

Guide d'annotation du projet POSEPI

Lausanne: Université de Lausanne

V 1.1 – 9.7.2024

Jérôme Jacquin, Ana Claudia Keck, Clotilde Robin, Sabrina Roh

Modifications depuis la version précédente (V 1. – 2.11.2022) :

- Ajout de la référence aux données primaires stockées sur CLAPI
- Ajout de la section 6 : *Résultats des scores d'accords inter-annotateurs*
- Ajout de la section 7 : *Mise à disposition de la base de données*

Comment citer ce document

Jacquin, J., Keck A. C., Robin, C. & Roh, S. (2024 [2022]). *Guide d'annotation du projet POSEPI*. Version 1.1. Université de Lausanne – Fonds National Suisse de la recherche. doi: 10.5281/zenodo.12697145

Table des matières

Guide d'annotation du projet POSEPI	1
Table des matières	2
0. Contexte et objectif du guide	5
1. Opérations préliminaires à l'annotation	6
1.1 Inventaire des lemmes-types	6
1.2 Extraction des lemmes-token par expressions régulières	7
1.2.1 Formulation de l'expression régulière (ER)	7
1.2.2 Recherche dans ELAN et export des résultats	8
1.2.3 Tri : cas des faux positifs	9
1.2.4 Tri : cas des marqueurs complexes	10
1.2.5 Import des résultats dans la grille d'annotation	10
2. Prétraitement	11
2.1 Prétraitement du marqueur	11
2.1.1 Identification du marqueur	11
2.1.1.1 Identification du marqueur dans ELAN	11
2.1.1.2 Identification du marqueur dans Excel (col. A)	11
2.1.2 Apparition ou non dans un discours rapporté (col. B)	11
2.1.3 Absence ou non de la portée propositionnelle (col. C)	12
2.1.4 Intersection avec d'autres marqueurs	12
2.1.4.1 Marqueur(s) ayant la même portée (col. D)	12
2.1.4.2 Marqueur portant sur un autre marqueur (col. E et F)	13
2.2 Prétraitement de la portée	16
2.2.1 Identification de la portée	16
2.2.1.1 Identification de la portée dans ELAN	18
2.2.1.2 Identification de la portée dans Excel (col. AF)	18
3. Annotation du marqueur	18
3.1 Informations relatives à l'événement	18
3.1.1 événement (col. G)	18
3.1.2 année (col. H)	18
3.1.3 fichier vidéo (col. I)	18
3.1.4 genre (col. J)	19
3.2 Informations relatives au marqueur lui-même	19
3.2.1 Lien CLAPI (col. K)	19
3.2.2 Timing début (col. L) et timing fin (col. M) du marqueur	19
3.2.3 Réalisation du marqueur (col. N)	19
3.2.4 Lemme du marqueur (col. O)	19
3.2.5 Forme/expression (col. P)	20
3.2.6 POS-TAG (col. Q)	20
3.2.7 Type morpho-syntaxique (col. R)	20
3.3 Informations relatives au contexte interactionnel	24
3.3.1 Position du marqueur dans l'UCT (col. S)	25
3.3.2 Position de l'UCT dans le TDP (col. T)	28
3.3.3 Rôle de l'UCT dans la séquence (col. U)	30
3.4 Informations relatives au locuteur du marqueur	31
3.4.1 Locuteur du marqueur (col. V)	31
3.4.2 Place interactionnelle du locuteur du marqueur (col. W)	31
3.4.3 Statut socio-professionnel du locuteur du marqueur (col. X)	31

3.5 Informations relatives à l'/aux allocutaire(s) du marqueur	32
3.5.1 Allocutaire du marqueur (col. Y)	32
3.5.2 Place interactionnelle de l'allocutaire du marqueur (col. Z)	33
3.5.3 Statut socio-professionnel de l'allocutaire du marqueur (col. AA)	33
3.6. Informations relatives à la mimo-gestualité	34
3.6.1 Direction du regard (relatif au marqueur)	34
3.6.1.1 Regard avant le marqueur (col. AB)	39
3.6.1.2 Regard pendant le marqueur (col. AC)	39
3.6.1.3 Regard après le marqueur (col. AD)	39
3.6.2 Gestualité (relative au marqueur)	39
3.6.2.1 Présence ou non de geste(s) accompagnant le marqueur (col. AE)	39
3.6.2.2 Type de geste (col. AF)	40
4. Annotation de la portée	41
4.1 Informations relatives à la portée elle-même	41
4.1.1 Identification de la portée (col. AG)	41
4.1.2 Timing début (col. AH) et fin (col. AI) de la portée	42
4.1.3 Réalisation de la portée (col. AJ)	42
4.2 Informations relatives au locuteur de la portée	42
4.2.1 Locuteur de la portée (col. AK)	42
4.2.2 Place interactionnelle du locuteur de la portée (col. AL)	42
4.2.3 Statut socio-professionnel du locuteur de la portée (col. AM)	42
4.3 Informations relatives à l'/aux allocutaire(s) de la portée	43
4.3.1 Allocutaire de la portée (col. AN)	43
4.3.2 Place interactionnelle de l'allocutaire de la portée (col. AO)	43
4.3.3 Statut socio-professionnel de l'allocutaire de la portée (col. AP)	43
5. Annotation de la relation	43
5.1 Direction (col. AQ)	43
5.2 Source (col. AR)	44
5.3 Polarité (col. AS)	45
6. Résultats des scores d'accords inter-annotateurs	48
7. Mise à disposition de la base de données	49
Références	50
Remerciements	51

Légende et convention de transcription des exemples

Convention de transcription (ICOR¹)

[début du chevauchement
]	fin du chevauchement
=	enchaînement rapide
&	continuation du tour après chevauchement
(.)	micro-pause de moins de 0.2 seconde
(2.1)	pause mesurée en secondes et dixièmes de secondes
mo-	troncation d'un mot
mot:	allongement syllabique
/	intonation montante
\	intonation descendante
MOT	voix plus forte
°mot°	voix moins forte
.h	aspiration
(mot)	transcription incertaine

Convention de transcription multimodale (suivant Mondada, 2019)

