CCL 2025

The 24th China National Conference on Computational Linguistic

Proceedings of the 24th China National Conference on Computational Linguistics

August 11 - August 14, 2025

Jinan, China

©The 24th China National Conference on Computational Linguistic

Order copies of this and other CCL proceedings from:

China National Conference on Computational Linguistic (CCL)

Courtyard 4, South Fourth Street, Zhongguancun , Haidian District, Beijing

100190, China

Tel: + 010-62562916

Fax: + 010-62661046

cips@iscas.ac.cn

Introduction

Welcome to the proceedings of the 24th China National Conference on Computational Linguistics (24th CCL). The conference and symposium were hosted and co-organized by Qilu University of Technology, China.

CCL is an annual conference (bi-annual before 2013) that started in 1991. It is the flagship conference of the Chinese Information Processing Society of China (CIPS), which is the largest NLP academic and industrial community in China. CCL is a premier nation-wide forum for disseminating new scholarly and technological work in computational linguistics, with a major emphasis on computer processing of the languages in China. The Program Committee selected 87 papers (60 Chinese papers and 27 English papers) out of 228 submissions for publication. For each of the submission, 4 reviewers were assigned in double-blind. More reviewers joined, when severe controversy appeared, aiding domain chair to give final decision. The final acceptance rate of CCL2025 is 38.16%. The 87 papers cover the following topics:

- Large Language Models (14)
- Machine Translation and Multilingual Information Processing (6)
- Text Generation, Dialogue and Summarization (8)
- Information Retrieval, Text Classification and QA (5)
- Language Resource and Evaluation (12)
- NLP Applications (19)
- ACL ARR Fast Track (3)
- Minority Language Information Processing (4)
- Knowledge Graph and Information Extraction (7)
- Social Computing and Sentiment Analysis (4)
- Fundamental Theory and Method of Language Computing and Cognition (5)

The final program for the 24th CCL represents the culmination of extensive efforts by our dedicated colleagues. We would like to express our sincere gratitude, first and foremost, to the authors who submitted their papers, thereby enriching the program with their contributions of high quality. We are profoundly indebted to all members of the Program Committee for their thorough and insightful reviews, executed within a constrained timeframe. Additionally, we extend our heartfelt appreciation to the conference sponsors for their generous support.

Furthermore, we wish to convey our special thanks to all members of the Organizing Committee and the secretariat for their diligent efforts in coordinating the conference, as well as to Springer for their assistance in ensuring the timely publication of the proceedings.

We hope that all participants found the CCL conference in Jinan to be an enjoyable and enriching experience.

August 2025

Maosong Sun, Peiyong Duan,
Zhiyuan Liu, Ruifeng Xu, Weiwei Sun

Organizers

24th CCL Conference Chairs

Maosong Sun	Tsinghua University, China
Peiyong Duan	Qilu University of Technology, China

24th CCL Program Chairs

Zhiyuan Liu	Tsinghua University, China
Ruifeng Xu	Harbin Institute of Technology, Shenzhen, China
Weiwei Sun	University of Cambridge

24th CCL Local Arrangement Chairs

Xiaoming Wu	Qilu University of Technology, China
Wenpeng Lu	Qilu University of Technology, China
Weiyu Zhang	Qilu University of Technology, China

24th CCL Evaluation Chairs

Hongfei Lin	Dalian University of Technology, China
Hongye Tan	Shanxi University, China
Bin Li	Nanjing Normal University, China

24th CCL Publication Chairs

Yubo Chen	Institute of Automation, CAS, China
Zhiliang Tian	National University of Defense Technology, China
Zhenghao Liu	Northeastern University, China

24th CCL Chair of Frontier Forum

Tong Xiao	Northeastern University, China
Xiaocheng Feng	Harbin Institute of Technology, China

24th CCL Workshop Chairs

Tong Xu	University of Science and Technology of China
Xu Chen	Renmin University, China

24th CCL Sponsorship Chairs

Kang Liu	Institute of Automation, CAS, China
Xianpei Han	Institute of Software, CAS, China

24th CCL Forum Chairs:

Baotian Hu	Harbin Institute of Technology, Shenzhen, China
Zhouhan Lin	Shanghai Jiao Tong University, China

24th CCL Hot Paper Chairs:

Hongyu Lin	Institute of Software, CAS, China
Zhuosheng Zhang	Shanghai Jiao Tong University, China
Wenqiang Lei	Sichuan University, China

24th CCL Publicity Chair

Peng Li	Tsinghua University, China
Xin Xin	Beijing Institute of Technology, China

