Curriculum Vitae (Zhijian Ou)

Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

Email: ozj@tsinghua.edu.cn

Homepage: http://oa.ee.tsinghua.edu.cn/ouzhijian

EDUCATION

- Department of Electronic Engineering, Tsinghua University, Beijing
  9/1998 - 7/2003   Ph.D. Thesis: Temporal Dependent Models for Speech Recognition
- Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai
  9/1994 - 7/1998   Bachelor

PROFESSIONAL EXPERIENCES

- 12/2024 – present Professor, Department of Electronic Engineering, Tsinghua University
- 12/2018 - present Co-founder and Chief Scientist of TasiTech
- 12/2008 - 11/2024 Associate Professor, Department of Electronic Engineering, Tsinghua University
- 08/2014 - 08/2015 Visiting Scholar, Beckman Institute, University of Illinois at Urbana-Champaign
- 08/2003 - 11/2008 Assistant Professor, Department of Electronic Engineering, Tsinghua University

HONORS AND AWARDS

1. Second Prize, *China Broadcast and Internet Media AI Innovation Challenge*, for the project "Innovative Application of Efficient Vietnamese Speech Translation Technology for Intelligent Program Content Analysis", 2025. (ranking 2nd in 10 members)
2. *Zheng-Junli Teaching and Education Outstanding Award*, Electronic Engineering Department, Tsinghua University, 2023.
3. *Zhang-Yujin Book Award for the monograph* "Energy-Based Models with Applications to Speech and Language Processing", Electronic Engineering Department, Tsinghua University, 2023.
4. *China Guangxi Provincial Science and Technology Award*, for the project "Key Technology Innovations and Applications for Intelligent IPTV Supervision", 2022. (ranking 4th in 10 members)
5. First Prize, *China Broadcast and Internet Media AI Innovation Challenge*, for the project "Applying AI Speech Technology to Intelligent Broadcasting Supervision of Fujian Dialects", 2021. (ranking 2nd in 10 members)
6. Best Paper Finalist, IEEE Spoken Language Technology Workshop (SLT) 2021, "Efficient Neural Architecture Search For End-To-End Speech Recognition Via Straight-Through Gradients" (Huahuan Zheng, Keyu An, Zhijian Ou).
7. *Advisor for Beijing Outstanding Undergraduate Thesis* "Mixed-feature Random Field Language Models with Applications to Speech Recognition" (Silin Gao), 2020.
8. *Advisor for Best Student Paper, ISCSLP (International Symposium on Chinese Spoken Language Processing)*, "Angular Softmax Loss for End-to-end Speaker Verification" (Yutian Li, Feng Gao, Zhijian Ou, and Jiasong Sun), 2018.
9. *Advisor for Tsinghua University Outstanding Ph.D. Thesis* "Statistical Language Models Based on Trans-dimensional Random Fields" (Bin Wang), 2018.
10. *China NRTA Science and Technology Innovation Award*, for the project "Regulatory System for Mobile TV over Digital Mobile Communication Networks", 2012. (ranking 4th in 8 members)

11. *First-class Award, Tsinghua University Young Teachers Teaching Contest*, 2012.
12. *Award for Mentoring Tsinghua University Outstanding Undergraduate Class*, 2010.
13. *Awards for Mentoring Tsinghua University Outstanding Undergraduate Thesis*, 2009, 2011, 2014, 2019, 2020, 2021, 2022.
14. *Awards for Mentoring Tsinghua University Outstanding Student Research Training (SRT) Project*, 2005, 2012.
15. *Best Paper, NCMMSC (National Conference on Man Machine Speech Communication)*, "Application of maximum likelihood linear regression to linear prediction HMMs" (Jinyin Wang, Zuoying Wang, and Zhijian Ou), 2005.
16. Led to achieve the best result for Chinese syllable recognition in the *863 Program evaluation* in 2003.
17. *Best Paper, Ph.D. Student Academic Forum*, Department of EE, Tsinghua University, 2002.

RESEARCH INTERESTS

Speech and language processing (particularly speech recognition and dialog systems) and machine intelligence (particularly with graphical models and deep learning). The broad goals are - proposing new models of speech and language, developing new methods for learning and inference with those models, and applying them to real-world tasks in speech and language processing.

RESEARCH PROJECTS AND GRANTS (SELECTED)

Government Funding
1. PI: Highly specialized vertical-domain large model data construction and training, *National Science and Technology Major Sub-Project 2023ZD0121401*, 2024-2026.
2. PI: Research and Development of Efficient AI Speech and Language Technologies for Minorities in Guangxi, *Guangxi Science and Technology funding*, 2023-2025.
3. PI: Strongly-generalizable human-machine dialogue technology for cross-domain large-scale applications, *China Mobile*, 2021-2023.
4. PI: Integration of Array Perception and Speech Recognition for Far-field Sound Sources, *National Science Foundation of China (NSFC)*, 2020-2023.
5. Co-PI: Constructing domain knowledge graph and building intelligent dialog system for customer service, *China Electric Power Research Institute*, 2019-2020.
6. PI: Customer service dialog system based on deep reinforcement learning and knowledge graph, *China Ministry of Education and China Mobile Joint Funds*, 2018-2019.
7. PI: Single-channel speech separation based on Probabilistic Acoustic Tube model, *National Science Foundation of China (NSFC)*, 2015-2018.
8. PI: Brain-inspired computer audition, *Tsinghua University Initiative*, 2015.
9. PI: Study of cognition about uncertainty in speech and language phenomena, *Tsinghua University Initiative*, 2012-2015.
10. PI: Robust speech recognition with Bayesian voice modeling and adaptive noise compensation, *National Science Foundation of China (NSFC)*, 2011-2013.
11. Co-PI: Novel research on content-based high performance audio search, *National High Technology Research and Development Program of China (863 Project)*, 2006-2008.
12. PI: Research on Refined-Structural Acoustic Modeling for Speech Recognition Based on Bayesian Networks, *National Science Foundation of China (NSFC)*, 2005-2007.
13. PI: Internet speech/audio information retrieval system, *China Ministry of Information Industry*,

2004-2005.

