

计算语言学教育部重点实验室 2013 成果清单

一、论文

Mairgup Mansur, Wenzhe Pei, Baobao Chang, Feature-based Neural Language Model and Chinese Word Segmentation, 6th International Joint Conference on Natural Language Processing (IJCNLP-2013), 2013.

Wenzhe Pei, Dongxu Han, Baobao Chang, A refined HDP-based model for unsupervised Chinese word segmentation, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) ,8202 LNAI ,2013

Tao Ge, Baobao Chang, Sujian Li, Zhifang Sui, Event-based Time Label Propagation for Automatic Dating of News Articles, the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP 2013), 2013

马尔胡甫·曼苏尔, 裴文哲, 常宝宝, 基于深层学习的词义自动归纳, 第十二届全国计算语言学会议(CCL 2013), 2013

羅鳳珠、砂岡和子、姜柄圭、俞士汶、王雷、常寶寶，分階多語成語典故知識庫教學設計，《臺灣華語教學研究》，2013年，6期

俞士汶、羅鳳珠、朱學鋒、王雷、常寶寶、砂岡和子、姜柄圭，漢語成語及典故知識庫在語文學習中的應用，臺灣華語教學研究, 7(2), pp 13-36, 2013.12。

Tao Ge, Zhifang Sui, Baobao Chang, Exploiting collaborative filtering techniques for automatic assessment of student free-text responses, International Conference on Information and Knowledge Management, Proceedings (CIKM 2013), 2013

Dongxu Han, Baobao Chang, A Maximum Entropy Approach to Chinese Spelling Check, the 7th SIGHAN Workshop on Chinese Language Processing (SIGHAN-7), 2013.

董秀芳.“未 X”式副词的委婉用法及其由来，《语言科学》2012年第5期：477-488页。又收入张谊生、宗守云主编《汉语副词研究论集（第一辑）》，上海三联书店，2013年8月。

董秀芳.话题标记来源补议，《古汉语研究》2012年第3期：28-36页。又收入吴福祥、邢向东主编《语法化与语法研究（六）》，商务印书馆，2013年9月。

董秀芳. 领属转喻与汉语的句法和语篇, 《汉语学习》2012年第6期, 32-38页。又转载于人大复印资料《语言文字学》2013年第3期: 58-64页。英文版: Possessive metonymy in Chinese syntax and discourse, in *Language Communication in a Multicultural Context.*, North-Eastern Federal University, 2012.

董秀芳. 从现代汉语个体量词的名性特征看其内部差异, 《世界汉语教学》2013年第1期, 18-26页。又收入邵敬敏、黄宝珊主编《汉语语法研究的新拓展(六)》, 上海教育出版社, 2013年8月。

董秀芳. 趋向动词语法化为体标记: 来自汉语历史和方言之证据, 《现代中国语研究》第15期: 10-24页, 朝日出版社(日本), 2013年10月。

董秀芳. 词汇与句法的关联: 词义聚合与句法结构义聚合的平行性, 《语文研究》2013年第4期: 5-9页。转载于人大复印资料《语言文字学》2014年第3期: 35-41页。

董秀芳. 辭彙雙音化對動詞論元結構的影響, 《汉语史学报》第十三辑: 17-25页, 2013年12月。

Yansong Feng*, Mirella Lapata, Automatic Caption Generation for News Images. *IEEE Transaction on Pattern Analysis and Machine Intelligence* Vol35(4), pp.797-812, 2013 (WOS:000314931000003) (IF=5.694)

Junyang Rao, Aixia Jia, Yansong Feng*, Dongyan Zhao, Taxonomy Based Personalized News Recommendation: Novelty and Diversity. WISE 2013: 209-218 (EI: 20134716995887)

陈立玮, 冯岩松*, 赵东岩, 基于弱监督学习的海量网络数据关系抽取, 计算机研究与发展, 2013 年, 50 卷(第 9 期), 页码: 1825-1835. (EI: 20134116841929)

Li Sujian, Li Jiwei, Song Tao, Li Wenjie, 2013. A Novel Topic Model for Automatic term extraction, ACM SIGIR 2013, short paper. July 28, Dublin, 2013.

