

## Heiga ZEN

Google UK

6 Pancras Road, London, UK

Email: heigazen@google.com

Web: <https://research.google.com/pubs/HeigaZen.html>

## RESEARCH INTERESTS

text-to-speech synthesis, automatic speech recognition, voice conversion, natural language processing, machine learning, deep learning

## PROFESSIONAL EXPERIENCE

- **Research Scientist, Google, July 2011–Present**

- November 2017–Present: Senior Staff Research Scientist
- November 2015–November 2017: Staff Research Scientist
- June 2013–November 2015: Senior Research Scientist
- July 2011–June 2013: Research Scientist

Text-to-speech synthesis research and development.

- **Research Engineer, Speech Technology Group, Toshiba Research Europe Ltd. Cambridge Research Laboratory, July 2008 – July 2011**

Speech synthesis research.

- **Postdoctoral Research Associate, EC FP7 EMIME project, Nagoya Institute of Technology, April 2008 – July 2008**

Develop and release a software for HMM-based speech synthesis.

- **Postdoctoral Research Associate, MEXT e-Society Project, Nagoya Institute of Technology, April 2006 – March 2008**

Developed and released a software for HMM-based speech synthesis. Participated to the open evaluation of corpus-based speech synthesis systems, Blizzard Challenge, with the state-of-the-art HMM-based speech synthesis systems. Investigated new acoustic models for statistical speech recognition and synthesis.

- **Intern/Co-op Researcher, IBM T.J. Watson Research Center, June 2004 – May 2005**

Developed a hybrid system between the HMM-based and unit selection-based text-to-speech synthesis systems. Investigated a weighted finite state transducer-based search algorithm for unit selection-based text-to-speech synthesis system.

- **Intern Researcher, ATR Spoken Language Translation Research Lab., April 2003 – March 2004**

Integrated HMM-based speech synthesis as a prosody prediction module of ATR's text-to-speech synthesis system (XIMERA). Investigated better training procedure and contextual factors (Japanese and Mandarin) for the HMM-based speech synthesis.

- **Intern student, ATR Spoken Language Translation Research Lab., October 2002 – November 2002**

Developed a statistical prompt text selection algorithm for corpus-based speech synthesis.

## UNIVERSITY DEGREES

- **Doctor of Engineering, Nagoya Institute of Technology, Nagoya, Japan, March 2006**  
Concentration: Statistical speech recognition and synthesis, Prof. Keiichi Tokuda Lab.
- **MS in Electrical and Computer Engineering, Nagoya Institute of Technology, Nagoya, Japan, March 2003**  
Concentration: Statistical speech recognition and synthesis, Prof. Keiichi Tokuda Lab.
- **BS in Computer Science, Nagoya Institute of Technology, Nagoya, Japan, March 2001**  
Concentration: Spoken dialog system, Prof. Tadashi Kitamura Lab.
- **AS in Electronic and Information Engineering, Suzuka National College of Technology, Suzuka, Japan, March 1999**  
Concentration: Speech signal processing, Prof. Yuji Kuwabara Lab.

## PUBLICATION

### BOOK CHAPTERS

1. **Heiga Zen**, Keiichi Tokuda, 7.3 *The HMM-based speech synthesis system (HTS)*, In: S. Itahashi, C.-Y. Tseng (Eds.), *Computer processing of Asian spoken languages*, Consideration Books, 2010. (ISBN 978-0-935047-72-1).
2. Keiichi Tokuda, **Heiga Zen**, Alan W. Black, *HMM-based approach to multilingual speech synthesis*, In: S. Narayanan, A. Alwan (Eds.), *Text to speech synthesis: New paradigms and advances*, Prentice Hall, 2004 (ISBN: 0-13-145661-X).

### REFEREED JOURNAL PAPERS

1. **Heiga Zen**, Keiichi Tokuda, Tadashi Kitamura, *Decision tree based simultaneous clustering of phonetic contexts, dimensions, and state positions for acoustic modeling*, IEICE Trans. Inf. & Syst., vol. J87-D-II, no. 8, pp. 1593–1602, August 2004.
2. Amaro de Lima, **Heiga Zen**, Yoshihiko Nankaku, Chiyomi Miyajima, Keiichi Tokuda, Tadashi Kitamura, *On the use of kernel PCA for feature extraction of speech recognition*, IEICE Trans. Inf. & Syst., vol. E87-D, no. 12, pp. 2802–2811, December 2004.

3. Hiroyuki Suzuki, **Heiga Zen**, Yoshihiko Nankaku, Chiyomi Miyajima, Keiichi Tokuda, Tadashi Kitamura, *Continuous speech recognition based on general factor dependent acoustic models*, IEICE Trans. Inf. & Syst., vol. E88-D, no. 3, pp. 410–417, March 2005.
4. Yohei Itaya, **Heiga Zen**, Yoshihiko Nankaku, Chiyomi Miyajima, Keiichi Tokuda, Tadashi Kitamura, *Deterministic annealing EM algorithm in acoustic modeling for speaker and speech recognition*, IEICE Trans. Inf. & Syst., vol. E88-D, no. 3, pp. 425–431, March 2005.
5. Amaro de Lima, **Heiga Zen**, Yoshihiko Nankaku, Keiichi Tokuda, Tadashi Kitamura, Fernand Gil Resende Jr, *Applying sparse KPCA for feature extraction in speech recognition*, IEICE Trans. Inf. & Syst., vol. E88-D, no. 3, pp. 401–409, March 2005.
6. **Heiga Zen**, Tomoki Toda, Masaru Nakamura, Keiichi Tokuda, *Details of Nitech HMM-based speech synthesis system for the Blizzard Challenge 2005*, IEICE Trans. Inf. & Syst., vol. E90-D, no. 1, pp. 325–333, January 2007.
7. **Heiga Zen**, Keiichi Tokuda, Tadashi Kitamura, *Reformulating the HMM as a trajectory model by imposing explicit relationships between static and dynamic feature vector sequences*, Computer Speech & Language, vol. 21, no. 1, pp. 153–173, January 2007.
8. **Heiga Zen**, Takashi Masuko, Keiichi Tokuda, Takao Kobayashi, Tadashi Kitamura, *A hidden semi-Markov model-based speech synthesis system*, IEICE Trans. Inf. & Syst., vol. E90-D, no. 5, pp. 825–834, May 2007.
9. **Heiga Zen**, Tomoki Toda, Keiichi Tokuda, *The Nitech-NAIST HMM-based speech synthesis system for the Blizzard Challenge 2006*, IEICE Trans. Inf. & Syst., vol. E91-D, no. 6, pp. 1764–1773, June 2008.
10. Keiichiro Oura, **Heiga Zen**, Yoshihiko Nankaku, Akonobu Lee, Keiichi Tokuda, *A fully consistent hidden semi-Markov model-based speech recognition system*, IEICE Trans. Inf. & Syst., vol. E91-D, no. 11, pp. 2693–2700, November 2008.
11. Junichi Yamagishi, Takashi Nose, **Heiga Zen**, Zhen-Hua Ling, Tomoki Toda, Keiichi Tokuda, Simon King, Steve Renals, *A robust speaker-adaptive HMM-based text-to-speech synthesis*, IEEE Trans. Audio Speech & Language Processing, vol. 17, no. 6, pp. 1208–1230, July 2009.
12. **Heiga Zen**, Alan W. Black, Keiichi Tokuda, *Statistical parametric speech synthesis*, Speech Communication, vol. 51, no. 11, pp. 1039–1064, November 2009.
13. Keiichiro Oura, **Heiga Zen**, Yoshihiko Nankaku, Akinobu Lee, Keiichi Tokuda, *A tied covariance technique for HMM-based speech synthesis*, IEICE Trans. Inf. & Syst. vol. E93-D, no. 3, pp. 595–601, March 2010.
14. Ryuta Terashima, **Heiga Zen**, Yoshihiko Nankaku, Keiichi Tokuda, *A frame-based context-dependent acoustic modeling for speech recognition*, Journal of IEEJ Trans. Elect. Inf. & Syst, vol. 130, no. 10, pp.1856-1864, Oct. 2010 (in Japanese).

15. **Heiga Zen**, Yoshihiko Nankaku, Keiichi Tokuda, *Continuous stochastic feature mapping based on trajectory HMMs*, IEEE Trans. Audio Speech & Language Processing, vol. 19, no. 2, pp.417–430, February 2011.
16. Kei Hashimoto, **Heiga Zen**, Yoshihiko Nankaku, Akinobu Lee, Keiichi Tokuda, *Bayesian context clustering using cross validation for speech recognition*, IEICE Trans. Inf. & Syst., vol. E94-D, no. 3, pp. 668–678, March 2011.
17. Kai Yu, **Heiga Zen**, Francois Mairesse, Steve Young, *Context adaptive training with factorized decision trees for HMM-based statistical parametric speech synthesis*, Speech Communication, vol. 32, no. 6, pp. 236–243, November 2011.
18. Sayaka Shiota, Kei Hashimoto, **Heiga Zen**, Yoshihiko Nankaku, Akinobu Lee, Keiichi Tokuda, *Speech recognition based on statistical models including multiple model structures*, Acoustical Science & Technology, vol. 32, no. 6, pp. 236–243, November 2011.
19. **Heiga Zen**, Mark J. F. Gales, Yoshihiko Nankaku, Keiichi Tokuda, *Product of experts for statistical parametric speech synthesis*, IEEE Trans. Audio Speech & Language Processing, vol. 20, no. 3, pp. 794–805, March 2012.
20. **Heiga Zen**, Norbert Braunschweiler, Sabine Buchholz, Mark J. F. Gales, Kate Knill, Sacha Krstulovic, Javier Latorre, *Statistical parametric speech synthesis based on speaker and language factorization*, IEEE Trans. Audio Speech & Language Processing, vol. 20, no. 6, pp. 1713–1724, August 2012.
21. Matt Shannon, **Heiga Zen**, William Byrne, *Autoregressive models for statistical parametric speech synthesis*, IEEE Trans. Audio Speech & Language Processing, vol. 21, no. 3, pp. 587–597, November 2012.
22. Zhen-Hua Ling, Shi-Yin Kang, **Heiga Zen**, Andrew Senior, Mike Schuster, Xiao-Jun Qian, Helen Meng, Li Deng, *Deep learning for acoustic modeling in parametric speech generation: A systematic review of existing techniques and future trends*, IEEE Signal Processing Magazine, vol. 32, no. 3, pp.35–52, 2015.