Enterprise Funding

1. PI: *TasiTech* funding, 2022 - 2025.

2. PI: *Meituan* funding, 2021 - 2022.

3. PI: *Apple* funding, 2020-2021.

4. PI: Research for highly accurate ultra large vocabulary continuous speech recognition technology, *Toshiba*, 2012-2016.

5. PI: Language Assessment for CET (College English Test) Spoken English Test, *IBM*, 2008-2009.

6. PI: Discriminative training for robust speech recognition, *Panasonic*, 2007-2008.

7. PI: Person Indexing and Retrieval for movies and TVs - audio, *Intel*, 2006-2008.

8. PI: Multimodal Soccer Video Analysis - audio, *Intel*, 2006-2008.

PUBLICATIONS (* corresponding authors)

Book

1. **Zhijian Ou**. Energy-Based Models with Applications to Speech and Language Processing. Foundations and Trends® in Signal Processing, Now publishers, 2024.

2. **Zhijian Ou**, Gang Li. Probability Theory and Stochastic Process (in Chinese). Tsinghua University Press, 2022.

Journal papers

1. Saierdaer Yusuyin, Te Ma, Hao Huang, **Zhijian Ou***. Pronunciation-Lexicon Free Training for Phoneme-based Crosslingual ASR via Joint Stochastic Approximation. *IEEE Transactions on Audio, Speech and Language Processing*, 2026.

2. Saierdaer Yusuyin, Te Ma, Hao Huang, Wenbo Zhao, **Zhijian Ou***. Whistle: Data-Efficient Multilingual and Crosslingual Speech Recognition via Weakly Phonetic Supervision. *IEEE Transactions on Audio, Speech and Language Processing*, 2025.

3. Ying Hu, Qin Yang, Wenbing Wei, Li Lin, Liang He, **Zhijian Ou**, Wenzhong Yang. MN-Net: Speech Enhancement Network via Modeling the Noise. *IEEE Transactions on Audio, Speech and Language Processing*, 2025.

4. Qiyan Song, **Zhijian Ou**. Modified Frequency-Sliding Generalized Cross Correlation for Time Delay Difference Estimation of Microphone Array. *IEEE Sensors Journal*, 2023.

5. Xinwei Zhang, Zhiqiang Tan*, **Zhijian Ou***. Persistently Trained, Diffusion-assisted Energy-based Models. *STAT*, Vol.12, Issue 1, 2023.

6. Hong Liu, Yucheng Cai, Zhenru Lin, **Zhijian Ou***, Yi Huang, Junlan Feng. Variational Latent-State GPT for Semi-Supervised Task-Oriented Dialog Systems. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, 2023, Vol.31, Page 970-984.

7. Dai, Yinpei, Yichi Zhang, Hong Liu, **Zhijian Ou***, Yi Huang, and Junlan Feng. Elastic CRFs for Open-Ontology Slot Filling. *Applied Sciences*, vol.11, 2021.

8. Yunfu Song, **Zhijian Ou***. Semi-supervised Seq2seq Joint-stochastic-approximation Autoencoders with Applications to Semantic Parsing. *IEEE Signal Processing Letters*, vol. 27, p.31-35, 2019.

9. Zhiqiang Tan*, Yunfu Song, **Zhijian Ou***. Calibrated Adversarial Algorithms for Generative Modeling. *STAT*, 2019.

10. Bin Wang, **Zhijian Ou***, Zhiqiang Tan*. Learning Trans-dimensional Random Fields with Applications to Language Modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2018, vol.40, no.4, pp.876-890.

11. Haotian Xu, **Zhijian Ou***. Scalable Discovery of Audio Fingerprint Motifs in Broadcast Streams with Determinantal Point Process Based Motif Clustering. *IEEE Transactions on Audio, Speech and Language Processing*, 2016, vol.24, no.5, pp.978-989.

12. Hui Lin, **Zhijian Ou***. Switching Auxiliary Chains for Speech Recognition. *IEEE Signal Processing Letters*, 2007, 14(8): 568~571.

13. Xianyu Zhao, **Zhijian Ou***. Closely Coupled Array Processing and Model-Based Compensation for Microphone Array Speech Recognition. *IEEE Transactions on Audio, Speech and Language Processing*, 2007, vol.15, no. 3, pp.1114-1122.

14. Hui Lin, **Zhijian Ou***, Xi Xiao. Generalized Time-series Active Search with Kullback-Leibler Distance for Audio Fingerprinting. *IEEE Signal Processing Letters*, 2006, vol.13, no. 8, pp.464-468.

15. Jun Luo, **Zhijian Ou***, Zuoying Wang. Eigenvoice-based MAP adaptation within correlation subspace. *Frontiers of Electrical and Electronic Engineering in China - Selected Publications from Chinese Universities*, Higher Education Press and Springer-Verlag, 2006, Vol.1, No.2: 130~134.

16. Jun Luo, **Zhijian Ou***. Efficient keyword spotting system for information retrieval. *Journal on Communication (in Chinese)*, 2006, 27(2): 113~118.

17. Xianyu Zhao, **Zhijian Ou***, Zuoying Wang. Using Vector Taylor Series with Noise Clustering for Speech Recognition in Non-stationary Noisy Environments. *High Technology Letters*, 2006, 12(1): 18~23.

18. Xianyu Zhao, **Zhijian Ou***, Zuoying Wang. Space Discriminative Function for Microphone Array Robust Speech Recognition. *High Technology Letters*, 2005, 11(4): 351~354.

19. Jun Luo, **Zhijian Ou***, Zuoying Wang. Two-Stage Keyword Spotting System based on Syllable Graphs. *Journal of Tsinghua University (in Chinese)*, 2005, 45(10): 1356~1359.