Li Jiwei, Li Sujian, Claire Cardie. 2013. Evolutionary Hierarchical Dirichlet Process for Timeline Summarization, ACL short paper, Aug. 3, 2013, Sophia, Bulgaria.

Li Jiwei, Claire Cardie, Li Sujian. 2013. A Topic-Model based Approach for Spam Detection, ACL short paper, Aug. 3, 2013, Sophia, Bulgaria.

Li Jiwei, Li Sujian*. 2013. A Novel Feature-based Bayesian Model for Query Focused Multi-document Summarization, Transactions of the Association for Computational Linguistics, 1 (2013) 89–98.

Wang Xun, Zhu Feida, Jiang Jing, Li Sujian. Real Time Event Detection in Twitter. WAIM 2013. Baoding, Hebei.

Ouyang Y., Li W.J., Li, S.J. and Lu Q., A Progressive Sentence Selection Strategy for Document Summarization, Information Processing and Management, Volume 49 Issue 1, pp 213-221, 2013. (EI, SCI)

陆勤, 谌贻荣, 李素建, “自下而上”与“自上而下”本体构建方法的探讨, 《当代语言学》第 15 卷, 2013 年第 2 期, 129-141 页.

Wei Wang, Sujian Li*, Jiwei Li, Wenjie Li, Furu Wei, Exploring Hypergraph-based Semi-supervised Ranking for Query-Oriented Summarization, Information sciences 237 (2013). 271-286. doi: 10.1016/j.ins.2013.03.012 (EI, SCI)

Sujian Li, Chen Wang, Jiawei Gu, A Novel Pairwise Learning to Rank Algorithm in Key Phrase Extraction, the Eighth International Conference on Innovative Computing, Information and Control (ICICIC2013), Kumamoto, Japan. Sep. 2013.(ICIC express, Vol.8, 2014) (EI)

Tao Ge, Baobao Chang, Sujian Li, Zhifang Sui, Event-based Time Label Propagation for Automatic Dating of News Articles, EMNLP 2013.

李素建, 王荀, 王宇昕, 内容标签和关系标签相结合的汉语篇章标注, CLSW 2013, 2013 年 5 月, 郑州, 河南.

王宇昕, 李素建, 篇章标注在医学领域问答系统中的应用, 苏州. CNCCL 2013.(accepted)

曹自强, 李素建, HDP 与互信息相结合的中文无指导分词, 苏州, CNCCL 2013 .(accepted)

陆顾婧, 陈刚, 刘扬. 汉语构词分析与词义知识表示研究. 第十四届汉语词汇语义学研讨会(CLSW2013), 郑州, 2013.

Mengxiang, Liu Yang, Wang Houfeng, Zhang Longkai. The Acquisition of Chinese Ergative Verbs and Its Usage in Semantic Role Labeling, CLSW2013, Zhengzhou, 2013.

Tao Ge, Baobao Chang, Sujian Li, Zhifang Sui. Event-based Time Label Propagation for Automatic Dating of News Articles. EMNLP 2013

Tao Ge, Zhifang Sui, Baobao Chang. Exploiting collaborative filtering techniques for automatic assessment of student free-text responses. CIKM 2013

Tao Ge, Wenjie Li, Zhifang Sui. A Semi-supervised Method for Opinion Target Extraction. WWW 2014

Xingxing Zhang, Tao Ge, Zhifang Sui. Learning to Extract Attribute Values from a Search Engine with Few Examples. CNCCL 2013

Sunyu Duan, Zhe Huang, Bingyan Liu, Bing Swen, “Distributional Similarity on Large Chinese Web Collection”, The 19th China Conference on Information Retrieval, 2013.

Sunyu Duan, Zhe Huang, Bingyan Liu, Bing Swen, “Improving Performance of Chinese Word Segmentation on Large Dataset”, The 19th China Conference on Information Retrieval, 2013.

黄劲松.一种基于搭配词多重聚类的人名消歧方法。第十四届汉语词汇语义学研讨会 CLSW2013, 2013年5月.

