

Prosodic Analysis of a German Read Corpus for a CALL System for Rehabilitation Purposes

Zaheer Hussain, Rüdiger Hoffmann

Dresden University of Technology (Germany)

zaheer.hussain@mailbox.tu-dresden.de, ruediger.hoffmann@tu-dresden.de

Abstract

Now-a-days Computer Assisted Language Learning (CALL) systems for rehabilitation purposes involving rhythmic analysis and prosodic analysis which are inter-related have gained much importance in speech synthesis and recognition systems. The main objective is to develop a language learning system by monitoring and evaluating the prosodic variations and content data corresponding to the pronunciation training of German and Slavic languages. The CALL system AzAR (German acronym for 'automat for accent reduction') which was developed at our institute laboratory provides adequate feedback regarding the pronunciation of the learner as well as the prosodic quality [1]. With regard to learners and teachers, the most important new feature is the integration of large speech corpus and multilingual speech data bases[2]. The main objective of the present work is to demonstrate the significance of limited german phoneme rhythmic based prosodic variations in frequency f_0 , duration and intensity and to develop the evaluation criteria for speech synthesis. This work starts with rhythmic based prosodic analysis, an accent and de-accent experiment using resynthesis and a perceptual test. The main work involves in rhythmic factors related to prosodic analysis which results for f_0 , duration and intensity and their significance. The improved perceptual quality of the duration and mean frequency modified phonemes proved to be a promising result for perception as demonstrated in the subjective evaluation test with resynthesis stimuli. Stress is considered as a prominent factor in this analysis because it is the basis of rhythm in all languages and it was well proposed a rhythmic continuum that does not stretch from phoneme to stress timing [5]. The speech database involves 50 male and 60 female speakers aged from 55 to 75 years. Each speaker read 105 german sentences. The data from 5 speakers was recorded in laboratory studio at 16 kHz, 16 bit PCM. Our results will discuss the significance of stress and intonation, phoneme duration, intensity variation, relation among duration, f_0 , intensity and perception evaluation. The main application leads to develop a multimodel data-based assistance system for self learning activities for old aged people affected with brain related diseases e.g., Parkinson's disease. This work is a part of ongoing project titled "RehaVox", funded by Federal Ministry of Economics and Technology, Germany.

References

- [1].Jäckel, R., Jokisch, O., Hoffmann, R." Evaluation of speaker proficiency in a pronunciation tutoring system". *Proc. 12th International Conference (SPECOM)*, Moscow, pp. 772 – 777 October 2007.
- [2].Jäckel, R.,Hoffmann, R.:Sprachsignalanalyse für Menschen mit Deutsch als Zweitsprache - eine intelligente Lerner-Software zur Aneignung der deutschen Standardaussprache. In: *Deutsch als Zweitsprache 3*, pp.37 - 46 2008
- [3].K.S.Rao and B.Yegnanarayana, "Duration modification using glottal closure instants and vowel onset points," *Speech Communication* Vol.40, no.3, pp.1263-1269, *Elsevier* 2009
- [4].Jokisch, O.,Birhanu, Y., Hoffmann, R. "Syllable-based prosodic analysis of Amharic read speech " *IEEE Spoken Language Technology Workshop (SLT)*, pp. 252 – 257, 2012
- [5].Felicitas Kleber, Nadine Klippahn "An acoustic investigation of secondary stress in German" working papers at *Institute of Phonetics and Digital Speech Processing*, Christian-Albrechts University Kiel,Germany, 2011