Samuele Cornell

Postdoctoral researcher
Carnegie Mellon University
Language Technologies Institute
5000 Forbes Avenue
Pittsburgh, PA
15213-3891

GitHub: https://github.com/popcornell
Website: https://popcornell.github.io | ORCID: https://orcid.org/0000-0002-5358-1844
Google Scholar: https://scholar.google.com/citations?user=A3lfL0QAAAAJ&hl=en&oi=ao (4,000+ citations, 26 h-index)
Email: cornellsamuele@gmail.com, scornell@andrew.cmu.edu, samuele.cornell@ieee.org

Short Bio

Samuele Cornell is currently a postdoctoral research associate at Carnegie Mellon University at the Language Technologies Institute within Prof. Shinji Watanabe's research group (WAVLab).
He got a Master's degree in electronic engineering (summa cum laude) at Università Politecnica delle Marche in 2019 and, in 2023, at the same institution, a doctoral degree in Information Engineering.
His research interests are mainly in the area of robust speech processing (speech enhancement, speech separation, diarization, automatic speech recognition) for distant multi-talker conversational scenarios, and also in the broader field of machine listening (sound event detection and classification) with over 60 publications in these fields.
He is also author and has significant contributions in several open-source speech-processing toolkits (e.g. SpeechBrain, ESPnet, Asteroid source separation) and has organized and co-organized impactful audio processing challenges in the fields of sound event detection, robust speech processing and speech enhancement such as DCASE Task 4 (2022, 2021, 2024), CHiME (CHiME-7/8 DASR lead organizer) and URGENT (2024 and 2025). Core architectures he co-developed, such as SepFormer and TF-GridNet, have been widely adopted by industrial research labs including Google, Microsoft, MERL, and Samsung. He is an elected member of the IEEE Speech and Language Processing Technical Committee (SLTC, Speech Recognition area) for the 2027–2029 term. In Fall 2026 he serves as instructor of CMU's graduate course 11-751/18-781 Speech Recognition and Understanding.

Professional Experience

Carnegie Mellon University, Pittsburgh, PA, USA – Post-doctoral researcher at Language Technologies Institute, WAVLab research group (October 2023 – present)

- JSALT 2025 EMMA team co-lead organizer (distant multi-talker speech recognition and diarization) — co-wrote the funded proposal; co-led a team of 17 researchers
- CHiME-7/8 DASR (Task 1, distant automatic speech recognition) challenge lead organizer — 9 teams, 32 systems across the two editions

- URGENT 2024 and 2025 challenge organizer (universal speech enhancement) — 21 teams in the 2024 blind test phase; co-organized with researchers from Google and Meta
- DCASE 2024 challenge lead organizer (sound event detection) — 23 submitted systems
- LLM + Conversational TTS synthetic data augmentation for multi-talker speech recognition

Carnegie Mellon University, Pittsburgh, PA, USA – Invited Researcher at Language Technologies Institute (April 2022 – August 2022)

- Multi-talker far-field ASR and diarization (SLIDAR paper, ICASSP 2024)
- Speech separation (TFGridNet, Multi-IRIS)
- CHiME-7 DASR (Task 1) lead organizer

Amazon Alexa, Wakeword team, Cambridge, MA, USA – Applied Research Scientist (Intern) (June 2021 - September 2021)

- Device-directed speech detection, keyword spotting
    - Implicit Acoustic Echo Cancellation

Fortell Research (formerly Chromatic), New York, NY, USA – Consulting (July 2024 – July 2025)

- Neural speech separation methods for hearing aid devices; the company has since raised $163M at a $740M valuation (Forbes, 2026).

AlterEgo, USA – Consulting (May 2025 – July 2025)

- Silent speech recognition from myoelectric signals.

Awards & Notable Achievements

- ASRU 2025 hackathon prize for best technical quality and correctness
- JSALT 2025 workshop co-lead of EMMA team (with funding and support through JHU and BUT)
- Meta Audiobox Research Grant
- 1st place at the 2nd Clarity Enhancement Challenge (2022) (results; demo)
- 1st place on L3DAS22 Speech Enhancement Challenge
- Judges' award for most innovative system for DCASE 2020 Task 4 Challenge
- 1st place (Developer Tools track), Global PyTorch Summer Hackathon 2020 — DeMask, built with Asteroid
- Best student paper at IEEE SLT in 2022 (co-author)
- Ranked 2nd place at the 2nd Clarity ICASSP 2023 Grand Challenge (2023)
- 3rd place on track 2 on DIHARD II diarization challenge (SPEED team)
- Finalist on Xilinx Open Hardware Contest 2018
- Outstanding Reviewer Recognition at IEEE ICASSP in 2023

Education

PhD (Dr. Eng.), Information Engineering

Università Politecnica delle Marche, Ancona, June 2023

Thesis: Front-End Processing for Speech Applications with Deep Learning Techniques

Supervisor: Prof. Stefano Squartini

Master of Science, Electronic Engineering

Università Politecnica delle Marche, Ancona, 25 October 2019

Thesis: Detection of Speakers Activity in Challenging Acoustic Environments

Supervisor: Prof. Stefano Squartini

Other Experience

Fred Jelinek Memorial Summer Workshop, Brno – co team leader, Team: EMMA end-to-end multi channel multi talker ASR.

