

12月26日,11:20—11:40

第三分会场

认知视角下的小学生语文作文话语分析

摘要:在我国第一语言习得研究相对于第二语言习得研究要滞后许多,第一语言习得研究者多数为小学语文教师和小学教育学者,研究缺乏理论指导和实证性数据支持,因此第一语言习得研究多数为经验式的教学总结性研究。本研究以话语分析为理论基础,从认知语言/义学视角,以中国地质大学附属学校小学三年级学生150篇记叙文为分析语料,用AntConc对该语料进行检索,以期了解小学生写作认知图示和心智特点。

本文分为两部分:(1)根据语篇/话语分析方法,对作文的内容、文字结构、字数、词类进行简单数据分析,得出小学生作文的普遍特点。(2)着重从认知角度分析语篇背后的认知特点。即概念化—词法/用词,突显—句法结构,动态分析—语篇连贯,心理空间—篇章

结构。

本研究旨在从表面意义中寻找学生幕后认知，幕后认知决定了语言形式的意义。作者认为小学生在记叙文体的写作中其心理空间仍然是线性的故事空间，还未上升到知识空间和想象空间。小学语文教师如何激活学生的知识空间和想象空间是小学生记叙文体作文提高的关键。

关键词：小学生语文作文、话语分析、认知视角

张红燕

中国地质大学（武汉）

2-3-2

12月26日，11:40—12:00

第三分会场

基于语料库的通用大学英语教材对比研究

摘要：本文在回顾了国内外教材评估研究的现状和我国大学英语教材评估之困境的基础上，尝试以《大学英语课程教学要求》为参照彼岸准，才用基于语料库的定量与定性相结合的方法，从词汇、多词单位和题材三个方面对我国高校通用的四套大学英语教材进行对比分析。研究发现，四套教材在上述三个方面不同程度上存在欠缺，难以保证学生获得学习目标语言的资源。本文期望能为高校大学英语教师及其学习者根据自身特点和目的选择适合的大学英语教材提供参考。

关键词：大学英语教材、语料库、《大学英语课程教学要求》、教材评估

刘泽权、刘艳红

燕山大学、燕山大学里仁学院

第四组

2-4-1

12月26日，11:20—11:40

第四分会场

基于 SECOPEETS 语料库的学习者英语口语失误分析

摘要：以新建《公共英语等级考试口语语料库》为基础，本研究对中国学习者英语口语中言语失误进行了调查分析。结果发现，1) 中国学习者英语口语中，失误频数较高的是动词词组、名词词组、句子和介词;2) 代词和介词失误明显多于书面语;3) 口语言语失误可划分为句法性失误和语义性失误两大因子，不同于书面语的是，句法性失误明显多于语义性失误;4) 学习者英语水平和言语失误类型之间呈现明显的对应关系，5) 母语迁移是口语中言语失误的重要原因。

关键词：英语口语语料库、失误类型、分布型式、错误成因

A SECOPEETS-based Study on Speech Errors in Chinese EFL Learners' Spoken English

Abstract: Based on the Spoken English Corpus of Public English Test System (SECOPEETS), the present study investigated the errors in Chinese EFL learners' spoken English. Results indicate that 1) the frequent errors in Chinese EFL learners' spoken English are related to verbs, nouns, sentences and prepositions, 2) the errors related to pronouns and prepositions are more frequent in speech than in writing, 3) all the errors are syntactic and semantic in nature; different from learners' written language, the weight of the syntactic errors is greater than that of the semantic errors, 4) a correspondence is observed between the learners' English proficiency levels and speech error types, 5) language transfer from mother tongue is one of the major causes for speech errors.

Keywords: spoken English corpus, error type, distribution pattern, causes for errors

邓耀臣、罗卫华
大连海事大学

2-4-2

12月26日, 11:40—12:00
第四分会场

基于语料库的中国英语专业学生口语 *there* 存在句使用研究

摘要: 由于语法结构和语篇功能的特殊性, *there* 存在句受到学者们的广泛关注。然而, 多数研究都是凭借内省和直觉展开对 *there* 存在句的理论研究, 这些研究成果因未得到真实语料的支持而缺乏可靠性和有效性。此外, 相关研究关注的几乎都是书面语, 忽视了该结构在口语中的大量使用。

本文以中介语对比分析和计算机辅助错误分析为理论基础, 采用语料库语言学研究方法, 基于三大语料库—SWECCCL, BASE 和 FLOB, 研究中国英语专业学生对 *there* 存在句在口语和笔语中的习得和使用情况。

经量化和质化分析本文得出以下结论:

1) 总体上, 中国英语专业学生已经掌握了 *there* 存在句的使用。语料显示中国英语专业学生不仅能正确使用 *there* 存在句的典型结构而且还可以灵活地运用其变体。

2) 研究显示: 与本族语者相比, 中国英语专业学习者倾向于使用简单 *there* 存在句结构, 较少使用复杂的句型; 中国英语专业学生有过度使用 *there* 存在句的现象。

3) 中国英语专业学习者没有语体意识。在口语中, 像 *there is, there are* 等正式的表达形式远远多于 *there's, there're* 等的缩略形式, 而在书面语中, 中国英语专业学习者大量使用缩略形式。

4) 中国英语专业学习者倾向于犯三种错误。(1) 中国英语专业学习者容易将 *there* 存在句和表示“有”的 *have* 混用; (2) 主谓一致问题; (3) 由于语音和拼写的相似性, 中国英语学习者容易将 *there, their* 和 *they* 混淆。

关键词: *there* 存在句、语料库、口语、笔语

陈佳
中山大学

第五组

2-5-1

12月26日, 16:00—16:20
第一分会场

基于学习者英语语料库的扩展语义单位研究

摘要: Sinclair (1996) 提出扩展语义单位 (又称词汇项) 这一概念, 并总结出相应的模型。本研究以扩展语义单位模型为框架、利用 CLEC (中国学习者英语语料库) 和 ICLE 中的 LOCNESS 子库, 对核心词 happen、take place 和 occur 的搭配类、搭配词语义选择趋向和语义韵分布进行比较分析, 力图探讨中国英语学习者书面表达中扩展语义单位的分布特点、发展模式以及各层面之间的关系。研究发现: 1) 各类学习者都掌握了 happen 和 take place 的常用类联接, 但中国学习者很少使用 occur、普遍出现语义韵混淆现象; 2) 中国学习者从高中、大学低年级到高年级阶段在搭配词的丰富程度与语义韵和谐程度上呈上升趋势。但与英语本族语学习者之间存在显著差异; 3) 扩展语义单位的各个层面相互限制、相互依赖。中国学习者搭配词选择能力不高、倾向于按照字面意思直译, 构成影响其英语表达语义韵和谐的主要因素; 4) 学习者语言搭配词分布离散性大, 需要对提取方法进行个性化设计。文章最后讨论了上述问题产生的根源以及本研究对英语教学的启示。

关键词: 语料库、学习者英语、扩展语义单位

陆军
扬州大学

2-5-2

12月26日, 16:20—16:40
第一分会场

“CAUSE” ——基于语料库的中国学习者学术写作词汇能力研究

摘要: 本文从东南大学自建高级学习者学术英语写作语料库 CAWNNS (Corpus of Academic Writing of Non-Native Speakers)¹ 以及自建本族语者学术英语语料库 CAWNS (Corpus of Academic Writing of Native Speakers)¹ 中抽取相应语料, 构成两个组成相似、大小接近的对比语料库, 以 “CAUSE” 为例, 从类连接、搭配、语义韵三个角度对中国学习者学术写作中词汇能力进行了对比研究, 结果如下: (1) 学习者能够掌握大部分类连接, 但受母语影响, 在 CAUSE 使用中出现了 V-V/V-V-ing 和 V-CLAUSE 错误类连接 (2) 本族语者最常使用的

¹本研究所用语料库由东南大学外国语学院张萍博士创建, 容量分别为 276,993 字和 17,498,810 字。

典型搭配词中，只有约 31%出现在学习者搭配词中，且 MI 值有明显差异，同时学习者搭配词中有错误搭配词出现（3）本族语者和学习者在 CAUSE 的语义韵上没有显著差异，均表现出消极语义韵。本研究结果对提高学习者学术写作能力有一定的参考意义。

关键词：词汇能力、类连接、搭配、语义韵

“CAUSE”: A Corpus-Based Study on Vocabulary Usage in Academic Writing of Chinese EFL Learners

Abstract: This paper aims to investigate Chinese EFL learner's lexical competence of "CAUSE" in academic context through two contrastive corpora based on CAWNNS and CAWNS. The findings are as follows: 1. Learners can basically use colligation properly, and misused colligation V-V/V-V-ing and V-CLAUSE can be found; 2. 31% of typical collocates retrieved from native corpus are used by learners, though with different MI values. Atypical and inappropriate collocates are found in learner corpus; 3. There is no significant difference between learners and native speakers in terms of semantic prosody.

Keywords: lexical competence, colligation, collocation, semantic prosody

何晶、张萍
东南大学

2-5-3

12月26日，16:40—17:00

第一分会场

中国英语学习者写作中副词与形容词搭配的使用特点

摘要：搭配能力对语言水平的提高具有重要意义，尤其是对语言水平和搭配使用能力较低的外语学习者。因此对于学习者写作中搭配特点的研究有助于帮助学习者发现自身问题从而采取一些有效的措施提高搭配运用能力并最终实现语言水平的提高。

针对中国英语学习者写作中副词和形容词搭配使用的特点，本研究利用 LOCNESS 和 CLEC 的子语料库数据着重回答三个问题：1) 非本族语者和本族语者之间在副词和形容词搭配的使用方面是否有显著性差异？2) 中国英语学习者对副词和形容词搭配的使用特点是否与其语言水平有显著关联？3) 中国英语学习者搭配中的主要问题是什么？

研究表明，中国英语学习者对副词和形容词搭配的使用显著不同于本族语者。1) 学习者对搭配的使用基本上局限在比较狭窄的范围内，对位于前十位高频搭配词的使用过于频繁，其使用频率显著高于本族语者。2) 通过对副词使用情况的进一步观察发现中国英语学习者过量使用封闭性副词，占总数的 90%以上，从而过少使用了那些语义上更加灵活多变的开放性副词。3) 这种过多使用高频搭配词和封闭性副词的总体趋势并没有随着学习者写作水平的提高而得到改善。通过对词条的逐条检查发现中国英语学习者写作中的副词和形容词搭配在形式和语义两个方面都存在一些问题，具体表现在：形容词比较级 more 的使用，词汇拼写，高频副词 very 的过分强调，quite 和 rather 的混用，以及 too 的不恰当使用等。这些具体的问题使其表达不够不地道。

关键词：副词、形容词、搭配特点、中国英语学习者、语言水平

Adverb-adjective Collocation Features in Chinese EFL Learners' Writing

Abstract: Collocation is important to language proficiency, especially to foreign language learners who are usually weak at using the appropriate collocations. Studies on collocation features in learners' writing might be of great value to help learners realize the non-native problems in their language so that some possible strategies might be employed to overcome the difficulties in collocation learning and eventually improve the language proficiency.

With a purpose to find some typical features of Adv-adj Collocations in Chinese EFL learners' writing, this study makes use of sub-corpora from LOCNESS (Louvain Corpus of Native English Essays) and CLEC (Chinese Learners of English Corpus) as data sources in an attempt to answer three questions: 1) Is there any significant difference between Chinese EFL learners and native speakers of English in terms of the features of Adv-adj Collocations? 2) Is the learners' use of Adv-adj Collocations related to their language proficiency? 3) What are the major problems of Adv-adj Collocations in Chinese EFL learners' writing?

Results reveal that Chinese EFL learners' use of Adv-adj Collocations is significantly different from that of the native speakers. 1) The learners are mostly confined within a narrower range of collocations and heavily rely on the top ten frequently used items. 2) A further investigation into adverbs as adjective modifiers reveals that Chinese EFL learners overuse the closed-class adverbs, with a high percentage of more than 90%. The semantically more variable open-class adverbs are on the contrary significantly underused. 3) The general tendency of overusing the high-frequency items and closed-class adverbs does not get improved for learners at a higher language proficiency level. In addition, an item-by-item examination of matches in the data reveals that collocation problems in Chinese EFL learners' writing involve both form and meaning.

Keywords: adverbs, adjectives, collocation features, Chinese EFL learners, language proficiency

何影
河北大学

2-5-4

12月26日, 17:00—17:20

第一分会场

专业物流英语词块特色研究

摘要: 自20世纪90年代以来,以词汇为中心的教学法受到国外语言学界的广泛重视,由此也掀起了词块研究的热潮,研究者们不仅意识到单词在阅读中的重要意义,同时也意识到词块(chunk)在文本分布特点及其在语篇中的重要意义。“词汇教学的内容早已超出了单词的范围还应包括短语(phrase)、习语(idioms)”(Willis 1990; Lewis 1993, 2000; Wray 2000; Gavilio 2005; Biber 2006; Anne Okeeffe, 2007等)。在此热潮影响下,国内语教学界也从理论和实践两方面开展了一些研究(丁言仁 2005; 濮建忠 2005; 卫乃兴 2007; 许家金、许宗瑞 2007; 郑晶 2008; 刘晓玲 2009等)。相对国外研究而言,国内词块研究多注重通识英语教学实践和理论探讨,而在专业英语层面的研究并不多见,少数几篇:李太志(2006),词块在外贸

英语写作教学中的优势及产出性；宋淑敏，2008 词块教学法及其在商务英语信函教学中的应用），基于专业英语教材语料开展的词块研究，几乎凤毛麟角。有鉴于此，本研究将以词块教学理论为理论依据，运用语料库调查方法，从 4 字语入手，对专业物流英语教材高频词块作质性和量性分析，以期在物流英语教学和学习上获得启示。

本研究将对立信会计出版社 08 年版《专业物流英语》中的短语词块做调查分析。其目的是调查分析该套教材中关于物流主题的词块在教材文本中的分布位置及其相应语篇功能。

研究问题：1) 在教材语料中是否有多次反复出现的 4 字语？如果有这类词块出现，它们的频率是高还是低？2) 高频词块在教材语篇中有意义吗？它们的文本分布规律和语篇作用又如何？

调查方法将以小型教材语料库调查为主。语料由两部分组成：教材语料《专业物流英语》和专业学术参照语料（学术书籍节选文章和期刊杂志专业学术论文）。小型语料库调查方法中强调基于频数分析法(a frequency-driven approach)和关键词观察分析法（KW analysis approach）。其中频数分析法帮助研究者判断被调查词块出现频率，关键词分析法则有助于研究者对高频词块的文本分布位置和语境做出观察和分析，从而判断其相应的语篇功能。采用的研究工具：AntConc3.2 软件（免费工具）和 Concgram 1.0 软件（正在购买中）。前者对词频和关键词分析很有帮助，后者则有助于观察词块不同搭配范式并判断其功能。判断词块语篇功能的理论依据是 2006 年 Biber 对大学课堂教学和教材的词块研究中提出的关于学术语篇词块功能分类标准。根据这一分类标准，学术语篇词块有三类篇章功能：1) 表写作者立场；2) 有语篇组织功能；3) 表指代关系。运用以上方法和理论依据，研究者将对两类语料中出现频数高且有语篇意义的 4 字语作对比分析，以揭示物流英语教材这一特定语域学术语篇中，出现频数高且有语篇意义的词块文本分布特点及其语篇功能，并希望对物流专业英语教学和学习提供指导意义。

关键词：词块、专业物流英语、4 字语、小型教材语料库

胡玉迎

广东外语外贸大学

2-5-5

12 月 26 日，17:20—17:40

第一分会场

语料库在同义词搭配习得过程中的应用研究

摘要：提高学生自主学习能力，是大学英语教学模式改革的新亮点。语料库的兴起和应用研究推动了大学英语自主学习的发展。词汇作为二语习得的中心任务，词汇习得水平是判断学习者语言整体水平的一个不可忽视的重要指标。在此指导思想下，英语教学研究重点开始从如何拓展学生词汇的广度向关注词汇发展的深度转变。然而，如何提高学生词汇习得的效率，特别是如何更准确的对同义词进行辨析并精确使用，一直被认为是在词汇深度学习过程中的一个难点。

本研究以语料库为基础的数据驱动学习为理论基础，以同义词搭配习得实验为量化研究数据支撑，探究如何将同义词辨析习得应用于大学英语词汇的自主学习。该研究以词义、语义韵为搭配考察角度对目标词进行选定，且实验设计中不仅关注了实验班与参照班即时测试的效果，对延时测试也进行了详尽的对比分析。研究发现数据驱动学习不仅仅可以提高学习

者自主学习的兴趣、有效的提高学习者词汇习得质量，并具有较强的记忆优势，在大脑中的长时记忆保存更加完整与清晰。因此，该研究不是单纯的对数据驱动学习应用的验证性研究，而是对如何将语料库引入课堂教学的探索性研究。

随着计算机网络的高速发展，运用先进的网络资源与技术提高学生的自主学习兴趣与能力已经成为现实。语料库资源庞大，有真实性的优势，且经研究大量真实语言输入有利于语言的产出，所以将其引入课堂是非常必要的。但是如何有效的选取语料库资源运用于教学实践尚处于研究阶段。本文旨在探究语料库应用过程中的特点及应用方法，从而有效提高学习者英语学习能力。

关键词：数据驱动学习、同义词搭配、自主学习

Applied Study of Corpus on the Collocation Acquisition of Synonyms

Abstract: Corpus linguistics is a new area which seems to be developing at quite an amazing speed in the decades. The development of corpora leads to a new approach to language description and has significant pedagogical consequences. As the application approach of corpus, Data-driven learning (DDL) promotes the development of independent learning.

