A Thematicity-based Prosody Enrichment Tool for CTS

Mónica Domínguez monica.dominguez@upf.edu
Mireia Farrús mireia.farrus@upf.edu
Leo Wanner leo.wanner@upf.edu

Universitat Pompeu Fabra Barcelona
TALN Grup de Recerca en Tractament Automàtic del Llenguatge Natural

Abstract

This paper presents a demonstration of a thematicity-based prosody module for enrichment of synthesized speech using SSML prosody tags in the context of a CTS application. The motivation for using hierarchical thematicity is exemplified, together with the capabilities of the module to generate a variety of SSML prosody tags within a controlled range of data-driven values depending on the input thematicity label.

Motivation

Traditional Information Structure (IS) interpretations [1, 2, 3] establish a deterministic correlation between **theme–rheme (i.e.,** "what is being talked about" and "what is being said") and **rising–falling** intonation patterns in text-to-speech (TTS) applications.

However, such methodology has several drawbacks:

(i) it **fails to describe longer sentences** with complex syntactic structures;
(ii) it **ignores other prosodic elements**, such as rhythm and intensity (that also relate to information and prosody structure [5, 6]) and;
(iii) it **presupposes a fully deterministic mapping** between intonation labels and acoustic parameters.

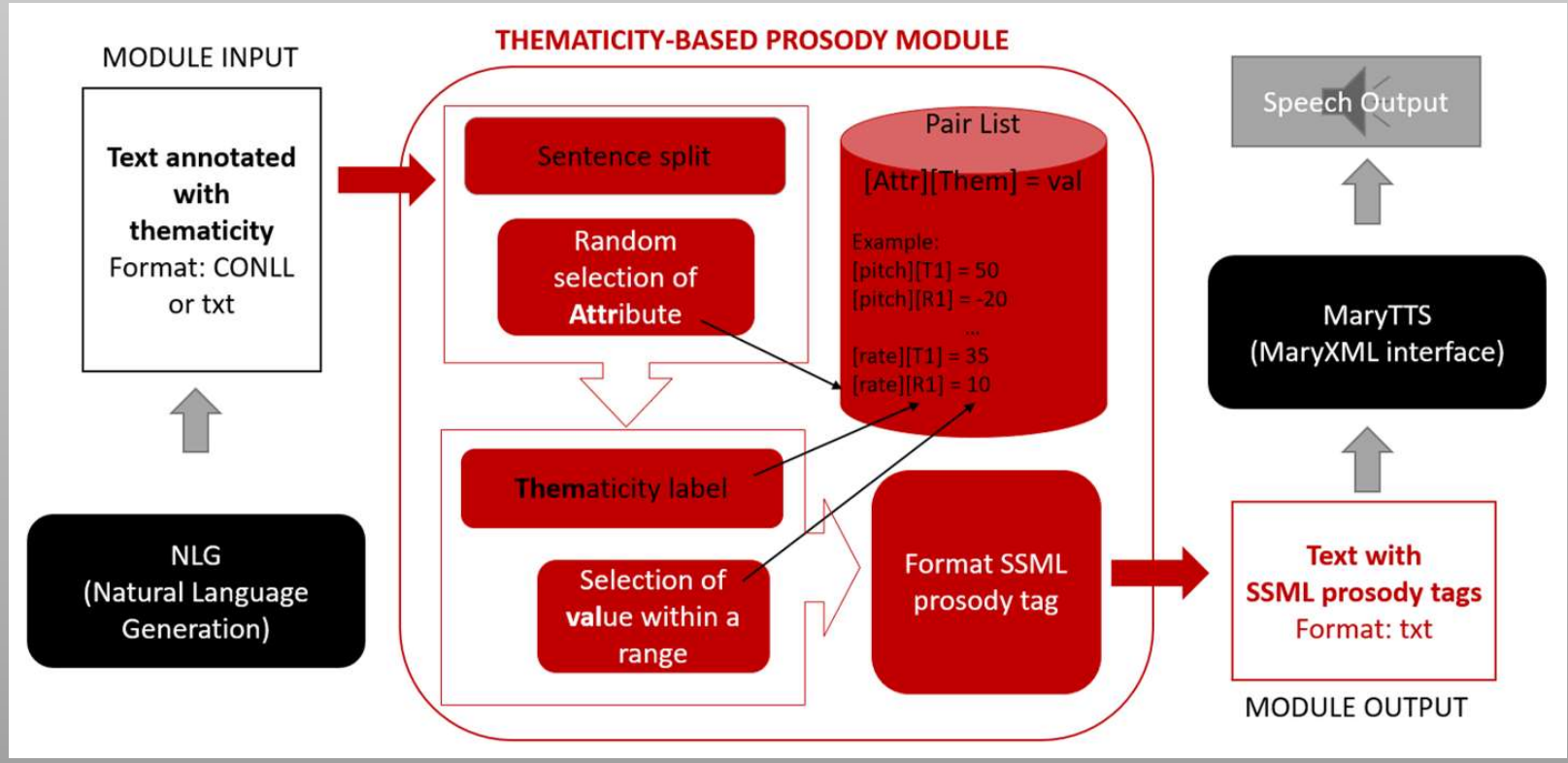
Our Approach

Our approach to **thematicity** is based upon **Mel'čuk's [4]** formal proposal of **communicative structure representation** within the **Meaning-Text Theory (MTT)**.

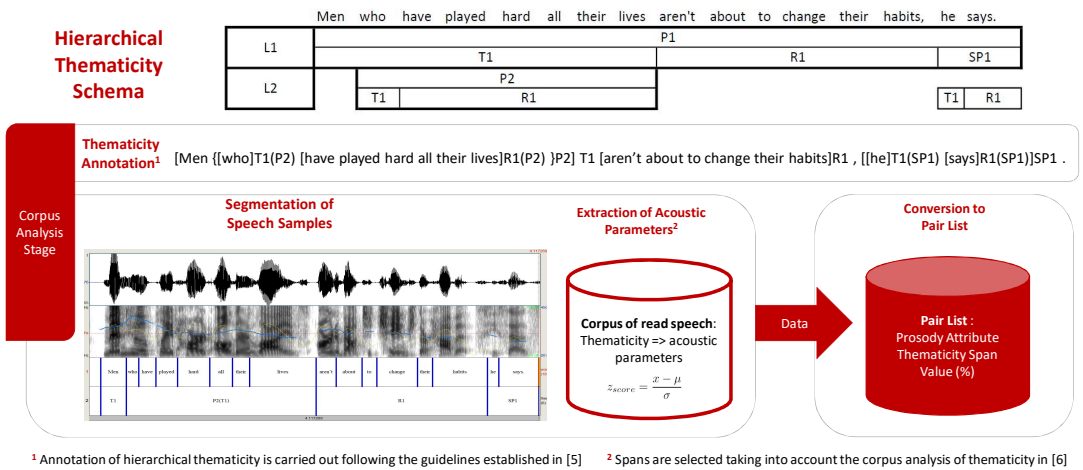
- Apart from theme and rheme, a third element, namely 'specifier' (which sets the utterance's context), is introduced;
- Thematicity is defined over propositions (i.e., a theme or a rheme can contain another theme/rheme/specifier division) and, thus, be hierarchical.