20. Xianyu Zhao, **Zhijian Ou***, Zuoying Wang. VTS-based robust speech recognition. *Journal of Tsinghua University (in Chinese)*, 2005, 45(7): 892~895.

21. Shucai Xiao, **Zhijian Ou***, Zuoying Wang. Speaker Clustering Algorithm for Speech Recognition. *Journal of Chinese Information Processing (in Chinese)*, 2005, 19(4): 84~88.

22. Xianyu Zhao, **Zhijian Ou***, Zuoying Wang. Using Word-Graph in Unsupervised Vector Taylor Series Algorithm for Robust Speech Recognition. *High Technology Letters (in Chinese)*, 2004, 14: 17~22.

23. Jun Luo, **Zhijian Ou***, Zuoying Wang. Speaker adaptive training in continuous speech recognition. *Journal of Chinese Information Processing (in Chinese)*, 2004, 18(3): 61~65.

24. **Zhijian Ou***, Zuoying Wang. Research on polynomial-fitting speech-trajectory model in Chinese continuous speech recognition. *Acta Electronica Sinica (in Chinese)*, 2003, 31(4): 608~611.

25. **Zhijian Ou***, Zuoying Wang. From linear prediction HMM to a new combined model for speech recognition. *Acta Electronica Sinica (in Chinese)*, 2002, 30(9): 1313~1316.

Conference papers (refereed)

1. Yucheng Cai, Yuxuan Wu, Yi Huang, Junlan Feng, **Zhijian Ou***. Knowledge Augmented Finetuning Matters in both RAG and Agent Based Dialog Systems. 2025 National Conference on Man-Machine Speech Communication (NCMMSC).

2. Te Ma, Nanjie Li, Hao Huang, **Zhijian Ou***. Phoneme-based speech recognition driven by large language models and sampling marginalization. 2025 National Conference on Man-Machine Speech Communication (NCMMSC).

3. Te Ma, Min Bi, Saierdaer Yusuyin, Hao Huang*, **Zhijian Ou***. LLM-based phoneme-to-grapheme for phoneme-based speech recognition. INTERSPEECH 2025.

4. Xiangzhu Kong, Hao Huang, **Zhijian Ou***. Lightweight and Robust Multi-Channel End-to-End Speech Recognition with Spherical Harmonic Transform. INTERSPEECH 2025.

5. Jiabo Jing, Ying Hu, Hao Huang, Liang He, **Zhijian Ou**. A Joint Network for Singing Melody Extraction from Polyphonic Music with Attention Aggregation and Self-Consistency Training. INTERSPEECH 2025.

6. Yucheng Cai, Ke Li, Yi Huang, Junlan Feng, **Zhijian Ou**∗. Entriever: Energy-based Retriever for Knowledge-Grounded Dialog Systems. ACL 2025 Findings.

7. Yucheng Cai, Si Chen, Yuxuan Wu, Yi Huang, Junlan Feng, **Zhijian Ou**∗. The 2nd FutureDial Challenge: Dialog Systems with Retrieval Augmented Generation (FutureDial-RAG). SLT, 2024.

8. Yuetong Zhao, Hongyu Cao, Xianyu Zhao, **Zhijian Ou**∗. An Empirical Study of Retrieval Augmented Generation with Chain-of-Thought. ISCSLP, 2024.

9. Lukuan Dong, Donghong Qin, Fengbo Bai, Fanhua Song, Yan Liu, Chen Xu, **Zhijian Ou**∗. Low-Resourced Speech Recognition for Iu Mien Language via Weakly-Supervised Phoneme-based Multilingual Pre-training. ISCSLP, 2024.

10. Wenbo Zhao, Ziwei Li, Chuan Yu, **Zhijian Ou**∗. CUSIDE-T: Chunking, Simulating Future and Decoding for Transducer based Streaming ASR. ISCSLP, 2024.

11. Xiangzhu Kong, Tianqi Ning, Hao Huang, **Zhijian Ou**∗. A Streaming Multi-Channel End-to-End Speech Recognition System with Realistic Evaluations. ISCSLP, 2024.

12. Yucheng Cai, Wentao Ma, Yuchuan Wu, Shuzheng Si, Yuan Shao, Zhijian Ou and Yongbin Li. UniPCM: Universal Pre-trained Conversation Model with Task-aware Automatic Prompt. LREC-COLING, 2024.

13. Hong Liu, Yucheng Cai, Yuan Zhou, **Zhijian Ou**∗, Yi Huang, Junlan Feng. Prompt Pool Based Class-Incremental Continual Learning for Dialog State Tracking. ASRU, 2023.

14. Yucheng Cai, Hong Liu, **Zhijian Ou**∗, Yi Huang, Junlan Feng. Knowledge-Retrieval Task-Oriented Dialog Systems with Semi-Supervision. INTERSPEECH, 2023.

15. Hong Liu, Zhaobiao Lv, **Zhijian Ou**∗, Wenbo Zhao, Qing Xiao. Exploring Energy-based Language Models with Different Architectures and Training Methods for Speech Recognition. INTERSPEECH, 2023.

16. Hong Liu, Yucheng Cai, **Zhijian Ou**∗, Yi Huang, Junlan Feng. Building Markovian Generative Architectures over Pretrained LM Backbones for Efficient Task-Oriented Dialog Systems. SLT, 2022.

17. Hong Liu, Hao Peng, **Zhijian Ou**∗, Juanzi Li, Yi Huang, Junlan Feng. Information Extraction and Human-Robot Dialogue towards Real-life Tasks: A Baseline Study with the MobileCS Dataset. EMNLP 2022 SereTOD Workshop.

18. Hong Liu, Yucheng Cai, **Zhijian Ou**∗, Yi Huang, Junlan Feng. A Generative User Simulator with GPT-based Architecture and Goal State Tracking for Reinforced Multi-Domain Dialog Systems. EMNLP 2022 SereTOD Workshop.