This paper focuses on the collocation acquisition of synonyms. It's quite a problem for the second language learner to acquire collocation knowledge in the language learning process. Learners with a large vocabulary still have problems— they may know a lot of words, but their collocational competence with them is very limited. “Collocation runs through the whole of the English language. No piece of natural spoken or written English is totally free of collocation” (Crowther *et al.* 2001: vii). In hence, this paper illustrates the effectiveness of DDL in the process of promoting collocation acquisition.

From the aspect of methodology, the present thesis, choose 3 pairs tests words from the synonym test by use of CLEC. Then put into three groups of pedagogical practice. Through comparison work, testify the effectiveness of concordance in the depth of lexicon development. The powerful statistical software SPSS has been used to process data. The statistical techniques employed for analysis are some hypothetical tests to confirm assumptions such as T-test, one way ANOVA and some descriptive statistics to draw a general picture of collocation. From the statistics, we can summaries that DDL can not only raise the interests of lexical learning but also keep a long time memory and high expertness in the learners mind.

As the development of computer technology and statistical development, corpus will become a new learning tool of second language learners. The priorities of corpus like abundant real materials will give a great help in the procedure of language teaching and learning.

Keywords: Data-driven Learning, collocation, synonym, independent learning

黄瑾、崔艳嫣
聊城大学

2-5-6

12月26日，17:40—18:00
第一分会场

论机切口语词块的元话语性

摘要: 本文通过引述相关文献探讨机切口语词块的元话语性。在理论上, 语用化理论为机切口语词块的元话语性提供了依据; 在实证证据上, 已有研究表明机切口语词块在形式、语义和功能上富有元话语特征。机切口语词块的元话语性使它们成为交际的互动单位, 它们的使用可以给信息接收者的话语理解提供明示的语用导向, 使话语具备最佳关联性。机切口语词块的元话语性对我国的英语教学具有重要启示。

关键词: 机切口语词块、元话语、语料库

On the Metadiscursivity of Computer-Retrieved Oral Chunks

Abstract: The present paper discusses the metadiscursivity of computer-retrieved oral chunks (CROCs) based on existing literature. Theoretically, the theory of pragmaticalization renders support for the metadiscursivity of CROCs; and empirically, previous studies have shown that CROCs are characterized with metadisocurse forms, meanings functions. The metadiscursivity of CROCs makes them units of interaction, and their use gives the audience explicit pragmatic guidance to the correct interpretation of the utterance, thus implementing the presumption of optimal relevance. The metadiscursivity of CROCs has important implications to ELT in China.

Keywords: computer-retrieved oral chunks, metadiscourse, corpus

林维燕
暨南大学

第六组

2-6-1

12月26日, 16:00—16:20

第二分会场

一个社群、三类语境：中国学者期刊论文摘要的多维度体裁分析

摘要: 作为全文精华的论文摘要能够让作者建构专业形象、展现相关的专业背景及推理能力, 同时提供读者概略理解方向、激起专业认同及参与沟通的意愿 (Hyland 2000; Swales & Feak 2009; 刘贤轩 2003)。由于信息爆炸、科技交流高度国际化与同侪竞争加剧更提升了该类 (亚) 学术体裁的重要性并促进了其促销性特征的形成。根据体裁分析的探讨, 论文摘要的言步 (move) 结构一般包括: 背景、目的、方法、结果、结论。非英美学者受制于语言障碍、个人写作经验和国内外学术话语社群的规约差异, 故此在摘要写作时可能偏离主流学术范式, 常遭遇如语句长度、过渡词语使用、时态/语态选择、人称突显与否、言步结构缺省等多种困难 (Martin-Martin 2003; Van Bonn & Swales 2007; 刘贤轩 2003)。以往的中外以及其他非英语国家同英美期刊论文摘要的对比分析多为传统的对比修辞研究, 强调宏观言语社群之间的差异, 忽视了现代专业化社会里亚文化、体制化实践 (如同一言语社群下的不同出版语境) 等相对中观的交际元素。

中国材料科学家是中国科学家群体中相对国际化程度最高的学术社群,近十年来被 SCI 收录论文 70,353 篇,居世界第二位;同时,他们还是内部高度异质化的中国科学家群体,参与多重出版语境(如国内中文、国内英文、国际英文)已成为该社群的常态实践。但以上三类语境的体制化实践却在社群规模、交际强度与认同程度、物质丰沛度等方面差异明显,论文摘要的话语表现也会因此随之有所变化。

本研究选取中国材料科学家三类出版语境以及英美科学家期刊论文摘要各 20 篇(合计 80 篇),同时结合基于话语的焦点专家访谈(discourse-based interviews of specialist informants),进行多维度的对比体裁分析。初步研究结果表明,中国科学家在国际英文语境中的论文摘要总体上更接近英美文本,且由于材料科学的学科特性,‘方法’言步的出现率在中外科学家的四类文本中并无差异。但国内的中英两个语境很少出现‘背景’言步以及显性的作者存在,且国内英文语境在时态和语态上更为依赖传统认知的过去时和被动态。需要指出的是,多个焦点专家的基于话语的访谈反映了机制化实践中不同目标语境与或然性因素(contingency)对摘要写作与发表过程的重要影响。

关键词: 中国学者、期刊论文摘要、出版语境、多维度体裁分析

One Discourse Community, Three Publication Contexts: A Multidimensional Genre Analysis of Research Article Abstracts by Chinese Academics

Abstract: The genre of research article (RA) abstracts has been playing an important dual role (i.e., informative and promotional) in contemporary academia, which is even strengthened by current information explosion, academic globalization and fierce competition among academics (Hyland 2000; Swales & Feak 2009). However, non-Anglophone academics are likely to deviate from mainstream paradigms of writing abstracts mainly owing to their language barriers, personal writing experience, and existing differences between the practices of local and international discourse communities. Previous literature comparing Chinese or other non-Anglophone RA abstracts with those of Anglophone scholars (Martin-Martin 2003; Van Bonn & Swales 2007; 刘贤轩 2003) were mostly traditional contrastive rhetorical studies, stressing disparities between macro speech communities but neglecting such meso communicative factors as small cultures and institutional practices (e.g., different publication contexts within a speech community).

Chinese materials scientists are a special disciplinary community which enjoys highest globalization and heterogeneity as compared with other scientists in China. Amongst their regularly visited three target publication contexts such as local Chinese, local English and international English, differences in their community size, interaction intensity, identification degree and materiality will accordingly lead to variation in discursal representation of RA abstracts.

Our research aims to conduct a multi-dimensional genre analysis on RA abstracts of Chinese materials scientists by compiling a corpus of 80 RA abstracts (20 each in the three contexts of Chinese scientists and in Anglo-American scientists' international English texts). This corpus analysis is supplemented by discourse-based interviews of specialist informants. Our preliminary results show that, ENG texts of Chinese scientists were mostly close to NSE texts of their Anglophone colleagues; the disciplinary nature of materials science resulted in no significant differences in ‘Method’ move among the four contexts of publication. Also, ‘Background’ moves and explicit authorial presence seldom occurred in the two local contexts of Chinese scientists, while simple past tense and passive voice were most heavily used in the local English-medium

context. It is to note that, several discourse-based interviews indicated the distinct compact of various target publication contexts and contingencies in institutional practice upon the writing and publishing of RA abstracts.

Keywords: Chinese academics, research article abstracts, publication contexts, multidimensional genre analysis

黄大网、徐亮、杨廷君

宁波大学

2-6-2

12月26日, 16:20—16:40

第二分会场

基于语料库的中国英语变体研究

摘要: 中国英语的客观存在及其研究必要性已经得到国内不少学者(葛传槩, 汪榕培, 李文中, 谢之君, 贾冠杰, 向明友等)的肯定, 也有从中国英语变体与文化认同的关系进行论述(潘章仙, 《中国英语变体中的语言和文化认同》2005), 但基于语料的实证研究比较少见。本研究把中国英语分为四类, 第一类是中国名家的英语作品, 比如林语堂、辜鸿铭的英语作品; 第二类是经典汉译英翻译作品, 里面包括以汉语为母语或者以英语为母语的译者进行的中翻英译著; 第三类是国内英语时事报道杂志和期刊上英语文章; 第四类是中国英语学习者的练习作品。将以上四类按一定比例抽取, 制作成语料, 应用语料检索方法, 研究中国英语在词汇选择上的比例, 并将之与标准英语语料库进行对比, 从中得出一些规律, 找出中国英语变异特点及其趋向, 并进一步分析产生此种变异的深刻原因。限于人力财力, 本研究仅集中于词汇选择上的语料对比研究, 对于在语音、语义、搭配和语法等方面, 留待后续研究。
关键词: 中国英语、语料库、词汇选择

邓海龙

赣南师范学院

2-6-3

12月26日, 16:40—17:00

第二分会场

语体研究与研究生科技论文英文摘要写作教学

摘要: 科技论文摘要语篇属于正式书面语体, 语篇结构、句式和词汇都要求体现正式语体意义, 符合科学推理逻辑。英语论文摘要写作是研究生英语写作教学的重点。本科大学英语教学注重语法和词汇, 学生语体意识欠缺, 在具有鲜明学术语体特征的科技摘要写作中, 容易出现语体错位现象, 这将严重影响摘要的效果。词汇、句子和语篇等都是语体意义标记, 是语体意义的具体体现成分。本文基于自建的 SCI 和学生论文摘要语料库, 对 SCI 摘要和本科生摘要从语篇图式结构、句式和词汇方面进行了语体对比分析。结果显示学生摘要信息不

完整，内容不具体，偏重于研究背景及知识性介绍，有的无结果，有的无结论，随意性较强；在句子层面上不太会使用隐喻表达式，句子不够书面化，有碍于论证展开的逻辑性、严谨性及客观性，语言也欠简洁；使用核心词汇较多，对学术词汇的掌握欠佳，专业词汇和低频词缺乏。所以提高学生在写作中的语体意识，使其掌握宏观语篇图示结构、隐喻式句子表达法以及一定量的学术词汇，避免语体错位，是研究生英文摘要写作教学的重点。

关键词：语体、图示结构、隐喻式、学术词汇、科技论文英语摘要

Stylistic Research and Graduate Science English Abstract Writing

Abstract: Problems in undergraduates' science English abstracts are studied from stylistic perspective. SCI and student abstract corpora are established and a comparative research is done in schematic structure, sentence structure and vocabulary. The result shows that student abstracts lack necessary and concrete information as well as academic words, with spoken sentences in poor logic, conciseness and objectivity. The conclusion is that improving students' stylistic sense in writing and avoiding stylistic confusion within a text are the focus in postgraduate English abstract teaching.

Keywords: stylistic, schematic structure, incongruent, academic words, science English abstract

江淑娟

中国石油大学（北京）

2-6-4

12月26日，17:00—17:20

第二分会场

多模态信息检索工具 ——MCA 及其在外语教学中的应用

摘要：随着多媒体信息技术的发展，语料库中的信息呈现多模态的特点。因此，传统的文本检索工具，例如 WordSmith 已经不能够充分满足教学和研究的需要。目前，在国外 SF-MDA 研究界推出了多款多模态信息检索工具，并在语言教学中得到了应用。本文将介绍 MCA 的语言学原理、功能，并说明如何利用 MCA 开展外语教学的设计。

关键词：多模态、信息检索、MCA、语言教学

李学宁

江南大学

2-6-5

12月26日，17:20—17:40

第二分会场

理工科大学生英语口语多模态学习者语料库的构建研究

摘要：语料库的建立和研究始于上世纪后期，已成为语言学的热门领域。随着英语学习者国际语料库（ICLE）的建立，语料库已正式步入第二语言习得的研究领域。在口语语料库方面，较有代表性的是英国国家语料库（BNC）中的口语部分以及国际英语语料库（ICE）和 London-Lund 语料库。我国较有代表性的学习者口语语料库是“中国学习者英语口语语料库”（COLSEC）以及“中国学习者英语口语语料库”（SECCL）。但现有口语语料库的基本功能是基于转写文本来分析言语错误，对在日常交流中扮演重要角色的非言语交际的标注、符码和分析却鲜有涉及。本研究基于话语多模态分析的理论，尝试构建我国理工科大学生英语口语多模态学习者语料库，从语音、词汇、句法、篇章、非言语交际等多个维度提炼标注指标，以期对该类学生的口语能力作全方位、多层次的剖析，为英语口语课堂教学和学生课后自主学习提供有价值的反馈。

关键词：理工科大学生；英语口语；多模态；学习者语料库

The Construction of a Multimodal Learner Corpus of Oral English for Chinese Science and Engineering Majors

Abstract: The establishment of and research on corpora started in late 20th century and has since then become an important field in linguistics. With the construction of International Corpus of Learner English (ICLE), research on corpora has been recognized as a section of Second Language Acquisition. The representatives of oral corpora are the Oral Session of British National Corpus (BNC), International Corpus of English (ICE) and London-Bund worldwide compared with College Learners' Spoken English Corpus (COLSEC) and Spoken English Corpus of Chinese Learners (SECCL) in China. However, the basic function of current oral corpora is to analyze linguistic mistakes based on transcripts. Annotation, tagging and analysis of non-linguistic features, which play an important role in daily communication, have seldom been touched upon. The authors of the present paper attempt to construct a multimodal learner corpus of oral English for Chinese Science and Engineering Majors based on multimodal discourse analysis. Traits from various dimensions such as pronunciation, vocabulary, syntax, discourse coherence, and non-linguistic features are distilled as annotation parameters to analyze the oral proficiency of this repertoire of students in an all-round way. The authors sincerely hope that by doing so, valuable feedback can be provided to enhance oral English classroom instruction as well as students' after-class learning autonomy.

