Shinji Watanabe

Office Address
Johns Hopkins University
3400 North Charles Street, Baltimore, MD
21218-2608, USA
https://sites.google.com/view/shinjiwatanabe/
shinjiw@ieee.org,shinjiw@jhu.edu

EDUCATION

PhD, Computer Engineering
Waseda University, Tokyo, Japan   February 2006
Thesis: Speech Recognition based on a Bayesian approach
Supervisor: Prof. Tetsunori Kobayashi

Master of Science, Theoretical Physics
Waseda University, Tokyo, Japan   March 2001
Thesis: Evaluation of T Violation in Neutrino Oscillation using Geological Layers
Supervisor: Prof. Ichiro Ohba

Bachelor of Science, Theoretical Physics
Waseda University, Tokyo, Japan   March 1999
Thesis: Solar Neutrino Oscillation and CP Violation
Supervisor: Prof. Ichiro Ohba

PROFESSIONAL EXPERIENCE

Johns Hopkins University, Baltimore, MD, USA   July 2017 –
  Associate Research Professor (July 2017 –)

- Noise robust/distant talk speech recognition, speaker recognition, and language recognition
  - Organization of noise robust/distant talk speech recognition challenges (CHiME-5 and CHiME-6)
  - Development of benchmark/state-of-the-art speech recognition, speaker recognition, and language recognition systems
  - Deep learning based speech separation/enhancement and beamforming
  - Development of state-of-the-art speaker diarization

- End-to-end speech recognition
  - Development of opne source end-to-end ASR tool (ESPnet: https://github.com/espnet/espnet)
  - Novel architecture for end-to-end ASR
  - Multilingual, multichannel, and multispeaker extensions of end-to-end ASR
  - Text to speech
  - Speech translation

Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA, USA   January 2012 – June 2017
  Visiting Member (January 2012 – March 2013)
  Principal Member (April 2013 - September 2014)
  Senior Principal Member (October 2014 - June 2017)
  - Noise robust/distant talk speech recognition

- Organization/participation of noise robust/distant talk speech recognition challenges (CHiME2/3/4 and REVERB challenges)
- Development of benchmark/state-of-the-art speech recognition system
- Deep learning based speech enhancement and beamforming
- End-to-end multichannel speech recognition

- Speech recognition in general
  - End-to-end speech recognition using hybrid attention/CTC
  - Black box optimization of neural network based speech recognition
  - Uncertainty issues in speech recognition

- Machine learning for speech and language processing
  - Bayesian approach
  - Deep learning

- Spoken dialog systems
  - Statistical dialog manager
  - Spoken language understanding based on Recurrent neural network
  - Application of word embedding to spoken language understanding

NTT Communication Science Laboratories, Kyoto, Japan   April 2001 – November 2011

Researcher (April 2001 - May 2011)

Research Scientist (June 2011 - November 2011)

- Acoustic and language models in speech recognition
  - Training/Adaptation of hidden Markov models (HMMs)
  - Topic based N-gram language models
  - Inference algorithm (EM algorithm, Gibbs sampling)
  - Finite State Transducer based speech and language processing

- Machine learning for speech and language processing
  - Variational Bayes
  - Markov chain Monte Carlo
  - Discriminative method

AWARD, NOTABLE ACHIEVMENT

1. Google Faculty Research Award in 2019
2. Tencent AI Lab Rhino-Bird Gift Fund in 2018
3. Facebook Research Award (Speech and Audio Technology for Voice Interaction and Video Understanding) in 2018
4. Co-authored paper achieved the ISCA Interspeech Best Student Paper Award in 2018
5. Ranked 1st place at The First DIHARD Speech Diarization Challenge track1 in 2018
6. Best paper candidates (2 papers) and best student paper candidate at the IEEE ASRU in 2017
7. Ranked 3rd place at 4th CHiME Challenge, 6 channel track in 2016
8. Ranked 2nd place at 5th Dialog State Tracking Challenge (DSTC5) in 2016
9. Ranked 2nd place at 3rd CHiME Challenge in 2015
10. Ranked 1st place at REVERB Challenge Official Data Track (2nd overall) in 2014
11. Ranked 1st place at 2nd CHiME Challenge Track 2 in 2013
12. Co-authored paper achieved the IEEE ICASSP 2012 Student Paper Award in 2012
13. Best In-Category Nominee at the International Semantic Web Conference in 2011
14. Ranked 1st place at 1st CHiME Challenge in 2011
15. Co-authored paper achieved the IEEE Signal Processing Society Japan Chapter Student Paper Award in 2010
16. The TELECOM System Technology Award from the Telecommunications Advancement Foundation in 2006

17. The Itakura Award from the Acoustical Society of Japan (ASJ) in 2006
18. The Best Paper Award from the Institute of Electronics, Information, and Communication Engineers of Japan (IEICE) in 2004
19. The Awaya Award from the Acoustical Society of Japan (ASJ) in 2003

PROGRAMING SKILL

- Experienced in Linux programming based on C/C++, Python, Matlab, and Bash/Csh
- Experienced in speech recognition toolkit (Kaldi, ESPnet, HTK, and NTT SOLON) and deep learning toolkit (Theano, Chainer, PyTorch, and CNTK)

LANGUAGE SKILL

- Japanese (native)
- English (fluent)

TEACHING

- "EN. 601.667: Introduction to Human Language Technology", Johns Hopkins University, USA, Spring 2018-
- "End-to-end speech recognition" JHU Summer School on Human Language Technology (June 20, 2018)
- "EN. 520.666: Information Extraction", Johns Hopkins University, USA, Spring 2018
- "Machine Learning for Speech Processing," Department of Electronic Engineering, Tsinghua University, China, December 2012

KEYNOTE TALK

1. Shinji Watanabe, "End-to-End Speech Processing: From Pipeline to Integrated Architecture," Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) 2018, Honolulu, Hawaii, USA (2018)
2. Shinji Watanabe, "Neural End-to End Architectures for Speech Recognition in Adverse Environments," ITG Conference on Speech Communication, Oldenburg, Germany (2018)
3. Shinji Watanabe, "Neural End-to End Architectures for Speech Recognition in Adverse Environments," 30th annual Conference on Computational Linguistics and Speech Processing (ROCLING), Hsinchu, Taiwan (2018)

SEMINAR/INVITED TALK

1. "End-to-End Speech Processing: From Pipeline to Integrated Architecture," CMU, USA, October 2019 (host: Prof. Florian Metze)
2. "End-to-End Speech Processing: From Pipeline to Integrated Architecture," IBM, USA, August 2019 (host: Dr. Xiaodong Cui)
3. "End-to-End Speech Processing: From Pipeline to Integrated Architecture," University of Sheffield, UK, May 2019 (host: Prof. Jon Barker)
4. "Multichannel End-to-End Speech Recognition, Toyota Technological Institute at Chicago (TTIC), USA, November 2018 (host: Prof. Karen Livescu)
5. "Multichannel End-to-End Speech Recognition, Paderborn University, Germany, October 2018 (host: Prof. Reinhold Hb-Umbach)

6. "Hybrid CTC/Attention Architecture for End-to-End Speech Recognition," National Chiao Tung University, Taiwan, October 2018 (host: Prof. Jen-Tzung Chien)
7. "Multichannel End-to-End Speech Recognition, Academia Sinica, Taiwan, October 2018 (host: Prof. Yu Tsao)
8. "Distant-Talk Speech Processing Toward Natural Conversation Understanding, Johns Hopkins University, USA, January 2017 (host: Prof. Najim Dehak)
9. "Pushing the envelope at both ends beamforming acoustic models and joint CTC/Attention schemes for end-to-end ASR, Tokyo Institute of Technology, Japan, December. 2016 (host: Prof. Koichi Shinoda)
10. "Recent activities of distant talk speech recognition, Nanyang Technological University, Singapore, March. 2016 (host: Dr. Xiong Xiao)
11. "Recent activities of distant talk speech recognition, Shanghai Jiao Tong University, China, March. 2016 (host: Prof. Kai Yu)
12. Practical Bayesian methods for speech and language processing, Brno University of Technology, Czech Republic, September. 2015 (host: Profs. Jan "Honza" Cernocky)
13. Recent trends in far-field speech recognition, Nara Institute of Science and Technology, Japan, March. 2015 (host: Profs. Kevin Duh and Satoshi Nakamura)
14. Recent trends on noise robust speech recognition, Tokyo Institute of Technology, Japan, October. 2014 (host: Prof. Takahiro Shinozaki)
15. Bayesian Learning for Speech and Language Processing, The Institute of Statistical Mathematics, Japan, March. 2013 (host: Prof. Tomoko Matsui)
16. Bayesian Learning for Speech and Language Processing, Ruhr-Universitt Bochum, Germany, Sep. 2013 (host: Prof. Dorothea Kolossa)
17. Bayesian Learning for Speech and Language Processing, National Taiwan University, Taiwan, July. 2013 (host: Prof. Lin-shan Lee)
18. Structural Bayesian linear regression for hidden Markov models Speech Signal Processing Workshop 2013, Taiwan, June. 2013 (host: Prof. Jen-Tzung Chien)
19. The 4th Young Researchers Forum on ALAGIN Speech Processing Session, Japan, Dec. 2012 (host: Prof. Daisuke Saito)
20. Bayesian Learning for Speech and Language Processing, BBN Technology, USA, Jun. 2012 (host: Dr. Ivan Bulyko)
21. Speech Recognition by Tracking Acoustical and Linguistic Environments, MIT CSAIL Seminar, USA, April 2012 (host: Prof. Jim Glass)
22. Bayesian Linear Regression For Hidden Markov Model Based On Optimizing Variational Bounds, NICT, Japan, Nov. 2011 (host: Dr. Chiori Hori)
23. Bayesian Linear Regression For Hidden Markov Model Based On Optimizing Variational Bounds, Tsinghua University, China, Sep 2011 (host: Prof. Zhijian Ou)
24. Incremental adaptation of speech recognition based on macroscopic time evolution system, Seminar on Human Language Technology and Pattern Recognition, RWTH Aachen University, Germany, Sep 2011 (host: Prof. Ralf Schlueter)
25. Incremental adaptation of speech recognition based on macroscopic time evolution system, Seminar on Tokuda & Lee Laboratory, Nagoya Institute of Technology, Japan, June 2011 (host: Prof. Keiichi Tokuda)
26. Incremental adaptation of speech recognition based on macroscopic time evolution system, Seminar on Machine Intelligence Laboratory, Cambridge University Department of Engineering, UK, May 2011 (host: Dr. Kai Yu)
27. Speech recognition based on a Bayesian approach, Seminar on NHK Science & Technology Laboratories, Japan, September 2009 (host: Dr. Toru Imai)
28. Variational Bayesian estimation and clustering for speech recognition, Georgia tech. CSIP seminar, USA, February 2009 (host: Prof. Fred Juang).

