Title of Work:

Applications in accessibility of text-to-speech synthesis for South African languages: Initial System integration and user engagement

Conference:

Proceedings of the South African Institute of Computer Scientists and Information Technologists

Rationale to ensure venue quality:

There are several metrics that help ensure venue quality. First, this conference has been running since 1985. While this is not always necessarily an indicator of quality, I feel that we can be reasonably confident that a conference that has been running for over 30 years has some credibility and validity to its name. Additionally, this journal is an "accredited specialist ICT academic journal" backed by South African governmental support. Finally, this venue is supported by the ACM, with its papers and proceedings being published in the ACM Digital Library. The combination of these factors indicates that this venue is sufficiently reputable and of sufficient quality.

Problem Statement:

Users with certain disabilities struggle with barriers to information access and communication. Assistive technologies work to mitigate these issues in various capacities: screen readers convey information to vision-impaired users, captioning systems aid hearing impaired users in understanding information, and many more. In particular, Human Language Technology plays an important role in the accessibility domain, as many users are capable of speech but have some condition that otherwise restricts them from interacting with a system. The authors explore several use cases in HLT Technology, specifically focusing on offering accessibility to literate South African users.

Paper Synopsis:

In this study, the authors examine three use cases where multilingual individuals using some form of Augmentative and alternative communication were observed to measure a "baseline integration of the existing Qfrency TTS voices into a selected AAC system and to evaluate the user experience." Grid 3, an AAC system sold and commonly used in South Africa was integrated with the Qfrency TTS voices and customized to build text interfaces with simple South African sentences. Literate AAC users were recruited to perform acceptance testing on this new tool, focusing on how natural and intelligible the TTS voices were when using the application. Users were asked to utilize closed-form answering machines to rank the prosody, pronunciation, and intelligibility of the system. Intelligibility was scored high fairly consistently, but naturalness ratings were "more spread out between the two poles of robotic and human-like synthetic speech." This shows that

current technology is capable of making systems that can be understood, there are still steps to be taken to improve how natural utilizing such a system feels.

Future Work:

As I've referenced in many of my papers, the study did not make any references towards identifying the appropriateness of a given data type as a variable in the success or failure of the application. This motivates a line of inquiry into determining how the use of these different data types could affect the naturalness of NLP, NLG, and TTS Systems. Additionally, the paper notes that while all users were literate, the actual levels of literacy were relatively low. I would like to see a study that focuses on analyzing the intelligibility and naturalness ratings across users with various levels of literacy. This could aid in drawing some correlation to a user's education or familiarity with a language and their preferences in NLG and TTS Applications.