* *	indication (début/fin) du regard du loc. A
+ +	indication (début/fin) d'une conduite mimo-gestuelle du loc. A
‡ ‡	indication (début/fin) du regard du loc. B
Δ Δ	indication (début/fin) d'une conduite mimo-gestuelle du loc. B
.....	préparation de l'action
~~~~~	rétractation de l'action
*---->	l'action décrite se poursuit sur les lignes suivantes jusqu'au même symbole suivant
---->*	fin de l'action décrite
#	moment auquel la capture d'écran a été prise

### Marqueurs épistémiques

xx	marqueur épistémique retenu pour l'analyse
yy	portée du marqueur considéré
zz	portée d'un marqueur sous la portée du marqueur considéré

### Référencement des exemples

(n) XXX_YYY / DD:DD:DD.DDD - FF:FF:FF.FFF [ID NNNNN]	
(n)	Numéro de l'exemple
XXX	Type d'événement (DPU = Débat public ; DTV = Débat télévisé ; REU = réunion prof.)
YYY	Code de l'événement
DD: ...	Timing du début de l'exemple (heure:minute:seconde.milliseconde)
FF: ...	Timing de la fin de l'exemple (heure:minute:seconde.milliseconde)
NNNNN	Identifiant (ID) du ou des marqueur(s) présent(s) dans l'exemple

¹ [http://icar.cnrs.fr/projets/corinte/bandeau_droit/convention_icor.htm](http://icar.cnrs.fr/projets/corinte/bandeau_droit/convention_icor.htm), consulté le 19 janvier 2022.



## 0. Contexte et objectif du guide

« Prendre une position épistémique dans l'interaction : Les marqueurs du savoir, du non-savoir et du doute en français » est un projet de recherche de 4 ans (2020-2024) financé par le Fonds National Suisse de la recherche [100012_188924] et hébergé à la section des Sciences du Langage et de l'Information de l'Université de Lausanne (<http://www.unil.ch/sli/posepi>). L'objectif du projet est de proposer un panorama systématique, à la fois quantitatif et qualitatif, des marqueurs épistémiques/évidentiels du français, tels qu'ils émergent dans un corpus de 28h d'interactions naturelles vidéo-enregistrées (Jacquin, Etienne, et al., 2024), 14h étant tirées de débats politiques (publics ou télévisés) et 14h de réunions professionnelles en entreprise.

Le corpus a été intégralement transcrit et révisé dans ELAN (2023), selon les conventions ICOR (voir note 1).

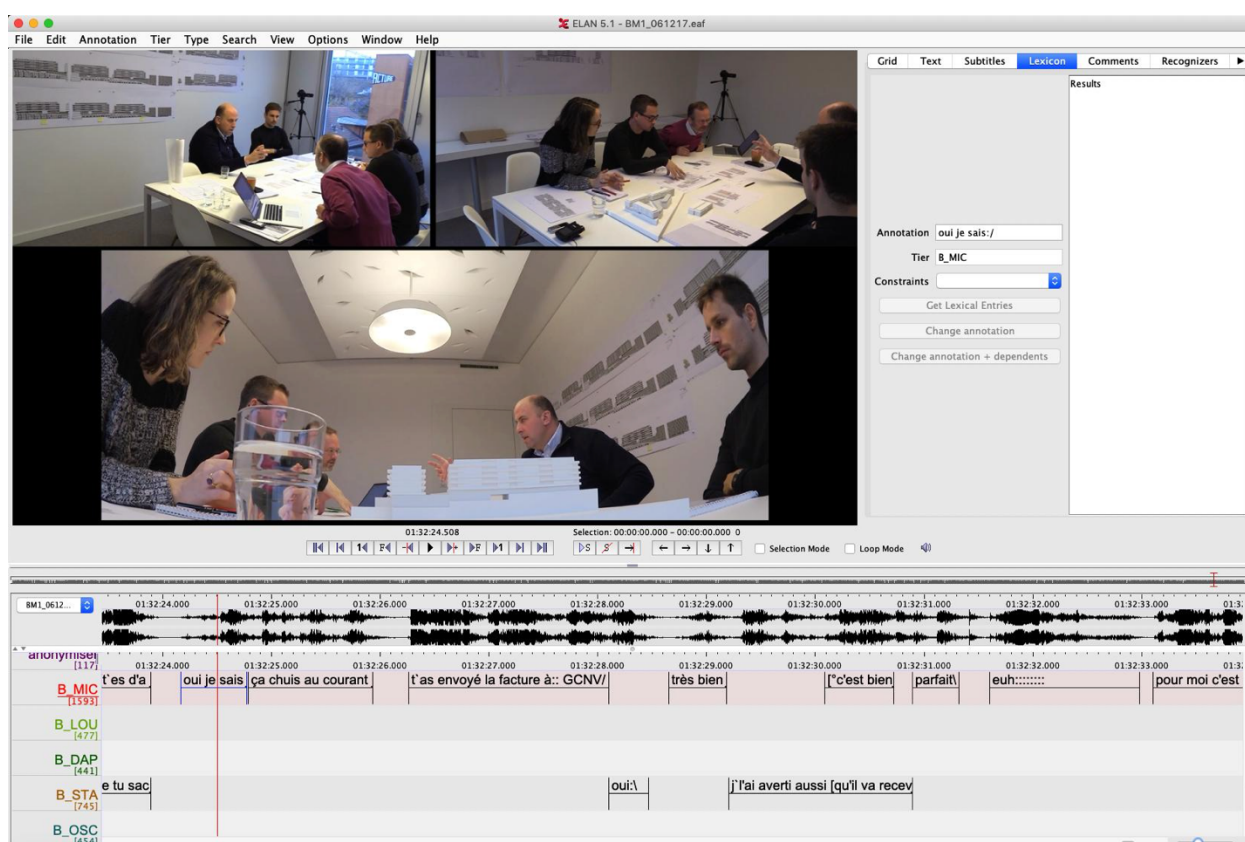


Figure 1 : Capture d'écran du logiciel ELAN.

Le présent guide d'annotation s'inspire en partie des catégories retenues par le projet MODAL (Pietrandrea & Cervoni, 2016) et par l'étude de Pekarek Doehler sur « je sais pas » (2016, 2019). Il introduit les catégories et plus généralement la démarche d'annotation des marqueurs épistémiques dans un tableau Excel à double entrée, qui renseigne ainsi les différentes informations pertinentes relatives à chaque occurrence (par ex. tel « peut-être » intervenant à tel

endroit du corpus). Chaque occurrence est liée aux données primaires par un identifiant permettant de le resituer dans son contexte original (c'est-à-dire dans ELAN²).

Le guide sert à expliciter les choix d'annotation en vue de (i) favoriser au maximum l'accord inter-annotateurs et (ii) faciliter l'appréhension du schéma d'annotation par toute personne intéressée par le projet ou le corpus.

## 1. Opérations préliminaires à l'annotation

### 1.1 Inventaire des lemmes-types

La démarche adoptée dans ce projet pour l'analyse des marqueurs épistémiques – au sens large, c'est-à-dire relevant tant de la modalité épistémique que de l'évidentialité – est de nature sémasiologique, c'est-à-dire que le point de départ est constitué d'une liste de mots ou d'expressions qui expriment un degré de certitude ou une source de l'information, pour conceptualiser dans un deuxième temps leur valeur et fonctionnement dans l'interaction (lemmes > sens). La liste des lemmes épistémiques/évidentiels est élaborée : (1) sur la base de mots ou expressions à caractère épistémique/évidentiel déjà étudiés dans la littérature existante, comme l'adverbe *apparemment* ou encore le verbe *croire*, mais également (2) par recherche sémantique autour des mots ou expressions en question, et finalement (3) par un enrichissement au travers de l'accès au corpus. Cette triple approche permet ainsi de proposer une liste la plus exhaustive possible des lemmes possédant un sémantisme épistémique/évidentiel en français.

On notera que le projet se concentre sur ce que Gosselin (2010) appelle les « marqueurs de modalités *extrinsèques* » (nous soulignons), c'est-à-dire ce qui relève essentiellement du « modus » dans l'analyse modale traditionnelle et qui agit comme opérateur externe sur une portée propositionnelle (voir Figure 2 ci-dessous, branche du milieu). Il ne s'agit pas d'écarter forcément les autres moyens d'expression relevés par l'auteur ou étudiés sur corpus dans d'autres projets (par ex. le projet MODAL, voir Pietrandrea, 2018), mais de débiter par ce qui relève le plus intuitivement, immédiatement et réalistement (sur grand corpus) des *marqueurs* épistémiques/évidentiels, venant travailler une portée propositionnelle au niveau du degré de certitude et de l'origine énonciative.

---

² Un système de renvoi supplémentaire permet de retrouver l'occurrence en contexte dans la base de données en ligne CLAPI (voir *infra*, 3.2.1).

Fig.4

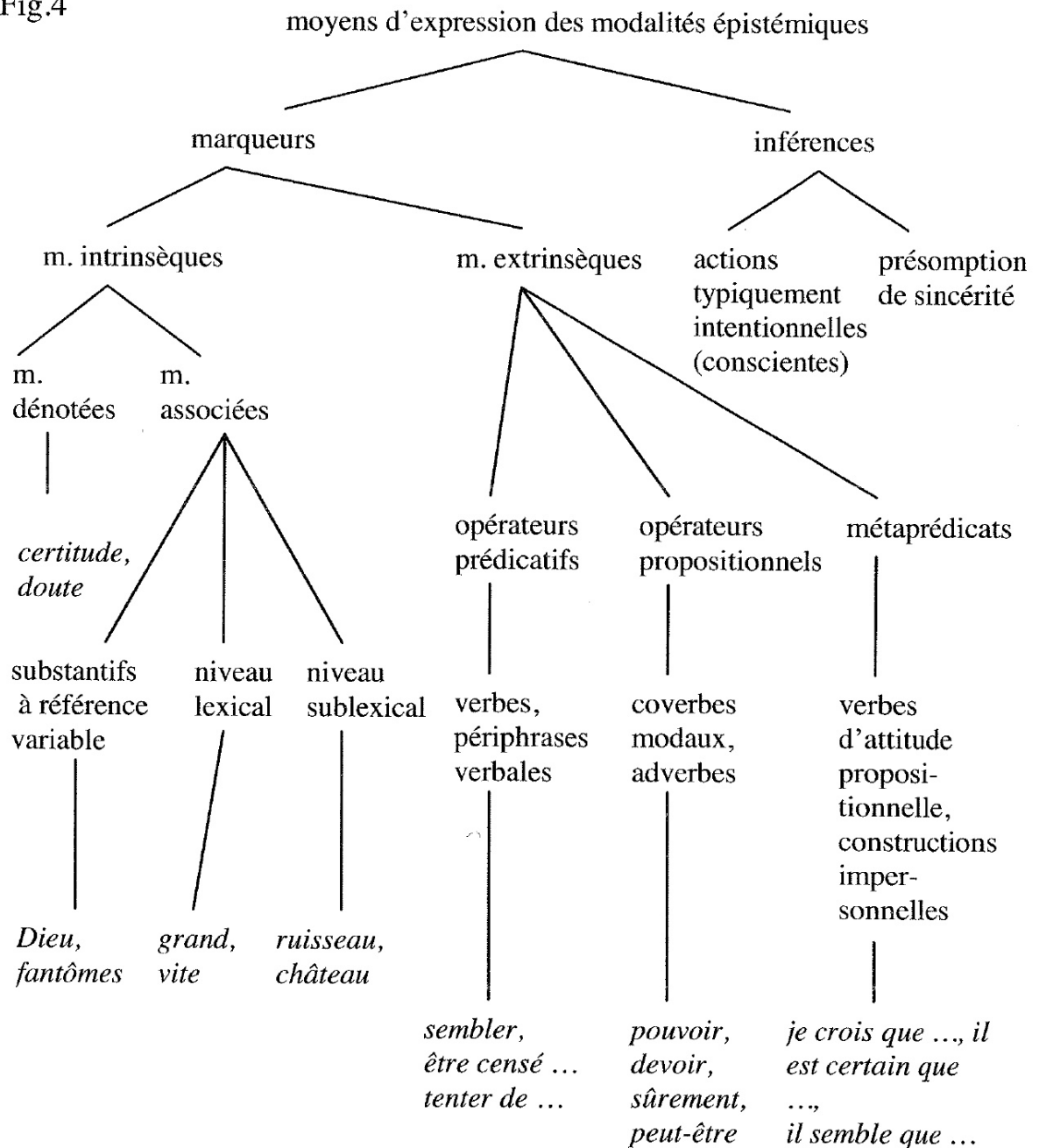


Figure 2 : Les moyens d'expression des modalités épistémiques selon Gosselin (2010, p. 331).

## 1.2 Extraction des lemmes-token par expressions régulières

### 1.2.1 Formulation de l'expression régulière (ER)

Les conventions de transcriptions ICOR utilisées pour le projet exigent d'utiliser des expressions régulières (désormais ER) pour extraire les occurrences de lemmes à partir des transcriptions réalisées dans ELAN. En effet, dans le cas où on cherche simplement l'expression *probablement*, ELAN ne sera pas capable de reconnaître le cas, par exemple, de « proba[blement] » ou « pro:bablement ».

On liste ci-dessous les éléments qui peuvent s'intercaler au sein du lemme recherché, accompagnés de l'ER correspondante :

- Un espace : \s
- Un apostrophe : \'
- Une ouverture de parenthèse : \(
- Une fermeture de parenthèse : \)
- Une barre oblique : \/
- Une barre oblique inversée : \\\
- Un point : \.
- Un caractère numérique : \d
- Un crochet ouvrant : \[
- Un crochet fermant : \]
- Un signe égal : \=
- Deux points superposés : \:
- Les éléments ci-dessus autant de fois que nécessaire : {0, }

Cela donne l'ER suivante, à copier-coller entre chaque caractère du lemme recherché :  
 (\s|\'|\\(|\\)|\\\\|\\.|\\d|\\[|\\]|\\=|\\:){0, }

Pour « probablement », cela donne ainsi :  
 p(\s|\'|\\(|\\)|\\\\|\\.|\\d|\\[|\\]|\\=|\\:){0, }r(\s|\'|\\(|\\)|\\\\|\\.|\\d|\\[|\\]|\\=|\\:){0, }o(\s|\'|\\(|\\)|\\\\|\\.|\\d|\\[|\\]|\\=|\\:){0, }b((\s|\'|\\(|\\)|\\\\|\\.|\\d|\\[|\\]|\\=|\\:){0, } ...

Chaque ER est testée sur regex101.com. Le *Test String* peut être l'export d'une transcription complète d'un fichier ELAN.

**Attention** : en général, on veille à utiliser une ER trop puissante plutôt que pas assez. Par exemple, on cherchera plus volontiers « être » pour trouver les occurrences du lemme « peut-être », même si cela exige ensuite de trier les résultats pour écarter les faux positifs (voir 1.2.3 ci-dessous). De même, on cherchera « affirm » pour trouver tous les cas basés sur le lemme « affirmer » (« tu affirmes », « Xavier lui a affirmé », « il a peut-être affirmé », etc.).

## 1.2.2 Recherche dans ELAN et export des résultats

Pour opérer une recherche dans ELAN, on sélectionne *Rechercher EAF multiples* dans le menu *Rechercher* de la barre supérieure. Dans la fenêtre qui s'ouvre, on commence par *Définir le domaine de recherche* en sélectionnant et chargeant le dossier où la recherche sera effectuée (dans le cas présent, le dossier partagé intitulé « 1 CORPUS »).

On utilise ensuite l'ER en cochant l'option *expression régulière* avant de lancer la fonction *Rechercher*. Dans la fenêtre de résultats, on vérifie que le nombre de fichiers examinés est correct (26 fichiers dans le cas d'une recherche POSEPI).

On choisit ensuite d'*Exporter* et de sauvegarder l'export dans un dossier de traitement provisoire. Le fichier exporté doit être manuellement renommé en ajoutant l'extension « .txt ». Dans un tableau Excel vierge, on utilise l'option *Importer* du menu *Fichier*. On choisit alors l'option « Fichier texte » ainsi que « Unicode (UTF-8) », avant de poursuivre et d'enregistrer le résultat de l'import dans le dossier de traitement provisoire.