## REFEREED JOURNAL LETTER

1. **Heiga Zen**, Keiichi Tokuda, Takashi Masuko, Takayoshi Yoshimura, Takao Kobayashi, Tadashi Kitamura, *State duration modeling for HMM-based speech synthesis*, IEICE Trans. Inf. & Syst., vol. E90-D, no. 3, pp. 692–693, Mar. 2007.

## REFEREED CONFERENCE AND WORKSHOP PAPERS

1. **Heiga Zen**, Keiichi Tokuda, Tadashi Kitamura, *Decision tree distribution tying based on a dimensional split technique*, Proc. of ICSLP 2002, pp.1257–1260, Sept. 2002.
2. Keiichi Tokuda, **Heiga Zen**, Alan W. Black, *An HMM-based speech synthesis system applied to English*, Proc. of IEEE Speech Synthesis Workshop, Sept. 2002.
3. Hiroyuki Suzuki, **Heiga Zen**, Yoshihiko Nankaku, Chiyomi Miyajima, Keiichi Tokuda, Tadashi Kitamura, *Speech recognition using voice-characteristic dependent acoustic model*, Proc. of ICASSP 2003, vol.1, pp.740–743, Apr. 2003.

4. Takahiro Hoshiya, Shinji Sako, **Heiga Zen**, Keiichi Tokuda, Takashi Masuko, Takao Kobayashi, Tadashi Kitamura, *Improving the performance of HMM-based very low bitrate speech coding*, Proc. of ICASSP 2003, vol.1, pp.800–803, Apr. 2003.
5. **Heiga Zen**, Keiichi Tokuda, Tadashi Kitamura, *Decision tree based simultaneous clustering of phonetic contexts, dimensions, and state positions for acoustic modeling*, Proc. of Eurospeech 2003, pp.3189–3192, Sept. 2003.
6. Keiichi Tokuda, **Heiga Zen**, Tadashi Kitamura, *Trajectory modeling based on HMMs with the explicit relationship between static and dynamic features*, Proc. of Eurospeech 2003, pp.865–868, Sept. 2003.
7. Ranniery S. Maia, **Heiga Zen**, Keiichi Tokuda, Tadashi Kitamura, *Toward the development of a Brazilian Portuguese text-to-speech system based on HMM*, Proc. of Eurospeech 2003, pp.2465–2468, Sept. 2003.
8. Amaro Lima, **Heiga Zen**, Yoshihiko Nankaku, Chiyomi Miyajima, Keiichi Tokuda, Tadashi Kitamura, *On the use of kernel PCA for feature extraction in speech recognition*, Proc. of Eurospeech 2003, pp.2625–2628, Sept. 2003.
9. **Heiga Zen**, Keiichi Tokuda, Tadashi Kitamura, *A Viterbi algorithm for a trajectory model derived from HMM with explicit relationship between static and dynamic features*, Proc. of ICASSP 2004, pp.837–840, Montreal, May 2004.
10. **Heiga Zen**, Keiichi Tokuda, Tadashi Kitamura, *An introduction of trajectory model into HMM-based speech synthesis*, Proc. of 5th ISCA Speech Synthesis Workshop, Pittsburgh, June 2004.
11. **Heiga Zen**, Keiichi Tokuda, Takashi Masuko, Takao Kobayashi, Tadashi Kitamura, *Hidden semi-Markov model based speech synthesis*, Proc. of ICSLP 2004, vol.II, pp.1397–1400, Jeju, Oct. 2004.
12. Yohei Itaya, **Heiga Zen**, Yoshihiko Nankaku, Chiyomi Miyajima, Keiichi Tokuda, Tadashi Kitamura, *Deterministic annealing EM algorithm in parameter estimation for acoustic model*, Proc. of ICSLP 2004, vol.I, pp.433–436, Jeju, Oct. 2004.
13. Ryosuke Tsuzuki, **Heiga Zen**, Keiichi Tokuda, Tadashi Kitamura, Murtaza Bulut, Shrikanth S. Narayanan, *Constructing emotional speech synthesizers with limited speech database*, Proc. of ICSLP 2004, vol.II, pp.1185–1188, Jeju, Oct. 2004.
14. Keiichi Tokuda, **Heiga Zen**, Tadashi Kitamura, *Reformulating the HMM as a trajectory model*, Proc. of Beyond HMM – Workshop on statistical modeling approach for speech recognition, Kyoto, Dec. 2004.
15. Amaro Lima, **Heiga Zen**, Yoshihiko Nankaku, Keiichi Tokuda, Tadashi Kitamura, *Sparse KPCA for feature extraction in speech recognition*, Proc. of ICASSP2005, vol.I, pp.353–356, Philadelphia, PA, Mar. 2005.
16. **Heiga Zen**, Tomoki Toda, *An overview of Nitech HMM-based speech synthesis system for Blizzard Challenge 2005*, Proc. of Interspeech2005 (Eurospeech), pp.93–96, Lisbon, Sept. 2005.

17. Wael Hamza, Raimo Bakis, Zhang-Wei Shuang, **Heiga Zen**, *On building a concatenative speech synthesis system from the Blizzard Challenge speech databases*, Proc. of Interspeech2005 (Eurospeech), pp.97–100, Lisbon, Sept. 2005.
18. **Heiga Zen**, Yoshihiko Nankaku, Keiichi Tokuda, Tadashi Kitamura, *Estimating trajectory HMM parameters using Monte Carlo EM with Gibbs sampler*, Proc. of ICASSP2006, pp.1173–1176, Toulouse, France, May 2006.
19. Keiichiro Oura, **Heiga Zen**, Yoshihiko Nankaku, Akinobu Lee, Keiichi Tokuda, *Hidden semi-Markov model based speech recognition system using weighted finite-state transducer*, Proc of ICASSP2006, pp.33–34, Toulouse, France, May 2006.
20. **Heiga Zen**, Yoshihiko Nankaku, Keiichi Tokuda, Tadashi Kitamura, *Speaker adaptation of trajectory HMMs using feature-space MLLR*, Proc. of Interspeech2006 (ICSLP), pp.2274–2277, Pittsburgh, PA, Sept. 2006.
21. Keijiro Saino, **Heiga Zen**, Yoshihiko Nankaku, Akinobu Lee, Keiichi Tokuda, *HMM-based singing voice synthesis system*, Proc. of Interspeech2006 (ICSLP), pp.1141–1144, Pittsburgh, PA, Sept. 2006.
22. **Heiga Zen**, Tomoki Toda, Keiichi Tokuda, *The Nitech-NAIST HMM-based speech synthesis system for the Blizzard Challenge 2006*, Proc. of Blizzard Challenge 2006 workshop, Pittsburgh, PA, Sept. 2006.
23. Alan W. Black, **Heiga Zen**, Keiichi Tokuda, *Statistical parametric speech synthesis*, Proc. of ICASSP2007, pp.1229–1232, Honolulu, Hawaii, Apr. 2007.
24. **Heiga Zen**, Yoshihiko Nankaku, Keiichi Tokuda, *Model-space MLLR for trajectory HMMs*, Proc. of Interspeech2007, pp.2065–2068, Antwerp, Belgium, Aug. 2007.
25. Ranniery Maia, Tomoki Toda, **Heiga Zen**, Yoshihiko Nankaku, Keiichi Tokuda, *A trainable excitation model for HMM-based speech synthesis*, Proc. of Interspeech2007, pp.1909–1912, Antwerp, Belgium, Aug. 2007.
26. **Heiga Zen**, Takashi Nose, Junichi Yamagishi, Shinji Sako, Takashi Masuko, Alan W. Black, Keiichi Tokuda, *The HMM-based speech synthesis system version 2.0*, Proc. of ISCA SSW6, pp.294–299, Bonn, Germany, Aug. 2007.
27. Ranniery Maia, Tomoki Toda, **Heiga Zen**, Yoshihiko Nankaku, Keiichi Tokuda, *An excitation model for HMM-based speech synthesis based on residual modeling*, Proc. of ISCA SSW6, pp.131–136, Bonn, Germany, Aug. 2007.
28. Junichi Yamagishi, Takao Kobayashi, Steve Renals, Simon King, **Heiga Zen**, Tomoki Toda, Keiichi Tokuda, *Improved average-voice-based speech synthesis using gender-mixed modeling and a parameter generation algorithm considering GV*, Proc. of ISCA SSW6, pp.125–130, Bonn, Germany, Aug. 2007.
29. Junichi Yamagishi, **Heiga Zen**, Tomoki Toda, Keiichi Tokuda, *Speaker-Independent HMM-based Speech Synthesis System – HTS-2007 System for the Blizzard Challenge 2007*, Proc. of BLZ3-2007, paper-008, Bonn, Germany, Aug. 2007.