(June 2025 - August 2025)

- Together with Lukas Burget, we submitted a proposal for the workshop (organized each year by JHU) and received funding from US government and industry partners via JHU. I co-led a team of 17 people towards advancing the state of the art of robust speech recognition and diarization.
    - Website: https://jsalt2025.fit.vut.cz/summer-workshop
    - Results include 5 papers submitted to ICASSP 2026, new SotA results across several benchmarks and a new multi-talker ASR framework based on shuffle automata.
    - We are still collaborating regarding end-to-end target speaker ASR and diarization integration as well as on text-to-speech (TTS) based multi speaker synthetic data generation.

CHiME-7 and 8 Distant Automatic Speech Recognition (DASR) Challenges – lead organizer

(September 2022 - August 2024)

- Was responsible for challenge design and organization as well as baseline systems and leading the organizing team.
    - Codebase of baseline system: https://github.com/espnet/espnet/tree/master/egs2/chime7_task1

DCASE Task 4 Sound Event Detection in Home Environments – lead organizer (2021, 2022, 2024 editions)

(June 2022 - August 2024)

- Was responsible for challenge design and organization as well as baseline systems and leading the organizing team.
    - Codebase of baseline system: https://github.com/DCASE-REPO/DESED_task/tree/master

SpeechBrain Developer Team, Remote – Core Developer, co-author

(June 2020 - June 2021)

- SpeechBrain (10,000+ GitHub stars, one of most used speech processing toolkits) dataloading (dynamic batching), data io, data format and attention-based encoder decoder ASR recipes. Collaborated with a software development team of 15+ developers.

Fred Jelinek Memorial Summer Workshop, Montreal – participant, Team: Using Cooperative Ad-hoc Microphone Arrays for ASR, organized by JHU

(June 2019 - August 2019)

- Worked on pre-processing front-end systems for multi-microphone Distant ASR applications: DNN-based beamforming, Voice Activity Detection and Overlapped Speech Detection.

Programming Skills

Languages

Daily use: Python, Bash
was well versed (different projects in the past), now may use occasionally: C/C++, Matlab, VHDL.

DevOps

Git, GitHub, Docker, AWS

Machine Learning/ Frameworks

PyTorch, TensorFlow, NumPy, pandas

Speech and Audio Processing Frameworks

I have contributed or I am author of many of the most used speech processing toolkits in the field, collectively exceeding 23,000 GitHub stars and backed by industry sponsors including Samsung, NVIDIA, and Dolby:
SpeechBrain (**co-author**, 23 merged PRs)
ESPnet (top-10 all time contributors, **co-author** of ESPnet-SE++ and ESPnet3, 25 merged PRs incl. CHiME-7/8 DASR baselines)
Asteroid Source Separation Toolkit (**co-author**, 25 merged PRs incl. DPRNN, DPTNet, DeMask; natively integrated on the Hugging Face Hub — 340+ models)
TorchAudio 2.1 (**co-author**, merged contribution; consulted by the PyTorch core team on complex-tensor and beamforming API design (2020–2021))
Lhotse (contributor, 2 merged PRs)
DCASE-REPO (main contributor/**author**, DCASE Task 4 baselines)
Pyroomacoustics (contributed bug fix to adaptive filtering example, merged via PR #277)
NVIDIA NeMo (2 merged PRs, CHiME-8 baseline code)
Kaldi (1 merged PR)

HPC Experience

SLURM, Sun GridEngine;
Served as main sys admin during my PhD for our group computational infrastructure at UNIVPM.

