

# POLITOLINGUISTICS THROUGH SPOKEN LANGUAGE PROCESSING: A METHODOLOGICAL FRAMEWORK FOR GERMAN AND ITALIAN POLITICAL SPEECHES

*Marcella Palladino*

*Università degli Studi di Modena e Reggio Emilia  
marcella.palladino@unimore.it*

**Abstract:** This paper presents the methods used for the doctoral project of the author, which aims to combine politolinguistics with spoken language processing. The purpose is to offer a methodological framework that can be used for further studies on political spoken language. The illustration of the methods used for the doctoral project shows how politolinguistic analyses can be enriched by selected tools of the Bavarian Archive for Speech Signals (BAS) and how Rapid Prosody Transcription (RPT) can be a suitable option to relate prosody with politolinguistic units of analysis using a large corpus of political spoken language involving two different languages (i.e. German and Italian). Eventually, despite the limitations of this pilot methodology, the paper lays the foundations for future studies which can further improve the framework. The exploratory character of the methodology, which binds quantitative and qualitative research, paves the way to further improvements and application in the field of politolinguistics.

## 1 Introduction

This paper illustrates the methods used in the doctoral project of the author, undertaken at the University of Modena and Reggio Emilia (Italy), which aimed to apply the methods of spoken language processing to enrich and expand a politolinguistic analysis. In this contribution, the doctoral project is not delved into in depth, but only the research methods are presented. The purpose of binding politolinguistics and spoken language processing comes from the state of the art of politolinguistic analyses, which tend to focus either on written or already transcribed spoken texts, presumably also due to the lack of methods to reduce the time-consuming task of transcribing. Moreover, since political language is mostly spoken language, it seemed reasonable to propose a methodology grounded on spoken language processing to consider the acoustic signal also in politolinguistics and especially to bind politolinguistic units with prosodic prominence. This methodological proposal has an exploratory character and represents a pilot study that needs to be further expanded and whose limitations need to be further addressed.

As far as the theoretical background is concerned, scholarship that touches upon politolinguistics [1], political rhetoric [2], and prosody [3] has been considered. However, it seems that no study applied the same methodology before. The contribution of Draxler [4] discussed thoroughly and extensively some web tools that will be presented in this study, but they were not applied to politolinguistics. These tools, which come from the spoken language processing, did prove to be helpful in the field of politolinguistics [5] and this doctoral project paves the way for their application to enhance studies on spoken language in politolinguistics. In the following sections, first the corpus of the project is described, then the units of analysis are briefly explained and eventually the methods are delved into, focusing also on their limitations and further steps that can be useful for the analysis.

## 2 Corpus and transcription of the speeches

The corpus built for the doctoral project consists of 44 political speeches divided into 22 Italian speeches (11 male and 11 female speeches) and 22 German speeches (11 male and 11 female speeches). The choice of comparing German and Italian was linked to the political orientation of the selected speeches for the study and the attitude towards traditional gender-related roles. The speeches were collected on YouTube, where the videos are publicly available, and they were converted into audio recordings in WAV format with the software Audacity. The Italian recordings last approximately 06:48 hours in total and the German recordings last approximately 05:58 hours in total. Automatic Speech Recognition (ASR) was applied for the orthographic transcription of the speeches. While the Italian recordings were processed by an ASR system offered by the team of Daniele Falavigna at Fondazione Bruno Kessler (FBK) in Trento, the German recordings were processed by the IBM Watson ASR service directly available on OCTRA [6]. The Italian transcripts were then corrected<sup>1</sup> by the author using OCTRA [7] and the German transcripts were corrected by German students.

The time frame of the speeches is from 2015 to 2022, and they belong to different genres, i.e. they are either parliamentary or non-institutional<sup>2</sup> speeches. The public availability of the speeches constituted one of the criteria for selection. All speeches were delivered by a single speaker, and no continuous interaction is present. Collecting speeches of different genres gives the possibility to observe whether the transcribing process shows different problems related to the genre of the speech. Moreover, all selected speeches contain background noise, regardless of the speech type. This surely poses a challenge for the ASR.

