

## Yuki SAITO (齋藤 佑樹), Ph.D.

Room #140, Eng. bldg. #6, 7-3-1 Hongo, Bunkyo, Tokyo 113-8656, Japan

E-mail: [yuuki\\_saito@ipc.i.u-tokyo.ac.jp](mailto:yuuki_saito@ipc.i.u-tokyo.ac.jp), Website: <http://sython.org/>

### MEMBERSHIPS

---

- Member: Institute of Electrical and Electronics Engineers (IEEE) Signal Processing Society  
#94511943 (SPS)
- Member: International Speech Communication Association (ISCA)  
#18636
- Member: Acoustical Society of Japan (ASJ)  
#20364
- Member: Institute of Electronics, Information and Communication Engineers (IEICE)  
#1912414
- Member: Information Processing Society of Japan (IPSJ)  
#201803618

### RESEARCH INTERESTS

---

Machine learning, signal processing, speech synthesis, and voice conversion

### EDUCATIONAL BACKGROUNDS

---

- Mar. 2021 Ph.D. degree in Information Science and Technology, Department of Information Physics and Computing, Graduate School of Information Science and Technology, The University of Tokyo, Japan.  
*Adviser: Professor Hiroshi Saruwatari*
- Mar. 2018 M.S. degree in Information Science and Technology, Department of Creative Informatics, Graduate School of Information Science and Technology, The University of Tokyo, Japan.  
*Adviser: Professor Hiroshi Saruwatari*
- Mar. 2014 B.S. degree in Engineering, Advanced Course of Electronic and Information Systems Engineering, National Institute of Technology, Kushiro College, Japan.  
*Adviser: Assistant Professor Hiroshi Tenmoto*
- Mar. 2012 A.S. degree in Engineering, Department of Information Engineering, National Institute of Technology, Kushiro College, Japan.  
*Adviser: Assistant Professor Hiroshi Tenmoto*

### RESEARCH AND WORK EXPERIENCES

---

- Apr. 2024—Current Lecturer of The University of Tokyo, Japan.
- Apr. 2023—Mar. 2024 Assistant Professor of The University of Tokyo, Japan.

- Apr. 2021—Mar. 2023 Project Research Associate of The University of Tokyo, Japan.  
*“Research and Development on Acoustic Information Processing and Voice Conversion,” Moonshot Research & Development Program of Japan Science and Technology Agency, Representative: Hiroshi Saruwatari*
- Jun. 2019—Mar. 2020 Short-time researcher in DeNA Co., Ltd., Japan.  
*Mentor: Kentaro Tachibana*
- Apr. 2019—Mar. 2021 Research assistant of The University of Tokyo, Japan.  
*“Stress-free, real-time, and full-band voice conversion based on perceptual models,” executed under the Commissioned Research of MIC SCOPE 182103104, Representative: Shinnosuke Takamichi*
- Oct. 2018—Mar. 2019 Short-time researcher in DeNA Co., Ltd., Japan.  
*Mentor: Kentaro Tachibana*
- Apr. 2018—Mar. 2021 Research fellow (DC1) of Japan Society for the Promotion of Science (JSPS), Japan.  
*“Active speech synthesis based on listener perceptual modeling,” JSPS KAKENHI 18J22090, Representative: Yuki Saito*
- Aug. 2017—Oct. 2017 Short-time researcher in NTT Media Intelligence Laboratories, NTT Corporation, Japan.  
*Mentor: Yusuke Ijima*
- Aug. 2016—Sep. 2016 Short-time researcher in NTT Communication Science Laboratories, NTT Corporation, Japan.  
*Mentor: Hirokazu Kameoka*

#### PROFESSIONAL ACTIVITIES

---

- Apr. 2024—Current Paper Review for Acoustical Science and Technology (AST)
- Jun. 2023—Current Paper Reviews for Computer Speech and Language (CSL)
- Dec. 2022—Current Paper Reviews for Journal of Audio Engineering Society (AES)
- Dec. 2022—Current Paper Reviews for IEICE Transactions on Information and Systems
- Aug. 2022—Current Paper Reviews for Journal of Information Processing
- Jun. 2021—Current Paper Reviews for APSIPA Transactions on Signal and Information Processing (TSIP)
- Apr. 2021—Current Paper Reviews for EURASIP Journal on Audio Speech and Music Processing (ASMP)
- Apr. 2021—Current Paper Reviews for INTERSPEECH
- Mar. 2021—Current Paper Reviews for IEEE Access
- Sep. 2020—Current Paper Reviews for IEEE/ACM Transactions on Audio, Speech, and

