

---

## Prosodic and syntactic structures in spontaneous speech: a wavelet-based approach to prosodic modelling

---

Mittwoch  
08.03.2017  
15:15 – 15:45  
B4 1, 0.25

Anna Dannenberg<sup>1</sup>, Stefan Werner<sup>2</sup>, Martti Vainio<sup>1</sup> & Antti Suni<sup>1</sup>

<sup>1</sup>University of Helsinki, <sup>2</sup>University of Eastern Finland

{anna.dannenberg|martti.vainio|antti.suni}@helsinki.fi, stefan.werner@uef.fi

For analysing the prosodic structure of spoken language, especially of spontaneous speech, effective automatic applications have been infrequent. For instance, prosody is often assumed to have some kind of a hierarchical structure, but prosodic patterns have been hard to observe or visualise directly.

Our solution is a Continuous Wavelet Transform (CWT). The method applies the weighted sum of f0, energy and segmental durations to represent prosodic signals in a two-dimensional time-scale plane akin to spectrograms. This can be further enhanced with lines of maximum amplitude to produce a visual representation of the prosodic hierarchies of speech.

We have used a CWT based tool to analyse the prosodic structure of spontaneous speech. The results have been compared with grammatical analysis of the same data, achieved both automatically and manually, to examine the relation of prosody and syntax in spoken language.

Our study of English and Finnish speech data demonstrates a significantly high rate of co-occurrence between prosodic and syntactic boundaries. This result proposes a connection between prosody and syntax. The internal hierarchical structures of prosodic and syntactic units, on the contrary, appear to have practically no resemblance to each other. This may be due to our choice of syntactic model, based on traditional phrase structure grammar, that is not well suited for analysing unplanned spontaneous speech.

It thus seems that CWT is a plausible method for modelling prosodic structures of spoken language, but new grammatical approaches are required to deal with syntactic characteristics of spontaneous speech. We are looking forward to see if CWT based models of prosodic hierarchy can help to find new perspectives to the grammar of spoken language.

**References:** • Mallat, S. (1999): *A Wavelet Tour of Signal Processing*. Academic Press. • Suni, A., Aalto, D. and Vainio, M. (2015): Hierarchical representation of prosody for statistical speech synthesis. arXiv:1510.01949. • Dannenberg, A., Werner, S. and Vainio, M. (2016): Prosodic and syntactic structures in spontaneous English speech. In: *Proceedings of Speech Prosody 2016*, Boston, USA, 59–63.

AG6