<h1 style="text-align:center">Publications</h1>

Thesis

1. Kai Yu. *Adaptive Training for Large Vocabulary Continuous Speech Recognition.* PhD thesis, Cambridge University, Cambridge, U.K., 2006.

2. Kai Yu. *Study of Several Topics in Digital Image Processing and Analysis & Their Applications in Optical Coherence Tomography (OCT) System.* Master thesis, Tsinghua University, Beijing, China, 2002.

Book Chapter

1. 俞凯等. 听觉信息处理研究前沿. 上海交通大学出版社 2021

2. 俞栋、邓力、俞凯等. 人工智能：语音识别理解与实践. 电子工业出版社 2020

3. 俞凯等译. 强化学习（第二版）. 电子工业出版社 2019

4. 俞凯、吴科译. 人工智能中的深度结构学习. 机械工业出版社 2017

5. 俞凯等译. 解析深度学习：语音识别实践. 电子工业出版社 2016

6. Kai Yu. Statistical Models for Dealing with Discontinuity of Fundamental Frequency. in *Keikichi Hirose and Jianhua Tao (eds.), Speech Prosody in Speech Synthesis: Modeling and generation of prosody for high quality and flexible speech synthesis, Prosody, Phonology and Phonetics.* Springer-Verlag, Berlin Heidelberg 2015: 123-144.

7. Kai Yu.Automatic Speech Recognition:A Deep Learning Approach in*Dong Yu and Li Deng (eds.)*Springer, Berlin Heidelberg 2014

8. Peilu Wang, Ruihua Sun, Hai Zhao and Kai Yu. A New Word Language Model Evaluation Metric For Character Based Languages. in *Chapter: Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data. Volume 8202 of the series Lecture Notes in Computer Science.* Springer-Verlag, Berlin Heidelberg 2013: 315-324.

9. Kai Yu, Mark Gales, Lan Wang and Philip Woodland. Chapter 3.3.3: Unsupervised Training and Directed Manual Transcriptions for Mandarin Speech-to-Text. In *Handbook of Natural Language Processing and Machine Translation: DARPA Global Autonomous Language Exploitation.* Springer, Berlin Heidelberg 2011: 652-663.

10. Marcus Tomalin, William Byrne, Adria Gispert, Mark Gales, Xunying Liu, Phillip Woodland, and Kai Yu. Chapter 3.5.5: Improving Speech Transcription for Mandarin-English Translation. In *Handbook of Natural Language Processing and Machine Translation: DARPA Global Autonomous Language Exploitation.* Springer, Berlin Heidelberg 2011.

Journal

1. Xuenan Xu, Zeyu Xie, Mengyue Wu and Kai Yu. Beyond the Status Quo: A Contemporary Survey of Advances and Challenges in Audio Captioning. *IEEE/ACM Transactions on Audio, Speech, and Language Processing,*vol. 32, pp. 95-112, 2024

2. Ruisheng Cao, Lu Chen, Jieyu Li, Hanchong Zhang, Hongshen Xu, Wangyou Zhang and Kai Yu. A Heterogeneous Graph to Abstract Syntax Tree Framework for Text-to-SQL. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI),*vol. 45, no. 11, pp. 13796-13813, 2023

3. Wenbin Jiang and Kai Yu. Speech Enhancement With Integration of Neural Homomorphic Synthesis and Spectral Masking. *IEEE/ACM Transactions on Audio, Speech, and Language Processing,*vol. 31, pp. 1758-1770, 2023

4. Chenpeng Du, Yiwei Guo, Xie Chen and Kai Yu. Speaker Adaptive Text-to-Speech With Timbre-Normalized Vector-Quantized Feature. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 31, pp. 3446-3456, 2023

5. Jieyu Li, Zhi Chen, Lu Chen, Zichen Zhu, Hanqi Li, Ruisheng Cao, and Kai Yu. DIR: A Large-Scale Dialogue Rewrite Dataset for Cross-Domain Conversational Text-to-SQL. *Applied Sciences*, vol. 13(4), pp. 2262, 2023

6. Bo Chen, Chenpeng Du and Kai Yu. Neural Fusion for Voice Cloning. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 30, pp. 1993-2001, 2022

7. 徐志航, 陈博, 张辉, 俞凯. 小数据下的音素级别说话人嵌入的语音合成自适应方法. 计算机学报,Vol.45, No.2, pp. 1003-1017, 2022

8. 刘奇, 马饶, 俞凯. 马尔可夫决策过程和先验控制向量在弱约束自然语言生成中的应用. 计算机学报,Vol.45, No.2, pp. 289-301,2022

9. Zhi Chen, Yuncong Liu, Lu Chen, Su Zhu, Mengyue Wu and Kai Yu. OPAL: Ontology-Aware Pretrained Language Model for End-to-End Task-Oriented Dialogue. *Transactions of the Association for Computational Linguistics*, vol.11, pp. 68-84, 2022

10. Heinrich Dinkel, Shuai Wang, Xuenan Xu, Mengyue Wu and Kai Yu. Voice activity detection in the wild: A data-driven approach using teacher-student training. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 1542-1555, 2021

11. Heinrich Dinkel, Mengyue Wu and Kai Yu. Towards Duration Robust Weakly Supervised Sound Event Detection. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 887-900, 2021

12. Bo Chen, Zhihang Xu and Kai Yu. Data augmentation based non-parallel voice conversion with frame-level speaker disentangler. *Speech Communication*, vol. 136, pp. 14-22, 2021

13. Chenpeng Du and Kai Yu. Phone-Level Prosody Modelling with GMM-Based MDN for Diverse and Controllable Speech Synthesis. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 30, pp. 190-201, 2021

14. Chen Zhang, Daihui Peng,Lu Lv, Kaiming Zhuo, Kai Yu, Tian Shen, Yifeng Xu and Zhen Wang. Individual Perceived Stress Mediates Psychological Distress in Medical Workers During COVID-19 Epidemic Outbreak in Wuhan *Neuropsychiatric Disease and Treatment*,vol.16,pp.2529-2537,2020

15. Kai Yu, Rao Ma, Kaiyu Shi and Qi Liu. Neural Network Language Model Compression With Product Quantization and Soft Binarization *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 2438-2449, 2020

16. Zhi Chen, Lu Chen, Xiaoyuan Liu and Kai Yu. Distributed Structured Actor-Critic Reinforcement Learning for Universal Dialogue Management *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp.2400-2411, 2020

17. Shuai Wang, Yexin Yang, Zhanghao Wu, Yanmin Qian and Kai Yu. Data Augmentation using Deep Generative Models for Embedding based Speaker Recognition *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp.2598-2609, 2020

18. Qi Liu, Zhehuai Chen, Hao Li, Mingkun Huang, Yizhou Lu and Kai Yu. Modular End-to-end Automatic Speech Recognition Framework for Acoustic-to-word Model *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 2174-2183, 2020

19. Su Zhu, Ruisheng Cao and Kai Yu. Dual Learning for Semi-Supervised Natural Language Understanding *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1936-1947, 2020

20. Su Zhu, Zijian Zhao, Rao Ma and Kai Yu. Prior Knowledge Driven Label Embedding for Slot Filling in Natural Language Understanding *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1440-1451, 2020

21. Zhehuai Chen, Wenlu Zhen, Yongbin You, Yanmin Qian, Kai Yu. Label Synchronous Decoding for Speech Recognition(In Chinese) *Chinese Journal of Comouter*, Vol.42.No.7.pp.1511-1523.July 2019

22. Shuai Wang, Zili Huang, Yanmin Qian, Kai Yu. Discriminative Neural Embedding Learning for Short-Duration Text-Independent Speaker Verification *IEEE/ACM Transactions on Audio, Speech, and Language Processing*,vol. 27, no. 11, pp. 1686-1696, Nov. 2019

23. Lu Chen, Zhi Chen, Bowen Tan, Sishan Long, Milica Gasic and Kai Yu. Agent-Graph: Toward Universal Dialogue Management With Structured Deep Reinforcement Learning *IEEE/ACM Transactions on Audio, Speech, and Language Processing*,vol. 27, no. 9, pp. 1378-1391, Sep. 2019

24. Zhehuai Chen, Yanmin Qian and Kai Yu. Sequence Discriminative Training for Deep Learning based Acoustic Keyword Spotting *Speech Communication*,vol. 102, 100-111, 2018

25. Heinrich Dinkel, Yanmin Qian and Kai Yu. Investigating Raw Wave Deep Neural Networks for End-to-End Speaker Spoofing Detection *IEEE/ACM Transactions on Audio, Speech, and Language Processing*,vol. 26, no. 11, pp. 2002-2014, Nov. 2018

26. Tian Tan, Yanmin Qian, Hu Hu, Wen Ding, Ying Zhou and Kai Yu, Adaptive very deep convolutional residual network for noise robust speech recognition *IEEE/ACM Transactions on Audio, Speech, and Language Processing*,vol. 26, no. 8, pp. 1393-1405, Aug. 2018

27. Kai Yu, Zijian Zhao, Xueyang Wu, Hongtao Lin and Xuan Liu, Rich Short Text Conversation Using Semantic Key Controlled Sequence Generation *IEEE/ACM Transactions on Audio, Speech, and Language Processing*,vol. 26, no. 8, pp. 1359-1368, Mar. 2018

28. Zhehuai Chen, Yimon Zhuang, Yanmin Qian and Kai Yu. Phone Synchronous Speech Recognition with CTC Lattices. *IEEE/ACM Transactions on Audio*,vol. 25, no. 1, 86-97, 2017.