---

	Language Processing (TASLP)
Jul. 2019—Current	Paper Reviews for IEEE International Workshop on Machine Learning for Signal Processing (MLSP)
Sep. 2018—Current	Paper Reviews for IEEE Signal Processing Letters (SPL)
May 2018—Current	Paper Reviews for IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)

---

#### ACADEMIC ACTIVITIES

---

Apr. 2024—Current	Board member of IEICE Technical Committee on Speech (SP)
Apr. 2024—Current	Board member of IPSJ SIG-SLP Committee
Sep. 2023—Current	Vice-chairman of ASJ Meeting
Mar. 2017—Mar. 2023	Students and Young Researchers Forum of ASJ, Japan. (Official organization of ASJ)

---

#### PATENTS

---

Filled in Feb. 5, 2021	Shinnosuke Takamichi, <b>Yuki Saito</b> , Takaaki Saeki, and Hiroshi Saruwatari, "VOICE CONVERSION DEVICE, VOICE CONVERSION METHOD, AND VOICE CONVERSION PROGRAM," PCT/JP2021/004367
Filled in Aug. 18, 2020	Shinnosuke Takamichi, <b>Yuki Saito</b> , Takaaki Saeki, and Hiroshi Saruwatari, "VOICE CONVERSION DEVICE, VOICE CONVERSION METHOD, AND VOICE CONVERSION PROGRAM," PCT/JP2020/031122
Filled in Aug. 19, 2019	Shinnosuke Takamichi, <b>Yuki Saito</b> , Takaaki Saeki, and Hiroshi Saruwatari, "VOICE CONVERSION DEVICE, VOICE CONVERSION METHOD, AND VOICE CONVERSION PROGRAM," JP2021032940
Filled in May 21	Kentaro Tachibana, <b>Yuki Saito</b> , Kei Akuzawa, "SPEECH PROCESSING APPARATUS AND SPEECH PROCESSING PROGRAM," JP2020190605

---

#### COMPETITIVE FUNDS

---

Nov. 2023—Oct. 2024	Google-initiated Research Grant, 30,000 USD (Representative: <b>Yuki Saito</b> )
Oct. 2023—Mar. 2026	ACT-X, Japan Science and Technology Agency (JST), 4,500,000 JPY (Representative: <b>Yuki Saito</b> )
Aug. 2023	Travel Grant Award for INTERSPEECH2023, 750 EUR
Apr. 2023—Mar. 2026	Research Grant (S) from Tateisi Science and Technology Foundation, 30,000,000 JPY (Representative: Hiroshi Saruwatari)
Apr. 2022—Mar. 2025	Grant-in-Aid for Young Scientists, JSPS, 3,600,000 JPY (Representative: <b>Yuki Saito</b> )
Apr. 2022—Mar. 2023	Research Grant (A) from Tateisi Science and Technology Foundation, 2,200,000 JPY (Representative: <b>Yuki Saito</b> )

Sep. 2021—Mar. 2023	Grant-in-Aid for Research Activity Start-up, JSPS, 2,400,000 JPY (Representative: <b><u>Yuki Saito</u></b> )
Jun. 2021—Mar. 2022	KIOXIA Incentive Research, 1,000,000 JPY (Representative: <b><u>Yuki Saito</u></b> )
May 2018—Mar. 2021	Grant-in-Aid for JSPS Fellows, JSPS, 2,500,000 JPY (Representative: <b><u>Yuki Saito</u></b> )
Apr. 2018	Grants for Researchers Attending International Conferences from NEC C&C, 250,000 JPY

---

 INVITED/VISITING TALKS
 

---

Dec. 2022	<b><u>Yuki Saito</u></b> , "Towards human-in-the-loop DNN-based speech synthesis technologies," Seminar by IEEE NZ Signal Processing / Information Theory Joint Chapter and Acoustics Research Center, The University of Auckland, Dec. 2022. <a href="#">[Invited]</a>
Aug. 2022	<b><u>Yuki Saito</u></b> , "Towards human-in-the-loop speech synthesis technologies," Seminar by IEEE Systems, Man and Cybernetics Singapore Chapter, Chinese and Oriental Languages Information Processing Society Teochew Doctorate Society, Singapore, and Human Language Technology Lab., National University of Singapore, Aug. 2022. <a href="#">[Invited]</a>

