INTERSPEECH 2023
20-24 August 2023, Dublin, Ireland

Providing interpretable insights for neurological speech and cognitive disorders from interactive serious games

Mario Zusag[1]*, Laurin Wagner*[1]

[1]myReha GmbH, Austria

mario@myreha.ai, laurin@myreha.ai

Abstract

We propose an automated pipeline for robustly identifying neurological disorders from interactive therapeutic exercises, which are gathered via the mobile therapy app *myReha*. The app captures speech and cognitive parameters from over 30.000 tasks in various scenarios. Users get immediate and highly accurate feedback for pronunciation and coherency for language tasks, while voice recordings are fed to a feature extraction pipeline in the backend. These features are then used to construct speech characteristics, which are highly indicative of different neurological disorders, such as acquired aphasia after stroke. The data is visually presented in a web application *nyra.insights*, which allows medical professionals to quickly derive recommendations for treatment and closely monitor outcomes. During the Show and Tell session, users can experiment with the interactive myReha app and experience the real-time speech analysis capabilities via the nyra.insights web platform.

Index Terms: speech analysis, neurological disease classification, interpretable speech biomarkers

1. Introduction

Rehabilitation after neurological deficits is essential for restoring function and improving the quality of life for affected individuals. The efficacy of therapeutic interventions is influenced by their intensity and individualization, which are key determinants of therapy success [1].

Acquired language and cognitive disorders resulting from brain injury are varying in types and characteristics. Accurate classification of these deficits is crucial for implementing effective therapeutic strategies [2]. However, evaluating speech and language subtleties can be time-consuming, and automated approaches using speech recordings could expedite rehabilitation. Previous research has attempted to classify for example aphasia types and other speech pathologies using manual transcripts [3, 4] and various models, such as the XLSR-53 model [5, 6], which extracts language-invariant linguistic features.

We propose a novel approach for not only aphasia but multiple neurological deficits. By combining the outputs of the XLSR-53 and the encoder-decoder Whisper model [7], we apply carefully selected feature extraction methods on aligned and annotated transcripts to build prototypes of healthy speech. These prototypes are used to normalize features for inference, creating interpretable insights for anomalous speech recordings. The automated pipeline achieves human-level accuracy on distinguishing speech recordings from aphasic patients and healthy controls, as well as over 90% accuracy in distinguishing speech from the control group and various aphasia types. Furthermore, this is the first approach that exploits the different decoding mechanisms of CTC and encoder-decoder-based methods to create rich annotations for automatic feature extraction. This fully automated approach can potentially benefit the assessment and rehabilitation of multiple neurological deficits, offering a more efficient and accurate classification method for various speech disorders. Together with the large amount of evidence-based therapeutic exercises of the showcased myReha app and the intuitive visualisation of therapy progress in the nyra.insights web platform, we provide tools for conducting a highly intense and individual neuro-therapy.

Future research should focus on developing and validating innovative approaches to deliver personalized and intensive therapy, with the aim of improving rehabilitation outcomes for patients with neurological deficits.

2. Methods

2.1. The myReha app

The myReha mobile app is a modern and intuitive medical device software for treating speech and cognitive disorders after brain damage. It contains over 30.000 tasks presented in more than 45 evidence-based therapeutic exercises, which are constantly adapting to the user's progress. The app can be used in clinical and home settings and provides highly accurate and individual feedback. For gathering speech samples, the app provides interactive voice-based chats, free image description exercises, audio diary and mood entries as well as instructed pronunciation, syntax and semantic exercises.

2.2. Feature construction pipeline

The pipeline for assessing speech impairments uses a two-phase process: prototype construction and classification. The pipeline utilizes feature extraction methods in three steps. The first step involves constructing acoustic and clean transcripts using XLSR-53 and Whisper speech recognition models. The second step enriches transcripts with annotations for pauses, filler words, and part-of-speech (POS) tags. The third step calculates scores for coherence, fluency, syntax, lexical richness, and pronunciation. For classification, speech recordings are processed through the pipeline to compute the distance between each score and the score prototypes. These distances are then used for training simple machine learning classifiers and conducting experiments. The process is outlined in Figure 1.

2.3. The nyra.insights web platform

The nyra.insights platform provides detailed analyses and smart documentation with seamless integration of myReha data. Treating medical professionals can quickly assess each patient's speech and cognitive capabilities, their therapy progress and whether and how interventions worked in the past. The platform