29. Yanmin Qian, Nanxin Chen and Kai Yu. Deep Features for Automatic Spoofing Detection. *Speech Communication*, vol. 85, 43-52, 2016.

30. Yanmin Qian, Mengxiao Bi, Tian Tan and Kai Yu. Very Deep Convolutional Neural Networks for Noise Robust Speech Recognition. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 24, no. 12, 2263-2276, 2016.

31. Kai Sun, Qizhe Xie and Kai Yu. Recurrent Polynomial Network for Dialogue State Tracking. *Dialogue & Discourse*, vol. 7, no. 37, 65-68, 2016.

32. Kai Yu, Lu Chen, Kai Sun, Qizhe Xie and Su Zhu. Evolvable dialogue state tracking for statistical dialogue management. *Frontiers of Computer Science*, vol. 10, no. 2, 201-215, 2016.

33. Tian Tan, Yanmin Qian and Kai Yu. Cluster Adaptive Training for Deep Neural Network Based Acoustic Model. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 24, no. 3, 459-468, 2016.

34. Kai Yu, Kai Sun, Lu Chen and Su Zhu. Constrained Markov Bayesian Polynomial for Efficient Dialogue State Tracking. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 23, no. 12, 2177-2188, 2015.

35. Kai Yu, Lu Chen, Bo Chen, Kai Sun and Su Zhu. Cognitive Technology in Task-oriented Dialogue Systems –Concepts, Advances and Future. *Chinese Journal of Computers*, vol. 38, no. 12, 2333-2348, 2015. (In Chinese)

36. Yuan Liu, Yanmin Qian, Nanxin Chen, Tianfan Fu and Kai Yu. Deep Feature for Text-dependent Speaker Verification. *Speech Communication*, vol. 73, 1-13, 2015.

37. Kai Yu, Heiga Zen, Francois Mairesse and Steve Young. Context adaptive training with factorized decision trees for HMM-based statistical parametric speech synthesis. *Speech Communication*, vol.53, no.6, 914-923, 2011.

38. Kai Yu and Steve Young. Continuous F0 modelling for HMM based statistical parametric speech synthesis. *IEEE Transactions on Audio, Speech and Language Processing*, vol.19, no.5, 1071-1079, 2011.

39. Kai Yu, Mark Gales, Lan Wang and Philip Woodland. Unsupervised training and directed manual transcription for LVCSR. *Speech Communication*, vol. 52, 652-663, 2010.

40. Steve Young, Milica Gasic, Simon Keizer, Francois Mairesse, Jost Schatzmann, Blaise Thomson and Kai Yu. The hidden information state model: a practical framework for POMDP-based spoken language management. *Computer Speech and Language*, vol. 24, no. 2, 150-174, 2009.

41. Kai Yu, Mark Gales, and Philip Woodland. Unsupervised adaptation with discriminative mapping transforms. *IEEE Transactions on Audio, Speech and Language Processing*, vol. 17, no. 4, 714-723, 2009.

42. Kai Yu and Mark Gales. Bayesian adaptive inference and adaptive training. *IEEE Transactions on Audio, Speech and Language Processing*, vol. 15, no. 6, 1932-1943, 2007.

43. Kai Yu and Mark Gales. Discriminative cluster adaptive training. *IEEE Transactions on Audio, Speech and Language Processing*, vol. 14, no. 5, 1694-1703, 2006.

44. Kai Yu and Liang Ji. Karyotyping of CGH human metaphases using kernel nearest-neighbor algorithm. *Cytometry*, vol. 48, no. 4, 202-208, 2002.

45. Kai Yu and Liang Ji. Kernel nearest-neighbor algorithm. *Neural Processing Letters*, vol. 15, no. 2, 147-156, 2002.

46. Kai Yu, Liang Ji, Lei Wang and Ping Xue. How to optimize OCT image. *Optics Express*, vol. 9, no. 1, 24-35, 2001.

Conference

1. Hanchong Zhang, Ruisheng Cao, Lu Chen, Hongshen Xu, and Kai Yu. ACT-SQL: In-Context Learning for Text-to-SQL with Automatically-Generated Chain-of-Thought. In *Findings of the Association for Computational Linguistics: EMNLP*, Singapore, 2023, 3501-3532

2. Danyang Zhang, Lu Chen, Situo Zhang, Hongshen Xu, Zihan Zhao and Kai Yu. Large Language Models Are Semi-Parametric Reinforcement Learning Agents. In *NeurIPS*, New Orleans, USA, 2023

3. Chenpeng Du, Qi Chen, Tianyu He, Xu Tan, Xie Chen, Kai Yu, Sheng Zhao, Jiang Bian. DAE-Talker: High Fidelity Speech-Driven Talking Face Generation with Diffusion Autoencoder.In *31st ACM International Conference on Multimedia (ACM-MM)*, Ottawa, Canada, 2023, 4281-4289

4. Guangpeng Li, Lu Chen and Kai Yu. "How ChatGPT is Robust for Spoken Language Understanding?" In *24th Annual Conference of the International Speech Communication Association (InterSpeech)*, Dublin, Ireland, 2023, 2163-2167

5. Pingyue Zhang, Mengyue Wu and Kai Yu. ReCLR: Reference-Enhanced Contrastive Learning of Audio Representation for Depression Detection. In *24th Annual Conference of the International Speech Communication Association (InterSpeech)*,Dublin, Ireland, 2023, 2998-3002

6. Sen Liu, Yiwei Guo, Chenpeng Du, Xie Chen, Kai Yu. DSE-TTS: Dual Speaker Embedding for Cross-Lingual Text-to-Speech. In *24th Annual Conference of the International Speech Communication Association (InterSpeech)*, Dublin, Ireland, 2023, 616-620

7. Zheng Liang, Zheshu Song, Ziyang Ma, Chenpeng Du, Kai Yu, Xie Chen. Improving Code-Switching and Named Entity Recognition in ASR with Speech Editing based Data Augmentation. In *24th Annual Conference of the International Speech Communication Association (InterSpeech)*, Dublin, Ireland, 2023, 919-923

8. Zeyu Xie, Xuenan Xu, Mengyue Wu and Kai Yu. Enhance Temporal Relations in Audio Captioning with Sound Event Detection. In *24th Annual Conference of the International Speech Communication Association (InterSpeech)*, Dublin, Ireland, 2023, 4179-4183

9. Xuenan Xu, Mengyue Wu, and Kai Yu. Investigating Pooling Strategies and Loss Functions for Weakly-Supervised Text-to-Audio Grounding via Contrastive Learning. In *IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW)*,Rhodes Island, Greece, 2023

10. Hanchong Zhang, Jieyu Li, Lu Chen, Ruisheng Cao, Yunyan Zhang, Yu Huang, Yefeng Zheng, and Kai Yu. CSS: A Large-scale Cross-schema Chinese Text-to-SQL Medical Dataset. In *Findings of the Association for Computational Linguistics: ACL*,Toronto, Canada,2023,6970–6983

11. Jieyu Li, Lu Chen, Ruisheng Cao, Su Zhu, Hongshen Xu, Zhi Chen, Hanchong Zhang, and Kai Yu. Exploring Schema Generalizability of Text-to-SQL. In *Findings of the Association for Computational Linguistics: ACL*,Toronto, Canada,2023,1344–1360

12. Yiwei Guo, Chenpeng Du, Xie Chen and Kai Yu. EmoDiff: Intensity-Controllable Emotional Text-to-Speech with Soft-Label Guidance. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*,Rhodes Island, Greece, 2023

13. Zhijun Liu, Yiwei Guo, Kai Yu. DiffVoice: Text-To-Speech With Latent Diffusion. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, 2023