Attention : ELAN exporte les annotations qui contiennent l'ER recherchée, et non les occurrences de l'ER. Cela signifie qu'une annotation contenant plus qu'une fois l'ER recherchée (par ex. deux fois « peut-être ») n'est exportée qu'une seule fois et qu'il faut manuellement dupliquer les lignes concernées pour avoir autant de lignes que d'occurrences.

### 1.2.3 Tri : cas des faux positifs

Selon l'ER choisie pour extraire les occurrences d'un lemme, il se peut que des faux positifs (occurrences prises en considération par l'ER mais qui ne correspondent pas à un marqueur épistémique/évidentiel) apparaissent dans les résultats, auquel cas un tri manuel est nécessaire.

Plusieurs cas sont distingués :

- a. Les faux positifs au niveau lemmatique sont systématiquement supprimés. Ex. une occurrence de *savoir-être* ou de *peut être parmi* les résultats de la recherche relative au marqueur *peut-être* (par l'ER ciblant « être »), ou alors *distinction* qui apparaît lors de la recherche de *dis*, etc.
- b. De la même manière, les cas d'homophonie claire sont systématiquement écartés. Ex. *savoir* (ou *paraître*) comme nom alors que se sont les marqueurs contenant le lemme *savoir* (ou *paraître*) en tant que verbe qui sont recherchées.
- c. De manière cruciale, sont également exclues à ce stade toutes les occurrences ne portant pas sur une proposition complète. Sont donc exclus les cas de portée nominale (« je crois au Père Noël »). S'agissant de portées propositionnelles partielles ou tronquées, elles sont conservées et annotées (e.g. « je crois qu'il est parti faire - » ; « c'était je crois - »). Lorsque la portée propositionnelle est totalement inexistante, par exemple du fait d'une interruption, l'entrée est conservée dans la base, mais non annotée (voir *infra*, 2.1.3).
- d. Sont exclues également les occurrences modales qui ne sont pas utilisées de manière m-performative, c'est-à-dire qui ne modulent pas la prise en charge énonciative d'une proposition par le locuteur au moment où il parle (par ex. utilisation de l'imparfait dans « je croyais qu'il était sorti ») (Nuyts, 2001; voir également Keck, en préparation, pour une discussion).
- e. Sont aussi enlevées à ce stade les occurrences lemmatiques qui ne respectent pas les critères de définition d'un marqueur évidentiel, tels que proposés dans Robin (en préparation). Ainsi, sont exclus les cas qui (i) n'ont pas comme valeur primaire une valeur évidentielle (ex. « oui mais tu vois quoi » ou « j'entends » au sens de *comprendre*), (ii) ne sont pas imputables au locuteur comme expérimentateur (L=E), seul ou accompagné d'autrui ((L=E) + (A=E)) (ex. « il a vu que »), (iii) se présentent comme des antécédents de conditionnel (ex. « si je vois que »), traduisent des projections hypothétiques (ex. « imagine là il dit que », « il a l'air de dire ») et, de manière cruciale, sont employés au futur (ex. « il expliquera que »), exception faite des futurs dits « réalisés » et (iv) sont accompagnés d'une particule négative, lorsque celle-ci annule la valeur évidentielle (ex. « j'ai pas vu que Marie était enceinte »).
- f. Les cas de polysémie non traités en amont par une méthodologie rigoureuse et discutée collectivement sont conservés (par ex. *peut-être* utilisé non pas comme atténuateur de certitude mais comme atténuateur d'acte directif, comme dans « on va peut-être passer la parole à X »).

### 1.2.4 Tri : cas des marqueurs complexes

Il peut arriver que plusieurs lemmes épistémiques/évidentiels se combinent et prennent la forme d'un marqueur complexe. Dans ce cas, une seule entrée, correspondante au lemme noyau, est introduite dans la base de données, tandis que les lemmes associés sont écartés et apparaîtront donc dans les colonnes correspondant à la réalisation et à la forme du marqueur. On portera donc une attention particulière aux situations suivantes :

- a. Sont écartés à ce stade les cas de verbes modaux (*pouvoir, devoir*) portant sur des verbes à l'infinitif qui font eux-mêmes partie des lemmes recherchés. En effet, suivant par exemple Willems (2011), ces associations de verbes modaux avec des infinitives fonctionnent comme des « prédicats complexes » et ne peuvent donc être traités comme deux marqueurs indépendants l'un de l'autre. Ainsi, on conservera et annotera en tant que marqueur autonome (= 1 entrée dans la base) l'occurrence de *pouvoir* (« il a pu ») dans « il a pu tomber ». À l'inverse, l'expression *on pourrait penser* identifiée à partir d'une recherche autour du lemme *pouvoir* est écartée au profit du seul lemme *penser*, qui sera considéré comme le lemme-noyau du marqueur « on pourrait penser ». Il en va de même pour le conditionnel associé à ce « pourrait », qui ne fait pas non plus l'objet d'une entrée autonome dans la base (voir b. ci-dessous).
- b. De même, est plus généralement écarté le cas du conditionnel associé à un lemme verbal lui-même épistémique/évidentiel ou associé à une construction hypothétique en « si ». Ainsi, le conditionnel sera considéré comme un marqueur annoté dans « il aurait des ennuis », mais pas dans « on pourrait penser », « j'y croirais » ou « s'il venait j'aurais des ennuis ». Pareillement, dans le cas où le conditionnel est associé à un verbe modal seul (non associé à un verbe épistémique à l'infinitif, voir ci-dessus, a.), comme dans « je pourrais avoir réussi », le conditionnel n'est pas entré comme entrée autonome, au profit du verbe modal.
- c. Attention : l'adverbe ou syntagme adverbial épistémique/évidentiel fait toujours l'objet d'une entrée autonome, quel que soit son contexte d'apparition (« X semble manifestement Y » fera donc l'objet de deux entrées ; pareil pour « X semble clairement démontrer que Y »).

Les cas d'articulation entre lemmes traités comme marqueurs autonomes sont traités ci-dessous, 2.1.3.

### 1.2.5 Import des résultats dans la grille d'annotation

Il s'agit ensuite de copier-coller les colonnes exportées dans les colonnes correspondantes dans la grille d'annotation (par exemple, locuteur, événement, lemme, etc.).



## 2. Prétraitement

### 2.1 Prétraitement du marqueur

#### 2.1.1 Identification du marqueur

##### 2.1.1.1 Identification du marqueur dans ELAN

Dans ELAN, sur le **tier** (ou « acteur ») MQR, introduire une annotation correspondant à la totalité (du début à la fin) de production du marqueur (Ex. la portion correspondant à l'émission de « peut-être » ou de « on pourrait penser »). Dans l'annotation elle-même, introduire un code d'identification unique à 5 chiffres, entre 10000 et 99999³.

Sont considérés comme intégrant le marqueur, tous les éléments qui lui sont syntaxiquement associés. Cela comprend les cas :

- de dislocations à gauche ou à droite (ex. « moi je **pense** que », « ça ça me paraît **évident** », « là là c'est **écrit** »).
- d'auto-réparations (« c'est pas c'est pas **envisageable** de »), à l'exception des cas où deux marqueurs syntaxiquement complets se suivent, qui sont dès lors distingués (« c'est pas **possible** c'est pas **envisageable** de » => « c'est pas **possible** » + « c'est pas **envisageable** de », ayant la même portée ; voir aussi l'exemple ci-dessous en 2.1.3).
- des circonstants spatiaux ou temporels (« pour le moment j'ai l'**impression** »), mais pas des connecteurs ou subordonnants régisseurs, qui sont supprimés (« si vous avez l'**impression** », « alors il **paraît** que »).

##### 2.1.1.2 Identification du marqueur dans Excel (col. A)

Introduire le code d'identification unique du marqueur utilisé dans ELAN dans la colonne A.

##### 2.1.2 Apparition ou non dans un discours rapporté (col. B)

Il peut arriver que des marqueurs apparaissent au sein de discours rapportés, comme c'est le cas dans l'exemple ci-dessous. Ces cas ne sont pas étudiés dans le cadre du projet, qui se concentre sur les positionnements épistémiques des participants à l'interaction. Lorsque l'annotateur choisit « oui », le reste de la grille ne doit donc pas être rempli, à l'exception des colonnes remplies suite à l'import des résultats (événement, locuteur, lemme, etc.) (voir 1.2.5 *supra*).