19. Keyu An, Ji Xiao, **Zhijian Ou**∗. Exploiting Single-Channel Speech for Multi-Channel End-to-End Speech Recognition: A Comparative Study. ISCSLP, 2022.

20. Yucheng Cai, Hong Liu, **Zhijian Ou**∗, Yi Huang, Junlan Feng. Advancing Semi-Supervised Task Oriented Dialog Systems by JSA Learning of Discrete Latent Variable Models. SIGDIAL, 2022.

21. Huahuan Zheng, Keyu An, **Zhijian Ou**∗, Chen Huang, Ke Ding, Guanglu Wan. An Empirical Study of Language Model Integration for Transducer based Speech Recognition. INTERSPEECH, 2022.

22. Keyu An, Huahuan Zheng, **Zhijian Ou**∗, Hongyu Xiang, Ke Ding, Guanglu Wan. CUSIDE: Chunking, Simulating Future Context and Decoding for Streaming ASR. INTERSPEECH, 2022.

23. Chengrui Zhu, Keyu An, Huahuan Zheng, **Zhijian Ou**∗. Multilingual and crosslingual speech recognition using phonological-vector based phone embeddings. ASRU, 2021.

24. Yunfu Song, Huahuan Zheng, **Zhijian Ou**∗. An empirical comparison of joint-training and pre-training for domain-agnostic semi-supervised learning via energy-based models. MLSP, 2021.

25. Keyu An, Yi Zhang, **Zhijian Ou**∗. Deformable TDNN with adaptive receptive fields for speech

recognition. INTERSPEECH, 2021.

26. Huahuan Zheng, Keyu An, **Zhijian Ou**∗. Efficient Neural Architecture Search for End-to-end Speech Recognition via Straight-Through Gradients. SLT, 2021. [Best Paper Finalist]

27. Fan Yu, Zhuoyuan Yao, Xiong Wang, Keyu An, Lei Xie∗, **Zhijian Ou**∗, Bo Liu, Xiulin Li, Guanqiong Miao. The SLT 2021 children speech recognition challenge: Open datasets, rules and baselines. SLT, 2021.

28. Yichi Zhang, **Zhijian Ou**∗, Huixin Wang, Junlan Feng. A Probabilistic End-To-End Task-Oriented Dialog Model with Latent Belief States towards Semi-Supervised Learning. EMNLP Long Paper, 2020.

29. Keyu An, Hongyu Xiang, **Zhijian Ou**∗. CAT: A CTC-CRF based ASR Toolkit Bridging the Hybrid and the End-to-end Approaches towards Data Efficiency and Low Latency. INTERSPEECH, 2020.

30. Yichi Zhang, Yinpei Dai, **Zhijian Ou**∗, Huixin Wang, Junlan Feng. Improved Learning of Word Embeddings with Word Definitions and Semantic Injection. INTERSPEECH, 2020.

31. **Zhijian Ou**, Yunfu Song. Joint Stochastic Approximation and Its Application to Learning Discrete Latent Variable Models. UAI, 2020.

32. Silin Gao, Yichi Zhang, **Zhijian Ou**∗ and Zhou Yu. Paraphrase Augmented Task-Oriented Dialog Generation. ACL Long Paper, 2020.

33. Silin Gao, **Zhijian Ou**∗, Wei Yang, Huifang Xu. Integrating discrete and neural features via mixed-feature trans-dimensional random field language models. ICASSP, Barcelona, Spain, 2020.

34. Yunfu Song, **Zhijian Ou**∗, Zitao Liu, Songfan Yang. Upgrading CRFs to JRFs and its benefits to sequence modeling and labeling. ICASSP, Barcelona, Spain, 2020.

35. Yichi Zhang, **Zhijian Ou**∗, Zhou Yu. Task-Oriented Dialog Systems that Consider Multiple Appropriate Responses under the Same Context. AAAI, New York, USA, 2020.

36. Hongyu Xiang, **Zhijian Ou**∗. CRF-based Single-stage Acoustic Modeling with CTC Topology. ICASSP, Brighton, UK, 2019.

37. Kai Hu, **Zhijian Ou**∗, Min Hu, Junlan Feng. Neural CRF Transducers for Sequence Labeling. ICASSP, Brighton, UK, 2019.

38. Bin Wang, **Zhijian Ou**∗. Improved training of neural trans-dimensional random field language models with dynamic noise-contrastive estimation. IEEE Workshop on Spoken Language Technology (SLT), Athens, Greece, 2018.

39. Yutian Li, Feng Gao, **Zhijian Ou**∗, Jiasong Sun. Angular Softmax Loss for End-to-end Speaker Verification. International Symposium on Chinese Spoken Language Processing (ISCSLP), Taipei, 2018. [Best Student Paper Award]

40. Zhangyu Xiao, **Zhijian Ou**∗, Wei Chu, Hui Lin. Hybrid CTC-Attention based End-to-End Speech Recognition using Subword Units. International Symposium on Chinese Spoken Language Processing (ISCSLP), Taipei, 2018.

41. Yichi Zhang, **Zhijian Ou**∗. Learning Sparse Structured Ensembles With Stochastic Gradient MCMC Sampling and Network Pruning. IEEE Machine Learning for Signal Processing (MLSP) Workshop, Aalborg, Denmark, 2018.

42. Yinpei Dai, **Zhijian Ou**∗, Dawei Ren, Pengfei Yu. Tracking of enriched dialog states for flexible conversational information access. ICASSP, Calgary, Canada, 2018.

43. Bin Wang, **Zhijian Ou**∗. Learning neural trans-dimensional random field language models with noise-contrastive estimation. ICASSP, Calgary, Canada, 2018.

44. Bin Wang, **Zhijian Ou**∗. Language modeling with neural trans-dimensional random fields. IEEE Automatic Speech Recognition and Understanding (ASRU) Workshop, Okinawa, Japan, 2017.