Weiwei Sun and Xiaojun Wan. 2013. Data-driven, PCFG-based and Pseudo-PCFG-based Models for Chinese Dependency Parsing. TACL.

Weiwei Sun, Xiaochang Peng and Xiaojun Wan. 2013. Capturing Long-distance Dependencies in Sequence Models: A Case Study of Chinese Part-of-speech Tagging. IJCNLP.

Xu Sun, Hisashi Kashima, Naonori Ueda. Large-Scale Personalized Human Activity Recognition using Online Multi-Task Learning. IEEE Transactions on Knowledge and Data Engineering (TKDE). 25(11): 2551-2563. IEEE. 2013.

Xu Sun, Takuya Matsuzaki, Wenjie Li. Latent Structured Perceptrons for Large-Scale Learning with Hidden Information. IEEE Transactions on Knowledge and Data Engineering (TKDE). 25(9): 2063-2075. IEEE. 2013.

Xinjie Zhou, Xiaojun Wan and Jianguo Xiao. Collective Opinion Target Extraction in Chinese Microblogs. In EMNLP2013. (Full Oral paper)

Xiaojun Wan and Fang Liu. Are Literature Citations Equally Important? Automatic Citation Strength Estimation and Its Applications. Journal of the American Society for Information Science and Technology (JASIST).

Weiwei Sun and Xiaojun Wan. Data-driven, PCFG-based and Pseudo-PCFG-based Models for Chinese Dependency Parsing. Transactions of the Association for Computational Linguistics(TACL). (Oral Presentation in ACL2013)

Xiaojiang Huang, Xiaojun Wan, Jianguo Xiao. Comparative News Summarization Using Concept-based Optimization. Knowledge and Information Systems (KAIS).

Xiaojun Wan. Subtopic-Based Multi-Modality Ranking for Topic-Focused Multi-Document Summarization. Computational Intelligence, In Press.

Yue Hu and Xiaojun Wan. PPSGen: Learning to Generate Presentation Slides for Academic Papers. In IJCAI2013.

Shanshan Huang, Xiaojun Wan, Xuewei Tang. AMRec: An Intelligent System for Academic Method Recommendation. In AAAI2013. (Late-breaking short paper)

Jiwei Tan, Xiaojun Wan, Jianguo Xiao. Learning to Order Natural Language Texts. In ACL2013. (short oral paper)

Xiaojun Wan. Co-Regression for Cross-Language Review Rating Prediction. In ACL2013. (short paper)

Shanshan Huang, Xiaojun Wan. AKMiner: Domain-Specific Knowledge Graph Mining from Academic Literatures. In WISE2013. (Full paper)

Weiwei Sun, Xiaochang Peng and Xiaojun Wan. Capturing Long-distance Dependencies in Sequence Models: A Case Study of Chinese Part-of-speech Tagging. In IJCNLP2013.

徐戈;蒙新泛;王厚峰;采用无标注语料的动词和形容词主观性评级, 软件学报,2013(05)

Mengxiang Wang, Yang Liu, Houfeng Wang, Longkai Zhang, The Acquisition of Chinese Ergative Verbs and Its Usage in Semantic Role Labeling. *Proceedings of the 14th Chinese Lexical Semantics Workshop (CLSW2013)*, LNAI 8229,2013

Zhengyan He, Shujie Liu, Mu Li, Ming Zhou, Longkai Zhang, Houfeng Wang. Learning Entity Representation for Entity Disambiguation. *Proceedings of the conference of the 51st Annual Meeting of the Association for Computational Linguistics (ACL 2013)*, 2013, Pp30-34

Longkai Zhang, Li Li, Zhengyan He, Houfeng Wang, Ni Sun. Improving Chinese Word Segmentation on Micro-blog Using Rich Punctuations. *Proceedings of the conference of the 51st Annual Meeting of the Association for Computational Linguistics (ACL 2013)*, 2013, Pp177-182

Mengxiang Wang, Houfeng Wang, Longkai Zhang, Judgment, Extraction and Selective Restriction of Chinese Eventive Verb, *2013 International Conference on Asian Language Processing (IALP 2013)*, 17-19, Aug, 2013, pp7-10, Urumqi, China