30. Yi-Jian Wu, **Heiga Zen**, Yoshihiko Nankaku, Keiichi Tokuda, *Minimum generation error criterion considering global/local variance for HMM-based speech synthesis*, Proc. of ICASSP2008, pp.4621–4624, Las Vegas, NV, Mar. 2008.
31. Junichi Yamagishi, Takashi Nose, **Heiga Zen**, Tomoki Toda, Keiichi Tokuda, *Performance evaluation of the speaker-independent HMM-based speech synthesis system “HTS-2007” for the Blizzard Challenge 2007*, Proc. of ICASSP2008, pp.3957–3960, Las Vegas, NV, Mar. 2008.
32. Yoshihiko Nankaku, Kazuhiro Nakamura, **Heiga Zen**, Keiichi Tokuda, *Acoustic modeling with contextual additive structure for HMM-based speech recognition*, Proc. of ICASSP2008, pp.4469–4472, Las Vegas, NV, Mar. 2008.
33. Junichi Yamagishi, **Heiga Zen**, Yi-Jian Wu, Tomoki Toda, Keiichi Tokuda, *HTS-2008: Yet another evaluation of speaker adaptive HMM-based speech synthesis system*, Proc. of Blizzard Challenge Workshop 2008, Brisbane, Australia, Sept. 2008.
34. Sayaka Shiota, Kei Hashimoto, **Heiga Zen**, Yoshihiko Nankaku, Akinobu Lee, Keiichi Tokuda, *Acoustic modeling based on model structure annealing for speech recognition*, Proc. of Interspeech2008, pp.932–935, Brisbane, Australia, Sept. 2008.
35. Kei Hashimoto, **Heiga Zen**, Yoshihiko Nankaku, Akinobu Lee, Keiichi Tokuda, *Bayesian context clustering using cross valid prior distribution for HMM-Based speech recognition*, Proc. of Interspeech2008, pp.936–939, Brisbane, Australia, Sept. 2008.
36. **Heiga Zen**, Yoshihiko Nankaku, Keiichi Tokuda, *Probabilistic feature mapping based on trajectory HMMs*, Proc. of Interspeech2008, pp.1068–1071, Brisbane, Australia, Sept. 2008.
37. Simon King, Keiichi Tokuda, **Heiga Zen**, Junichi Yamagishi, *Unsupervised adaptation for HMM-based speech synthesis*, Proc. of Interspeech2008, pp.1869–1872, Brisbane, Australia, Sept. 2008.
38. Zhi-Peng Yu, Yi-Jian Wu, **Heiga Zen**, Yoshihiko Nankaku, Keiichi Tokuda, *Analysis of stream-dependent tying structure for HMM-based speech synthesis*, Proc. of ICSP2008, Beijing, China, Oct. 2008.
39. **Heiga Zen**, Yoshihiko Nankaku, Keiichi Tokuda, *Stereo-based stochastic noise compensation based on trajectory GMMs*, Proc. of ICASSP2009, pp.4577–4580, Taipei, Taiwan, April 2009.
40. Kei Hashimoto, **Heiga Zen**, Yoshihiko Nankaku, Takashi Masuko, Keiichi Tokuda, *A Bayesian approach to HMM-based speech synthesis*, Proc. of ICASSP2009, pp.4029–4033, Taipei, Taiwan, April 2009.
41. **Heiga Zen**, Norbert Braunschweiler, *Context-dependent additive log  $F_0$  model for HMM-based speech synthesis*, Proc. of Interspeech2009, pp.2091–2094, Brighton, UK, Sept. 2009.

42. Keiichiro Oura, **Heiga Zen**, Yoshihiko Nankaku, Akinobu Lee, Keiichi Tokuda, *Tying covariance matrices to reduce the footprint of HMM-based speech synthesis systems*, Proc. of Interspeech2009, pp.1759–1762, Brighton, UK, Sept. 2009.
43. **Heiga Zen**, Keiichiro Oura, Takashi Nose, Junichi Yamagishi, Shinji Sako, Tomoki Toda, Takashi Masuko, Alan W. Black, Keiichi Tokuda, *Recent development of the HMM-based speech synthesis system (HTS)*, Proc. APSIPA ASC 2009, Sapporo, Japan, Oct. 2009 (accepted).
44. **Heiga Zen**, Mark J. F. Gales, Yoshihiko Nankaku, Keiichi Tokuda, *Statistical parametric speech synthesis based on product of experts*, Proc. of ICASSP2010, pp.4242–4245, Dallas, TX, U.S.A., March 2010.
45. Ranniery Maia, **Heiga Zen**, Mark J. F. Gales, *Statistical parametric speech synthesis with joint estimation of acoustic and excitation model parameters*, Proc. of ISCA SSW7, pp. 88–93, Kyoto, Japan, Sept. 2010.
46. **Heiga Zen**, Norbert Braunschweiler, Sabine Buchholz, Kate Knill, Sacha Krstulovic, Javier Latorre, *HMM-based polyglot speech synthesis by speaker and language adaptive training*, Proc. of ISCA SSW7, pp. 186–191, Kyoto, Japan, Sept. 2010.
47. **Heiga Zen**, *Speaker and language adaptive training for HMM-based polyglot speech synthesis*, Proc. of Interspeech2010, pp.410–413, Makuhari, Japan, Sept. 2010.
48. Kai Yu, **Heiga Zen**, Francois Mairesse, Steve Young, *Context adaptive training with factorized decision trees for HMM-based speech synthesis*, Proc. of Interspeech2010, pp.414–417, Makuhari, Japan, Sept. 2010.
49. Nicholas Pilkington, **Heiga Zen**, *An implementation of decision tree-based context clustering on graphics processing units*, Proc. of Interspeech2010, pp.833–836, Makuhari, Japan, Sept. 2010.
50. Javier Latorre, M.J.F. Gales, **Heiga Zen**, *Training a parametric-based log F0 model with the minimum generation error criterion*, Proc. of Interspeech2010, pp.2174–2177, Makuhari, Japan, Sept. 2010.
51. **Heiga Zen**, Mark J. F. Gales, *Decision tree-based context clustering based on cross validation and hierarchical priors*, Proc. of ICASSP2011, pp.4560–4563, Prague, Czech, May 2011.
52. Matt Shannon, **Heiga Zen**, William Byrne, *The effect of using normalized models in statistical speech synthesis*, Proc. of Interspeech2011, pp.121–124, Florence, Italy, August 2011.
53. Ranniery Maia, **Heiga Zen**, Kate Knill, Mark J. F. Gales, Sabine Buchholz, *Multipulse sequences for residual signal modeling*, Proc. of Interspeech2011, pp.1833–1836, Florence, Italy, August 2011.
54. Ling-Hui Chen, Yoshihiko Nankaku, **Heiga Zen**, Keiichi Tokuda, Zhen-Hua Ling, Li-Rong Dai, *Estimation of window coefficients for dynamic feature extraction for HMM based speech synthesis*, Proc. of Interspeech2011, pp.1801–1804, Florence, Italy, August 2011.



55. Nicholas Pilkington, **Heiga Zen**, Mark J. F. Gales, *Gaussian process experts for voice conversion*, Proc. of Interspeech2011, pp.2761–2764, Florence, Italy, August 2011.
56. Cassia Valentini-Botinhao, Ranniery Maia, Junichi Yamagishi, Simon King, **Heiga Zen**, *Cepstral analysis based on the Glimpse proportion measure for improving the intelligibility of HMM-based synthetic speech in noise*, Proc of ICASSP2012, pp.3997–4000, Kyoto, Japan, March 2012.
57. Vincent Wan, Javier Latorre, K. K. Chin, Langzhou Chen, Mark J. F. Gales, **Heiga Zen**, Kate Knill, Masami Akamine, *Combining multiple high quality corpora for improving HMM-TTS*, Proc. of Interspeech, pp.1135–1138, Portland, OR, U.S.A., September 2012.
58. **Heiga Zen**, Andrew Senior, Mike Schuster, *Statistical parametric speech synthesis using deep neural networks*, Proc. of ICASSP2013, pp.7962–7966, Vancouver, Canada, May 2013.
59. **Heiga Zen**, Andrew Senior, *Deep mixture density networks for acoustic modeling in statistical parametric speech synthesis*, Proc. of ICASSP2014, pp.3872–3876, Florence, Italy, May 2014.
60. **Heiga Zen**, Hasim Sak, *Unidirectional long short-term memory recurrent neural network with recurrent output layer for low-latency speech synthesis*, Proc. of ICASSP2015, pp.4470–4474, Brisbane, Australia, April 2015.
61. Keiichi Tokuda, **Heiga Zen**, *Directly modeling speech waveforms by neural networks for statistical parametric speech synthesis*, Proc. of ICASSP2015, pp.4215–4219, Brisbane, Australia, April 2015.
62. **Heiga Zen**, *Acoustic modeling in statistical parametric speech synthesis - From HMM to LSTM-RNN*, Proc. of MLSLP, Aizu-Wakamatsu, Japan, September 2015.
63. Keiichi Tokuda, **Heiga Zen**, *Directly modeling voiced and unvoiced components in speech waveforms by neural networks*, Proc. of ICASSP2016, pp. 5640–5644, Shanghai, P.R. China, April 2016.
64. **Heiga Zen**, Yannis Agiomyrgiannakis, Niels Egberts, Fergus Henderson, Przemyslaw Szczepaniak, *Fast, compact, and high quality LSTM-RNN based statistical parametric speech synthesizers for mobile devices*, Proc. of Interspeech2016, San Francisco, CA, U.S.A., September 2016.
65. Bo Li, **Heiga Zen**, *Multi-language multi-speaker acoustic modeling for LSTM-RNN based statistical parametric speech synthesis*, Proc. of Interspeech2016, San Francisco, CA, U.S.A., September 2016.
66. Hideki Kawahara, Yannis Agiomyrgiannakis, **Heiga Zen**, *Using instantaneous frequency and aperiodicity detection to estimate F0 for high-quality speech synthesis*, Proc. ISCA SSW9, Sunnyvale, CA, U.S.A., September 2016.

## UNREFEREED CONFERENCE AND WORKSHOP PAPERS

More than 60 papers have been published. For the complete list, see this page.