24th CCL Website Chair

Jiali Zuo	Jiangxi Normal University, China
Jiasheng Si	Qilu University of Technology, China

24th CCL System Demonstration Chairs

Jinsong Su	Xiamen University, China
Fei Long	Inner Mongolia University, China

24th CCL Student Seminar Chairs

Shujian Huang	Nanjing University, China
Weixiang Zhao	Harbin Institute of Technology, China

24th CCL Finance Chair

Yuxing Wang	Tsinghua University, China

Table of Content

难度可控的词义辨析选择题自动生成

刘廷超，王雨，荀恩东.....1

基于多样性数据重组增强的藏汉神经机器翻译

薛嘉怡，陈锦明，陈波，鲍薇，赵小兵.....16

DeepSeek等大语言模型幽默生成能力及其特征的评测分析

蒋彦廷，应以周.....28

面向法律事件检测的大模型协同主动学习框架

崔婷婷，昝红英，籍欣萌，宋金旺，张坤丽，贾玉祥.....37

基于多维度答案筛选的低资源语言开放域问答方法

王新阳，关昕，张利飞，余正涛，黄于欣.....50

基于区域顶点标注的司法文本实体关系联合抽取

乐滢滢，孙媛媛，林鸿飞.....66

基于大模型增强的两阶段高效事件共指消解方法

吴耀宗，齐帅，王方圆，鲍琛龙，唐晋韬.....77

MIDF：基于模态交互和关系引导决策融合的多模态知识图谱补全

曾啸尘，赵晖，英迪.....89

基于大语言模型的中文医学命名实体识别

吕腾啸，罗凌，吕慧怡，孙媛媛，王健，林鸿飞.....102

例句质量评估体系构建及大语言模型例句生成能力评估

方明炜，朱君辉，鲁鹿鸣，杨尔弘，杨麟儿.....117

大语言模型汉字富语义能力评测

余艺喆，董明，何婷婷.....131

面向工艺规范的树结构检索增强生成方法研究

姜禹辰，王裴岩，冯煜博，余卓，纪贵阳.....143

基于检索增强生成的两阶段常识推理方法
李东洋，袁志勇，车超.....156
基于思维链和知识迁移的多语言问答推理研究
罗健，孙媛.....170
基于检索增强思维提示的汉语框架语义解析方法
李迎旭，陈涛，黎议泽，李斌阳.....184
基于数据合成的多模态讽刺隐喻理解大模型的构建
戴凌睿，李浩，吴云芳.....210
基于构式知识本体抽取评价名词的评价对象
周红照.....227
引入反思机制的机器译文质量估计方法
万洁，李茂西.....237
现代汉语同音字家族系统的效率与易学性：来自计算与模拟的证据
肖哲.....247
基于古汉语大语言模型的多任务学习探究
姚欣宇，王梦笛，高原，高歌，陈波，赵小兵.....259
TibLex：一种基于拉丁编码的藏文词表优化策略
更尕多杰，孙媛.....273
多模态嵌入的全局对齐增强下的基于强化学习的扩散模型
尤昊辰，刘宝静.....285
目标自适应的可解释立场检测：新任务及大模型实验
蓝伊，王子豪，陈波，赵小兵.....298
基于提示探针的大模型知识掌握能力评测
王淳昱，陈波，徐洋，赵小兵.....311
基于关系结构感知增强的知识图谱规则挖掘方法
徐会亲，黄琪，章程，刘祥棋，罗文兵，王明文.....323

AntIF:大语言模型抗干扰能力评估
罗雅晶，侯钰涛，陈云，陈冠华.....335
K-CoT：基于关键词思维链提示的中文排比句生成研究
钟茂生，甘家其，张鹤君，谢林康，李宏伟.....363
面向信息处理的中国手语音系学标注加工规范
赵源，金澎，敬思远，姚登峰，孟子文.....377
基于证据理论和局部语义区分的嵌套命名实体识别
徐波波，叶娜，蒋明翀.....390
基于多模型协同的儿童互联网新闻风险管理与价值观引导框架
梁宇蓝，王悦，于东，刘鹏远，康晨.....402
大语言模型和知识图谱协同的查询扩展方法
张旷，涂新辉，刘晗.....422
基于自提示多模态大语言模型和语义感知离散扩散模型的图像描述生成算法
陈宇峰，江爱文，黄琪，王明文.....434
融合MOE的多任务学习文档级企业新闻事件抽取
郑傲泽，张坤丽，王影，袁颂瑞，田怡豪，昝红英.....448
基于大语言模型多维度特征增强的医学命名实体识别方法
蒋明翀，叶娜，徐波波.....460
传统价值观成语当代语境表现分析———基于BCC语料库的计量研究
孙浩，刘洋洋，杜惠东，刘鹏远，于东，康晨.....474
基于LLM与跨语言嵌入的中亚低资源语言平行语料库构建方法
袁琦，阿力木木拉提.....488
基于双系统推理框架的法律判决研究
尹圣迪，白泽文，林鸿飞，杨亮.....500
面向对话场景的构式数据集
薛旭晶，李俊材，苏雪峰，杨沛渊，柴清华，李茹.....513

