

Category: 04  
School of Computer Engineering

Student: CHAITANYA MALAVIYA

URECA  
Undergraduate Research Experience on Campus

Project ID: SCE14065

## Input Text

The quick brown fox jumps over.....

### Existing Research

- TTS : Generation of synthesized speech from text
- Two components of a TTS system :
  - **Text Processing** : Processing of input text and generation of smaller segments like words, phonemes etc.
  - **Speech Synthesis**: Synthesis of segments by selection of a unit from a large speech corpus
- 2 primary quality parameters of a TTS system: Intelligibility and naturalness of synthesized speech

### Applications

- Popular examples of TTS systems: Microsoft Sam, Mac's VoiceOver, ChromeVox
- Communication aid for disabled and dyslectic people
- Car navigation systems and transit schedule announcements
- Spoken e-books, mobile gaming and car navigation systems
- Natural user interface for complex consumer products

## Speech



**Linguistic Analysis =**

**Phrasing**

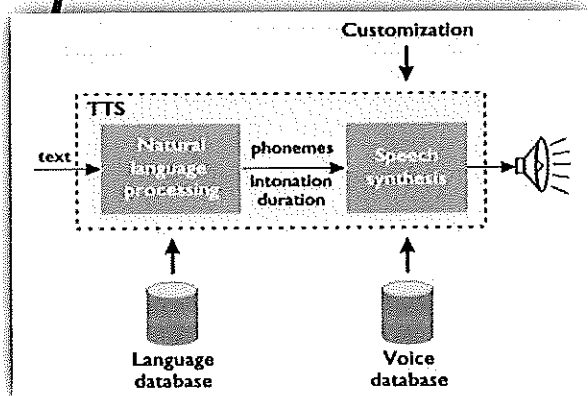
**Intonation**

**Duration**

## TEXT TO SPEECH



### Customization



## Waveform Generation



### Methodology and approach to problem

- **Goal** : To build a TTS system which takes customized input and synthesizes speech using a stored database
- **Approach**: Using a **concatenative TTS Synthesis**: concatenation of segments of recorded speech from a stored database
- Implementing a **unit selection algorithm** : Index of units is created in database with pitch, duration, and position information.
- The target utterance is selected at runtime by selecting the best unit using a decision tree.

### Conclusions and Future Steps

- Implement **dynamic time warping** algorithm to compare the prosodies and phonetics of the desired speech from the database and the required speech
- Some more challenges: Text normalization, emotional quality of generated speech.
- **Selection cost**: During unit selection, 2 costs should be defined in synthesis: unit cost (how close a selected unit to the desired unit) and connection cost (degree of continuity between the selected units).

## Segments

Project Title: Text to Speech Application  
Supervisor: Assoc Prof Chng Eng Siong

Co-Supervisor\*/Collaborator\*: XXX

www.ntu.edu.sg/ureca