Keywords: Science and Engineering Majors, oral English, multimodality, learner corpus

刘芹、潘鸣威

上海理工大学、上海外国语大学

2-6-6

12月26日，17:40—18:00

第二分会场

英语专业写作教学语料库的建设与应用研究

摘要: 英语写作教学语料库是以学生的考试和平时作文为内容,以建库者的后期标注为参考,以教师课堂写作教学和科研为目的的专门用途语言教学语料库。该语料库的建设可以为一线的教师提供大量生动的、可用于教学的范例,为课件的编写、教学内容更新、课堂教学活动的组织安排等提供新的教学资源,从而有助于提供课堂写作教学的有效性。

本库在建设前期有代表性地选取了英语专业学习者语料库 (CEM) 中的部分语料,库容约为 50 万词次,并按照文章类型、组织结构、句子和词等层面进行人工及机器自动标注。在建设中期,我们在语料与写作教材之间搭建接口,使得教师不仅能应用于写作教学研究,也能在课堂教学中发挥其作用,以此进行讲课分析。在建设后期,我们实验性地将部分接口运用到实际写作教学,并向使用者提供自编软件用于学生写作语料分析,使得教师和学生以互动的方式来开展英语写作的教与学,进一步实现以学生为中心的写作教学理念。

关键词: 教学语料库、英语写作教学、英语专业学习者

The Construction and Application of Pedagogical Corpus for English Writing

Abstract: Pedagogical corpus for English writing, with a view to both facilitating the process of teaching writing and rendering more insight in writing research, is characterized by the learners' writing output both in exam and classroom practice environments as well as the annotation by corpus-builders. Such corpus construction makes it possible to provide teachers with fresh teaching resources as they may get the access to authentic, teaching-oriented writing samples, thus conducive to their compilation of teaching prompts, upgrading of teaching material and organization of classroom activities.

In the initial stage of building the corpus, we selected a portion (0.5 million tokens) from *Corpus for English Majors* (CEM) featuring enough representativeness, and then moved on to manual and machine-automatic annotations from the layers of genre, organization, sentence and word. Afterwards, the interfaces between our corpus and the teaching materials were established so that teachers are capable of applying both to research and classroom teaching. In the final stage, we tentatively employed a part of the interfaces to teaching writing and supplied the users with the self-designed software for composition analysis. In this way, both teaching and learning of English writing can be interactively conducted so that the philosophy of learner-centeredness in teaching English writing might be fulfilled.

Keywords: pedagogical corpus, teaching English writing, English majors

潘鸣威
上海外国语大学

第七组

2-7-1

12月26日, 16:00—16:20
第三分会场

作为中介语的“翻译腔”问题初探——一项基于学习者语料库的研究

摘要: 对中介语的研究在国际和国内都取得了一定的进展,研究者从理论和实证层面都进行了广泛的讨论,但从语料库语言学角度进行的研究还很少。学习者双语对应语料库的研制和应用将为中介语的研究提供十分重要的方法论和数据来源。对“翻译腔”问题的研究无论是从理论上和实践中都有待进行更深入的探讨。正是在这种理论和实证研究的背景之下,本文旨在从中介语的理论视角讨论英语专业学生汉英翻译中存在的“翻译腔”问题。本文认为,学生在翻译中出现的“翻译腔”可以作为一种典型的中介语现象进行研究,而基于学习者双语对应语料库又为中介语和“翻译腔”这两个问题的研究都提供了一种可行的方法论和可靠的数据来源。本文以文秋芳和王金铨等研制的中国大学生英汉汉英口笔译语料库为基础,抽取其中的汉英翻译语料库的语料4(D01CE)作为数据来源,以ParaConc和AntConc为分析工具,对语料中作为中介语的“翻译腔”问题在主要词汇和句法层面上表现出来的特征做出初步的归纳和分析。在分析过程中,以“经济全球化”、“缩小南北差距”、“经济全球化使南北差距进一步扩大”等词语和句子翻译中出现的“翻译腔”问题为中心,探讨作为中介语的“翻译腔”在学习者双语对应语料库体现出的特征和规律。

关键词: 学习者语料库、中介语、“翻译腔”

“Translationese” as Interlanguage —A Study Based on the Parallel Corpus of Chinese EFL Learners

Abstract: The study on interlanguage has made considerable progress both at home and abroad. While researchers have studied the issue from various perspectives, the approach from the corpus linguistic in this aspect remains extremely inadequate. The establishment and application of the parallel corpus of Chinese EFL learners provides very important methodology and data for the study of interlanguage, and the research on “translationese” calls for further investigation from theoretical and empirical perspective. In this theoretical and practical context, the present paper attempts to study the “translationese” of the Chinese EFL learners from the theoretical framework of interlanguage. The author maintains that the “translationese” of Chinese students can be viewed as a kind of typical interlanguage, and the parallel corpus of Chinese EFL learners provides feasible methodology and reliable data for the study of both interlanguage and “translationese”. Based on the parallel corpus of Chinese EFL learners established by Wen Qiufang and Wang Jinquan, this paper draws the 4th sub-corpus (D01CE) as the data for analysis, and utilizes ParaConc and AntConc as analytical tools, to probe into the semantic and syntactic features of “translationese” represented in the students’ translation versions. These features can be studied from the theoretical perspective of interlanguage. In the analysis of the features and particular rules of “translationese”, this study focuses on the phrases of “economic globalization”, “to bridge the North-South gap”, and the sentence “Economic globalization further widens the gap between the North and the South”, to investigate the semantic and syntactic features of “translationese” as interlanguage.

Keywords: Learners corpus, interlanguage, translationese

赵勇
大连外国语学院

2-7-2

12月26日, 16:20—16:40

第三分会场

情态动词在教学大纲中的语法排序对比 ——以中外语料库为根据

摘要: 二语语法学习中,情态动词(涉及13项含过去式)是二语学习者感觉最难把握的语法项目之一,涉及语言态度、态势与语言情感的英语情态动词既要表示可能、许可、义务、必然、推测、意愿、意图与决心,还有更多的诸如:惋惜、忧虑、欢乐、委婉、胆敢等比较复杂细腻的感情与情绪表达方式。其细微与复杂程度堪称观止,如果能在实际的语言语法教学中遵循科学规律,利用频率、语境及语用等规律廓清各词项关系,厘清各层面区别,将对于学生的学习会是一种极大的帮助。本研究就想对所有情态动词项的在各大国外权威口、笔语(LLC, Frown/Brown, LOB/FLOB,)以及在国内、外的教材语料库(XBZ, ECEC, NHW)出现频率做一对比,同时就各大词项出现先后顺序做一比较,力图发现它们之间的差异,同时找出其中的语言文化原因所在,最后对此现象提出自己独特的教学大纲修正建议及实际教学的内容编排示范,希望达到对于各学生阶层情态动词语法教学的启发帮助作用。这种大规模的语料库调查这在国内还尚属第一次。

关键词: 情态动词、语料库、语法排序、教学大纲

龙绍赞
江西师范大学

2-7-3

12月26日, 16:40—17:00

第三分会场

促进学习者自主性: 基于学习者翻译语料库的翻译评估

摘要: 培养学习者自主性是近年来教育研究中的一个重要论题。本文认为,学习者翻译语料库对于促进学习者自主性,推动翻译评估模式的革新有着特殊的意义。相对于平行语料库和比较语料库,学习者翻译语料库为学习者提供一个较为实际的参照,有助于树立学习者信心,促进学习者自主性;同时,学习者翻译语料库对于观察翻译这一特殊的中介语形式不可或缺,教师可以较为真实地了解学生的翻译过程从而进行相应的课堂设计,进一步提高学习者自主性。

基于上述假设,我们进行了实验。实验对象为英语专业3年级学生,掌握必要的翻译理论与技能,具备一定的计算机操作能力。实验分为两个阶段:第一阶段,我们建构了小型学习者翻译语料库,包含同一原文的10篇不同译文,通过课堂引导与讲解对学生进行实验前的培训。第二阶段,我们利用教学平台在线提交功能获取61名学生的122篇译文,对每篇译文的前500词进行错误标注,然后交由学生使用。学生在语言实验室借助计算机与相关软件进行语料检索与统计分析。实验的基本内容有:译文评估与标注,词对齐文本取样,句对齐文本取样,某一错误类型频度统计,个人错误类型统计。由此产生的结果将在课堂或课外

进行讨论或报告演示。

通过对实验过程的观测和调查问卷的结果统计,我们发现:较之传统的教师讲评,基于学习者翻译语料库的翻译自评和互评能够显著提高学习者自主性。绝大多数学生在此过程中积极主动,能自发开展小组讨论,少数学生能对教师的批注提出合理的质疑,学生整体来说对学习者翻译语料库感兴趣,肯定其应用价值。存在的问题有:电子阅读易造成视觉疲劳;学生对自己所做的评估缺乏信心和必要的依据;个别学生对于使用索引工具等有一定的畏惧心理。因此,今后研究中我们将加强语料库的相关培训,在实验中引入必要的平行语料库和比较语料库,更为全面的挖掘学习者翻译语料库的应用价值。

关键词: 学习者翻译语料库、翻译评估、学习者自主性

Towards Learner Autonomy: A Learner Translation Corpus-Based Approach to Translation Quality Assessment

Abstract: Learner autonomy has been a vital issue in recent studies on pedagogy. The paper proposed a hypothesis that learner translation corpora (LCT) can be of considerable significance for promoting learner autonomy in translation training with its application in translation quality assessment. To confirm the hypothesis, an experiment designed for English majors in their third years has been conducted.

For the purpose of tentative research, we used a LCT consisting 122 translated texts of two original texts by 61 students. Preliminary work has been done based on a LCT consisting ten translated texts of the same original text while teacher lectured and showed the procedure of the experiment.

Students in experiment are expected to: sample the translation of certain word, phrases or sentences; quantify the frequency of certain error types; quantify the frequency of error types for individual. There is a wide range of choices for them as which words, phrases, sentences or error types they will make an investigation based on their own interest. A challenging task for the students is to peer assess the quality of translation in LCT since we annotated only the first 500 words of each translated texts purposely to encourage the peer-assessment and enhance the learner autonomy. The findings of the above research will provide practical assistance for both the students and teachers with a proper presentation or discussion in class and an individual reflection on achievements in their translation journal.

Observation of the experiment together with questionnaire survey supports the above hypothesis. Great enhancement in learner autonomy can be perceived when most of students actively involved in the research and group discussion was held without any requirement from teachers. Some, even not many students even questioned the authority of some annotation on error. The pressing problems include the following: visual fatigue; lack of confidence and criteria in students for peer-assessment; technophobia such as that of handling the concordancer.

Keywords: Learner translation corpus, translation quality assessment, learner autonomy

邵莉

山东轻工业学院

12月26日, 17:00—17:20

第三分会场

运用美国当代英语语料库(COCA)提升学习者真实语感

摘要: 美国当代英语语料库(COCA)是美国杨伯翰大学的 Mark Davies 教授开发的免费在线大型英语语料库, 现已涵盖 1990—2009 年间的五大语域共 4 亿词汇, 并且会逐年增加最新数据。在本族语语料库与学习者语料库都并行发展的今天, 过分强调学习者语料库的价值有可能误导学习者, 从而使他们在语言的语感方面受到负面影响。学习者语料库的主要使用者应该是研究者而非学习者, 学习者应更多地接触目标语语料库, 因而为学习者引荐一个可以自行使用并分析和感知自己语言使用错误的语料库就有其实际意义。在线语料库使用者查询的内容可以很好地被收集, 从而成为真实的使用者进行语感校验的一个集合, 可以被教师用来指导学习者避免类似学习错误。

COCA 是一个真实美语的平衡语料库, 反映出当代美语的典型语言特征, 语言的使用规律在简单直观而且友好的查询界面内可以真实地反映出来, 其统计结果可以进一步引领使用者调阅查询内容的微语境信息。传统的错误分析无法囊括所有学习者的问题, 只有学习者的亲力亲为才能有效提升其真实语感。将学习变成一种以学习者为中心, 语料库为重心, 体验为核心的活动, 能更好地实现 3I 学习法模式(即 Illustration-Interaction-Induction)(Carter & McCarthy 1995:155), 这种模式使借助语料库进行的学习更具探索性。COCA 除了能有效解决搭配方面的疑惑外, 学习者对所查阅的节点词的微语境反映的文化以及社会的认识也会增强。运用 COCA 语料库进行自主学习能较好地解决语言学习材料的真实性和滞后性等问题, 将本族语语料库变成一个为学习者服务的语料库的价值便会凸显出来。

关键词: 学习者、语料库、3I 学习法、真实语感

Using COCA to Increase Sense of Language in English

Abstract: The Corpus of Contemporary American English (COCA) was created by Mark Davies, a professor of corpus linguistics at Brigham Young University. COCA is the largest freely-available corpus of English (<http://www.americancorpus.org/>), and the only large and balanced corpus of American English. The corpus contains more than 400 million words of text and is equally divided among spoken, fiction, popular magazines, newspapers, and academic texts. It includes 20 million words each year from 1990-2009 and the corpus is also updated twice a year. Because of its user-friendly interface, it is perhaps the only corpus of English that is suitable for looking at different words or collocations in the language with unlimited access.

Both the learner corpora and the natives' corpora are being developed well in recent years, but too much stress on using learner corpora might mislead learners and make them be exposed to negative transfers. Researchers and teachers had better confine the learner corpora to themselves. An optimal corpus should be the one which could be used by language learners by themselves whenever they want. Thus using COCA is practical and helpful for them to observe and perceive language used by native speakers, and these advantages could not be substituted by learner corpora. COCA can also show the users about 180 words of the contexts of the searched strings and they help the users to know more about the background, culture and society. Learning to use this corpus well is explorative and it could foster learners' language development. Traditional language error analyses could not include all learners' problems. To explore the native corpus

could train them in the 3I learning process (Illustration-Interaction-Induction). And COCA could help learners to throw away outdated materials and activate their potential in searching uncertain phrases.

