The Workshop on Automatic Assessment of Atypical Speech (AAAS-2025)

Proceedings of the Workshop

March 5, 2025
Tallinn, Estonia

Editors: Mikko Kurimo and Tamas Grosz

Message from the Workshop Chair

The workshop on Automatic Assessment of Atypical Speech (AAAS) explores the assessment of pronunciation and speaking skills of children, language learners, and speakers with speech sound disorders and methods to provide automatic rating and feedback using automatic speech recognition (ASR) and large language models (LLMs). The workshop takes place in Tallinn, Estonia, on March 5th, 2025, in a physical setting, allowing for potential hybrid participation.

Automatic speaking assessment (ASA) is a rapidly growing field that answers to the need for AI tools to self-practise second and foreign language skills. This is not limited to pronunciation assessment, but the AI systems can also provide more complex feedback about fluency, vocabulary and grammar of the recorded speech. ASA is also very relevant for the detection and quantification of speech disorders and for developing speech exercises that can be performed independently at home. The important applications of processing non-standard speech also include interfaces for children and elderly speakers as an alternative to using text input and output. The topic is timely, because the latest large speech models allow us now to develop ASR and classification methods for low-resourced data, such as atypical speech, where annotated training datasets are rarely available, expensive and difficult to transcribe, rate and share.

The idea to organize this workshop came during the last year of a 4-year long research project TEFLON with partners from Finland, Sweden and Norway. The project has been funded by NordForsk's programme for multidisciplinary research collaboration in Nordic countries and it focused on gamified pronunciation training and assessment for children learning Nordic languages. The goal of this workshop is to present the latest results in the field of ASA and discuss the future work and collaboration between the researchers in Nordic and Baltic countries.

In the call for papers, we invited students, researchers, and other experts and stakeholders to contribute papers and/or join the discussion on the following (and related) topics:

- Automatic speaking assessment (ASA) for L2 (second or foreign language) pronunciation
- ASA for spoken L2 proficiency
- ASA for speech sound disorders (SSD)
- Automatic speech recognition (ASR) for L2 learners

- ASR for children and young L2 learners
- ASA and ASR for Nordic and other low-resource languages and tasks
- Spoken L2 learning and speech therapy using games
- Automatic generation of verbal feedback for spoken L2 learners using LLMs

In total 7 submissions were received of which 4 were archival submissions. The programme committee (PC) consisted of 27 members (excluding the 3 program chairs), who served as reviewers providing at least 3 reviews for each archival submission. Based on the PC assessments regarding the content, and quality of the submissions, the program chairs decided to accept only 2 submissions for presentation and publication. The non-archival submissions were presentation abstracts of related projects in Finland and Estonia. These 3 project presentation submissions, 2 peer-reviewed research papers, 1 invited presentation from the TEFLON project and 2 other invited talks together compose our workshop programme consisting of 8 talks and a panel discussion.

To complete the programme we invited 3 keynote talks to strengthen the connection of the speaking assessment research to its main application fields: in speech therapy by Nina Benway (University of Maryland, USA) and second language proficiency assessment Ari Huhta (University of Jyväskylä, Finland). For the third keynote talk we invited all the partners of the TEFLON project to briefly present their key results.

Thank you to everyone for being part of AAAS-2025, and I wish you a wonderful workshop!

Mikko Kurimo, Workshop Chair
Espoo
February 2025

Organizers

Workshop Chair

Mikko Kurimo, Aalto University, Finland

Workshop Organizers

Mikko Kurimo, Aalto University, Finland
Giampiero Salvi, NTNU
Sofia Strömbergsson, Karolinska Institutet
Sari Ylinen, Tampere University
Minna Lehtonen, University of Turku
Tamas Grosz, Aalto University
Ekaterina Voskoboinik, Aalto University
Yaroslav Getman, Aalto University
Nhan Phan, Aalto University

Program Chairs

Mikko Kurimo, Aalto University, Finland
Tamas Grosz, Aalto University, Finland
Simon King, University of Edinburgh, UK