The orthographic transcripts obtained from the ASR and corrected in OCTRA were downloaded in TXT format in two versions for each transcript, namely one with time stamps and one without time stamps. Indications such as overlapping of the speakers or the changing of the speakers (although very limited, since the speeches are delivered by a main speaker who is rarely interrupted) are signaled with angle brackets. These transcripts were further processed using other web tools and the Rapid Prosody Transcription (see 4.2).

## 3 Units of analysis

The types of units of analysis used in the PhD dissertation are two. The first part of the analysis aims at considering quantitative data and making use of descriptive statistics [8], focusing on the acoustic signal. The second part of the analysis revolves around politolinguistics and prosodic prominence, offering a more qualitative perspective. For both the quantitative and the qualitative parts of the analysis, the doctoral project aimed to be an exploratory case study that compared two languages. For this reason, features were selected both on the basis of similar pre-existing literature and as samples to test the methods.

For the quantitative analysis, the selected units of analysis are pitch range, speaking rate and pauses (duration and frequency) for each category of speakers (German male politicians, German female politicians, Italian male politicians, Italian female politicians). They were extracted with the package emuR [9], organized in plots and compared. In the doctoral project, the observation of these features was brought into politolinguistics and specifically into a work that considered gender-related characterization in political language. The analysis of these features was carried out on the basis of previous literature. In the qualitative analysis, the units of analysis are lexical and rhetorical categories typical of politolinguistic literature, namely

---

<sup>1</sup> ‘Corrected’ means in this study that the errors in terms of words made by the ASR were corrected, but the mistakes made by the speakers were preserved in order to maintain typical traits of orality. Interjections and repetitions were also preserved, and added in case the ASR had left them out.

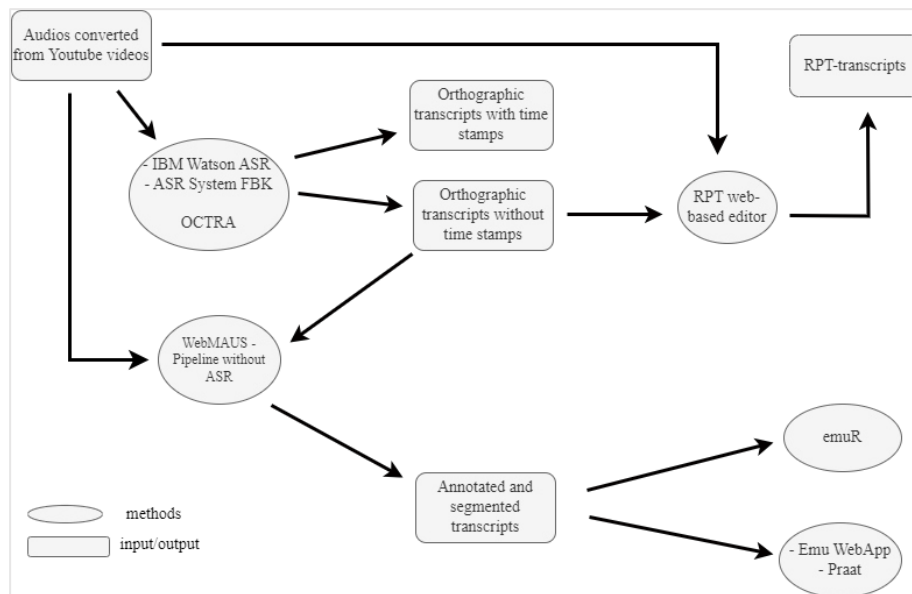
<sup>2</sup> ‘Non-institutional’ indicates in this work all speeches that are not delivered in a parliamentary setting, e.g. speeches at political conventions, party rallies, election speeches.

*Schlag*-<sup>3</sup>, *Fahnen*- and *Stigmawörter*, *Plastikwörter*, *Begriffe besetzen*, hyperbole/euphemism, metaphor, irony. These categories were related to prosodic prominence [11] through the method of Rapid Prosody Transcription (RPT) [12], which is explained in section 4.2.

The outcomes of the analysis are not the focus of this paper; the methodology used, on the contrary, represents the main point of discussion and will be illustrated extensively. For the quantitative analysis, the abovementioned features were selected because of the specific aims of the doctoral project. As far as the units of analysis for the qualitative analysis are concerned, they were chosen due to both their relevance in previous politolinguistic studies and the focus on the lexical level, since the perspective that was taken regarding the prosodic prominence was centered on the word level.