Xu Sun, N. Okazaki, J. Tsujii, Houfeng Wang Learning Abbreviations from Chinese and English Terms by Modeling Non-local Information. *ACM Transactions on Asian Language Information Processing (TALIP)*. 2013, 12(2)

Zhengyan He, Shujie Liu, Yang Song, Mu Li, Ming Zhou, Houfeng Wang. Efficient Collective Entity Linking with Stacking. *Proceedings of the conference on Empirical Methods in Natural Language Processing (EMNLP2013)*

Longkai Zhang, Houfeng Wang, Xu Sun, Mairgup Mansur. Exploring Representations from Unlabeled Data with Co-training for Chinese Word Segmentation . *Proceedings of the conference on Empirical Methods in Natural Language Processing (EMNLP2013)*.

李广一, 王厚峰, 基于多步聚类的汉语命名实体识别和歧义消解, 中文信息学报, 2013 (05)

万富强, 吴云芳. 2013. 基于中文维基百科的词语语义相关度计算. 《中文信息学报》 Vol 27(6): 31-37.

Yunfang Wu and Peng Jin. 2013. SemEval-2010: Disambiguating sentiment ambiguous adjectives. *Language Resources and Evaluation*, Vol. 47:743-755. (SCI Index)

Fei Wang, Yifeng Xu, Yunfang Wu and Xueqiang Lv. 2013. Predicting the semantic orientation of emoticons. *Journal of Computational Information Systems*, Pp 2391-2398. (EI Index)

Fei Wang, and Yunfang Wu. 2013. Exploiting hierarchical discourse structure for review sentiment analysis. *2013 International Conference on Asian Language Processing (IALP-2013)*, Pp 121-124.

吴云芳, 石静, 万富强, 吕学强. 2013. 汉语并列复句的自动识别方法. 《北京大学学报》(自然科学版)第 1 期, Vol.49(1):1-6.

石静, 吴云芳, 邱立坤. 2013. 基于大规模语料库的汉语词义相似度计算方法. 《中文信息学报》第 1 期, Vol.27(1):1-6.

赣语古全浊声母今读浊音的类型 《语言学论丛》第 47 辑, 商务印书馆 2013 年

汉语方言里的拾取义动词(上) 《民俗典籍文字研究》第 12 辑, 商务印书馆 2013 年

万年(石镇)方言同音字汇(合作; 第二作者: 詹澹) 《开篇》第 31 卷, [日本]好文出版 2013

汉语方言古全浊声母今读类型的地理分布: 分区角度的考察 《中国语言学》第 7 辑, 北京大学出版社 2013, 北京

说“冰雹” 《现代语言学》(OA 期刊)2013 年第 1 卷第 1 期

袁毓林、王明华。《作为方法论策略的零形式和零成分》, 《汉语学报》2013 年第 1 期, 第 19-30 页。中国人民大学复印报刊资料《语言文字学》2013 年第 6 期, 第 22-32 页。全文转载。

袁毓林。《“差点儿”中的隐性否定及其语法效应》, 《语言研究》2013 年, 第 2 期, 第 54-64 页。中国人民大学复印报刊资料《语言文字学》2013 年第 9 期, 第 46-57 页。全文转载。

魏雪、袁毓林。《基于语义类和物性角色建构名名组合的释义模板》, 《世界汉语教学》, 2013 年, 第 2 期, 第 172-181 页。

袁毓林。《试析“把”字句对述结式的选择限制》, 《木村英树教授还历纪念 中国语文法论丛》, 第 107-129 页。[日本]白帝社, 2013 年 5 月出版。

曹宏、袁毓林。《关于辞书中词性标注的几个问题》，《对外汉语研究》第9期，第142-153页，商务印书馆，2013年8月。

李新良、袁毓林。《面向计算的汉语动词蕴涵关系研究和型式库建设》，《中国社会科学》2013年第12期，第120-130页。中国人民大学复印报刊资料《语言文字学》2014年第3期，第116-127页。全文转载。

