

Characterizing discourse genres with prosodic features in a reference treebank of spoken French

Guillaume Desagulier ^{1,*,@}, Anne Lacheret-Dujour ^{2,@}, Frédéric Isel ^{2,@}, Seongmin Mun ^{3,4,@}

1 : Modèles, Dynamiques, Corpus (MoDyCo) - [Website](#)

CNRS : UMR7114, Université de Paris X - Nanterre

Université Paris 10 Bâtiment A - Bureau 402 A 200, avenue de la République 92001 Nanterre Cedex - France

2 : Modèles, Dynamiques, Corpus (MoDyCo) - [Website](#)

CNRS : UMR7114, Université Paris X - Paris Ouest Nanterre La Défense

Université Paris 10 Bâtiment A - Bureau 402 A 200, avenue de la République 92001 Nanterre Cedex - France

3 : Modèles, Dynamiques, Corpus (MoDyCo) - [Website](#)

Université Paris Nanterre : UMR7114, Centre National de la Recherche Scientifique : UMR7114

Université Paris 10 Bâtiment A - Bureau 402 A 200, avenue de la République 92001 Nanterre Cedex - France

4 : Ajou University - [Website](#)

*: Corresponding author

Rhapsodie is a 33000-word treebank of spoken French that is annotated for syntax and prosody. It breaks down into 57 five-minute long samples produced by 89 male and female speakers. The discourse profile of each sample is captured by six variables: event structure (dialogue vs. monologue), social context (public vs. private), genre (argumentation, description, narrative, oratory, and procedural), interactivity (interactive, non-interactive, and semi-interactive), channel (broadcasting and face-to-face), and planning type (planned, semi-spontaneous, and spontaneous).

The prosodic profile of each sample is captured by two sets of three variables. The first set consists of primary (i.e. structurally objective) variables, namely the mean number per second of pauses (fPauses), conversational overlaps (fOverlap), and gap fillers (fEuh). The second set is based on a model consisting of secondary variables determined a priori by the authors because they are likely to occur in certain discourse genres. They are the mean numbers per second of prosodic prominences (fProm), intonational periods (fIPE), intonation packages (fIPA).

Our main research question is whether discourse types in French can be characterized and ultimately predicted by prosodic features. We also address two side questions. First, does the fact that the corpus is relatively small, heterogeneous, and not necessarily balanced affect the representativeness of our results? Second, are the secondary prosodic features representative of discourse genres?

We compiled a data table that consists of 57 observations (the corpus samples) and the twelve above listed variables. We visualized the table with RhapVis, a tool we designed on purpose (Fig. 1), explored it with principal component analysis (Fig. 2), and looked for confirmed tendencies with non-parametric one-way ANOVAs (Kruskal-Wallis H tests).

Our exploration shows that argumentative and narrative sequences are prosodically marked, whereas descriptive and procedural sequences are not. A discourse genre is prosodically marked when it is characterized by a high frequency of prosodic features, namely the simultaneous occurrence of overlaps, prominences, and intonation packages. We also claim that a discourse genre is prosodically marked when it is atypical with respect to the other speech genres. This is the case with oratory speech, which is characterized by a high frequency of intonational periods and pauses and is consequently isolated from the other types.

These results were partially confirmed by the ANOVAs. Focusing on primary variables, running an ANOVA on fPause showed a significant main effect of Genre ($p < 0.05$). Further inspection indicates that while the lowest fPause score was found in Narration ($M = 0.32$; $SD = 0.04$), the highest score was observed in Oratory ($M = 0.42$; $SD = 0.01$). For fOverlap, the main effect of Genre reached the

level of significance ($p < 0.001$), indicating that $fOverlap$ also varies according to Genre. The descriptive data showed that the $fOverlap$ score was the highest for both Argumentation ($M = 0.05$, $SD = 0.04$) and Narration ($M = 0.02$, $SD = 0.01$). Conversely, no overlap was found in both Oratory and Procedural samples.

References

- Lindqvist, Christina. Corpus transcrits de quelques journaux télévisés français, Stockholm, Elanders Gotab, 2001, 289 pages
- Portele T, Heuft B, Widera C, Wagner P, Wolters M (2000) Perceptual Prominence In: Speech and Signals. Aspects of Speech Synthesis and Automatic Speech Recognition. Festschrift dedicated to Wolfgang Hess on his 60th birthday. Forum Phonetikum, 69. Hektor, Frankfurt a.M.: 97-116.
- Wagner, P. et al. (2015b), « Disentangling and connecting different perspectives on prosodic prominence », Communication à ICPL, International Conference Prominence in Language, 2015, Cologne, ICPH, 2015

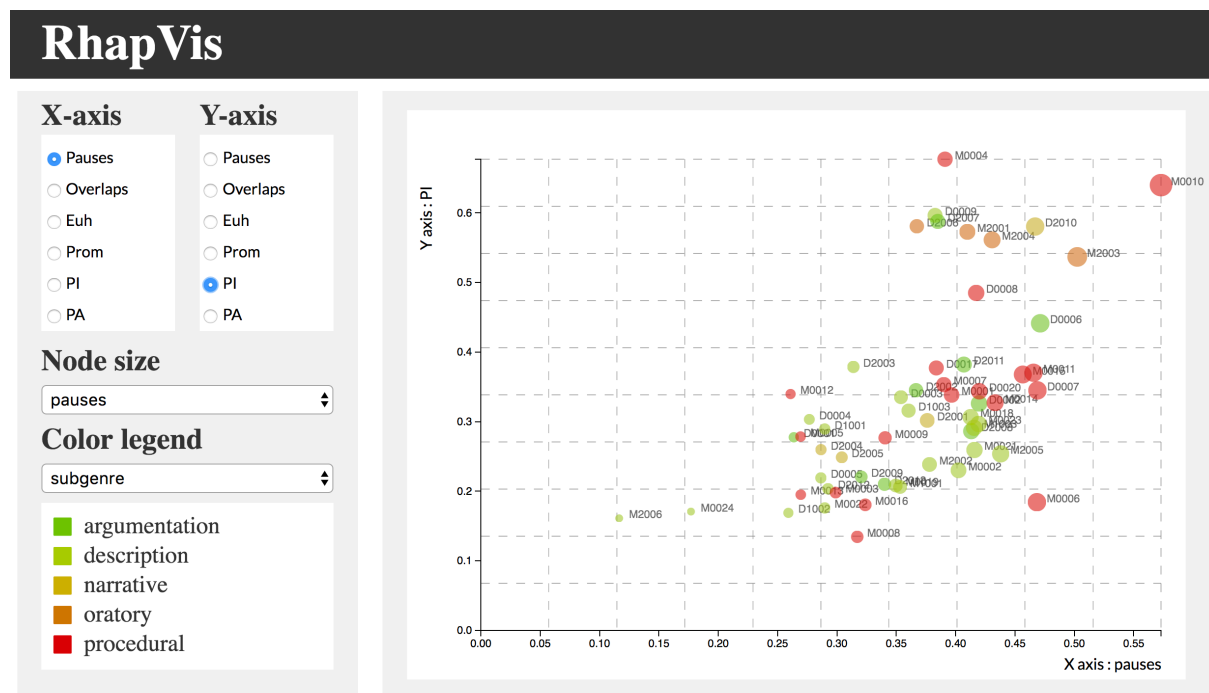


Figure 2. PCA plot of the discourse profiles in Rhapsodie based on prosodic features

