

Prendre une position épistémique dans le français-en-interaction : l'expression du degré de certitude et de la source de l'information

Introduction et questions de recherche

Ce poster présente les aspects théoriques et méthodologiques, ainsi que des premiers résultats du projet FNS « Prendre une position épistémique dans l'interaction. Les marqueurs du savoir, du non-savoir et du doute en français » (2020-2024). Ce dernier vise à étudier de manière systématique les marqueurs **épistémiques** et **évidentiels** du français, tels qu'ils émergent dans un corpus de 28h de données naturelles, et ce dans une triple approche **énonciative**, **interactionnelle** et **multimodale**. Le projet souhaite ainsi répondre aux questions suivantes :

- Quantitativement, quels sont les marqueurs épistémiques (au sens large) préférés et dans quel(s) contexte(s) d'apparition ?
- Qualitativement, comment les marqueurs épistémiques (au sens large) contribuent-ils à l'élaboration de positions épistémiques particulières dans l'interaction ?

Ancrage théorique

Dans l'interaction, la manière dont le locuteur manifeste à quel point et de quelle manière il dispose d'un savoir dans un certain domaine de connaissance vis-à-vis de ses interlocuteurs constitue une prise de **position épistémique**. Au moyen de ressources linguistiques, le locuteur peut encoder différents états de connaissance, attribuer ou négocier un statut épistémique dans un contexte interactionnel donné (Heritage, 2012). Les ressources lexicales et grammaticales qui marquent un **degré de certitude** relèvent de la catégorie de la **modalité épistémique**, tels que les expressions *je pense, peut-être, probablement*. Les marqueurs de la **source de l'information** relèvent de la catégorie de l'**évidentialité** tels que *il m'a dit, j'ai vu, apparemment*. On considère la modalité épistémique et l'évidentialité comme deux sous-catégories distinctes qui appartiennent au domaine de l'**épistémicité** (Boye, 2012).

Plus récemment, les études interactionnelles ont mis en avant la relation entre les particules épistémiques et l'organisation séquentielle des actions et la construction des relations sociales (Heritage & Raymond, 2005 ; Stivers, Mondada & Steensig, 2011 ; Sidnell, 2012). Ainsi, les analyses quantitative et qualitative de corpus attestés témoignent de la diversité de sens et de fonctions endossés par ces marqueurs.

Corpus et méthodes

Le corpus se divise en deux sous-corpus de taille égale, dont le premier traite de données plus « politiques », avec des **débats publics** (9h) et des **débats télévisés** (5h), tandis que le second rassemble des **réunions de travail** (14h). Les 28h de données totalisent environ 350'000 mots, produits par **145 locuteurs** différents.

Acronyme	Genre et description	Taille
DEBATTRE	Débats publics organisés à l'Université de Lausanne par des associations d'étudiants sur des sujets de société (2007-2009).	Nbr=8 H=9
INFRAROUGE	Débats télévisés à la Radio-Télévision publique Suisse, RTS (2007-2013). Proximité thématique avec les débats publics.	Nbr=5 H=5
E-STAMP	Réunions de travail dans trois entreprises (architecture, ingénierie, communication) de la région lausannoise (2017-2018).	Nbr=13 H=14

Les données ont été intégralement annotées dans ELAN (2020), en suivant les conventions de transcription ICOR. A l'issu du projet, les données seront partagées sur CLAPI. Guidées par une approche **sémasiologique** (mots>sens) de l'étude des marqueurs épistémiques, deux étapes méthodologiques ont été réalisées :

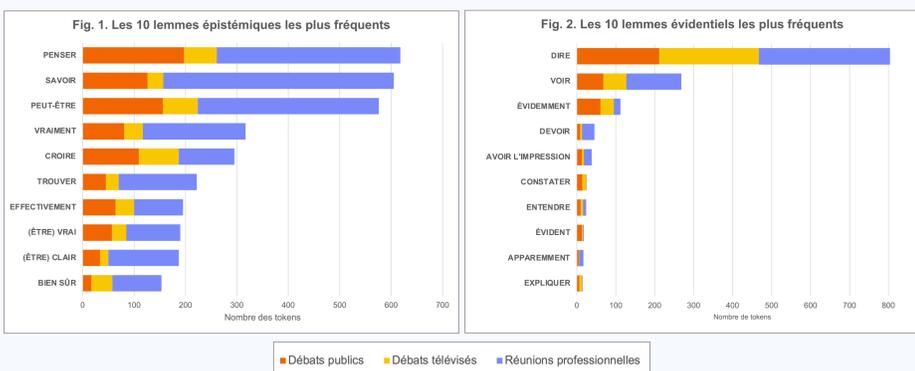
- 1 La première consiste à **rechercher** et **extraire** les occurrences lemmatiques des données, puis à **trier** celles-ci au regard de différents critères, propres au projet et/ou aux deux thèses.

