

**Intonation and stress generator for Arabic text-to-speech**  
**Hanady Mansour**  
**Qatar University, Arts and Sciences Faculty (CAS)**  
**hanadyma@qu.edu.qa**

The quality of a Text-To-Speech (TTS) synthesis system depends on the naturalness and the intelligibility of the generated speech. Therefore, an automatic generator of stress and intonation is necessary. We describe an approach to the task of generating speech from MSA text which provides the information required for imposing an appropriate intonation contour and stress placement.[Ramsay & Mansour 2008].

This abstract presents a theoretical and computational framework for building a Prosody generator for Arabic TTS. We developed a Natural Language Process (NLP) engine- which included prosody generator- for converting Arabic text into a phonetically transcribed and prosodically labeled text. The prosodic generator is the final module of the NLP engine that derives the pitch information automatically. We use MBROLA system which is a diphone-base synthesizer to produce the signal waves.

Our TTS system is designed for Modern Standard Arabic (MSA) [Mansour.H 2005]. This system generates speech from unrestricted Arabic text by carrying on the following tasks:

- Assigning diacritics to the written text.
- Automatic phonetic transcription.
- Syntactic analysis in order to assign a global intonation contour.
- Calculating and producing the local pitch contours for sentences.
- Connecting the NLP engine with MBROLA synthesizer

The first stage of computational framework aims to convert NLP output into a narrow phonetic transcription. We provide a syllable transducer for the syllabification process and assigning stress location. By applying a set of Prolog notations, the slopes of both top and bottom line for pitch contour are calculated. The final stage will be devoted to connecting the NLP engine and the phone-into-diphone database for creating a sound file by MBROLA.

Adopting and modifying Rifat's (2003) phonological rules, the marked tone sequence should be started by [HL] tone and is separated by a tune boundary [+] from the following tone where the phonological phrase boundary is allocated when the tune changes and the reader intends to focus on a certain concept.

The goal of the work described above was to use fine-grained linguistic analysis to help produce spoken output which was both intelligible and natural sounding from undiacriticised Arabic text. In order to assess whether we had achieved this goal we carried out a linked pair of experiments based on playing the output of two synthesizers for a set of 23 test sentences to a group of 14 subjects. The details and rationales of these experiments are given in the presentation.