Language Skills

Italian (native), English (fluent)

Academic Activity

Organizer/Committee Member

- Elected member, IEEE Speech and Language Processing Technical Committee (SLTC), Speech Recognition area, 2027–2029 term
- Secretary of ISCA Special Interest Group on Robust Speech Processing, 2023-2025
- JSALT 2025 EMMA team (co-lead organizer), 2025
- CHiME-8 DASR (Task 1) (lead organizer), 2024
- CHiME-7 DASR (Task 1) (lead organizer), 2023
- URGENT Challenge, (organizer), (NeurIPS, Interspeech, ICASSP, Interspeech) (2024, 2025)
- ICASSP 2023 Special Session on Resource Efficient Speech Enhancement, 2023
- DCASE Task 4 Challenge (organizer), (2021, 2022, 2024)
- Guest editor for Computer Speech and Language Special Issue on "Multi-Speaker, Multi-Microphone, and Multi-Modal Distant Speech Recognition" (2024)

Session Chair

- IEEE ICASSP (2023) Resource-efficient Speech Enhancement Special Session
- IEEE ICASSP (2024) Self-Supervised Learning for Conversational Speech Processing
- IEEE ICASSP (2024) Voice Conversion and Benchmarking
- IEEE ICASSP (2024) Multi-Modal Speech and Motion Generation
- NeurIPS 2024: URGENT Challenge Workshop

Membership

- Member, the Institute of Electrical and Electronics Engineers (IEEE)
- Member, the International Speech Communication Association (ISCA)
- Member, IEEE Computational Intelligence Society Membership
- Member, IEEE Signal Processing Society Membership

Grants & Funded Research Projects

- IC Postdoctoral Research Fellowship (ORAU/ODNI) (2023 – present)
- NSF ACCESS Discover allocation CIS261832, “CMU 11751/18781: Speech Recognition and Understanding, 2026 Fall” — PI; 750,000 ACCESS credits of computing for course assignments and student term projects (2026–2027)
- Meta Audiobox Research Grant (2024)

- CMU–Amazon Research Initiative, “Generative Speech Separation and Restoration for Conversational AI” (PI: S. Watanabe), $120,000 (2026) — helped secure the award
- JSALT 2025 workshop — co-wrote the funded proposal (funding via JHU from US government and industry partners); co-led the EMMA team
- WavLab – MIT Lincoln Laboratory collaboration (Year 5, $360k) — contributor
- Speech processing applications for call center data applications (AGEVOLA project, Fondazione Caritro) (2021-2023, UNIVPM, Italy) (participant)
- Non-intrusive load monitoring: Nilm clOud seRvices for ResIdential userS (NORRIS) (2021, UNIVPM, Italy) (participant)

Teaching

- **Instructor**, 11-751/18-781 Speech Recognition and Understanding, Carnegie Mellon University (Fall 2026) — graduate course cross-listed between the Language Technologies Institute and Electrical & Computer Engineering

- Lecture at JSALT 2025 summer school (June 2025, Brno, Czechia) on robust speech recognition.
- Guest lecture at 11492/11692/18495 Speech Technology for Conversational AI, Carnegie Mellon University (March 2024 and March 2025)
- Guest lecture at Université de Montréal (invited by Prof. Mirco Ravanelli) (February 2021)
- UNIVPM external course "AI & Behaviour Based Safety" for FairConnect S.p.A. (November 2021)
- Guest lectures at UNIVPM, Digital Adaptive Circuits and Learning Systems (DACLS) course (May 2020, May 2021, May 2022, May 2023)

Invited Talks

- JHU HLTCOE (May 2026)
- CMU LTI Colloquium (February 2026)
- MILA Conversational AI reading group (February 2026)
- MBZUAI (January 2026)
- Microsoft (October 2025)
- CMU Speech lunch (October 2025)
- CMU Sphinx lunch (November 2022)
- Fondazione Cluster Marche AI Webinar (November 2022)
- Korea Advanced Institute of Science and Technology (KAIST) (April 2024)
- UNIVPM (May 2024)

Press & Media

Quoted as an expert in *MIT Technology Review*: "Noise-canceling headphones use AI to let a single voice through" (R. Williams, May 2024) — covering the University of Washington's real-time "target speech hearing" system (CHI 2024), built on the TF-GridNet architecture I co-authored.
Quoted as an expert in *MIT Technology Review*: "A new AI translation system for headphones clones multiple voices simultaneously" (R. Williams, May 2025)