14. Chenpeng Du, Yiwei Guo, Feiyu Shen, Kai Yu. Multi-Speaker Multi-Lingual VQTTS System for LIMMITS 2023 Challenge. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*,Rhodes Island, Greece, 2023

15. Tao Liu, Zhengyang Chen, Yanmin Qian, Kai Yu. Multi-Speaker End-to-end Multimodal Speaker Diarization System for the MISP 2022 CHALLENGE. In*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*,Rhodes Island, Greece, 2023

16. Guangwei Li, Xuenan Xu, Lingfeng Dai, Mengyue Wu, Kai Yu. Diverse and Vivid Sound Generation from Text Descriptions. In*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*,Rhodes Island, Greece, 2023

17. Qi Chen, Ziyang Ma, Tao Liu, Xu Tan, Qu Lu, Kai Yu, Xie Chen. Improving Few-Shot Learning for Talking Face System with TTS Data Augmentation. In*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*,Rhodes Island, Greece, 2023

18. Tao Liu, Xu Xiang, Zhengyang Chen, Bing Han, Kai Yu, Yanmin Qian. The X-Lance Speaker Diarization System for the Conversational Short-phrase Speaker Diarization Challenge 2022. In*13th International Symposium on Chinese Spoken Language Processing (ISCSLP)*,Singapore, 2022, 498-501

19. Wenbin Jiang, Tao Liu, Kai Yu. Efficient Speech Enhancement with Neural Homomorphic Synthesis. In*23rd Annual Conference of the International Speech Communication Association (InterSpeech)*, Incheon, South Korea, 2022, 986-990

20. Chenpeng Du, Yiwei Guo, Xie Chen and Kai Yu. VQTTS: High-Fidelity Text-to-Speech Synthesis with Self-Supervised VQ Acoustic Feature. In*23rd Annual Conference of the International Speech Communication Association (InterSpeech)*, Incheon, South Korea, 2022, 1596-1600

21. Zihan Zhao, Lu Chen, Ruisheng Cao, Hongshen Xu, Xingyu Chen and Kai Yu. TIE: Topological Information Enhanced Structural Reading Comprehension on Web Pages. In*2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*, Seattle, United States, 2022, 1808-1821

22. Yu Xi, Tian Tan, Wangyou Zhang, Baochen Yang and Kai Yu. Text Adaptive Detection for Customizable Keyword Spotting. In*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, Singapore, 2022, 6652-6656

23. Guangwei Li, Xuenan Xu, Mengyue Wu and Kai Yu. Category-Adapted Sound Event Enhancement with Weakly Labeled Data. In*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, Singapore, 2022, 851-855

24. Guangwei Li, Xuenan Xu, Mengyue Wu and Kai Yu. Navigating Audio-Visual Event Detection Across Mismatched Modalities. In*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, Singapore, 2022, 1975-1979

25. Siyu Lou, Xuenan Xu, Mengyue Wu and Kai Yu. Audio-Text Retrieval in Context. In*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, Singapore, 2022, 4793-4797

26. Xuenan Xu, Mengyue Wu and Kai Yu. Diversity-controllable and Accurate Audio Captioning Based on Neural Condition. In*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, Singapore, 2022, 971-975

27. Yiwei Guo, Chenpeng Du and Kai Yu. Unsupervised word-level prosody tagging for controllable speech synthesis. In*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, Singapore, 2022, 7597-7601

28. Lingfeng Dai, Lu Chen, Zhikai Zhou and Kai Yu. LatticeBART: Lattice-to-Lattice Pre-Training for Speech Recognition. In*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, Singapore, 2022, 6112-6116

29. Wenbin Jiang, Zhijun Liu, Kai Yu and Fei Wen. Speech enhancement with neural homomorphic synthesis. In*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, Singapore, 2022, 376-380

30. Zhi Chen, Lu Chen, Bei Chen, Libo Qin, Yuncong Liu, Su Zhu, Jian-Guang Lou, and Kai Yu. UniDU: Towards A Unified Generative Dialogue Understanding Framework. In*23rd Annual Meeting of the Special Interest Group on Discourse and Dialogue*, Edinburgh, UK, 2022, 442-455

31. Zhi Chen, Bei Chen, Lu Chen, Kai Yu and Jian-Guang Lou. AdapterShare: Task Correlation Modeling with Adapter Differentiation. In *2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Abu Dhabi, United arab emirates, 2022, 10645-10651

32. Liangtai Sun, Xingyu Chen, Lu Chen, Tianle Dai, Zichen Zhu, and Kai Yu. META-GUI: Towards Multi-modal Conversational Agents on Mobile GUI. In *2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Abu Dhabi, United arab emirates, 2022, 6699-6712

33. Binwei Yao, Chao Shi, Likai Zou, Lingfeng Dai, Mengyue Wu, Lu Chen, Zhen Wang, and Kai Yu. 2022. D4: a Chinese Dialogue Dataset for Depression-Diagnosis-Oriented Chat. In *2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Abu Dhabi, United arab emirates, 2022, 2438-2459

34. Zhi Chen, Lu Chen, Yanbin Zhao, Su Zhu, Kai Yu. Dual Learning for Dialogue State Tracking. In *National Conference on Man-Machine Speech Communication (NCMMSC)*, Hefei, China, 2022, 293-305

35. Xingyu Chen, Zihan Zhao, Lu Chen, JiaBao Ji, Danyang Zhang, Ao Luo, Yuxuan Xiong and Kai Yu. WebSRC: A Dataset for Web-Based Structural Reading Comprehension. In *2021 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Online and Punta Cana, Dominican Republic, 2021, 4173–4185

36. Pingyue Zhang, Mengyue Wu, Heinrich Dinkel and Kai Yu. DEPA: Self-Supervised Audio Embedding for Depression Detection. In *29th ACM International Conference on Multimedia (ACM-MM)*, Virtual Event, China, 2021, 135-143

37. Yao Zhao, Lu Chen and Kai Yu. Relation-aware Multi-hop Reasoning for Visual Dialog. In *10th CCF International Conference on Natural Language Processing and Chinese Computing (NLPCC)*, Qingdao, China, 2021

38. Boer Lyu, Lu Chen and Kai Yu. Glyph Enhanced Chinese Character Pre-Training for Lexical Sememe Prediction. In *Findings of the Association for Computational Linguistics: EMNLP*, Punta Cana, Dominican Republic, 2021, 4549–4555

39. Xuenan Xu, Heinrich Dinkel, Mengyue Wu and Kai Yu. A Lightweight Framework for Online Voice Activity Detection in the Wild. In *22nd Annual Conference of the International Speech Communication Association (InterSpeech)*, Brno, Czechia, 2021, 371-375

40. Lingfeng Dai, Qi Liu and Kai Yu. Class-Based Neural Network Language Model for Second-Pass Rescoring in ASR. In *22nd Annual Conference of the International Speech Communication Association (InterSpeech)*, Brno, Czechia, 2021, 2022-2026

41. Chenpeng Du and Kai Yu. Rich Prosody Diversity Modelling with Phone-level Mixture Density Network. In *22nd Annual Conference of the International Speech Communication Association (InterSpeech)*, Brno, Czechia, 2021, 3136-3140

42. Zhi Chen, Lu Chen, Hanqi Li, Ruisheng Cao, Da Ma, Mengyue Wu and Kai Yu. Decoupled Dialogue Modeling and Semantic Parsing for Multi-Turn Text-to-SQL. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP* 2021, 3063–3074

43. Ruisheng Cao, Lu Chen, Zhi Chen, Yanbin Zhao, Su Zhu and Kai Yu. LGESQL: Line Graph Enhanced Text-to-SQL Model with Mixed Local and Non-Local Relations. In *59th Annual Meeting of the Association for Computational Linguistics (ACL) and the 11th International Joint Conference on Natural Language Processing*, 2021, 2541–2555

44. Xuenan Xu, Heinrich Dinkel, Mengyue Wu, Zeyu Xie and Kai Yu. Investigating Local and Global Information for Automated Audio Captioning with Transfer Learning. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, Ontario, Canada, 2021, 905-909

45. Xuenan Xu, Heinrich Dinkel, Mengyue Wu and Kai Yu. Text-to-Audio Grounding: Building Correspondence Between Captions and Sound Events. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, Ontario, Canada, 2021, 606-610

46. Chenpeng Du, Bing Han, Shuai Wang, Yanmin Qian and Kai Yu. SYNAUG: Synthesis-Based Data Augmentation for Text-Dependent Speaker Verification. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, Ontario, Canada, 2021, 5829-5833