```
(1) REU_BL2 / 00:21:16.869 - 00:21:22.768 [ID 40234]
1 FLO      il l'enquirlait (iks) fois/ tu vois il arrive vers lui
2          °pis il disait° <((change de timbre)) alors il paraît que::>
3          °h.: j'étais là° °mais° [QU'EST-ce:] qu'il lui °dit encore°
4 SOP                                [ah ouais/]
```

³ Les codes d'identification uniques sont répartis comme suit : 10000-19999 (disponibles pour le groupe, à discuter au cas par cas); 20000-39999 (disponibles pour Ana Keck); 40000-59999 (disponibles pour Clotilde Robin); 60000-79999 (disponibles pour Sabrina Roh); 80000-99999 (disponibles pour Jérôme Jacquin).

### 2.1.3 Absence ou non de la portée propositionnelle (col. C)

Il peut arriver que la portée propositionnelle d'un marqueur soit totalement absente, par exemple dans le cas d'une interruption. Dans ce cas, l'annotateur choisit « non » et le reste de la grille ne doit donc pas être rempli, à l'exception des colonnes remplies suite à l'import des résultats (événement, locuteur, lemme, etc.) (voir 1.2.5 *supra*).

Ainsi, dans le cas suivant, « je pense pas que » n'a pas de portée, contrairement au « je pense que » qui suit.

