Phonetics Colloquium at UdS

on Thursday, **05 February 2026**, 12:00 (**unusual day and time**)

location: building C7.2 in room 5.09

Valentin Kany

Universität des Saarlandes

On the automatic prediction of preschool children's speech fluency

Language Proficiency Assessments (LPAs) of preschool children are ubiquitous in Germany. They are mandatory in most federal states to ensure language learning support for children lagging behind in language development at an early stage [1]. Although speech fluency has proven to be a significant indicator of language proficiency (e.g. [2], [3]), it is not the focus of the LPA methods typically used in Germany.

We recorded spontaneous speech data from children with German as native and second language aged 4–6 years, using WUSCHEL, a game-based automated LPA method [4] and annotated it for various fluency-related phenomena (e.g., filler particles such as uh and uhm, repetitions, overlong pauses). Based on recordings from 10 children (5 L1, 5 L2), 32 listeners evaluated perceived fluency to analyse the relationship between annotated phenomena and fluency ratings, enabling feedback for the recorded children. The analysis revealed a significant negative effect of disfluent pauses and other disfluencies (such as repairs, truncations, repetitions, and lengthenings) in the recordings on the fluency rating. A significant positive effect was found for articulation rate [5].

However, to align with WUSCHEL's concept as an automatic LPA, this evaluation procedure needs to be fully automated to generate feedback immediately after recording a child. We therefore present an approach to automatically predict preschool children's speech fluency within LPAs. The human ratings and the fluency-related acoustic features extracted from the annotations were used to build Cumulative Link Mixed Models (CLMMs) with and without splines to test their ability to predict the human ratings with multiple metrics (Spearman's ρ, MAE, quadratic weighted κ). Results show that a parsimonious linear model already reaches near-human agreement (quadratic weighted $\kappa = 0.65$) and that incorporating non-linear spline effects does not improve predictive accuracy. These findings suggest that relatively simple CLMMs can substitute additional human raters in fine-grained fluency assessment of preschool children, which is a task that is already challenging for trained listeners.

References

[1] Faas, S., Götz, A., & Müller, C. (2021). *Sprachstandsfeststellung, Sprachförderung und sprachliche Bildung*. Berlin: Pädquis Stiftung b.R.

[2] Ginther, A., Dimova, S., & Yang, R. (2010). Conceptual and empirical relationships between temporal measures of fluency and oral English proficiency with implications for automated scoring. *Language Testing 27*(3), pp. 379–399.

[3] Révész, A., Ekiert, M., & Torgersen, E. (2016). The effects of complexity, accuracy, and fluency on communicative adequacy in oral task performance. *Applied Linguistics 37*, pp. 828–848.

[4] Roche, J., Haberzettl, S., Pagonis, G., Jessen, M., & Weidinger, N. (2019). Serious games in der Sprachstandsermittlung. In *Serious games in der Sprachbildung*. Tübingen: Narr Francke Attempto Verlag, pp. 340–358.

[5] Kany, V. (2025). From features to fluency: Predicting perceived speech fluency of preschool children for language proficiency assessments. In *Proceedings of SLaTE 2025*, pp. 118–122.