

Prosodic Stress Detection for Fixed Stress Languages Using Formal Atom Decomposition and a Statistical Hidden Markov Hybrid

György Szaszák, Máté Ákos Tündik, Branislav Gerazov

PII: S0167-6393(17)30408-9
DOI: [10.1016/j.specom.2018.06.005](https://doi.org/10.1016/j.specom.2018.06.005)
Reference: SPECOM 2575



To appear in: *Speech Communication*

Received date: 7 November 2017
Revised date: 8 May 2018
Accepted date: 26 June 2018

Please cite this article as: György Szaszák, Máté Ákos Tündik, Branislav Gerazov, Prosodic Stress Detection for Fixed Stress Languages Using Formal Atom Decomposition and a Statistical Hidden Markov Hybrid, *Speech Communication* (2018), doi: [10.1016/j.specom.2018.06.005](https://doi.org/10.1016/j.specom.2018.06.005)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Prosodic Stress Detection for Fixed Stress Languages Using Formal Atom Decomposition and a Statistical Hidden Markov Hybrid

György Szaszák^{a,*}, Máté Ákos Tündik^a, Branislav Gerazov^b

^a*Dept. of Telecommunications and Media Informatics*

Budapest University of Technology and Economics

2 Magyar Tudósok körútja, Budapest H-1117, Hungary

^b*Faculty of Electrical Engineering and Information Technologies,
University of Ss. Cyril and Methodius in Skopje, Macedonia*

Abstract

The detection of prosodic events, prosodic stress, and speech segmentation based on prosody have received much attention in the research community in the past decades. Prosody is relevant for both main areas of speech technology, text-to-speech synthesis and automatic speech recognition and understanding, and is exploited increasingly: besides providing redundancy, prosody is recognized to carry information unavailable from other sources and also contributes to the naturalness of the perceived speech. This paper addresses a recently proposed intonation analysis technique, called Weighted Correlation based Atom Decomposition (WCAD). The WCAD approach is inspired by the physiology of speech production and the Fujisaki-model used in speech synthesis, however, it is employed in an analytic, and not in a generative approach: the intonation contour is decomposed into a set of elementary components, called atoms, by a pattern matching algorithm. The obtained atom decomposition is used for prosodic stress detection and automatic phonological phrasing. We compare and also combine the WCAD approach to a phonological approach, which relies on automatic segmentation for phonological phrases using a Gaussian Mixture

*Corresponding author

Email addresses: szaszak@tmit.bme.hu (György Szaszák), tundik@tmit.bme.hu (Máté Ákos Tündik), gerazov@feit.ukim.edu.mk (Branislav Gerazov)

Download English Version:

<https://daneshyari.com/en/article/6960432>

Download Persian Version:

<https://daneshyari.com/article/6960432>

[Daneshyari.com](https://daneshyari.com)