---

 TEACHING EXPERIENCES
 

---

Jul. 2024—Current	Instructor of "Information System Laboratory III: Signal Processing and Machine Learning," Department of Mathematical Engineering and Information Physics, The University of Tokyo, Japan. <i>Lectured fundamentals of machine learning.</i>
Jun. 2024—Current	Instructor of "Academics Frontier Lecture (Introduction to Cybernetics – Advanced Information Science Connecting Physics, People, and Society—): Signal Processing Technologies for sound analysis and synthesis," College of Arts and Sciences (Junior Division), The University of Tokyo, Japan. <i>Lectured fundamentals of source separation and sound media synthesis.</i>
Jun. 2024—Current	Instructor of "Applied Acoustics," Department of Mathematical Engineering and Information Physics, The University of Tokyo, Japan. <i>Lectured fundamentals of spoken language processing techniques.</i>
May 2023	Instructor of "Information System Laboratory: Project Practice,"

- Department of Mathematical Engineering and Information Physics, The University of Tokyo, Japan.  
*Lectured how to build speech signal processing systems.*
- Apr. 2023 Instructor of “Mathematical Engineering and Information Physics: Digital Signal Processing and Acoustic Systems,” Department of Mathematical Engineering and Information Physics, The University of Tokyo, Japan.  
*Lectured basic digital signal processing and its application to audio/speech signals.*
- Jun. 2022 Guest Presenter of "Advanced Signal Processing," Graduate School of Information Science and Technology, The University of Tokyo, Japan.  
*Lectured neural speech synthesis & voice conversion technologies.*
- Nov. 2021 Guest Presenter of "Applied Gaussian Process and Machine Learning," Graduate School of Information Science and Technology, The University of Tokyo, Japan.  
*Lectured generative adversarial network-based speech synthesis technologies.*
- May 2016—Jul. 2016 Teaching Assistant of Project Practice, Department of Mathematical & Engineering and Information Physics, The University of Tokyo, Japan.
- May 2017—Jul. 2017 *Lectured how to build speech signal processing systems.*

#### AWARDS

---

- Sep. 2024. Winners of The INTERSPEECH2024 Discrete Speech Challenge (TTS Track)
- Jul. 2024. 2024 IPSJ Yamashita SIG Research Award
- Feb. 2024 The 40th Inoue Research Award for Young Scientists
- Aug. 2023 Travel Grant Award for INTERSPEECH2023
- Jun. 2023 2023 Otagaku Symposium Best Presentation Award
- May 2023 The 22nd Funai Information Technology Award for Young Researchers
- Jun. 2022 The 2021 The IEICE Journal Paper Award
- Mar. 2022 The 2021 IPSJ SIG-SLP Best Student Paper Award (Yahoo! JAPAN Award)
- Jun. 2021 The 2020 IEEE SPS Young Author Best Paper Award
- Mar. 2021 Dean’s Award, Department of Information Physics and Computing, Graduate School of Information Science and Technology, The University of Tokyo, Japan.
- Mar. 2021 The 49th Awaya Prize Young Researcher Award of ASJ
- Jan. 2019 Outstanding Paper Award for Young C&C Researchers
- Nov. 2018 The 12th IEEE SPS Japan Student Journal Paper Award

- Aug. 2018 The Information and Systems Society (ISS) Young Researcher's Award in Speech Field of IEICE
- May 2018 Partial Exemption from Repayment of Scholarship Loan for Students with Outstanding Results, Japan Student Services Organization (JASSO)
- Mar. 2018 The 34th Telecom System Technology Student Award, The Telecommunication Advancement Foundation (TAF)
- Nov. 2017 The 1st IEEE SPS Tokyo Joint Chapter Student Award
- Mar. 2017 The 14th Best Student Presentation Award of ASJ
- Mar. 2017 Spoken Language Processing Student Grant Award of IEEE ICASSP
- Jan. 2017 The ISS Student Poster Award of the Institute of IEICE
- Feb. 2016 Graduation Research Award, Advanced Course of Electronic and Information Systems Engineering, National Institute of Technology, Kushiro College
- Mar. 2014 Dean's Award, Department of Information Engineering, National Institute of Technology, Kushiro College

