

## Prediction of human listeners' speech recognition performance based on automatic speech recognition

Mahdie Karbasi<sup>(1)</sup>, Dorothea Kolossa<sup>(2)</sup>

<sup>(1)</sup>Ruhr University Bochum, Germany, mahdie.karbasi@rub.de

<sup>(2)</sup>Ruhr University Bochum, Germany, dorothea.kolossa@rub.de

### Abstract

There have been many past and recent efforts to improve the accuracy of instrumental measures in predicting the intelligibility of speech signals. However, modeling human auditory perception and, more specifically, human speech recognition, is a challenging task and it is still difficult to achieve a reliable prediction that is consistently close to human listening results. In addition, most of the intelligibility measures introduced so far are relying on the information acquired using both, the distorted signal and its correspondent clean version, which constitutes a serious drawback for such systems. Automatic speech recognizers have recently been introduced as a useful tool for predicting the intelligibility directly, without requiring the clean version of the distorted input signal. In this work, we will consider some discriminative measures that can be extracted utilizing the models trained for an ASR and investigate the accuracy of these measures in predicting the human listening test outcomes. The ASR-based discriminative measures is computed for small units of speech like words or phonemes and can be used to predict the intelligibility of such units of speech. This study will present a detailed analysis of the performance of the proposed model-based measures in comparison to some well-known signal-based intelligibility measures.