



HAL
open science

Annotation en discours représenté dans un corpus de conversations orales : paramètres, critères et contraintes

Igor Ilić

► **To cite this version:**

Igor Ilić. Annotation en discours représenté dans un corpus de conversations orales : paramètres, critères et contraintes. Posters Day, Feb 2024, Louvain-la-Neuve, Belgique. , 2024. hal-04485519

HAL Id: hal-04485519

<https://hal.science/hal-04485519v1>

Submitted on 1 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Question de **repérage** :

- ➔ ressources lexicales et/ou syntaxiques, mais quid des formes non marquées ?
- ➔ ne retenir que les caractéristiques principales, mais décalage entre théories

réf :
Buvet, P.-A. (2012). Traitement Automatique du Discours Rapporté.

Question d'**interprétation** :

- ➔ annotation de haut niveau – dimension interprétative importante vs nécessité de trancher
- ➔ modèles linguistiques différents

réf :
Lefebvre, J. & Sitri, F. (2023). L'annotation de la RDA : enjeux et problèmes.

C-Oral-Rom : fnatpd01

RUI ben voi- voilà moi aussi je l'ai reçu le samedi:/ et il fallait voter pour le mar- que ça [soit mardi à Paris

PAU [eh faut

OUvrir ta boîte à lettres

RUI ((rires)) oui **mais moi** je l'ouvre/ ma boîte aux lettres

[...]

BRD tu vas pas l'obliger tu vas pas l'obliger à [voter s'il a pas envie de voter/

GRA [ben un peu ouais

ARB non non moi je: je vote pas

CRU non c'est pas **pas envie de voter** c'est plus important que ça\ (0.7) c'est pas l'histoire de **pas envie de voter** sur un mouvement d'humeur/ je trouve que c'est PLUS important que ça la question posée



exemple de Lefebvre et al. (2023)

Spécificités du corpus :

- ➔ interactions verbales – processus dynamiques de (re)négociation avec un contexte évolutif
- ➔ ressources verbales, mais aussi sonores ?

réf :
Auchlin, A., & Grobet, A. (2006). Polyphonie et prosodie: Contraintes et rendements de l'approche modulaire du discours.

Vers un **guide d'annotation** :

- ➔ conception d'un schéma d'annotation avec un jeu d'étiquettes opérationnel
- ➔ prise en compte des variations voco-prosodiques ?

MO [non j'ai pas dit JJB [ceux qui disent ONCEUX [ni l'un ni l'autre] se trompent] j'ai pas dit ça JJB [MO [] j'ai dit MO [les deux se trompent]]

JJB [MO [les deux se trompent]]

MO [oui parce que à gauche il y a cette idée que c'est une chance]

exemple tiré de Ilić (2021)

réf :
Roulet, E., Fillietaz, L. & Grobet, A. (avec la participation de Burger M.) (2001). Un modèle et un instrument d'analyse de l'organisation du discours. Peter Lang.

Quel **protocole** ?

- ➔ phase 1 : annotation manuelle à partir de la transcription
- ➔ phase 2 : étude des contributions de la prosodie sur les occurrences annotées

ANNOTATION EN DISCOURS REPRÉSENTÉ

dans un corpus de conversations orales : paramètres, critères et contraintes

Laboratoire
Linguistique, Langues, Parole
LILPA | UR 1339
Université de Strasbourg

Igor ILIĆ
igor.ilic@unistra.fr