#### CO-AUTHOR'S AWARDS

---

- Sep. 2024 YANS2024 IVRy Award (Awardee: Taisei Takano)
- Sep. 2024 The 28th Best Student Presentation Award of ASJ (Awardee: Kazuki Yamauchi)
- Aug. 2024 Shortlisted for the ISCA Best Student Paper Award 2024 (Awardee: Dong Yang)
- Jun. 2024 2024 Otagaku Symposium Best Presentation Award (Awardee: Kazuki Yamauchi)
- Mar. 2024 2024 IEICE ISS Student Poster Award (Awardee: Kazuki Yamauchi)
- Mar. 2024 2023 IPSJ SIG-SLP Best Student Paper Award (Fairy Devices Award) (Awardee: Ryunosuke Hirai)
- Mar. 2024 The 27th Best Student Presentation Award of ASJ (Awardee: Aya Watanabe)
- Jul. 2022 Google Travel Grants for Students in East Asia (Awardee: Yuto Nishimura)
- Mar. 2021 National Institute of Technology Student Award (Awardee: Kazuki Fujii)
- Jun. 2020 IPSJ Special Interest Group on Music and Computer (SIG-MUS) / Spoken Language Processing (SLP) Student Poster Award (Awardee: Kazuki Fujii)
- Jun. 2020 FujiSankei Business i Awards (Awardee: Kazuki Fujii)
- Mar. 2020 IPSJ Yamashita Special Interest Group (SIG) Research Award (Awardee: Shinnosuke Takamichi)
- Dec. 2019 The 3rd IEEE SPS Tokyo Joint Chapter Student Award (Awardee: Hiroki Tamaru)
- Mar. 2019 The 18th Best Student Presentation Award of ASJ (Awardee: Satoshi Mizoguchi)
- Jun. 2018 IPSJ SIG-MUS / SLP Presentation Award (Awardee: Shinnosuke Takamichi)

#### LANGUAGES

---

Japanese (native) and English (conversant)

## COMPUTER SKILLS

---

 Programming Python, C, C#, C++, R, Java, JavaScript, Lisp (Scheme), Matlab

Web design Html, CSS

Other LaTeX, Word, Excel, PowerPoint, Unity, XML

 PUBLICATIONS (Google Scholar page: [\[here\]](#))
 

---

**Original Journal Papers**

1. Detai Xin\*, Junfeng Jiang\*, Shinnosuke Takamichi, **Yuki Saito**, Akiko Aizawa and Hiroshi Saruwatari, “JVNV: A corpus of Japanese emotional speech with verbal content and nonverbal expressions,” IEEE Access, Vol. 12, pp. 19752—19764, Feb. 2024. (\*equal contribution)
2. **Yuki Saito\***, Kohei Yatabe\*, and Shogun, “Does controller sound contain valuable information for video game scene analysis? Case study by character identification of Super Smash Bros. Ultimate,” AST, vol. 45, no. 2, pp. 113—116, Feb. 2024. (\*equal contribution)
3. Xuan Luo, Shinnosuke Takamichi, **Yuki Saito**, Tomoki Koriyama, and Hiroshi Saruwatari, “Emotion-controllable speech synthesis using emotion soft label, utterance-level prosodic factors, and word-level prominence,” APSIPA TSIP, Vol. 13, No. 1, 30 pages, Feb. 2024.
4. Satoshi Mizoguchi, **Yuki Saito**, Shinnosuke Takamichi, and Hiroshi Saruwatari, “DNN-based low-musical-noise single-channel speech enhancement based on higher-order-moments matching,” IEICE Transactions on Information and Systems, vol. 104-D, no. 11, pp. 1971—1980, Nov. 2021.
5. Takaaki Saeki, **Yuki Saito**, Shinnosuke Takamichi, and Hiroshi Saruwatari, “Real-time full-band voice conversion with sub-band modeling and data-driven phase estimation of spectral differentials,” IEICE Transactions on Information and Systems, vol. 104-D, no. 7, pp. 1002—1016, Jul. 2021. [[The 2021 IEICE Journal Paper Award](#)]
6. **Yuki Saito**, Shinnosuke Takamichi, and Hiroshi Saruwatari, “Perceptual-similarity-aware deep speaker representation learning for multi-speaker generative modeling,” IEEE/ACM TASLP, vol. 29, pp. 1033—1048, Feb. 2021.
7. **Yuki Saito**, Taiki Nakamura, Yusuke Ijima, Kyosuke Nishida, and Shinnosuke Takamichi, “Non-parallel and many-to-many voice conversion using variational autoencoders integrating speech recognition and speaker verification,” AST, vol. 42, no. 1, pp. 1—11, Jan. 2021.
8. **Yuki Saito**, Kei Akuzawa, and Kentaro Tachibana, “Joint adversarial training of speech recognition and synthesis models for many-to-one voice conversion using phonetic posteriorgrams,” IEICE Transactions on Information and Systems, vol. 103-D, no. 9, pp. 1978—1987, Sep. 2020.
9. Shinnosuke Takamichi, Ryosuke Sonobe, Kentaro Mitsui, **Yuki Saito**, Tomoki Koriyama, Naoko Tanji, and Hiroshi Saruwatari, “JSUT and JVS: free Japanese voice corpora for accelerating speech synthesis research,” AST, vol. 41, no. 5, pp. 761—768, Sep. 2020.

