

Sara Papi

✉ sara95papi@gmail.com

✉ @sarapapi

in sara-papi

🌐 <https://sarapapi.github.io/>

🎓 Google Scholar (h-index 11)



Experience

- 2024 – now 📌 **Researcher in Speech Processing and Multimodal LLMs** (WP Technical Lead in the MEETWEEN Horizon European Project), Fondazione Bruno Kessler (Italy).
- 2023 – 2023 📌 **Research Intern in Streaming Speech Recognition and Translation** (Supervised by Jinyu Li), Microsoft (USA).
- 2019 – 2019 📌 **Research Intern in Automatic Speech Scoring**, Fondazione Bruno Kessler (Italy).

Education

- 2020 – 2024 📌 **Ph.D. cum Laude in Information Engineering and Computer Science**, Fondazione Bruno Kessler & University of Trento (Italy).
- 2017 – 2020 📌 **M.Sc. cum Laude in Computer and Automation Engineering**, University of Siena (Italy).
- 2014 – 2017 📌 **B.Sc. in Information Engineering**, University of Siena (Italy).

Awards and Achievements

Awards

- 2025 📌 **Best PhD Graduate 2024 in Information and Communication Technology**, University of Trento.
- 📌 **PhD Excellence Award**, Fondazione Bruno Kessler.
- 📌 **Highly Commended 2024 Anthony C Clarke Award** for the **EAMT Best Thesis**, EAMT.
- 2024 📌 **Special Mention for the Best Italian PhD Thesis**, AIXIA 2024.

Achievements

- 📌 **Outstanding Paper and SAC Award**, ACL 2024.
- 📌 **Social Impact Paper Award**, EMNLP 2024.

Invited Talks

Seminars/Keynote Talks



- 2025 📌 **SlatorCon Remote**. Talk Title: “How “Real” is your Real-Time Simultaneous Speech-to-Text System?”.
- 📌 **MT Marathon**. Talk Title: “Speech Translation in the LLM Era: Between Old and New Challenges”.

Panels

- 2024 📌 **IWSLT 2024**. Panel Title: “Emerging trends in Speech Translation”.



Mentorship

Advisor





- 2025  **Amirbek Djanibekov** (PhD Student). Internship in Simultaneous Translation Agents for Long-form Speech.
-  **Nishant Singh** (MSc Student). Internship in Subtitle Segmentation using Multimodal LLMs.












Research Publications











Journal Articles

- 1 **S. Papi**, P. Polák, D. Macháček, and O. Bojar, “How “real” is your real-time simultaneous speech-to-text translation system?” *Transactions of the Association for Computational Linguistics*, vol. 13, pp. 281–313, Apr. 2025, ISSN: 2307-387X.  DOI: 10.1162/tacl_a_00740. eprint: https://direct.mit.edu/tacl/article-pdf/doi/10.1162/tacl_a_00740/2513610/tacl_a_00740.pdf.
- 2 **S. Papi**, M. Gaido, A. Karakanta, M. Cettolo, M. Negri, and M. Turchi, “Direct speech translation for automatic subtitling,” *Transactions of the Association for Computational Linguistics*, vol. 11, pp. 1355–1376, Nov. 2023, ISSN: 2307-387X.  DOI: 10.1162/tacl_a_00607. eprint: https://direct.mit.edu/tacl/article-pdf/doi/10.1162/tacl_a_00607/2197786/tacl_a_00607.pdf.

Conference Proceedings


- 1 N. R. Koluguri, M. Sekoyan, G. Zelenfroynd, *et al.*, “Granary: Speech recognition and translation dataset in 25 european languages,” in *Accepted at INTERSPEECH*, 2025.
- 2 T. K. Lam*, M. Gaido*, **S. Papi**, L. Bentivogli, and B. Haddow, “Prepending or cross-attention for speech-to-text? an empirical comparison,” in *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, Albuquerque, New Mexico: Association for Computational Linguistics, Apr. 2025, pp. 2994–3006, ISBN: 979-8-89176-189-6.  URL: <https://aclanthology.org/2025.naacl-long.153/>.
- 3 F. Verdini*, P. Melucci*, S. Perna*, *et al.*, “How to connect speech foundation models and large language models? what matters and what does not,” in *Accepted at INTERSPEECH*, 2025.
- 4 I. S. Ahmad, A. Anastasopoulos, O. Bojar, *et al.*, “FINDINGS OF THE IWSLT 2024 EVALUATION CAMPAIGN,” in *Proceedings of the 21st International Conference on Spoken Language Translation (IWSLT 2024)*, Bangkok, Thailand (in-person and online): Association for Computational Linguistics, Aug. 2024, pp. 1–11.  DOI: 10.18653/v1/2024.iwslt-1.1.
- 5 M. Cettolo, A. Piergentili, **S. Papi**, M. Gaido, M. Negri, L. Bentivogli, *et al.*, “Magnet-machines generating translations: A calamita challenge,” in *Proceedings of the Tenth Italian Conference on Computational Linguistics (CLiC-it 2024)*, 2024.
- 6 M. Gaido, **S. Papi**, M. Cettolo, *et al.*, “Automatic subtitling and subtitle compression: FBK at the IWSLT 2024 subtitling track,” in *Proceedings of the 21st International Conference on Spoken Language Translation (IWSLT 2024)*, Bangkok, Thailand (in-person and online): Association for Computational Linguistics, Aug. 2024, pp. 86–96.  DOI: 10.18653/v1/2024.iwslt-1.13.
- 7 M. Gaido, **S. Papi**, M. Negri, and L. Bentivogli, “Speech translation with speech foundation models and large language models: What is there and what is missing?” In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Bangkok, Thailand: Association for Computational Linguistics, Aug. 2024, pp. 14 760–14 778.  DOI: 10.18653/v1/2024.acl-long.789.