45. Yiyan Wang, Haotian Xu, **Zhijian Ou**∗. Joint Bayesian Gaussian discriminant analysis for speaker verification. ICASSP, New Orleans, USA, 2017.

46. Hongyu Xiang, Bin Wang, **Zhijian Ou**∗. The THU-SPMI CHiME-4 system: Lightweight design with advanced multi-channel processing, feature enhancement, and language modeling. CHiME Workshop, San Francisco, USA, 2016.

47. Haotian Xu, **Zhijian Ou**∗. Joint Stochastic Approximation Learning of Helmholtz Machines. International Conference on Learning Representations (ICLR) 2016 Workshop Track, Puerto Rico, USA, 2016.

48. Jinye Zhang, **Zhijian Ou**∗. Block-Wise MAP Inference for Determinantal Point Processes with Application to Change-Point Detection. IEEE Workshop on Statistical Signal Processing (SSP), Palma de Mallorca, Spain, 2016.

49. Ruobai Wang, Yang Zhang, **Zhijian Ou**∗, Mark Hasegawa-Johnson. Use of Particle Filtering and MCMC for Inference in Probabilistic Acoustic Tube Model. IEEE Workshop on Statistical Signal Processing (SSP), Palma de Mallorca, Spain, 2016.

50. Yang Zhang, **Zhijian Ou**∗, Mark Hasegawa-Johnson∗. Incorporating AM-FM effect in voiced speech for probabilistic acoustic tube model. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Paltz, NY, USA, 2015.

51. Bin Wang, **Zhijian Ou**∗, Zhiqiang Tan∗. Trans-dimensional Random Fields for Language Modeling. Annual Meeting of the Association for Computational Linguistics (ACL Long Paper), Beijing, China, 2015.

52. Bin Wang, **Zhijian Ou**∗, Jian Li, Akinori Kawamura. Joint-Character-POC N-gram Language Modeling for Chinese Speech Recognition. International Symposium on Chinese Spoken Language Processing (ISCSLP), Singapore, 2014.

53. Yang Zhang, **Zhijian Ou**∗, Mark Hasegawa-Johnson∗. Improvement of Probabilistic Acoustic Tube Model for Speech Decomposition. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 2014.

54. Xin He, **Zhijian Ou**∗, Jiasong Sun. Joint N-gram Chinese Language Modeling with an Application to Chinese Word Segmentation. IEEE International Conference on Audio, Language and Image Processing (ICALIP), Shanghai, 2012.

55. **Zhijian Ou**∗, Yang Zhang. Probabilistic Acoustic Tube: A Probabilistic Generative Model of Speech for Speech Analysis/Synthesis. International Conference on Artificial Intelligence and Statistics (AISTATS), La Palma, Spain, 2012.

56. **Zhijian Ou**∗, Huaqing Luo. CRF-based Confidence Measures of Recognized Candidates for Lattice-based Audio Indexing. ICASSP, Kyoto, Japan, 2012.

57. **Zhijian Ou**∗, Kan Deng. Combining Eigenvoice Speaker Modeling and VTS-based Environment Compensation for Robust Speech Recognition. ICASSP, Kyoto, Japan, 2012.

58. Yun Wang, **Zhijian Ou**∗. Combining HMM-based Melody Extraction and NMF-based Soft Masking for Separating Voice and Accompaniment from Monaural Audio. ICASSP, Prague, Czech, 2011.

59. **Zhijian Ou**∗, Ji Xiao. A Study of Large Vocabulary Speech Recognition Decoding Using Finite-state Graphs. International Symposium on Chinese Spoken Language Processing (ISCSLP), Tainan, Taiwan, 2010.

60. Yimin Tan, **Zhijian Ou**∗. Topic-weak-correlated Latent Dirichlet Allocation. International Symposium on Chinese Spoken Language Processing (ISCSLP), Tainan, Taiwan, 2010.

61. Yi Sun, **Zhijian Ou**∗, Wei Hu, Yimin Zhang. Excited Commentator Speech Detection with Unsupervised Model Adaptation for Soccer Highlight Extraction. IEEE International Conference on Audio, Language and Image Processing (ICALIP), Shanghai, 2010.

62. Qin Shi, Kun Li, Shilei Zhang, Stephen Chu, Ji Xiao, **Zhijian Ou**. Spoken English assessment system for

non-native speakers using acoustic and prosodic features. Annual Conference of the International Speech Communication Association (INTERSPEECH), Makuhari, Japan, 2010.

63. Nan Ding, **Zhijian Ou***. Variational nonparametric Bayesian hidden Markov model. ICASSP, Dallas, USA, 2010.

64. **Zhijian Ou***, Jun Luo. Eigenspace Estimation with Missing Values and its Application to Eigenvoice Adaptation for Speech Recognition. IEEE International Conference on Audio, Language and Image Processing (ICALIP), Shanghai, 2008.

65. Cong Li, **Zhijian Ou***, Wei Hu, Tao Wang, Yimin Zhang. Caption-aided speech detection in videos. ICASSP, Las Vegas, USA, 2008,4.

66. **Zhijian Ou***, Jun Luo. Latent correlation analysis of HMM parameters for speech recognition. ICASSP, Hawaii, USA, 2007,4.

67. Hui Lin, **Zhijian Ou***. Switching Auxiliary Chains for Speech Recognition based on Dynamic Bayesian Networks. International Conference on Pattern Recognition (ICPR), HongKong, 2006.

68. Hui Lin, **Zhijian Ou***. Partial-tied-mixture Auxiliary Chain Models for Speech Recognition Based on Dynamic Bayesian Networks. IEEE International Conference on Systems, Man and Cybernetics (SMC), Taipei, Taiwan, 2006.