- 
10. Shinnosuke Takamichi, Yuki Saito, Norihiro Takamune, Daichi Kitamura, and Hiroshi Saruwatari, “Phase reconstruction from amplitude spectrograms based on directional-statistics deep neural networks,” *Signal Processing*, vol. 169, 12 pages, Apr. 2020.
  11. Hiroki Tamaru, Yuki Saito, Shinnosuke Takamichi, Tomoki Koriyama, and Hiroshi Saruwatari, “Generative moment matching network-based neural double-tracking for synthesized and natural singing voices,” *IEICE Transactions on Information and Systems*, vol. E103-D, no. 03, pp. 639—647, Mar. 2020.
  12. Yuki Saito, Shinnosuke Takamichi, and Hiroshi Saruwatari, “Vocoder-free text-to-speech synthesis incorporating generative adversarial networks using low-/multi-frequency STFT amplitude spectra,” *CSL*, vol. 58, pp. 343—363, Nov. 2019.
  13. Yuki Saito, Shinnosuke Takamichi, and Hiroshi Saruwatari, “Statistical parametric speech synthesis incorporating generative adversarial networks,” *IEEE/ACM TASLP*, vol. 26, no. 1, pp. 84—96, Jan. 2018. [[The 34th Telecom System Technology Student Award](#), [The Telecommunication Advancement Foundation \(TAF\)](#), [The 12th IEEE Signal Processing Society \(SPS\) Japan Student Journal Paper Award](#), [The 2020 IEEE SPS Young Author Best Paper Award](#)]
  14. Yuki Saito, Shinnosuke Takamichi, and Hiroshi Saruwatari, “Voice conversion using input-to-output highway networks,” *IEICE Transactions on Information and Systems*, vol. E100-D, no. 8, pp. 1925—1928, Aug. 2017.

#### Peer-Reviewed International Conference Proceedings

1. Emiru Tsunoo, Yuki Saito, Wataru Nakata, and Hiroshi Saruwatari, “Causal speech enhancement with predicting semantics based on quantized self-supervised learning features,” *Proc. ICASSP*, 5 pages, Hyderabad, India, Apr. 2025. (ACCEPTED)
2. Kazuki Yamauchi, Wataru Nakata, Yuki Saito, and Hiroshi Saruwatari, “Decoding strategy with perceptual rating prediction for language model-based text-to-speech synthesis,” *Proc. NeurIPS Audio Imagination Workshop*, 7 pages, Vancouver, Canada, Dec. 2024.
3. Wataru Nakata, Takaaki Saeki, Yuki Saito, Shinnosuke Takamichi, and Hiroshi Saruwatari, “NecoBERT: Self-supervised learning model trained by masked language modeling on rich acoustic features derived from neural audio codec,” *Proc. APSIPA ASC*, pp. xxxx—xxxx, Macau, China, Dec. 2024. (ACCEPTED) [[Best Student Paper Candidates](#)]
4. Yuto Ishikawa, Osamu Take, Tomohiko Nakamura, Norihiro Takamune, Yuki Saito, Shinnosuke Takamichi, and Hiroshi Saruwatari, “Real-time noise estimation for Lombard-effect speech synthesis in human--avatar dialogue systems,” *Proc. APSIPA ASC*, pp. xxxx—xxxx, Macau, China, Dec. 2024. (ACCEPTED)
5. Kaito Baba, Wataru Nakata, Yuki Saito, and Hiroshi Saruwatari, “The T05 system for The VoiceMOS Challenge 2024: Transfer learning from deep image classifier to naturalness MOS