- 8 M. Gaido, **S. Papi**, M. Negri, M. Cettolo, and L. Bentivogli, “SBAAM! eliminating transcript dependency in automatic subtitling,” in *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Bangkok, Thailand: Association for Computational Linguistics, Aug. 2024, pp. 3673–3691.  DOI: 10.18653/v1/2024.acl-long.201.
- 9 M. Gaido*, **S. Papi***, L. Bentivogli, *et al.*, “MOSEL: 950,000 hours of speech data for open-source speech foundation model training on EU languages,” in *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, Miami, Florida, USA: Association for Computational Linguistics, Nov. 2024, pp. 13 934–13 947.  DOI: 10.18653/v1/2024.emnlp-main.771.
- 10 M. Gaido*, **S. Papi***, M. Negri, and L. Bentivogli, “How do hyenas deal with human speech? speech recognition and translation with ConfHyena,” in *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, Torino, Italia: ELRA and ICCL, May 2024, pp. 8184–8191.  URL: <https://aclanthology.org/2024.lrec-main.717/>.
- 11 **S. Papi**, M. Gaido, M. Negri, and L. Bentivogli, “SimulSeamless: FBK at IWSLT 2024 simultaneous speech translation,” in *Proceedings of the 21st International Conference on Spoken Language Translation (IWSLT 2024)*, Bangkok, Thailand (in-person and online): Association for Computational Linguistics, Aug. 2024, pp. 72–79.  DOI: 10.18653/v1/2024.iwslt-1.11.
- 12 **S. Papi**, M. Gaido, M. Negri, and L. Bentivogli, “StreamAtt: Direct streaming speech-to-text translation with attention-based audio history selection,” in *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Bangkok, Thailand: Association for Computational Linguistics, Aug. 2024, pp. 3692–3707.  DOI: 10.18653/v1/2024.acl-long.202.
- 13 **S. Papi**, P. Wang, J. Chen, *et al.*, “Leveraging timestamp information for serialized joint streaming recognition and translation,” in *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2024, pp. 10 381–10 385.  DOI: 10.1109/ICASSP48485.2024.10447565.
- 14 **S. Papi***, M. Gaido*, A. Pilzer, and M. Negri, “When good and reproducible results are a giant with feet of clay: The importance of software quality in NLP,” in *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Bangkok, Thailand: Association for Computational Linguistics, Aug. 2024, pp. 3657–3672.  DOI: 10.18653/v1/2024.acl-long.200.
- 15 B. Savoldi, **S. Papi**, M. Negri, A. Guerbero-f-Arenas, and L. Bentivogli, “What the harm? quantifying the tangible impact of gender bias in machine translation with a human-centered study,” in *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, Miami, Florida, USA: Association for Computational Linguistics, Nov. 2024, pp. 18 048–18 076.  DOI: 10.18653/v1/2024.emnlp-main.1002.
- 16 D. Fucci, M. Gaido, **S. Papi**, M. Cettolo, M. Negri, and L. Bentivogli, “Integrating language models into direct speech translation: An inference-time solution to control gender inflection,” in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, Singapore: Association for Computational Linguistics, Dec. 2023, pp. 11 505–11 517.  DOI: 10.18653/v1/2023.emnlp-main.705.
- 17 M. Gaido, **S. Papi**, M. Negri, and M. Turchi, “Joint speech translation and named entity recognition,” in *Interspeech 2023*, 2023, pp. 47–51.  DOI: 10.21437/Interspeech.2023-1767.
- 18 **S. Papi**, M. Gaido, and M. Negri, “Direct models for simultaneous translation and automatic subtitling: FBK@IWSLT2023,” in *Proceedings of the 20th International Conference on Spoken Language Translation (IWSLT 2023)*, Toronto, Canada (in-person and online): Association for Computational Linguistics, Jul. 2023, pp. 159–168.  DOI: 10.18653/v1/2023.iwslt-1.11.
- 19 **S. Papi**, M. Negri, and M. Turchi, “Attention as a guide for simultaneous speech translation,” in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Toronto, Canada: Association for Computational Linguistics, Jul. 2023, pp. 13 340–13 356.  DOI: 10.18653/v1/2023.acl-long.745.

