

## Patient-friendly speech recognition feedback for aphasia patients

Sandra Wyss, Eliane Rickert, Sven Altermatt, Katrin Petra Kuntner, Markus Degen, Claire Reymond, Sandra Widmer, Anja Blechschmidt, Simone Hemm  
*University of Applied Sciences and Arts Northwestern Switzerland, Switzerland*

Prior studies have shown that personalized, tablet-based home practice is effective in the recovery of word finding difficulties in aphasia patients. Currently existing applications for such speech practice rarely include feedback or are not based on a recorded utterance. To improve training motivation, speech recognition might be used for the assessment of naming performance. In order to decrease incorrect feedback a speech recognition-based method is introduced avoiding misstatements, especially in case the spoken word is not determined with full confidence.

A speech recognition model based on Kaldi ASR toolkit was built with a lexicon of 32 standard German words from 162 healthy subjects (2352 audios). It was evaluated with a 10-fold cross-validation. The model-output gives the words back with the Kaldi-probability. To test false-positive and true-negative rates, the results of the cross-validation and 4352 additional audio-recordings of different German words were used.

A third response option “unsure” was introduced in the decision algorithm besides “correct” and “wrong”. For the categorization, the best three Kaldi-probabilities are considered. If they contain the searched word and if the corresponding probability is higher than a defined threshold, it is categorized as correct otherwise as unsure.

Thresholds between 0.60 and 0.99 have been investigated with an increment of 0.01. Results were compared to binary classification.

The binary classification resulted in an overall word-error-rate (WER) of 4.39%. For the different thresholds 0.65, 0.80, 0.99 with the ternary classification the WER was 2.49%, 2.39%, 2.03% and the unsure-rate 4.0%, 4.59%, 7.25%. This resulted in a decrease of the WER compared to the binary classification of 43.3%, 45.6%, 53.8%.

Between a threshold of [0.65, 0.74] the number of false-positives moved into the unsure group was constant, while the number of moved true-positives increased. Above a threshold of 0.70, the number of moved false-positives was higher compared to the moved true-positives.