Mentoring and Supervising

- during JSALT 2025 workshop (and ongoing):
    - Aug 2025 Rohan Phadke (master student/University of North Carolina, Chapel Hill, USA)
    - Aug 2025 Alexander Polok (PhD student, Brno University of Technology, Czechia)
    - Aug 2025 Darshan Prabhu (PhD student, Indian Institute of Technology Bombay, India)
    - Aug 2025 Dominik Klemens (PhD student, Brno University of Technology, Czechia)
- at CMU
    - June 2026 Thanapat Trachu (visiting researcher)
    - Jan 2026 Alexander Polok (visiting researcher)
    - Jan 2026 Dahee Yang (visiting researcher)
    - Oct 2023 Tyler Scaringella (external collaborator)
    - Oct 2024 Chyi-Jiunn Lin (master student)
    - Jan 2025 Shikhar Bharadwaj (PhD student)
    - Oct 2025 Masao Someki (master student)
- at UNIVPM
    - Jan 2022 Luca Serafini (non-PhD research associate)
    - Jun 2023 Silvio Osimi (master student)
        - Co-supervisor for his Master Thesis on "Loudspeaker Equalization in Reverberant Environments using Deep Convolutive Dereverberation".'
    - Jul 2021 Alessandro Bacà (master student)
        - Co-supervisor for his Master Thesis on "End-to-End Target Speaker Deep Learning Architecture for Automatic Speech Recognition".
    - Jun 2020 Niko Fioravanti (master student)

Review Activity

Grants

- MITACS

Journal

- IEEE/ACM Transactions on Audio, Speech, and Language Processing
- IEEE Signal Processing Letters
- IEEE Journal of Selected Topics in Signal Processing
- IEEE Transactions on Neural Networks and Learning Systems
- Journal of the Acoustical Society of America (JASA)
- Speech Communication
- IEEE Transactions on Mobile Computing
- IEEE Transactions on Emerging Topics in Computing (TETC)
- Neurocomputing
- Machine Learning with Applications
- ACM Computing Surveys
- Computer Speech and Language
- Transactions on Machine Learning Research

Conference

1. ICASSP
    - Meta Reviewer & Area Chair ICASSP 2026
2. ICLR
3. EMNLP
4. AAAI
    - PC member 2027
5. NeurIPS
    - regular & competition track reviewer
6. Interspeech
    - Area Chair 2026
7. IEEE IS2
    - PC member 2023
8. IEEE SLT
9. IEEE ASRU
10. WASPAA
11. DCASE
12. DAFx
13. CHiME Workshop
14. IJCNN (IEEE International Joint Conference on Neural Network)
    - Area chair IJCNN 2025
    - Area chair IJCNN 2026

Publications

Journal

1. **Cornell, S.,** Omologo, M., Squartini, S., & Vincent, E. (2022). Overlapped speech detection and speaker counting using distant microphone arrays. *Computer Speech & Language*, *72*, 101306.
2. Subakan, C., Ravanelli, M., **Cornell, S.,** Grondin, F., & Bronzi, M. (2022). On Using Transformers for Speech-Separation. IEEE/ACM Transactions on Audio, Speech, and Language Processing.
3. Wang, Z. Q., **Cornell, S.,** Choi, S., Lee, Y., Kim, B. Y., & Watanabe, S. (2022). TF-GridNet: Integrating Full-

and Sub-Band Modeling for Speech Separation. IEEE/ACM Transactions on Audio, Speech, and Language Processing.

4. Serafini L, **Cornell, S.,** Morrone G., Zovato E, Brutti A, Squartini S. (2022). An experimental review of speaker diarization methods with application to two-speaker conversational telephone speech recordings. *Computer Speech & Language.*
5. Lu, Y. J., Chang, X., Li, C., Zhang, W., **Cornell, S.,** Ni, Z., ... & Watanabe, S. (2022). Software Design and User Interface of ESPnet-SE++: Speech Enhancement for Robust Speech Processing. *Journal of Open Source Software (JOSS).*
6. Aironi, C., **Cornell, S.**, & Squartini, S. (2024). A Graph-Based Neural Approach to Linear Sum Assignment Problems. *International Journal of Neural Systems*, 2450011-2450011.
7. Morrone, G., **Cornell, S**., Serafini, L., Zovato, E., Brutti, A., & Squartini, S. (2023). End-to-End Integration of Speech Separation and Voice Activity Detection for Low-Latency Diarization of Telephone Conversations. *Speech Communication*.
8. Masuyama, Y., Chang, X., Zhang, W., **Cornell, S.**, Wang, Z. Q., Ono, N., ... & Watanabe, S. (2026). An end-to-end integration of speech separation and recognition with self-supervised learning representation. *Computer Speech & Language*, *95*, 101813.
9. Aironi, C., Gabrielli, L., **Cornell, S.**, & Squartini, S. (2024). Complex-bin2bin: A Latency-Flexible Generative Neural Model for Audio Packet Loss Concealment. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*.
10. **Cornell, S.**, Boeddeker, C., Park, T., Huang, H., Raj, D., Wiesner, M., ... & Watanabe, S. (2025). Recent trends in distant conversational speech recognition: A review of CHiME-7 and 8 DASR challenges. *Computer Speech & Language*, 101901.