PUBLICATIONS (Book/Book Chapter)

1. **Shinji Watanabe**, Tuomas Virtanen, Dorothea Kolossa, "Application of Source Separation to Robust Speech Analysis and Recognition," in "Audio Source Separation and Speech Enhancement," eds. Emmanuel Vincent, Tuomas Virtanen, Sharon Gannot, John Wiley & Sons Ltd (2018)
2. **Shinji Watanabe**, Marc Delcroix, Florian Metze, and John R. Hershey, "New Era for Robust Speech Recognition -Exploiting Deep Learning-," Springer (2017)
3. **Shinji Watanabe** and Jen-Tzung Chien, "Bayesian Speech and Language Processing," Cambridge University Press (2015)
4. Marc Delcroix, **Shinji Watanabe**, and Tomohiro Nakatani, " Variance Compensation for Recognition of Reverberant Speech with Dereverberation Preprocessing, " in Robust Speech Recognition of Uncertain or Missing Data, eds. Dorothea Kolossa and Reinhold Haeb-Umbach, Springer, pp. 225–256, (2011)

PUBLICATIONS (TUTORIAL/OVERVIEW)

1. Takaaki Hori Tomoki Hayashi, Shigeki Karita, and **Shinji Watanabe**, Advanced Methods for Neural End-to-End Speech Processing - Unification, Integration, and Implementation,"Interspeech'19 (2019)
2. **Shinji Watanabe**, "Hybrid CTC/Attention Architecture for End-to-End Speech Recognition in Challenging Environments," INTERNATIONAL WORKSHOP ON SPEECH PROCESSING IN CHALLENGING ENVIRONMENTS (SPICE) (2018)
3. **Shinji Watanabe**, "Multi-microphone Speech Recognition –Recent Advances in Distant Speech Recognition," Speech and Language Symposium (2016)
4. **Shinji Watanabe**, Xiong Xiao, and Marc Delcrioix "Multi-microphone Speech Recognition," APSIPA ASC'16 (2016)
5. **Shinji Watanabe**, "Pushing the envelope at both ends: beamforming acoustic models and joint CTC/Attention schemes for end-to-end ASR," SANE 2016 - Speech and Audio in the Northeast (2016)
6. Marc Delcroix and **Shinji Watanabe**, "Recent Advances in Distant Speech Recognition," Interspeech'16 (2016)
7. **Shinji Watanabe** and Jen-Tzung Chien, " Bayesian Learning for Speech and Language Processing," ICASSP12 2012, Tutorial T-10, (2012)
8. **Shinji Watanabe**, "Bayesian approaches in speech recognition," APSIPA ASC 2011, Plenary Overview Sessions, (2011)
9. **Shinji Watanabe**, Resent Topics in Acoustic Modeling for Speech recognition: A Machine Learning Perspective based on Generative and Discriminative Approaches," IEICE Technical Report, SP2011-31, pp. 710 (2011.6)
10. **Shinji Watanabe**, Recent topics in speech research, Tutorial Session 1-2, the 2nd Young Researchers Forum on ALAGIN Speech Processing Session (2011.3)
11. **Shinji Watanabe**, Bayesian training of acoustic models, IPSJ SIG Technical Report, Vol.2009-SLP-77, No. 9, (2009. 7)
12. **Shinji Watanabe**, Speech recognition based on a Bayesian approach and its application, Japan Techno Center, Lecture on Speech recognition technology application and future trend (2007.2.)
13. **Shinji Watanabe**, Yasuhiro Minami, Atsushi Nakamura, and Naonori Ueda,Speech recognition using variational Bayes, 2005 Workshop on Information-Based Induction Sciences (IBIS2005), pp. 269274, (2005.11)
14. **Shinji Watanabe**, Selection of shared-state hidden Markov model structure using Bayesian criterion," IEICE Technical Report, SP2004-149, pp. 2530 (2005)
15. **Shinji Watanabe**, Speech recognition based on the Bayesian approach, Tutorial lecture on IEICE Technical Report, SP2004-74, pp. 13–20, (2004)
16. **Shinji Watanabe**, VBEC: robust speech recognition based on Bayesian approach," 5th young researcher workshop, ASJ Kansai section, I-2, (2003)

PUBLICATIONS (REVIEW PAPER)

1. Reinhold Haeb-Umbach, **Shinji Watanabe**, Tomohiro Nakatani, Michiel Bacchiani, Bjrn Hoffmeister, Michael L. Seltzer, Heiga Zen, and Mehrez Souden, "Speech Processing for Digital HomeAssistants," IEEE Signal Processing Magazine, vol. 36, no. 6, pp. 111–124 (2019)
2. Takahiro Shinozaki and **Shinji Watanabe**, "Automatic speech recognition and black-box optimization," The Journal of the Acoustical Society of Japan, vol. 72, number 10, pp. 644–652, (2016). (in Japanese)
3. Felix Weninger, Hakan Erdogan, **Shinji Watanabe**, Emmanuel Vincent, Jonathan Le Roux, John R. Hershey, and Bjoern Schuller, "Speech enhancement with LSTM recurrent neural networks and its application to noise-robust ASR," Proc. Latent Variable Analysis and Signal Separation, pp. 91–99 (2015)
4. **Shinji Watanabe** and Atsushi Nakamura, "Bayesian Approaches to Acoustic Modeling: A Review," APSIPA Transactions on Signal and Information Processing, Volume 1, e5 (11 pages), (2012)
5. Mark Gales, **Shinji Watanabe**, Eric Fossler-Lussier, Structured Discriminative Models for Speech Recognition, IEEE Signal Processing Magazine, vol. 29, no.6, pp. 70–82 (2012).
6. **Shinji Watanabe** and Atsushi Nakamura, "Tutorial: Discriminative Training in Speech Recognition," (in Japanese) The Journal of the Institute of Electronics, Information and Communication Engineers (IEICE), vol. 94(10), pp. 920–922, (2011)
7. **Shinji Watanabe**, "Acoustic models in speech recognition," The Journal of the Acoustical Society of Japan, vol. 66, number 1, pp. 18–22, (2010). (in Japanese)
8. Atsushi Nakamura, **Shinji Watanabe**, Takaaki Hori, Erik, McDermott, and Shigeru Katagiri, "Advanced Computational Models and Learning Theories for Spoken Language Processing," IEEE Computational Intelligence Magazine, vol. 1, issue 2, pp. 5–9, (2006)
9. **Shinji Watanabe**, "Speech recognition based on a Bayesian approach," The Journal of the Acoustical Society of Japan, vol. 62, number 8, pp. 599–604, (2006). (in Japanese)

PUBLICATIONS (JOURNAL, REFREED)

1. Ruizhi Li, Xiaofei Wang, Sri Harish Mallidi, **Shinji Watanabe**, Takaaki Hori, and Hynek Hermansky, "Multi-Stream End-to-End Speech Recognition," IEEE Transactions on Audio, Speech and Language Processing (accepted)
2. Jonathan Le Roux, Gordon Wichern, **Shinji Watanabe**, Andy Sarroff, and John R. Hershey, Phasebook and Friends: Leveraging discrete representations for source separation," IEEE Journal of Selected Topics in Signal Processing, pp. 370–382 (2019)
3. Takafumi Moriya, Tomohiro Tanaka, Takahiro Shinozaki, **Shinji Watanabe**, and Kevin Duh, "Evolution-Strategy-Based Automation of System Development for High-Performance Speech Recognition," IEEE Transactions on Audio, Speech and Language Processing, vol. 27, no. 1, pp. 77–88 (2019)
4. Chungwei Lin, Tim K. Marks, Milutin Pajovic, **Shinji Watanabe**, and Chih-kuan Tung, "Model parameter learning using KullbackLeibler divergence," Physica A, vol. 491, pp. 549-559 (2018)
5. Tsubasa Ochiai, **Shinji Watanabe**, Takaaki Hori, John R. Hershey, and Xiong Xiao, "A Unified Architecture for Multichannel End-to-End Speech Recognition with Neural Beamforming," IEEE Journal of Selected Topics in Signal Processing, vol. 11, no. 8, pp. 1274–1288 (2017)
6. **Shinji Watanabe**, Takaaki Hori, Suyoun Kim, John R. Hershey and Tomoki Hayashi, "Hybrid CTC/Attention Architecture for End-to-End Speech Recognition," IEEE Journal of Selected Topics in Signal Processing, vol. 11, no. 8, pp. 1240–1253 (2017)
7. Tomoki Hayashi, **Shinji Watanabe**, Tomoki Toda, Takaaki Hori, Jonathan Le Roux, Kazuya Takeda, "Duration-Controlled LSTM for Polyphonic Sound Event Detection," IEEE Transactions on Audio, Speech and Language Processing (accepted)
8. Takaaki Hori, Zhuo Chen, Hakan Erdogan, John Hershey, Jonathan Le Roux, Vikramjit Mitra, and **Shinji Watanabe**, Multi-Microphone Speech Recognition Integrating Beamforming, Robust Feature Extraction, and Advanced DNN/RNN Backend, " Computer Speech & Language, pp. 401–418 (2017)

9. Jon Barker, Ricard Marxer, Emmanuel Vincent, **Shinji Watanabe**, "The third CHIME speech separation and recognition challenge: Analysis and outcomes," Computer Speech & Language, pp. 605–626 (2017)

10. Emmanuel Vincent, **Shinji Watanabe**, Aditya Nugraha, Jon Barker, Ricard Marxer, "An analysis of environment, microphone and data simulation mismatches in robust speech recognition," Computer Speech & Language, pp. 535–557 (2017)

11. Naohiro Tawara, Tetsuji Ogawa, **Shinji Watanabe**, Atsushi Nakamura, and Tetsunori Kobayashi, "A sampling-based speaker clustering using utterance-oriented Dirichlet process mixture model and its evaluation on large scale data," APSIPA Transactions on Signal and Information Processing, vol. 4, e16 (2015).

12. Yuuki Tachioka, Tomohiro Narita, and **Shinji Watanabe**, "Effectiveness of dereverberation, feature transformation, discriminative training methods, and system combination approach for various reverberant environments," EURASIP Journal on Advances in Signal Processing, vol. 2015:52 (2015).

13. Masayuki Suzuki, Takuya Yoshioka, **Shinji Watanabe**, Nobuaki Minematsu, and Keikichi Hirose, "Feature Enhancement with Joint Use of Consecutive Corrupted and Noise Feature Vectors with Discriminative Region Weighting," IEEE Transactions on Audio, Speech & Language Processing, vol. 21, no. 10, pp. 2172–2181 (2013).

14. **Shinji Watanabe**, Atsushi Nakamura, and Biing-Hwang (Fred) Juang, "Structural Bayesian linear regression for hidden Markov models," Journal of Signal Processing Systems, pp. 1–18 (2013).

15. Seong-Jun Hahm, **Shinji Watanabe**, Atsunori Ogawa, Masakiyo Fujimoto, Takaaki Hori, and Atsushi Nakamura, "Prior-Shared Feature and Model Space Speaker Adaptation by Consistently Employing MAP Estimation," Speech Communication, vol. 55, Issue 3, pp. 415–431 (2013).

16. Tomoharu Iwata and **Shinji Watanabe**, "Influence Relation Estimation based on Lexical Entrainment in Conversation," Speech Communication, vol. 55, issue 2, pp 329-339 (2013)

17. Marc Delcroix, Keisuke Kinoshita, Tomohiro Nakatani, Shoko Araki, Atsunori Ogawa, Takaaki Hori, **Shinji Watanabe**, Masakiyo Fujimoto, Takuya Yoshioka, Takanobu Oba, Yotaro Kubo, Mehrez Souden, Seong-Jun Hahm, and Atsushi Nakamura, "Speech recognition in living rooms: Integrated speech enhancement and recognition system based on spatial, spectral and temporal modeling of sounds," Computer Speech & Language, vol. 27, Issue 3, pp. 851–873 (2013).