Last update: January 2nd, 2018

## MISC

1. **Heiga Zen**, Yoshihiko Nankaku, Tomoki Toda, *ICSLP 2006 Summary - Acoustic Modeling and Speech Synthesis* -, Technical Report of IEICE, SP2006-115, pp.1-6, Dec. 2006 (in Japanese).
2. Kiyohiro Shikano, Kazuya Takeda, Tatsuya Kawahara, Hideki Kawahara, Hiroshi Saruwatari, Keiichi Tokuda, Akinobu Lee, Hiromichi Kawanami, Ryuichi Nishimura, Randy Gomez, Tomoki Toda, Takanobu Nishiura, Toru Takahashi, Hideki Banno, **Heiga Zen**, *E-Society Software Development Project for Speech Recognition and Synthesis (Technical Report)*, Journal of IEICE, vol.92, no.6, pp.475-491, June 2009 (in Japanese).
3. **Heiga Zen**, Keiichi Tokuda, *TechWare: HMM-Based Speech Synthesis Resources*, IEEE Signal Processing Magazine, vol. 26, no. 4, pp. 95–97, July 2009.
4. Keiichi Tokuda, **Heiga Zen**, *Fundamentals and recent advances in HMM-based speech synthesis*, Tutorial given at Interspeech, Sept. 2009.
5. Keiichi Oura, **Heiga Zen**, Shinji Sako, Keiichi Tokuda, *Building HMM-based speech synthesis systems using HTS*, Journal of Human Interface Society, Vol.12, No.1, pp.35-40, 2010 (in Japanese).
6. **Heiga Zen**, *Research and life at Google London*, IEICE Inf. & Syst. Society Magazine, vol.21, no.1, pp.22–23, May 2016.
7. Aäron van den Oord, Sander Dieleman **Heiga Zen**, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, Koray Kavukcuoglu, *WaveNet: A generative model for raw audio*, arXiv:1609.03499, September 2016.
8. Aäron van den Oord, Yazhe Li, Igor Babuschkin, Karen Simonyan, Oriol Vinyals, Koray Kavukcuoglu, George van den Driessche, Edward Lockhart, Luis Carlos Cobo Rus, Florian Stimberg, Norman Casagrande, Dominik Grewe, Seb Noury, Sander Dieleman, Erich Elsen, Nal Kalchbrenner, **Heiga Zen**, Alexander Graves, Helen King, Thomas Walters, Dan Belov, Demis Hassabis, *Parallel WaveNet: Fast High-Fidelity Speech Synthesis*, arXiv:1711.10433, November 2017.

## RESEARCH TALKS

1. **Heiga Zen**, *A trajectory model derived from the HMM by imposing explicit relationship between static and dynamic features for statistical speech recognition and synthesis*, Research Seminar, IBM T.J.Watson Research Center, Yorktown Heights, NY, USA, Dec. 2004.
2. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationship between static and dynamic features*, SSLI Seminar, University of Washington, Seattle, WA, USA, June 2005.
3. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationship between static and dynamic features*, Research Seminar, Microsoft Research, Redmond, WA, USA, June 2005.

4. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationship between static and dynamic features*, CSTR Seminar, University of Edinburgh, Edinburgh, UK, July 2005.
5. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationship between static and dynamic features*, MIL Speech Seminar, University of Cambridge, Cambridge, UK, July 2005.
6. **Heiga Zen**, *Reformulating an HMM as a trajectory model by imposing explicit relationships between static and dynamic features*, Research Seminar, Microsoft Research Asia, Beijing, People's Republic of China, March 2006.
7. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationships between static and dynamic features*, CLSP Seminar, Johns Hopkins University, Baltimore, MD, USA, Sept. 2006.
8. **Heiga Zen**, *Statistical parametric speech synthesis*, MIL Speech Seminar, University of Cambridge, Cambridge, UK, January 2009.
9. **Heiga Zen**, *Continuous stochastic feature mapping based on trajectory HMMs*, 2nd One Day Meeting on Unified Models for Speech Recognition and Synthesis, University of Birmingham, Birmingham, UK, March 2009.
10. **Heiga Zen**, *Fundamentals and recent advances in HMM-based speech synthesis*, Keynote speech given at FALA2010, Nov. 2010.
11. **Heiga Zen**, *Statistical speech synthesis*, Speech Synthesis Seminar Series, University of Cambridge, UK, January 2011.
12. Matt Shannon, **Heiga Zen**, *Modelling trajectories in statistical speech synthesis*, Speech Synthesis Seminar Series, University of Cambridge, Cambridge, UK, January 2011.
13. **Heiga Zen**, *Statistical parametric speech synthesis based on product of experts*, ICCS/HCRC seminar, University of Edinburgh, Edinburgh, UK, March 2011.
14. **Heiga Zen**, *Statistical parametric speech synthesis based on speaker and language factorization*, Speech Synthesis Seminar Series, University of Cambridge, Cambridge, UK, June 2011.
15. **Heiga Zen**, *Text-to-speech synthesis and machine learning*, 9th London Machine Learning Meetup, July 2013.
16. **Heiga Zen**, *Deep learning in speech synthesis*, Keynote speech given at ISCA SSW8, Barcelona, Spain, August 2013.
17. **Heiga Zen**, *Statistical parametric speech synthesis*, Tutorial given at UK Speech Conference, Edinburgh, UK, June 2014.
18. **Heiga Zen**, *Statistical parametric speech synthesis: from HMM to LSTM-RNN*, Lecture given at RTTH Summer School on Speech Technology, Barcelona, Spain, July 2015.
19. **Heiga Zen**, *Statistical parametric speech synthesis: from HMM to LSTM-RNN*, Lecture given at Speech Processing Courses in Crete (SPCC), Crete, Greece, July 2015.

20. **Heiga Zen**, *Acoustic modeling for speech synthesis*, Invited talk given at ASRU 2015, Scottsdale, Arizona, USA, December 2015.
21. **Heiga Zen**, *Neural Network based TTS*, Lecture given at Speech Processing Courses in Crete (SPCC), Crete, Greece, July 2016.
22. **Heiga Zen**, *Generative model-based text-to-speech synthesis*, Invited talk given at CBMM workshop on speech representation, perception, & recognition, Cambridge, Massachusetts, USA, February 2017. shihiko Nankaku, Akinobu Lee, Keiichi Tokuda, *Tying covariance matrices to reduce the footprint of HMM-based speech synthesis systems*, Proc. of Interspeech2009, pp.1759–1762, Brighton, UK, Sept. 2009.
23. **Heiga Zen**, Keiichiro Oura, Takashi Nose, Junichi Yamagishi, Shinji Sako, Tomoki Toda, Takashi Masuko, Alan W. Black, Keiichi Tokuda, *Recent development of the HMM-based speech synthesis system (HTS)*, Proc. APSIPA ASC 2009, Sapporo, Japan, Oct. 2009 (accepted).
24. **Heiga Zen**, Mark J. F. Gales, Yoshihiko Nankaku, Keiichi Tokuda, *Statistical parametric speech synthesis based on product of experts*, Proc. of ICASSP2010, pp.4242–4245, Dallas, TX, U.S.A., March 2010.
25. Ranniery Maia, **Heiga Zen**, Mark J. F. Gales, *Statistical parametric speech synthesis with joint estimation of acoustic and excitation model parameters*, Proc. of ISCA SSW7, pp. 88–93, Kyoto, Japan, Sept. 2010.
26. **Heiga Zen**, Norbert Braunschweiler, Sabine Buchholz, Kate Knill, Sacha Krstulovic, Javier Latorre, *HMM-based polyglot speech synthesis by speaker and language adaptive training*, Proc. of ISCA SSW7, pp. 186–191, Kyoto, Japan, Sept. 2010.
27. **Heiga Zen**, *Speaker and language adaptive training for HMM-based polyglot speech synthesis*, Proc. of Interspeech2010, pp.410–413, Makuhari, Japan, Sept. 2010.
28. Kai Yu, **Heiga Zen**, Francois Mairesse, Steve Young, *Context adaptive training with factorized decision trees for HMM-based speech synthesis*, Proc. of Interspeech2010, pp.414-417, Makuhari, Japan, Sept. 2010.
29. Nicholas Pilkington, **Heiga Zen**, *An implementation of decision tree-based context clustering on graphics processing units*, Proc. of Interspeech2010, pp.833–836, Makuhari, Japan, Sept. 2010.
30. Javier Latorre, M.J.F. Gales, **Heiga Zen**, *Training a parametric-based log F0 model with the minimum generation error criterion*, Proc. of Interspeech2010, pp.2174–2177, Makuhari, Japan, Sept. 2010.
31. **Heiga Zen**, Mark J. F. Gales, *Decision tree-based context clustering based on cross validation and hierarchical priors*, Proc. of ICASSP2011, pp.4560–4563, Prague, Czech, May 2011.
32. Matt Shannon, **Heiga Zen**, William Byrne, *The effect of using normalized models in statistical speech synthesis*, Proc. of Interspeech2011, pp.121–124, Florence, Italy, August 2011.