Conference

1. Severini, M., Principi, E., **Cornell, S.,** Gabrielli, L., & Squartini, S. (2020, July). Who Cried When: Infant Cry Diarization with Dilated Fully-Convolutional Neural Networks. In *2020 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.
2. **Cornell, S**., Principi, E., & Squartini, S. (2020, July). A Novel Adversarial Training Scheme for Deep Neural Network based Speech Enhancement. In *2020 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.
3. Pariente, M., **Cornell, S.**, Deleforge, A., & Vincent, E. (2020, May). Filterbank design for end-to-end speech separation. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 6364-6368). IEEE.
4. **Cornell, S**., Omologo, M., Squartini, S., & Vincent, E. (2020, October). Detecting and counting overlapping speakers in distant speech scenarios. In *Interspeech 2020*.
5. Pariente, M., **Cornell, S.**, Cosentino, J., Sivasankaran, S., Tzinis, E., Heitkaemper, J., ... & Vincent, E. (2020). Asteroid: The PyTorch-Based Audio Source Separation Toolkit for Researchers. *Interspeech 2020*.
6. **Cornell, S**., Olvera, M., Pariente, M., Pepe, G., Principi, E., Gabrielli, L., & Squartini, S. (2020, November). Domain-adversarial training and trainable parallel front-end for the dcase 2020 task 4 sound event detection challenge. In *DCASE 2020-5th Workshop on Detection and Classification of Acoustic Scenes and Events*.
7. **Cornell, S.,** Olvera, M., Pariente, M., Pepe, G., Principi, E., Gabrielli, L., & Squartini, S. (2020, November). Task-aware separation for the DCASE 2020 task 4 sound event detection and separation challenge. In *DCASE 2020-5th Workshop on Detection and Classification of Acoustic Scenes and Events*.
8. Subakan, C., Ravanelli, M., **Cornell, S.**, Bronzi, M., & Zhong, J. (2021, June). Attention is all you need in speech separation. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 21-25). IEEE.
9. Aironi, C., **Cornell, S.,** Principi, E., & Squartini, S. (2021, August). Graph-based Representation of Audio signals for Sound Event Classification. In *2021 29th European Signal Processing Conference (EUSIPCO)* (pp. 566-570). IEEE.