18. Marc Delcroix, **Shinji Watanabe**, Tomohiro Nakatani, and Atsushi Nakamura, "Cluster-based dynamic variance adaptation for interconnecting speech enhancement pre-processor and speech recognizer," Computer Speech & Language, vol. 27, Issue 1, pp. 350–368 (2013)

19. Yotaro Kubo, **Shinji Watanabe**, Takaaki Hori, and Atsushi Nakamura, "Structural Classification Methods based on Weighted Finite-State Transducers for Automatic Speech Recognition," IEEE Transactions on Audio, Speech & Language Processing, vol. 20, issue 8, pp. 2240–2251 (2012).

20. Daisuke Saito, **Shinji Watanabe**, Atsushi Nakamura, and Nobuaki Minematsu, "Statistical Voice Conversion Based on Noisy Channel Model, " IEEE Transactions on Audio, Speech & Language Processing, vol. 20, issue 6, pp. 1784–1794 (2012). **ASJ Itakura Award**

21. Takuya Maekawa and **Shinji Watanabe**, Training Data Selection with User's Physical Characteristics Data for Acceleration-based Activity Modeling, Personal and Ubiquitous Computing, vol. 17, issue 3, pp. 451–463 (2013).

22. Masakiyo Fujimoto, **Shinji Watanabe**, and Tomohiro Nakatani, "Frame-wise model re-estimation method based on Gaussian pruning with weight normalization for noise robust voice activity detection," Speech Communication, vol. 54(2), pp. 229–244 (2012)

23. Takaaki Hori, Shoko Araki, Takuya Yoshioka, Masakiyo Fujimoto, **Shinji Watanabe**, Takanobu Oba, Atsunori Ogawa, Kazuhiro Otsuka, Dan Mikami, Keisuke Kinoshita, Tomohiro Nakatani, Atsushi Nakamura, and Junji Yamato, "Low-latency Real-time Meeting Recognition and Understanding Using Distant Microphones and Omni-directional Camera," IEEE Transactions on Audio, Speech & Language Processing, vol. 20, issue 2, pp. 499–513 (2012).

24. Hideyuki Watanabe, Shigeru Katagiri, Kouta Yamada, Erik McDermott, Atsushi Nakamura, **Shinji Watanabe**, and Miho Ohsaki, "Minimum Classification Error Training Using Geometric-Margin-Based Misclassification Measure," (in Japanese) IEICE Transactions on Information and Systems, Vol.J94-D No.10 pp.1664–1675 (2011)

25. Hideyuki Watanabe, Shin'ichi TANIGUCHI, Shigeru Katagiri, Kouta Yamada, Atsushi Nakamura, Erik McDermott, **Shinji Watanabe**, and Miho Ohsaki, "Incremental Minimum Classification Error Training for Pattern Recognition," (in Japanese) IEICE Transactions on Information and Systems vol. J94-D, no. 4, pp. 702–711, (2011)

26. **Shinji Watanabe**, Tomoharu Iwata, Takaaki Hori, Atsushi Sako, and Yasuo Ariki, "Topic Tracking Language Model for Speech Recognition," Computer Speech & Language, vol. 25, issue 2, pp. 440–461, (2011)

27. Yotaro Kubo, **Shinji Watanabe**, Atsushi Nakamura, Erik McDermott, and Tetsunori Kobayashi, "A Sequential Pattern Classifier Based on Hidden Markov Kernel Machine and Its Application to Phoneme Recognition," IEEE Journal of Selected Topics in Signal Processing Volume 4, Issue 6, pages 974–984 (2010). **IEEE Signal Processing Society Japan Chapter Student Paper Award**

28. David Cournapeau, **Shinji Watanabe**, Atsushi Nakamura, and Tatsuya Kawahara, "Online Unsupervised Classification with Model Comparison in the Variational Bayes Framework for Voice Activity Detection," IEEE Journal of Selected Topics in Signal Processing, volume 4, issue 6, pp. 1071–1083 (2010)

29. Tomoharu Iwata, **Shinji Watanabe**, Takeshi Yamada and Naonori Ueda, "Topic Tracking Model for Purchase Behavior Analysis," (in Japanese) IEICE Transactions on Information and Systems, vol. J93-D, No. 6, pp. 978–987 (2010)

30. Kenta Nishiki, Yousuke Izumi, **Shinji Watanabe**, Takuya Nishimoto, Nobutaka Ono, And Shigeki Sagayama, "Stereo-input speech recognition using sparseness-based blind source separation," (in Japanese) IEICE Transactions on Information and Systems vol. J93-D, no. 3, pp. 303–311, (2010)

31. **Shinji Watanabe** and Atsushi Nakamura, "Predictor-Corrector Adaptation by using Time Evolution System with Macroscopic Time Scale," IEEE Transactions on Audio, Speech & Language Processing, vol. 18, issue 2, pp. 395–406 (2010)

32. Marc Delcroix, Tomohiro Nakatani, and **Shinji Watanabe**, "Static and dynamic variance compensation for recognition of reverberant speech with dereverberation pre-processing, " IEEE Transactions on Audio, Speech & Language Processing, vol. 17, issue 2, pp. 324–334, (2009)

33. **Shinji Watanabe** and Atsushi Nakamura, "Speech recognition based on Student's t-distribution derived from total Bayesian framework," IEICE Transactions on Information and Systems, vol.E89-D, no. 3, pp. 970–980, (2006)

34. **Shinji Watanabe**, Atsushi Sako and Atsushi Nakamura, "Automatic Determination of Acoustic Model Topology using Variational Bayesian Estimation and Clustering for Large Vocabulary Continuous Speech Recognition," IEEE Transactions on Speech and Audio Processing,, vol. 14, issue 3, pp. 855–872, (2006). **The TELECOM System Technology Award from the Telecommunications Advancement Foundation in 2006**

35. **Shinji Watanabe** and Atsushi Nakamura, "Acoustic Model Adaptation based on Coarse/Fine Training of Transfer Vectors," (in Japanese), Information Technology Letters (2004)

36. **Shinji Watanabe**, Yasuhiro Minami, Atsushi Nakamura, and Naonori Ueda, "Variational Bayesian Estimation and Clustering for Speech Recognition," IEEE Transactions on Speech and Audio Processing, vol. 12, pp. 365–381, (2004)

37. **Shinji Watanabe**, Yasuhiro Minami, Atsushi Nakamura, and Naonori Ueda, "Selection of Shared-State Hidden Markov Model Structure Using Bayesian Criterion,"IEICE D-II, vol. J86-D-II, no. 6, pp. 776–786, (2003), **the Best Paper Award from IEICE**. The English translation version is in IEICE Transactions on Information and Systems, vol.E88-D, no. 1, pp. 1–9, (2005))

38. Hisakazu Minakata and **Shinji Watanabe**, "Solar Neutrinos and Leptonic CP Violation," Phys. Lett. B, 468, p. 256, (1999).

PUBLICATIONS (CONFERENCE, REFREED)

1. Naoyuki Kanda, Shota Horiguchi, Yusuke Fujita, Yawen Xue, Kenji Nagamatsu, **Shinji Watanabe**, "SIMULTANEOUS SPEECH RECOGNITION AND SPEAKER DIARIZATION FOR MONAURAL DIALOGUE RECORDINGS WITH TARGET-SPEAKER ACOUSTIC MODELS," Proc. ASRU'19 (accepted)

2. Hirofumi Inaguma, Kevin Due, Tatsuya Kawahara, **Shinji Watanabe**, "MULTILINGUAL END-TO-END SPEECH TRANSLATION," Proc. ASRU'19 (accepted)

3. Shigeki Karita, Nanxin Chen, Tomoki Hayashi, Takaaki Hori, Hirofumi Inaguma, Ziyan Jiang, Masao Someki, Nelson Enrique Yalta Soplin, Ryuichi Yamamoto, Xiaofei Wang, **Shnji Watanabe**, Takenori Yoshimura, Wangyou Zhang, "A COMPARATIVE STUDY ON TRANSFORMER VS RNN IN SPEECH APPLICATIONS," Proc. ASRU'19 (accepted)

4. Emiru Tsunoo, Yosuke Kashiwagi, Toshiyuki Kumakura, **Shinji Watanabe**, "TRANSFORMER ASR WITH CONTEXTUAL BLOCK PROCESSING," Proc. ASRU'19 (accepted)

5. Xuankai Chang, Wangyou Zhang, Yanmin Qian, Jonathan Le Roux, **Shinji Watanabe**, "MIMO-SPEECH: End-to-End Multi-Channel Multi-Speaker Speech Recognition," Proc. ASRU'19 (accepted)

6. Yiming Wang, Tongfei Chen, Hainan Xu, Shuoyang Ding, Hang Lv, Yiwen Shao, Nanyun Peng, Lei Xie, **Shinji Watanabe**, Sanjeev Khudanpur, "ESPRESSO: A FAST END-TO-END NEURAL SPEECH RECOGNITION TOOLKIT," Proc. ASRU'19 (accepted)

7. Yusuke Fujita, Naoyuki Kanda, Shota Horiguchi, Yawen Xue, Kenji Nagamatsu, **Shinji Watanabe**, "END-TO-END NEURAL SPEAKER DIARIZATION WITH SELF-ATTENTION," Proc. ASRU'19 (accepted)

8. Matthew K Maciejewski, Gregory Sell, Yusuke Fujita, Leibny Paola Garcia Perera, **Shinji Watanabe**, and Sanjeev Khudanpur, "Analysis of Robustness of Deep Single-Channel Speech Separation Using Corpora Constructed from Multiple Domains," Proc. WASPAA'19 (accepted)

9. Toru Taniguchi, Aswin Shanmugam Subramanian, Xiaofei Wang, Dung Tran, Yuya Fujita, and **Shinji Watanabe**, "Generalized Weighted-Prediction-Error Dereverberation with Varying Source Priors for Reverberant Speech Recognition," Proc. WASPAA'19 (accepted)

10. Aswin Shanmugam Subramanian, Xiaofei Wang, Murali Karthick Baskar, **Shinji Watanabe**, Toru Taniguchi, Dung Tran, and Yuya Fujita, "Speech Enhancement Using End-to-End Speech Recognition Objectives," Proc. WASPAA'19 (accepted)

11. Hiroshi Seki, Takaaki Hori, **Shinji Watanabe**, Jonathan Le Roux and John Hershey, "End-to-End Multilingual Multi-Speaker Speech Recognition," Proc. Interspeech'19, pp. 3755–3759 (2019)

12. Matthew Wiesner, Adithya Renduchintala, **Shinji Watanabe**, Chunxi Liu, Najim Dehak and Sanjeev Khudanpur, "Pretraining by Backtranslation for End-to-end ASR in Low-Resource Settings," Proc. Interspeech'19, pp. 4375–4379 (2019)

13. Tomoki Hayashi, **Shinji Watanabe**, Tomoki Toda, Kazuya Takeda, Shubham Toshniwal and Karen Livescu, "Pre-trained Text Embeddings for Enhanced Text-to-Speech Synthesis," Proc. Interspeech'19, pp. 4430–4434 (2019)

14. Yusuke Fujita, Naoyuki Kanda, Shota Horiguchi, Kenji Nagamatsu and **Shinji Watanabe**, "End-to-End Neural Speaker Diarization with Permutation-Free Objectives," Proc. Interspeech'19, pp. 4300–4304 (2019)

15. Martin Karafiat, Murali Karthick Baskar, **Shinji Watanabe**, Takaaki Hori, Matthew Wiesner and Jan ernock, "Analysis of Multilingual Sequence-to-Sequence speech recognition systems," Proc. Interspeech'19, pp. 2220–2224 (2019)