47. Zhi Chen, Lu Chen, Yanbin Zhao, Ruisheng Cao, Zihan Xu, Su Zhu and Kai Yu. ShadowGNN: Graph Projection Neural Network for Text-to-SQL Parser. In *2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*, online, 2021, 5567–5577

48. Boer Lyu, Lu Chen, Su Zhu and Kai Yu. LET: Linguistic Knowledge Enhanced Graph Transformer for Chinese Short Text Matching. In *the Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)*, online, 2021, 2934-2942

49. Xuenan Xu, Heinrich Dinkel, Mengyue Wu and Kai Yu. Audio Caption in a Car Setting with a Sentence-Level Loss. In *12th International Symposium on Chinese Spoken Language Processing (ISCSLP)*, Hong Kong, China, 2021, 1-5

50. Shuai Wang, Yexin Yang, Yanmin Qian and Kai Yu. Revisiting the Statistics Pooling Layer in Deep Speaker Embedding Learning. In *12th International Symposium on Chinese Spoken Language Processing (ISCSLP)*, Hong Kong, China, 2021, 1-5

51. Yefei Chen, Heinrich Dinkel, Mengyue Wu and Kai Yu. Voice activity detection in the wild via weakly supervised sound event detection. In *21st Annual Conference of the International Speech Communication Association (InterSpeech)*, Shanghai, China, 2020, 3665-3669

52. Chen Liu, Su Zhu, Zijian Zhao, Ruisheng Cao, Lu Chen and Kai Yu. Jointly Encoding Word Confusion Network and Dialogue Context with BERT for Spoken Language Understanding. In *21st Annual Conference of the International Speech Communication Association (InterSpeech)*, Shanghai, China, 2020, 871-875

53. Hongji Wang, Heinrich Dinkel, Shuai Wang, Yanmin Qian and Kai Yu. Dual-Adversarial Domain Adaptation for Generalized Replay Attack Detection. In *21st Annual Conference of the International Speech Communication Association (InterSpeech)*, Shanghai, China, 2020, 1086-1090

54. Zihan Zhao, Yuncong Liu, Lu Chen, Qi Liu, Rao Ma, and Kai Yu. An Investigation on Different Underlying Quantization Schemes for Pre-trained Language Models. In *9th CCF International Conference on Natural Language Processing and Chinese Computing (NLPCC)*, Zhengzhou, China, 2020, 359-371

55. Zihan Xu, Zhi Chen, Lu Chen, Su Zhu and Kai Yu. Memory Attention Neural Network For Multi-Domain Dialogue State Tracking. In *9th CCF International Conference on Natural Language Processing and Chinese Computing (NLPCC)*, Zhengzhou, China, 2020, 41–52

56. Chen Liu, Su Zhu, Lu Chen and Kai Yu. Robust Spoken Language Understanding with RL-Based Value Error Recovery. In *9th CCF International Conference on Natural Language Processing and Chinese Computing (NLPCC)*, Zhengzhou, China, 2020, 78-90

57. Xuenan Xu, Heinrich Dinkel, Mengyue Wu and Kai Yu. A CRNN-GRU Based Reinforcement Learning Approach to Audio Captioning. In *5th Workshop on Detection and Classification of Acoustic Scenes and Events (DCASE)*, Tokyo, Japan, 2020, 225-229

58. Lu Chen, Yanbin Zhao, Boer Lv, Lesheng Jin, Zhi Chen, Su Zhu and Kai Yu. Neural Graph Matching Networks for Chinese Short Text Matching. In *58th Annual Meeting of the Association for Computational Linguistics (ACL)*, Seattle, USA, 2020, 6152-6158

59. Yanbin Zhao, Lu Chen, Zhi Chen, Ruisheng Cao, Su Zhu and Kai Yu. Line Graph Enhanced AMR-to-Text Generation with Mix-Order Graph Attention Networks. In *58th Annual Meeting of the Association for Computational Linguistics (ACL)*, Seattle, USA, 2020, 732-741

60. Ruisheng Cao, Su Zhu, Chenyu Yang, Chen Liu, Rao Ma, Yanbin Zhao, Lu Chen and Kai Yu. Unsupervised Dual Paraphrasing for Two-stage Semantic Parsing. In *58th Annual Meeting of the Association for Computational Linguistics (ACL)*, Seattle, USA, 2020, 6806-6817

61. Su Zhu, Jieyu Li, Lu Chen and Kai Yu. Efficient Context and Schema Fusion Networks for Multi-Domain Dialogue State Tracking. In *Findings of the Association for Computational Linguistics: EMNLP*, Online, 2020, 766-781

62. Jieyu Li, Su Zhu and Kai Yu. A Hierarchical Tracker for Multi-Domain Dialogue State Tracking. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, 2020, 8014-8018

63. Heinrich Dinkel and Kai Yu. Duration Robust Weakly Supervised Sound Event Detection. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, 2020, pp. 311-315

64. Yexin Yang, Shuai Wang, Xun Gong, Yanmin Qian and Kai Yu. Text Adaptation for Speaker Verification with Speaker-Text Factorized Embeddings. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, 2020, 6454-6458

65. Zhengyang Chen, Shuai Wang, Yanmin Qian and Kai Yu. Channel Invariant Speaker Embedding Learning with Joint Multi-Task and Adversarial Training. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, 2020, 6574-6578

66. Chenpeng Du, and Kai Yu. Speaker Augmentation for Low Resource Speech Recognition. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, 2020, 7719-7723

67. Shuai Wang, Johan Rohdin, Lukáš Burget, Oldřich Plchot, Kai Yu and Jan Černocký. Investigation of Specaugment for Deep Speaker Embedding Learning. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, 2020, 7139-7143

68. Rao Ma, Hao Li, Qi Liu, Lu Chen and Kai Yu. Neural Lattice Search for Speech Recognition. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, 2020, 7794-7798

69. Rao Ma, Lesheng Jin, Qi Liu, Lu Chen and Kai Yu. Addressing the Polysemy Problem in Language Modeling with Attentional Multi-Sense Embeddings. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, 2020, 8129-8133

70. Lu Chen, Boer Lv, Chi Wang, Su Zhu, Bowen Tan and Kai Yu. Schema-Guided Multi-Domain Dialogue State Tracking with Graph Attention Neural Networks. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20)*, New York, USA, 2020, 7521-7528

71. Yanbin Zhao, Lu Chen, Zhi Chen and Kai Yu. Semi-Supervised Text Simplification with Back-Translation and Asymmetric Denoising Autoencoders. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20)*, New York, USA, 2020, 9668-9675

72. Xu Xiang, Shuai Wang, Houjun Huang, Yanmin Qian and Kai Yu. Margin Matters: Towards More Discriminative Deep Neural Network Embeddings for Speaker Recognition In *APSIPA Annual Summit and Conference*, Lanzhou, China, 2019, 1652-1656

73. Mingkun Huang, Yizhou Lu, Lan Wang, Yanmin Qian, Kai Yu. Exploring Model Units and Training Strategies for End-to-End Speech Recognition. In *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, Sentosa, Singapore, 2019, 524-531

74. Rao Ma, Qi Liu, Kai Yu. Highly Efficient Neural Network Language Model Compression Using Soft Binarization Training. In *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, Sentosa, Singapore, 2019, 62-69

75. Zijian Zhao, Su Zhu and Kai Yu. Data Augmentation with Atomic Templates for Spoken Language Understanding. In *2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP 2019)*, Hong Kong, China, 2019, 3628–3634

76. Yefei Chen, Shuai Wang, Yanmin Qian and Kai Yu. End-to-End Speaker-Dependent Voice Activity Detection. In *The 15th National Conference on Man-Machine Speech Communication (NCMMSC2019)*, Xining, Qinghai, China, 2019

77. Hao Li, Zhehuai Chen, Qi Liu, Yanmin Qian and Kai Yu. OOV Words Extension for Modular Neural Acoustics-to-Word Model. In *The 15th National Conference on Man-Machine Speech Communication (NCMMSC2019)*, Xining, Qinghai, China, 2019

78. Hao Li, Chen Liu, Su Zhu and Kai Yu. Robust Spoken Language Understanding with Acoustic and Domain Knowledge. In *2019 International Conference on Multimodal Interaction (ICMI 2019)*, Suzhou, China, 2019, 531-535

79. Su Zhu, Zijian Zhao, Tiejun Zhao, Chengqing Zong and Kai Yu. CATSLU: The 1st Chinese Audio-Textual Spoken Language Understanding Challenge. In *2019 International Conference on Multimodal Interaction (ICMI 2019)*, Suzhou, China, 2019

80. Jiaqi Guo, Yongbin You, Yanmin Qian and Kai Yu. Joint Decoding of CTC Based Systems for Speech Recognition. In *20th Annual Conference of the International Speech Communication Association (InterSpeech)*, Graz, Austria, 2019, 2205-2209

81. Zhanghao Wu, Shuai Wang, Yanmin Qian and Kai Yu. Data Augmentation using Variational Autoencoder for Embedding based Speaker Verification.

