

# SPOKEN LANGUAGE TRAITS IN THE POLITICAL DOMAIN: A CORPUS-BASED STUDY

DANIELA TROTTA  
SARA TONELLI  
ANNIBALE ELIA

*Studi Italiani di Linguistica Teorica e Applicata, anno L, 2021, numero 2*

## ABSTRACT

In this paper, we present the first available corpus of Italian political inter-views with multimodal annotation, consisting of 56 face-to-face interviews (14 hours) taken from a political talk show. We first detail the annotation scheme that besides the metadata (i.e. information on tools used for the transcription, link to the interview etc.) includes some speech constants like pauses (used to mark a pause either between or within utterances), vocal expressions (marking non-lexical expressions such as burp and semi-lexical expressions such as primary interjections), deletions (false starts, repetitions and truncated words) and overlaps. Furthermore, we present several statistical analyses to understand the relation between these mul-timodal traits and language complexity. We also exploit the corpus to test the validity of existing studies on political orientation and language use, showing that results on our data are not as clear-cut as on English ones.