- 20 **S. Papi**, M. Turchi, and M. Negri, “Alignatt: Using attention-based audio-translation alignments as a guide for simultaneous speech translation,” in *Interspeech 2023*, 2023, pp. 3974–3978.  DOI: 10.21437/Interspeech.2023-170.
- 21 **S. Papi**, P. Wang, J. Chen, J. Xue, J. Li, and Y. Gaur, “Token-level serialized output training for joint streaming asr and st leveraging textual alignments,” in *2023 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, 2023, pp. 1–8.  DOI: 10.1109/ASRU57964.2023.10389715.
- 22 M. Gaido*, **S. Papi***, D. Fucci, G. Fiameni, M. Negri, and M. Turchi, “Efficient yet competitive speech translation: FBK@IWSLT2022,” in *Proceedings of the 19th International Conference on Spoken Language Translation (IWSLT 2022)*, Dublin, Ireland (in-person and online): Association for Computational Linguistics, May 2022, pp. 177–189.  DOI: 10.18653/v1/2022.iwslt-1.13.
- 23 **S. Papi**, M. Gaido, M. Negri, and M. Turchi, “Does simultaneous speech translation need simultaneous models?” In *Findings of the Association for Computational Linguistics: EMNLP 2022*, Abu Dhabi, United Arab Emirates: Association for Computational Linguistics, Dec. 2022, pp. 141–153.  DOI: 10.18653/v1/2022.findings-emnlp.11.
- 24 **S. Papi**, M. Gaido, M. Negri, and M. Turchi, “Over-generation cannot be rewarded: Length-adaptive average lagging for simultaneous speech translation,” in *Proceedings of the Third Workshop on Automatic Simultaneous Translation*, Online: Association for Computational Linguistics, Jul. 2022, pp. 12–17.  DOI: 10.18653/v1/2022.autosimtrans-1.2.
- 25 **S. Papi**, A. Karakanta, M. Negri, and M. Turchi, “Dodging the data bottleneck: Automatic subtitling with automatically segmented ST corpora,” in *Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, Online only: Association for Computational Linguistics, Nov. 2022, pp. 480–487.  DOI: 10.18653/v1/2022.aacl-short.59.
- 26 A. Karakanta*, **S. Papi***, M. Negri, and M. Turchi, “Simultaneous speech translation for live subtitling: From delay to display,” in *Proceedings of the 1st Workshop on Automatic Spoken Language Translation in Real-World Settings (ASLTRW)*, Virtual: Association for Machine Translation in the Americas, Aug. 2021, pp. 35–48.  URL: <https://aclanthology.org/2021.mtsummit-asltrw.4/>.
- 27 **S. Papi**, M. Gaido, M. Negri, and M. Turchi, “Dealing with training and test segmentation mismatch: FBK@IWSLT2021,” in *Proceedings of the 18th International Conference on Spoken Language Translation (IWSLT 2021)*, Bangkok, Thailand (online): Association for Computational Linguistics, Aug. 2021, pp. 84–91.  DOI: 10.18653/v1/2021.iwslt-1.8.
- 28 **S. Papi**, M. Negri, and M. Turchi, “Visualization: The missing factor in simultaneous speech translation,” in *Proceedings of the Seventh Italian Conference on Computational Linguistics (CLiC-it 2021)*, 2021.
- 29 **S. Papi***, M. Gaido*, M. Negri, and M. Turchi, “Speechformer: Reducing information loss in direct speech translation,” in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Online and Punta Cana, Dominican Republic: Association for Computational Linguistics, Nov. 2021, pp. 1698–1706.  DOI: 10.18653/v1/2021.emnlp-main.127.
- 30 **S. Papi**, E. Trentin, R. Gretter, M. Matassoni, and D. Falavigna, “Mixtures of deep neural experts for automated speech scoring,” in *Interspeech 2020*, 2020, pp. 3845–3849.  DOI: 10.21437/Interspeech.2020-1055.

Community Service


Organizer/Committee Member

- 2025  **IWSLT 2025 - Evaluation Campaigns**, Instruction Following (Chair) and Simultaneous Translation (Organizer). Conference.

Community Service (continued)

- 2024  **EACL 2024 Student Research Workshop** (Student Program Chair). Conference.
  **IWSLT 2024 - Evaluation Campaigns**, Subtitling (Organizer). Conference.

Area Chair

- 2025  **ACL Rolling Review**. February and May cycles (ACL, EMNLP). Conference.

Reviewer

-  **NeurIPS 2025**. Conference.
 **CLIC-it 2025**. Conference.
 **Interspeech 2025**. Conference.
 **IWSLT 2025**. Conference.
 **LANS-TTS**. Journal.
 **ICASSP 2025**. Satellite Workshop: Workshop on Speech Pathology Analysis and DEtection (SPADE).
2024  **ACL Rolling Review**. February (ACL), April, and June (EMNLP) cycles. Conference.
  **ICASSP 2024**. Conference.
  **IWSLT 2024**. Conference.
  **Interspeech 2024**. Conference.
2023  **HITLAML 2023**. Conference.
2022  **AMTA 2022**. Conference.