
Speech Applications: Development, Status, Challenges and Quality

Ruvan Weerasinghe
(Slides due to Asanka Wasala)

Regional Conference on Localized ICT Development & Dissemination Across Asia
Vientiane, Laos

11th - 16th January, 2009



Language Technology Research Laboratory
University of Colombo School of Computing, Sri Lanka

Speech Applications

Speech Applications

Sinhala Text-to-Speech Engine

- Open-source, platform independent, “*Commercial Grade*”
- Support Screen Readers
- One of the main deliverables of Pan L10N project

Sinhala Speech Recognition System

- In collaboration with Dialog & Telisma
- Commercial Product
- Resources: Speech Corpus, Pronunciation Lexicon



Development

Development : Text-to-Speech System (Strategy)

Diphone Database Construction

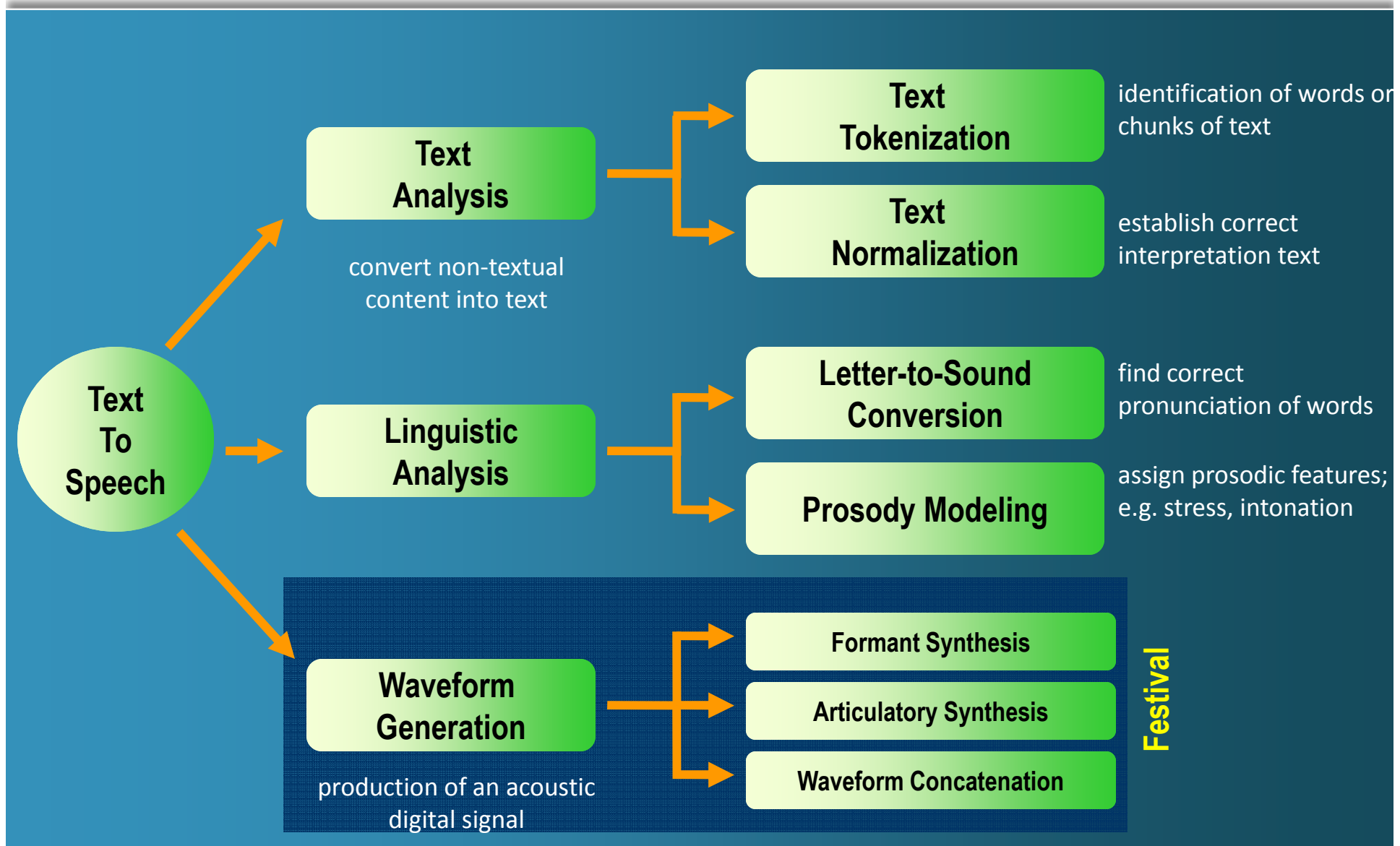
Development of Natural Language Processing Modules

- Definition of Phone set
- **Tokenization and Text Normalization**
- **Letter-to-sound conversion**
- Syllabification & Stress Assignment
- Phrase Breaking Algorithm
- Prosodic Analysis

Integration with different platforms/screen reader software

Evaluation

Development : Text-to-Speech System (Modules Implemented)



Status

Status : Sinhala Text-to-Speech Engine (**Research**)

Publications

NLP Applications of Sinhala: TTS & OCR. The Third International Joint Conference on Natural Language Processing. (IJCNLP). Exhibitions & Demonstration Session. January 7-12, 2008, Hyderabad, India. (2008).

Festival-si: A Sinhala Text-to-Speech System. 10th Conference on Text, Speech and Dialogue (TSD 2007), Plzeň, Czech Republic. (2007).

Facilitating Information Accessibility for the Print Disabled. Diriya 2007 - a conference on "Mainstreaming Disability into Development". Colombo, Sri Lanka. (2007).

Sinhala Grapheme-to-Phoneme Conversion and Rules for Schwa Epenthesis. Proceedings of the COLING/ACL 2006 Main Conference Poster Sessions, Sydney, Australia. pp. 890—897

A Rule Based Syllabification Algorithm for Sinhala. Proceedings of the 2nd International Joint Conference on Natural Language Processing (IJCNLP-05), Jeju Island, Korea. pp. 438-449

Status : Sinhala Text-to-Speech Engine (**Product**)

Sinhala voice can be easily integrated with Linux applications

Most of the Linux distributions now come with Festival pre-installed



Festival-si

Made accessible for screen reader software GNOME-Orca* and Gnopernicus**

Integrated Festival along with Sinhala voice into Microsoft Speech Application Programming Interface (MS-SAPI) .



MS-SAPI compliant Sinhala voice is available via any speech enabled Windows application.

Proved to work well with, screen readers such as Thunder, Non-Visual Desktop Access (NVDA)

* <http://live.gnome.org/Orca>

** <http://www.baum.ro/gnopernicus.html>

Challenges

Challenges

- **Lack of linguistic resources**

Previous studies on phonetics, phonology, syllabification, stress assignment, intonation

- **Lack of a well defined POS tag set, POS Tagger and Tagged Corpus**

- **Lack of a comprehensive lexicon / pronunciation dictionary**

- **Building and Annotation of a Speech Corpus**

 - Designing of prompts/scripts

 - Data collection strategy

 - Dealing with noise (different environments)



Quality

Sinhala Text-to-Speech Engine Won “Most Innovative Product” at the “Infotel 2008”, ICT Exhibition



Sinhala Speech Recognition System: Status

- **Text Corpus Creation**
- **Development of Phonetic Lexicon**
- **Designing Prompt Sheets**
- **Recruitment of the panel of Callers**
- Collection of Calls (Recording)
- Preparation of Speech Data for Testing and Training
- Development of Acoustic Models
- Development and Integration of Recognition System
- Testing and Fine tuning
- Documentation



thank you