69. Jinyin Wang, Zuoying Wang*, **Zhijian Ou***, "Application of maximum likelihood linear regression to linear prediction HMMs", National Conference on Man-Machine Speech Communication (NCMMSC), Beijing, China, 2005. [Best Paper Award]

70. Jun Luo, **Zhijian Ou***, Zuoying Wang*. Discriminative Speaker Adaptation with Eigenvoices. European Conference on Speech Communication and Technology (EUROSPEECH), Lisbon, Portugal, 2005.

71. Xianyu Zhao, **Zhijian Ou***, Minhua Chen, Zuoying Wang. Closely coupled array processing and model-based compensation for microphone array speech recognition. ICASSP, Philadelphia, USA, 2005.

72. **Zhijian Ou***, Zuoying Wang. Discriminative combination of multiple linear predictions for speech recognition. International Conference on Spoken Language Processing (ICSLP), Jeju, Korea, 2004.

73. **Zhijian Ou***, Zuoying Wang. A combined model of statics-dynamics of speech optimized using maximum mutual information. ICSLP, Denver, USA, 2002.

74. **Zhijian Ou***, Zuoying Wang. A new combined model of statics-dynamics of speech. ICASSP, Orlando, USA, 2002.

75. **Zhijian Ou***, Zuoying Wang. A New DP-like Speaker Clustering Algorithm. EUROSPEECH, Aalborg, Denmark, 2001.

PROFESSIONAL ACTIVITIES (SELECTED)

Memberships and Services

1. IEEE Senior Member (2014-present)
2. Senior Area Editor (01/2025-01/2028), Associate Editor (05/2019-05/2024), IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)
3. Editorial Board Member (11/2021-07/2025), Computer Speech and Language
4. Senior Area Editor (03/2024-03/2026), IEEE Signal Processing Letters
5. Chair of Workshops Sub-Committee (since 2022) and Committee Member (2020 - 2022, 2023 - 2025), IEEE Speech and Language Processing Technical Committee (SLTC)
6. Committee Member (2019-present), APSIPA Machine Learning and Data Analytics (MLDA) Technical Committee
7. Distinguished Member (2022-present), China Computer Federation (CCF)
8. Standing Committee Member (2018-present), CCF Speech Conversation and Auditory Technical

Committee

9. Committee Member (2021-present), Speech Acoustics and Hearing Branch, The Acoustical Society of China (ASC)

10. Steering Committee Member (2017-present), National Conference on Man-Machine Speech Communication (NCMMSC)

11. Proposal reviewer (2020-present), National Science Foundation of China (NSFC)

Conference Organization Committee Member/Seminar Organizer

1. General Chair, IEEE Spoken Language Technology Workshop (SLT) 2021

2. Tutorial Chair, INTERSPEECH 2020

3. SPS Liaison, IEEE Automatic Speech Recognition and Understanding Workshop (ASRU) 2021

4. SPS Liaison, IEEE Spoken Language Technology Workshop (SLT) 2022

5. General Chair, SereTOD (Towards Semi-Supervised and Reinforced Task-Oriented Dialog Systems) Workshop, Co-located with EMNLP 2022

6. Advisory board, IEEE Spoken Language Technology Workshop (SLT) 2024

7. General Chair, FutureDial-RAG (Retrieval Augmented Generation) Challenge and Special Session, Co-located with SLT 2024

8. Technical Program Chair, International Symposium on Chinese Spoken Language Processing (ISCSLP) 2024

9. Youth Forum Chair, National Conference on Man-Machine Speech Communication (NCMMSC) 2023, 2024, 2025.

10. Oct. 18, 2025: Organize Special Session at NCMMSC 2025 "The 2nd Seminar on Semi-Supervised Speech and Language AI Technologies" (SALT-Semi 2nd), Zhenjiang, China.

11. Aug. 16, 2024: Organize Special Session at NCMMSC 2024 "The 2nd Seminar on Multilingual Speech and Language AI Technologies" (SALT-Multilingual 2nd), Urumqi, China.

12. Dec. 9, 2023: Organize Special Session at NCMMSC 2023 "The 1st Seminar on Semi-Supervised Speech and Language AI Technologies" (SALT-Semi), Suzhou, China.

13. Sep. 23, 2023: Organize CCF workshop "The 1st Seminar on Multilingual Speech and Language AI Technologies" (SALT-Multilingual), Nanning, China.

Conference Senior Program Committee member/Area Chair/Meta reviewer

1. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2020, 2021, 2022, 2023, 2024, 2025, 2026)

2. International Joint Conference on Artificial Intelligence (IJCAI 2021)

3. Empirical Methods in Natural Language Processing (EMNLP 2022 Track on Dialogue and Interactive Systems)

4. IEEE Spoken Language Technology Workshop (SLT 2022, 2024)

5. IEEE Automatic Speech Recognition and Understanding Workshop (ASRU 2023, 2025)

Conference Session Chair

1. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2020, 2021, 2022, 2024)

2. Annual Conference of the International Speech Communication Association (INTERSPEECH 2020, 2022, 2025)

3. IEEE Spoken Language Technology Workshop (SLT 2024)

4. IEEE Automatic Speech Recognition and Understanding Workshop (ASRU 2021)

5. International Symposium on Chinese Spoken Language Processing (ISCSLP 2010, 2018)

Conference Technical Program Committee Member/Scientific Committee Member/Reviewer

Curriculum Vitae (Zhijian Ou)

