

## Automatic Analysis of Speech Prosody in Dutch

*Na Hu<sup>1</sup>, Berit Janssen<sup>2</sup>, Judith Hanssen<sup>3</sup>, Carlos Gussenhoven<sup>4</sup>, Aoju Chen<sup>1</sup>*

<sup>1</sup> Utrecht University, the Netherlands

<sup>2</sup> Digital Humanities Lab, Utrecht University, the Netherlands

<sup>3</sup> Avans University of Applied Sciences, the Netherlands

<sup>4</sup> Radboud University, the Netherlands

{n.hu, b.d.janssen, aoju.chen}@uu.nl, judithhanssen@gmail.com, c.gussenhoven@let.ru.nl

In this talk we present the first publicly available tool for automatic analysis of speech prosody (AASP) in Dutch. Incorporating the state-of-the-art analytical frameworks, AASP enables users to analyze prosody from two different theoretical perspectives. Structurally, AASP analyzes prosody in terms of prosodic events within the auto-segmental metrical framework, hypothesizing prosodic labels in accordance with Transcription of Dutch Intonation (ToDI). Holistically, by means of the Functional Principal Component Analysis (FPCA) AASP generates mathematical functions that capture changes in the shape of a pitch contour. Regarding ToDI, AASP performs four tasks including pitch accent detection, pitch accent classification, prosodic boundary detection, and prosodic boundary tone classification. Using SVM, AASP performs with accuracy comparable to similar tools for other languages for pitch accent detection, prosodic boundary detection, and prosodic boundary tone classification. Notably, we have found that by combining functional features extracted from FPCA with conventional acoustic features, AASP can attain a higher accuracy for pitch accent classification (76.87%) than AuToBI for English using conventional acoustic features (71.6%). Regarding FPCA, AASP outputs the weights of principal components that capture core variations in the shape of pitch contours in a .csv file, which can be directly used for further statistical analysis.

Published as a Docker container, AASP can be set up on various operating systems in only two steps. Moreover, the tool is accessed through a graphic user interface, making it accessible to users with limited programming skills. It has also the potential to be adapted for prosodic analysis in other languages.