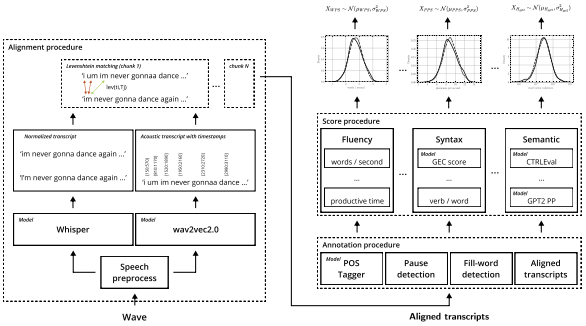

Figure 1: *Prototype pipeline*

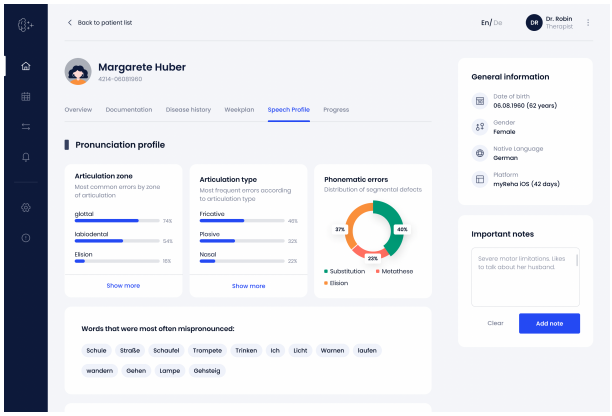

Figure 2: *nyra.insights screenshot*

shows granular speech parameters by averaging the features of different speech domains from the pipeline presented in Section 2.2, which is portrayed in Figure 2.

3. Demo description

We will provide multiple mobile tablets with the myReha app already installed and a larger screen connected to a laptop showcasing the nyra.insights platform. In terms of language we support English and German.

Depending on the type of account used to log into the myReha app, either the patient or clinic version is shown. The clinic version contains an admin dashboard, where medical professionals can quickly switch between already created patient accounts and can add new patient profiles. For the demo, we are creating a clinic account with some dummy profiles. The credentials of this clinic demo account will also be used in the web application. Any changes that are done in one platform immediately transfer to the other.

Users of the Show and Tell demo can create profiles in a matter of seconds or start with existing ones using the myReha mobile app. Via interactive speech exercises through various scenarios, we are gathering speech samples. We deliberately are not providing microphones to showcase the robustness of the transcription and analysis capabilities in real-world settings. Next to real-time pronunciation feedback, accurate transcriptions and coherence checking, we are constructing speech characteristics indicative of different neurological profiles and portray them in nyra.insights. Some of the scores used to construct these characteristics from our pipeline 2.2 need a minimum number of approximately 100 words to be computed robustly. After that amount is gathered, a snapshot of the user's speech characteristics is built by stitching multiple recordings together and feed them through our pipeline.

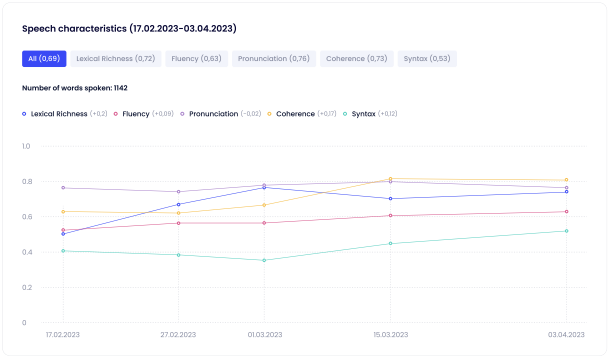

Figure 3: *nyra.insights screenshot*

4. Conclusions

The integration of a mobile app like myReha offering personalized neuro-therapy with speech analysis capabilities, along with a real-time patient progress monitoring web platform such as nyra.insights, transforms rehabilitation for individuals with speech and cognitive impairments. The mobile app delivers tailored therapy while providing advanced speech analysis, enabling accurate assessment of patients' speech patterns and allowing for targeted interventions. The web platform offers healthcare professionals valuable insights to optimize therapy plans and maximize recovery potential. This innovative approach, leveraging speech analysis in a user-friendly context, significantly improves therapy outcomes and redefines rehabilitation for neurological deficits.

5. References

[1] H. Jokinen, S. Melkas, R. Ylikoski, T. Pohjasvaara, M. Kaste, T. Erkinjuntti, and M. Hietanen, "Post-stroke cognitive impairment is common even after successful clinical recovery," *European journal of neurology*, vol. 22, no. 9, pp. 1288–1294, 2015.

[2] J.-F. Landrigan, F. Zhang, and D. Mirman, "A data-driven approach to post-stroke aphasia classification and lesion-based prediction," *Brain*, vol. 144, no. 5, pp. 1372–1383, 2021.

[3] K. C. Fraser, J. A. Meltzer, N. L. Graham, C. Leonard, G. Hirst, S. E. Black, and E. Rochon, "Automated classification of primary progressive aphasia subtypes from narrative speech transcripts," *cortex*, vol. 55, pp. 43–60, 2014.

[4] D. Fromm, J. Greenhouse, M. Pudil, Y. Shi, and B. MacWhinney, "Enhancing the classification of aphasia: a statistical analysis using connected speech," *Aphasiology*, vol. 36, no. 12, pp. 1492–1519, 2022.

[5] A. Conneau, A. Baevski, R. Collobert, A. Mohamed, and M. Auli, "Unsupervised cross-lingual representation learning for speech recognition," *arXiv preprint arXiv:2006.13979*, 2020.

[6] G. Chatzoudis, M. Plitsis, S. Stamouli, A.-L. Dimou, A. Katsamanis, and V. Katsouros, "Zero-shot cross-lingual aphasia detection using automatic speech recognition," 2022. [Online]. Available: https://arxiv.org/abs/2204.00448

[7] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust speech recognition via large-scale weak supervision," 2022. [Online]. Available: https://arxiv.org/abs/2212.04356