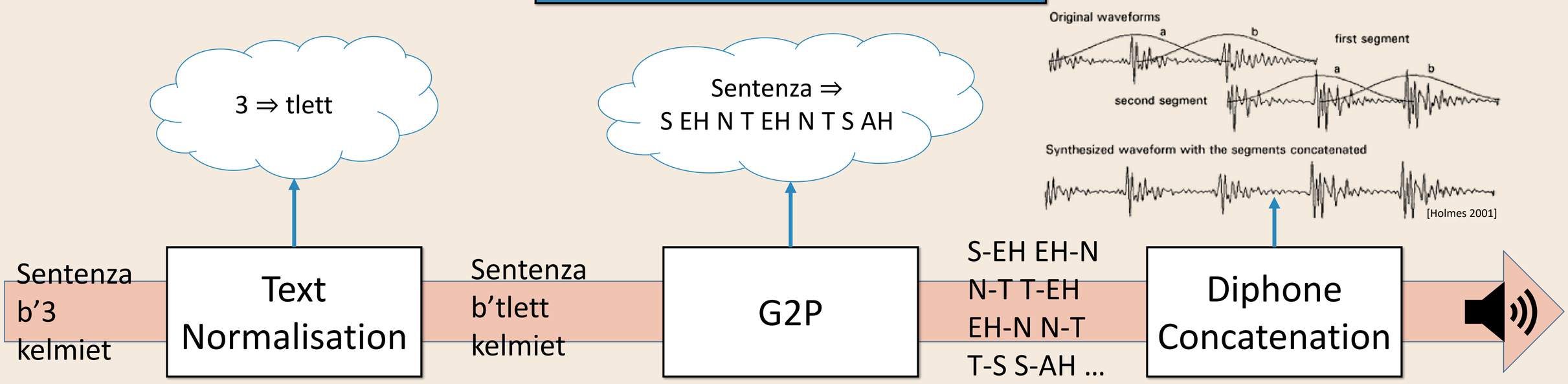
A Diphone-Based Maltese Speech Synthesis System

By Daniel Magro
Supervised by Dr Claudia Borg and Dr Andrea DeMarco

Aim

The principal aim of this thesis is to develop a speech synthesis system which given any Maltese text, can produce an intelligible audio output for the vision-impaired and illiterate (totalling to 8.3% of Maltese Speakers).

Design of the Speech Synthesiser



Building the Diphone Database

Start with a list of desired Diphones: AH-AH, AH-AO, AH-EH, AH-IH, AH-UH, AH-JH, …
A Corpus of word, phonemic transcription pairs: kaotiku – [K AH AO T IH K UH], ħuta – [HH UH T AH], dawl – [D AH UH L], …

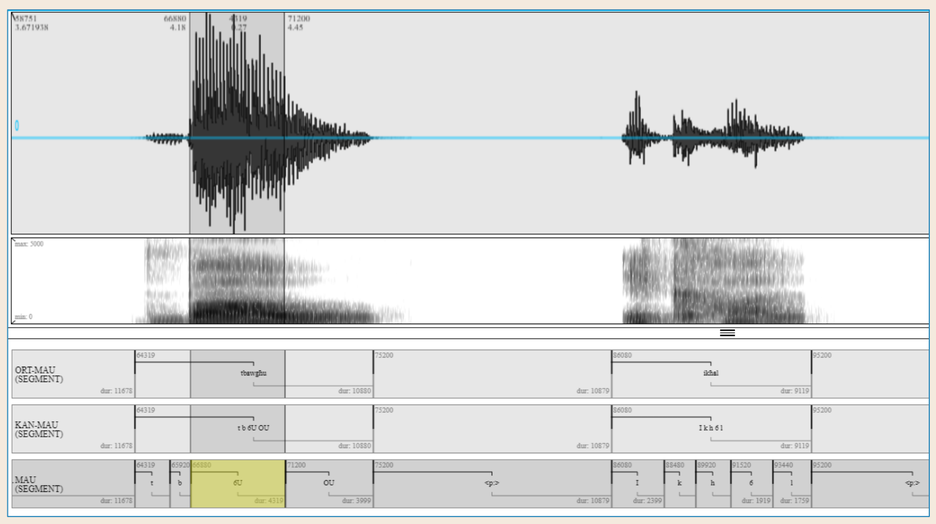
Choose a Carrier Word for each Diphone:

AH-AO – kaotiku
AH-UH – dawl
AH-JH – raġel
⋮

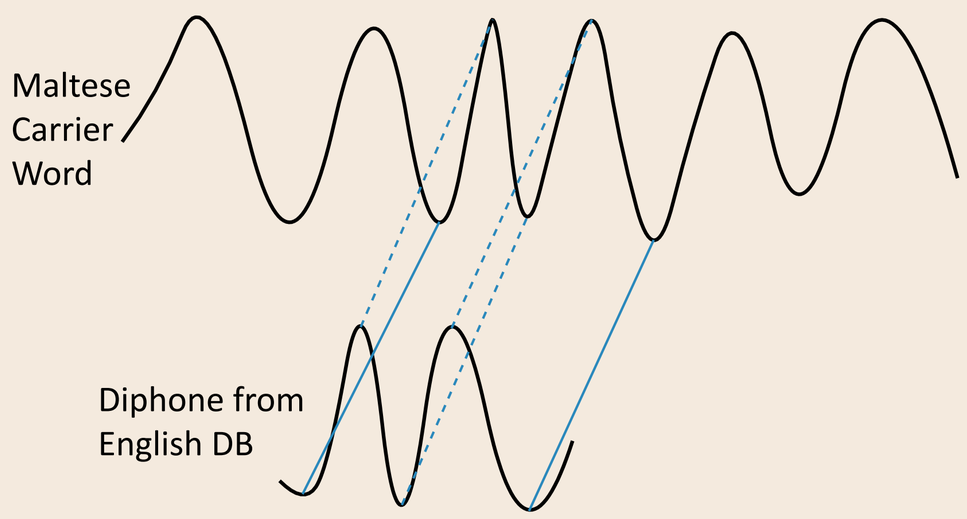
} Record every carrier word

Diphones are Extracted from the Carrier Words:

Manually, by listening for Diphone Boundaries

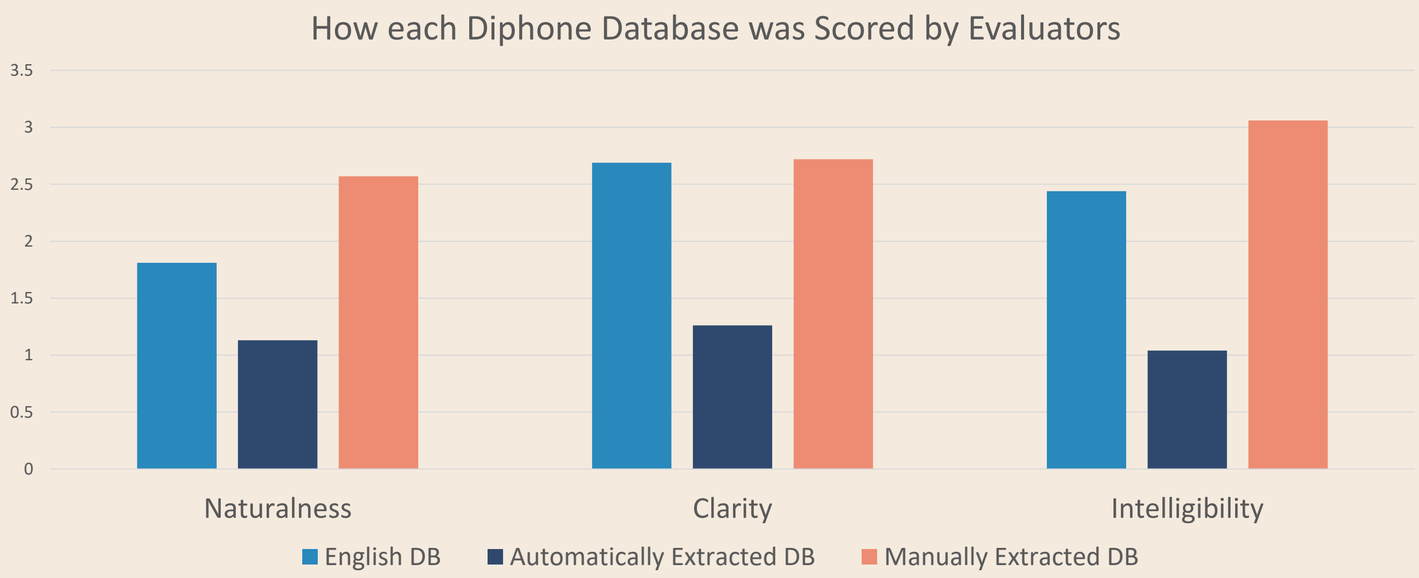


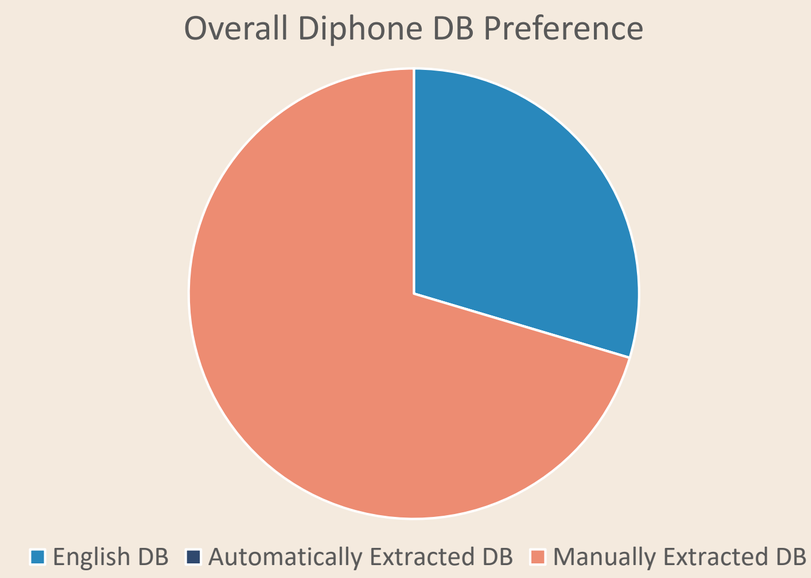
Automatically, using the DTW Algorithm



Results and Conclusions

A set of phrases were synthesised using each of the developed Diphone Databases, and presented to evaluators to score. From the responses, the Diphone DB made by manually extracting diphones from Maltese Carrier Phrases outperformed a professionally recorded English Diphone DB in all criteria.





Future Work

1. *Improve the Text Normalisation component to handle more Non-Standard Word types.*
2. *Refine the Diphone Concatenation Algorithm so that the flow of speech sounds more natural.*
3. *Extract more Diphones to obtain greater coverage of Maltese.*
4. *Develop a web-plugin so that users can have the contents of a website read out to them by issuing a voice command.*