```
(2) REU_BL4 / 00:19:02.034 - 00:19:10.015 [ID 23393]
1 LUC     pis au final chacun a commen- a continué à travailler un peu
2         comme il travaillait avant::/ et [j` pense]&
3 CHR                                           [mh]
4 LUC     &pas qu- j` pense qu'on a amené deux TROIS choses
5         mais (0.6) on n'a pas révolutionné non plus/
```

Attention : l'absence de portée ne doit pas être confondu avec la situation où deux marqueurs se suivent pour la même portée propositionnelle, que celle-ci soit avant (« [P] c'est reconnu c'est prouvé ») ou après (« je sais je crois que [P] ») les marqueurs en question.

### 2.1.4 Intersection avec d'autres marqueurs

Il se peut que des marqueurs soient liés les uns aux autres. Deux cas sont distingués :

#### 2.1.4.1 Marqueur(s) ayant la même portée (col. D)

Il peut arriver que le marqueur considéré ait une portée préalablement déjà identifiée comme portée pour un autre marqueur. Deux marqueurs (au moins) se retrouvent donc à avoir la même portée. Dans cette situation, les deux marqueurs sont le plus souvent sur le même plan syntaxique (pas d'enchâssement de l'un dans l'autre), comme c'est visible dans l'exemple ci-dessous.