- prediction of high-quality synthetic speech,” Proc. SLT, pp. xxxx—xxxx, Macau, China, Dec. 2024. (ACCEPTED)
6. Kazuki Yamauchi, **Yuki Saito**, and Hiroshi Saruwatari, “Cross-dialect text-to-speech in pitch-accent language incorporating multi-dialect phoneme-level BERT,” Proc. SLT, pp. xxxx—xxxx, Macau, China, Dec. 2024. (ACCEPTED)
  7. Dong Yang, Tomoki Koriyama, and **Yuki Saito**, “Frame-wise breath detection with self-training: An exploration of enhancing breath naturalness in text-to-speech,” Proc. INTERSPEECH, pp. 4928—4932, Kos, Greece, Sep. 2024. [[Shortlisted for the ISCA Best Student Paper Award 2024](#)]
  8. Takuto Igarashi, **Yuki Saito**, Kentaro Seki, Shinnosuke Takamichi, Ryuichi Yamamoto, Kentaro Tachibana, and Hiroshi Saruwatari, “Noise-robust voice conversion by conditional denoising training using latent variables of recording quality and environment,” Proc. INTERSPEECH, pp. 2750—2754, Kos, Greece, Sep. 2024.
  9. **Yuki Saito**, Takuto Igarashi, Kentaro Seki, Shinnosuke Takamichi, Ryuichi Yamamoto, Kentaro Tachibana, and Hiroshi Saruwatari, “SRC4VC: Smartphone-recorded corpus for voice conversion benchmark,” Proc. INTERSPEECH, pp. 1825—1829, Kos, Greece, Sep. 2024.
  10. Kentaro Seki, Shinnosuke Takamichi, Norihiro Takamune, **Yuki Saito**, Kanami Imamura, and Hiroshi Saruwatari, “Spatial voice conversion: Voice conversion preserving spatial information and non-target signals,” Proc. INTERSPEECH, pp. 177—181, Kos, Greece, Sep. 2024.
  11. Kazuki Yamauchi, Yusuke Ijima, and **Yuki Saito**, “StyleCap: Automatic speaking-style captioning from speech based on speech and language self-supervised learning models” Proceedings of the 49th ICASSP, Seoul, South Korea, 5 pages, Apr. 2024.
  12. Aya Watanabe, Shinnosuke Takamichi, **Yuki Saito**, Wataru Nakata, Detai Xin, and Hiroshi Saruwatari, “Coco-Nut: Corpus of Japanese utterances and voice characteristics description for prompt-based control,” Proceedings of the 9th IEEE Workshop on Automatic Speech Recognition Understanding (ASRU), pp. 781—788 Taipei, Taiwan, Dec. 2023.
  13. Ryunosuke Hirai, **Yuki Saito**, and Hiroshi Saruwatari, “Federated learning for human-in-the-loop many-to-many voice conversion,” Proceedings of the 12th ISCA Speech Synthesis Workshop, 6 pages, Grenoble, France, Aug. 2023.
  14. **Yuki Saito**, Eiji Iimori, Shinnosuke Takamichi, Kentaro Tachibana, and Hiroshi Saruwatari, “CALLS: Japanese empathetic dialogue speech corpus of complaint handling and attentive listening in customer center,” Proceedings of the 24th INTERSPEECH, pp. 5561—5565, Dublin, Ireland, Aug. 2023. [[Travel Grant Award for INTERSPEECH2023](#)]
  15. Yota Ueda, Shinnosuke Takamichi, **Yuki Saito**, Norihiro Takamune, and Hiroshi Saruwatari,

- “HumanDiffusion: diffusion model using perceptual gradients,” Proceedings of the 24th INTERSPEECH, pp. 4264—4268, Dublin, Ireland, Aug. 2023.
16. **Yuki Saito**, Shinnosuke Takamichi, Eiji Iimori, Kentaro Tachibana, and Hiroshi Saruwatari, “ChatGPT-EDSS: empathetic dialogue speech synthesis trained from ChatGPT-derived context word embeddings,” Proceedings of the 24th INTERSPEECH, pp. 3048—3052, Dublin, Ireland, Aug. 2023. [[Travel Grant Award for INTERSPEECH2023](#)]
  17. Dong Yang, Tomoki Koriyama, **Yuki Saito**, Takaaki Saeki, Detai Xin, and Hiroshi Saruwatari, “Duration-aware pause insertion using pre-trained language model for multi-speaker text-to-speech,” Proceedings of the 48th ICASSP, 5 pages, Rhodes, Greece, Jun. 2023.
  18. Aya Watanabe, Shinnosuke Takamichi, **Yuki Saito**, Detai Xin, and Hiroshi Saruwatari, “Mid-attribute speaker generation using optimal-transport-based interpolation of Gaussian mixture models,” Proceedings of the 48th ICASSP, 5 pages, Rhodes, Greece, Jun. 2023.
  19. Kazuki Fujii, **Yuki Saito**, and Hiroshi Saruwatari, “Adaptive end-to-end text-to-speech synthesis based on error correction feedback from humans,” Proceedings of the 14th APSIPA ASC, pp. 1699—1674, Chiang Mai, Thailand, Nov. 2022.
  20. Yusuke Nakai, **Yuki Saito**, Kenta Udagawa, and Hiroshi Saruwatari, “Multi-task adversarial training algorithm for multi-speaker neural text-to-speech,” Proceedings of the 14th APSIPA ASC, pp. 744—749, Chiang Mai, Thailand, Nov. 2022.
  21. **Yuki Saito**, Yuto Nishimura, Shinnosuke Takamichi, Kentaro Tachibana, and Hiroshi Saruwatari, “STUDIES: Corpus of Japanese empathetic dialogue speech towards friendly voice agent,” Proceedings of the 23rd INTERSPEECH, pp. 5155—5159, Incheon, South Korea, Sep. 2022.
  22. Wataru Nakata, Tomoki Koriyama, Shinnosuke Takamichi, **Yuki Saito**, Yusuke Ijima, Ryo Masumura, and Hiroshi Saruwatari, “Predicting VQVAE-based character acting style from quotation-annotated text for audiobook speech synthesis,” Proceedings of the 23rd INTERSPEECH, pp. 4551—4555, Incheon, South Korea, Sep. 2022.
  23. Yuto Nishimura, **Yuki Saito**, Shinnosuke Takamichi, Kentaro Tachibana, and Hiroshi Saruwatari, “Acoustic modeling for end-to-end empathetic dialogue speech synthesis using linguistic and prosodic contexts of dialogue history,” Proceedings of the 23rd INTERSPEECH, pp. 3373—3377, Incheon, South Korea, Sep. 2022.
  24. Kenta Udagawa, **Yuki Saito**, and Hiroshi Saruwatari, “Human-in-the-loop speaker adaptation for DNN-based multi-speaker TTS,” Proceedings of the 23rd INTERSPEECH, pp. 2968—2972, Incheon, South Korea, Sep. 2022.
  25. Xuan Luo, Shinnosuke Takamichi, Tomoki Koriyama, **Yuki Saito**, and Hiroshi Saruwatari, “Emotion-controllable speech synthesis using emotion soft labels and fine-grained prosody factors,” Proceedings of the 13th APSIPA ASC pp. 794—799, Tokyo, Japan, Dec. 2021.