16. Marc Delcroix, **Shinji Watanabe**, Tsubasa Ochiai, Keisuke Kinoshita, Shigeki Karita, Atsunori Ogawa and Tomohiro Nakatani, "End-to-end SpeakerBeam for single channel target speech recognition," Proc. Interspeech'19, pp. 451–455 (2019)

17. Murali Karthick Baskar, **Shinji Watanabe**, Ramn Astudillo, Takaaki Hori, Lukas Burget and Jan ernock, "Semi-supervised Sequence-to-sequence ASR using Unpaired Speech and Text," Proc. Interspeech'19, pp. 3790–3794 (2019)

18. Laureano Moro Velazquez, Jaejin Cho, **Shinji Watanabe**, Mark Hasegawa-Johnson, Odette Scharenborg, Kim Heejin and Najim Dehak, "Study of the performance of automatic speech recognition systems in speakers with Parkinson's Disease," Proc. Interspeech'19, pp. 3875–3879 (2019)

19. Hiroshi Seki, Takaaki Hori, **Shinji Watanabe**, Niko Moritz and Jonathan Le Roux, "Vectorized Beam Search for CTC-Attention-based Speech Recognition," Proc. Interspeech'19, pp. 3825–3829 (2019)

20. Daniel Garcia-Romero, David Snyder, **Shinji Watanabe**, Gregory Sell, Alan McCree, Dan Povey and Sanjeev Khudanpur, "Speaker recognition benchmark using the CHiME-5 corpus," Proc. Interspeech'19, pp. 1506–1510 (2019)

21. Shigeki Karita, Nelson Yalta, **Shinji Watanabe**, Marc Delcroix, Atsunori Ogawa and Tomohiro Nakatani, "Improving Transformer Based End-to-End Speech Recognition with Connectionist Temporal Classification and Language Model Integration," Proc. Interspeech'19, pp. 1408–1412 (2019)
22. Naoyuki Kanda, Shota Horiguchi, Ryoichi Takashima, Yusuke Fujita, Kenji Nagamatsu and **Shinji Watanabe**, "Interference Speaker Loss for Target-Speaker Speech Recognition," Proc. Interspeech'19, pp. 236–240 (2019)
23. Nelson Enrique Yalta Soplin, **Shinji Watanabe**, Takaaki Hori, Kazuhiro Nakadai, and Tetsuya Ogata, "CNN-based Multichannel End-to-End Speech Recognition for Everyday Home Environments," Proc. EUSIPCO'19 (accepted)
24. Ashish Arora, Chun-Chieh Chang, Babak Rekabdar, Daniel Povey, David Etter, Desh Raj, Hossein Hadian, Jan Trmal, Paola Garcia, **Shinji Watanabe**, Vimal Manohar, Yiwen Shao and Sanjeev Khudanpur, "Using ASR methods for OCR," Proc. ICDAR'19 (accepted)
25. Nelson Enrique Yalta Soplin, **Shinji Watanabe**, Kazuhiro Nakadai and Tetsuya Ogata, "Weakly-Supervised Deep Recurrent Neural Networks for Basic Dance Step Generation," Proc. IJCNN'19 (accepted)
26. Oliver Adams, Matthew Wiesner, **Shinji Watanabe** and David Yarowsky, "Massively Multilingual Adversarial Speech Recognition," Proc. NAACL-HLT'19, pp. 96–108 (2019)
27. Murali Karthick Baskar, Luk Burget, **Shinji Watanabe**, Martin Karafit, Takaaki Hori, Jan Honza ernock, "PROMISING ACCURATE PREFIX BOOSTING FOR SEQUENCE-TO-SEQUENCE ASR," Proc. ICASSP'19, pp. 5646–5650 (2019)
28. Hirofumi Inaguma, Jaejin Cho, Murali Karthick Baskar, Tatsuya Kawahara, **Shinji Watanabe**, "Transfer learning of language-independent end-to-end ASR with language model fusion," Proc. ICASSP'19, pp. 6096–6100 (2019)
29. Hainan Xu, Shuoyang Ding, **Shinji Watanabe**, "Improving End-to-end Speech Recognition with Pronunciation-assisted Sub-word Modeling," Proc. ICASSP'19, pp. 7110–7114 (2019)
30. Jaejin Cho, **Shinji Watanabe**, Takaaki Hori, Murali Karthick Baskar, Hirofumi Inaguma, Jesus Villalba, Najim Dehak, "LANGUAGE MODEL INTEGRATION BASED ON MEMORY CONTROL FOR SEQUENCE TO SEQUENCE SPEECH RECOGNITION," Proc. ICASSP'19, pp. 6191–6195 (2019)
31. Xiaofei Wang, Ruizhi Li, Sri Harish Mallidi, Takaaki Hori, **Shinji Watanabe**, Hynek Hermansky, "STREAM ATTENTION-BASED MULTI-ARRAY END-TO-END SPEECH RECOGNITION," Proc. ICASSP'19, pp. 7105–7109 (2019)
32. Vimal Manohar, Szu-Jui Chen, Zhiqi Wang, Yusuke Fujita, **Shinji Watanabe**, Sanjeev Khudanpur, "ACOUSTIC MODELING FOR OVERLAPPING SPEECH RECOGNITION: JHU CHIME-5 CHALLENGE SYSTEM," Proc. ICASSP'19, pp. 6665-6669 (2019)
33. Takaaki Hori, Ramfon Astudillo, Tomoki Hayashi, Yu Zhang, **Shinji Watanabe**, Jonathan Le Roux, "CYCLE-CONSISTENCY TRAINING FOR END-TO-END SPEECH RECOGNITION," Proc. ICASSP'19, pp. 6271–6275 (2019)
34. Sandeep Kothinti, Keisuke Imoto, Debmalya Chakrabarty, Gregory Sell, **Shinji Watanabe**, Mounya Elhilali, "JOINT ACOUSTIC AND CLASS INFERENCE FOR WEAKLY SUPERVISED SOUND EVENT DETECTION," Proc. ICASSP'19, pp. 36–40 (2019)
35. Jonathan Le Roux, Gordon Wichern, **Shinji Watanabe**, Andy Sarroff, John R. Hershey, "THE PHASEBOOK: BUILDING COMPLEX MASKS VIA DISCRETE REPRESENTATIONS FOR SOURCE SEPARATION," Proc. ICASSP'19, pp. 66–70 (2019)
36. Xuankai Chang, Yanmin Qian, Kai Yu, **Shinji Watanabe**, "End-to-End Monaural Multi-speaker ASR System without Pretraining," Proc. ICASSP'19, pp. 6256–6260 (2019)
37. Shigeki Karita, **Shinji Watanabe**, Tomoharu Iwata, Marc Delcroix, Atsunori Ogawa, Tomohiro Nakatani, "SEMI-SUPERVISED END-TO-END SPEECH RECOGNITION USING TEXT-TO-SPEECH AND AUTOENCODERS," Proc. ICASSP'19, pp. 6166–6170 (2019)
38. Naoyuki Kanda, Yusuke Fujita, Shota Horiguchi, Rintaro Ikeshita, Kenji Nagamatsu, **Shinji Watanabe**, "ACOUSTIC MODELING FOR DISTANT MULTI-TALKER SPEECH RECOGNITION WITH SINGLE- AND MULTI-CHANNEL BRANCHES," Proc. ICASSP'19, pp. 6630–6634 (2019)

39. Chunxi Liu, Matthew Wiesner, **Shinji Watanabe**, Craig Harman, Jan Trmal, Najim Dehak, and Sanjeev Khudanpu, "Low-resource contextual topic identification on speech," Proc. IEEE SLT'18, pp. 656–663 (2018)

40. Tomoki Hayashi, **Shinji Watanabe**, Yu Zhang, Tomoki Toda, Takaaki Hori, Ramon Astudillo, and Kazuya Takeda, "Back-translation-style data augmentation for end-to-end ASR" Proc. IEEE SLT'18, pp. 426–433 (2018)

41. Ruizhi Li, Jaejin Cho, Murali Karthick Baskar, Matthew Wiesner, Sri Harish Mallidi, Nelson Yalta, Martin Karafiat, **Shinji Watanabe**, and Takaaki Hori "'Multilingual sequence-to-sequence speech recognition: Architecture, transfer learning, and language modeling," Proc. IEEE SLT'18, pp. 521–527 (2018)

42. Takaaki Hori, Jaejin Cho, and **Shinji Watanabe**, "End-to-end speech recognition with word-based RNN language models," Proc. IEEE SLT'18, pp. 389–396 (2018)

43. Szu-Jui Chen, Aswin Shanmugam Subramanian, Hainan Xu and **Shinji Watanabe**, "Building state-of-the-art distant speech recognition using the CHiME-4 challenge with a setup of speech enhancement baseline," Proc. Interspeech'18, pp. 1571–1575 (2018)

44. **Shinji Watanabe**, Takaaki Hori, Shigeki Karita, Tomoki Hayashi, Jiro Nishitoba, Yuya Unno, Nelson Enrique Yalta Soplin, Jahn Heymann, Matthew Wiesner, Nanxin Chen, Adithya Renduchintala and Tsubasa Ochiai, "ESPnet: End-to-End Speech Processing Toolkit," Proc. Interspeech'18, pp. 2207–2211 (2018)

45. Tomoki Hayashi, **Shinji Watanabe**, Tomoki Toda and Kazuya Takeda, "Multi-Head Decoder for End-to-End Speech Recognition," Proc. Interspeech'18, pp. 801–805 (2018)

46. Shigeki Karita, **Shinji Watanabe**, Tomoharu Iwata, Atsunori Ogawa and Marc Delcroix, "Semi-Supervised End-to-End Speech Recognition," Proc. Interspeech'18, pp. 2–6 (2018)

47. Jon Barker, **Shinji Watanabe**, Emmanuel Vincent and Jan Trmal, "The fifth 'CHiME' Speech Separation and Recognition Challenge: Dataset, task and baselines," Proc. Interspeech'18, pp. 1561–1565 (2018)

48. Aswin Shanmugam Subramanian, Szu-Jui Chen and **Shinji Watanabe**, "Student-Teacher Learning for BLSTM Mask-based Speech Enhancement," Proc. Interspeech'18, pp. 3249–3253 (2018)

49. Adithya Renduchintala, Shuoyang Ding, Matthew Wiesner and **Shinji Watanabe**, "Multi-Modal Data Augmentation for End-to-end ASR," Proc. Interspeech'18 , pp. 2394-2398 (2018) best student paper award

50. Peter Frederiksen, Jesus Villalba, **Shinji Watanabe**, Zheng-Hua Tan and Najim Dehak, "Effectiveness of Single-Channel BLSTM Enhancement for Language Identification," Proc. Interspeech'18, pp. 1823–1827 (2018)

51. Gregory Sell, David Snyder, Alan McCree, Daniel Garcia-Romero, Jesus Villalba, Matthew Maciejewski, Vimal Manohar, Najim Dehak, Daniel Povey, **Shinji Watanabe** and Sanjeev Khudanpur, "Diarization is Hard: Some Experiences and Lessons Learned for the JHU Team in the Inaugural DIHARD Challenge," Proc. Interspeech'18, pp. 2808–2812 (2018)

52. Marc Delcroix, **Shinji Watanabe**, Atsunori Ogawa, Shigeki Karita and Tomohiro Nakatani, "Auxiliary feature based adaptation of end-to-end ASR systems," Proc. Interspeech'18, pp. 2444–2448 (2018)