## 4 Methods

Methodologically, this study integrates tools from BAS – Bavarian Archive for Speech Signals [13] – and the method of Rapid Prosody Transcription with politolinguistic and rhetorical categories of analysis. The methods are visualized in a scheme that is included in the doctoral dissertation (Figure 1).



**Figure 1** - Visualization of the methodological framework (the scheme is taken from the doctoral dissertation of the author)

### 4.1 Spoken Language Processing

The tools that were used for the doctoral project belong to BAS and were selected on the basis of the objectives of the research: orthographic transcription with ASR; annotation, segmentation and time alignment of the transcripts with the signal; extraction of data and organization of the data into plots. OCTRA and the systems used for the transcription of the speeches were already mentioned in section 2. The downloaded transcripts without time stamps were then used for RPT (see 4.2). The transcripts and the corresponding audios were uploaded in WebMAUS [14] to obtain time-aligned annotated and segmented transcripts. WebMAUS General appeared to be the most suitable option due to the presence of more features than WebMAUS Basic.

<sup>3</sup> The politolinguistic categories are in German, since the relevant literature on this topic used for the project is German. The term *Politolinguistik* ('politolinguistics') was introduced by Burkhardt [10].

However, since the formats available for the upload on MAUS were TXT and WAV files (and not BPF as requested by WebMAUS General), it was decided to use WebMAUS General inside the tool Pipeline without ASR available on BAS. The specific pipeline that was opted for was G2P→Chunker→MAUS and the output chosen was IPA (UTF-8). The outputs of MAUS were transcripts that showed four different tiers containing the segmented orthographic transcription, the SAMPA transcription, the IPA transcription and the whole orthographic transcription in the bottom tier. The outputs were saved both in annot.json and in TextGrid and they were visualized on Praat [15] and on EMU-webApp [16]. These outputs were also necessary to extract data using the R package emuR and to create plots depicting pitch range, speaking rate and pauses (duration and frequency). The box plots created were compared and used to answer the specific research question concerning gender-related characterization in political speeches. In this way, these types of quantitative data regarding the signal were brought into a politolinguistic analysis, and a relatively big amount of data, which could be representative for the specific analysis, could be considered. As already mentioned, the aim of this paper is to report the use of these methods in politolinguistic analysis. The results of the analysis are related to the research questions of the project and are not consistent with the purpose of this paper, therefore they will not be discussed in this contribution.

The mentioned BAS web tools were used independently from one another in the doctoral project, although they can be used directly one after the other in the BAS Transcription Portal [17] or used through emuR. The reason why it was decided to use the tools separately lies in the pilot nature of the project: The BAS web tools had never been applied before to (long) political speeches embodying the abovementioned challenges; therefore, the doctoral study also offered the chance to test the web tools for this type of data. Applying each tool separately could give a better perspective on their performance and their results. Moreover, the work was also intended as a first attempt to present the use of BAS web tools in politolinguistics, without focusing on programming languages for the moment.

## 4.2 Rapid Prosody Transcription for political speeches

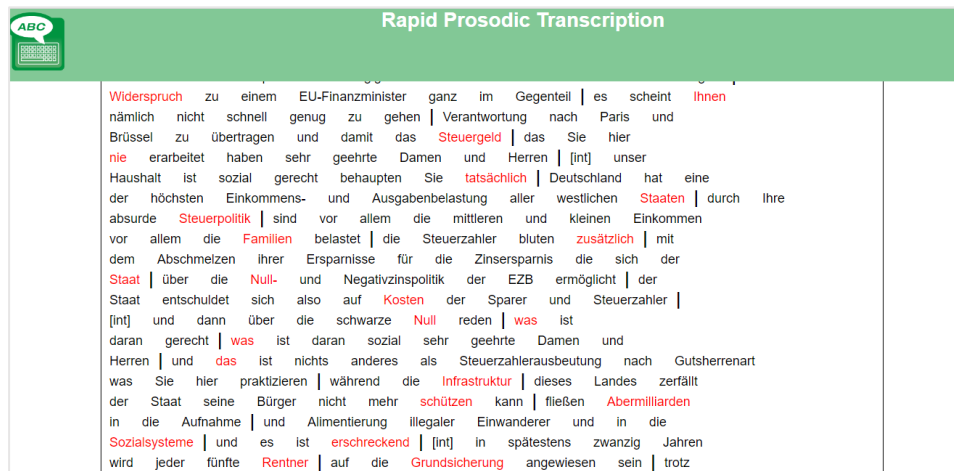
The second part of the methodology concerns the Rapid Prosody Transcription (RPT) and its relation to the mentioned politolinguistic and rhetorical units of analysis. RPT is based on prosodic prominence of words, thus it seems particularly suitable for the relation to the mentioned categories from politolinguistics and rhetoric, which involve word-level units of analysis. To conduct the prosodic transcription, a web-based editor for RPT<sup>4</sup> was developed in the BAS platform. This editor met the requirements specifically needed for the doctoral project, namely the possibility to pause and rewind the audios (considered the length of the speeches and the quality of the available audios) and to obtain an output in TXT. The prosodic transcription was carried out by the author of the project only, due to the scope of the research.