82. Yexin Yang, Hongji Wang, Heinrich Dinkel, Zhengyang Chen, Shuai Wang, Yanmin Qian and Kai Yu. The SJTU Robust Anti-spoofing System for the ASVspoof 2019 Challenge In *20th Annual Conference of the International Speech Communication Association (InterSpeech)*,Graz, Austria, 2019, 1038-1042

83. Hongji Wang, Heinrich Dinkel, Shuai Wang, Yanmin Qian and Kai Yu. Cross-domain Replay Spoofing Attack Detection using Domain Adversarial Training In *20th Annual Conference of the International Speech Communication Association (InterSpeech)*,Graz, Austria, 2019, 2938-2942 In *20th Annual Conference of the International Speech Communication Association (InterSpeech)*, Graz, Austria, 2019, 1163-1167

84. Shuai Wang, Johan Rohdin, Lukáš Burget, Oldřich Plchot, Yanmin Qian, Kai Yu and Jan Černocký. On the Usage of Phonetic Information for Text-independent Speaker Embedding Extraction. In *20th Annual Conference of the International Speech Communication Association (InterSpeech)*, Graz, Austria, 2019, 1148-1152.

85. Ruisheng Cao, Su Zhu, Chen Liu, Jieyu Li and Kai Yu. Semantic Parsing with Dual Learning In *57th Annual Meeting of the Association for Computational Linguistics (ACL)*,Florence, Italy, 2019, 51-64

86. Xuankai Chang, Yanmin Qian, Kai Yu, Shinji Watanabe. End-to-end Monaural Multi-speaker ASR System without Pretraining. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*,Brighton, UK, 2019, 6256-6260

87. Shuai Wang, Yexin Yang, Tianzhe Wang, Yanmin Qian and Kai Yu. Knowledge Distillation for Small Foot-print Deep Speaker Embedding In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*,Brighton, UK, 2019,6021-6025

88. Zijian Zhao, Su Zhu and Kai Yu. A Hierarchical Decoding Model for Spoken Language Understanding from Unaligned Data In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*,Brighton, UK, 2019,7305-7309

89. Mengyue Wu, Heinrich Dinkel and Kai Yu. Audio Caption: Listen and Tell In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, UK, 2019, 830-834

90. Liliang Ren, Kaige Xie, Lu Chen and Kai Yu. Towards Universal Dialogue State Tracking In *The 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP)*,Brussels, Belgium, 2018: 2780–2786

91. Kaiyu Shi and Kai Yu. Structured Word Embedding for Low Memory Neural Network Language Model In *19th Annual Conference of the International Speech Communication Association (InterSpeech)*, Hyderabad, India,2018, 1254-1258

92. Kuan Chen, Bo Chen, Jiahao Lai and Kai Yu. High-quality Voice Conversion Using Spectrogram-Based WaveNet Vocoder In *19th Annual Conference of the International Speech Communication Association (InterSpeech)* ,Hyderabad, India, 2018, 1993-1997

93. Mingkun Huang, Yongbin You, Zhehuai Chen, Yanmin Qian and Kai Yu. Knowledge Distillation for Sequence Model In *19th Annual Conference of the International Speech Communication Association (InterSpeech)*,Hyderabad, India, 2018, 3703-3707

94. Zili Huang, Shuai Wang and Kai Yu. Angular Softmax for Short-Duration Text-independent Speaker Verification In *19th Annual Conference of the International Speech Communication Association (InterSpeech)*,Hyderabad, India, 2018, 3623-3627.

95. Shuai Wang, Heinrich Dinkel, Yanmin Qian and Kai Yu. Covariance Based Deep Feature for Text-dependent Speaker Verification In *2018 International Conference on Intelligence Science and Big Data Engineering (IScIDE)*,Lanzhou, China, 2018, 228-239

96. Yexin Yang, Shuai Wang, Man Sun, Yanmin Qian and Kai Yu. Generative Adversarial Networks based X-vector Augmentation for Robust Probabilistic Linear Discriminant Analysis in Speaker Verification In *the 11th International Symposium on Chinese Spoken Language Processing (ISCSLP)*, Taipei, Taiwan, 2018,

97. Shuai Wang, Zili Huang, Yanmin Qian and Kai Yu. Deep Discriminant Analysis for i-vector based Robust Speaker Recognition In *the 11th International Symposium on Chinese Spoken Language Processing (ISCSLP)*, Taipei, Taiwan, 2018,

98. Lu Chen, Bowen Tan, Sishan Long and Kai Yu. Structured Dialogue Policy with Graph Neural Networks In *The 27th International Conference on Computational Linguistics (COLING)*,Santa Fe, New Mexico, USA, 2018, 1257–1268

99. Kaige Xie, Cheng Chang, Liliang Ren, Lu Chen and Kai Yu. Cost-Sensitive Active Learning for Dialogue State Tracking In *The 19th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*,Melbourne, Australia, 2018, 209-213.

100. Su Zhu and Kai Yu. Concept Transfer Learning for Adaptive Language Understanding In *19th Annual Meeting of the Special Interest Group on Discourse and Dialogue*,Melbourne, Australia, 2018, 391-399

101. Huifeng Zhang, Su Zhu, Shuai Fan and Kai Yu. Joint Spoken Language Understanding and Domain Adaptive Language Modeling In *2018 International Conference on Intelligence Science and Big Data Engineering (IScIDE)* ,Lanzhou, China, 2018: 305-318

102. Xuan Liu, Di Cao and Kai Yu. Binarized LSTM Language Model In *The 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics(NAACL)*,New Orleans, USA, 2018: 2123-2121

103. Lu Chen, Cheng Chang, Zhi Chen, Bowen Tan, Milica Gasic and Kai Yu. Policy Adaptation for Deep Reinforcement Learning-based Dialogue Management. In *IEEE International Conference on Acoustics Speech and Signal Processing(ICASSP)*,Calgary, Canada, 2018,6074-6078

104. Shuai Wang, Yanmin Qian and Kai Yu. Focal KL-Divergence based Dilated Convolutional Neural Networks for Co-channel Speaker Identification. In *IEEE International Conference on Acoustics Speech and Signal Processing(ICASSP)*,Calgary, Canada, 2018,5339-5343

105. Ruinian Chen and Kai Yu. Fast OOV Words Incorporation using Structured Word Embeddings for Neural Network Language Model. In *IEEE International Conference on Acoustics Speech and Signal Processing(ICASSP)*,Calgary, Canada, 2018,6119-6123

106. Zhehuai Chen, Qi Liu, Hao Li, Kai Yu. On Modular Training of Neural Acoustics-to-word Model for LVCSR. In *IEEE International Conference on Acoustics Speech and Signal Processing(ICASSP)*,Calgary, Canada, 2018, 4754-4758

107. Su Zhu, Ouyu Lan and Kai Yu. Robust Spoken Language Understanding with Unsupervised ASR-error Adaptation In *IEEE International Conference on Acoustics Speech and Signal Processing(ICASSP)*,Calgary, Canada, 2018, 6179-6183

108. Ouyu Lan, Su Zhu and Kai Yu. Semi-Supervised Training Using Adversarial Multi-Task Learning for Spoken Language Understanding. In *IEEE International Conference on Acoustics Speech and Signal Processing(ICASSP)*,Calgary, Canada, 2018, 6049-6053

109. Yue Wu，Qi Liu and Kai Yu. The adaptive adjustment of learning rate is applied in the language model. In *The 14th National Conference on Man-Machine Speech Communication NCMMSC2017)*, Lianyungang, China, 2017

110. Yuncong Hu and Kai Yu Research and improvement of speech model fault tolerance in speech recognition In *In Proc. National Conference on Man-Machine Speech Communication* (NCMMSC),Lianyungang, China, 2017

111. Xuan Liu and Kai Yu GLEU-guided Multiresolution Network for Short Text Conversation In *Proc. The 14th National Conference on Man-Machine Speech Communication* (NCMMSC), Lianyungang, China, 2017

112. Yue Wu, Tianxing He, Zhehuai Chen, Yanmin Qian, Kai Yu. Multi-view LSTM Language Model with Word-synchronized Auxiliary Feature for LVCSR In *The 16th China National Conference on Computational Linguistics* (CCL), Nanjing, China, 2017

113. Zhehuai Chen, Yanmin Qian, Kai Yu. A unified confidence measure framework using auxiliary normalization graph. In *IScIDE*, 2017