53. Hiroshi Seki, Takaaki Hori, **Shinji Watanabe**, Jonathan Le Roux and John R Hershey, "A Purely End-to-End System for Multi-speaker Speech Recognition," Proc. ACL'18 (Volume 1: long paper), pp. 2620–2630 (2018)

54. Tsubasa Ochiai, **Shinji Watanabe**, Shigeru Katagiri, Takaaki Hori, and John Hershey, "Speaker adaptation for multichannel end-to-end speech recognition," Proc. ICASSP'18, pp. 6707–6711 (2018)

55. Hiroshi Seki, **Shinji Watanabe**, Takaaki Hori, Jonathan Le Roux, and John Hershey, "An end-to-end language-tracking speech recognizer for mixed-language speech," Proc. ICASSP'18, pp. 4919–4923 (2018)

56. Shane Settle, Jonathan Le Roux, Takaaki Hori, **Shinji Watanabe**, and John R. Hershey, "End-to-end multi-speaker speech recognition," Proc. ICASSP'18, pp. 4819–4823 (2018)

57. Hayato Shibata, Taku Kato, Takahiro Shinozaki, and **Shinji Watanabe**, "Composite embedding systems for ZEROSPEECH2017 track1," Proc. ASRU'17, pp. 747–751 (2017), student best paper candidate

58. Takaaki Hori, **Shinji Watanabe**, and John Hershey, "Multi-level language modeling and decoding for open vocabulary end-to-end speech recognition," Proc. ASRU'17, pp. 287–291 (2017), best paper candidate

59. **Shinji Watanabe**, Takaaki Hori, and John Hershey, "Language independent end-to-end architecture for joint language identification and speech recognition," Proc. ASRU'17, pp. 265–269 (2017), best paper candidate

60. Tsubasa Ochiai, **Shinji Watanabe**, and Shigeru Katagiri, "Does Speech Enhancement Work With End-To-End ASR Objectives?: Experimental Analysis Of Multichannel End-To-End ASR," Proc. MLSP'17 (2017)

61. Yuuki Tachioka, Tomohiro Narita, Iori Miura, Takanobu Uramoto, Natsuki Monta, Shingo Uenohara, Ken'ichi Furuya, **Shinji Watanabe** and Jonathan Le Roux, "Coupled initialization of multi-channel non-negative matrix factorization based on spatial and spectral information," Proc. Interspeech'17, pp. 2461–2465 (2017)

62. Takaaki Hori, **Shinji Watanabe**, Yu Zhang and William Chan, "Advances in Joint CTC-Attention based End-to-End Speech Recognition with a Deep CNN Encoder and RNN-LM," Proc. Interspeech'17, pp. 949–953 (2017)

63. Takahiro Shinozaki, **Shinji Watanabe**, Daichi Mochihashi and Graham Neubig, "Semi-Supervised Learning of a Pronunciation Dictionary from Disjoint Phonemic Transcripts and Text," Proc. Interspeech'17, pp. 2546–2550 (2017)

64. Tsubasa Ochiai, **Shinji Watanabe**, Takaaki Hori, and John R. Hershey, "Multichannel End-to-End Speech Recognition" Proc. ICML'17, pp. 2632–2641 (2017)

65. Takaaki Hori, **Shinji Watanabe**, and John R. Hershey, "Joint CTC/attention decoding for end-to-end speech recognition" Proc. ACL'17 long paper, pp. 518–529 (2017)

66. Suyoun Kim, Takaaki Hori, and **Shinji Watanabe**, "Joint CTC-attention based end-to-end speech recognition using multi-task learning," Proc. ICASSP'17, pp. 4835–4839 (2017)

67. Tomoki Hayashi, **Shinji Watanabe**, Tomoki Toda, Takaaki Hori, Jonathan Le Roux, Kazuya Takeda, "BLSTM-HMM Hybrid System Combined with Sound Activity Detection Network for Polyphonic Sound Event Detection" Proc. ICASSP'17, pp. 766–780 (2017)

68. **Shinji Watanabe**, Takaaki Hori, Jonathan Le Roux, and John Hershey, "Student-Teacher Network Learning with Enhanced Features," Proc. ICASSP'17, pp. 5275–5279 (2017)

69. Zhong Meng, **Shinji Watanabe**, John R. Hershey, and Hakan Erdogan, "Deep Long Short-Term Memory Adaptive Beamforming Networks for Multichannel Robust Speech Recognition," Proc. ICASSP'17, pp. 271–275 (2017)

70. Xiong Xiao, **Shinji Watanabe**, Eng Siong Chng, and Haizhou LI, "Beamforming Networks Using Spatial Covariance Features for Far-field Speech Recognition", APSIPA ASC'16 (2016)

71. Yuuki Tachioka, **Shinji Watanabe**, and Takaaki Hori, "The MELCO/MERL System Combination Approach for the Fourth CHiME Challenge," Proc. CHiME 2016 Workshop pp, 1–3 (2016)

72. Xiong Xiao, Chenglin Xu, Zhaofeng Zhang, Shengkui Zhao, Sining Sun, and **Shinji Watanabe**, "A Study of Learning Based Beamforming Methods for Speech Recognition," Proc. CHiME 2016 Workshop, pp. 26–31 (2016)

73. Tomohiro Tanaka, Takahiro Shinozaki, **Shinji Watanabe**, and Takaaki Hori, "Evolution Strategy Based Neural Network Optimization and LSTM Language Model for Robust Speech Recognition," Proc. CHiME 2016 Workshop, pp. 32–35 (2016)

74. Hakan Erdogan, Tomoki Hayashi, John R. Hershey, Takaaki Hori, Chiori Hori, Wei-Ning Hsu, Suyoun Kim, Jonathan Le Roux, Zhong Meng, and **Shinji Watanabe** "Multi-Channel Speech Recognition: LSTMs All the Way Through," Proc. CHiME 2016 Workshop, pp. 45–48 (2016)

75. Tomoki Hayashi, **Shinji Watanabe**, Tomoki Toda1, Takaaki Hori, Jonathan Le Roux, and Kazuya Takeda, "Bidirectional LSTM-HMM Hybrid System for Polyphonic Sound Event Detection," Proc. DCASE2016 Workshop (2016)

76. Takaaki Hori, Hai Wang, Chiori Hori, **Shinji Watanabe**, Bret A. Harsham, Jonathan Le Roux, John R. Hershey, Yusuke Koji, Yi Jing, Zhaocheng Zhu, and Takeyuki Aikawa, Dialog state tracking with attention-based sequence-to-sequence learning, Proc. SLT16, (accepted). **Ranked 2nd at DSTC5 challenge**.

77. Tomohiro Tanaka, Takafumi Moriya, Takahiro Shinozaki, **Shinji Watanabe**, Takaaki Hori, and Kevin Duh, Automated structure discovery and parameter tuning of neural network language model based on evolution strategy, Proc. SLT16, (accepted)

78. Toshiaki Koike-Akino, Ruhi Mahajan, Tim K. Marks, Oncel C. Tuzel, Ye Wang, **Shinji Watanabe**, Philip V. Orlik, High-Accuracy User Identification Using EEG Biometrics, Proc. EMBC (IEEE Engineering in Medicine and Biology Society)16, (2016)

79. Yusuf Ik, Jonathan Le Roux, Zhuo Chen, **Shinji Watanabe** and John Hershey, Single-channel multi-speaker separation using deep clustering, Interspeech16, (accepted)

80. Chiori Hori, Takaaki Hori, **Shinji Watanabe** and John Hershey, Context-sensitive and role-dependent spoken language understanding using bidirectional and attention LSTMs, Interspeech16, (accepted)

81. Katerina Zmolikova, Martin Karafiat, Karel Vesely, Marc Delcroix, **Shinji Watanabe**, Lukas Burget and Jan Cernocky, Data selection by sequence summarizing neural network in mismatch condition training, Interspeech16, (accepted)

82. Hakan Erdogan, John Hershey, **Shinji Watanabe**, Michael Mandel and Jonathan Le Roux, Improved MVDR beamforming using single-channel mask prediction networks, Interspeech16, (accepted)

83. Hori Chiori, **Shinji Watanabe**, Takaaki Hori, Bret A. Harsham, John R. Hershey, Yusuke Koji, Youichi Fujii, and Yuki Furumoto, Driver confusion status detection using recurrent neural networks, ICME16, (accepted)

84. Xiong Xiao, **Shinji Watanabe**, Hakan Erdogan, Liang Lu, John Hershey, Michael L. Seltzer, Guoguo Chen, Yu Zhang, Michael Mandel, and Dong Yu, Deep beamforming networks for multi-channel speech recognition, ICASSP16, (accepted)

85. John R. Hershey, Zhuo Chen, Jonathan Le Roux, and **Shinji Watanabe**, Deep clustering: discriminative embeddings for segmentation and separation, ICASSP16, (accepted)

86. Scott Wisdom, John R. Hershey, Jonathan Le Roux, **Shinji Watanabe**, Deep unfolding for multi-channel source separation, ICASSP16, (accepted)

87. Karel Vesely, **Shinji Watanabe**, Katerina Zmolikova, Martin Karafiat, Lukas Burget, Jan Cernocky, Sequence summarizing neural network for speaker adaptation, ICASSP16, (accepted)

88. Takaaki Hori, Chiori Hori, **Shinji Watanabe**, John Hershey, Minimum word error training of long short-term memory recurrent neural network language models for speech recognition, ICASSP16, (accepted)

89. Takaaki Hori, Zhuo Chen, Hakan Erdogan, John Hershey, Jonathan Le Roux, Vikramjit Mitra, and **Shinji Watanabe**, The MERL/SRI system for the 3rd CHiME challenge using beamforming, robust feature extraction, and advanced speech recognition, ASRU15, (accepted). **ranked 2nd at 3rd CHiME Challenge among 26 submissions**

90. Takafumi Moriya, Tomohiro Tanaka, Takahiro Shinozaki, **Shinji Watanabe**, and Kevin Duh Automation of system building for state-of-the-art large vocabulary speech recognition using evolution strategy, Proc. ASRU15, (accepted).

91. Jon Barker, Ricard Marxer, Emmanuel Vincent, and **Shinji Watanabe**, The third 'CHiME' Speech Separation and Recognition Challenge: Dataset, task and baselines, ASRU15, (accepted).

92. Roger Hsiao, Jeff Ma, William Hartmann, Martin Karafiat, Frantisek Grezl, Lukas Burget, Igor Szoke, Jan Honza Cernocky, **Shinji Watanabe**, Zhuo Chen, Sri Harish Mallidi, Hynek Hermansky, Stavros Tsakalidis, and Richard Schwartz, Robust speech recognition in unknown reverberant and noisy conditions, Proc. ASRU15, (accepted).

93. Hiroki Kanagawa, Yuuki Tachioka, **Shinji Watanabe**, and Jun Ishii, Feature-space structural MAPLR with regression tree-based multiple transformation matrices for DNN, APSIPA ASC15, (accepted).

94. Chiori Hori, Takaaki Hori, **Shinji Watanabe**, and John R. Hershey, "Context Sensitive Spoken Language Understanding using Role Dependent LSTM layers", NIPS Workshop for Machine Learning for SLU & Interaction (2015)

95. Bret A. Harsham, **Shinji Watanabe**, Alan Esenther, John R. Hershey, Jonathan Le Roux, Yi Luan, Daniel N. Nikovski, and Vamsi K. Potluru, Driver prediction to improve interaction with in-vehicle HMI, Proc. DSP for In-Vehicle Workshop15, (accepted).