袁毓林。《基于生成词库论和论元结构理论的语义知识体系研究》，《中文信息学报》2013年第6期，第23-30页。

袁毓林。《形容词的语义特征和句式特点之间的关系》，《汉藏语学报》2013年，第7期，第147-166页。

袁毓林。《徜徉在应用需求和学理探究之间——数字化时代语言研究的困惑和蜕变》，《中国语言资源动态》2013年第2期（总第25期），第12-14页，教育部语言文字信息管理司指导，商务印书馆主办，内部刊物。

魏雪、袁毓林。《汉语名名组合的语义解释规律和释义模板库》，《语言学论丛》第48辑，第72-105页，商务印书馆，2013年。

Wang Tinghua, Zhao Dongyan and Liu Fulai, An Efficient Kernel Evaluation Criterion for Multiclass Classification, Chinese Journal of Electronics, 22(2):219-224 (2013)

陈立玮, 冯岩松, 赵东岩, 基于弱监督学习的海量网络数据关系抽取, 计算机研究与发展, 50(9): 1825-1835, 2013

Weiguo Zheng, Lei Zou, Yansong Feng, Lei Chen, Dongyan Zhao, Efficient SimRank-Based Similarity Join Over Large Graphs, 39th International Conference on Very Large Databases (VLDB), 2013

Weiguo Zheng, Lei Zou, Xiang Lian, Dong Wang, Dongyan Zhao, Graph Similarity Search with Edit Distance Constraint in Large Graph Databases, The 22nd ACM International Conference on Information & Knowledge Management (CIKM), 2013

Tinghua Wang, Junyang Rao, Dongyan Zhao, Using exponential kernel for word sense disambiguation. In: Proceedings of the 23rd International Conference on Artificial Neural Networks (ICANN), Sofia, Bulgaria, LNCS 8131, pp. 545–552, 2013.

Zhan, J., Yu, H., & Zhou, X. fMRI evidence for the interaction between orthography and phonology in reading Chinese compound words. *Frontiers in Human Neuroscience*. 2013, 7, 753.

Jiang, X., Li, Y., & Zhou, X. Is it over-respectful or disrespectful? Differential patterns of brain activity in perceiving pragmatic violation of social status information during utterance comprehension. *Neuropsychologia*. 2013, 51, 2210-2223.

Wei, P., Szameitat, A. J., Müller, H. J., Schubert, T. & Zhou, X. The neural correlates of perceptual load induced attentional selection: An fMRI study. *Neuroscience*. 2013, 250, 372-380.

Jiang, X., Li, Y., & Zhou, X. *Even a rich man can afford that expensive house*: ERP responses to construction-based pragmatic constraints during sentence comprehension. *Neuropsychologia*. 2013, 51, 1857-1866.

Yu, H., Zhou Z., & Zhou, X. The amygdalostriatal and corticostriatal effective connectivity in anticipation and evaluation of facial attractiveness. *Brain and Cognition*. 2013, 82, 291-300.

Luo, Y., Yan, M., & Zhou, X. Prosodic boundaries delay the processing of upcoming lexical information during silent sentence reading. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 2013, 39(3), 915-930.

Chen, Q., & Zhou, X. Vision dominates at the pre-response level and audition dominates at the response level in cross-modal interaction: Behavioral and neural evidence. *Journal of Neuroscience*. 2013, 33(17), 7109-7121.

Yu, H., Chen, X., Liu, J., & Zhou, X. Gum chewing inhibits the sensory processing and the propagation of stress-related information in a brain network. *PloS One*. 2013, 8(4), 557111

Wang, L., Yu, H., & Zhou, X. Interaction between value and perceptual salience in value-driven attentional capture. *Journal of Vision*. 2013, 13(3), 5, 1-13.

Wei, P., & Zhou, X. Attentional selection within and across hemispheres: Implications for the perceptual load theory. *Experimental Brain Research*. 2013, 225, 37-45.

Xu, X., Jiang, X., & Zhou, X. Processing biological gender and number information during Chinese pronoun resolution: ERP evidence for functional differentiation. *Brain and Cognition*. 2013, 81, 223-236.

Wu, Y., van Dijk, E., & Zhou, X. Evaluating self- vs. other-owned objects: The modulatory role of oxytocin. *Biological Psychology*. 2013, 92, 179-184.