- 2 La seconde consiste à **annoter** dans un tableau à double entrée (Excel) les propriétés morphosyntaxique, énonciatives-discursive, interactionnelle et multimodale (gestes et direction du regard) du **marqueur**, mais aussi la nature de sa **portée** et la relation entre les **deux**. La grille d'annotation est basée en partie sur le projet MODAL (Ghia et al. 2016 ; Pietrandrea & Cervoni 2016) et Pekarek Doehler (2016, 2019). Un échantillonnage a été effectué pour la partie modale du projet en raison du nombre important de tokens concernés.

Un **guide d'annotation**, rédigé collectivement, a été créé, exemplifiant et expliquant la procédure à suivre pour chaque variable.

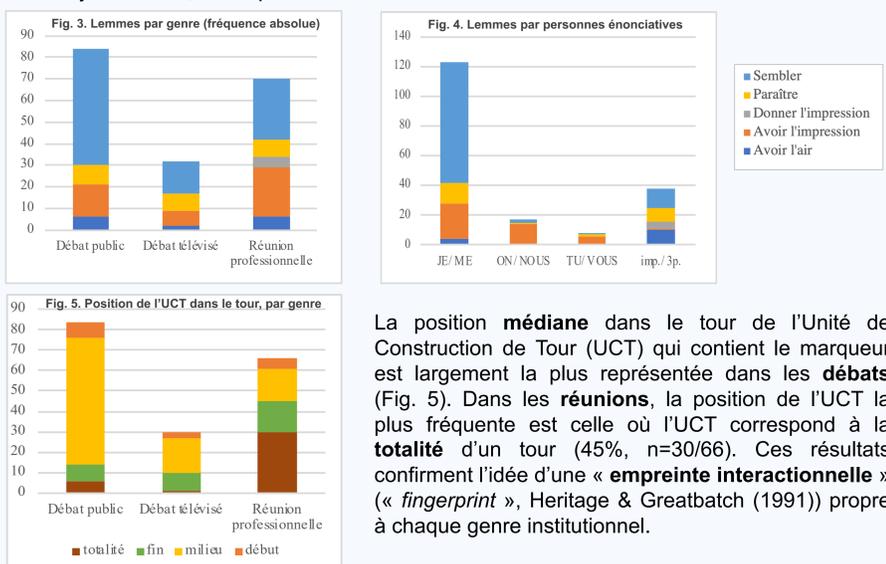
Panorama général

Le corpus recense environ **7'000 marqueurs épistémiques** (au sens large), dont **5'000** relèvent de la **modalité épistémique** et **2'000** de l'**évidentialité**. Les figures ci-dessous présentent les dix lemmes épistémiques (Fig. 1) et les dix lemmes évidentiels (Fig. 2) employés le plus fréquemment dans nos données, ainsi que leur fréquence absolue dans chaque genre institutionnel :



Zoom sur les verbes d'apparence (Jacquin et al., 2022)

Nous avons identifié et annoté **186 tokens** formés sur les verbes d'apparence *sembler, paraître, donner l'impression, avoir l'impression* et *avoir l'air*. On **constate** une nette prédominance du verbe **sembler** (52% du total, n=97/186), et ce dans chaque genre (Fig. 3). On **remarque par ailleurs** que **75%** (n=140/186) des tokens sont **associés à la première personne** (Fig. 4), marquant ainsi une tendance à utiliser les verbes d'apparence comme marqueurs d'un point de vue pris en charge et assumé subjectivement, à titre personnel.



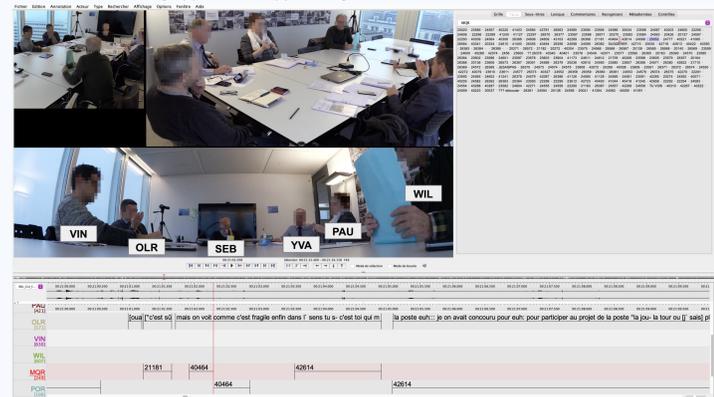
La position **médiane** dans le tour de l'Unité de Construction de Tour (UCT) qui contient le marqueur est largement la plus représentée dans les **débats** (Fig. 5). Dans les **réunions**, la position de l'UCT la plus fréquente est celle où l'UCT correspond à la **totalité** d'un tour (45%, n=30/66). Ces résultats confirment l'idée d'une « **empreinte interactionnelle** » (« *fingerprint* », Heritage & Greatbatch (1991)) propre à chaque genre institutionnel.