96. Yi Luan, **Shinji Watanabe**, and Bret Harsham, Efficient learning for spoken language understanding tasks with word embedding based pre-training, Interspeech15, pp. 1398–1402 (2015).

97. Zhuo Chen, **Shinji Watanabe**, Hakan Erdogan, and John R. Hershey, Integration of Speech Enhancement and Recognition using Long-Short Term Memory Recurrent Neural Network, Proc. Interspeech15, pp. 3274–3278 (2015).

98. Ahmed Hussen Abdelaziz, **Shinji Watanabe**, John R. Hershey, Emmanuel Vincent, and Dorothea Kolossa, Uncertainty Propagation through Deep Neural Networks, Proc. Interspeech15, pp. 3561–3565 (2015).

99. Yuuki Tachioka, **Shinji Watanabe**, Uncertainty training and decoding methods of deep neural networks based on stochastic representation of enhanced features, Proc. Interspeech15, pp. 3541–3545 (2015).

100. Yuuki Tachioka, **Shinji Watanabe**, A Discriminative Method for Recurrent Neural Network Language Models, Proc. ICASSP15, pp. 5386–5390 (2015).

101. Takahiro Shinozaki, **Shinji Watanabe**, Structure Discovery of Deep Neural Network Based on Evolutionary Algorithms, Proc. ICASSP15, pp. 4979–4983 (2015).

102. Erdogan Hakan, John R. Hershey, **Shinji Watanabe**, Jonathan Le Roux, Phase-sensitive and recognition-boosted speech separation using deep recurrent neural networks, Proc. ICASSP15, pp. 708–712 (2015).

103. Yuuki Tachioka, **Shinji Watanabe**, Jonathan Le Roux, and John Hershey, "Sequence Discriminative Training for Low-Rank Deep Neural Networks," Proc. GlobalSIP14-Machine Learning Applications in Speech Processing, pp. 735–739 (2014).

104. Felix Weninger, Jonathan Le Roux, John Hershey, and **Shinji Watanabe**, "Discriminative NMF and its Application to Single-Channel Source Separation," Proc. Interspeech'14, pp. 865–869 (2014)

105. **Shinji Watanabe**, John Hershey, Tim Marks, Youichi Fujii, and Yusuke Koji, "Cost-Level Integration of Statistical and Rule-Based Dialog Managers," Proc. Interspeech'14, pp. 323–327 (2014)

106. Yuuki Tachioka, **Shinji Watanabe**, Jonathan Le Roux, and John Hershey, "Sequential Maximum Mutual Information Linear Discriminant Analysis for Speech Recognition," Proc. Interspeech'14, pp. 2415–2419 (2014)

107. Yuuki Tachioka, Tomohiro Narita, **Shinji Watanabe**, Jonathan Le Roux, "Ensemble Integration of Calibrated Speaker Localization and Statistical Speech Detection in Domestic Environments," Proc. HSCMA (2014)

108. Yuuki Tachioka, Tomohiro Narita, Felix J Weninger, and **Shinji Watanabe**, "Dual System Combination Approach for Various Reverberant Environments with Dereverberation Techniques," Proc. REVERB Workshop (2014) **ranked 1st place at REVERB Challenge Official Data Track (2nd overall)**

109. Felix J Weninger, **Shinji Watanabe**, Jonathan Le Roux, John Hershey, Yuuki Tachioka, Jrgen T. Geiger, Bjrn W Schuller, and Gerhard Rigoll, "The MERL/MELCO/TUM system for the REVERB Challenge using Deep Recurrent Neural Network Feature Enhancement," Proc. REVERB Workshop (2014)

110. Hao Tang, **Shinji Watanabe**, Tim K. Marks, and John Hershey, "Log-Linear Dialog Manager," Proc. ICASSP'14, pp. 4120 – 4123 (2014)

111. **Shinji Watanabe**, Jonathan Le Roux, "Black Box Optimization for Automatic Speech Recognition," Proc. ICASSP'14, pp. 3280 – 3283 (2014)

112. Chao Weng, Dong Yu, **Shinji Watanabe**, and Biing-Hwang (Fred) Juang, "Recurrent Deep Neural Networks for Robust Speech Recognition," Proc. ICASSP'14, pp. 5569 – 5572 (2014)

113. Felix J Weninger, **Shinji Watanabe**, Yuuki Tachioka, and Bjrn W Schuller, "Deep Recurrent De-Noising Auto-Encoder and Blind De-Reverberation for Reverberated Speech Recognition," Proc. ICASSP'14, pp. 4656 – 4659 (2014)

114. Yuuki Tachioka, **Shinji Watanabe**, Jonathan Le Roux, and John Hershey, A Generalized Discriminative Training Framework for System Combination, Proc. ASRU13, pp. 43–48 (2013)

115. Emmanuel Vincent, Jon Barker, **Shinji Watanabe**, Jonathan Le Roux, Francesco Nesta, and Marco Matassoni, The Second 'CHiME' Speech Separation and Recognition Challenge: An Overview of Challenge Systems and Outcomes, Proc. ASRU13, pp. 162 – 167 (2013)

116. Jonathon Le Roux, **Shinji Watanabe**, and John R. Hershey, Ensemble Learning for Speech Enhancement, Proc. WASPAA13 (2013)

117. Koichiro Yoshino, **Shinji Watanabe**, Jonathon Le Roux, and John R. Hershey, Statistical dialogue management using intention dependency graph, Proc. IJCNLP13, pp. 962–966 (2013)

118. Naohiro Tawara, Tetsuji Ogawa, **Shinji Watanabe**, Atsushi Nakamura, and Tetsunori Kobayashi, Blocked Gibbs sampling based multi-scale mixture model for speaker clustering on noisy data, Proc. MLSP13 (accepted)

119. Yuuki Tachioka and **Shinji Watanabe**, Discriminative training of acoustic models for system combination, Proc. Interspeech13, pp. 2355–2359 (2013)

120. **Shinji Watanabe** and John R. Hershey, Stereo based feature enhancement based on dictionary learning, Proc. ICASSP13, pp. 7073–7077 (2013)

121. Yuuki Tachioka, **Shinji Watanabe**, and John Hershey, Effectiveness of discriminative training and feature transformation for reverberated and noisy speech, Proc. ICASSP13, pp. 6935–6939 (2013)

122. Emmanuel Vincent, Jon Barker, **Shinji Watanabe**, Jonathan Le Roux, Francesco Nesta, and Marco Matassoni, The second CHiME speech separation and recognition challenge: datasets, tasks and baselines, Proc. ICASSP13, pp. 126–130 (2013)

123. Yuuki Tachioka, **Shinji Watanabe**, Jonathan Le Roux, and John R. Hershey, Discriminative methods for noise robust speech recognition: A CHiME Challenge Benchmark, Proc. CHiME13, pp. 19–24, (2013) **ranked 1st place at 2nd CHiME Challenge Track 2**

124. Naohiro Tawara, Tetsuji Ogawa, **Shinji Watanabe**, Atsushi Nakamura, Tetsunori Kobayashi, Fully Bayesian speaker clustering based on hierarchically structured utterance-oriented Dirichlet process mixture model, Proc. Interspeech12, pp. 2166–2169 (2012)

125. Seong-Jun Hahm, **Shinji Watanabe**, Masakiyo Fujimoto, Atsunori Ogawa, Takaaki Hori, and Atsushi Nakamura, Normalization and Adaptation by Consistently Employing MAP Estimation, in: Proc. of International Workshop on Statistical Machine Learning for Speech Processing (IWSML) (2012).

126. **Shinji Watanabe**, Yotaro Kubo, Takanobu Oba, Takaaki Hori, Atsushi Nakamura, Bag of arcs: new representation of speech segment features based on finite state machines, Proc. ICASSP12, pp. 4201–4204 (2012).

127. Roland Roller, **Shinji Watanabe**, Tomoharu Iwata, Effect of dialog acts on word use in polylogue, Proc. ICASSP12, pp. 4969–4972 (2012).

128. Masayuki Suzuki, Takuya Yoshioka, **Shinji Watanabe**, Nobuaki Minematsu, Keikichi Hirose, MFCC enhancement using joint corrupted and noise feature space for highly non-stationary noise environments, Proc. ICASSP12, pp. 4109–4112 (2012). **ICASSP 2012 Student Paper Award**

129. Ekapol Chuangsuwanich, **Shinji Watanabe**, Hori Takaaki, Tomoharu Iwata, James Glass, Handling uncertain observations in unsupervised topic-mixture language model adaptation, Proc. ICASSP12, pp. 5033–5036 (2012).

130. Naohiro Tawara, Tetsuji Ogawa, **Shinji Watanabe**, Tetsunori Kobayashi, Fully bayesian inference of multi-mixture gaussian model and its evaluation using speaker clustering, Proc. ICASSP12, pp. 5253–5256 (2012).

131. Marc Delcroix, Atsunori Ogawa, **Shinji Watanabe**, Tomohiro Nakatani, Atsushi Nakamura, Discriminative feature transforms using differenced maximum mutual information, Proc. ICASSP12, pp. 4753–4756 (2012).

132. Yotaro Kubo, **Shinji Watanabe**, Atsushi Nakamura, Simon Wiesler, Ralf Schlueter, Hermann Ney, Basis vector orthogonalization for an improved kernel gradient matching pursuit method, Proc. ICASSP12, pp. 4753–4756 (2012).

133. Yotaro Kubo, **Shinji Watanabe**, Atsushi Nakamura, Decoding network optimization using minimum transition error training, Proc. ICASSP12, pp. 4197–4200 (2012).

134. Masakiyo Fujimoto, **Shinji Watanabe**, Tomohiro Nakatani, Noise suppression with unsupervised joint speaker adaptation and noise mixture model estimation, Proc. ICASSP12, pp. 4713 - 4716 (2012).

135. **Shinji Watanabe**, Atsushi Nakamura, and Biing-Hwang Juang, "Bayesian Linear Regression for Hidden Markov Model Based on Optimizing Variational Bounds," Proc. MLSP'11, pp. 1–6 (2011).

