

Dissertation press release

30.11.2020

# Analysis and conversion of speech signals using machine learning

<b>Title of the dissertation</b>	Machine learning methods for suprasegmental analysis and conversion in speech.
<b>Contents of the dissertation</b>	Speech technology is a field of technological research focusing on methods to process spoken language. Work in the area has largely relied on a combination of speech-specific scientific knowledge and signal processing and statistical models tailored to process speech data. In this context, machine learning (ML) has played a central role in estimating the parameters of such statistical models. The goal of this thesis is to investigate the applicability of recent developments in ML to the modelling and processing of speech at the so-called suprasegmental level—a level that corresponds to speech patterns spanning across multiple individual speech sounds (phones). The work tackles the following topical problems in speech research: 1) zero-resource speech processing, where the aim is to learn language patterns from speech without access to annotated ML training datasets, 2) automatic word and syllable count estimation which focus on quantifying the amount of linguistic content in audio recordings, and 3) speaking style conversion, which deals with the conversion of the speaking style of an utterance while retaining the linguistic content, speaking identity and sound quality. This thesis analyses several ML methods as well as develops novel ML architectures that outperform competing methods.
<b>Field of the dissertation</b>	Speech and Language Technology.
<b>Doctoral candidate</b>	Shreyas Seshadri, M.Sc. Born in Chennai, India on 25.03.1991.
<b>Time of the defence</b>	18.12.2020 at 12:00
<b>Place of the defence</b>	Aalto University, School of Electrical Engineering, Link to the Zoom-event: <a href="https://aalto.zoom.us/j/62070164113?pwd=bTV4cFZXRmcyUExESi81bGk3Nm45Zz09">https://aalto.zoom.us/j/62070164113?pwd=bTV4cFZXRmcyUExESi81bGk3Nm45Zz09</a>
<b>Opponent</b>	D.Sc. Ville Hautamäki, University of Eastern Finland, Finland.
<b>Custos</b>	Professor Paavo Alku, Aalto University School of Electrical Engineering, Department of Signal Processing and Acoustics.
<b>Electronic dissertation</b>	<a href="http://urn.fi/URN:ISBN:978-952-64-0167-6">http://urn.fi/URN:ISBN:978-952-64-0167-6</a>
<b>Doctoral candidate's contact information</b>	Shreyas Seshadri, Aalto University, <a href="mailto:shreyas.seshadri@aalto.fi">shreyas.seshadri@aalto.fi</a> , +358 504765155.