1. Neural Information Processing Systems (NeurIPS 2018, 2020, 2021, 2022, 2023, 2024)
2. International Conference on Learning Representations (ICLR 2019, 2020, 2021, 2022, 2023, 2024, 2025)
3. International Conference on Machine Learning (ICML 2021, 2022, 2023, 2024, 2025)
4. Association for the Advancement of Artificial Intelligence (AAAI 2021, 2022, 2023)
5. International Joint Conference on Artificial Intelligence (IJCAI 2021)
6. Association for Computational Linguistics (ACL 2020, 2021, 2023)
7. Empirical Methods in Natural Language Processing (EMNLP 2020, 2021)
8. ACL Rolling Review (Dec 2021, Jan 2022, 2025)
9. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026)
10. Annual Conference of the International Speech Communication Association (INTERSPEECH 2018, 2019, 2020, 2022, 2023, 2025, 2026)
11. IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2017, 2018, 2019, 2020, 2021)
12. International Workshop on Acoustic Signal Enhancement (IWAENC 2018)
13. International Symposium on Chinese Spoken Language Processing (ISCSLP 2010, 2012, 2014, 2016, 2018, 2020, 2022)
14. Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL-HLT 2021)
15. Conference on Language Modeling (COLM 2024)
16. International Workshop on Speech Processing in Everyday Environment (CHiME 2018, 2020, 2023)
17. Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2017, 2018, 2019)
18. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2008)
19. IEEE International Symposium on Circuits and Systems (ISCAS 2009)
20. Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA 2014, 2017)
21. IEEE International Conference on Audio, Language and Image Processing (ICALIP 2010, 2012, 2014)
22. IEEE Data Science and Learning Workshop (DSLW 2021)

Journal Reviewer

1. IEEE Transactions on Audio, Speech and Language Processing
2. IEEE Transactions on Systems, Man and Cybernetics
3. Journal of Machine Learning Research (JMLR)
4. IEEE Signal Processing Letters
5. The Journal of the Acoustical Society of America
6. Science China
7. Chinese Journal of Electronics
8. Acta Automatic Sinica
9. Journal of Computer Science and Technology
10. Journal of Chinese Information Processing

External PhD Thesis Review/Committee

1. Jilu Jin, 2026, Northwestern Polytechnical University
2. Shanzheng Guan, 2026, Northwestern Polytechnical University
3. Yuze Zhou, 2025, Rutgers University, USA
4. Hua HUA, 2025, University of Chinese Academy of Sciences
5. Shaojing ZHAO, 2025, University of Chinese Academy of Sciences

6.   Xiaowei Zhang, 2025, University of Chinese Academy of Sciences

7.   Zehui YANG, 2025, University of Chinese Academy of Sciences

8.   Yuan XIE, 2025, University of Chinese Academy of Sciences

9.   Shengchang XIAO, 2025, University of Chinese Academy of Sciences

10. Jiakun SHEN, 2025, University of Chinese Academy of Sciences

11. Han Zhu, 2024, University of Chinese Academy of Sciences

12. Yifan Chen, 2024, University of Chinese Academy of Sciences

13. Chengxin Chen, 2024, University of Chinese Academy of Sciences

14. Weihang Nie, 2024, University of Chinese Academy of Sciences

15. Zihao Cui, 2022, Beijing University of Technology

16. Yingke Zhao, 2022, Northwestern Polytechnical University

17. Xin Leng, 2022, Northwestern Polytechnical University

18. Yerbolat Khassanov, 2020, Nanyang Technological University, Singapore

19. Feng Zhou, 2020, The University of New South Wales, Australia

20. Peng Lin, 2017, The University of New South Wales, Australia

21. Yingying Liu, 2016, The University of New South Wales, Australia

Conference Tutorials

• Zhijian Ou, "Energy-Based Models with Applications to Speech and Language Processing", ICASSP, Singapore, 2022.

• Zhijian Ou, "State-of-the-Art of End-to-End Speech Recognition", The 6th Asian Conference on Pattern Recognition (ACPR), Jeju Island, South Korea, 2021.

• Zhijian Ou and Bin Wang, "Language Modeling: State of the Art", International Symposium on Chinese Spoken Language Processing (ISCSLP), Taipei, 2018 (Tutorial T-1).

• Zhijian Ou, "Undirected Graphical Models: Theory and Applications to Speech and Language Processing", International Symposium on Chinese Spoken Language Processing (ISCSLP), Tianjin, China, 2016 (Tutorial T-5).

Invited Talks

1.   Dec 10, 2025: Invited short course organized by State-owned Assets Supervision and Administration Commission of the State Council (国资委): "The Current Status, Trends, and Challenges of Artificial Intelligence Technology" (人工智能技术发展现状、趋势与挑战), Dalian, China.

2.   Nov. 7, 2025: Invited Talk "Phonemes as Speech-Language Interface for Data-Efficient Multilingual and Crosslingual Speech Recognition", at Huawei Media Technology Summit, Shenzhen, China.

3.   Nov. 2, 2025: Invited Talk "Are Large Models the Watt Steam Engine of the New Industrial Revolution?", at Video and Image Security in the Era of Large Models Forum, The 13th International Conference on Image and Graphics (ICIG 2025), Xuzhou, China.

4.   Oct. 18, 2025: Invited talk "Phonemes as Speech-Language Interface for Data-Efficient Multilingual and Crosslingual Speech Recognition", Special Session at NCMMSC 2025 "The 2nd Seminar on Multilingual Speech and Language AI Technologies" (SALT-Multilingual 2nd), Zhenjiang, China.

5.   Aug. 15, 2025: Invited short course at Guangxi Radio and Television Monitoring Center: "Reliable and Efficient Artificial Intelligence Large Model Speech and Language Technologies" (可靠高效的人工智能语音与语言大模型技术), Nanning, China.

6.   May 9, 2025: Invited short course at State-owned Assets Supervision and Administration Commission of the State Council (国资委): "The Development History and Trends of Artificial Intelligence" (人工智能发展历程及趋势), Beijing, China.

7.   Nov. 19, 2024: "High-Efficiency And High-Reliability Large Model Man-Machine Dialogue Technology

For Vertical Domains" (面向专业场景的高效高可靠大模型人机对话技术), China Information Communication Conference & Annual Conference for China Institute of Communications (CiCC) (中国信息通信大会暨中国通信学会学术年会), Chengdu, China.

8. Oct. 11, 2024: "Reliable and Efficient Artificial Intelligence Speech and Language Technologies (高可靠高效益的大模型语音与语言技术)", Nankai University, TianJin, China.