136. Marc Delcroix, Keisuke Kinoshita, Tomohiro Nakatani, Shoko Araki, Atsunori Ogawa, Takaaki Hori, **Shinji Watanabe**, Masakiyo Fujimoto, Takuya Yoshioka, Takanobu Oba, Yotaro Kubo, Mehrez Souden, Seong-Jun Hahm, and Atsushi Nakamura, "Speech recognition in the presence of highly non-stationary noise based on spatial, spectral and temporal speech/noise modeling combined with dynamic variance adaptation, " Proc. CHiME'11, pp. 1217 (2011)

137. **Shinji Watanabe**, Atsushi Nakamura, and Biing-Hwang Juang, "Model Adaptation for Automatic Speech Recognition Based on Multiple Time Scale Evolution," Proc. Interspeech'11, pp. 10811084 (2011)
138. Masakiyo Fujimoto, **Shinji Watanabe**, and Tomohiro Nakatani, "A Robust Estimation Method of Noise Mixture Model for Noise Suppression," Proc. Interspeech'11, pp. 697700 (2011)
139. Tomoharu Iwata and **Shinji Watanabe**, "Learning Influences from Word Use in Polylogue," Proc. Interspeech'11, pp. 3089-3092 (2011)
140. Naohiro Tawara, **Shinji Watanabe**, Tetsuji Ogawa, and Tetsunori Kobayashi, "Speaker Clustering Based on Utterance-oriented Dirichlet Process Mixture Model," Proc. Interspeech'11, pp. 2905-2908 (2011)
141. Tomoharu Iwata, **Shinji Watanabe** and Hiroshi Sawada, Fashion Coordinates Recommender System using Photographs from Fashion Magazines, Proc. IJCAI11, pp. 2262–2267 (2011)
142. Marc Delcroix, **Shinji Watanabe**, Tomohiro Nakatani, and Atsushi Nakamura, Discriminative approach to dynamic variance adaptation for noisy speech recognition, Proc. HSCMA11, pp. 7–12 (2011)
143. Shoko Araki, Takaaki Hori, Takuya Yoshioka, Masakiyo Fujimoto, **Shinji Watanabe**, Takanobu Oba, Atsunori Ogawa, Kazuhiro Otsuka, Dan Mikami, Keisuke Kinoshita, Tomohiro Nakatani, Atsushi Nakamura, Junji Yamato, Low-latency meeting recognition and understanding using distant microphones, Proc. HSCMA11, pp. 151–152 (2011)
144. Takuya Maekawa, **Shinji Watanabe**, Unsupervised Activity Recognition with User's Physical Characteristics Data, Proc. ISWC11, pp. 89–96 (2011). **Best In-Category Nominee**
145. **Shinji Watanabe**, Daichi Mochihashi, Takaaki Hori, and Atsushi Nakamura, Gibbs Sampling Based Multi-Scale Mixture Model for Speaker Clustering, Proc. ICASSP11, pp. 4524–4527 (2011).
146. Masakiyo Fujimoto, **Shinji Watanabe**, and Tomohiro Nakatani, Non-Stationary Noise Estimation Method Based on Bias-Residual Component Decomposition for Robust Speech Recognition, Proc. ICASSP11, pp. 4816 - 4819 (2011).
147. Daisuke Saito, **Shinji Watanabe**, Atsushi Nakamura, and Nobuaki Minematsu, High Accurate Model-Integration-Based Voice Conversion Using Dynamic Features and Model Structure Optimization, Proc. ICASSP11, pp. 4576 - 4579 (2011).
148. Yotaro Kubo, Simon Wiesler, Ralf Schlueter, Hermann Ney, **Shinji Watanabe**, Atsushi Nakamura, and Tetsunori Kobayashi, Subspace Pursuit Method for Kernel-Log-Linear Models, Proc. ICASSP11, pp. 4500 - 4503 (2011).
149. **Shinji Watanabe**, Tomoharu Iwata, Takaaki Hori, Atsushi Sako, and Yasuo Ariki, "Application of Topic Tracking Model to Language Model Adaptation and Meeting Analysis," Proc. IEEE Workshop on Spoken Language Technology (SLT'10), pp. 366–371 (2010)
150. Takaaki Hori, Shoko Araki, Takuya Yoshioka, Masakiyo Fujimoto, **Shinji Watanabe**, Takanobu Oba, Atsunori Ogawa, Kazuhiro Otsuka, Dan Mikami, Keisuke Kinoshita, Tomohiro Nakatani, Atsushi Nakamura, Junji Yamato, "Real-time Meeting Recognition and Understanding Using Distant Microphones and Omni-directional Camera," Proc. IEEE Workshop on Spoken Language Technology (SLT'10), pp. 412417 (2010)
151. **Shinji Watanabe**, Takaaki Hori, and Atsushi Nakamura, "Large Vocabulary Continuous Speech Recognition Using WFST-based Linear Classifier for Structured Data," Proc. Interspeech'10, pp. 346–349, (2010)
152. Masakiyo Fujimoto, **Shinji Watanabe**, and Tomohiro Nakatani, "Voice Activity Detection Using Frame-Wise Model Re-Estimation Method Based on Gaussian Pruning with Weight Normalization," Proc. Interspeech'10, pp. 3102–3105, (2010)
153. Yotaro Kubo, **Shinji Watanabe**, Atsushi Nakamura, and Tetsunori Kobayashi, "A Regularized Discriminative Training Method of Acoustic Models Derived by Minimum Relative Entropy Discrimination," Proc. Interspeech'10, pp. 2954–2957, (2010)
154. Daisuke Saito, **Shinji Watanabe**, Atsushi Nakamura, and Nobuaki Minematsu, "Probabilistic Integration of Joint Density Model and Speaker Model for Voice Conversion," Proc. Interspeech'10, pp. 1728–1731, (2010)

155. Takaaki Hori, **Shinji Watanabe**, and Atsushi Nakamura, "Improvements of Search Error Risk Minimization in Viterbi Beam Search for Speech Recognition," Proc. Interspeech'10, pp. 1962–1965, (2010)

156. **Shinji Watanabe**, Takaaki Hori, Erik McDermott, and Atsushi Nakamura, "A discriminative model for continuous speech recognition based on weighted finite state transducers," Proc. ICASSP'10, pp. 4922–4925, (2010)

157. David Cournapeau, **Shinji Watanabe**, and Atsushi Nakamura, Tatsuya Kawahara, "Using online model comparison in the variational Bayes framework for online unsupervised voice activity detection," Proc. ICASSP'10, pp. 4462–4465, (2010)

158. Takaaki Hori, **Shinji Watanabe**, and Atsushi Nakamura, "Search error risk minimization in Viterbi beam search for speech recognition," Proc. ICASSP'10, pp. 4934–4937, (2010)

159. Kazuo Aoyama, **Shinji Watanabe**, Hiroshi Sawada, Yasuhiro Minami, Naonori Ueda, and Kazumi Saito, "Fast similarity search on a large speech data set with neighborhood graph indexing," Proc. ICASSP'10, pp. 5358–5361, (2010)

160. Erik McDermott, **Shinji Watanabe**, and Atsushi Nakamura, "Discriminative training based on an integrated view of MPE and MMI in margin and error space," Proc. ICASSP'10, pp. 4894–4897, (2010)

161. Hideyuki Watanabe, Shigeru Katagiri, Kouta Yamada, Erik McDermott, Atsushi Nakamura, **Shinji Watanabe**, and Miho Ohsaki, "Minimum error classification with geometric margin control," Proc. ICASSP'10, pp. 49224925, (2010)

162. Erik McDermott, **Shinji Watanabe**, and Atsushi Nakamura, "Margin-Space Integration of MPE Loss via Differencing of MMI Functionals for Generalized Error-Weighted Discriminative Training," Proc. Interspeech'09, pp. 224–227, (2009)

163. Yosuke Izumi, Kenta Nishiki, **Shinji Watanabe**, Takuya Nishimoto, Nobutaka Ono, and Shigeki Sagayama, "Stereo-input Speech Recognition using Sparseness-based Time-frequency Masking in a Reverberant Environment," Proc. Interspeech'09, pp. 1955–1958, (2009)

164. Tomoharu. Iwata, **Shinji Watanabe**, Takeshi Yamada and Naonori Ueda, "Topic tracking model for analyzing consumer purchase behavior," Proc. IJCAI'09, pp. 1427–1432, (2009)

165. Atushi Nakamura, Erik McDermott, **Shinji Watanabe**, and Shigeru Katagiri, "A unified view for discriminative objective functions based on negative exponential of difference measure between strings, " Proc. ICASSP'09, pp. 1633-1636, (2009)

166. **Shinji Watanabe** and Atsushi Nakamura, "Speech recognition with incremental tracking and detection of changing environments based on a macroscopic time evolution system, " Proc. ICASSP'09, pp. 4373-4376, (2009)

167. Marc Delcroix, Tomohiro Nakatani, and **Shinji Watanabe**, "Combined static and dynamic variance adaptation for efficient interconnection of speech enhancement pre-processor with speech recognizer," Proc. ICASSP'08, pp. 4073–4076, (2008)

168. **Shinji Watanabe** and Atsushi Nakamura, "A unified interpretation of adaptation approaches based on a macroscopic time evolution system and indirect/direct adaptation approaches," Proc. ICASSP'08, pp. 4285–4288, (2008)

169. **Shinji Watanabe** and Atsushi Nakamura, "Incremental adaptation based on a macroscopic time evolution system," Proc. ICASSP'07, vol. 4, pp. 769–772, (2007)

170. **Shinji Watanabe** and Atsushi Nakamura, "Acoustic model adaptation based on coarse/fine training of transfer vectors using directional statistics," Proc. ICASSP'06, vol. 1, pp. 1005–1008, (2006)

171. **Shinji Watanabe** and Atsushi Nakamura, "Effects of Bayesian predictive classification using variational Bayesian posteriors for sparse training data in speech recognition," Proc. Interspeech'05, pp. 1105–1109, (2005)

172. **Shinji Watanabe** and Atsushi Nakamura, "Robustness of acoustic model topology determined by VBEC (Variational Bayesian Estimation and Clustering for speech recognition) for different speech data sets," Proc. Workshop on statistical modeling approach for speech recognition - Beyond HMM, pp. 55–60, (2004)

173. **Shinji Watanabe** and Atsushi Nakamura, "Acoustic model adaptation based on coarse-fine training of transfer vectors and its application to speaker adaptation task," Proc. ICSLP'04, vol. 4, pp. 2933–2936, (2004)

174. Parham Zolfaghari, **Shinji Watanabe**, Atsushi Nakamura and Shigeru Katagiri, "Bayesian Modelling of the Speech Spectrum Using Mixture of Gaussians," Proc. ICASSP'04, vol. 1, pp. 553–556, (2004)
175. **Shinji Watanabe**, Atsushi Sako and Atsushi Nakamura, "Automatic Determination of Acoustic Model Topology using Variational Bayesian Estimation and Clustering," Proc. ICASSP'04, vol. 1, pp. 813–816, (2004)
176. Parham Zolfaghari, Hiroko Kato, **Shinji Watanabe** and Shigeru Katagiri, "Speech Spectral Modelling using Mixture of Gaussians, " Proc. SWIM , (2004)
177. **Shinji Watanabe**, Yasuhiro Minami, Atsushi Nakamura, and Naonori Ueda, "Bayesian Acoustic Modeling for Spontaneous Speech Recognition," Proc. SSPR'03, pp. 47–50, (2003)
178. **Shinji Watanabe**, Yasuhiro Minami, Atsushi Nakamura, and Naonori Ueda, "Application of Variational Bayesian Estimation and Clustering to Acoustic Model Adaptation," Proc. ICASSP'03, vol. 1, pp. 568–571, (2003)
179. **Shinji Watanabe**, Yasuhiro Minami, Atsushi Nakamura, and Naonori Ueda, "Application of Variational Bayesian Approach to Speech Recognition," NIPS15 MIT Press, (2002)
180. **Shinji Watanabe**, Yasuhiro Minami, Atsushi Nakamura, and Naonori Ueda, "Constructing Shared-State Hidden Markov Models Based on a Bayesian Approach," Proc. ICSLP'02, vol. 4, pp. 2669–2672, (2002).