10. **Cornell, S.,** Brutti, A., Matassoni, M., & Squartini, S. (2021). Learning to Rank Microphones for Distant Speech Recognition. In *Interspeech 2021*.
11. Ronchini, F., Serizel, R., Turpault, N., & **Cornell, S.** (2021, November). The impact of non-target events in synthetic soundscapes for sound event detection. In *DCASE 2021-Detection and Classification of Acoustic Scenes and Events*.
12. Morrone, G., **Cornell, S.,** Zovato, E., Brutti, A., & Squartini, S. (2022, May). Conversational Speech Separation: an Evaluation Study for Streaming Applications. In *Audio Engineering Society Convention 152*. Audio Engineering Society.
13. Lu, Y. J., Chang, X., Li, C., Zhang, W., **Cornell, S.,** Ni, Z., ... & Watanabe, S. (2022). ESPnet-SE++: Speech enhancement for robust speech recognition, translation, and understanding. *ICASSP 2022*.
14. Lu, Y. J**., Cornell, S.,** Chang, X., Zhang, W., Li, C., Ni, Z., ... & Watanabe, S. (2022, May). Towards Low-Distortion Multi-Channel Speech Enhancement: The ESPNET-Se Submission to the L3DAS22 Challenge. *ICASSP 2022* (pp. 9201-9205). IEEE. (first in the L3DAS22 challenge)
15. **Cornell, S.,** Pariente, M., Grondin, F., & Squartini, S. (2022, May). Learning filterbanks for end-to-end acoustic beamforming. *ICASSP 2022* (pp. 6507-6511). IEEE.
16. Subakan, C., Ravanelli, M., **Cornell, S.,** & Grondin, F. (2022, May). REAL-M: Towards speech separation on real mixtures. *ICASSP 2022* (pp. 6862-6866). IEEE.
17. Aironi, C., **Cornell, S.,** Principi, E., & Squartini, S. (2022, August). Graph Node Embeddings for ontology-aware Sound Event Classification: an evaluation study. In *2022 30th European Signal Processing Conference (EUSIPCO)* (pp. 414-418). IEEE.
18. Ronchini, F**., Cornell, S**., Serizel, R., Turpault, N., Fonseca, E., & Ellis, D. P. DESCRIPTION AND ANALYSIS OF NOVELTIES INTRODUCED IN DCASE TASK 4 2022 ON THE BASELINE SYSTEM. DCASE 2022
19. **Cornell, S.,** Balestri, T., & Sénéchal, T. (2022). Implicit Acoustic Echo Cancellation for Keyword Spotting and Device-Directed Speech Detection. *SLT 2022*
20. Morrone, G**., Cornell, S**., Raj, D., Zovato, E., Brutti, A., & Squartini, S. (2022). Low-latency Speech Separation Guided Diarization for Conversational Telephone Speech. SLT 2022. (**equal contribution Cornell and Morrone**)
21. Masuyama, Y., Chang, X., **Cornell, S.,** Watanabe, S., & Ono, N. (2022). End-to-End Integration of Speech Recognition, Dereverberation, Beamforming, and Self-Supervised Learning Representation. *SLT 2022 (best student paper award)*
22. Serizel, R**., Cornell, S.,** & Turpault, N. (2022). PERFORMANCE ABOVE ALL? ENERGY CONSUMPTION VS. PERFORMANCE FOR MACHINE LISTENING, A STUDY ON DCASE TASK 4 BASELINE. ICASSP 2023.
23. Aironi, C., **Cornell, S.**, Serafini, L., & Squartini, S. (2023). Tackling the Linear Sum Assignment Problem with Graph Neural Networks. Applied Intelligence and Informatics: Second International Conference.
24. Wang, Z. Q., **Cornell, S.**, Choi, S., Lee, Y., Kim, B. Y., & Watanabe, S. (2023). TF-GridNet: Making time-frequency domain models great again for monaural speaker separation. ICASSP 2023.
25. Wang, Z. Q., **Cornell, S.**, Choi, S., Lee, Y., Kim, B. Y., & Watanabe, S. (2023). NEURAL SPEECH ENHANCEMENT WITH VERY LOW ALGORITHMIC LATENCY AND COMPLEXITY VIA INTEGRATED FULL- AND SUB-BAND MODELING. *ICASSP 2023.*
26. L. Della Libera, C. Subakan, M. Ravanelli, **S. Cornell**, F. Lepoutre and F. Grondin, "Resource-Efficient Separation Transformer," *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.
27. **Cornell, S.**, Jung, J. W., Watanabe, S., & Squartini, S. (2024, April). One Model to Rule Them All? Towards End-to-End Joint Speaker Diarization and Speech Recognition. In *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 11856-11860). IEEE.
28. Hwang, J., Hira, M., Chen, C., Zhang, X., Ni, Z., Sun, G., ... **Cornell, S.**...& Tao, Y. (2023, December). TorchAudio 2.1: Advancing speech recognition, self-supervised learning, and audio processing

components for PyTorch. In *2023 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)* (pp. 1-9). IEEE.

