

# Syllable Prominence and Prosodic Phrasing in Spoken Prose

## Prediction from Text and Validation

Isabelle Franz<sup>1,2</sup>, Christine Knoop<sup>1</sup>, Gerrit Kentner<sup>1,2</sup>, Vanessa Kegel<sup>1</sup>, Julia Vasilieva<sup>1</sup>, Sascha Rothbart<sup>1</sup>,  
Sanja Methner<sup>1</sup> & Winfried Menninghaus<sup>1</sup>

<sup>1</sup> Max-Planck-Institute for Empirical Aesthetics, Frankfurt; <sup>2</sup> Goethe University Frankfurt

## Introduction

**Prose rhythm:** the notion that prose, like poetry, may be marked by distinctive types of rhythm.

No conclusive evidence that the notion of 'prose rhythm' is anything more than a metaphor. [1]

Main reason: no replicable, phonetically validated coding system for rhythmic/prosodic grouping.

### Challenge:

- Relative syllable prominence - use of metrical grids limited to short units. [2]
- Prosodic boundaries of different strengths - existing systems depend on syntactically annotated texts.

## Methods

- Based on work in the fields of metrical phonology [2] and phonetics of speech rhythm, we developed a manual for coding syllable prominences and caesurae (Figure 1). Quality Criteria: Replicability – Broad/Simple Usability – Accuracy – Brevity.
- 3 annotators coded the beginning pages of four German novels.
- Inter-annotator agreement Cohen's  $\kappa$  .9 - .96.

### Phonetic analysis of readings

- 8 professional speakers read the text.
- 90.000 syllables, automatic annotation with MAUS [3].
- Spoken syllables were matched to, and aligned with, citation form syllables, extraction of duration, F0, and intensity for each syllable with praat [4], comparison with predicted syllable prominence and caesurae.

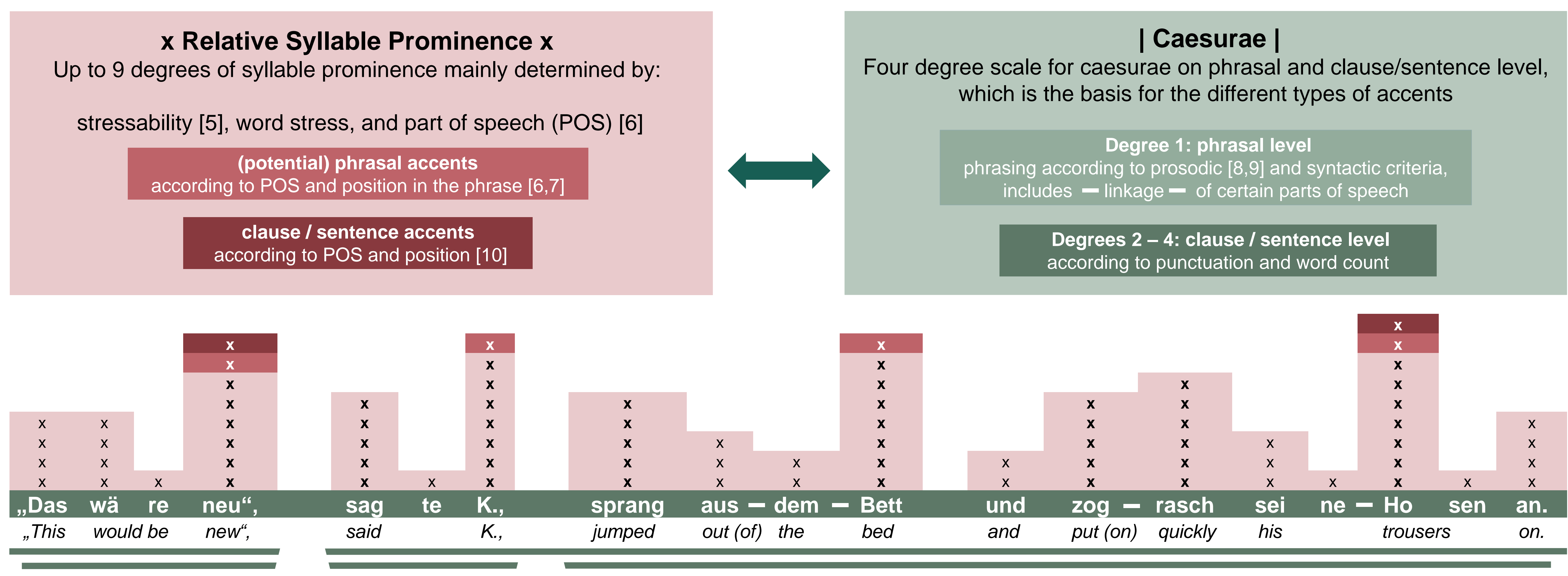


Figure 1: Basic workings of the coding system for the prediction of syllable prominence and caesurae

