Utkal University (in India)

Name of the Technology: TEXT-TO-SPEECH FOR ORIYA LANGUAGE

Nature of the Technology : APPLICATION & RESEARCH ORIENTED

Level: (Product / Technology / Sub-system): TECHNOLOGY & PRODUCT

Technical Description of the Technology / Product including Basic block diagram, Algorithm used, O/S used, Front-end / user interface, and Specification of the Technology / Product: ANNEXURE-I

Representative Snapshot/screenshot of the Technology / Product: ANNEXURE-II

Scalability / Portability / Expandability: ALL

Readiness of Transfer of Technology (ToT): YES

Availability of documentation: YES

Testing of the Product / Technology: SUBMITTED TO MCIT

IPR / Open-source: IPR - SW1179/2003

Potential beneficiaries: *COMMON MAN, BLIND & ILLITERATES*

User-agency tie-up: *MODULAR INFOTECH, PUNE*

Name and address of the Resource Person:
Dr. (Mrs) Sanghamitra Mohanty
RC-ILTS-ORIYA,
Department of CSA,
Utkal University,Bhubaneswar - 751004

ANNEXURE - I

TTS system provides an interface through which a user enters certain text/document and it is the software that reads it as natural as a human. The basic approach followed here is, first to analyse the document (language, font etc.), and then extract words from the text, try to parse individual words into vowels and consonants respectively. Then corresponding to these vowels and consonants existing (previously stored in the database) ".wav" files are concatenated and played.

Technologies Behind :

- Creating the wave file database: - For creation of such a database we studied a lot of recorded words and sentences and try to break them into vowels and consonants by minute hearing. Then we analyse those cut pieces and store the appropriate and generalised form in the database.
- Extraction of exact words of a given sentence:- The same words in different sentences have different stress due to its position in the sentence. Appropriate hidden vowels are detected form the words extracted. The format of vowel and consonant break point is shown in the figure 1.
- Choosing of appropriate '.wav' file from the database.

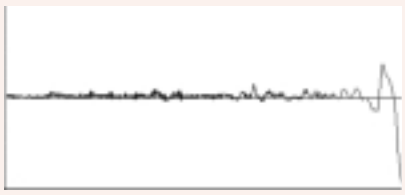
Considering the above example, may be the vowels we get after parsing are the same as '@' (for Oriya), but it is not exactly the same 'a.wav' we concat in every case. Thus, we analyse vowels broadly in three categories as ma:tra: in

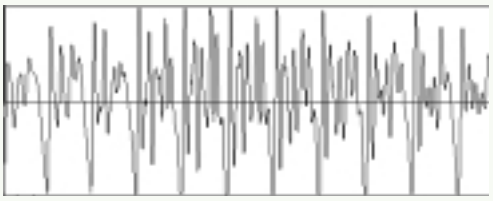
- Beginning
- Middle
- End

And this is observed that the duration of ma:tra:s say '@' here, varies from each other, i.e. '@' in middle is not the same as that in the end. Again accordingly we need to get the appropriate ".wav" files from the database. As observed in the example the durations of '@'(Oriya) are as follows:

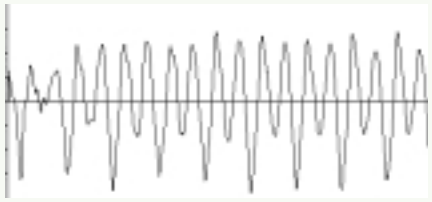
Starting ma:tra	:	- 0.065 sec
Middle ma:tra	:	- 0.105 sec
End ma:tra	:	- 0.116 sec

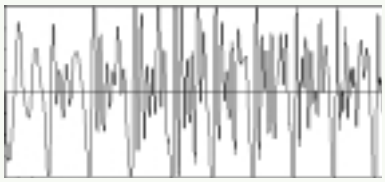
- It is observed that concatenation of the wave files is not that natural as expected. This is due to the certain transitions between the characters in the actual pronunciation. Thus we are developing a robust algorithm for the generation of naturalness in the TTS output.



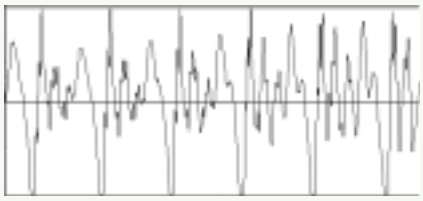


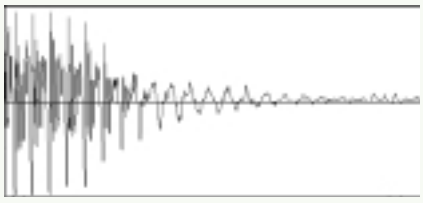
(starting ma:tra)





(middle ma:tra)

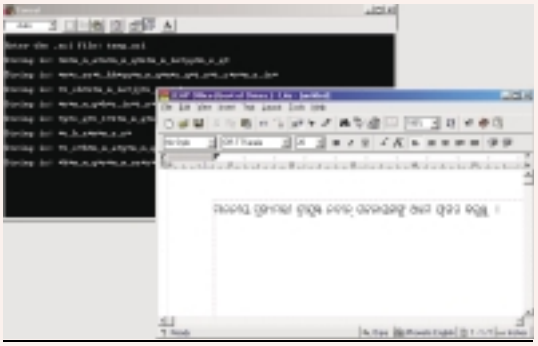




(end ma:tra:)

(Figure - 1)

ANNEXURE - II



(Screen shot of ORITTS peaks the .aci file provided by the user and utters)