33. Ranniery Maia, **Heiga Zen**, Kate Knill, Mark J. F. Gales, Sabine Buchholz, *Multipulse sequences for residual signal modeling*, Proc. of Interspeech2011, pp.1833–1836, Florence, Italy, August 2011.
34. Ling-Hui Chen, Yoshihiko Nankaku, **Heiga Zen**, Keiichi Tokuda, Zhen-Hua Ling, Li-Rong Dai, *Estimation of window coefficients for dynamic feature extraction for HMM based speech synthesis*, Proc. of Interspeech2011, pp.1801–1804, Florence, Italy, August 2011.
35. Nicholas Pilkington, **Heiga Zen**, Mark J. F. Gales, *Gaussian process experts for voice conversion*, Proc. of Interspeech2011, pp.2761–2764, Florence, Italy, August 2011.
36. Cassia Valentini-Botinhao, Ranniery Maia, Junichi Yamagishi, Simon King, **Heiga Zen**, *Cepstral analysis based on the Glimpse proportion measure for improving the intelligibility of HMM-based synthetic speech in noise*, Proc of ICASSP2012, pp.3997–4000, Kyoto, Japan, March 2012.
37. Vincent Wan, Javier Latorre, K. K. Chin, Langzhou Chen, Mark J. F. Gales, **Heiga Zen**, Kate Knill, Masami Akamine, *Combining multiple high quality corpora for improving HMM-TTS*, Proc. of Interspeech, pp.1135–1138, Portland, OR, U.S.A., September 2012.
38. **Heiga Zen**, Andrew Senior, Mike Schuster, *Statistical parametric speech synthesis using deep neural networks*, Proc. of ICASSP2013, pp.7962–7966, Vancouver, Canada, May 2013.
39. **Heiga Zen**, Andrew Senior, *Deep mixture density networks for acoustic modeling in statistical parametric speech synthesis*, Proc. of ICASSP2014, pp.3872–3876, Florence, Italy, May 2014.
40. **Heiga Zen**, Hasim Sak, *Unidirectional long short-term memory recurrent neural network with recurrent output layer for low-latency speech synthesis*, Proc. of ICASSP2015, pp.4470–4474, Brisbane, Australia, April 2015.
41. Keiichi Tokuda, **Heiga Zen**, *Directly modeling speech waveforms by neural networks for statistical parametric speech synthesis*, Proc. of ICASSP2015, pp.4215–4219, Brisbane, Australia, April 2015.
42. **Heiga Zen**, *Acoustic modeling in statistical parametric speech synthesis - From HMM to LSTM-RNN*, Proc. of MLSLP, Aizu-Wakamatsu, Japan, September 2015.
43. Keiichi Tokuda, **Heiga Zen**, *Directly modeling voiced and unvoiced components in speech waveforms by neural networks*, Proc. of ICASSP2016, pp. 5640–5644, Shanghai, P.R. China, April 2016.
44. **Heiga Zen**, Yannis Agiomyrgiannakis, Niels Egberts, Fergus Henderson, Przemyslaw Szczepaniak, *Fast, compact, and high quality LSTM-RNN based statistical parametric speech synthesizers for mobile devices*, Proc. of Interspeech2016, San Francisco, CA, U.S.A., September 2016.
45. Bo Li, **Heiga Zen**, *Multi-language multi-speaker acoustic modeling for LSTM-RNN based statistical parametric speech synthesis*, Proc. of Interspeech2016, San Francisco, CA, U.S.A., September 2016.

46. Hideki Kawahara, Yannis Agiomyrgiannakis, **Heiga Zen**, *Using instantaneous frequency and aperiodicity detection to estimate F0 for high-quality speech synthesis*, Proc. ISCA SSW9, Sunnyvale, CA, U.S.A., September 2016.

## UNREFEREED CONFERENCE AND WORKSHOP PAPERS

More than 60 papers have been published. For the complete list, see this page.

## MISC

1. **Heiga Zen**, Yoshihiko Nankaku, Tomoki Toda, *ICSLP 2006 Summary - Acoustic Modeling and Speech Synthesis -*, Technical Report of IEICE, SP2006-115, pp.1-6, Dec. 2006 (in Japanese).
2. Kiyohiro Shikano, Kazuya Takeda, Tatsuya Kawahara, Hideki Kawahara, Hiroshi Saruwatari, Keiichi Tokuda, Akinobu Lee, Hiromichi Kawanami, Ryuichi Nishimura, Randy Gomez, Tomoki Toda, Takanobu Nishiura, Toru Takahashi, Hideki Banno, **Heiga Zen**, *E-Society Software Development Project for Speech Recognition and Synthesis (Technical Report)*, Journal of IEICE, vol.92, no.6, pp.475-491, June 2009 (in Japanese).
3. **Heiga Zen**, Keiichi Tokuda, *TechWare: HMM-Based Speech Synthesis Resources*, IEEE Signal Processing Magazine, vol. 26, no. 4, pp. 95–97, July 2009.
4. Keiichi Tokuda, **Heiga Zen**, *Fundamentals and recent advances in HMM-based speech synthesis*, Tutorial given at Interspeech, Sept. 2009.
5. Keiichihiro Oura, **Heiga Zen**, Shinji Sako, Keiichi Tokuda, *Building HMM-based speech synthesis systems using HTS*, Journal of Human Interface Society, Vol.12, No.1, pp.35-40, 2010 (in Japanese).
6. **Heiga Zen**, *Research and life at Google London*, IEICE Inf. & Syst. Society Magazine, vol.21, no.1, pp.22–23, May 2016.
7. Aäron van den Oord, Sander Dieleman **Heiga Zen**, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, Koray Kavukcuoglu, *WaveNet: A generative model for raw audio*, arXiv:1609.03499, September 2016.
8. Aäron van den Oord, Yazhe Li, Igor Babuschkin, Karen Simonyan, Oriol Vinyals, Koray Kavukcuoglu, George van den Driessche, Edward Lockhart, Luis Carlos Cobo Rus, Florian Stimberg, Norman Casagrande, Dominik Grewe, Seb Noury, Sander Dieleman, Erich Elsen, Nal Kalchbrenner, **Heiga Zen**, Alexander Graves, Helen King, Thomas Walters, Dan Belov, Demis Hassabis, *Parallel WaveNet: Fast High-Fidelity Speech Synthesis*, arXiv:1711.10433, November 2017.

## RESEARCH TALKS

1. **Heiga Zen**, *A trajectory model derived from the HMM by imposing explicit relationship between static and dynamic features for statistical speech recognition and synthesis*, Research Seminar, IBM T.J.Watson Research Center, Yorktown Heights, NY, USA, Dec. 2004.

2. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationship between static and dynamic features*, SSLI Seminar, University of Washington, Seattle, WA, USA, June 2005.
3. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationship between static and dynamic features*, Research Seminar, Microsoft Research, Redmond, WA, USA, June 2005.
4. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationship between static and dynamic features*, CSTR Seminar, University of Edinburgh, Edinburgh, UK, July 2005.
5. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationship between static and dynamic features*, MIL Speech Seminar, University of Cambridge, Cambridge, UK, July 2005.
6. **Heiga Zen**, *Reformulating an HMM as a trajectory model by imposing explicit relationships between static and dynamic features*, Research Seminar, Microsoft Research Asia, Beijing, People's Republic of China, March 2006.
7. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationships between static and dynamic features*, CLSP Seminar, Johns Hopkins University, Baltimore, MD, USA, Sept. 2006.
8. **Heiga Zen**, *Statistical parametric speech synthesis*, MIL Speech Seminar, University of Cambridge, Cambridge, UK, January 2009.
9. **Heiga Zen**, *Continuous stochastic feature mapping based on trajectory HMMs*, 2nd One Day Meeting on Unified Models for Speech Recognition and Synthesis, University of Birmingham, Birmingham, UK, March 2009.
10. **Heiga Zen**, *Fundamentals and recent advances in HMM-based speech synthesis*, Keynote speech given at FALA2010, Nov. 2010.
11. **Heiga Zen**, *Statistical speech synthesis*, Speech Synthesis Seminar Series, University of Cambridge, UK, January 2011.
12. Matt Shannon, **Heiga Zen**, *Modelling trajectories in statistical speech synthesis*, Speech Synthesis Seminar Series, University of Cambridge, Cambridge, UK, January 2011.
13. **Heiga Zen**, *Statistical parametric speech synthesis based on product of experts*, ICCS/HCRC seminar, University of Edinburgh, Edinburgh, UK, March 2011.
14. **Heiga Zen**, *Statistical parametric speech synthesis based on speaker and language factorization*, Speech Synthesis Seminar Series, University of Cambridge, Cambridge, UK, June 2011.
15. **Heiga Zen**, *Text-to-speech synthesis and machine learning*, 9th London Machine Learning Meetup, July 2013.
16. **Heiga Zen**, *Deep learning in speech synthesis*, Keynote speech given at ISCA SSW8, Barcelona, Spain, August 2013.

17. **Heiga Zen**, *Statistical parametric speech synthesis*, Tutorial given at UK Speech Conference, Edinburgh, UK, June 2014.
18. **Heiga Zen**, *Statistical parametric speech synthesis: from HMM to LSTM-RNN*, Lecture given at RTTH Summer School on Speech Technology, Barcelona, Spain, July 2015.
19. **Heiga Zen**, *Statistical parametric speech synthesis: from HMM to LSTM-RNN*, Lecture given at Speech Processing Courses in Crete (SPCC), Crete, Greece, July 2015.
20. **Heiga Zen**, *Acoustic modeling for speech synthesis*, Invited talk given at ASRU 2015, Scottsdale, Arizona, USA, December 2015.
21. **Heiga Zen**, *Neural Network based TTS*, Lecture given at Speech Processing Courses in Crete (SPCC), Crete, Greece, July 2016.
22. **Heiga Zen**, *Generative model-based text-to-speech synthesis*, Invited talk given at CBMM workshop on speech representation, perception, & recognition, Cambridge, Massachusetts, USA, February 2017. shihiko Nankaku, Akinobu Lee, Keiichi Tokuda, *Tying covariance matrices to reduce the footprint of HMM-based speech synthesis systems*, Proc. of Interspeech2009, pp.1759–1762, Brighton, UK, Sept. 2009.
23. **Heiga Zen**, Keiichiro Oura, Takashi Nose, Junichi Yamagishi, Shinji Sako, Tomoki Toda, Takashi Masuko, Alan W. Black, Keiichi Tokuda, *Recent development of the HMM-based speech synthesis system (HTS)*, Proc. APSIPA ASC 2009, Sapporo, Japan, Oct. 2009 (accepted).
24. **Heiga Zen**, Mark J. F. Gales, Yoshihiko Nankaku, Keiichi Tokuda, *Statistical parametric speech synthesis based on product of experts*, Proc. of ICASSP2010, pp.4242–4245, Dallas, TX, U.S.A., March 2010.
25. Ranniery Maia, **Heiga Zen**, Mark J. F. Gales, *Statistical parametric speech synthesis with joint estimation of acoustic and excitation model parameters*, Proc. of ISCA SSW7, pp. 88–93, Kyoto, Japan, Sept. 2010.
26. **Heiga Zen**, Norbert Braunschweiler, Sabine Buchholz, Kate Knill, Sacha Krstulovic, Javier Latorre, *HMM-based polyglot speech synthesis by speaker and language adaptive training*, Proc. of ISCA SSW7, pp. 186–191, Kyoto, Japan, Sept. 2010.
27. **Heiga Zen**, *Speaker and language adaptive training for HMM-based polyglot speech synthesis*, Proc. of Interspeech2010, pp.410–413, Makuhari, Japan, Sept. 2010.
28. Kai Yu, **Heiga Zen**, Francois Mairesse, Steve Young, *Context adaptive training with factorized decision trees for HMM-based speech synthesis*, Proc. of Interspeech2010, pp.414-417, Makuhari, Japan, Sept. 2010.
29. Nicholas Pilkington, **Heiga Zen**, *An implementation of decision tree-based context clustering on graphics processing units*, Proc. of Interspeech2010, pp.833–836, Makuhari, Japan, Sept. 2010.