29. Aironi, C., **Cornell, S.**, Gabrielli, L., & Squartini, S. (2023, October). A Score-aware Generative Approach for Music Signals Inpainting. In *2023 4th International Symposium on the Internet of Sounds* (pp. 1-7). IEEE.
30. Masuyama, Y., Chang, X., Zhang, W., **Cornell, S**., Wang, Z. Q., Ono, N., ... & Watanabe, S. (2023, October). Exploring the Integration of Speech Separation and Recognition with Self-Supervised Learning Representation. In *2023 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)* (pp. 1-5). IEEE.
31. Aironi, C., **Cornell, S.**, Serafini, L., & Squartini, S. (2023, September). A time-frequency generative adversarial based method for audio packet loss concealment. In *2023 31st European Signal Processing Conference (EUSIPCO)* (pp. 121-125). IEEE.
32. **Cornell, S.**, Wiesner, M., Watanabe, S., Raj, D., Chang, X., Garcia, P., ... & Khudanpur, S. The CHiME-7 DASR Challenge: Distant Meeting Transcription with Multiple Devices in Diverse Scenarios. *CHiME-7 Workshop 2023.*
33. **Cornell, S.**, Wang, Z. Q., Masuyama, Y., Watanabe, S., Pariente, M., Ono, N., & Squartini, S. (2023, June). Multi-channel speaker extraction with adversarial training: The WAVLAB submission to the Clarity ICASSP 2023 grand challenge. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1-2). IEEE.
34. Silvio Osimi, Leonardo Gabrielli, **Samuele Cornell**, Stefano Squartini. (2024). Equalizing Loudspeakers in Reverberant Environments Using Deep Convolutive Dereverberation. *International Conference on Digital Audio Effects (DAFx) 2024*.
35. Someki, M., Choi, K., Arora, S., Chen, W., **Cornell, S.**, Han, J., ... & Watanabe, S. (2024). ESPnet-EZ: Python-only ESPnet for Easy Fine-tuning and Integration. SLT 2024.
36. **Cornell, S.**, Ebbers, J., Douwes, C., Martín-Morató, I., Harju, M., Mesaros, A., & Serizel, R. (2024). DCASE 2024 Task 4: Sound Event Detection with Heterogeneous Data and Missing Labels. DCASE Workshop 2024.
37. **Cornell, S.**, Park, T., Huang, S., Boeddeker, C., Chang, X., Maciejewski, M., ... & Watanabe, S. (2024). The chime-8 dasr challenge for generalizable and array agnostic distant automatic speech recognition and diarization. CHiME Workshop 2024.
38. **Cornell, S.**, Darefsky, J., Duan, Z., & Watanabe, S. Generating Data with Text-to-Speech and Large-Language Models for Conversational Speech Recognition. SynData4GenAI Interspeech Workshop 2024.
39. Li, C., **Cornell, S.**, Watanabe, S., & Qian, Y. (2024). Diffusion-based Generative Modeling with Discriminative Guidance for Streamable Speech Enhancement. *SLT 2024.*
40. Zhang, W., Scheibler, R., Saijo, K., **Cornell, S.**, Li, C., Ni, Z., ... & Qian, Y. (2024). URGENT Challenge: Universality, Robustness, and Generalizability For Speech Enhancement. Interspeech 2024.
41. Yan, B., Hamed, I., Shimizu, S., Lodagala, V. S., Chen, W., Iakovenko, O.,..., **Cornell, S.,**... & Watanabe, S. (2025). CS-FLEURS: A Massively Multilingual and Code-Switched Speech Dataset. In *Proc. Interspeech 2025* (pp. 743-747).
42. Bando, Y., **Cornell, S.**, Fukayama, S., & Watanabe, S. (2025, April). Investigation of Spatial Self-Supervised Learning and Its Application to Target Speaker Speech Recognition. In *Proc. of ICASSP*.
43. Shi, J., Cheng, Y., Su, B. H., Shim, H. J., Tian, J., **Cornell, S.**, ... & Watanabe, S. (2025). ARECHO: Autoregressive Evaluation via Chain-Based Hypothesis Optimization for Speech Multi-Metric Estimation. NeurIPS. 2025
44. Tian, J., Shi, J., Chen, W., Arora, S., Masuyama, Y., Maekaku, T., ..., **Cornell, S.**... & Watanabe, S. (2025, April). ESPnet-SpeechLM: An Open Speech Language Model Toolkit. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (System Demonstrations)* (pp. 116-124).

45. Bharadwaj, S., **Cornell, S.**, Choi, K., Fukayama, S., Shim, H. J., Deshmukh, S., & Watanabe, S. (2025). OpenBEATs: A Fully Open-Source General-Purpose Audio Encoder. WASPAA 2025.
46. Saijo, K., Zhang, W., **Cornell, S.**, Scheibler, R., Li, C., Ni, Z., ... & Watanabe, S. (2025). Interspeech 2025 URGENT Speech Enhancement Challenge. Interspeech 2025.
47. Zhang, W., Saijo, K., **Cornell, S**., Scheibler, R., Li, C., Ni, Z., ... & Qian, Y. (2025). Lessons Learned from the URGENT 2024 Speech Enhancement Challenge. Interspeech 2025.
48. Sheikh, Z., Shimizu, S., Arora, S., Shi, J., **Cornell, S.**, Li, X., & Watanabe, S. (2025). Scalable Spontaneous Speech Dataset (SSSD): Crowdsourcing Data Collection to Promote Dialogue Research. In *Proc. Interspeech 2025* (pp. 3963-3967).
49. Li, C., Zhang, W., Wang, W., Scheibler, R., Saijo, K., **Cornell, S**., ... & Qian, Y. (2025). Less is More: Data Curation Matters in Scaling Speech Enhancement. *ASRU 2025*
50. Wang, J., Li, C., Wang, W., Zhang, W., **Cornell, S.**, Sach, M., ... & Qian, Y. (2025). URGENT-PK: Perceptually-Aligned Ranking Model Designed for Speech Enhancement Competition. *ASRU 2025*
51. *Polok, A., Klement, D., **Cornell, S.**, Wiesner, M., Černocký, J., Khudanpur, S., & Burget, L. (2026). SE-DiCoW: Self-Enrolled Diarization-Conditioned Whisper. ICASSP 2026*
52. *Li, C., Wang, W., Sach, M., Zhang, W., Saijo, K., **Cornell, S.**, ... & Qian, Y. (2026). ICASSP 2026 URGENT Speech Enhancement Challenge. ICASSP 2026*
53. *Ivry, Amir, **Samuele Cornell**, and Shinji Watanabe. "MAPSS: Manifold-based Assessment of Perceptual Source Separation." ICLR 2026*
54. *Polok, A., Han, J., Klement, D., **Cornell, S.,** Černocký, J., & Burget, L. (2025). BUT system for the MLC-SLM challenge. MLC-SLM Interspeech workshop*