Keywords: English learners, corpus, 3I learning process, sense of language

汪兴富、刘世平
重庆大学

2-7-5

12月26日, 17:20—17:40

第三分会场

中国英语学习者语料库在双语学习词典编纂中的应用

摘要: 词典编纂的素材按来源可分为三类: (一)其他词典。比如, 双语词典的编者通常会找若干本单语词典为蓝本; (二)依靠编者的内省和对词典使用者的调查; (三)对使用中的语言的调查, 即借助语料库。本文针对国内英汉双语词典编纂的一个不足之处, 即没有充分考虑学习者用户的产出需要, 没有利用语料库科学地统计出学习者常见的语言使用错误, 在词典的微观结构中加入相关设计, 从而促进学习者语言的产出。本文借鉴国际双语词典编纂研究的最新成果, 充分利用《中国英语学习者语料库》等国内外英语学习者语料库, 应用语料库语言学和二语习得理论等, 科学统计、深入分析、系统总结中国英语学习者常见的错误类型及其在双语学习词典中的处理措施, 切实把理论研究落实到词典的辅助编纂中, 推动我国英语教育事业的发展, 促进中国学习者英语水平的提高。

关键词: 学习者语料库、双语学习词典

王亚军
厦门理工学院

2-7-6

12月26日, 17:40—18:00

第三分会场

基于语料库的中美大学生议论文高频动词及物性研究

摘要: 本研究以系统功能语法的二维及物模式 (Peng 2005) 下的过程类别和隐喻理论为分析框架, 比较中美大学生议论文语料库中的两个高频动词 *make* 和 *find* 的四种过程的使用情况, 运用定量和定性分析相结合的方法, 旨在考察和揭示我国大学生英语议论文中高频动词的及物性过程使用特征。具体考察问题为 (1) 中美大学生使用 *make* 和 *find* 的及物性 (或意义) 分布情况; (2) 中美大学生使用 *make* 和 *find* 的一致式用法和隐喻用法的特点, 以及它们与类联接和词语搭配的关系。研究发现, 中国学生和本族语者在及物性过程的分布上没有显著差异, 以社会经验和认知经验为主要的经验类型。动词的一致式用法多于隐喻用法, 但隐喻现象仍较普遍。动词的类联接和搭配与及物性和隐喻方式有一定联系, 但这种联系比

较复杂，有待进一步研究。

关键词：及物性、过程、学习者语料库、高频动词

张莎

北京外国语大学

第八组

2-8-1

12月26日，16:00—16:20

第四分会场

中国大学生英语会话中的回指修补研究

摘要：相关研究发现，在自我修补的纠错阶段，回指现象较为普遍。回指表示话者在实施修补时，往往会回到修补成分开始处，而不是直截了当地在错词后实行修补。目前国内外对于回指修补的研究主要限于少量的L1的研究，对L2的回指修补研究还是空白，本研究旨在对这一普遍的语言现象做一些探究，来弥补这一空白领域。本研究试图回答以下两个问题：

- (1) 中国大学生在英语会话回指修补策略的使用上有何特征？
- (2) 中国大学生在英语会话回指修补上回指词的分布如何？与L1比较有何异同？

本研究应用中国学习者英语口语语料库(COLSEC)，采用随机抽样法从语料库中抽取共31个样本。鉴于国内外对会话回指修补的研究十分有限，在对COLSEC语料回指修补策略的分类上本文借鉴了姚剑鹏(2007)分类模型，并结合本语料对分类模型做了调整，将回指修补策略分为三大类：无显见纠错的回指(即自我重复)；纠错回指(包括替换、插入和删除)和重组回指。笔者按照上述分类模型对语料进行了手工标注，然后用AntConc3.2.2统计了各类修补的频数。

研究结果发现，大学生在英语口语回指修补上自我重复所占比例最高，重组和插入所占比例很少，反映出大学生的信息加工能力和语言在线处理能力以及交际技巧有待提高，他们在回指起始词使用上与本族语者有明显差异。本研究借用认知心理学的注意力理论来解释结果，并指出该研究对大学英语口语教学的启示。

关键词：英语会话、自我修补、回指、口语语料库

A Study of Anticipatory Retracing in English Conversation by Chinese University Students

Abstract: The relevant research shows anticipatory retracing is a common form of self-repair and it is conducted by tracing to and repeating some word prior to the trouble element. The present study explores the modes and features of anticipatory retracing by Chinese university students under the theoretical framework of cognitive psychology. It attempts to address two research questions: 1) what are the typical features of anticipatory retracing done by Chinese university students? 2) Are there any significant differences between Chinese university students and native speakers in their uses of initial words of retraced constituents?

Chinese Learner Spoken English Corpus (Yang *et al.* 2005) is used for the study. The data of

31 students has been selected, using random sampling method. The study follows Yao Jianpeng's model (2007) which is modified for the fulfilment of purpose of the present study. In the present study, the three main categories of anticipatory retracing are self-repetition, retracing with correction and retracing with reformulation. All the transcripts are tagged by hand. And AntConc 3.2.2 is then used to count frequencies of the categories.

The research has found: 1) There is a high frequency of repetition which outnumbered any other categories; 2) there are significant differences between Chinese university students and native speakers in their uses of initial words of retraced constituents. The results are discussed in the light of two-process theory of attention. Besides, the paper points out that anticipatory retracing merits our due attention, especially upon oral English teaching.

Keywords: English Conversation, self-repair, anticipatory retracing, spoken corpus

权立宏
广东外语外贸大学

2-8-2

12月26日, 16:20—16:40
第四分会场

连接副词的语用功能与使用情况调查分析

摘要: 本文依据关联理论, 从认知的角度对英语连接副词进行分析, 指出连接副词在话语交际中的各种语用标记作用, 并用语料库的研究方法, 以英语本族语者所写的议论文作为参照, 调查中国英语学习者在英语议论文写作中使用连接副词的情况, 以考察中国英语学习者在议论文写作中的语用认知意识。

关键词: 连接副词、语料库、语用功能、关联理论

An Investigation of Pragmatic Functions of Adverbial Connectors and English Writing

Abstract: This paper makes an analysis of English adverbial connectors from a cognitive perspective based on the framework of Relevance Theory. It points out that adverbial connectors perform different pragmatic functions in communication. The paper goes further to conduct a comparative, corpus-based study of English adverbial connectors in the essays written by the Chinese EFL learners and native speakers of English, so as to examine the Chinese EFL learners' pragmatic cognitive awareness in their English essay writing.

Keywords: adverbial connectors, procedural meaning, pragmatic function, relevance theory

邓飞
华南农业大学

2-8-3

12月26日, 16:40—17:00

第四分会场

A Corpus-based Study on Written Vocabulary by First-year English Majors

Abstract: Based on the Written English Corpus of Chinese Learners (SWECCL), this paper investigates the lexical richness and features of some first-year English majors' writings with the same title. The tools used in this research are Nation's vocabulary profile software Range32 and Anthony's corpus concordance and statistics software AntConc3.2.1. With Range32 the study analyzes the lexical richness in terms of the word-frequency distribution, the lexical variation, the lexical sophistication and the lexical density. The 69 English learners' writing corpus is also compared with the native corpus BNC to examine the use of the first 50 high-frequency words. At last the use of some high-frequency verbs are analyzed with the concordance tool and cluster of AntConc3.2.1. The findings show that students express themselves with too much high-frequency words and produce very limited content words for lack of vocabulary richness in terms of the four parameters mentioned above. They have overused the first and second personal pronouns and underused the third ones; simultaneously some modal verbs are overused or underused. They have a relatively good command of some high-frequency verbs and use some collocations and idioms correctly. But there also exist some mistakes, inappropriately chosen words and Chinese English, which might be caused by factors like limited productive vocabulary and depth of vocabulary knowledge, negative transfer of mother language and traditional teaching patterns. Therefore, it is suggested that measures be taken to emphasize language output, vocabulary culture lead-in and adopt corpus-based teaching methods so as to enhance the accuracy and nativeness of the English learners' language production and accelerate the development of productive vocabulary.

Keywords: WECCCL, written vocabulary, corpus, lexical richness, high-frequency words

付玉萍
海南琼州学院

2-8-4

12月26日, 17:00—17:20

第四分会场

大学英语写作中的词形错误分析 ——一项基于语料库的研究

摘要: 错误分析作为语料分析的一个重要方法, 无疑在第二语言的教与学中起到重要作用。它不仅能帮助我们了解学习者对第二语言掌握的程度以及学习的策略, 而且还可以帮助教师了解学习中大家共有的特点, 对促进教师和学生的教与学具有重要意义。

近年来, 在国内, 随着几大语料库的建立, 如桂诗春、杨惠中的《中国学习者英语语料库》(CLEC), 文秋芳、王立非的《中国学生英语口语语料库》, 许多研究者开始对这些语料库的语言错误进行研究。桂诗春, 杨惠中(2003) 在 CLEC 中把语言错误分为词形、动词词组、名词词组、代词、形容词词组、副词、介词词组、连词、词汇、搭配和句法 11 类错

误，其中词形失误包括拼写、构词及大小写失误。本文参照桂诗春的分类，采用定量和定性两种分析方法，对中国学习者语料库中 290 篇作文的词形错误进行分析。为了对不同学习层次的词形错误进行对比，其中的 130 篇来自非英语专业低年级学生，160 篇来自非英语专业高年级学生，且写作任务相同。

结果表明：（1）词形错误占语言错误的 29.41%，仅次于词汇错误；（2）词汇拼写错误占词形错误的 71.90%，构词错误占 10.38%，大小写错误占 17.72%；（3）作文的词形错误数量与写作质量负相关（ $r = -.352^{**}$ ）；（4）不同学习层次学生以及不同写作水平学生犯词形错误的数量差异非常显著。

关键词：英语写作、词形错误、学习者语料库

何华清
西华师范大学

2-8-5

12月26日，17:20—17:40

第四分会场

从英语专业限时议论文看被动语态习得 ——基于 SWECCL 的对比研究

摘要：本文基于 SWECCL 作文子库的大一、大三限时作文，研究中国英语专业大学生议论文中被动句的应用，结果显示：1、中国英语专业大学生使用被动语态类型相对单一，大一、大三差别不显著。2、使用最频繁的被动语态为“情态动词+be”的形式和“Subject + be + V-en”形式的一般时态，但完成时、进行时被动使用相对较少。3、大一英语专业学生过度使用一些实意动词，变化缺乏多样性，而大三使用的种类相对丰富。4、受汉语被动“遭受”意“迁移”影响，学生所用被动句明显呈现“消极”色彩。

关键词：被动语态、语料库、限时议论文

黄斌兰、刘士祥、朱兵艳
广西大学、广西大学、海南大学

2-8-6

12月26日，17:40—18:00

第四分会场

中国学习者语域意识的多维度分析

摘要：本研究采用语料库语言学方法，把语域变异多维度分析模式（MDA）应用到中国英语学习者书面语的语域意识研究中。所用 120 篇学习者语料来自 WECCL，由一所重点大学学生所写议论文，用于对比的本族语者语料库来自 LOCNESS 的子库，共 175 篇，是美国大学生所写议论文。为了对比学习者内部语域发展的特征，按学习者作文得分，将其划分为高低分组。本研究分别在 MDA 的第一、三、四、五等四不同维度，系统地分析了中国大

学生在作文中所呈现出的语域特点。研究发现：（1）我国非英语专业大学生在作文中呈现混合语域特征，在四个不同维度上分别具有不同特征。在第一维度上，学习者表现出明显的口语倾向和参与性特征；在第三维度上，学习者的语域特征和本族语者并无明显差异；在第四维度上，学习者与本族语者语域意识虽然也无明显差异，但是学习者却出乎意料地表现出更明显的议论性特征；在第五维度上，学习者表现出明显抽象性特征；（2）在学习者内部，随着学习者英语水平的提高，在第一、三维度上的语域意识有了较为显著的发展，分别具有显著信息性和详尽指代性特点；但是，在第四、五维度上却并无明显提高。该研究结果一方面从全新角度进一步证明了学习者中介语的混合语体特点，同时发现学习者语域意识并非总是和本族语者存在明显差异，他们有时会表现出相似性，学习者的语域意识在各个维度的发展并非同步进行。

关键词：语域变异多维度分析、语域意识、二语学习者

A Multi-dimensional Analysis to Chinese EFL Learners' Register-awareness

Abstract: Previous studies on L2 learners' lack of register-awareness have often focused on one or two aspects; little attention has been paid to multi-dimensional and systematic analysis of the problem. This study employs corpus analysis techniques to reveal the register characteristics of Chinese L2 learners in their written English on four dimensions by applying the MDA technique. Findings of the study indicate: (1) Chinese L2 learners generally display hybrid register. On the four different dimensions, their register respectively are more involved, elaborated, overt argumentative and abstract; (2) with the progress of their English proficiency, higher-level learners' register-awareness has been improved significantly. Their language is more informative and elaborated on Dimension one and three. However, no significant progress can be found on Dimension four and five.

Keywords: multi-dimensional analysis, register-awareness, L2 learners

刘小燕
徐州工程学院

第九组

2-9-1

12月27日，10:00—10:20

第一分会场

中国研究生英语学术论文摘要中的词块研究

摘要：英语学术摘要写作是我国研究生和科研工作者在英语使用上的重点和难点。本文对中国研究生英语学术摘要中的词块进行了基于语料库的对比研究，从而为我国科研工作者在实际的英语学术摘要写作提供参考。本研究通过建立两个平行语料库——国际贸易方向国际权威期刊的英语摘要和中国国际贸易专业研究生毕业论文的英语摘要，使用 AntConc 进行非预定的词块检索，结合手工排除，确定对比词块，然后对其进行分类和比较。研究结果表明

中国研究生在词块使用的丰富性，结构性和互动性方面均与国际水平存在较大差距。本文在最后提出了基于该研究的教学建议。

关键词：英语学术摘要、词块、语料库

杨华

北京第二外国语学院

2-9-2

12月27日，10:20—10:40

第一分会场

基于语料库的同义词语义韵辨析及对外语教学的启示 ——以 *effect* 为例

摘要：作为语言研究的一种重要研究方法，语料库语言学一直把词语搭配作为一个研究重点，而词语搭配多表现为一些特殊的语义韵倾向。通过观察检索行中所呈现的同义词搭配特征，可以更好区分同义词的类连接、搭配关系和语义韵等语言特征。本文以 *effect* 为例，分析了其与同义词 *influence* 和 *impact* 在 LOB 和 WECCL 中的语义韵特征和搭配特征，试图找出了母语使用者和中国英语学习者在这些同义词的语义韵使用特征上的区别，旨在为同义词学习以及外语教学提出一些建设性的意见。

关键词：语料库、同义词、语义韵、搭配

Semantic Prosody of Synonyms and its Implications to Foreign Language Teaching——A Case Study of *effect* in LOB and WECCL

Abstract: As one of the approaches in linguistic research, corpus linguistics places one of its main focuses on the concordance of the words. The fact is that most of English words have semantic prosody, so it is important to differentiate the colligation, concordance and semantic prosody of the synonyms in English learning. This article analyses the noun *effect* and its synonyms on the basis of two corpora: LOB and WECCL. It also makes a comparison on the semantic prosody of synonyms between the native language users and Chinese English learners. Some suggestions are proposed for language learning, especially the ones concerning synonym learning.

Keywords: corpus, synonym, semantic prosody, concordance

刘浩

天津师范大学津沽学院

2-9-3

12月27日，10:40—11:00

第一分会场

英语专业大一、大三学生英语议论文中后置定语“语块输出”特点

——基于 SWECCCL 语料库的对比研究

摘要: 本文基于 SWECCCL 语料库 (1.0 和 2.0) 作文子库的大一、大三限时作文, 研究中国英语专业大学生议论文后置定语的使用特点, 并对其作进一步对比研究。结果显示: 1), 中国英语专业学生使用后置定语语块类型相对单一, 且过度使用 2 词的线性语块, 缺乏必要修饰语; 2), 大三学生使用名词后置语块种类相对丰富, 典型特征不明显; 3), 中国英语专业学生使用后置定语语块呈现典型线性句法特征, 映射中式英语翻译思维。本研究对二语习得与外语教学具有启发意义。

关键词: 语料库、语块、后置定语、议论文

The Features of Post-attributive “Chunks” by Freshmen and Juniors of English Majors in English Argumentation —A Comparative Study Based on Corpus SWECCCL

Abstract: Based on sub-corpus of SWECCCL (1.0 and 2.0), timed compositions by freshmen and juniors, the paper aims to study the post-attributive characteristics of Chinese English Majors and to give a further comparative study between them. The results showed that: 1), The post-attributive formulaic language used by Chinese English majors are relatively homogeneous with an excessive use of two linear formulaic language, lacking some necessary modifiers; 2), The post-attributive type of formulaic language used by juniors are relatively abundant, without obvious typical features; 3), The post-attributive formulaic language used by Chinese-English majors showed typical linear syntactic features, suggesting strongly Chinese translating ways of thinking in English output. This study would be instructive for the SLA and FLT.

