

Odette Scharenborg

Bias in Automatic Speech Recognition

Automatic speech recognition (ASR) is increasingly used, e.g., in emergency response centers, domestic voice assistants, and search engines. Because of the paramount relevance spoken language plays in our lives, it is critical that ASR systems are able to deal with the variability in the way people speak (e.g., due to speaker differences, demographics, different speaking styles, and differently abled users). ASR systems promise to deliver objective interpretation of human speech. Practice and recent evidence however suggests that the state-of-the-art SotA ASRs struggle with the large variation in speech due to e.g., gender, age, speech impairment, race, and accents. Many factors can cause the bias of an ASR system. The overarching goal in our project is to uncover bias in ASR systems to work towards proactive bias mitigation in ASR. In this talk, I will provide an overview of possible factors that can cause this bias; I will present systematic experiments aimed at quantifying the bias of state-of-the-art ASRs on speech from different groups of speakers; and I will present analyses that shed light onto the cause of this bias. In our work, we focus on bias against gender, age, regional accents and non-native accents.