Politzer-Ahles, S., Fiorentino, R., Jiang, X., & Zhou, X. Distinct neural correlates for pragmatic and semantic meaning processing: An event-related potential investigation of scalar implicature processing using picture-sentence verification. *Brain Research*. 2013, 1490, 134-152.

Weiguo Zheng, Lei Zou, Yansong Feng, Lei Chen, Dongyan Zhao. Efficient SimRank-based Similarity Join Over Large Graphs. International Conference on Very Large Data Bases, VLDB, 2013. Liwei Chen, Yansong Feng, Lei Zou, Dongyan Zhao. Explore Person Specific Evidence in Web Person Name Disambiguation

Lei Wang, Shiwen Yu, Shujing Li, Weiguang Qu. Construction and Application of the Knowledge Base of Chinese Multi-word Expressions. CLSW 2013, LNAI 8229 Springer-Verlag Berlin Heidelberg 2013, pp. 564–571.

Tongfei Chen, Weimeng Zhu, Xueqiang Lv, Junfeng Hu*. A Kalman Filter Based Human-Computer Interactive Word Segmentation System for Ancient Chinese Texts: Proc. The Twelfth China National Conference on Computational Linguistics (CCL 2013) .

Tongfei Chen, Xiaojun Zou, Weimeng Zhu, Junfeng Hu*. Human-Computer Interactive Chinese Word Segmentation: An Adaptive Dirichlet Process Mixture Model Approach. Proc. IJCNLP 2013.

Xiaojun Zou, Ni Sun, Junfeng Hu*. Diachronic Corpus Based Word Semantic Variation and Change Mining. Proc. International Conference Language Processing and Intelligent Information Systems (LP&IIS 2013).

Weimeng Zhu, Ni Sun, Xiaojun Zou, Junfeng Hu*. The Application of Kalman Filter Based Human-Computer Learning Model to Chinese Word Segmentation, Proc. 14th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing 2013). p218-230, 2013.

周志, 邹小军, 胡俊峰*. 基于带权复杂网络的关键词抽取研究, 第十四届汉语词汇语义学国际研讨会(CLSW 2013).

二、专著

语方言地理学——入门与实践 (项梦冰、曹晖合著), 中国书籍出版社, 2013 年, 北京。

《著名中年语言学家自选集·袁毓林卷》, 上海教育出版社, 2013 年, 第 1-526 页, 袁毓林。

三、科研项目

类型	批准部门	编号	负责人	时间	经费
----	------	----	-----	----	----

教育部人文社会科学研究项目	教育部	13YJA740060	吴云芳	2013.01- 2015.12	10 万
国家社科基金重大项目	国家社科基金	12&ZD227	吴云芳(子 课题)	2013.01- 2017.12.	18 万
国家自然科学基金面上项目	国家自然科学基金委	61273278	李素建	2013.1 - 2016.12	80 万
自然科学基金	基金委	61202233 基于叙事 模式分析的无监督 新闻事件语义抽取 研究	冯岩松	2013.1- 2015.12	25 万
国家自然科学基金	基金委	61273318	常宝宝	2013.1- 2016.12	80 万 元

自然科学基金	基金委	61272344 中文动态语义网构建技术研究	赵东岩	2013.1-2016.12	80 万
国家社科基金重大项目	全国哲学社会科学规划办	12&ZD227	王厚峰(首席)	2013-2017	80 万元
重大研究计划培育项目	国家自然科学基金委	91232708	周晓林	2013 年 1 月-2015 年 12 月	70 万元
省部级	北京市大学英语研究会	大学生英语写作冠词应用错误自动识别研究 (2013)	王雷	2013.7	2000 元
省部级	教育部计算语言学重点实验室 (北京大学)	汉语和英语多词表达中的隐喻研究 (201302)	王雷	2013.5	30000 元

自然科学基金	科技部	M1321005	项目申请书-评审人 计算机辅助指派关键技术研究	胡俊峰	2013 年 11 月 — 2014 年 10 月
--------	-----	----------	----------------------------	-----	--