The prosodic transcription was done using the orthographic transcripts without time stamps downloaded from OCTRA. Since one of the requirements for RPT is to consider texts without punctuation, these transcripts appeared to be adequate for this method. The outputs of the RPT are in TXT and the prosodically prominent words are inside angle brackets in the form of <ITM ..> (ITM stands for ‘item’). The boundaries that divide the prosodic units within the text are indicated with a bar ‘|’ in the TXT-transcripts. The editor on which the prosodic transcription was carried out is shown in Fig. 2. In the editor, the bar that divides the prosodic units is inserted

---

<sup>4</sup> The RPT editor was developed by PD Dr. Christoph Draxler at the Institute of Phonetics and Speech Processing, LMU Munich (Germany): [https://webapp.phonetik.uni-muenchen.de/WebTranscribe/politoprominence\\_de.html](https://webapp.phonetik.uni-muenchen.de/WebTranscribe/politoprominence_de.html) (accessed July 08, 2024). The editor for the original RPT method was developed by the Prosody & Speech Dynamics Lab of Northwestern University (USA): [https://github.com/ProSD-Lab/NU\\_LMEDS](https://github.com/ProSD-Lab/NU_LMEDS) (accessed July 08, 2024).

with a click. To indicate the prosodic prominence of a word, the click on this word turns it into red.



**Figure 2** - Example of a German speech in the RPT editor developed at the IPS as part of the BAS services

Furthermore, RPT seems to be a promising method to be applied to large corpora and it does not require either a specific expertise in prosodic annotation or an extensive time consumption. Considering two languages, as the case of this doctoral project, poses a challenge regarding prosodic transcription, especially if the different prosodic specificities of German and Italian [18] are considered. The use of RPT allows the comparison between these two languages, since they both have words available as units and thus the prosodic prominence can be indicated on the word level [19]. Moreover, RPT seemed to have never been applied to the Italian language when the doctoral project started, which made this project also a first attempt at using this method for Italian language.

## 5 Outcomes and limitations

As already mentioned, this paper does not aim to present the research questions and the analyses of the doctoral project in detail, but rather to briefly illustrate a methodology to bring spoken language processing into politolinguistics. The reason is that the illustrated tools could facilitate and speed up the gathering and the transcription of data and thus enhance studies on political spoken language (even for non-institutional speeches that do not have an official transcription). The outputs and results from the mentioned tools allow to consider quantitative data and create plots that accompany the qualitative politolinguistic analysis, offering an overview of the signal.

As far as the prosodic transcription and its relation to the mentioned politolinguistic and rhetorical features are concerned, the doctoral project showed that the results of the qualitative analysis are characterized by a large degree of variability for all the categories considered. No recurrent prosodic patterns for prosodic prominence were found in the politolinguistic and rhetorical categories, but the attribution of the prosodic prominence seemed instead to be dependent on the context and the co-text of the speech. Even in the same speech, the same word sometimes carried prosodic prominence and some other times it did not. Therefore, it was observed that a fixed scheme for prosodic prominence of each feature considered did not appear to exist. Moreover, a certain variability was observed also across the politolinguistic categories: For instance, if a word was a *Fahnenwort* in a specific instance of the speech, not only was the same word not a *Fahnenwort* in another instance in the other speeches considered, but it could change its function even within the same speech (see also [20] and [21]).