30. Javier Latorre, M.J.F. Gales, **Heiga Zen**, *Training a parametric-based log F0 model with the minimum generation error criterion*, Proc. of Interspeech2010, pp.2174–2177, Makuhari, Japan, Sept. 2010.
31. **Heiga Zen**, Mark J. F. Gales, *Decision tree-based context clustering based on cross validation and hierarchical priors*, Proc. of ICASSP2011, pp.4560–4563, Prague, Czech, May 2011.
32. Matt Shannon, **Heiga Zen**, William Byrne, *The effect of using normalized models in statistical speech synthesis*, Proc. of Interspeech2011, pp.121–124, Florence, Italy, August 2011.
33. Ranniery Maia, **Heiga Zen**, Kate Knill, Mark J. F. Gales, Sabine Buchholz, *Multipulse sequences for residual signal modeling*, Proc. of Interspeech2011, pp.1833–1836, Florence, Italy, August 2011.
34. Ling-Hui Chen, Yoshihiko Nankaku, **Heiga Zen**, Keiichi Tokuda, Zhen-Hua Ling, Li-Rong Dai, *Estimation of window coefficients for dynamic feature extraction for HMM based speech synthesis*, Proc. of Interspeech2011, pp.1801–1804, Florence, Italy, August 2011.
35. Nicholas Pilkington, **Heiga Zen**, Mark J. F. Gales, *Gaussian process experts for voice conversion*, Proc. of Interspeech2011, pp.2761–2764, Florence, Italy, August 2011.
36. Cassia Valentini-Botinhao, Ranniery Maia, Junichi Yamagishi, Simon King, **Heiga Zen**, *Cepstral analysis based on the Glimpse proportion measure for improving the intelligibility of HMM-based synthetic speech in noise*, Proc of ICASSP2012, pp.3997–4000, Kyoto, Japan, March 2012.
37. Vincent Wan, Javier Latorre, K. K. Chin, Langzhou Chen, Mark J. F. Gales, **Heiga Zen**, Kate Knill, Masami Akamine, *Combining multiple high quality corpora for improving HMM-TTS*, Proc. of Interspeech, pp.1135–1138, Portland, OR, U.S.A., September 2012.
38. **Heiga Zen**, Andrew Senior, Mike Schuster, *Statistical parametric speech synthesis using deep neural networks*, Proc. of ICASSP2013, pp.7962–7966, Vancouver, Canada, May 2013.
39. **Heiga Zen**, Andrew Senior, *Deep mixture density networks for acoustic modeling in statistical parametric speech synthesis*, Proc. of ICASSP2014, pp.3872–3876, Florence, Italy, May 2014.
40. **Heiga Zen**, Hasim Sak, *Unidirectional long short-term memory recurrent neural network with recurrent output layer for low-latency speech synthesis*, Proc. of ICASSP2015, pp.4470–4474, Brisbane, Australia, April 2015.
41. Keiichi Tokuda, **Heiga Zen**, *Directly modeling speech waveforms by neural networks for statistical parametric speech synthesis*, Proc. of ICASSP2015, pp.4215–4219, Brisbane, Australia, April 2015.
42. **Heiga Zen**, *Acoustic modeling in statistical parametric speech synthesis - From HMM to LSTM-RNN*, Proc. of MLSLP, Aizu-Wakamatsu, Japan, September 2015.

43. Keiichi Tokuda, **Heiga Zen**, *Directly modeling voiced and unvoiced components in speech waveforms by neural networks*, Proc. of ICASSP2016, pp. 5640–5644, Shanghai, P.R. China, April 2016.
44. **Heiga Zen**, Yannis Agiomyrgiannakis, Niels Egberts, Fergus Henderson, Przemyslaw Szczepaniak, *Fast, compact, and high quality LSTM-RNN based statistical parametric speech synthesizers for mobile devices*, Proc. of Interspeech2016, San Francisco, CA, U.S.A., September 2016.
45. Bo Li, **Heiga Zen**, *Multi-language multi-speaker acoustic modeling for LSTM-RNN based statistical parametric speech synthesis*, Proc. of Interspeech2016, San Francisco, CA, U.S.A., September 2016.
46. Hideki Kawahara, Yannis Agiomyrgiannakis, **Heiga Zen**, *Using instantaneous frequency and aperiodicity detection to estimate F0 for high-quality speech synthesis*, Proc. ISCA SSW9, Sunnyvale, CA, U.S.A., September 2016.

## UNREFEREED CONFERENCE AND WORKSHOP PAPERS

More than 60 papers have been published. For the complete list, see this page.

## MISC

1. **Heiga Zen**, Yoshihiko Nankaku, Tomoki Toda, *ICSLP 2006 Summary - Acoustic Modeling and Speech Synthesis -*, Technical Report of IEICE, SP2006-115, pp.1-6, Dec. 2006 (in Japanese).
2. Kiyohiro Shikano, Kazuya Takeda, Tatsuya Kawahara, Hideki Kawahara, Hiroshi Saruwatari, Keiichi Tokuda, Akinobu Lee, Hiromichi Kawanami, Ryuichi Nishimura, Randy Gomez, Tomoki Toda, Takanobu Nishiura, Toru Takahashi, Hideki Banno, **Heiga Zen**, *E-Society Software Development Project for Speech Recognition and Synthesis (Technical Report)*, Journal of IEICE, vol.92, no.6, pp.475-491, June 2009 (in Japanese).
3. **Heiga Zen**, Keiichi Tokuda, *TechWare: HMM-Based Speech Synthesis Resources*, IEEE Signal Processing Magazine, vol. 26, no. 4, pp. 95–97, July 2009.
4. Keiichi Tokuda, **Heiga Zen**, *Fundamentals and recent advances in HMM-based speech synthesis*, Tutorial given at Interspeech, Sept. 2009.
5. Keiichiro Oura, **Heiga Zen**, Shinji Sako, Keiichi Tokuda, *Building HMM-based speech synthesis systems using HTS*, Journal of Human Interface Society, Vol.12, No.1, pp.35-40, 2010 (in Japanese).
6. **Heiga Zen**, *Research and life at Google London*, IEICE Inf. & Syst. Society Magazine, vol.21, no.1, pp.22–23, May 2016.
7. Aäron van den Oord, Sander Dieleman **Heiga Zen**, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, Koray Kavukcuoglu, *WaveNet: A generative model for raw audio*, arXiv:1609.03499, September 2016.

8. Aäron van den Oord, Yazhe Li, Igor Babuschkin, Karen Simonyan, Oriol Vinyals, Koray Kavukcuoglu, George van den Driessche, Edward Lockhart, Luis Carlos Cobo Rus, Florian Stimberg, Norman Casagrande, Dominik Grewe, Seb Noury, Sander Dieleman, Erich Elsen, Nal Kalchbrenner, **Heiga Zen**, Alexander Graves, Helen King, Thomas Walters, Dan Belov, Demis Hassabis, *Parallel WaveNet: Fast High-Fidelity Speech Synthesis*, arXiv:1711.10433, November 2017.

## RESEARCH TALKS

1. **Heiga Zen**, *A trajectory model derived from the HMM by imposing explicit relationship between static and dynamic features for statistical speech recognition and synthesis*, Research Seminar, IBM T.J.Watson Research Center, Yorktown Heights, NY, USA, Dec. 2004.
2. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationship between static and dynamic features*, SSLI Seminar, University of Washington, Seattle, WA, USA, June 2005.
3. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationship between static and dynamic features*, Research Seminar, Microsoft Research, Redmond, WA, USA, June 2005.
4. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationship between static and dynamic features*, CSTR Seminar, University of Edinburgh, Edinburgh, UK, July 2005.
5. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationship between static and dynamic features*, MIL Speech Seminar, University of Cambridge, Cambridge, UK, July 2005.
6. **Heiga Zen**, *Reformulating an HMM as a trajectory model by imposing explicit relationships between static and dynamic features*, Research Seminar, Microsoft Research Asia, Beijing, People's Republic of China, March 2006.
7. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationships between static and dynamic features*, CLSP Seminar, Johns Hopkins University, Baltimore, MD, USA, Sept. 2006.
8. **Heiga Zen**, *Statistical parametric speech synthesis*, MIL Speech Seminar, University of Cambridge, Cambridge, UK, January 2009.
9. **Heiga Zen**, *Continuous stochastic feature mapping based on trajectory HMMs*, 2nd One Day Meeting on Unified Models for Speech Recognition and Synthesis, University of Birmingham, Birmingham, UK, March 2009.
10. **Heiga Zen**, *Fundamentals and recent advances in HMM-based speech synthesis*, Keynote speech given at FALA2010, Nov. 2010.
11. **Heiga Zen**, *Statistical speech synthesis*, Speech Synthesis Seminar Series, University of Cambridge, UK, January 2011.