Keywords: Corpus, formulaic language, post-attributives, argumentation

刘士祥、韦保丞、朱兵艳
广西大学、广西大学、海南大学

2-9-4

12 月 27 日, 11:00—11:20
第一分会场

组块在英汉平行语料库中的对齐和提取

摘要: 在汉语语料库语言学研究中, 基于词的研究取得了令人满意的成果, 如中国科学院张华群的 ICTCLAS 汉语词性标注。英语单词之间有空格, 分词相对简单, 词性标注有很多软件工具可供选择。但是对句子进行句法分析的软件, 特别是对汉语进行句法分析的软件, 非缺乏常, 成为当前计算语言学研究中的瓶颈。令人庆幸的是, Stanford 大学的 NLP 研究小组为我们免费提供了建立英语、汉语 Penn 树库的一系列工具, 包括汉语分词、词性标注和句法分析以及英语词性标注和句法分析。基于词的分析有一定的局限性, 而基于更大粒度的 Chunk 级分析应用范围更广。本文简要介绍了组块 (Chunk) 的研究历史和现状, 探讨运用 Stanford NLP 工具进行 Chunk 抽取方法和运用免费语料库工具 Uplug 进行英语、汉语 Chunk 级对齐以及英语、汉语 Chunk 级的对应语抽取方法。

关键词：组块、英汉平行语料库、组块级对齐、组块提取

Chunks Alignment and Extracting in English and Chinese Parallel Corpora

Abstract: In the study of Chinese Corpus Linguistics, satisfactory results are achieved based on the study of words, for example, ICTCLAS, the Zhang Huaqun's Chinese part-of-speech tagger at Chinese Academy of Sciences. Due to the spaces between the spaces between words, word segmentation in English is relatively simple, and there are a lot of part-of-speech taggers available. But the software for syntax analysis, especially for Chinese syntax analysis, is very rare, which is the bottleneck in the study of Computational linguistics. To our rejoice, the NLP group of Stanford University supplies us with a series of free software tools to build the Penn tree bank of English, Chinese and so on, including Chinese word segmentation, the Chinese and English part-of-speech tagging and syntax analysis. Since there are limitations based on the analysis of words, it is widely used based on larger granularity, the chunk. This article briefly introduces the history and current situation of the study of "Chunk" and then mainly focuses on the methods of extracting the parallel Chunks of Chinese and English, using the free software tools of Stanford NLP group and Uplug of the free corpus tools.

Keywords: chunk, parallel corpora, chunks alignment and extracting

卢美华

北京农业职业学院

2-9-5

12月27日, 11:20—11:40

第一分会场

基于语料库的英语专业教材词块分析

摘要：“教什么”和“怎样教”一直以来是教学界的重点研究问题。近年来，随着学者们对语块研究的不断深入，在外语教学中，不同层次的外语学习者应该学习哪些词块、词块教学应该如何在日常的教学中实施等问题已经成为“教什么”和“怎样教”的具体分支问题。国内外对词块的研究可谓方兴未艾。然而，迄今，尽管教材和课堂导入是外语教学输入的重要途径，国内对英语教材以及课堂教学所呈现的词块研究却不多。本文尝试通过分析国内的英语专业教材《综合教程》，试图为“教什么词块”和“怎样教词块”提供一点教学启示。本研究的具体做法如下：1.采用问卷调查的方式了解到学生认为带有介词的动词、名词和形容词短语是他们习得的难点；2.使用语料库检索工具找出《综合教程》中高频的带有介词的动词、名词和形容词短语；3.以美国当代语料库(COCA)作为参考语料库，找出《综合教程》最高频的带有介词的动词、名词和形容词短语的语义搭配和语义韵；4.进行语料库驱动的词块教学实验。5.采用问卷调查的方式了解学生对语料库驱动的词块教学的感受。本研究的结果显示：1.英语专业教材《综合教程》的确呈现了一些学习者自认为需要掌握的词块；2.语料库教学加工在一定程度上能帮助突显教材中的词块。

关键词：词块、英语专业教材、词块教学、语料库

唐洁仪
华南师范大学

2-9-6

12月27日, 11:40—12:00
第一分会场

基于语料库的增强语程度副词语义韵认知研究

摘要: 语义韵指的是节点词习惯性的吸引某一类具有相同或相似语义特点的搭配词, 其语义相互感染、渗透, 从而在语境中形成一种特定的语义氛围 (Louw 1993; Sinclair 1996; Stubbs 2001; 卫乃兴 2002)。实际上, 语义韵是超越词界的具有联想色彩的搭配, 是一种集文体意义、感情意义、修辞意义和民族文化意义于一体的搭配。语义韵有积极语义韵、消极语义韵与中性语义韵之分。英语学习者在语言使用中如果忽视语义韵, 脱离语境, 单纯套用词语的字面意义或字典意义, 势必错误百出。

本研究通过定量定性相结合的方法, 利用中国学生英语口语语料库、北京外国语大学通用中英对应语料库和 Corpus of Contemporary American English, 通过检索工具 AntConc 和统计学中的卡方检验考察节点词在学习者语料库、专业译者语料库和英语本族语语料库的搭配、语义倾向和翻译中的使用情况。本研究中节点词欲选取 more, very, too, quite, much, far, highly, fully, completely, entirely 等 20 个程度增强语 (Quirk *et al.* 1985; Alexander 1988; 张道真 1995; 章振邦 2003), 考察他们的语义韵在学习者语料库、专业译者语料库及本族语语料库的使用中是否存在显著性差异, 并从认知的角度探索产生差异的原因。

借助语料库学习者或译者将极大的提高学习研究的自主性和灵活性, 语料库的最大价值在于唤起学习者或译者进行思考, 而非仅仅提供现成的答案。

关键词: 语义韵、认知、学习者语料库、中英对应语料库、英语母语语料库

赵秋荣、张红燕
中国地质大学 (武汉)

第十组

2-10-1

12月27日, 10:00—10:20
第二分会场

A Corpus-based Study on Vowel Pronunciation of Chinese EFL Learners

Abstract: Through carrying out serious research project and systematically analyzing data, this study aims at observing the actual variations of /ai/ pronounced by the second year students major in English based on the Spoken English Corpus of Chinese Learners (SECCL). Additionally, according to the research results, the distribution of the variations and the analysis on the reasons

with the help of the theories of second language acquisition, phonetics and phonology will be put forward, which is expected to play a guide role for teachers of English major or college English, so as to improve pronunciation teaching effectively. In this dissertation, Praat software will be used to get the spectrograms of the variations in SECCL comparing with that of British Component of the International Corpus of English (ICE-GB), the International Phonetic Alphabet (IPA) and Standard Mandarin Chinese ai.

The research on oral pronunciation of English as the interlanguage for Chinese learners per se is a blind spot in second language acquisition as well as interlanguage study in China. According to Dai & Cai (2006), “in terms of linguistic level, our Chinese and English interlanguage research still lacks the theoretical or empirical exploration on pronunciation and discourse”. “As for transfer, it is much more obvious on phonetic level than syntactic level (Ellis, 1999). But in China, we really need a push on phonetic research”. Therefore, this dissertation stands on the side of empirical study, using Praat to judge the actual vowel pronunciation of second year English majors in SECCL. Corpus-based research now is an emerging trend in linguistic study and a perfect combination between computer and second language acquisition. Based on Wang Lifei and Sun’s statistics (2005), SECCL is the only corpus which contains the oral sound files among those having been issued. The results based on it will be relatively general and convincing. This study is different from some previous studies based on acoustic analysis, which concentrate on learners’ natural property. With the scientific methods, the conclusion in this dissertation is more reliable.

The specific research questions to be addressed by this study are as follows:

1. Does English diphthong /ai/ pronounce the same with Chinese Pinyin ai?
2. What are the distributions of the diphthong /ai/ pronounced by the native speakers in ICE-GB corpus?
3. What are the distributions of the variations of /ai/ in the interlanguage of the Chinese learners in SECCL?
4. Does the influence of first language transfer play an important role in second language pronunciation?
5. How to improve English majors’ pronunciation teaching and learning, especially their vowel production?

Keywords: pronunciation, Chinese EFL learner, English diphthong /ai/, first language transfer

蒋凤霞
北京外国语大学

2-10-2

12月27日, 10:20—10:40
第二分会场

售书广告对书籍序言的语类侵蚀
——一项基于语料库的研究

摘要: 本文通过研究售书广告和书籍序言这两个语类的互文性在上个世纪中的变迁来发现售书广告对书籍序言进行语类侵殖的机制。语类侵殖是指一种语类的完整性被另一种语类或语类规约所侵犯的过程 (Bhatia 2004: 58)。它反映了现实语篇的动态性和复杂性, 是语类研究领域中的一个新课题。

本研究提出了一个综合的分析框架以探讨语类侵殖现象。该框架利用 Halliday 关于语旨和人际意义的功能派理论改进了 Bhatia 提出的语类变异分析模式。本文的分析针对三个电子语料库, 即英文售书广告语料库 (CBA)、写于 1900 年前的英文书籍序言语料库 (CBP1) 和写于 2000 年后的英文书籍序言语料库 (CBP2) 展开。文章在语类层面及与语类密切相关的语域、文本层面上对 CBP1 和 CBP2 与 CBA 的互文性进行了比较, 并发现了一个清晰的趋势: 与写于 1900 年前的书籍序言相比, 写于 2000 年后的书籍序言与售书广告间有更大的互文性。研究还在社会现实层面上对这一趋势进行解释, 指出它是书籍序言语类在现代社会高度商业化的新挑战下对售书广告的一种“屈从”。本文的结论是: 在当今高度商业化的社会中, 售书广告这种势力强大的语类已经全线入侵书籍序言的领土, 同时将其自身的促销性渗透到书籍序言的语类、人际意义和语言特征的建设中。

关键词: 语类侵殖、语类互文性、售书广告、书籍序言

Genre Colonization of Written Book Advertising over Book Prefaces —A Corpus-based Study

Abstract: This thesis aims to discover mechanisms for genre colonization of book advertising over book prefaces through an examination of changes in generic intertextuality between these two genres over the past century. Genre colonization, a new issue in genre studies responding to the dynamic and complex realities of discourse, refers to the process which “involves invasion of the integrity of one genre by another genre or genre convention” (Bhatia 2004: 58).

The present research sets out to investigate genre colonization based on an integrated analytical framework which modifies Bhatia’s model of generic variation by drawing on Halliday’s functional approach to tenor and interpersonal meaning. Three electronic corpora, namely, the corpus of English book advertisements (CBA), the corpus of English book prefaces written before the year 1900 (CBP1), and the corpus of English book prefaces written after the year 2000 (CBP2) are exploited for analysis. By comparing CBP1 and CBP2 in terms of their intertextuality with CBA on the plane of genre and two of its closely related planes – register and text, a clear tendency is found: post-2000 book prefaces share a much greater degree of intertextuality with book advertising than pre-1900 book prefaces. This tendency is then interpreted on the plane of social practice as a result of the bending of book preface genre towards book advertising under novel challenges of highly commercialized modern society. The thesis concludes that in today’s highly commercialized society, book advertising, a compulsively powerful genre, has been invading the territorial integrity of book preface genre on an enormous front, spreading its promotional elements into generic, interpersonal, and linguistic constructions of book prefaces.

Keywords: genre colonization, generic intertextuality, book advertising, book prefaces

王晓雯
广东外语外贸大学

2-10-3

12月27日, 10:40—11:00

第二分会场

教学资源优化与微型文本创建

摘要: 丰富的语料库及网络英语资源为中国英语教学提供了新的资源。在教学中, 参照语料库及网络英语资源优化教学资源成为一种必要。然而, 语料库及网络英语资源的庞然身形, 常使英语学习者和教学者却步, 为使学生学到地道的英语, 可以针对教学课件中语言实例匮乏或薄弱的环节, 建立微型文本, 并积累成文本库。

关键词: 语料库及网络英语资源、教学资源、微型文本

Improving Teaching Materials and Creating Mini-text

Abstract: The rich online English resources and all kinds of corpora can be a great help to English teachers in improving their teaching materials. However, the large size of the resource may hinder its application in teaching and discourage learners' direct use of it. The creating of mini-text in teaching where reference materials is lacking is a constructive step forward both in using corpora and internet resource, and refreshing the old teaching materials as well.

Keywords: corpora and internet resource, teaching materials, mini-text

叶拉美
深圳大学

2-10-4

12月27日, 11:00—11:20

第二分会场

针对特定领域对外汉语教学的语料库建设¹

摘要: 随着 ESP、移动学习和汉语国际传播的迅速发展, 语言教学的个性化、国别化要求日益凸显。如何针对不同领域的学习者提供个性化的教学语料库(包括字、词、语、篇章等), 成为具挑战性的研究课题。国家语言资源监测与研究中心的动态流通语料库已经支持了四年的《中国语言生活状况》的发布, 我们尝试将这些大规模的真实语料运用于教学语料库的建设。

我们先由统计得到最基础、最稳定的教学语料库, 然后从领域性、时间性两个方面, 针对不同领域的学习者需求, 建成特定领域的教学语料库。

(1) 在领域上, 我们以首都报纸动态流通语料库的 10 份主流报纸为基础, 共计 213 609 505 字次。对其进行文本自动分类, 共有经济、教育、科技等 15 个领域, 通过对这些领域的字词语不同使用情况的统计分析, 得到各个领域的语言文字特征, 如自动获得商务汉语、旅游汉语、工程汉语等对外汉语教学资源。

¹ 本文受北京市教育委员会共建项目“汉语国际推广背景下的首都留学生教育研究”资助, 谨表谢意!

(2) 在时间上, 我们利用汉语史语料库(包括自先秦到现代的不同的时期语料)及国家语言监测语料库的不同时间段语料, 分析其语言特征, 从而为语言教学提供文言词、新词语、流行词语等时效性强的教学资源。

语言教学的一个重要目的, 是确保语言学习高度情境化、个性化、协作式, 实现持续学习甚至终身学习。动态流通语料库亦可以支持不同时段、不同领域的教学语料库建设, 从而为语言教学(如外语学习、二语习得)在时间、空间(地域)、领域等多个层面提供具针对性的教学资源, 真正实现以学习者为中心, 并充分发挥语料库优势的教学互助模式。

关键词: 教学语料库建设、动态流通语料库、对外汉语教学

Teaching Corpus Construction for Teaching Chinese as a Foreign Language in Specific Fields

Abstract: With the rapid development of the ESP, M-learning and the international spread of Chinese language, the individualized and Country-Specific Teaching are increasingly concerned. How to build individualized learning corpus (including characters, words, expressions, discourses, etc.) for the learners from different fields and geographical regions becomes a challenging research item.

Moreover, the National Language Resource Monitoring & Research Center (NLRC) has supported the researches of “Chinese Language Situation Report” in the past 4 years based on the Dynamic Circulating Corpus. This paper tried to apply the large-scale real corpus to language learning. We first selected the most basic of learning corpus by the statistics methods, and expand it from the field and timeliness aspects, which meet the special demands of learners. Firstly, in terms of field, we selected a corpus of 10 major capital newspapers, then did automatic classification, including economy, education, science and technology, etc. we used statistical analysis and obtained the characteristics of the languages in various fields. Such as the business Chinese, Travel in Chinese, engineering Chinese, etc.