114. Di Cao, Kai Yu. Deep Attentive Structured Language Model Based on LSTM. In *IScIDE*, 2017

115. Qi Liu, Yanmin Qian, Kai Yu. Future Vector Enhanced LSTM Language Model for LVCSR In *IEEE Automatic Speech Recognition and Understanding Workshop* (ASRU), Okinawa, Japan, 2017

116. Lu Chen, Xiang Zhou, Cheng Chang, Runzhe Yang, Kai Yu. Agent-Aware Dropout DQN for Safe and Efficient On-line Dialogue Policy Learning In *The 2017 Conference on Empirical Methods in Natural Language Processing* (EMNLP), Copenhagen, Denmark, 2017

117. Cheng Chang, Runzhe Yang, Lu Chen, Xiang Zhou, Kai Yu. Affordable On-line Dialogue Policy Learning In *The 2017 Conference on Empirical Methods in Natural Language Processing* (EMNLP), Copenhagen, Denmark, 2017

118. Xu Xiang, Yanmin Qian, Kai Yu. Binary Deep Neural Networks for Speech Recognition In *18th Annual Conference of the International Speech Communication Association* (InterSpeech), Stockholm, Sweden, 2017

119. Bo Chen, Jiahao Lai, Kai Yu. Comparison of Modeling Target in LSTM-RNN Duration Model In *18th Annual Conference of the International Speech Communication Association* (InterSpeech), Stockholm, Sweden, 2017

120. Bo Chen, Tianling Bian, Kai Yu. Discrete Duration Model For Speech Synthesis In *Annual Conference of the International Speech Communication Association* (InterSpeech), Stockholm, Sweden, 2017

121. Shuai Wang, Yanmin Qian, Kai Yu. What Does the Speaker Embedding Encode? In *18th Annual Conference of the International Speech Communication Association* (InterSpeech), Stockholm, Sweden, 2017

122. Heinrich Dinkel, Yanmin Qian, Kai Yu. Small-footprint convolutional neural network for spoofing detection In *2017 International Joint Conference on Neural Networks* (IJCNN), Anchorage, AK, 2017

123. Lu Chen, Runzhe Yang, Cheng Chang, Zihao Ye, Xiang Zhou, Kai Yu. On-line Dialogue Policy Learning with Companion Teaching In *The 15th Conference of the European Chapter of the Association for Computational Linguistics* (EACL), Valencia, Spain, 2017

124. Heinrich Dinkel, Nanxin Chen, Yanmin Qian, Kai Yu. End-To-End Spoofing Detection With Raw Waveform Cldnns In *IEEE International Conference on Acoustics Speech and Signal Processing* (ICASSP), New Orleans, USA, 2017

125. Su Zhu, Kai Yu. Encoder-decoder with Focus-mechanism for Sequence Labelling Based Spoken Language Understanding In *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), New Orleans, USA, 2017

126. Zhehuai Chen, Yimeng Zhuang, Kai Yu. Confidence Measures for CTC-based Phone Synchronous Decoding In *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), New Orleans, USA, 2017

127. Jiahao Lai, Bo Chen, Tian Tan, Sibo Tong, Kai Yu. Phone-Aware LSTM-RNN for Voice Conversion. In *Proc. 13th IEEE International Conference on Signal Processing* (ICSP), Chengdu, China, 2016.

128. Qi Liu, Tian Tan, Kai Yu. An Investigation on Deep Learning with Beta Stabilizer. In *Proc. 13th IEEE International Conference on Signal Processing* (ICSP), Chengdu, China, 2016.

129. Xueyang Wu, Su Zhu, Yue Wu and Kai Yu. Rich Punctuations Prediction Using Large-scale Deep Learning. In *Proc. International Symposium on Chinese Spoken Language Processing* (ISCSLP), Tianjin, China, 2016.

130. Tianxing He, Yu Zhang, Jasha Droppo and Kai Yu. On Training Bi-directional Neural Network Language Model with Noise Contrastive Estimation. In *Proc. International Symposium on Chinese Spoken Language Processing* (ISCSLP), Tianjin, China, 2016.

131. Yimeng Zhuang, Sibo Tong, Maofan Yin, Yanmin Qian, Kai Yu. Multi-Task Joint-Learning for Robust Voice Activity Detection. In *Proc. International Symposium on Chinese Spoken Language Processing* (ISCSLP), Tianjin, China, 2016.

132. Da Zheng, Zhehuai Chen, Yue Wu, Kai Yu. Directed Automatic Speech Transcription Error Correction Using Bidirectional LSTM. In *Proc. International Symposium on Chinese Spoken Language Processing* (ISCSLP), Tianjin, China, 2016.

133. Kai Sun, Su Zhu, Lu Chen, Siqiu Yao, Xueyang Wu, Kai Yu. Hybrid Dialogue State Tracking for Real World Human-to-Human Dialogues. In *Proc. 17th Annual Conference of the International Speech Communication Association* (InterSpeech), San Francisco, America, 2016.

134. Yimeng Zhuang, Xuankai Chang, Yanmin Qian, Kai Yu. Unrestricted Vocabulary Keyword Spotting Using LSTM-CTC. In *Proc. 17th Annual Conference of the International Speech Communication Association* (InterSpeech), San Francisco, America, 2016.

135. Zhehuai Chen, Wei Deng, Tao Xu, Kai Yu. Phone Synchronous Decoding with CTC Lattice. In *Proc. 17th Annual Conference of the International Speech Communication Association* (InterSpeech), San Francisco, America, 2016.

136. Maofan Yin, Sunil Sivadas, Kai Yu, Bin Ma. Discriminatively Trained Joint Speaker and Environment Representations for Adaptation of Deep Neural Network Acoustic Models. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), Shanghai, China, 2016.

137. Sibo Tong, Hao Gu and Kai Yu. A Comparative Study of Robustness of Deep Learning Approaches for VAD. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), Shanghai, China, 2016.

138. Wengong Jin, Tianxing He, Yanmin Qian and Kai Yu. Paragraph Vector based Topic Model for Language Model Adaptation. In *Proc. 16th Annual Conference of the International Speech Communication Association* (InterSpeech), Dresden, Germany, 2015.

139. Nanxin Chen, Yanmin Qian and Kai Yu. Multi-Task Learning for Text-dependent Speaker Verication. In *Proc. 16th Annual Conference of the International Speech Communication Association* (InterSpeech), Dresden, Germany, 2015.

140. Nanxin Chen, Yanmin Qian, Heinrich Dinkel, Bo Chen and Kai Yu. Robust Deep Feature for Spoofing Detection - The SJTU System for ASVspoof 2015 Challenge. In *Proc. 16th Annual Conference of the International Speech Communication Association* (InterSpeech), Dresden, Germany, 2015.

141. Bo Chen, Zhehuai Chen, Jiachen Xu and Kai Yu. An Investigation of Context Clustering for Statistical Speech Synthesis with Deep Neural Network. In *Proc. 16th Annual Conference of the International Speech Communication Association* (InterSpeech), Dresden, Germany, 2015.

142. Mengxiao Bi, Yanmin Qian, Kai Yu. Very Deep Convolutional Neural Networks for LVCSR. In *Proc. 16th Annual Conference of the International Speech Communication Association* (InterSpeech), Dresden, Germany, 2015.

143. Qizhe Xie, Kai Sun, Su Zhu, Lu Chen and Kai Yu. Recurrent Polynomial Network for Dialogue State Tracking with Mismatched Semantic Parsers. In *Proc. Annual SIGdial Meeting on Discourse and Dialogue* (SIGDIAL), Prague, Czech Republic, 2015.

144. Tianxing He, Xu Xiang, Yanmin Qian and Kai Yu. Recurrent Neural Network Language Model with Structured Word Embeddings for Speech Recognition. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Brisbane, Australia, 2015.

145. Tian Tan, Yanmin Qian, Maofan Yin, Yimeng Zhuang and Kai Yu. Cluster Adaptive Training for Deep Neural Network. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Brisbane, Australia, 2015.

146. Yanmin Qian, Maofan Yin, Yongbin You and Kai Yu. Multi-Task Joint-Learning of Deep Neural Networks for Robust Speech Recognition. In *Proc. IEEE Automatic Speech Recognition and Understanding Workshop* (ASRU), Scottsdale, Arizona, USA, 2015.

147. Yanmin Qian, Tianxing He, Wei Deng and Kai Yu. Automatic Model Redundancy Reduction for Fast Back-Propagation for Deep Neural Networks in Speech Recognition. In *Proc. IEEE International Joint Conference on Neural Networks* (IJCNN), Killarney, Ireland, 2015.