ACADEMIC ACTIVITY

- Membership
    1. Senior Member of the Institute of Electrical and Electronics Engineers (IEEE)
    2. Member of the International Speech Communication Association (ISCA)
    3. Member of the Acoustical Society of Japan (ASJ)
    4. Member of the Institute of Electronics, Information, and Communication Engineers of Japan (IEICE)

- Organizer/Committee Member
    1. Area Chair, the Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20) (2019.8 – 2020.2)
    2. APSIPA Distinguished Lecturers (2019-2020)
    3. IEEE Signal Processing Society Liaison for 2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU) (2018.9 - 2019.12)
    4. Organizing Committee, the 2018 SANE 2018 - Speech and Audio in the Northeast (2018.5-2018.10)
    5. Member, the IEEE Signal Processing Society Machine Learning for Signal Processing Technical Committee (IEEE MLSP) (2018.1 - 2020.12)
    6. Area Chair "Vision, Robotics, Multimodal, Grounding and Speech" in ACL 2018
    7. Special session co-organizer of "Multi-Microphone Speech Recognition" accepted at ICASSP 2018.
    8. Program Chair of 2017 IEEE International Workshop on Machine Learning for Signal Processing (MLSP'17) (2016.9 - 2017.9)
    9. Technical Chair of IEEE ASRU 2017 (2016.5–2017.12)
    10. Organizing Committee of The 4th International Workshop on Speech Processing in Everyday Environments (CHiME 2016), (2016.5 - 2016.9)
    11. Senior Affiliates of "Building Speech Recognition System from Untranscribed Data" at 2016 Jelinek Summer Workshop on Speech and Language Technology (2015.12–2016.8)
    12. Chair of Industrial Membership Committee, APSIPA (2016-)
    13. Organizing Committee of The 4th CHiME Speech Separation and Recognition Challenge (2016.4-2016.9)
    14. Member of the IEEE CIS Task Force on Computational Audio Processing (2015.12-)
    15. Interactive Peer Reviewer at the 2016 Frederick Jelinek Memorial Summer Workshop on Speech and Language Technology (JSALT 2016), (2015.11)
    16. Senior Member of "Far-Field Speech Enhancement and Recognition in Mismatched Settings" at 2015 Jelinek Summer Workshop on Speech and Language Technology (2014.12–2015.8)

17. Special session co-organizer of "Robust Speech Processing using Observation Uncertainty and Uncertainty Propagation" at Interspeech (2015)
18. Organizing Committee of The 3rd CHiME Speech Separation and Recognition Challenge at IEEE ASRU 2015 (2014.7-2015.12)
19. Sponsorship Chair of IEEE ASRU 2015 (2014.6-2015.12)
20. TPC member of Machine Learning Applications in Speech Processing Symposium in the 2nd IEEE Global Conference on Signal and Information Processing (Global SIP) (2013.12 - 2014.12)
21. Industrial Membership Committee, APSIPA (2014.1 - 2017.12)
22. Member of the APSIPA Speech, Language, and Audio Technical Committee (APSIPA SLA TC) (2014.1 - 2019.12)
23. Member of the IEEE Signal Processing Society Speech and Language Technical Committee (IEEE SLTC) (2014.1 - 2019.12)
24. Organizing Committee of The 2012 SANE 2012 - Speech and Audio in the Northeast (2012.7-2012.10)
25. Organizing Committee of The 2012 IEEE AASP 'CHiME' SpeechSeparation and Recognition Challenge (2012.3-2013.3)
26. Publication Chair of the 2012 International Workshop on Statistical Machine Learning for Speech Processing (IWSML2012)
27. Executive Committee Member of the 2nd Young Researchers Forum on ALAGIN Speech Processing Session (2011.2 - 2011.3)
28. Panelist in IPSJ/2009-SLP-79, Tokyo, Japan (2009.12)
29. Program Committee Member of the 12th Workshop on Information-Based Induction Sciences (IBIS 2009) (2008.12 - 2009.11)
30. Review Committee Member of the Acoustical Society of Japan (2005.9-)

- Editor
    1. Subject Editor, Elsevier Speech Communication (2019.4 – 2021.4)
    2. Lead Guest Editor, IEEE JSTSP Special issue on Far-Field Speech Processing in the Era of Deep Learning (2018.5 - 2019.8)
    3. Associate Editor, APSIPA Transactions on Signal and Information Processing (2018.4 –)
    4. Associate Editor of the IEICE Transactions on Information and Systems (2016.7–2020)
    5. Guest Editor of Computer Speech and Language (2015.12–2016.11)
    6. Guest Associate Editor of IEICIE Transactions on Information and Systems (2013.4-2014.6)
    7. Associate Editor of the IEEE Transaction on Audio, Speech, and Language Processing (ASLP). (2012.4-2015.4)

- Session Chair
    1. IEEE MLSP (2017)
    2. ISCA Interspeech (2015, 2016, 2018, 2019)
    3. IEEE ICASSP (2014, 2015, 2016, 2017, 2018, 2019)
    4. IEEE ASRU (2015, 2017)
    5. Acoustical Society of Japan Meeting (2004.3-2010)

REVIEWER

- Conference Review
    1. Asian Conference on Machine Learning (ACML)
    2. Association for Computational Linguistics (ACL)
    3. DCASE workshop 2018
    4. Empirical Methods in Natural Language Processing (EMNLP)
    5. European Signal Processing Conference (EUSIPCO)
    6. IEEE Global Conference on Signal and Information Processing (GlobalSIP)
    7. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)
    8. IEEE International Workshop on Machine Learning for Signal Processing (MLSP)

9. IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)
10. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
11. International Conference on Learning Representations (ICLR)
12. International Conference on Machine Learning (ICML)
13. International Conference on Pattern Recognition (ICPR)
14. International Conference on Artificial Neural Networks (ICANN)
15. International Joint Conferences on Artificial Intelligence (IJCAI)
16. International Joint Conference on Neural Networks (IJCNN)
17. International Workshop on Acoustic Signal Enhancement (IWAENC)
18. International Workshop on Statistical Machine Learning for Speech Processing (IWSML)
19. ISCA INTERSPEECH
20. Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA)
21. Neural Information Processing Systems (NeurIPS)
22. REVERB workshop

- Journal/Transaction Review
    1. APSIPA Transactions on Signal and Information Processing
    2. Computer Speech and Language
    3. EURASIP Journal on Advances in Signal Processing
    4. Information Processing Society of Japan (IPSJ) Transactions
    5. IEEE Journal of Selected Topics in Signal Processing
    6. IEEE Signal Processing Letters
    7. IEEE Signal Processing Magazine
    8. IEEE Transaction on Audio, Speech and Language Processing
    9. IEEE Transaction on Emerging Topics in Computational Intelligence
    10. IEEE Transaction on Signal Processing
    11. Institute of Electronics, Information and Communication Engineers (IEICE) Transactions
    12. Journal of Signal Processing Systems
    13. Journal of Statistical Planning and Inference
    14. Journal of the Acoustical Society of America
    15. Journal of the Acoustical Society of Japan
    16. Speech Communication

- PhD Thesis Review
    1. Suyoun Kim (CMU)
    2. Tomoki Hayashi (Nagoya University)
    3. Dung Tien Tran (INRIA)
    4. Antti Hurmalainen (Tampere University of Technology)

JOINT RESEARCH PROJECTS

1. Sony, Tokyo, Japan (2019.7–)
2. Yahoo! Japan, Tokyo, Japan (2018.9–)
3. Hitachi Corporation, Tokyo, Japan (2018.7–)
4. NTT Communication Science Laboratories, Kyoto, Japan. Dr. Marc Delcroix (2017.12 –)
5. Tokyo Institute of Technology, Tokyo, Japan. Professor Takahiro Shinozaki (2014.4 -)
6. University of Tokyo, Tokyo, Japan. Professor Shigeki Sagayama (2008.7 - 2011.11)
7. Wased University, Tokyo, Japan. Professor Tetsunori Kobayashi (2009.9 - 2010.2)
8. Doshisha University, Kyoto, Japan. Professor Shigeru Katagiri (2010. 4 - 2011.11)
9. University of Tokyo, Tokyo, Japan. Professor Nobuyuki Minematsu (2010.7 - 2011.3)
10. Georgia Institute of Technology, Atlanta, USA. Professor Biing-Hwang (Fred) Juang (2010.11 - 2011.11)

<div align="center">**SUPERVISOR FOR STUDENT**</div>

- JHU (JHU student)
  1. 2019. 09 - : Jiatong Shi
  2. 2019. 09 - : Xuankai Chang
  3. 2018. 07 - 2019. 06: Zhiqi Wang
  4. 2017. 12 - : Matthew Wiesner (co-supervising)
  5. 2017. 10 - : Matthew Maciejewski (co-supervising)
  6. 2017. 09 - 2018. 12: Szu-Jui Chen
  7. 2017. 09 - : Aswin Shanmugam Subramanian

- JHU (Visitor)
  1. 2019. 07 - 2019. 10: Katsuki Inoue (Okayama University)
  2. 2018. 11 - 2019. 05: Murali Karthick Baskar (Brno University of Technology)
  3. 2018. 09 - 2018. 12: Xuankai Chang (Shanghai Jiao Tong University)
  4. 2018. 08 - 2018. 09: Hirofumi Inaguma (Kyoto University)
  5. 2018. 07 - 2019. 06: Yusuke Fujita (Hitachi Ltd.)
  6. 2018. 04 - 2018. 09: Nelson Enrique Yalta Soplin (Waseda University)c

- MERL
  1. 2016. 11 - 2016. 01: Tsubasa Ochiai (Doshisha University)
  2. 2016. 08 - 2016. 11: Tomoki Hayashi (Nagoya University)
  3. 2016. 05 - 2016. 08: Zhong Meng (Georgia Institute of Technology)
  4. 2016. 05 - 2016. 08: Suyoun Kim (Carnegie Mellon University)
  5. 2015. 09 - 2015. 12: Morteza Shahriari (University of Florida)
  6. 2014. 11 - 2015. 05: Zhuo Chen (Columbia University → Microsoft)
  7. 2014. 09 - 2014. 11: Ahmed Hussen Abdelaziz (Ruhr-Universitt Bochum → ICSI → Apple)
  8. 2014. 06 - 2014. 09: Yi Luan (University of Washington)
  9. 2013. 09 - 2014. 02: Felix Weninger (Technische Universitt Mnchen → Nuance)
  10. 2013. 06 - 2013. 08: Hao Tang (Toyota Technological Institute at Chicago)
  11. 2012. 12 - 2013. 02: Koichiro Yoshino (Kyoto University → NAIST)

- NTT
  1. 2011. 05 - 2011. 08: Ekapol Chuangsuwanich (MIT)
  2. 2011. 01 - 2011. 03: Yasuhisa Fujii (Toyohashi University of Technology → Google)
  3. 2010. 01 - 2010. 08: Jose Domingo Esparza Garcia (Technical University of Cartagena → Ulm University)
  4. 2010. 01 - 2010. 03: Daisuke Saito (University of Tokyo)
  5. 2009. 08: Denis Babani (NAIST → OwnerIQ, Inc.)
  6. 2009. 08 - 2009. 09, 2010. 08 - 2010. 09: Naohiro Tawara (Waseda University)
  7. 2008. 10 - 2008. 12: Yotaro Kubo (Waseda University → NTT Labs → Amazon)
  8. 2007. 12 - 2008. 11: David Cournapeau (Kyoto University → Enthought)
  9. 2007. 08: Kenta Nishiki (University of Tokyo → NTT Labs)
  10. 2006. 01 - 2006. 08: Rasa Narkeviciute (Kaunas University of Technology → AEA Technology)
  11. 2006. 01 - 2006. 02: Hirokazu Kameoka (University of Tokyo → NTT Labs/University of Tokyo)
  12. 2005. 07 - 2006. 05: Wesley William Arnquist (University of Washington → Costco Wholesale)
  13. 2003. 08: Toshiaki Kubo (Waseda University → Mitsubishi Electric)
  14. 2003. 04 - 2004. 03, 2008.10: Atsushi Sako (Ryukoku University → Kobe University → Nintendo)