Program Committee

Adriana Stan, Anna Smolander, Catia Cucchiarini, Ekaterina Voskoboinik, Geza Nemeth, Giampiero Salvi, Helmer Strik, Horia Cucu, Hugo Van Hamme, Katalin Mády, Klaus Zechner, Mathew Magimai Doss, Mikko Kurimo (chair), Mikko Kuronen, Minna Lehtonen, Mittul Singh, Nhan Phan, Paavo Alku, Pablo Riera, Ragheb Al-Ghezi, Raili Hilden, Riikka Ullakonoja, Sari Ylinen, Shahin Amiriparian, Simon King (co-chair), Sneha Das, Sofia Strömbergsson, Tamás Grósz (co-chair), Tanel Alumäe, Yaroslav Getman

Invited Talk: What is so hard about AI Speech Therapy? Evidence from Efficacy Trials

Nina R Benway
University of Maryland, College Park

Artificial intelligence (AI) speech therapy systems hold significant potential for individuals seeking to acquire and generalize speech sound motor plans through targeted intervention. Research indicates that approximately 5,000 speech therapy practice trials are required to generalize a newly acquired speech motor plan to continuous speech [1, 2]. However, access to sufficiently intensive intervention remains a challenge worldwide [3-6], with the actual dosage of therapy often falling well below evidence-based recommendations [7-9]. AI speech therapy systems could help bridge this gap by enabling at-home, independent practice and automated feedback that aligns with best-practice intensity levels and therapeutic paradigms [10]. While recent technological advancements have begun to overcome key technical barriers like child speech data scarcity and limited technical transparency, important questions remain regarding the therapeutic efficacy of AI clinicians. High-quality clinical trials are now emerging, offering critical insights into the real-world effectiveness and therapeutic impact of AI-driven speech therapy tools [11, 12].

It is important to critically examine the rigor of these clinical trials and the broader implications they pose for the future of AI speech therapy. Key questions include: What clinical results do developers need to report to show that their systems are fit-for-purpose? How do AI-driven speech analysis and intervention systems compare to human clinician judgment in real-world settings? Which speech errors are most appropriate for automated analysis? What child profiles are best suited for ethical independent speech practice? What preferences and difficulties do users have with regard to AI clinicians? This presentation explores both the promise and limitations of AI-driven speech therapy through the lens of systematic review [13-16], recently completed small-n and randomized efficacy trials [11, 12], and our ongoing randomized controlled trials.

References

[1] Koegel, L.K., R. Koegel, L., and J.C. Ingham, *Programming Rapid Generalization of Correct Articulation through Self-Monitoring Procedures.* Journal of Speech and Hearing Disorders, 1986. 51(1): p. 24-32.

[2] Koegel, R., L., et al., *Within-Clinic versus Outside-of-Clinic Self-Monitoring of Articulation to Promote Generalization.* Journal of Speech and Hearing Disorders, 1988. 53(4): p. 392-399.

[3] Health Workforce Australia, *Speech Pathologists in Focus*, in *Australia's Health Workforce Series*, Department of Health, Editor. 2014.

[4] Verdon, S., et al., *An investigation of equity of rural speech-language pathology services for children: A geographic perspective.* International Journal of Speech-Language Pathology, 2011. 13(3): p. 239-250.

[5] Sugden, E., et al., *Service delivery and intervention intensity for phonology-based speech sound disorders.* Int J Lang Commun Disord, 2018. 53(4): p. 718-734.

[6] Pring, T., et al., *The working practices and clinical experiences of paediatric speech and language therapists: A national UK survey.* International Journal of Language & Communication Disorders, 2012. 47(6): p. 696-708.

[7] Preston, J.L., et al., *A Randomized Controlled Trial of Treatment Distribution and Biofeedback Effects on Speech Production in School-Age Children With Apraxia of Speech.* Journal of Speech, Language, and Hearing Research, 2024. 67(9s).