Prosody is analyzed on a corpus of read speech using normalized values (z-scores) of acoustic parameters for:

- **F0**
- **Intensity**
- **Speech rate**



Corpus-based Approach to Thematicity-based Prosody Enrichment



[1] Annotation of hierarchical thematicity is carried out following the guidelines established in [5]

[2] Spans are selected taking into account the corpus analysis of thematicity in [6]

Publication:

M. Domínguez, M. Farrús, L.Wanner, "A Thematicity-based Prosody Enrichment Tool for CTS", in *Proceedings of Interspeech 2017: system demonstrations*, Stockholm, Sweden, 2017.

Code:

http://github.com/TalnUPF/thematicitytoSSML

References:

[1] M. Steedman, "Information structure and the syntax-phonology interface," in *Linguistic inquiry*. Cambridge, Massachussetts: The MIT Press, vol. 31, no. 4, 2000, pp. 649–689.

[2] I. Kruijff-Korbayova, et al., "Producing Contextually Appropriate Intonation in an Information-State Based Dialogue System," in *Proceedings of the 10th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2003, pp. 227–234.

[3] M. Haji-Abdolhosseini, "A Constraint-Based Approach to nformation Structure and Prosody Correspondence," in *Michigan State University, East Lansing*, S. Muller, Ed. CSLI Publications, 2003, pp. 143–162.

[4] I. A. Mel'čuk, *Communicative Organization in Natural Language: The semantic-communicative structure of sentences*. Amsterdam, Philadephia: Benjamins, 2001.

[5] B. Bohnet, A. Burga, and L. Wanner, "Towards the annotation of Penn Treebank with information structure," in *Proceedings of the Sixth IJCNLP*, Nagoya, Japan, 2013, pp. 1250–1256.

[6] M. Domínguez, M. Farrús, L.Wanner, "A Data-driven Approach toThematicity-based Prosody Enrichment", in *Proceedings of SLSP2017, Le Mans, France, 2017.*