Exemple : Source de l'information et incertitude dans une demande de confirmation

L'extrait ci-dessous est tiré d'une réunion de travail entre membres du comité de direction d'un bureau d'ingénieurs. Aux lignes 1 à 8, Olrich (OLR), directeur associé de l'entreprise, demande confirmation à l'administrateur William (WIL) : est-ce bien lui qui lui a communiqué une information et celle-ci est-elle correcte ?

```

REU_CL4 - 00:20:57.740-00:21:20.238
01 OLR  mais **on voit ocomme c'est fragile enfin dans l` sens tu s- c'est toi qui m` disais
      *regarde WIL----->
      **pointe WIL----->
      øfig.6
02  l'aut` jour** la poste euh:: je on avait concouru pour euh: pour participer au projet
      -----**
03  de la poste `la jou- la tour ou [j` sais] plus c` que c'était là`
      ----->
04 SEB  [ouais/]
05 WIL  `mm/°
06 OLR  .h: `et p-° pis qu` *c'était pas *encore fait non plus `c- c'est pas un projet qui s'est`
      -----*regarde YVA-*regarde WIL----->
07  (0.3) `il m` semble qu` tu m'as dit qu` ça ça c'était aussi: euh:° (0.4) [on on sait pas
      ----->
08  c` qu'i`] s'est pa[ssé depuis/]`
      -----*
09 YVA  [remis en
10  question/]
11 WIL  [<((souffle )) H:;>]
  
```



Différentes marques d'hésitations, d'autoréparations ainsi que la présence d'autres marques d'incertitudes cooccurentes au **il me semble que** (ligne 7) témoignent d'un locuteur diminuant fortement sa prétention épistémique (Heritage, 2012). En d'autres termes, Olrich se met en position épistémique basse (K-) vis-à-vis de William (« **c'est toi qui me disais l'autre jour** », ligne 1-2 ; « **je sais plus** », ligne 3) et lui demande confirmation en associant **il me semble que** (p) à une intonation montante (lignes 7-8), élevant ainsi William au statut de source fiable permettant d'assurer le caractère intersubjectif, partagé, de la mémoire en jeu.

Références

Boye, K. (2012). Epistemic meaning: A crosslinguistic and functional-cognitive study. De Gruyter Mouton.
 Ghia et al. (2016). A Construction-centered Approach to the Annotation of Modality. Proceedings of the 12th ISO Workshop on Interoperable Semantic Annotation. Portoroz, 29 may 2016.
 Heritage, J. (2012). Epistemics in Action: Action Formation and Territories of Knowledge. Research on Language & Social Interaction, 45 (1), 1-29.
 Heritage, J. & Raymond, G. (2005). The Terms of Agreement: Indexing Epistemic Authority and Subordination in Talk-in-Interaction. Social Psychology Quarterly, 68 (1), 15-36.
 Heritage, J. and Greatbatch, D. (1991). On the Institutional Character of Institutional Talk: The Case of News Interviews. In Dierdre Boden and Don H Zimmerman (ed.) Talk and Social Structure. Berkeley: University of California Press, 93-137.
 Jacquin, J., Keck, A. C., Robin, C., & Roh, S. (2022). Les verbes d'apparence dans le français-en-interaction. Formes, fonctions et distributions de sembler, paraître, avoir l'air, avoir l'impression et donner l'impression dans un corpus de débats politiques et de réunions d'entreprise. SHS Web Conf., 138, 01012.
 Pekarek Doehler, S. (2016). More than an epistemic hedge: French je sais pas 'I don't know' as a resource for the sequential organization of turns and actions. Journal of Pragmatics, 106, 148-162.
 Pekarek Doehler, S. (2019). At the Interface of Grammar and the Body: Chais pas ('I don't know') as a Resource for Dealing with Lack of Recipient Response. Research on Language and Social Interaction, 52 (4), 365-387.
 Pietrandrea, P., & Cervoni, V. (2016). Modal - annotation guidelines. Version 1.0.
 Sidnell, J. (2012). "Who knows best?": Evidentiality and epistemic asymmetry in conversation. Pragmatics and Society, 3 (2), 294-320.
 Stivers, T., Mondada, L., & Steensig, J. (2011). Knowledge, Morality and Affiliation in Social Interaction. In: T. Stivers, L. Mondada, & J. Steensig (Eds.), The Morality of Knowledge in Conversation. Cambridge: Cambridge University Press, 3-24.
 Stivers, T., Mondada, L., & Steensig, J. (Eds.). (2011). The Morality of Knowledge in Conversation. New York: Cambridge University Press.