## Validation of coding system with phonetic realization

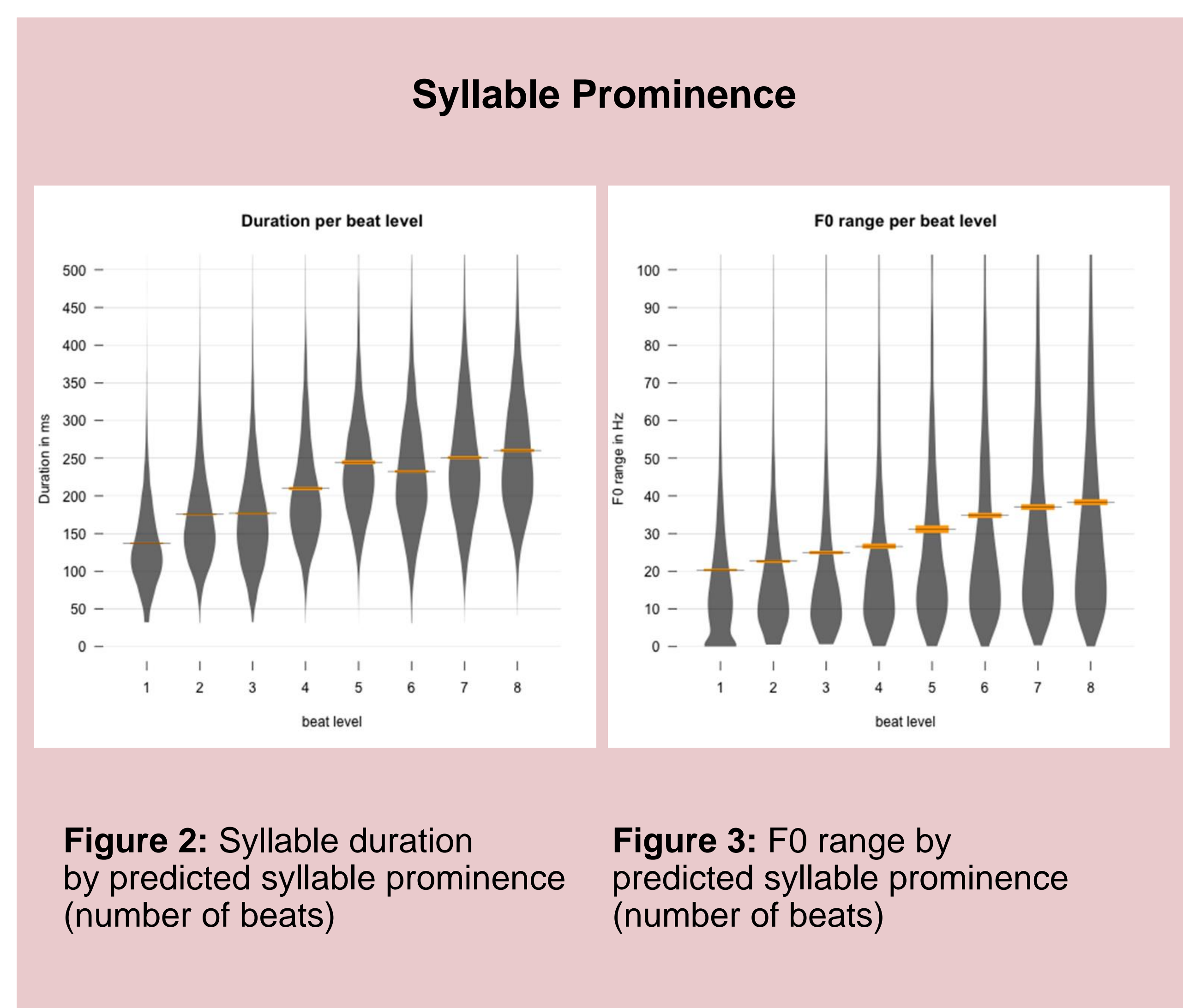


Figure 2: Syllable duration by predicted syllable prominence (number of beats)

Figure 3: F0 range by predicted syllable prominence (number of beats)