26. Detai Xin, **Yuki Saito**, Shinnosuke Takamichi, Tomoki Koriyama, and Hiroshi Saruwatari, “Cross-lingual speaker adaptation using domain adaptation and speaker consistency loss for text-to-speech synthesis,” Proceedings of the 22nd INTERSPEECH, pp. 1614—1618, Brno, Czech Republic, Sep. 2021.
27. Yota Ueda, Kazuki Fujii, **Yuki Saito**, Shinnosuke Takamichi, Yukino Baba, and Hiroshi Saruwatari, “HumanACGAN: conditional generative adversarial network with human-based auxiliary classifier and its evaluation in phoneme perception,” Proceedings of the 46th ICASSP, pp.6468—6472, Toronto, Canada, Jun. 2021.
28. Yuki Yamashita, Tomoki Koriyama, **Yuki Saito**, Shinnosuke Takamichi, Yusuke Ijima, Ryo Masumura, and Hiroshi Saruwatari, “Investigating effective additional contextual factors in DNN-based spontaneous speech synthesis,” Proceedings of the 21st INTERSPEECH, pp. 3201—3205, Shanghai, China, Oct. 2020.
29. Detai Xin, **Yuki Saito**, Shinnosuke Takamichi, Tomoki Koriyama, and Hiroshi Saruwatari, “Cross-lingual text-to-speech synthesis via domain adaptation and perceptual similarity regression in speaker space,” Proceedings of the 21st INTERSPEECH, pp. 2947—2951, Shanghai, China, Oct. 2020.
30. Shunsuke Goto, Kotaro Onishi, **Yuki Saito**, Kentaro Tachibana, and Koichiro Mori, “Face2Speech: towards multi-speaker text-to-speech synthesis using an embedding vector predicted from a face image,” Proceedings of the 21st INTERSPEECH, pp. 1321—1325, Shanghai, China, Oct. 2020.
31. Takaaki Saeki, **Yuki Saito**, Shinnosuke Takamichi, and Hiroshi Saruwatari, “Real-time, full-band, online DNN-based voice conversion system using a single CPU,” Proceedings of the 21st INTERSPEECH, pp. 1021—1022, Shanghai, China, Oct. 2020.
32. **Yuki Saito**, Shinnosuke Takamichi, and Hiroshi Saruwatari, “SMASH corpus: a spontaneous speech corpus recording third-person audio commentaries on gameplay,” Proceedings of the 12th LREC, pp. 6573—6579, Marseille, France, May 2020.
33. Yuki Yamashita, Tomoki Koriyama, **Yuki Saito**, Shinnosuke Takamichi, Yusuke Ijima, Ryo Masumura, and Hiroshi Saruwatari, “DNN-based speech synthesis using abundant tags of spontaneous speech corpus,” Proceedings of the 12th LREC, pp. 6440—6445, Marseille, France, May 2020.
34. Kazuki Fujii, **Yuki Saito**, Shinnosuke Takamichi, Yukino Baba, and Hiroshi Saruwatari, “HumanGAN: generative adversarial network with human-based discriminator and its evaluation in speech perception modeling,” Proceedings of the 45th ICASSP, pp. 6239—6243, Barcelona, Spain, May 2020. [[Main contribution paper for FujiSankei Business i Awards](#)]
35. Takaaki Saeki, **Yuki Saito**, Shinnosuke Takamichi, and Hiroshi Saruwatari, “Lifter training and sub-band modeling for computationally efficient and high-quality voice conversion using