55. *Wang, W., Zhang, W., Li, C., Wang, J., **Cornell, S.**, Sach, M., Saijo, K., Fu, Y., Ni, Z., Han, B., Gong, X., Bi, M., Fingscheidt, T., Watanabe, S., & Qian, Y. (2026). UrgentMOS: Unified Multi-Metric and Preference Learning for Robust Speech Quality Assessment. ICASSP 2026.*

56. *Sach, M., Fu, Y., Saijo, K., Zhang, W., **Cornell, S.**, Scheibler, R., Li, C., Ni, Z., Kumar, A., Wang, W., Qian, Y., Watanabe, S., & Fingscheidt, T. (2026). 2025 URGENT Speech Enhancement Challenge Multilingual P.808 Listening Tests: Approach and Results. ICASSP 2026.*

57. *Maciejewski, M., & **Cornell, S.** (2026). Ring Mixing with Auxiliary Signal-to-Consistency-Error Ratio Loss for Unsupervised Denoising in Speech Separation. SLT 2026.*

58. *Someki, M., Polok, A., Carvalho, C., Lin, C., Yang, D., Shi, J., Tian, J., Yalta Soplin, N., **Cornell, S.**, Arora, S., Teixeira, F., Wang, W., Chen, W., Abad, A., Li, C., Watanabe, S., & Zhang, W. (2026). ESPnet3: Infrastructure for Scalable Speech and Audio Research in the Foundation Model Era. Interspeech 2026.*

59. *Tawara, N., **Cornell, S.**, Polok, A., Delcroix, M., Burget, L., & Watanabe, S. (2026). Who Spoke What When? Evaluating Spoken Language Models for Conversational ASR with Semantic and Overlap-Aware Metrics. Interspeech 2026.*

60. *Polok, A., **Cornell, S.**, Udupa, S., Černocký, J., Watanabe, S., & Burget, L. (2026). Grounding Spoken LLMs in Multi-Speaker Audio via Diarization Conditioning. Interspeech 2026.*

61. *Wiesner, M., **Cornell, S.**, Polok, A., Yang, L., Burget, L., & Khudanpur, S. (2026). Modeling Overlapped Speech with Shuffles. Interspeech 2026.*

62. *Maciejewski, M., & **Cornell, S.** (2026). Exploiting Noise Inseparability for Weakly-Supervised Discriminative Speech Denoising Using Noisy Targets. IWAENC 2026.*

63. *Takizawa, D., Nakamura, T., **Cornell, S.**, Chen, W., Fukayama, S., & Watanabe, S. (2026). Dissecting Sensitivity to Training Language in Self-Supervised Speech Learning Using Neural Audio Codec Tokens. Interspeech 2026.*

Other (Challenges/Datasets/Toolkits)

1. Ravanelli, M., Parcollet, T., Plantinga, P., Rouhe, A., **Cornell, S.**, Lugosch, L., ... & Bengio, Y. (2021). SpeechBrain: A general-purpose speech toolkit. *arXiv preprint arXiv:2106.04624*.
2. Cosentino, J., Pariente, M., **Cornell, S.**, Deleforge, A., & Vincent, E. (2020). Librimix: An open-source dataset for generalizable speech separation. *arXiv preprint arXiv:2005.11262*.
3. Sahidullah, M., Patino, J., **Cornell, S.**, Yin, R., Sivasankaran, S., Bredin, H., ... & Barras, C. (2019). The speed submission to DIHARD II: Contributions & lessons learned. *arXiv preprint arXiv:1911.02388*.
4. **Cornell, S**., Pepe, G., Principi, E., Pariente, M., Olvera, M., Gabrielli, L., & Squartini, S. (2020). The univpm-inria systems for the dcase 2020 task 4. *DCASE2020 Challenge, Tech. Rep.* (system description for DCASE 2020 Task 4, won jury award)
5. **Cornell, S.,** Wang, Z. Q., Masuyama, Y., Watanabe, S., Pariente M & Ono, N. (2022). Multi-Channel Target Speaker Extraction with Refinement: The WAVLab Submission to the Second Clarity Enhancement Challenge. *Second Clarity Enhancement Challenge Workshop*
6. *Bharadwaj, S., **Cornell, S.**, Choi, K., Shim, H. J., Deshmukh, S., Fukayama, S., & Watanabe, S. (2026). The CMU-AIST submission for the ICME 2025 Audio Encoder Challenge.*