```
(3) DPU_ETU / 01:17:16.538 - 01:17:23.248 [ID 41357; 42004]
1 MAR     ils ont dû abandonner\ parce que ça n'est pas possible\
2         déjà ça leur coûte\ administrativement beaucoup plus que
3         ça ne rapporte/ ça c'est prouvé c'est pas moi qui le dis\
```

Dans ce cas de figure, l'annotateur indique dans la colonne D l'identifiant unique du ou des marqueur(s) concerné(s). Dans le cas où plus de deux marqueurs possèdent la même portée, des points-virgules, sans espace, sont utilisés pour séparer les identifiants (ex. 12309;23467). Attention : symétriquement, l'annotateur doit introduire les identifiants pertinents dans la col. D pour tous les marqueurs déjà identifiés.

Colonne A	Colonne D
ID marqueur	Marqueur(s) ayant la même portée
41357	42004
42004	41357



La situation est analogue dans l'exemple ci-dessous, où « ça c'est évident » et « ça ça me paraît tout à fait clair »⁴ constituent deux marqueurs ayant la même portée.

(4) DPU_UDC / 00:32:55.873 - 00:33:05.943 [ID 41143; 28474]  
 1 CH1 alors oui/ on va c'est une tendance heu systémique pour mille  
 2 raisons: (0.3) heu: y a une tendance de personnalisation: de mise  
 3 en spectacle de la politique ça c'est évident\ (.) ça ça me paraît  
 4 tout à fait clai:r

Voici un dernier exemple, où la portée se trouve cette fois-ci entre les deux marqueurs :

(5) REU_BM2 / 00:10:29.146 - 00:10:32.484 [ID 25166 ; 42433]  
 1 CLA °j`trouve° c'est pas:: c'est pas une bonne idée:/ mais: je l'ai  
 2 déjà dit\

#### 2.1.4.2 Marqueur portant sur un autre marqueur (col. E et F)

Il peut arriver qu'un marqueur porte sur un autre marqueur. C'est généralement le cas lorsqu'il y a enchâssement syntaxique entre deux marqueurs (hors cas des marqueurs complexes traités en 1.2.4, qui font de toute manière l'objet d'une seule entrée). Suivant les cas, un marqueur peut porter sur un autre marqueur uniquement, ou sur un autre marqueur associé à sa portée. C'est le cas ci-dessous, où le marqueur « je pense pas que » porte sur le marqueur « ça pourrait » et sa portée « être la seule idée ».

(6) REU_AC1 / 00:18:16.837 - 00:18:26.424 [ID 23263; 27583]  
 1 JEA on: eh moi] je m'excuse mais je vais directement au but parce que  
 2 j'ai eu une idée (0.5) j`pense pas que ça pourrait être la seule  
 3 idée mais sérieusement j'y crois beaucoup à cette idée en fait  
 4 c'est le mot bulle

La colonne E permet d'indiquer l'identifiant du marqueur englobé (dans l'exemple : 27583 équivalant à « ça pourrait ») et, symétriquement, la colonne F permet d'indiquer l'identifiant du marqueur englobant (dans l'exemple : 23263 équivalant à « je pense pas que »). Dans le cas où plus de deux marqueurs sont sous la portée d'un autre marqueur, des points-virgules, sans espace, sont utilisés pour séparer les identifiants (ex. 12406;24567).

Colonne A	Colonne E	Colonne F
ID marqueur	Marqueur portant sur	Marqueur sous la portée de
23263 (je pense pas que)	27583	
27583 (ça pourrait)		23263

NB : comme on l'a vu plus haut (1.2.4), le conditionnel de « pourrait » n'est pas autonomisé en tant que marqueur dans ce cas du fait qu'il est associé au verbe modal « pouvoir », qui sert dès lors à enregistrer l'entrée du marqueur « ça pourrait ».

(7) DTV_DEM / 00:04:15.483 - 00:04:43.617 [ID 20079; 40238]  
 1 MAI i` faut simplement essayer d`comprendre ce qui s` passe\ (.) .h  
 2 c` qui se passe c'est qu` y a un parti: qui semble manifestement  
 3 aux abois\ (.) parce qu` y a un certain nombre de choses qui sont

⁴ On parle bien ici du marqueur « ça ça me paraît tout à fait clair », organisé autour du lemme « clair », et non du marqueur « me paraît » qu'il contient.

4 dites dans un RApport Officiel: (.) de la commission d`gestion

De même, dans l'exemple ci-dessus, « semble » et « manifestement » sont considérés comme deux marqueurs autonomes, « manifestement » portant sur l'association du marqueur « semble » à sa portée (« un parti qui semble {manifestement} aux abois »).

Ci-dessous, on trouvera d'autres exemples où un marqueur porte sur un autre marqueur et sa portée :

(8) REU_BL2 / 00:20:58.508 - 00:21:00.113 [ID 23352 ; 20869]  
1 FLO °donc ch-° (.) °p-° j` pense vraiment qu` [c'est une] piste à  
2 creuser/ .h  
3 SOP [°hm°]

Le marqueur « vraiment » porte sur « j` pense » et sa portée « qu' c'est une piste à creuser ».

(9) REU_BL1 / 00:12:55.630 - 00:12:58.288 [ID 23336; 40252]  
1 ANT [moi j` pense/ vu l'en]vironnement c'est plutôt un peu quelque  
2 °chose de rustique/ quoi\ [(hein/ )°]

Le marqueur « vu l'environnement » porte sur « moi j` pense » et sa portée « c'est plutôt un peu quelque chose de rustique ».

(10) DTV_VEH / 00:51.41.386 - 00:51:57.248 [ID 20047 ; 41155]  
1 PET le véhicule a plus de chances de se retourner et heu: (.) plus  
2 de risques de décès pour ses occupants.h et il a plus de  
3 risques pour les piétons ou: en cas de collision ou les vélos  
4 ou les motards .h donc juste pour cette raison ça nous paraît  
5 évident que ce type de véhicules pas tous (.) mais ceux qui  
6 sont dangereux: ou trop dangereux doivent être retirés de la  
7 vente/

Le marqueur « nous paraît » [20047] porte sur un autre marqueur associé à sa portée (« ça nous paraît évident que ce type de véhicules pas tous... »).

Il peut aussi arriver qu'un marqueur porte non pas sur l'association d'un autre marqueur et de sa portée, mais uniquement sur un autre marqueur. Attention : il s'agit donc d'un cas distinct des cas de combinaisons de lemmes épistémiques/évidentiels traités ci-dessus, 1.2.4.).

(11) DPU_REC / 00:27:47.850- 00:28:03.210 [ID 40122 ; 28701]  
1 P22 ce qui était pas l` cas il y avait quand même trois sites un à  
2 singapour un en irlande et un en [suisse] qui étaient en  
3 concurrence/ c'est celui irlandais qui a été choisi\ donc les  
4 mille cinq cent emplois étaient hypothétiques c'est pas::[:  
5 NA2 [non non  
6 et (.) c'était évidemment hypothétique\ (.) mais quand on est en  
7 concurrence internationale sur un projet/ (0.4) euh (.) si les dans  
8 les trois villes qui existent on fait une proposition fiscale  
9 intéressante pour faire venir le l'entreprise/

Dans cet exemple, « c'était évidemment hypothétique » (marqueur 1 sous la forme d'une proposition complète autour du lemme « hypothétique »), porte sur la proposition soulignée (lignes 3-4), tandis que « évidemment » (marqueur 2) porte sur le marqueur 1. La portée de ce « évidemment » est indiquée « c'était {évidemment} hypothétique ».

Ci-dessous sont encore proposés deux exemples complexes, qui articulent les deux cas qui viennent d'être discutés.

(12) DPU_UDC / 00:15:42.324 - 00:16:09.342 [ID 42042; 21502; 40088]  
1 CHE il souvent il fallait lire entre les lignes heu tout n'était pas  
2 extrêmement clair difficile d'en déduire un positionnement général  
3 du journal c'est pas la nature du journal d'ailleurs/ (0.4) et son  
4 rédacteur en chef\ (0.2) heu::: roth **disait** à **juste** titre me **semble-**  
5 t-il qu'au fond/ (.) c'est heu (.) le temps a beaucoup fonctionné  
6 dans cette campagne dans l` structure d'accueil de positions an-  
7 de de de positions de (0.3) de personnes qui: se sont EXPRIMÉES à  
8 l'intérieur des colonnes du temps (.) sans que ce soit l'avis du  
9 temps\

Trois marqueurs peuvent être identifiés :

1. « son rédacteur en chef roth disait à juste titre me semble-t-il que » [42042] porte sur « au fond c'est heu... ».

(12) DPU_UDC / 00:15:42.324 - 00:16:09.342 [ID 42042]  
1 CHE il souvent il fallait lire entre les lignes heu tout n'était pas  
2 extrêmement clair difficile d'en déduire un positionnement général  
3 du journal c'est pas la nature du journal d'ailleurs/ (0.4) et son  
4 **rédacteur en chef\ (0.2) heu::: roth disait à juste titre me semble-**  
5 **t-il qu'au fond/ (.) c'est heu (.) le temps a beaucoup fonctionné**  
6 **dans cette campagne dans l` structure d'accueil de positions an-**  
7 **de de de positions de (0.3) de personnes qui: se sont EXPRIMÉES à**  
8 **l'intérieur des colonnes du temps (.) sans que ce soit l'avis du**  
9 **temps**

2. « à juste titre » [21502] porte sur le marqueur 1 et sa portée

(12) DPU_UDC / 00:15:42.324 - 00:16:09.342 [ID 21502]  
1 CHE il souvent il fallait lire entre les lignes heu tout n'était pas  
2 extrêmement clair difficile d'en déduire un positionnement général  
3 du journal c'est pas la nature du journal d'ailleurs/ (0.4) et son  
4 **rédacteur en chef\ (0.2) heu::: roth disait à juste titre me semble-**  
5 **t-il qu'au fond/ (.) c'est heu (.) le temps a beaucoup fonctionné**  
6 **dans cette campagne dans l` structure d'accueil de positions an-**  
7 **de de de positions de (0.3) de personnes qui: se sont EXPRIMÉES à**  
8 **l'intérieur des colonnes du temps (.) sans que ce soit l'avis du**  
9 **temps**

3. « me semble-t-il » [40088] porte sur le marqueur 2 uniquement (« à juste titre »), dont on admet dès lors qu'il possède une portée propositionnelle du fait qu'il est lui-même associé à une portée propositionnelle.

(12) DPU_UDC / 00:15:42.324 - 00:16:09.342 [ID 40088 ; 21502]  
1 CHE il souvent il fallait lire entre les lignes heu tout n'était pas  
2 extrêmement clair difficile d'en déduire un positionnement général  
3 du journal c'est pas la nature du journal d'ailleurs/ (0.4) et son  
4 **rédacteur en chef\ (0.2) heu::: roth disait à juste titre me semble-**  
5 **t-il qu'au fond/ (.) c'est heu (.) le temps a beaucoup fonctionné**

6 dans cette campagne dans l` structure d'accueil de positions an-  
 7 de de de positions de (0.3) de personnes qui: se sont EXPRIMÉES à  
 8 l'intérieur des colonnes du temps (.) sans que ce soit l'avis du  
 9 temps\

Colonne A	Colonne E	Colonne F
ID marqueur	Marqueur portant sur	Marqueur sous la portée de
42042		21502
21502	42042	40088
40088	21502	

L'exemple suivant articule deux marqueurs.

(13) DPU_UDC / 00:13:55.500 - 00:14:36.701 [ID 20118 ; 26100]  
 1 CH1 (.) alors\ (0.4) soyons concrets maintenant/ segmentation  
 2 par exemple dans la dernière campagne suisse romande/ (.) suisse  
 3 alémanique\ (0.4) on est là en::: suisse romande aujourd'hui/  
 4 (0.7) vous avez probablement aussi (.) peut-être vu c` qui s`  
 5 passait ou C` QUI S`PASSAIT PAS en suisse alémanique (0.4)  
 6 soyons très clairs en suisse alémanique/ (0.6) y a eu pratiquement  
 7 (.) aucune (.) discussion de FOND\ (0.5) sur:: (.) la campagne de  
 8 l'u d c est-elle acceptable est-elle xénophobe ne l'est-elle pas  
 9 etcétera\ UN SEUL JOURNAL/ (0.5) a fait réellement\ (0.6) un  
 10 travail ASSEZ orienté qu'on pourrait qualifier (0.5) de d'anti-u d  
 11 c ou d'anti-bloch- ou de très critique à l'égard des thèses de  