Secondly, in terms of diachronic research, we used the historical Chinese language corpus (including the corpus from the pre-Qin dynasty to the modern) and the National Language Research Corpus of different periods, then analysis their linguistic features. So it can provide the new words, popular expressions and other learning resources.

The important target of the language teaching is to ensure a high degree of context-based, as well as individualized and collaborative language learning, even to achieve the continuous learning and lifelong learning. The Dynamic Circulating Corpus can support individualized language teaching on a timeline (such as foreign language learning, second language acquisition), also can provide well targeted resources in different periods, (geographical)spaces and fields, then realize the interactive Teaching, which focuses on learner-centred and exploits corpus' advantages to the full, to some extend.

Keywords: teaching corpus construction, dynamic circulating corpus, teaching Chinese as a foreign language

曾小兵、杨尔弘、邱丽娜、张普
北京语言大学

12月27日, 11:20—11:40

第二分会场

IT 英语语料库统计数据对大学英语教学改革的启示

摘要: 随着国际交流的日益频繁, 各行各业对英语水平的要求日益专业化。本文分析了大学英语教学的现状和 IT 英语语料库的统计数据, 探索了大学英语教学改革的一些新理念, 提出“以学习为中心”和变“一般英语”为“行业英语”的设想, 希望能帮助非英语专业大学毕业生用英语交流他们的专业信息。

关键词: IT 英语语料库、以学习为中心、行业英语

A Statistical Analysis of the IT English Corpus and Its Implications for College English Teaching Reform

Abstract: With the rapid development of international communication, English for specific purposes is demanded by more and more organizations and enterprises. Based on the analysis of the status of college English education and the statistics of the IT English Corpus, the writer explored some new ideas of college English education, suggested the transfer from general English to English for professionals, hoping to help college graduates to exchange their technical information in English.

Keywords: IT English Corpus, learning-centeredness, English for professionals

程世禄

广州大学华软软件学院

2-10-6

12月27日, 11:40—12:00

第二分会场

中美大学生英语定题演讲句式对比研究 ——一项基于语料库的研究

摘要: 本文分析了英语定题演讲研究的现状和存在的一些问题, 进一步阐明基于语料库作研究的优势、理论基础和可行性, 并以此为基础, 对中美大学生英语定题演讲句式, 进行了系统的对比分析, 这对于了解中美大学生英语定题演讲的异同, 量化实证这种句式的稳定性, 以及对提高我国英语教师教授这种体裁的能力等方面都具有十分重要的意义。

关键词: 体裁、英语定题演讲、中美大学生、语料库

A Syntactic Comparison of English Prepared Speeches between American and Chinese College Students—A Corpus-based Analysis

Abstract: In the paper, the author not only summarizes the previous researches and applications of

English prepared speeches in and out of China, but also analyses its current researches and questions. The paper further states the advantages, rationale, and feasibility of the corpus-based analysis to do the research. The author continues by making a systematic corpus-based comparison of English prepared speeches between American and Chinese college students. In this case, the ideas in this paper are significant to understand the similarities and dissimilarities of English prepared speeches of the American and Chinese college students on the syntactic level, prove the stability of English prepared speeches based on the quantitative analysis, and enhance our teachers' ability of teaching.

Keywords: genre, English prepared speeches, American and Chinese College students, corpus

张媛媛
中国矿业大学银川学院

第十一组

2-11-1

12月27日, 10:00—10:20

第三分会场

中国英语学习者与本族英语学习者情态动词 CAN 的对比研究 ——一项基于语料库的调查

摘要: 情态性是人类思维和语言的最基本的概念范畴, 对其研究无论在哲学领域抑或是在语言学范畴进行研究都是必要的。

传统的情态研究更注意研究客观的情态性, 即不考虑说话者的因素, 而现代的情态研究更注重情态的主观性, 从一种功能和语用的角度去来分析和研究人类存在的情态概念。

目前, 基于语言学的情态动词研究有三种研究模型, 即模糊意义模型, 多义模型, 和单义(或称之为核心意义)模型。本研究选取了核心意义模型, 基于 Pekins 的对 CAN 的分类标准, 把 CAN 根据语境划分为“能力”、“允许”“条件性可能”“建议”四个义项, 使用中国学生英语学习者语料库(CLEC)和国际英语学习者语料库(ICLE)中的两个子语料库即 ST6 语料库和 NATIVE 语料库进行考察。结果发现, 在四个义项中两个语料库的使用并不完全一致。这说明了语言系统的异同以及文化价值观念的差异会影响中国英语学习者对情态概念的掌握。本项研究不仅仅是一个基于语料库方法的基础性对比研究, 而且提出语料库在外语学习以及外语教学中的应用, 最后探讨了基于语料库的研究对中国英语教学的启示。

关键词: 情态性、情态动词 can、语料库方法、语义、外语教学

A Comparative Study of the Use of Modal Verb “CAN” between Chinese English Learners and Native English Learners—A Corpus-based Investigation

Abstract: Since modality is an essential category involved in human language and concept, it's quite necessary to work on it in the both fields of philosophy and linguistics.

Traditional study of modality focuses on objective modality rather than “speaker”, but current

study pays more attention to subjectivity of modality, and thus, analyses and studies the modal concepts of human beings from the functional and pragmatic perspective.

Three kinds of models are adopted in the current treatment of modality, i.e. ambiguity view, polysemy view and monosemy view (also called core-meaning approach). In this study, the monosemy view is proposed and then Perkins's classification of semantic categories of "can" is adopted as theoretical framework. Two sub-corpora deriving respectively from CLEC (Chinese Learner English Corpus) and ICLE (International Corpus of Learner English) are investigated. It is discovered that the uses of the four semantic categories of "can" are not always absolutely consistent, which reveals that the difference of two language systems and cultural values influences Chinese English learner's performance in SLA. The research is not only a fundamental research based on corpus, but also puts forward the application of corpus as a powerful tool to foreign language learning and teaching. At the end of this paper, the author discusses the implication of corpus-based study for the EFL learning and teaching in China.

Keywords: modality, modal verb *can*, corpus-based approach, semantics, foreign language teaching

朱红强
暨南大学

2-11-2

12月27日, 10:20—10:40

第三分会场

语料库在同义词辨析中的作用

摘要: 成年人二语词汇的发展依赖于第一语言的语义系统。成年二语学习者学习的新单词通常缺乏足够的语义内容, 他们通常会将一语单词的语义内容迁移到与之对应的二语单词中。因而成年二语学习者的心理词汇中充满了这样的二语单词, 它们的语义内容并非是学习者从二语学习的大环境中获得的, 而只是从第一语言的对应词汇中迁移过来的。这种捷径可以使学习者通过记忆单词表而迅速扩展二语词汇量, 但在运用这些单词时很有可能在语义方面出现错误。成年二语学习者在区分语义相近的词汇(尤其是在第一语言中具有相同对应翻译的二语单词)时情况更是如此。本研究使用句子填空任务, 进一步证实了在成人二语词汇习得的过程中存在着词汇发展与语义发展分离的现象, 并在实验中引入语料库来提高实验对象的同义词辨析能力。结果表明, 需要学习者进行深度语义加工的任务和明确地针对同义词的语义内容进行外显的教学有助于学习者更好地辨析同义词。

关键词: 语料库、同义词、语义迁移、句子填空任务

A Corpus-based Approach to Vocabulary Teaching: Helping EFL Students Differentiate between Near Synonyms

Abstract: L2 vocabulary development is to some extent dependent on the first language (L1) semantic system. L2 learners usually don't have enough semantic content in their newly gained words and therefore may use a "parasitic" strategy to transfer semantic information from their L1

words to the L2 equivalents. Thus, the L2 learner's mental lexicon is often rife with L2 word forms filled with L1 semantic content. This shortcut, while allowing a quick expansion of L2 vocabulary through learning glosses or word lists by rote, may cause many errors concerning semantic properties of the words in question. The difficulty of EFL learners in differentiating semantically close-related items (especially near synonyms in L2 with same translation in L1) is a case in point. In order to account for this difficulty, Jiang (2004b) used a sentence completion task to validate the separation of lexical and semantic development in adult L2 vocabulary acquisition. The present study replicates and builds on his study by introduction of linguistic corpora in an experiment intended to improve the participants' ability to differentiate between the near synonyms with L1 same translation. The results showed that tasks requiring depth of processing from the learners and explicit teaching on the semantic contents of these words indeed help learners make better differentiations.

Keywords: corpus, near synonym, semantic transfer, sentence completion task

王勇、杨晶

山东大学、山东建筑大学

2-11-3

12月27日, 10:40—11:00

第三分会场

基于语料库的连贯关系比较研究

摘要: 作为篇章的语义特征之一, 连贯性体现了文本内在的逻辑关系, 在很大程度上决定了文章的质量。连贯关系是实现文章连贯性的重要手段之一。然而, 以往对学习语篇连贯性的研究多拘泥于对文本中衔接手段这一表层特征的研究, 具有很大的局限性。本研究以 Mann 和 Thompson 的修辞结构理论 (RST) 为依据, 以《中国学生英语口语语料库》中的议论文作文为研究对象, 从语篇的局部连贯和整体连贯两个层面出发, 旨在研究: 在连贯关系的使用上, 高分作为和低分作文之间以及中高级学习者之间是否存在差异?

通过对文本的研究分析和统计计算, 得出以下结果:

首先, 在局部连贯层面 RST 理论中提出的 32 种关系在所选样本中都有所运用。高分作文中运用的连贯关系多于低分作文, 两者之间存在显著差异。但是, 这种数量差异在中高级学习者之间并不显著。另外, 不同组别所使用的高频率连贯关系非常接近。使用频率较高的连贯关系有阐述关系、非意愿性原因关系、非意愿性结果关系、评价关系、对照关系和证据关系。

其次, 在整体连贯层面, 中国英语学习者使用连贯关系的总体水平较高。但相对来说, 高分作文的学习者使用的连贯关系的水平要高于低分作文的学习者, 两者之间存在显著差异。通过纵向比较, 高级学习者使用连贯关系的水平明显优于中级学习者, 两者之间差异显著。

关键词: 修辞结构理论、连贯、英语作文

A Corpus-based Study of Coherence

Abstract: As a property of discourse, coherence represents the logic in a text, which, to a great extent, decides the quality of the text. Coherence is mainly realized through coherence relations (CRs). However, previous studies of coherence in EFL learners' writing have mainly focused on cohesive ties. With Mann and Thompson's Rhetorical Structure Theory (RST) as the theoretical foundation and argumentative compositions from the *Spoken and Written English Corpus of Chinese Learners* (SWECCCL) as the research data, the present study investigates CRs at both the local level and the global level, aiming to examine Chinese EFL learners' use of CRs in their English writing.

Through descriptive analysis and statistical calculation of the data, the findings are listed as follows:

Firstly, almost all the CRs in RST appear in the sampled compositions at the local level. More CRs are used in the good compositions than in the poor ones. This leads to significant difference between these two groups. However, the quantitative difference in using CRs between the advanced learners and the intermediate learners is not significant. Moreover, the high-frequency CRs used in different groups are quite close. *Elaboration*, *Non-volitional cause and result*, *Evaluation*, *Antithesis* and *Evidence* are the overwhelming relations.

Secondly, Chinese EFL learners have good mastery of CRs at the global level in general. However, significant difference exists both between the good compositions and the poor compositions, and between the advanced learners and the intermediate learners.

Keywords: RST, coherence, English composition

张利杰
安阳师范学院

2-11-4

12月27日, 11:00—11:20

第三分会场

基于语料库的外语教研平台 Direct Data

摘要: Direct Data 是一款专门针对国内英文教学和研究设计的以语料库为核心的教研平台, 语料库资源部分包括超过千万句对的双语句库、中国英文学习者语料库 CLEC、英文语料库 BNC/WSJ/NYT、搭配库、语段库、四元语言模型库、译文对照库、网络术语库等, 语料库工具集部分包括句子切分、词法分析、语段分析、搭配抽取、语料库统计和中国式英文识别等; 支持用户在平台上进行以语料库为支撑的教学和研究, 如在对比 CLEC 和 NYT 搭配库中 knowledge 前面的动词分布, 可以发现 CLEC 用户最常用的 10 个搭配用法中, 有约 7 个不标准, 如“learn ... knowledge”, “study ... knowledge”等, 明显的受到了汉语思维习惯的影响, 属典型的中国式英语。

关键词: 语料库、搭配、句法分析、中国式英语、教研平台

张跃
北京词网科技

2-11-5

12月27日, 11:20—11:40

第三分会场

基于语料库的《红楼梦》两个英译本的译者风格研究

摘要: 传统译者风格的研究大都基于个人的主观印象和直觉判断, 集中讨论译作的风格在何种程度上再现了原作的风格等。这种主观、印象式和规约性的方法存在诸多局限性, 如: 只重视译文与原文的风格对应, 而忽视文学翻译中译者特定的风格。随着描写翻译学的发展, 译者风格研究开始借助语料库的研究方法。通过大规模的语料分析, 来考察译者传承原作风格之外所表现出的风格, 这可以有效弥补主观判断的不足。本文以《红楼梦》的杨译本和霍译本为研究对象, 利用包含原著和两个译本全文的平行语料库, 基于大量语料的客观分析, 试图详细分析译者在词汇、句子和语篇三个层面上表现出来的独特风格。

语料库的实证考察表明, 杨宪益和夫人以及霍克斯在词汇、句子和语篇层面的确流露出原作风格之外的“痕迹”。总体上看, 在词汇层面, 杨宪益夫妇的用词更加复杂多样, 而霍克斯则更多地使用固定短语以符合目的语文化的规范。这使得霍译本呈现出明显的简单化倾向。其次, 在句子层面, 霍克斯倾向于使用长句来阐释原文, 并在 say 和 tell 等引述词后面使用 that 以及较多使用人称代词来增加句子之间的衔接度。这无疑增加了霍译本的明晰化程度。第三, 在语篇层面, 霍译本和杨译本都显示出了外显化的倾向, 但霍克斯倾向于使用介词短语作为衔接词, 而杨宪益夫妇主要使用连词。

最后, 本文试图从译者的母语文化和翻译规范等角度来解释二者表现出的迥异风格的成因。母语文化和翻译规范对译者风格的影响是无意识的。在翻译过程中, 译者通过无意识地利用自己在母语文化中形成的语言习惯以及有意识地遵从或违背特定历史时期的翻译规范而形成自己特定的风格。

关键词: 《红楼梦》、语料库翻译研究、译者风格、母语文化、翻译规范

赵晴

重庆邮电大学

2-11-6

12月27日, 11:40—12:00

第三分会场

基于语料库的中国学习者写作中单词“TRUE”的使用研究

摘要: 尽管中外研究者对单词“TRUE”有一些零散的研究, 如利用语料库分析“TRUE”和“REAL”的区别(王春艳 2009), 但是深入研究学习者如何使用单词“TRUE”则很少见。本研究采用对比分析学习者语料的方法, 对比分析中国英语学习者使用单词“TRUE”的基本特征和使用功能, 比较了中国英语专业学生、非英语专业学生和本族语学生在写作中使用单词“TRUE”上的差异, 并探讨母语知识和中国学生英语发音错误特征在使用和书写 TRUE 过程中的影响。

研究所用学习者英语语料为: 中国学习者英语语料库(CLEC)中四六级学生的作文子库、

语料均为英语专业的中国学生英语书面语语料(WECCL)、国际学生的英语书面语语料(ICLE)以及英美大学二、三年级学生的作文语料库(LOCNESS)。四者具有较强的可比性。对比结果可以提供关于中国学习者使用单词“TRUE”的规律和特点。