主题感知的多意图识别与槽位填充联合建模方法
罗晶，王炜华，曹越，飞龙，高光来.....525
中文阅读中的信息密度与认知资源动态分配研究
梁玉豪，钟晓路，杨泉.....537
Ti-MISO：基于TiLamb的藏文多模态生成式文本摘要
巩鑫，闫晓东，常浩远，田金超.....551
基于自监督表征蒸馏的Whisper低资源语音识别优化方法
胡剑，董凌，王文君，相艳，高盛祥，余正涛.....563
大语言模型可以分析花园幽径句吗？—基于跨语言数据集的实证研究
李琦，纪悦，李洪政.....574
基于依存树库的越南语依存距离研究
罗云，闫丹辉，马延周.....587
基于特征融合的大模型生成文本作者检测
赵晰莹，白梓萌，张妍，袁彩霞，王小捷.....598
基于个性化记忆策略的小参数语言模型高效对齐方法
朱孟笑，唐沛林，沙九，冯冲，拉马杰，闫丹智草.....613
人机价值观驱动的对话情绪生成模型
马志强，叶浩然，刘佳，吕凯.....627
基于强化学习的大语言模型古文释义选择研究
徐维潞，黄书剑.....640
历时演变视角下的古汉语分词：时期嵌入与大规模语料库的应用
柯永红.....651
控制句长的句子可读性研究：大语言模型驱动的数据集构建与评估
李罗希，李炜，邵艳秋.....666
基于细粒度时空建模的语音驱动手势生成模型
万浩聪，刘长红，杨海，江爱文，王明文.....684

跨语言方位词对“左-右”的语义衍化与语义关联模式探究
王梦焰，安纪元，杨麟儿，杨尔弘.....696
多领域翻译中语义消歧的话题方向盘方法
满志博，张玉洁，陈圆梦，陈钰枫，徐金安.....709
藏汉篇章机器翻译研究及语料库构建
田佳乐，江静，李亚超.....722
基于关联神经元识别的知识编辑方法
吴钰璋，穆永誉，王成龙，何荞至，肖桐，马安香，张春良，朱靖波.....733
基于大语言模型的立法文本可读性研究
胡钦谙，崔玉珍.....745
言行不一：大语言模型决策中的隐性偏见
林莘茹，李璐旸，刘湘婷.....756
语音相似性调节听觉词汇语义加工的脑电研究
耿立波，赵悦，陈茜.....769
HMUM：面向仇恨模因检测的多阶段多模态理解模型
裴淑娟，左家莉，何乐，万剑怡，王明文.....783
基于动态子空间重构的跨语言词向量对齐及应用
顾晓洋，胡玲，徐月梅.....792
Self-Supervised Contrastive Learning for Content-Centric Speech Representation
Jinlong Li, Ling Dong, Wenjun Wang, Zhengtao Yu, Shengxiang Gao.....807
RankLLM: A Multi-Criteria Decision-Making Method for LLM Performance Evaluation in Sentiment Analysis
Huzhi Xue, Butian Zhao, Haihua Xie, Zeyu Sun.....818
A Chunk-based Chain of Thought Prompting Method for Mitigating Over-Correction in Chinese Grammatical Error Correction
Xinquan Chang, Junguo Zhu.....831

Linguistic Differences between AI and Human Comments in Weibo: Detect AI-Generated Text through Stylometric Features

Ziqi Li, Qi Zhang.....842

Unveiling the Linguistic Acceptability Judgments of Large Language Models in Multilingual Contexts

Fuyu Xing, Haoyu Huang, Dawei Mo, Xinzhuo Yang, Zixuan Gao, Wei Wang, Zimu Wang, Haiyang Zhang.....852

Self-Preference: An Automated Method for Preference-Aligned Data Constructed from Business Metrics

