Abstract

Prosody modeling has become the backbone of TTS synthesis systems. Amongst all the prosodic modeling approaches, phonetic methods aiming to predict duration and F0 contour are being very praised, thanks to the development of regression tools, such as neural networks (NN). Besides, parametric representations like Fujisaki model for F0 contour generation help to reduce the problem into the approximation of parameters only. But, prior to the prediction process, text analysis should be carried out first, to select and encode the necessary input features. In our purpose to promote Arabic TTS synthesis, an Integrated Model of Arabic Prosody for Speech Synthesis (IMAPSS) tool has been designed to integrate our developed models for text analysis, NN-based phonemic duration prediction and Fujisaki-inspired F0 contour. Hence, the yielding parameters provide a command file to be read by speech synthesis systems, like MBROLA.
Reference

- Fujisaki, H. 2003 Prosody, information and modeling with emphasis on tonal features of speech, in Proceedings of Workshop on spoken language processing, ISCA-supported event, Mumbai, India.
- Baloul, S. 2003. Développement d’un système automatique de synthèse de la parole à

Index Terms

Computer Science  Signal Processing

Key words

Arabic TTS  prosodic parameters  text analysis
phonemic duration
F0 contour
neural networks
Fujisaki model