Although the described pilot methodology provided results and demonstrated how the mentioned tools can actually be used for politolinguistics, there are some limitations and further steps that can be implemented, as well as alternative tools that can be used in the future. One of the limitations concerns the automatization in the extraction of the quantitative data and the lack of the observation of inter-transcriber agreement for RPT. In the case of the first limitation, due to the size of the corpus and the time requirements of the project, the data were gathered through emuR but not double-checked manually. Hence, the character of the project is exploratory also in the sense that a first methodological process is proposed, but needs further investigation and further tests, as well as a further discussion on the parameters that can be selected. Regarding the inter-transcriber agreement, the RPT on the speeches of the corpus needs to be conducted by other transcribers in order to compare the results. Also, it has to be determined which language requirement (native/non-native speakers) the transcribers should meet and if they should master both German and Italian or if the transcribers should be divided into two groups according to their native language.

## 6 Conclusion

This paper focused on the methods used in the doctoral project of the author, which aimed to make use of tools from the spoken language processing to enhance the investigation of spoken language in politolinguistics. As mentioned, the specific research questions of the project, as well as their answers, did revolve around a gender-related analysis of political speeches, which also served as a case study to apply the selected methods. The limitations have been examined as well and future studies are planned to further delve into the topic.

The originality of this work lies in the application of BAS web tools to politolinguistics; hence, this contribution mainly focused on the description of these tools. The application of ASR systems based on Artificial Intelligence (AI) is a further step for future studies in order to test which system appears to be the most suitable for the political speeches considered; for instance, Whisper OpenAI [22] may be taken into account. Whisper is now available on OC-TRA and on the App aTrain [23], as well as with a package on Python. Enough material seems thus to be there to further implement and expand the methodology of the doctoral dissertation.

Future studies can involve the further application of RPT and some tests on inter-transcriber agreement. A further perspective may also take into account which level of linguistic proficiency the transcribers should have. For this reasons, future studies based on RPT transcription by language learners are already planned, focusing especially on Italian L2 learners of German in Italian university's BA and MA courses, using the RPT editor developed on the BAS platform.

The doctoral project showed that methods from spoken language processing can enrich a politolinguistic study and facilitate the gathering and organization of data. With the ASR systems, it was possible to build a representative corpus that included numerous and also long speeches. The Rapid Prosody Transcription applied to politolinguistics offered the possibility to carry out a prosodic annotation based on prosodic prominence of a large amount of data and to relate the results to selected politolinguistic categories. The research can be expanded in the future to include further languages, as already foreseen by the doctoral dissertation.

## Acknowledgments

I would like to thank Prof. Dr. Ulrike A. Kaunzner, PD Dr. Christoph Draxler, Prof. Dr. Jennifer Cole, M.A. Vanessa Reichel, Dr. Daniele Falavigna and his team at Fondazione Bruno Kessler, the German students who transcribed and corrected the German transcripts, the University of Modena and Reggio Emilia, the Institute of Phonetics and Speech Processing in Munich and the Prosody & Speech Dynamics Lab of Northwestern University (USA).