12. Matt Shannon, **Heiga Zen**, *Modelling trajectories in statistical speech synthesis*, Speech Synthesis Seminar Series, University of Cambridge, Cambridge, UK, January 2011.
13. **Heiga Zen**, *Statistical parametric speech synthesis based on product of experts*, ICCS/HCRC seminar, University of Edinburgh, Edinburgh, UK, March 2011.
14. **Heiga Zen**, *Statistical parametric speech synthesis based on speaker and language factorization*, Speech Synthesis Seminar Series, University of Cambridge, Cambridge, UK, June 2011.
15. **Heiga Zen**, *Text-to-speech synthesis and machine learning*, 9th London Machine Learning Meetup, July 2013.
16. **Heiga Zen**, *Deep learning in speech synthesis*, Keynote speech given at ISCA SSW8, Barcelona, Spain, August 2013.
17. **Heiga Zen**, *Statistical parametric speech synthesis*, Tutorial given at UK Speech Conference, Edinburgh, UK, June 2014.
18. **Heiga Zen**, *Statistical parametric speech synthesis: from HMM to LSTM-RNN*, Lecture given at RTTH Summer School on Speech Technology, Barcelona, Spain, July 2015.
19. **Heiga Zen**, *Statistical parametric speech synthesis: from HMM to LSTM-RNN*, Lecture given at Speech Processing Courses in Crete (SPCC), Crete, Greece, July 2015.
20. **Heiga Zen**, *Acoustic modeling for speech synthesis*, Invited talk given at ASRU 2015, Scottsdale, Arizona, USA, December 2015.
21. **Heiga Zen**, *Neural Network based TTS*, Lecture given at Speech Processing Courses in Crete (SPCC), Crete, Greece, July 2016.
22. **Heiga Zen**, *Generative model-based text-to-speech synthesis*, Invited talk given at CBMM workshop on speech representation, perception, & recognition, Cambridge, Massachusetts, USA, February 2017. shihiko Nankaku, Akinobu Lee, Keiichi Tokuda, *Tying covariance matrices to reduce the footprint of HMM-based speech synthesis systems*, Proc. of Interspeech2009, pp.1759–1762, Brighton, UK, Sept. 2009.
23. **Heiga Zen**, Keiichiro Oura, Takashi Nose, Junichi Yamagishi, Shinji Sako, Tomoki Toda, Takashi Masuko, Alan W. Black, Keiichi Tokuda, *Recent development of the HMM-based speech synthesis system (HTS)*, Proc. APSIPA ASC 2009, Sapporo, Japan, Oct. 2009 (accepted).
24. **Heiga Zen**, Mark J. F. Gales, Yoshihiko Nankaku, Keiichi Tokuda, *Statistical parametric speech synthesis based on product of experts*, Proc. of ICASSP2010, pp.4242–4245, Dallas, TX, U.S.A., March 2010.
25. Ranniery Maia, **Heiga Zen**, Mark J. F. Gales, *Statistical parametric speech synthesis with joint estimation of acoustic and excitation model parameters*, Proc. of ISCA SSW7, pp. 88–93, Kyoto, Japan, Sept. 2010.

26. **Heiga Zen**, Norbert Braunschweiler, Sabine Buchholz, Kate Knill, Sacha Krstulovic, Javier Latorre, *HMM-based polyglot speech synthesis by speaker and language adaptive training*, Proc. of ISCA SSW7, pp. 186–191, Kyoto, Japan, Sept. 2010.
27. **Heiga Zen**, *Speaker and language adaptive training for HMM-based polyglot speech synthesis*, Proc. of Interspeech2010, pp.410–413, Makuhari, Japan, Sept. 2010.
28. Kai Yu, **Heiga Zen**, Francois Mairesse, Steve Young, *Context adaptive training with factorized decision trees for HMM-based speech synthesis*, Proc. of Interspeech2010, pp.414-417, Makuhari, Japan, Sept. 2010.
29. Nicholas Pilkington, **Heiga Zen**, *An implementation of decision tree-based context clustering on graphics processing units*, Proc. of Interspeech2010, pp.833–836, Makuhari, Japan, Sept. 2010.
30. Javier Latorre, M.J.F. Gales, **Heiga Zen**, *Training a parametric-based log F0 model with the minimum generation error criterion*, Proc. of Interspeech2010, pp.2174–2177, Makuhari, Japan, Sept. 2010.
31. **Heiga Zen**, Mark J. F. Gales, *Decision tree-based context clustering based on cross validation and hierarchical priors*, Proc. of ICASSP2011, pp.4560–4563, Prague, Czech, May 2011.
32. Matt Shannon, **Heiga Zen**, William Byrne, *The effect of using normalized models in statistical speech synthesis*, Proc. of Interspeech2011, pp.121–124, Florence, Italy, August 2011.
33. Ranniery Maia, **Heiga Zen**, Kate Knill, Mark J. F. Gales, Sabine Buchholz, *Multipulse sequences for residual signal modeling*, Proc. of Interspeech2011, pp.1833–1836, Florence, Italy, August 2011.
34. Ling-Hui Chen, Yoshihiko Nankaku, **Heiga Zen**, Keiichi Tokuda, Zhen-Hua Ling, Li-Rong Dai, *Estimation of window coefficients for dynamic feature extraction for HMM based speech synthesis*, Proc. of Interspeech2011, pp.1801–1804, Florence, Italy, August 2011.
35. Nicholas Pilkington, **Heiga Zen**, Mark J. F. Gales, *Gaussian process experts for voice conversion*, Proc. of Interspeech2011, pp.2761–2764, Florence, Italy, August 2011.
36. Cassia Valentini-Botinhao, Ranniery Maia, Junichi Yamagishi, Simon King, **Heiga Zen**, *Cepstral analysis based on the Glimpse proportion measure for improving the intelligibility of HMM-based synthetic speech in noise*, Proc of ICASSP2012, pp.3997–4000, Kyoto, Japan, March 2012.
37. Vincent Wan, Javier Latorre, K. K. Chin, Langzhou Chen, Mark J. F. Gales, **Heiga Zen**, Kate Knill, Masami Akamine, *Combining multiple high quality corpora for improving HMM-TTS*, Proc. of Interspeech, pp.1135–1138, Portland, OR, U.S.A., September 2012.
38. **Heiga Zen**, Andrew Senior, Mike Schuster, *Statistical parametric speech synthesis using deep neural networks*, Proc. of ICASSP2013, pp.7962–7966, Vancouver, Canada, May 2013.

39. **Heiga Zen**, Andrew Senior, *Deep mixture density networks for acoustic modeling in statistical parametric speech synthesis*, Proc. of ICASSP2014, pp.3872–3876, Florence, Italy, May 2014.
40. **Heiga Zen**, Hasim Sak, *Unidirectional long short-term memory recurrent neural network with recurrent output layer for low-latency speech synthesis*, Proc. of ICASSP2015, pp.4470–4474, Brisbane, Australia, April 2015.
41. Keiichi Tokuda, **Heiga Zen**, *Directly modeling speech waveforms by neural networks for statistical parametric speech synthesis*, Proc. of ICASSP2015, pp.4215–4219, Brisbane, Australia, April 2015.
42. **Heiga Zen**, *Acoustic modeling in statistical parametric speech synthesis - From HMM to LSTM-RNN*, Proc. of MLSP, Aizu-Wakamatsu, Japan, September 2015.
43. Keiichi Tokuda, **Heiga Zen**, *Directly modeling voiced and unvoiced components in speech waveforms by neural networks*, Proc. of ICASSP2016, pp. 5640–5644, Shanghai, P.R. China, April 2016.
44. **Heiga Zen**, Yannis Agiomyrgiannakis, Niels Egberts, Fergus Henderson, Przemyslaw Szczepaniak, *Fast, compact, and high quality LSTM-RNN based statistical parametric speech synthesizers for mobile devices*, Proc. of Interspeech2016, San Francisco, CA, U.S.A., September 2016.
45. Bo Li, **Heiga Zen**, *Multi-language multi-speaker acoustic modeling for LSTM-RNN based statistical parametric speech synthesis*, Proc. of Interspeech2016, San Francisco, CA, U.S.A., September 2016.
46. Hideki Kawahara, Yannis Agiomyrgiannakis, **Heiga Zen**, *Using instantaneous frequency and aperiodicity detection to estimate F0 for high-quality speech synthesis*, Proc. ISCA SSW9, Sunnyvale, CA, U.S.A., September 2016.

## UNREFEREED CONFERENCE AND WORKSHOP PAPERS

More than 60 papers have been published. For the complete list, see this page.

## MISC

1. **Heiga Zen**, Yoshihiko Nankaku, Tomoki Toda, *ICSLP 2006 Summary - Acoustic Modeling and Speech Synthesis* -, Technical Report of IEICE, SP2006-115, pp.1-6, Dec. 2006 (in Japanese).
2. Kiyohiro Shikano, Kazuya Takeda, Tatsuya Kawahara, Hideki Kawahara, Hiroshi Saruwatari, Keiichi Tokuda, Akinobu Lee, Hiromichi Kawanami, Ryuichi Nishimura, Randy Gomez, Tomoki Toda, Takanobu Nishiura, Toru Takahashi, Hideki Banno, **Heiga Zen**, *E-Society Software Development Project for Speech Recognition and Synthesis (Technical Report)*, Journal of IEICE, vol.92, no.6, pp.475-491, June 2009 (in Japanese).
3. **Heiga Zen**, Keiichi Tokuda, *TechWare: HMM-Based Speech Synthesis Resources*, IEEE Signal Processing Magazine, vol. 26, no. 4, pp. 95–97, July 2009.

