

EEE 6110 Speech Processing.

Dr. Ciira Maina
ciira.maina@dkut.ac.ke

17th April, 2019

Text to speech synthesis

- ▶ Converting text to speech
- ▶ The speech should be as naturally sounding as possible
- ▶ Consists of text analysis and speech synthesis

Text Analysis

- ▶ Determine the phonemes to be spoken
- ▶ Determine where to place pauses, speaking rate, where to place emphasis
- ▶ Resolve ambiguities - words spelt the same but pronounced differently, foreign words
- ▶ Text analysis depends on a dictionary of word pronunciations

Text Analysis

See Figure 8.2 Rabiner and Schafer

Text Analysis

- ▶ Document structure detection
- ▶ Text normalisation
- ▶ Linguistic Analysis
- ▶ Phonetic Analysis
- ▶ Homograph disambiguation
- ▶ Letter to sound conversion
- ▶ Prosodic Analysis

Speech Synthesis

- ▶ Word concatenation - ignores co-articulation
- ▶ Unit selection methods - use prerecorded natural speech
 - ▶ Words
 - ▶ Syllables
 - ▶ Phonemes
 - ▶ demi-syllables
 - ▶ diphones

Festival Demo <http://www.festvox.org/festival/>

Readings

- ▶ HAH - Chapter 8
- ▶ RS - Chapter 14