9. Sep. 20, 2024: "Reliable and Efficient Artificial Intelligence Speech and Language Technologies (高可靠高效益的大模型语音与语言技术)", CUHK, HongKong, China.

10. Sep. 19, 2024: "Reliable and Efficient Artificial Intelligence Speech and Language Technologies (高可靠高效益的大模型语音与语言技术)", CUHK-ShenZhen, China.

11. Aug. 16, 2024: Invited talk "Whistle: Data-Efficient Multilingual and Crosslingual Speech Recognition via Weakly Phonetic Supervision", Special Session at NCMMSC 2024 "The 2nd Seminar on Multilingual Speech and Language AI Technologies" (SALT-Multilingual 2nd), Urumqi, China.

12. Jan. 5, 2024: "The success of large models and principled semi-supervised learning (大模型的成功与原则性半监督学习)", South China University of Technology, GuangZhou.

13. May. 30, 2024: Invited short course for China Mobile: "Artificial Intelligence Development Trends and Industry Application Practices (人工智能发展趋势与行业应用实践)" (2 hour), Beijing, China.

14. Nov. 19, 2023: Invited Talk at SpeechHome Conference on Speech Technology 2023: "Some Thoughts and Projections on Speech Foundation Models", Beijing, China.

15. Oct. 14, 2023: Invited Talk at CCF International Conference on Natural Language Processing and Chinese Computing (NLPCC) 2023: "Knowledge-Retrieval Dialog Systems with Semi-Supervision", Foshan, China.

16. July 8, 2023: Invited talk at 2023 World Artificial Intelligence Conference (WAIC) (世界人工智能大会): "The philosophy of semi-supervision and alignment behind ChatGPT", Shanghai, China.

17. June 9, 2023: Panel Discussion at 2023 BAAI Conference for the session "Large-scale model new infrastructure and intellectual operation", Beijing, China.

18. Apr 26, 2023: Invited talk at 2023 China Mobile Cloud Conference (移动云大会): "Progress and Future Challenges of Large-scale Human-Machine Dialog", Suzhou, China.

19. Dec 20, 2022: Invited Talk "Data-efficient Multilingual and Crosslingual Speech Recognition" at CCF Advanced Disciplines Lectures (ADL) (计算机学会学科前沿讲习班), Beijing, China.

20. April 29, 2022: "Introduction to EMNLP 2022 Workshop - Towards Semi-Supervised and Reinforced Task-Oriented Dialog Systems" and its Challenge", SereTOD Challenge Kickoff Seminar, virtual.

21. Dec 18, 2021: Invited Talk "Data-efficient Multilingual and Crosslingual Speech Recognition" at CCF Seminar on Southeast-Asia Language Technology and Application, Nanning, China.

22. Sep 24, 2021: Invited Talk "Semi-Supervised Task-Oriented Dialog Systems and Natural Language Labeling" at UIUC NLP Seminar.

23. Mar 29, 2021: Invited talk "On exploring the third generation of speech recognition technology" at National Conference on Acoustics (全国声学大会), Shanghai.

24. Jan 30, 2021: Invited Talk "Neural Architecture Search for End-to-end Speech Recognition via Straight-Through Gradients" at SLT2021 Children Speech Recognition Challenge (CSRC) Workshop.

25. Nov. 2019: "Marrying Neural Networks and Undirected Graphical Models for Efficient Speech Recognition", at ASRU 2019 Code-Switching Challenge Seminar, Beijing.

26. July 2019: a short course "Theory and Applications of Probabilistic Graphical Models" (1 credit) at Xinjiang University, China.

27. April 2019: "Toward Simple Speaker Recognition and Speech Recognition", at The Symposium on Speaker Recognition Research and Application, Kun-shan, China.

28. Dec. 2018: "THU-SPMI System For NIST 2018 Speaker Recognition Evaluation" at NIST SRE-2018 Workshop, Athens, Greece, 2018.
29. Oct. 2018: "Neural random fields with applications to modeling of languages and images" at Apple Boston, Toyota Technological Institute at Chicago (TTIC), and UIUC.
30. Oct. 2017: "Future Artificial Intelligence with Efficient Learning" at Tsinghua University Science and Technology Forum, SuZhou, China.
31. Dec. 2016: "The THU-SPMI SRE-16 System with Joint Bayesian Scoring and Ladder Network based Feature Learning" at 2016 NIST Speaker Recognition Workshop (SRE 2016), San Diego, USA.
32. Jan. 2016: "Machine Intelligence with Speech and Language" at Tsinghua Data Science Institute.
33. Oct. 2015: "Trans-dimensional Random Fields for Sequence Modeling" at SANE workshop, held at Google in New York City, NY.
34. June 2015: "Probabilistic Modeling of Speech and Language" at IBM Watson Research Center, New York.
35. June 2015: "Probabilistic Modeling of Speech and Language" at Mitsubishi Electric Research Laboratories (MERL), Boston.
36. May 2015: "Probabilistic Modeling of Speech and Language" at ICSI, Berkeley.
37. March 2015: "Probabilistic Modeling of Speech" at the DSP Seminar, UIUC.

TEACHING

Undergraduate courses in Tsinghua University
- Probability Theory
  - Spring, 2012 - 2014, 2016 - present
- Stochastic Process
  - Fall, 2012 - 2013, 2015 - present
- Mathematics and engineering applications
  - Spring, 2005 - 2011.

Graduate courses in Tsinghua University
- Theory and Applications of Probabilistic Graphical Models
  - Fall, 2004 – 2012;
  - Spring, 2017 - 2019, 2024

Short courses I hosted in Tsinghua University
- Machine Learning for Speech Processing, 12/17/2012 - 12/21/2012,
  By Dr. Shinji Watanabe, Mitsubishi Electric Research Laboratories (MERL), USA.
- Advanced Monte Carlo methods, 8/20/2018 - 8/31/2018,
  By Prof. Zhiqiang Tan, Rutgers University, USA.