4. Keiichi Tokuda, **Heiga Zen**, *Fundamentals and recent advances in HMM-based speech synthesis*, Tutorial given at Interspeech, Sept. 2009.
5. Keiichiro Oura, **Heiga Zen**, Shinji Sako, Keiichi Tokuda, *Building HMM-based speech synthesis systems using HTS*, Journal of Human Interface Society, Vol.12, No.1, pp.35-40, 2010 (in Japanese).
6. **Heiga Zen**, *Research and life at Google London*, IEICE Inf. & Syst. Society Magazine, vol.21, no.1, pp.22–23, May 2016.
7. Aäron van den Oord, Sander Dieleman **Heiga Zen**, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, Koray Kavukcuoglu, *WaveNet: A generative model for raw audio*, arXiv:1609.03499, September 2016.
8. Aäron van den Oord, Yazhe Li, Igor Babuschkin, Karen Simonyan, Oriol Vinyals, Koray Kavukcuoglu, George van den Driessche, Edward Lockhart, Luis Carlos Cobo Rus, Florian Stimberg, Norman Casagrande, Dominik Grewe, Seb Noury, Sander Dieleman, Erich Elsen, Nal Kalchbrenner, **Heiga Zen**, Alexander Graves, Helen King, Thomas Walters, Dan Belov, Demis Hassabis, *Parallel WaveNet: Fast High-Fidelity Speech Synthesis*, arXiv:1711.10433, November 2017.

## RESEARCH TALKS

1. **Heiga Zen**, *A trajectory model derived from the HMM by imposing explicit relationship between static and dynamic features for statistical speech recognition and synthesis*, Research Seminar, IBM T.J.Watson Research Center, Yorktown Heights, NY, USA, Dec. 2004.
2. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationship between static and dynamic features*, SSLI Seminar, University of Washington, Seattle, WA, USA, June 2005.
3. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationship between static and dynamic features*, Research Seminar, Microsoft Research, Redmond, WA, USA, June 2005.
4. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationship between static and dynamic features*, CSTR Seminar, University of Edinburgh, Edinburgh, UK, July 2005.
5. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationship between static and dynamic features*, MIL Speech Seminar, University of Cambridge, Cambridge, UK, July 2005.
6. **Heiga Zen**, *Reformulating an HMM as a trajectory model by imposing explicit relationships between static and dynamic features*, Research Seminar, Microsoft Research Asia, Beijing, People's Republic of China, March 2006.
7. **Heiga Zen**, *Reformulating the HMM as a trajectory model by imposing explicit relationships between static and dynamic features*, CLSP Seminar, Johns Hopkins University, Baltimore, MD, USA, Sept. 2006.

8. **Heiga Zen**, *Statistical parametric speech synthesis*, MIL Speech Seminar, University of Cambridge, Cambridge, UK, January 2009.
9. **Heiga Zen**, *Continuous stochastic feature mapping based on trajectory HMMs*, 2nd One Day Meeting on Unified Models for Speech Recognition and Synthesis, University of Birmingham, Birmingham, UK, March 2009.
10. **Heiga Zen**, *Fundamentals and recent advances in HMM-based speech synthesis*, Keynote speech given at FALA2010, Nov. 2010.
11. **Heiga Zen**, *Statistical speech synthesis*, Speech Synthesis Seminar Series, University of Cambridge, UK, January 2011.
12. Matt Shannon, **Heiga Zen**, *Modelling trajectories in statistical speech synthesis*, Speech Synthesis Seminar Series, University of Cambridge, Cambridge, UK, January 2011.
13. **Heiga Zen**, *Statistical parametric speech synthesis based on product of experts*, ICCS/HCRC seminar, University of Edinburgh, Edinburgh, UK, March 2011.
14. **Heiga Zen**, *Statistical parametric speech synthesis based on speaker and language factorization*, Speech Synthesis Seminar Series, University of Cambridge, Cambridge, UK, June 2011.
15. **Heiga Zen**, *Text-to-speech synthesis and machine learning*, 9th London Machine Learning Meetup, July 2013.
16. **Heiga Zen**, *Deep learning in speech synthesis*, Keynote speech given at ISCA SSW8, Barcelona, Spain, August 2013.
17. **Heiga Zen**, *Statistical parametric speech synthesis*, Tutorial given at UK Speech Conference, Edinburgh, UK, June 2014.
18. **Heiga Zen**, *Statistical parametric speech synthesis: from HMM to LSTM-RNN*, Lecture given at RTTH Summer School on Speech Technology, Barcelona, Spain, July 2015.
19. **Heiga Zen**, *Statistical parametric speech synthesis: from HMM to LSTM-RNN*, Lecture given at Speech Processing Courses in Crete (SPCC), Crete, Greece, July 2015.
20. **Heiga Zen**, *Acoustic modeling for speech synthesis*, Invited talk given at ASRU 2015, Scottsdale, Arizona, USA, December 2015.
21. **Heiga Zen**, *Neural Network based TTS*, Lecture given at Speech Processing Courses in Crete (SPCC), Crete, Greece, July 2016.
22. **Heiga Zen**, *Generative model-based text-to-speech synthesis*, Invited talk given at CBMM workshop on speech representation, perception, & recognition, Cambridge, Massachusetts, USA, February 2017.
23. **Heiga Zen**, *Generative model-based text-to-speech synthesis*, Nagoya Institute of Technology, April 2017.
24. **Heiga Zen**, *Generative model-based text-to-speech synthesis*, University of Tokyo, October 2017.



## AWARDS

- Minister of Education Promotional Award, English Technical Writing Test, 1998
- Special Student Award, Suzuka National College of Technology, 1998
- The Awaya Prize Young Researcher Award, Acoustic Society of Japan (ASJ), 2006
- The TELECOM System Technology Prize, Telecommunications Advancement Foundation (TAF), Japan, 2008
- The Itakura Prize Innovative Young Researcher Award, Acoustic Society of Japan (ASJ), 2008
- The Best Paper Award, Institute of Electronics Information and Communication Engineers (IEICE), Information and Systems Society, Japan, 2008
- The Yamashita SIG Research Award, Information Processing Society of Japan (IPSJ), 2009
- Best paper invited to Speech Communication, Interspeech 2010, 2010
- The Kiyasu Special Industrial Achievement Award, Information Processing Society of Japan (IPSJ), 2013

## GRANTS

- Japan Society for the Promotion of Science (JSPS), Grant No. 18800019, April 2006 – March 2008
- Hori Information Science Promotion Foundation, April 2007 – March 2008

## PROFESSIONAL ACTIVITIES

- ASJ Spring Meeting, Session Sub-Chair (Acoustic Modeling), March 2007.
- ASJ Autumn Meeting, Session Sub-Chair (Speech Synthesis), September 2007.
- ASJ Sprint Meeting, Session Sub-Chair (Adaptation & Transformation), March 2008.
- 7th, 8th, and 9th ISCA Speech Synthesis Workshop, Program Committee Member & Session Chairs.
- Session Chairs at ICASSP.
- Member of IEEE Speech & Language Processing Technical Committee, 2012–2014.
- Area Chair (Speech Synthesis) at Interspeech 2009, 2015, 2016, and 2018.
- Reviewer of IEEE, IEICE, Speech Communication, Computer Speech & Language, and Neural Computation.
- Associate Editor of IEEE/ACM Trans. Acoustics, Speech, and Language Processing from 2017.

## CAREER ACHIEVEMENTS

- **Original author of the HMM-based speech synthesis system (HTS)**

HTS is an open source software toolkit for HMM-based speech synthesis research and development. Almost all research institute and companies which are doing HMM-based speech synthesis research or business (including U. Cambridge, U. Edinburgh, CMU, U. Tokyo, Thinghua U., USTC, IBM, Microsoft, Nuance, Toshiba, VoiceWare, Yamaha, etc.) are using this software.

- **Original author of the hts\_engine API**

The hts\_engine is an open source software for synthesizing speech from HMMs trained by HTS. This software provides C API to use HMM-based speech synthesizer in other software. This software is released under the New & Simplified BSD license. Many research systems and commercial products using HMM-based speech synthesis are based on this software.

- **One of the main developers of the Festival speech synthesis system**

Festival speech synthesis system is a platform to do speech synthesis research and development. It offers a general framework for building speech synthesis systems as well as including examples of various modules. I was one of the main developers and worked to integrate HMM-based speech synthesis functionality to this software.

- **Active contributor to the HMM toolkit (HTK)**

The HMM toolkit (HTK) has been released from U. Cambridge and widely used in speech recognition research and development area. I have been an active contributor to this toolkit by reporting bugs and sending patch codes.

- **Active participant of the Blizzard Challenge events**

The Blizzard Challenge is an annual evaluation to better understand and compare research techniques in building corpus-based speech synthesizers on the same data. The basic challenge is to take the released speech database, build a synthetic voice from the data and synthesize a prescribed set of test sentences. The sentences from each synthesizer will then be evaluated through listening tests. I participated the Blizzard Challenge from the first event in 2005 until I joined to Toshiba in 2008. Our HMM-based speech system achieved the best performance in 2005 and good results in succeeding events. Especially, the winning of HMM-based speech synthesis in 2005 made a paradigm shift in speech synthesis area from concatenative speech synthesis approach to HMM-based approach. Nowadays, almost 60% presentations in speech synthesis sessions in international conferences related to speech research are about HMM-based speech synthesis. It was about 5% before 2005.

- **WaveNet TTS**

WaveNet is a deep generative model for raw audio signals. This is a breakthrough in generative model-based TTS; it achieved higher naturalness than existing concatenative TTS systems.

## COMPUTER SKILLS

- **Language:** C/C++, NVIDIA CUDA, Python
- **Platforms:** Linux, Windows

## **LANGUAGES**

Japanese (native) & English.

## **REFERENCES**

References are available upon request.