 12 l'u d c et de et de christoph blocher\ c'est le blick\ (0.7)

1. Marqueur 1 [20118] : « probablement » qui porte sur le marqueur 1 uniquement : « vous avez {probablement} aussi peut-être vu »
2. Marqueur 2 [26100] : « peut-être » qui porte lui aussi sur le marqueur 1 uniquement : « vous avez probablement aussi {peut-être} vu »

## 2.2 Prétraitement de la portée

### 2.2.1 Identification de la portée

La portée renvoie à la proposition dont la prise en charge est travaillée par le marqueur. Le projet MODAL (Pietrandrea & Cervoni, 2016) définit la portée de la manière suivante: « *The scope of an epistemic construction corresponds to the element of the context to which the marker attributes the truth value. As a truth value bearer, the scope of an epistemic construction expresses, by definition, a proposition, i.e., an abstract semantic object capable of being evaluated in terms of truth. From a formal standpoint, the proposition to which the truth value is attributed can be expressed by either a sentence or an utterance.* »

Lorsque le marqueur porte syntaxiquement sur un anaphorique (Ex. « ça » dans « ça j'en sais rien »), la portée est ce à quoi l'anaphorique fait référence, par exemple la proposition précédente.

Par ailleurs, sont considérés comme intégrant la portée d'un marqueur, tous les éléments syntaxiquement associés à la proposition identifiée comme portée. Cela comprend les cas :

- de dislocations à gauche ou à droite ou (pseudo-)clivage (ex. « elle elle vient » ; « il a pris cher Pierre » ; « le le deuxième phénomène, c'est {certainement} que le langage UDC

qui était en rupture totale avec la culture politique suisse, qui était soi-disant trop politiquement correcte ou cette fameuse démocratie trop sage, ça interpelle toujours les médias et ça marche toujours dans les médias » ; « c'est là {TRÈS clairement\} que votre bien commun... »).

- d'auto-réparations (ex. « il elle est venue en train »), à l'exception des hésitations pré- ou post-portée (ex. « tout à coup les gens vont changer: h-heu-heu »), qui sont supprimées.
- d'antécédent de propositions relatives (ex. « la nature qui serait menacée » ; « les politiciens qui semblent avoir tout faux »).
- des circonstants spatiaux ou temporels (ex. « hier elle est venue en train » ; « en angleterre c'est quelque chose comme... » ; « aujourd'hui on a dessiné ... »), à l'exception des connecteurs (ex. « donc moi je suis pour garder une part de disons pas forcément de rêve ») et subordonnants (ex. « si aujourd'hui près d'un suisse sur trois: qui vote u d c c'est c' que révèle à peu près le sondage »), qui sont enlevés, tout comme les modalisations d'énonciation (« à mon avis / peut-être il est arrivé »)⁵. Attention, cette exception ne s'applique toutefois pas aux portées qui sont régies par un introducteur de discours rapporté.
- des propositions associées à un marqueur subordonnant, qu'il s'agisse de relatives nominales (« c'est **évident** que on visera plutôt les gens qui sont à former, qui ont peut-être deux trois ans d'expérience ») ou de complétives (« y a probablement vingt ou trente ans: (.) on on **disait** les bon les glaciologues au fond ça sert à rien\ (.) qu'est-ce que ça ces ces gens-là/ »). Dans le cas d'une circonstancielle (ex. « y a terminer {**peut-être**} bouche de saltier si y a des choses à faire »), l'annotateur ou l'annotatrice décide, au cas par cas et sur la base du sens de l'énoncé, si la subordonnée circonstancielle intègre ou non la portée du marqueur identifié (ex. « les organisations de défense de la nature qui au{**raient**} le monopole/ .h: de la conscience écologique/ comme le parti socialiste français celui du cœur peut-[être\])

Les décisions ci-dessus peuvent avoir des conséquences plus ou moins importantes pour l'annotation. Ainsi, dans le cas suivant, le premier « c'est » fait partie de l'auto-réparation de la portée « c'est exceptionnel » et le marqueur « il paraît que » apparaît dès lors à l'intérieur de la portée.

(14) REU_AC1 / 00:23:01.135 - 00:23:04.741 [ID 40231]

1 CAR ah mais j'en ai entendu parler au[ssi]  
 2 BOR [c'est] il paraît que c'est  
 3 [exceptionnel\  
 3 CAR [oui c'est le truc hyper branché]

Finalement, dans le cas où la portée est discontinuée parce qu'une réaction d'un autre locuteur s'intercale (hors cas de chevauchement donc), l'élément intercalé est indiqué comme intégrant la portée et entouré de signes « ++ » (voir « [complém]entaires ouais\ » ci-dessous, ligne 3; ainsi que les « mh » des lignes 5 et 6).

(15) REU_BL1 / 00:02:41.611 - 00:02:55.758 [ID 40191]

1 ALE [°en° (.) palrallèle (.) on a aussi eu des demandes euh:: un peu  
 2 complémentai[res\  
 3 FRA [complém]entaires ouais\  
 4 ALE **évidemment** du client  
 5 ANT mh

⁵ Ces modalisations d'énonciation peuvent bien sûr être annotées séparément comme marqueurs ayant la même portée dans le cas où elles intègrent la liste des lemmes considérés dans le projet (« **visiblement** [marqueur1] il est arrivé **il me semble** [marqueur2] »).

6	FRA	mh[mh\]
7	ALE	[°qui°]: euh::: qui avait: au[ssi euh]:: (.) deux ou trois&
8	CHA	[°((rire))°]
9	ALE	&éléments\ typiquement: en termes de (0.5) de::: d'équipement: on

Mis à part les cas cités ci-dessus, ne doit être indiqué comme portée que ce qui constitue effectivement la portée même si l'association entre le marqueur indiqué en 2.1.1 et la portée donne un résultat qui paraît partiel ou tronqué (ex. « c'était je crois » ; ça me paraît » ; « vous avez l'air » ; « y a probablement vingt ou trente ans » ; « je sais pas qui/quoi/combien/où/pourquoi/quand », « pas unique bien évidemment »).

### 2.2.1.1 Identification de la portée dans ELAN

Dans ELAN, sur le *tier* POR, on crée une annotation couvrant le segment correspondant à la portée et on insère à l'intérieur de l'annotation le code d'identification unique de la portée, qui est identique à celui utilisé pour le marqueur correspondant.

Attention : dans le cas où le segment correspondant à la portée a déjà été utilisé (parce qu'il est concerné par un autre marqueur que celui considéré, notamment dans les cas d'une portée mobilisée par plus d'un marqueur), l'annotation correspondante doit être introduite dans un second *tier* POR (ne pas mettre deux identifiants dans la même annotation).

### 2.2.1.2 Identification de la portée dans Excel (col. AF)

Code identique à celui introduit pour le marqueur (étape 2.1.1.2 ci-dessus).

## 3. Annotation du marqueur

### 3.1 Informations relatives à l'événement

#### 3.1.1 événement (col. G)

L'événement correspond au code de l'événement parmi les 26 événements qui composent le corpus. En partant de l'import (trié) des résultats de la recherche faite dans ELAN (étapes 1.2.2 à 1.2.4 ci-dessus), il s'agit de remplacer le chemin de fichier par l'événement correspondant dans la liste déroulante.

Ex. [/.../POSEPI partagé/CORPUS/REUNIONS/REU_BM2/REU_BM2.eaf] → REU_BM2

#### 3.1.2 année (col. H)

L'*année* correspond à l'année d'enregistrement ou de diffusion de l'événement. L'information est automatiquement récupérée au moment de la saisie de l'événement (col. G).

Ex: 2007

#### 3.1.3 fichier vidéo (col. I)

Le *fichier vidéo* correspond au nom de la vidéo de référence de l'événement. L'information est automatiquement récupérée au moment de la saisie de l'événement (col. G).

Ex: DTV_SM1.mov

### 3.1.4 genre (col. J)

Le *genre* correspond au type d'événement (débat public, débat télévisé ou réunion professionnelle). L'information est automatiquement récupérée au moment de la saisie de l'événement (col. G).

Ex: débat public

## 3.2 Informations relatives au marqueur lui-même

### 3.2.1 Lien CLAPI (col. K)

Le *lien CLAPI* consiste en un lien URL pointant vers le marqueur dans la base de données CLAPI⁶, qui agrège, synchronise et met à disposition en ligne les données primaires du projet (fichiers vidéo et audio et transcriptions).

### 3.2.2 Timing début (col. L) et timing fin (col. M) du marqueur

Le *timing début* et le *timing fin* du marqueur renvoient aux coordonnées temporelles du marqueur au sein de l'événement. Il s'introduit en millisecondes à partir des informations stockées dans ELAN à l'étape 2.1.1.1 ci-dessus)⁷.

Ex. le marqueur ayant l'ID 42475 a comme timing début 1529435 et comme timing fin 1530060.

### 3.2.3 Réalisation du marqueur (col. N)

La *réalisation du marqueur* indique le marqueur tel qu'il a été effectivement et fidèlement transcrit dans ELAN. Il suffit dès lors de copier-coller la portion de transcription dans ELAN correspondant à la portion identifiée comme marqueur (col. A).

### 3.2.4 Lemme du marqueur (col. O)

Le *lemme* correspond à la forme générique ou l'entrée lexicale à laquelle le marqueur appartient. Il est choisi dans une liste déroulante alimentée à partir de l'inventaire des marqueurs établi à l'étape 1.1.

Ex. « il m'a dit » → DIRE

Dans le cas d'un marqueur complexe au sens discuté en 1.2.4, c'est donc le lemme-noyau qui est ici introduit (par ex. « évident » dans « ça me paraît évident » ou « penser » dans « on pourrait penser »).

---

⁶ CLAPI Corpus de LAngue Parlée en Interaction, <http://clapi.ish-lyon.cnrs.fr>, consulté le 1 septembre 2022.

⁷ Pour récupérer l'information, il suffit de faire un clic droit sur l'annotation correspondant au marqueur dans ELAN, puis « copy annotation » et la coller dans un document texte. « B=nnnnnn » correspond au timing début et « E=nnnnn » correspond au timing fin.

### 3.2.5 Forme/expression (col. P)

La *forme/expression* constitue un niveau intermédiaire d'abstraction entre le lemme d'une part (col. O) et la réalisation du marqueur de l'autre (col. N). Cela peut correspondre à des expressions telles que « je sais bien », « semble-t-il », etc. Si des propositions peuvent être faites au premier round d'annotation, il est très probable que cette colonne évolue au fil de l'examen du corpus.

### 3.2.6 POS-TAG (col. Q)

Le POS-TAG correspond à la catégorie morpho-syntaxique du lemme. Il est renseigné automatiquement à partir de la sélection du lemme (col. O).

Ex. DIRE → verbe

### 3.2.7 Type morpho-syntaxique (col. R)

Pour ce qui est des types morpho-syntaxiques, le projet se base sur les catégories suivantes, pour leur grande partie sélectionnées et traduites à partir des propositions du projet MODAL (Pietrandrea, 2022) :

- a. Morphème (« affixed element that attributes a truth value to the proposition encoded by the element to which it is affixed »). Il s'agit essentiellement du cas du conditionnel, lorsqu'il n'est pas associé à un lemme épistémique/évidentiel (voir 1.2.4. ci-dessus).