148. Yongbin You, Yanmin Qian, Tianxing He and Kai Yu. An Investigation on DNN-Derided Bottleneck Features for GMM-HMM Based Robust Speech Recognition. In *Proc. IEEE China Summit and International Conference on Signal and Information Processing* (ChinaSIP), Chengdu, China, 2015.

149. Yongbin You, Yanmin Qian and Kai Yu. Local Trajectory Based Speech Enhancement for Robust Speech Recognition With Deep Neural Network. In *Proc. IEEE China Summit and International Conference on Signal and Information Processing* (ChinaSIP), Chengdu, China, 2015.

150. Yuan Liu, Tianfan Fu, Yuchen Fan, Yanmin Qian and Kai Yu. Speaker Verification with Deep Features. In *Proc. IEEE International Joint Conference on Neural Networks* (IJCNN), Beijing, China, 2014.

151. Kai Sun, Lu Chen, Su Zhu and Kai Yu. The SJTU System for Dialog State Tracking Challenge 2. In *Proc. Annual Meeting of the Special Interest Group on Discourse and Dialogue* (SIGDIAL), Pennsylvania, USA, 2014.

152. Su Zhu, Lu Chen, Kai Sun, Da Zheng and Kai Yu. Semantic parser enhancement for dialogue domain extensions with little data. In *Proc. IEEE Spoken Language Technology Workshop* (SLT), South Lake Tahoe, USA, 2014.

153. Kai Sun, Lu Chen, Su Zhu and Kai Yu. A Generalized Rule Based Tracker for Dialogue State Tracking. In *Proc. IEEE Spoken Language Technology Workshop* (SLT), South Lake Tahoe, USA, 2014.

154. Zhehuai Chen and Kai Yu. An Investigation of Implementation and Performance Analysis of DNN Based Speech Synthesis System. In *Proc. IEEE International Conference on Signal Processing* (ICSP), Hangzhou, China, 2014.

155. Sibo Tong, Nanxin Chen, Yanmin Qian and Kai Yu. Evaluating VAD For Automatic Speech Recognition. In *Proc. IEEE International Conference on Signal Processing* (ICSP), Hangzhou, China, 2014.

156. Suliang Bu, Yanmin Qian and Kai Yu. A Novel Dynamic Parameters Calculation Approach For Model Compensation. In *Proc. 15th Annual Conference of the International Speech Communication Association* (InterSpeech), Singapore, 2014.

157. Tianfan Fu, Yanmin Qian, Yuan Liu and Kai Yu. Tandem Deep Features for Text-Dependent Speaker Verification. In *Proc. 15th Annual Conference of the International Speech Communication Association* (InterSpeech), Singapore, 2014.

158. Jianwei Niu, Yanmin Qian and Kai Yu. Acoustic Emotion Recognition using Deep Neural Network. In *Proc. International Symposium on Chinese Spoken Language Processing* (ISCSLP), Florence, Italty, 2014.

159. Tianxing He, Yuchen Fan, Yanmin Qian, Tian Tan and Kai Yu. Reshaping Deep Neural Network For Fast Decoding By Node Pruning. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Florence, Italty, 2014.

160. Suliang Bu, Yanmin Qian, Khe Chai Sim, Yongbin You, Kai Yu. Second order vector taylor series based robust speech recognition. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Florence, Italty, 2014.

161. Wei Deng, Yanmin Qian, Yuchen Fan, Tianfan Fu and Kai Yu. Stochastic data sweeping for fast DNN training. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Florence, Italty, 2014.

162. Yanmin Qian, Kai Yu and Jia Liu. Combination of Data Borrowing Strategies for Low-Resource LVCSR. In *Proc. IEEE Automatic Speech Recognition and Understanding Workshop* (ASRU), Olomouc, Czech Republic, 2013.

163. Kai Yu and Hainan Xu. Cluster Adaptive Training With Factorized Decision Trees For Speech Recognition. In *Proc. 14th Annual Conference of the International Speech Communication Association* (InterSpeech), Lyon, France, 2013.

164. Peilu Wang, Ruihua Sun, Hai Zhao, Kai Yu. A New Word Language Model Evaluation Metric For Character Based Languages In *The 12th China National Conference on Computational Linguistics*(CNNCL), Suzhou, China, 2013

165. Hainan Xu, Yuchen Fan and Kai Yu. Development of the 2012 SJTU HVR System. In *Proc. International Conference on Multimodal Interaction* (ICMI), Santa Monica, U.S.A., 2012.

166. Kai Yu. Review of F0 modelling and generation in HMM based speech synthesis. In *Proc. The 11th International Conference on Signal Processing* (ICSP), Beijing, China, 2012.

167. Matthew Henderson, Milica Gasic, Blaise Thomson, Pirros Tsiakoulis, Kai Yu, Steve Young. Discriminative Spoken Language Understanding Using Word Confusion Networks. In *Proc. IEEE Spoken Language Technology Workshop* (SLT), Miami, U.S.A., 2012.

168. Milica Gasic, Pirros Tsiakoulis, Matt Henderson, Blaise Thomson, Kai Yu, Eli Tzirkel and Steve Young. The effect of cognitive load on a statistical dialogue system. In *Proc. SIGDIAL Workshop on Discourse and Dialogue* (SIGDIAL), Seoul, South Korea, 2012.

169. Khechai Sim, Shengdong Zhao, Kai Yu and Hank Liao. ICMI'12 Grand Challenge - Haptic Voice Recognition. In *Proc. The 14th ACM International Conference on Multimodal Interaction* (ICMI), Santa Monica, U.S.A., 2012.

170. Milica Gasic, Filip Jurcicek, Blaise Thomson, Kai Yu and Steve Young. On-line policy optimisation of spoken dialogue systems via live interaction with human subjects. In *Proc. IEEE Automatic Speech Recognition and Understanding Workshop* (ASRU), New York, U.S.A., 2011.

171. Pirros Tsiakoulis, Milica Gasic, Matthew Henderson, Jorge Prombonas, Blaise Thomson, Kai Yu, Steve Young. Statistical Methods for Building Robust Spoken Dialogue Systems in an Automobile. In *Proc. The 4th International Conference on Applied Human Factors and Ergonomics* (AHFE), San Francisco, U.S.A., 2012.

172. Alan W Black, Susanne Burger, Alistair Conkie, Helen Hastie, Simon Keizer, Oliver Lemon, Nicolas Merigaud, Gabriel Parent, Gabriel Schubiner, Blaise Thomson, Jason D. Williams, Kai Yu, Steve Young, and Maxine Eskenazi. Spoken Dialog Challenge 2010: Comparison of Live and Control Test Results . In *Proc. SIGDIAL Workshop on Discourse and Dialogue* (SIGDIAL), Portland, Oregon, 2011.

173. Filip Jurcicek, Simon Keizer, Milica Gasic, Francois Mairesse, Blaise Thomson, Kai Yu, and Steve Young. Real user evaluation of spoken dialogue systems using Amazon Mechanical Turk. In *Proc. The International Speech Communication Association* (InterSpeech), Florence, Italy, 2011.

174. Kai Yu and Steve Young. Joint modelling of voicing label and continuous F0 for HMM based speech synthesis. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Prague, Czech, 2011.

175. Lei Jia, Kai Yu and Bo Xu. Structured precision modelling with Cholesky basis superposition for speech recognition. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Prague, Czech, 2011.

176. Blaise Thomson, Kai Yu, Simon Keizer, Milica Gasic, Filip Jurcicek, Francois Mairesse and Steve Young. Bayesian dialogue system for the Let's Go spoken dialogue challenge. In *Proc. IEEE Spoken Language Technology Workshop* (SLT), Berkeley, U.S.A., 2010.

177. Blaise Thomson, Filip Jurcicek, Milica Gasic, Simon Keizer, Francois Mairesse, Kai Yu and Steve Young. Parameter learning for POMDP spoken dialogue models. In *Proc. IEEE Spoken Language Technology Workshop* (SLT), Berkeley, U.S.A., 2010.

178. Kai Yu, Blaise Thomson and Steve Young. From discontinuous to continuous F0 modelling In HMM-based speech synthesis. In *The 7th International Society for Compute Association Speech Synthesis Workshop* (ISCA SSW7), Kyoto, Japan, 2010.

179. Kai Yu, Heiga Zen, Francois Mairesse and Steve Young. Context adaptive training with factorized decision trees for HMM-based speech synthesis. In *Proc. The International Speech Communication Association* (InterSpeech), Makuhari, Japan, 2010.

180. Mark Gales and Kai Yu. Canonical state models for automatic speech recognition. In *Proc. The International Speech Communication Association* (InterSpeech), Makuhari, Japan, 2010.

