Implementing high level Xhosa text-to-speech synthesis in mobile learning systems

Justus Chr. Roux

Stellenbosch University – South Africa

This presentation reflects on

- *results obtained from a joint international research project on speech synthesis for selected African languages, and*
- *on the implementation of a synthesizer in a mobile learning system as an aid, inter alia, to non-literates.*

Given the traditional approach of "unit selection" in generating speech from text, an alternative approach based on Hidden Markov Modeling (HMM) will be exemplified. This approach is not dependent on formalized rules of intonation, but derives intonation contours from a restricted set of speech data through statistical processes. In setting up a database for synthesis for a particular language, specific attention is however paid to include different sentence types, i.e. grammatical and morpho-syntactic information. Parts-of-speech (POS) tagging is required to delineate particular word types (and their concomitant tonal make-up), whilst experimentation with the assignment of syllable counts, and assigning an H to ante-penultimate syllables, contribute to prosodic phrasing.

The second part of the presentation focuses on embedding a Xhosa synthesizer onto a mobile phone or personal digital assistant (PDA) as a component of a mobile learning system. To the best of our knowledge this is the first application of its kind that incorporates an African language as a speech output option on a mobile phone. The presentation will include a demonstration of the system.