## References

- [1] NIEHR, T.: *Einführung in die Politolinguistik. Gegenstände und Methoden*. Vandenhoeck & Ruprecht, Göttingen, 2014.
- [2] BURKHARDT, A.: *Handbuch Politische Rhetorik*. De Gruyter, Berlin/Boston, 2019.
- [3] COLE, J.: *Prosody in context: A review*. In *Language, Cognition and Neuroscience*, 30, P. 1 – 31. 2015.
- [4] DRAXLER, C.: *Automatic Transcription of Spoken Language Using Publicly Available Web Services*. In *Fare linguistica applicata con le digital humanities, Studi AIItLA*, Vol. 14, P. 27 – 49. 2022.
- [5] PALLADINO, M.: *Webbasierte Tools für die Transkription und Analyse von Reden. Hilfreiche Instrumentarien für die (Polito)Linguistik*. In *Lingue e Linguaggi*. Università del Salento, P. 413 – 437. 2024.
- [6] DRAXLER, C.: *Analysis of transcriptions using Octra – a pilot study*. In: C. DRAXLER (Ed.): *Elektronische Sprachsignalverarbeitung 105*, P. 17 – 23. TUDpress, Dresden, 2023.
- [7] PÖMP, J., and C. DRAXLER: *OCTRA - A Configurable Browser-Based Editor for Orthographic Transcription*. In *Proceedings of Phonetik und Phonologie im deutschsprachigen Raum*. P. 145 – 148. Berlin, 2017.
- [8] MEINDL, C.: *Methodik für Linguisten. Eine Einführung in Statistik und Versuchsplanung*. Narr Francke Attempto, Tübingen, 2011.
- [9] JOCHIM, M., R. WINKELMANN, K. JAENSCH, S. CASSIDY, and J. HARRINGTON: *emuR: Main Package of the EMU Speech Database Management System*. R package version 2.5.0. 2024. <https://CRAN.R-project.org/package=emuR> [17.01.2025].
- [10] BURKHARDT, A.: *Politolinguistik: Versuch einer Ortsbestimmung*. In: J. KLEIN, and H. DIEKMANN-SHENKE (Eds.): *Sprachstrategien und Dialogblockaden: Linguistische und politikwissenschaftliche Studien zur politischen Kommunikation*, P. 75 – 100. De Gruyter, Berlin/Boston, 1996.
- [11] COLE, J., Y. MO, and M. HASEGAWA-JOHNSON: *Signal-based and expectation-based factors in the perception of prosodic prominence*. In *Laboratory Phonology 1*, P. 425 – 452. 2010.
- [12] COLE, J., and S. SHATTUCK-HUFNAGEL: *New Methods for Prosodic Transcription: Capturing Variability as a Source of Information*. In *Journal of the Association for Laboratory Phonology 7(1)*, P. 1 – 29. 2016.
- [13] KISLER, T., U. D. REICHEL, and F. SCHIEL: *Multilingual processing of speech via web services*. In *Computer Speech & Language*, Vol. 45, P. 326 – 347. 2017.
- [14] SCHIEL, F.: *Automatic Phonetic Transcription of Non-Prompted Speech*. In *Proceedings of the ICPHS 1999*. P. 607 – 610. San Francisco, 1999.

- [15] BOERSMA, P.: *Praat, a system for doing phonetics by computer*. In *Glott International* 5:9/10, P. 341 – 345. 2001.
- [16] WINKELMANN, R., J. HARRINGTON, and K. JÄNSCH: *EMU-SDMS: Advanced speech database management and analysis in R*. In *Computer Speech & Language*, Vol. 45, P. 392 – 410. 2017.
- [17] DRAXLER, C., and J. PÖMP: *Transcription Portal – A Zero-configuration Workbench for Transcribing Spoken Language Recordings*. In: C. DRAXLER (Ed.): *Elektronische Sprachsignalverarbeitung, 105*, P. 170 – 172. TUDpress, Dresden, 2023.
- [18] KAUNZNER, U. A.: *Prosodie im Sprachvergleich Deutsch – Italienisch. Praktische Relevanz und ausgewählte Problembereiche*. In *Annali Online Lettere IV 2*. Università di Ferrara. 2009.
- [19] LADD, D. R., and A. ARVANITI: *Prosodic Prominence Across Languages*. In *Annual Review of Linguistics*, P. 171 – 193. 2023.
- [20] NIEHR, T.: *Schlagwörter und Leerformeln in der politischen Rede*. In: A. BURKHARDT (Ed.): *Handbuch Politische Rhetorik*, P. 671 – 688. De Gruyter, Berlin/Boston, 2019.
- [21] WENGELER, M.: *Wortschatz I: Schlagwörter, politische Leitvokabeln und der Streit um Worte*. In: K. ROTH, M. WENGELER, and A. ZIEM (Eds.): *Handbuch Sprache in Politik und Gesellschaft*, P. 22 – 46. De Gruyter, Berlin/Boston, 2017.
- [22] RADFORD A., J. W. KIM, T. XU, G. BROCKMAN, C. MCLEAVEY, and I. SUTSKEVER: *Robust speech recognition via large-scale weak supervision*. In *International Conference on Machine Learning*, P. 28492 – 28518. 2023.
- [23] HABERL A., J. FLEIB, D. KOWALD, and S. THALMANN: *Take the aTrain. Introducing an interface for the Accessible Transcription of Interviews*. In *Journal of Behavioral and Experimental Finance*, Vol. 41. 2024.