- spectral differentials,” Proceedings of the 45th ICASSP, pp. 7784—7788, Barcelona, Spain, May 2020.
36. Taiki Nakamura, **Yuki Saito**, Shinnosuke Takamichi, Yusuke Ijima, and Hiroshi Saruwatari, “V2S attack: building DNN-based voice conversion from automatic speaker verification,” Proceedings of the 10th ISCA Speech Synthesis Workshop, pp. 161—165, Vienna, Austria, Sep. 2019.
  37. **Yuki Saito**, Shinnosuke Takamichi, and Hiroshi Saruwatari, “DNN-based speaker embedding using subjective inter-speaker similarity for multi-speaker modeling in speech synthesis,” Proceedings of the 10th ISCA Speech Synthesis Workshop, pp. 51—56, Vienna, Austria, Sep. 2019.
  38. Hiroki Tamaru, **Yuki Saito**, Shinnosuke Takamichi, Tomoki Koriyama, and Hiroshi Saruwatari, “Generative moment matching network-based random modulation postfilter for DNN-based singing voice synthesis and neural double-tracking,” Proceedings of the 44th ICASSP, pp. 7070—7074, Brighton, UK, May 2019.
  39. Masakazu Une, **Yuki Saito**, Shinnosuke Takamichi, Daichi Kitamura, Ryoichi Miyazaki and Hiroshi Saruwatari, “Generative approach using the noise generation models for DNN-based speech synthesis trained from noisy speech,” Proceedings of the 10th APSIPA ASC, pp. 99—103, Hawaii, U.S.A., Nov. 2018. [[Invited Special Session](#)]
  40. Shinnosuke Takamichi, **Yuki Saito**, Norihiro Takamune, Daichi Kitamura, and Hiroshi Saruwatari, “Phase reconstruction from amplitude spectrograms based on von-Mises-distribution deep neural network,” Proceedings of the 16th IWAENC, pp. 286—290, Tokyo, Japan, Sep. 2018.
  41. **Yuki Saito**, Yusuke Ijima, Kyosuke Nishida, and Shinnosuke Takamichi, “Non-parallel voice conversion using variational autoencoders conditioned by phonetic posteriorgrams and d-vectors,” Proceedings of the 43rd ICASSP, pp. 5274—5278, Alberta, Canada, Apr. 2018. [[Grants for Researchers Attending International Conferences from NEC C&C, Outstanding Paper Award for Young C&C Researchers](#)]
  42. **Yuki Saito**, Shinnosuke Takamichi, and Hiroshi Saruwatari, “Text-to-speech synthesis using STFT spectra based on low-/multi-resolution generative adversarial networks,” Proceedings of the 43rd ICASSP, pp. 5299—5303, Alberta, Canada, Apr. 2018.
  43. Hiroyuki Miyoshi, **Yuki Saito**, Shinnosuke Takamichi, and Hiroshi Saruwatari, “Voice conversion using sequence-to-sequence learning of context posterior probabilities,” Proceedings of the 18th INTERSPEECH, pp. 1268—1272, Stockholm, Sweden, Aug. 2017.
  44. **Yuki Saito**, Shinnosuke Takamichi, and Hiroshi Saruwatari, “Training algorithm to deceive anti-spoofing verification for DNN-based speech synthesis,” Proceedings of the 42nd ICASSP, pp. 4900—4904, New Orleans, U.S.A., Mar. 2017. [[SLP Student Grant Award](#)]

45. **Yuki Saito**, and Hiroshi Tenmoto, “Construction of highly interpretable classification rule based on linear SVM,” Proceedings of the 4th ISTS, Taipei, Taiwan, Nov. 2014.

MISC.

---

- Nov. 2018 I was invited to Google Research Talk, Google Tokyo, Japan.
- May 2018 I was invited to Google Speech Technology Summit 2018, Google London, UK. I talked about two research papers that were accepted to ICASSP 2018 at the poster session.
- Jan. 2018 A figure taken from our paper was used on the cover of the IEEE/ACM TASLP [\(January/February issue\)](#).