[8] Kaipa, R. and A.M. Peterson, *A systematic review of treatment intensity in speech disorders.* Int J Speech Lang Pathol, 2016. 18(6): p. 507-520.

[9] Allen, M.M., *Intervention efficacy and intensity for children with speech sound disorder.* J Speech Lang Hear Res, 2013. 56(3): p. 865-77.

[10] Maas, E., et al., *Principles of motor learning in treatment of motor speech disorders.* Am J Speech Lang Pathol, 2008. 17(3): p. 277-98.

[11] Hair, A., et al., *A Longitudinal Evaluation of Tablet-Based Child Speech Therapy with Apraxia World.* ACM Trans. Access. Comput., 2021. 14(1).

[12] Benway, N.R. and J.L. Preston, *Artificial Intelligence Assisted Speech Therapy for /ɹ/ using Speech Motor Chaining and the PERCEPT Engine: a Single Case Experimental Clinical Trial with ChainingAI.* American Journal of Speech-Language Pathology, 2024. 33(5): p. 2461-2486.

[13] Deka, C., et al., *AI-based automated speech therapy tools for persons with speech sound disorder: a systematic literature review.* Speech, Language and Hearing, 2024: p. 1-22.

[14] McKechnie, J., et al., *Automated speech analysis tools for children's speech production: A systematic literature review.* International Journal of Speech Language Pathology, 2018. 20(6): p. 583-598.

[15] Furlong, L., et al., *Mobile apps for treatment of speech disorders in children: An evidence-based analysis of quality and efficacy.* PLOS ONE, 2018. 13(8): p. e0201513.

[16] Chen, Y.-P.P., et al., *Systematic review of virtual speech therapists for speech disorders.* Computer Speech & Language, 2016. 37: p. 98-128.

Invited Talk: Automatic assessment of second/foreign language speaking: Review of developments for examination and teaching/learning purposes

Ari Huhta
University of Jyväskylä

The presentation focuses on describing how automatic assessment of second/foreign language (L2) speech has advanced due to innovations in speech processing, machine learning, and natural language processing (NLP). Modern systems evaluate pronunciation, fluency, prosody, and intelligibility using a combination of acoustic, linguistic, and prosodic features (e.g.[3, 5]).

Advances in artificial intelligence (AI) have led to improvements in assessment accuracy. These advancements have been integrated into commercial applications, including TOEFL®, Pearson's Versant, Duolingo English Test, and AI-driven tutoring systems like ELSA Speak. While English is still the most common language assessed automatically, significant developments are taking place also for many other languages (e.g. [1, 2, 4]).

Despite progress, challenges remain in data scarcity, accent robustness, bias mitigation, and adaptive feedback. Current systems still struggle with diverse L2 accents and ensuring fairness in automated scoring. Future developments are likely to focus on multimodal integration (speech, facial expressions, gestures), explainable AI feedback, and personalized adaptive learning models to improve language learning experiences.

References

[1] Al-Ghezi, R., Voskoboinik, K., Getman, Y., von Zansen, A., Kallio, H., Kurimo, M., Huhta, A., & Hildén, R. (2023). Automatic Speaking Assessment of Spontaneous L2 Finnish and Swedish. *Language Assessment Quarterly, 20*(4-5), 421-444.
[2] Farrús, M. 2023. Automatic Speech Recognition in L2 learning: A review based on PRISMA methodology. Languages 8:242. https://doi.org/10.3390/languages8040242
[3] Wu, X-Y. 2024. Artificial Intelligence in L2 learning: A meta-analysis of contextual, instructional, and social-emotional moderators. System 126, 103498.
[4] Wu, Y. & Shen, X. 2024. The assessment of automated rating of L2 Mandarin prosody in lexical tone recognition and pauses. Speech Prosody, Leiden, The Netherlands. 250-254. https://www.isca-archive.org/speechprosody_2024/wu24_speechprosody.pdf
[5] Zecher, K. & Evanini, K. (Eds.) 2020. Automated speaking assessment. Using language technologies to score spontaneous speech. Routledge.

Table of Contents