Feng Gao, Xuan Zhang, Boyi Ni, Chunping Wang, Lei Chen.....864

MQM-MSC: Enhancing Translation Quality Estimation Interpretability with Mask-Driven Self-Correction in Large Language Models

Guanghui Cai, Junguo Zhu.....880

EDGE: Enhanced Debiased Gradient Extraction for Robust Fine-tuning

Jinglong Li, Kun Zhang, Chenyu Zou, Wei Shi, Xin Li, Si Wei.....890

Improving Abstract Reasoning Ability of Large Language Models through Mixture Program-based Data Synthesis

Yile Wang, Hui Huang.....904

Fine-tuning GEC Model Based on Language Family Corpus

Yitao Liu, Dras Mark.....922

Cross-modal Ambiguity Learning with Heterogeneous Interaction Analysis For Rumor Detection

Zhuo Fan, Qing Zhu, Yang Xiao.....934

BiSaGA: A Novel Bidirectional Sparse Graph Attention Adapter for Evidence-Based Fact-Checking

Junfeng Ran, Weiyao Luo, Zailong Tian, Guangxiang Zhao, Dawei Zhu, Longyun Wu, Hailiang Huang, Sujian Li.....946

RJAG: Retrieval Judgment Augmented Generation

Kuangzhi Wang, Zhenhua Hu, Min Ren, Xiangzhi Tao.....960

Statistically Optimized SGNS Model: Enhancing Word Vector Representation with Global Semantic Weight

Yulin Liu, Feng Xiong, Wanwei Liu, Minghui Wu.....972

Towards Coarse-to-Fine Evaluation of Inference Efficiency for Large Language Models

Yushuo Chen, Tianyi Tang, Erge Xiang, Linjiang Li, Wayne Xin Zhao, Jing Wang, Yunpeng Chai, Ji-Rong Wen.....985

MASP: A Multilingual Dataset for Probing Scalar Modifier Understanding in LLMs

Xinyu Gao, Nai Ding, Wei Liu.....1003

HFSD-V2C: Zero-Shot Visual Voice Cloning Via Hierarchical Face-Styled Diffusion Model

Yaping Liu, Linqin Wang, Shengxiang Gao, Zhengtao Yu, Ling Dong.....1020

CRAF:Cross-Modal Representation Alignment and Fusion for Speech Translation

Zhenbei Guo, Wenzhou Wu, Hua Lai, Yan Xiang, Yuxin Huang, Zhengtao Yu....1031

Beyond Instruction Following: Evaluating Inferential Rule Following of Large Language Models

Wangtao Sun, Chenxiang Zhang, Xueyou Zhang, Xuanqing Yu, Ziyang Huang, Haotian Xu, Shizhu He, Jun Zhao, Kang Liu.....1043

Lao-English Code-Switched Speech Synthesis Via Neural Codec Language Modeling

Yaping Liu, Linqin Wang, Shengxiang Gao, Zhengtao Yu, Ling Dong, Tian Tian.....1067

UMAD: Enhancing LLM Debiasing via Multi-Agent Debate and Token-Level Bias Interpretation

Hanwen Gu, Jie Ma, Ying Qin, Ling Hu.....1078

Instruction-Driven In-Context Learning for Domain-Specific Chinese Spelling Correction

Hyunsoo Park, Hongqiu Wu, Hai Zhao.....1095

TAG: Dialogue Summarization Based on Topic Segmentation and Graph Structures

Yatian Shen, Qichao Hao, Guosong Deng, Songyang Wang, Eryan Zhang.....1108

DualReward: A Dynamic Reinforcement Learning Framework for Cloze Tests Distractor Generation

Tianyou Huang, Xinglu Chen, Jingshen Zhang, Xinying Qiu, Ruiying Niu.....1122

Enabling Real-Time Conversations with Minimal Training Costs

Wang Xu, Haoyu Wang, Shuo Wang, Weilin Zhao, Xu Han, Yukun Yan, Haiyan Zhao, Yudi Zhang, Zhe Tao, Zhiyuan Liu, Wanxiang Che.....1136

DSMR-SQL: Enhancing Text-to-SQL with Dual-Strategy SQL Generation and Multi-Role SQL Selection

Yiming Huang, Jiyu Guo, Jichuan Zeng, Cuiyun Gao, Peiyi Han, Chuanyi Liu.1148

Quality-aware Neural Machine Translation with Self-evaluation

Jiajia Cui, Lingling Mu, Qiuhui Liu, Hongfei Xu.....1178