结果表明中国英语专业学生和非英语专业学生在 TRUE 使用上差别不大。与国际学习者和本族语者相比,中国学习者更倾向于使用 True 做定语,而国际学习者差异较大,TRUE 主要用在“it is true that…”和作状语的“True, …”。卡方检验显示三者有显著性差异($\chi^2=26$, $df=2$, $p=.0000$),中国学习者对“it is true that…”和作状语的“True, …”的使用远不如后两类学习者使用得多,存在过少使用的问题,说明中国学习者在习得英语本族语化的 TRUE 用法方面还比较欠缺。同时,中国英语学习者把 TRUE 拼错为“TURE”的频数远远高于本族语者。该研究还发现,非英语专业的学生拼写错误的频数亦远远高于英语专业学习者,说明拼写错误与学习者英语水平有着一定的正相关关系。

中国学习者的 TRUE 这种过度使用和过少使用可能是由于学习者所接触的不均衡的语言输入造成的。同时造成 True 拼写错误的原因比较复杂,求证于中国学习者英语口语语料库(COLSEC)和 SECOPETS 语料库中中国英语学习者错音现象,研究表明中国学生英语发音错误特征是造成 TRUE 拼错的主要原因,而且以江浙沪方言区为主。发音错误产生了负迁移,导致了学习者将 TRUE 误拼成 TURE 的现象。

对于 TRUE 的用法,一般语法书中都很少提及,学习者无法得到恰当的语言输入或提示,因此过多和过少使用反映了学习者内在语言知识的一个方面。也许由于我国英语教材对此类 TRUE 短语的输入不均造成学习者还没能较好地掌握它们,所以在语言表达中使用得较少或避免使用它们。

关键词: 中国学习者、英语写作、单词 TRUE 的使用

郑洪波
浙江万里学院

第十二组

2-12-1

12月27日,10:00—10:20

第四分会场

中国英语学习者的时态习得现状分析 ——基于 CEM 和 CLEC 的分析

摘要: 本文是基于英语专业四、八级考试写作和翻译语料库基础上的时态习得错误分析。通过对比分析 2004、2005、2006 年三年的英语专业学习四、八级写作和翻译语料,看出时态错误构成了学习者动词错误的主要部分,对于写作而言,最多的时态错误类型是主谓一致;对于翻译来说,最多的是时态选择错误;二者最小频率的时态错误类型都是时态形式本身错误(包括非目的语的时态形式错误和动词曲折变化上的错误)。此外,文章将上述结果与中国学习者英语语料库(CLEC)中学习者英语时态使用情况进行对比,发现学习者的时态使用情况似乎也没有与其不断提升的语言能力成正比,表现出一种较为淡薄的时态概念。

关键词: 时态、中国英语专业学习者语料库、中国英语学习者语料库、英语专业四八级考试、

写作、翻译

On the Tense Acquisition of Chinese English Learners —A Study Based on CEM and CLEC

Abstract: This paper aims to probe into the acquisition of English tenses by Chinese English learners based on the CEM (Corpus of English Majors) and CLEC. The paper first presents the analysis of tense errors in CEM especially with special reference to TEM 8 writing and translation. The findings show that for writing the most frequent tense error is the subject-verb agreement, and for translation it is the wrong tense choices, while for both the least frequent errors are non-target language tense forms and tense-related wrong verb forms. In addition, the results are further compared with the tense uses in CLEC to obtain a complete picture of Chinese English learners' acquisition of tenses. However the comparison reveals that Chinese English learners' use of English tenses does not progress accordingly along their overall language proficiency increase.

Keywords: English tenses, CEM, CLEC, TEM4 & TEM8, writing, translation

张雪梅、杨滢滢、冯辉
上海外国语大学

2-12-2

12月27日, 10:20—10:40
第四分会场

中国学生习得虚拟语气特征分析——基于语料库研究

摘要: 本研究通过以专业英语四级试卷和六级试卷中这两项考试真题及相关模拟试题为原始语料自建语料库,对虚拟语气的用法及考点出现的频率进行归类、统计、分析和总结,对比分析语料库中的虚拟语气出现频数及分布模式,学生问题主要集中在对虚拟语气形式分类及语言结构代表的语义和什么样的语义特征匹配,符号的表达和认知机制之间的联系等的认识上,造成问题的原因可能是语内迁移、语际迁移,教师的讲解及语法课堂教学模式设计,各种习题的设计等等。因此文章从认知的角度探讨英语虚拟语气的本质,句法关系的从属性,虚拟语气的语用内涵及运作机制。最后探讨了基于语料库的研究对英语虚拟语气语法教学带来的启示。

关键词: 语料库、虚拟语气、标准频数、问题

马艳姿、谢晓莺
南华大学、广西大学

2-12-3

12月27日, 10:40—11:00
第四分会场

A Study of the Use of the Conjunction *BECAUSE* in L2 Hong Kong University Students' Writings

Abstract: English writing is a way for information exchange. In order to achieve effective communication, language in the written texts needs to be clear and coherent. Writers usually use conjunctions as linguistic devices to clarify their ideas in a logical way, thus facilitating text coherence. Conjunctions can also help readers to organize, interpret, and evaluate the given information (Kopple 1985), and serve what Halliday (1973) calls the textual and interpersonal functions of language. If conjunctions are appropriately used, they can contribute greatly to the clarity and comprehensibility of discourse (Tyler 1994; Flowerdew & Tauroza 1995; Leech & Svartvik 1994). Despite the great importance of conjunctions, they can be easily neglected by language learners due to their optionality yet complexity in using. Previous studies have shown that the use of conjunctions is problematic for ESL and EFL learners in terms of overuse, underuse and misuse (Granger & Tyson 1996; Altenberg & Tapper 1998). Therefore, further investigations are necessary for finding out how L2 learners actually use certain conjunctions in their writings.

The present study will focus on the situations in Hong Kong where English is learnt as a second language. It will examine the use of a particular conjunction *BECAUSE* in university students' writings and compare it with the use in British university students' essays. A corpus-based approach is used for data collection and data analysis. Hopefully, the study will shed some light on Hong Kong L2 learners' actual use of conjunction *BECAUSE*, make some explanations towards learners' errors, and give some pedagogical implications.

Keywords: conjunction, corpus-based, L2 learners, writing

潘不寒
华中农业大学

2-12-4

12月27日, 11:00—11:20
第四分会场

中国学习者英语口语中人称代词与互动的研究

摘要: 近年来英语口语引起了越来越多的话语分析者的浓厚兴趣。大型口语语料库的出现无疑给英语口语的研究提供了新的方法和思路。代词和指示词能够向听者提供参照点,使其更好地理解言语事件。前人基于英语本族语口语语料库的研究表明人称代词“we”、“I”和“you”在出现频率、指称和话语功能上有很大差异,而这种差异与互动有关。在此基础上,本研究对中国学习者语料库中上述人称代词的使用情况进行调查,并与本族语语料库进行比较,以期发现中国学习者人称代词的使用特点及其与互动的关系。研究采用定量研究与定性研究相结合的方法。定量研究中将中国学习者英语口语语料库(COLSEC)、中国学生英语口语语料库(SECCL)和MICASE语料库构成A组语料库,对A组语料库中人称代词“we”、“I”和“you”的使用频率进行了比较分析;定性研究中在三个语料库里分别随机抽取5个口语语篇构成B组语料库,对B组语料库中上述人称代词的指称和话语功能进行了比较分析。研究结果表明英语口语中中国学习者与本族语者在人称代词“we”、“I”和“you”的使用频率、指

称和话语功能上都有很大差异，另外不同水平的中国学习者之间也存在差异。而这两种差异都与口语互动有关，呈现的趋势是英语水平越高其口语互动性越强。

关键词：中国学习者、英语口语、互动

On Personal Pronouns and Interaction in Chinese Learner's Spoken English

Abstract: More and more discourse analysts have great interest in the study of spoken English. And large – scale spoken English corpus provides new approaches and perspectives for the study. Pronouns, together with demonstratives, offers the hearer reference points to understand a speech event. The previous research of the native speaker's spoken English corpus shows that the personal pronouns, *we*, *I* and *you*, have great difference in frequency, referents and discourse functions. The difference is closely related with interaction. This research investigates the use of these pronouns in Chinese learner's spoken English corpora and compares them with that of the native speaker's spoken English corpus. The purpose is to explore the characteristics of Chinese learner's use of personal pronouns and its relationship with interaction. Both quantitative and qualitative methods are used in this research. With quantitative method, Corpus A, which consists of COLSEC (Chinese Learner Spoken English Corpus), SECCL (Spoken English Corpus of Chinese Learners) and MICASE (Michigan Corpus of Academic Spoken English), is established and the frequency of the personal pronouns, *we*, *I* and *you*, is analyzed. With qualitative method, Corpus B is established and it consists of five spoken English texts which are selected at random from the above three corpora. The referents and discourse functions of the three personal pronouns are analyzed. In conclusion, this research shows that there is great difference in frequency, referents and discourse functions of the personal pronouns in spoken English between Chinese learners and native speakers. Difference also exists in Chinese learners of different levels. Both of the two kinds of differences are closely related with interaction. The better a person learns English, the more interaction he achieves.

Keywords: Chinese learners, spoken English, interaction

王岩、吴丽萍
沈阳航空工业学院

2-12-5

12月27日，11:20—11:40

第四分会场

学生会话话轮转换处的会话填充语 ——一项基于语料库的研究

摘要：学生会话话轮转换技巧的提高是会话流畅的关键。本研究在介绍会话、话轮、话轮转换理论的基础上，基于《中国学生英语口语笔语语料库（2.0）》并抽取其中2003—2006年专业四级口试的48个样本（共958个话轮），重点研究学生会话中话轮转换处的会话填充语。研究发现，专业四级学生会话话轮转换处的会话填充语主要有四大功能：延缓回答、回应/应答、要求发话、表明观点。研究证明，专业四级学生已基本掌握捕捉话轮转换关联位置的能力，

同时也具备一定的话轮转换技巧,例如话轮获取技巧和话轮放弃技巧。但是交谈过程还欠缺自然和流畅性,主要表现在学生对母语语音词“Er, um, em, en, mmm, eh”进行直接迁移和过度使用语气词“Oh”。数据证明,学生还不能自如使用话轮转换处的会话填充语。对以后的口语教学,本研究建议教师多教授具体的会话填充语、或者自创相关的本族语者应用会话填充语的小型语料库,并提供情景让学生练习,使学生提高话轮转换技巧和运用英语的能力。

关键词: 学生会话、话轮转换处的会话填充语、话轮转换技巧

Fillers at the Juncture of Turns in Chinese EFL Learners' English Conversation: A Corpus-based Study

Abstract: Students should improve their turn-taking strategies in order to create a natural and fluent conversation. Based on 48 TEM-4 oral test samples from the *Spoken and Written English Corpus of Chinese Learners (2.0)*, this paper researches the conversation fillers in turn transitional places in students' conversation. This research finds out that there are four basic functions for the conversation fillers used by TEM-4 students: postponing responses, sending responses, requesting utterances, expressing opinions. The findings prove that TEM-4 students have got the basic skills of catching TRP (Transition Relevance Places), and also got some turn-taking skills, especially for turn-obtaining skills and turn-abandoning skills. However, TEM-4 students still cannot take conversations naturally and fluently, for example, they use too many Chinese phonological words “Er, um, em, en, mmm, eh” which are negative transfers of mother tongue and use too many mood words like “Oh”. It tells us that TEM-4 students don't know how to choose better words for conversation fillers in turn transitional places. Therefore, this research suggests that teachers provide more real uses of conversation fillers, or establish a mini-corpus which displays the uses of conversation fillers by native speakers, so that they provide students more input of turn-taking strategies.

Keywords: Chinese EFL learners' conversation, conversation fillers in turn transitional places, turn-taking skills

张祺

广东外语外贸大学

2-12-6

12月27日, 11:40—12:00

第四分会场

中国英语学习者英语介词 *at* 的使用研究 ——基于语料库的分析

摘要: 介词是英语词汇中的一个重要词类。而英语介词丰富的语意和多变的用法造成学习者学习的困难,如空间介词 *over, in, on*。本文以 *at* 为研究对象,借助英语本族语语料库 *Brown* 和中国学习者英语语料库 *CLEC* 来揭示两者在介词 *at* 使用上的差异,讨论差异产生的原因,提出教学上的建议。

我们借助语料库索引软件 *Antconc3.1.302*,对两个语料库进行检索和分析后发现:

1. 与本族语者相比, 中国英语学习者总体上明显少用介词 *at*; 但初级学习者有明显超用 *at* 的现象。

2. 两个语料库中 *at* 的类型符比显示中国英语学习者和本族语者对 *at* 的使用没有显著差异; 但对 *at* 的搭配词分析和语意分类却揭示出学习者与本族语者在 *at* 的使用上有明显差异。

造成学习者对介词 *at* 的使用与本族语者间差异的原因有: 母语的干扰, 教师课堂用语的影响, 对搭配规则的忽视等。

最后我们得出教学启示: 基于语料库的学习者中介语研究可以帮助教师掌握学习者与本族语者相比所出现的误用、超用和少用现象, 并借此确定教学的重点和难点。语料库的语境共现功能可以使学生透过大量的语言输入而增加对语言点的感知认识。

关键词: 中国英语学习者、介词 *at*、语料库

王颖

北京协和医学院

第十三组

2-13-1

12月27日, 14:00—14:20

第一分会场

Semi-automatic Retrieval of Collocation Errors based on Taiwanese Learner Corpora

Abstract: Collocation knowledge is difficult for second language learners to acquire, and L2 students often committed various collocation errors (cf. Nesselhauf 2003). The most common collocation error made by many ESL learners is the V-N collocation. To better understand various collocation errors made by ESL learners, it is essential to collect various collocation errors from learners' interlanguage. Traditionally, collocation errors were collected manually by language teachers and researchers. Though some collocation errors have been uncovered with this method, the manual error tagging method is a slow and inefficient way of uncovering ESL learners' collocation errors.

In this paper, we introduce a semi-automatic method of retrieving collocation errors based on computerized corpora (both a large native corpus and a learner corpus). A well-known corpus processing tool called MonoConc Pro was used to process the tagged BNC corpus. Various native collocation patterns can be uncovered quickly with the help of the tagged corpus. In addition, a 1.5 million word Taiwanese learner corpus was also used. The availability of this POS-tagged learner corpus allows researchers to retrieve various learners' collocations. After MonoConc Pro was used to successfully retrieve various collocational patterns from both BNC and the Taiwanese learner corpus, some Perl programs were also used to compare these two large sets of collocation patterns. The learner collocation patterns which were not found in BNC corpus were treated as possible collocation errors. These possible collocation errors were then examined by researchers to confirm if these collocations were indeed collocation errors.

With this new approach, more ESL learners' collocation patterns can be identified and collected. The large number of collocation errors retrieved from computerized learner corpus is very useful data for language teachers who want to create collocation teaching materials and for CALL software programs developers who want to develop collocation checkers or grammar checkers.