181. Filip Jurcicek, Blaise Thomson, Simon Keizer, Francois Mairesse, Milica Gasic, Kai Yu and Steve Young. Natural Belief-Critic: a reinforcement algorithm for parameter estimation in statistical spoken dialogue systems. In *Proc. The International Speech Communication Association* (InterSpeech), Makuhari, Japan, 2010.

182. Milica Gasic, Filip Jurcicek, Simon Keizer, Francois Mairesse, Blaise Thomson, Kai Yu and Steve Young. Gaussian processes for fast policy optimisation of POMDP-based dialogue managers. In *The Annual SIGdial Meeting on Discourse and Dialogue* (SIGDIAL), Tokyo, Japan, 2010.

183. Simon Keizer, Milica Gasic, Filip Jurcicek, Francois Mairesse, Blaise Thomson, Kai Yu and Steve Young. Parameter estimation for agenda-based user simulation. In *The Annual SIGdial Meeting on Discourse and Dialogue* (SIGDIAL), Tokyo, Japan, 2010.

184. Francois Mairesse, Milica Gasic, Filip Jurcicek, Simon Keizer, Jorge Prombonas, Blaise Thomson, Kai Yu and Steve Young. Phrase-based statistical language generation using graphical models and active learning. In *Proc. The Annual Meeting on Association for Computational Linguistics* (ACL), Uppsala, Sweden, 2010.

185. Kai Yu, Francois Mairesse and Steve Young. Word-level emphasis modelling in HMM-based speech synthesis. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Dallas, U.S.A., 2010.

186. Milica Gasic, Fabrice Lefevre, Filip Jurcicek, Simon Keizer, Francois Mairesse, Blaise Thomson, Kai Yu and Steve Young. Back-off action selection in summary space-based POMDP dialogue systems. In *Proc. IEEE Automatic Speech Recognition and Understanding Workshop* (ASRU), Merano, Italy, 2009.

187. Fabrice Lefevre, Milica Gasic, Filip Jurcicek, Simon Keizer, Francois Mairesse, Blaise Thomson, Kai Yu and Steve Young. K-nearest neighbor Monte-Carlo control algorithm for POMDP-based dialogue systems. In *Proc. SIGDIAL Workshop on Discourse and Dialogue* (SIGDIAL), London, U.K., 2009.

188. Filip Jurcicek, Milica Gasic, Simon Keizer, Francois Mairesse, Blaise Thomson, Kai Yu and Steve Young. Transformation-based learning for semantic parsing. In *Proc. The Annual Conference of the International Speech Communication Association* (InterSpeech), Brighton, U.K., 2009.

189. Kai Yu, Tomiki Toda, Milica Gasic, Simon Keizer, Francois Mairesse, Blaise Thomson and Steve Young. Probabilistic modelling of F0 in unvoiced regions in HMM based speech synthesis. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Taipei, Taiwan, 2009.

190. Francois Mairesse, Milica Gasic, Filip Jurcicek, Simon Keizer, Blaise Thomson, Kai Yu and Steve Young. Spoken language understanding from unaligned data using discriminative classification models. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Taipei, Taiwan, 2009.

191. Simon Keizer, Milica Gasic, Francois Mairesse, Blaise Thomson, Kai Yu and Steve Young. Modelling user behaviour in the HIS-POMDP dialogue manager. In *Proc. IEEE Workshop on Spoken Language Technology* (SLT), Goa, India, 2008.

192. Chandra K. Raut, Kai Yu and Mark Gales. Adaptive training using discriminative mapping transforms. In *Proc. The Annual Conference of the International Speech Communication Association* (InterSpeech), Brisbane, Australia, 2008.

193. Blaise Thomson, Kai Yu, Milica Gasic, Simon Keizer, Francois Mairesse, Jost Schatzmann and Steve Young. Evaluating semantic-level confidence scores with multiple hypotheses. In *Proc. The Annual Conference of the International Speech Communication Association* (InterSpeech), Brisbane, Australia, 2008.

194. Blaise Thomson, Milica Gasic, Simon Keizer, Francois Mairesse, Jost Schatzmann, Kai Yu and Steve Young. User study of the Bayesian Update of Dialogue State approach to dialogue management. In *Proc. The Annual Conference of the International Speech Communication Association* (InterSpeech), Brisbane, Australia, 2008.

195. Milica Gasic, Simon Keizer, Francois Mairesse, Jost Schatzmann, Blaise Thomson, Kai Yu and Steve Young. Training and evaluation of the HIS POMDP dialogue system in noise. In *Proc. SIGDIAL Workshop on Discourse and Dialogue* (SIGDIAL), Columbus, U.S.A., 2008.

196. Kai Yu, Mark Gales, and Philip Woodland. Unsupervised discriminative adaptation using discriminative mapping transforms. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Las Vegas, U.S.A., 2008.

197. Xunying Liu, William Byrne, Mark Gales, Adria Gispert, Marcus Tomalin, Phillip Woodland and Kai Yu. Discriminative language model adaptation for Mandarin broadcast speech transcription and translation. In *Proc. IEEE Automatic Speech Recognition and Understanding Workshop* (ASRU), Kyoto, Japan, 2007.

198. Mark Gales, Frank Diehl, Chandra Raut, Marcus Tomalin, Phillip Woodland and Kai Yu. Development of a phonetic system for large vocabulary Arabic speech recognition. In *Proc. IEEE Automatic Speech Recognition and Understanding Workshop* (ASRU), Kyoto, Japan, 2007.

199. Kai Yu, Mark Gales, and Philip Woodland. Unsupervised training with directed manual transcription for recognizing Mandarin broadcast audio. In *Proc. The Annual Conference of the International Speech Communication Association* (Inter-Speech), Antwerp, Belgium, 2007.

200. Mark Gales, Xunying Liu, Rohit Sinha, Philip Woodland, Kai Yu, Spyridon Matsoukas, Tim Ng, Kham Nguyen, Long Nguyen, Jean-Luc Gauvain, Lori Lamel, Abdelkhalek Messaoudi. Speech recognition system combination for machine translation. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Honolulu, U.S.A., 2007.

201. Marcus Tomalin, Mark Gales, Xunying Liu, Khe Chai Sim, Rohit Sinha, Lan Wang, Phillip Woodland and Kai Yu. Improving speech transcription for Mandarin-English translation. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Honolulu, U.S.A., 2007.

202. Kai Yu and Mark Gales. Incremental adaptation using Bayesian inference. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Toulouse, France, 2006.

203. Kai Yu and Mark Gales. Bayesian adaptation and adaptively trained systems. In *Proc. IEEE Automatic Speech Recognition and Understanding Workshop* (ASRU), Cancun, Mexico, 2005.

204. Gunnar Evermann, Ho Yin Chan, Mark Gales, Bin Jia, David Mrva, Philip Woodland, and Kai Yu. Training LVCSR systems on thousands of hours of data. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Philadelphia, U.S.A., 2005.

205. Mark Gales, Bin Jia, Xunying Liu, Khe Chai Sim, Philip Woodland, and Kai Yu. Development of the CUHTK 2004 Mandarin conversational telephone speech transcription system. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Philadelphia, U.S.A., 2005.

206. Xunying Liu, Mark Gales, Khe Chai Sim, and Kai Yu. Investigation of acoustic modelling techniques for LVCSR systems. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Philadelphia, U.S.A., 2005.

207. Kai Yu and Mark Gales. Adaptive training using structured transforms. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Montreal, Canada, 2004.

208. Sue Tranter, Kai Yu, Gunnar Evermann, and Philip Woodland. Generating and evaluating segmentations for automatic speech recognition of conversational telephone speech. In *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* (ICASSP), Montreal, Canada, 2004.

Technical Report

1. Kai Yu and Lu Chen. Dialogue state tracking in statistical dialogue management. Technical Report SP2014-108(2014-12), The Institute of Electronics, Information and Communication Engineers (IEICE), 2014.

2. Kai Yu and Mark Gales. Bayesian adaptation and adaptive training. Technical Report CUED/F-INFENG/TR542, Cambridge University Engineering Department, 2006.

3. Kai Yu and Mark Gales. Discriminative cluster adaptive training. Technical Report CUED/F-INFENG/TR486, Cambridge University Engineering Department, 2004.

4. Sue Tranter, Kai Yu, Douglas Reynolds, Gunnar Evermann, Do Yeong Kim, and Philip Woodland. An investigation into the interactions between speaker diarization systems and automatic speech transcription. Technical Report CUED/F-INFENG/TR464, Cambridge University Engineering Department, 2003.