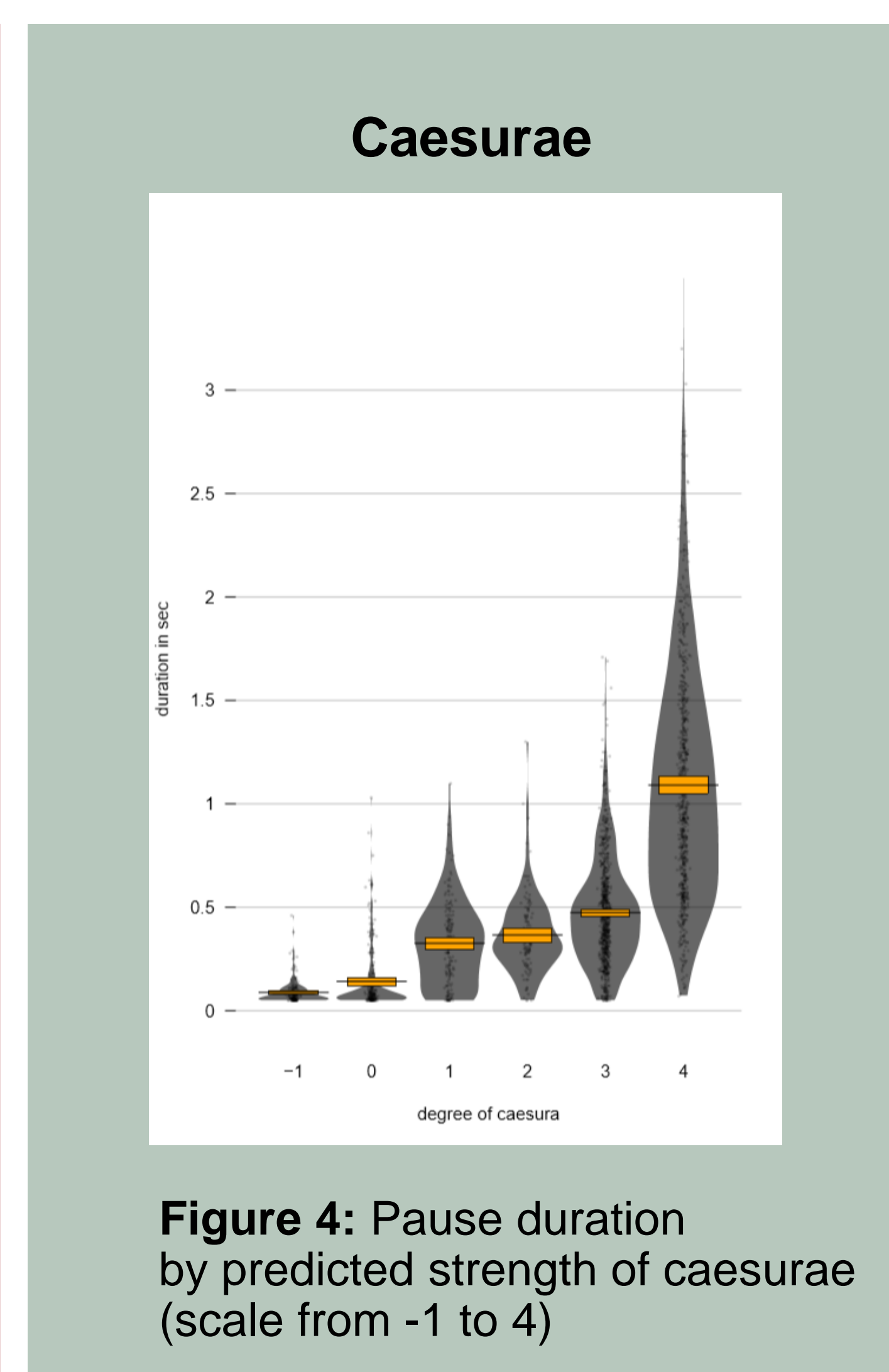


Figure 4: Pause duration by predicted strength of caesurae (scale from -1 to 4)

## Conclusion

- The data (~90 000 spoken syllables) confirm that up to 8 levels of cumulative prominence and 6 levels of caesurae can be distinguished.
- We present a phonetically validated and manually applicable tool for predicting syllable prominences and pauses in spoken prose.
- This instrument makes it possible to systematically analyze rhythmic patterns in prose or single sentences.

### Potential applications of the system

- Author profiling and genre/style recognition.
- Prosodic predictions for synthetic speech.
- Automatization of the annotation.