Keywords: collocation errors, MonoConc Pro, learner corpus, collocation teaching

陈浩然
国立台湾师范大学

2-13-2

12月27日, 14:20—14:40
第一分会场

英语增强词的语料库对比研究 ——以 very much 为例

摘要: 强势语在语言交际中起着强化语义、突出信息中心的重要作用。Quirk *et al.* (1985:589) 将强势语分为增强词和减弱词两大类。增强语的使用受到句法、语义、词汇和文本等方面的制约, 其复杂性使之成为有价值的探索课题。在文献查阅的基础上, 以 very much 为例, 运用语料库研究方法 (基于 CLEC 和 BNC), 从多维度调查分析学习者使用增强语的情况及问题所在。数据统计结果表明: (1) 学习者书面语中存在过多使用 very much 的趋势; (2) 语域差异显著。学习者书面语中使用 very much 的标准频次远远高于本族语者书面语, 与本族语者口语中使用的频率几近相同; (3) 类联接类型迥异。学习者使用的类联接类型数量少于本族语者, 后者频繁使用的 ADV N、ADV PREP 以及 ADV ADJR 的用法模式在学习者语料库中几乎未出现过, 学习者使用的某些模式为本族语者所不用; (4) 语义韵不同。学习者使用的 very much 呈明显错综语义韵, 而本族语者使用 very much 倾向于积极的语义韵; (5) 句法位置相左。学习者习惯将 very much 置于被修饰语之后, 而本族语者在多数情况下将其放置于被修饰语之前。综上所述, 学习者中介语状态下的增强语总体使用模式偏离目标语较远, 这与外语习得环境的缺乏、教学方式的不当、学习策略的误用以及母语负迁移干扰等因素密切相关。鉴于此, 应在外语教学和研究中积极引入语料库数据驱动学习理念和方法, 鼓励学习者在真实语境中去探索并发现语言事实, 从而习得地道、准确的语言。

关键词: 增强语、语料库、语域、类联接、语义韵、句法位置

陈颖
江苏工业学院

第十四组

2-14-1

12月27日, 14:00—14:20

第二分会场

中国大学生德语中介语语料库的创建及应用

摘要: 鉴于语料库语言学的快速发展和语料库建设的理论意义与应用价值不断上升, 我们于 2006 年开始为中国德语界的教与学及其与科研的互动构思和创建一个与中国汉语界、英语界、日语界以及德语国家的日耳曼学界等其它语料库具有相同共性, 但本身又自有构建特色的语料库。

出于对今后长期的研究和教学需要以及德语和德语国家背景特点的考虑, 我们构建的整个语料库将不仅汇集德语和汉语两种语言的语料数据, 而且涵盖普通语和专用语两大类。由这两个类别中的前者构成的子库——普通语语料库——的第一个模块是德语中介语语料库, 它将跨越不同的教学阶段和包含多个专业方向领域。

在以往取得的教学经验和科研成果基础上, 我们希望通过这一学习者语料库的创建和应用, 能够在诸多方面做出新的建树: 继续在研究著述中提出有理论意义和实践价值的观点; 在教学上有力推动新的方案和课程的设立; 培养出更多的精良人才和后继之秀; 争取在中国德语界发挥引领作用, 在校际和国际交流中展示创新实力。同时也为德语界提供一个促进教学与科研密切结合的服务平台。为此, 我们将在本报告中具体介绍这个学习者语料库的总体设计、建设进程和已开始的研究工作, 阐述为何和如何将至今语料库比较擅长的应用方式用于篇章语言学的研究领域, 并提出有关尝试设想。

关键词: 德语学习者语料库、普通语语料库、德语中介语

钱敏汝、李婧
北京外国语大学

2-14-2

12 月 27 日, 14:20—14:40

第二分会场

中医药实验性学位论文的英文摘要探析

摘要: 随着中医药走向国门, 中医药类论文的英语摘要在国内外学术交流中显的尤其重要。本文采用语料库研究方法, 分析我国中医药实验型学位论文的英文摘要。利用李文中教授主持的国家哲学社科基金项目中国英语语料库“中医摘要”部分建构子语料库, 该小型语料库包括: 2000 年至 2008 年在 CNKI 优秀硕、博论文库中发表的中医药类实验论文的英文摘要, 每年每库 30 篇 (共 540 篇)。本文主要进行如下研究: 1. 从宏观篇章语言学入手, 对常用语步进行定量分析并随机在硕博库中提取每年的两篇结构式英文摘要 (共 36 篇), 对话步内具体步骤进行观察, 建立基本语步模式并确定语步的具体意图。2. 对结构式英文摘要运用 XML 语言进行标注, 然后提取每个语步用 CLAWS 赋码。分析每个语步的语言特征: ①语法层面的时态、语态、人称和情态动词数量的定量分析, ②词簇分析, 以 3~5 词词簇为分析焦点。3. 每个语步的线性语言结构分析: 以高频动词为核心, 向左右跨距展开, 研究其搭配、类连接、语义倾向和话语韵, 探索发现较大的线性语言结构, 建构结构式摘要中每个语步的写作模板。通过从宏观语篇分析语步模式细化到微观词汇层面语法词簇分析, 再抽象到线性语言结构模式, 本文较全面细致地分析了中医药类学位论文的英文摘要; 既为教学提供了依据,

也使中医药学者能够更准确快捷地完成实验型论文的英文摘要,更好地进行国内外的学术交流。

关键词: 中医药类学位论文、英文摘要、语料库研究、语言结构模式

常哲
河南师范大学

第十五组

2-15-1

12月27日,14:00—14:20

第三分会场

基于语料库的翻译教学与学习者译本评析

摘要: 语料库应用于翻译研究至今已16年。在语料库翻译研究遍地开花的今天,反观肩负着翻译人才培养使命的翻译教学,大部分高校的英语专业仍然沿用传统的“理论—举例—结论”教科书式的教学方法。为了打破这种传统,本研究尝试以语料库翻译研究为理论根基,以翻译教学实践为切入点,探讨语料库在学生译本评析与翻译教学中的具体应用方法。

本研究利用的资源包括多个自建学习者翻译语料库和通用型语料库两种。自建学习者翻译语料库有:作者任教的三个年级5个班共133名学生的学习者的译文单语语料库(131,000词/字)、学习者译文一对多平行语料库(142,000词/字)、与专业译者译文语料库。直接使用的通用型语料库包括:《红楼梦》中英文双语平行语料库(2,781,166词/字)¹、《查泰来夫人和她的情人》英汉一对三平行语料库(706,000词/字),中国大学生英汉汉英口笔译语料库(PACCEL)(2,100,000词)与中国学习者英语语料库(CLEC)(1,000,000词)、英国国家语料库(BNC)(100,000,000词)。

本研究通过语料库检索与统计工具对学习者的译文语料库进行统计分析,从语义选择、句法结构、衔接成分、语义韵、词义褒贬、惯用搭配、高频词、意识形态及翻译策略等方面,探讨了学习者译文中存在的翻译共性特征,包括翻译过程的直观化、译文僵硬化与口语化等现象。除此之外,本文还通过翻译文本语料的呈现以及学习者译文的相互比较,制订了一套行之有效的基于语料库的翻译教学路线与译作讲评模式,使学习者能够在观察专业译者译文语料以及学习者译文语料之后,反思自己的翻译过程与翻译产品,归纳和总结出具有启发性的翻译策略及翻译方法,以便达到掌握翻译理论与翻译技巧的目的。

三年的实践表明,基于语料库的翻译教学方法,摒弃了以教师为中心的传统翻译教学方法语料虚假匮乏、过程单调乏味,讲评抽象表面的弊端。它符合学习者的认知方式,营造了以真实的翻译者及其翻译产品为中心的翻译培训环境,使翻译理论与翻译实践、翻译教学与翻译现实紧密结合,不仅有助于提高学生的翻译能力,而且有助于学生培养直观、科学的翻译批评思想、模式与方法,从而明显地提高翻译教学的效率与质量。

关键词: 翻译教学;语料库;学习者译文;评价

¹ 《红楼梦》中英文平行语料库为国家社科基金资助项目成果,包括《红楼梦》原文及其三个英语全译本,即Joly(1892-3)的前56回译本、霍克斯和闵福德(Hawkes & Minford 1973-86)的120回译本,以及Yang & Yang(1978-80)的120回全译本。

Translation Teaching and Learner Translation Assessment: A Corpus-based Methodology

Abstract: Translation Studies has seen the introduction of the corpus-based approach into its field of research for sixteen years. Albeit the fact that growing application has been made of the approach for various Translations Studies projects, seldom has it found its way into the translation classroom whereby the translation teaching has been following a “theory-demonstration-conclusion” routine. In order to make a breakthrough from this time-honoured practice, the authors of the present study attempted to experiment with a methodology which borrows the strengths of corpus-based Translation Studies theories to translation teaching, especially with emphasis on the assessment of the translations student produce.

This study reports how the experiment was conducted and what insights have been gained. The data of the experiment comprise two types of corpora: self-built learner translation corpora on the one hand, and readily-available or commercial corpora on the other. The first type of corpora consists of (a) a learner translation corpus (131,000 words in all) made up of the translations produced by three cohorts (totalling 133) of students attending the authors’ translation class, (b) a parallel multiple-version learner translation corpus consisting of both the source texts and their target texts produced by the same students (142,000 words). The second category of corpora is composed of two sub-types. The first includes two self-built parallel/comparable corpora: (a) the eighteenth-century Chinese classic novel *Hong Lou Meng* (*Dream of the Red Manison*) and its three English versions (2,781,000 words), and (b) D.H. Lawrence’s (in)famous novel *Lady Chatterley’s Lover* and its three target texts in Chinese (706,000 words). The second sub-type is made up of commercial corpora and they are used primarily for the purpose of contrast. These corpora are: the Parallel Corpus of Chinese EFL Learners (PACCEL, 2,100,000 words), Chinese Learner English Corpus (CLEC, 1,000,000 words) and British National Corpus (BNC, 100,000,000 words)

With the assistance of corpus-driving and statistical tools, learner translations are concordanced, categorized and analyzed for universal normative and behavioral features. These features range from semantic choice, syntactic structure, cohesive devices, semantic prosody, collocation, high frequency words, to choices of translation strategy and ideological orientation. In addition, this paper also established an effective corpus-based translation teaching process and translation assessment method based on in-class demonstration and comparison of learner translations. A comparison and contrast between learner translations and professional translations will facilitate the learners to reflect on their own translation process and translation product, which in turn leads to the learners’ cognition and summarization a set of instructive translation theories, strategies and techniques.

Three years experiment shows that the corpus-based translation teaching not only overcomes the shortcomings which characterize traditional translation instruction, such as the use of inauthentic teaching materials, the vapidity of the teaching process, and the impressionistic nature of the translation assessment, etc. More importantly, the new approach which goes in accord with the learner’s cognitive style construes a genuine learner- and product-oriented translation training environment where theory integrates with practice, and classroom teaching with the real world. The experiment proves beyond doubt that this approach is conducive to nurturing the learners’

translation ability, raising their awareness to a scientific and verifiable method of translation assessment, which all combines to contribute to the improvement of the quality and efficacy of translation teaching.

Keywords: translation teaching, corpus-based, learner translation, assessment

刘泽权、刘鼎甲
燕山大学

2-15-2

12月27日, 14:20—14:40

第三分会场

四级机考形势下大学英语写作教学新模式的研究与探索

摘要: 英语写作体现了学生的英语综合运用能力。目前大学非英语专业写作教学存在学时有限、授课班级规模偏大、没有专门的写作教学课, 现有的大英写作教学模式不易操作等现实问题, 不能系统化地对学生的英语写作进行培训, 以至于大多数学生对英语写作思维和方式缺乏深刻的认识, 写作时往往受到汉语母语负迁移的影响, 写作兴趣不高。本文旨在探索如何有效地利用有限的课时和多媒体、网络资源, 尤其是利用自建语料库和已有的语料库将大英写作教学融入、贯穿于大英读写和听力教学过程中, 分步、分段地进行系统性大英写作教学, 以激发学生写作兴趣, 提高学生写作水平。

关键词: 大学英语写作、写作教学模式、语料库

李渝
西南石油大学

第十六组

2-16-1

12月27日, 14:00—14:20

第四分会场

Examining the Parameters of Comparison in Learner English Research —Beyond “overuse” and “underuse”

Abstract: In the pioneering work by the contributors to Granger’s 1998 collection *Learner English on Computer* the terms of comparison ‘overuse’ and ‘underuse’ are frequently used to refer to the disparity of learner English and native speaker (NS) English in frequency. Since then these terms have become part of standard parameters for learner/NS English comparison. They have undoubtedly played their role in the study of learner English for the last decade. However, while more advanced studies in learner language come into place, it is becoming apparent that the

current parameters of comparison can hardly cover the detailed and different proportional relationship between the corpora compared. Take ‘overuse’ for example, in two comparable and normalised corpora of a learner corpus and a NS corpus, if the frequency of a certain word is used 100 times in the learner corpus but 95 times in the NS corpus, the proportional relationship can be labelled as ‘overuse’ by learners. If the frequency of another word is used 100 times in the learner corpus but 5 times in the NS corpus, the proportional relationship remains the same, i.e. “overuse” by learners. But do these two cases of ‘overuse’ by learners carry the same implication to researchers? Probably not. This study attempts to point out this ignored problem and proposes a set of much detailed proportional relationship in comparison. Two learner corpora with the information of marked levels of competency will be used to compare with a NS corpus to test the eligibility of the proposition while I explore the implications of each proportional relationship to researchers. The implications of this study to English pedagogy and testing will be discussed in the end.

Keywords: learner English research, parameters of comparison, learner corpora, proportional relationship

郭孝田

英国伯明翰新视野语言中心

2-16-2

12月27日, 14:20—14:40

第四分会场

中国学生英语写作中语法化连接词的研究

摘要: 基于传统语法, 国内外学者对英语学习者作文中的连接词已做了大量的理论方面的探讨和实证方面的研究, 功能语法的兴起更是把对学习者的作文中作为衔接手段之一的连接词的研究推向了高潮。与前人研究不同, 本研究则试图在语法化理论的框架下, 专门针对中国大学生英语写作中语法化了的连接词展开研究。本文通过对比中国大学生和英语本族语的英语书面语语料, 重点考察了中国大学生在写作中使用语法化连接词的特点, 及其与本族语者的异同。研究结果显示, 中国学生与本族语者在语法化连接词的使用上存在显著差异; 在具体词汇使用上反映出过度使用和使用不足并存的特点。分析进一步反映出中国大学生在写作中英语连接词的丰富度不够, 对英语连接词的把握程度欠佳。说明中国英语学习者对语法化连接词的理解掌握及恰当使用上可能还有一定难度。本研究的结论对我国大学英语词汇及写作教学具有一定的启示作用, 并为大学英语词汇表的编制提供一定反馈。

关键词: 语法化连接词、二语写作、语料库、英语词汇及写作教学

Grammaticalized Connectors in Essays of Chinese EFL Learners —A Corpus-based Study

Abstract: Based on traditional grammar, researchers at home and abroad have been dedicating themselves to theoretical discussions and empirical researches about connectors in EFL learners’ writings, which was later quite prosperous due to the rise of functional grammar. And different

from previous studies, this research aims to examine, under the framework of grammaticalization theory, the grammaticalized connectors in English writings of Chinese college students. This paper mainly studies the features of the grammaticalized connectors in the essays of Chinese EFL learners, with a comparison to English native corpus. The results show that there is significant difference between Chinese EFL learners and native speakers in the use of grammaticalized connectors in their writings; and Chinese EFL learners tend to either underuse or overuse certain connectors. Further analysis reflects some problems of Chinese EFL learners in the use of grammaticalized connectors in their writings. For instance, they exhibit inadequate knowledge of English connectors and have poor grasp of English grammaticalized connectors. It shows that Chinese EFL learners may have difficulties in understanding and using grammaticalized connectors. The findings of this study may shed lights on college English vocabulary and writing teaching in China.

Keywords: grammaticalized connectors, L2 writing, corpus, English vocabulary teaching, English writing teaching

陈功
北京外国语大学

参会代表索引