```
(16) DTV_REC / 00:56:42.099 - 00:56:51.993 [ID 45001]
1  NID      on est (.) vraiment au cœur d'un débat qui apparaît pour ce
2          qu'il est/ c'est-à-dire un débat idéologique\ qui a d'un côté
2          les organisations de défense de la nature qui auraient le
3          monopole/.h: de la conscience écologique/ comme le parti
4          socialiste français celui du cœur peut-[être\]
```

- b. Construction attributive à verbe copule. Il s'agit de constructions attributives reposant sur un verbe copule épistémique tel que « sembler », « paraître » ou « avoir l'air » (hors cas de constructions non pas attributives mais impersonnelles ; voir plus bas).

```
(17) DTV_SM2 / 00:14:37.871 - 00:14:42.729 [ID 20005]
1  PEL      .h MAIS LA QUESTION qui a été posée tout à l'heure par monsieur
2          meystre me paraît importante c'est celle du partenariat social\ oui
```

```
(18) REU_BL1 / 00:18:25.197 - 00:18:30.230 [ID 40001]
1  ANT      mh[mh\]
2  ALE      [et] puis [à l'intérieur]r&
3  CHA      [à l'intérieur\]
4  ALE      &ben [c'est]&
5  CHA      [°ouais°]
6  ALE      &no[ir\ le même qu` y a un peu]&
7  FRA      [ouais ça m` semble logique\]
8  ALE      &quand on rentre [dans la s]alle de bain °après c'[est no][ir\
9          donc° voilà]
```

```
(19) REU_AC1 / 00:10:25.681 - 00:10:33.481 [ID 20048]
1  CAR      est-ce que isaline participe euh: pou:hous:hous:
2  JEA      si [vous:: avez envie/]
3  CAR      [c'est c'est toi qui] mais c'est toi qui sais c'est:::: c'est:
```



4 ton truc hein/  
 5 (0.3)  
 6 BOR [bah ça] me paraît pas mal ouais  
 7 CAR [xx xx]

(20) REU_CL1 / 00:03:34.869 - 00:03:45.721 [ID 40061]

1 YVA ça fonctionne au niveau planning j'avais r'gardé avec euh::  
 2 SEB qui est mieux payé:/ là-bas:/ chez auster/ ou bien c'est juste une  
 3 question de::: [((inaudible))]  
 4 YVA [non: il il est] il est mieux payé i` m'a dit mais  
 5 i- il a surtout un statut d` CHEF de projet c` qui m` semble assez  
 6 incroyable\

(21) DPU_BIE / 00:47:13.743 - 00:47:27.530 [ID 20086]

1 P03 et je prends un exemple/ (0.6) aujourd'hui/ (0.4) dans le conflit  
 2 israélo-palestinien/ (.) la majorité des peuples/ (0.4) des deux  
 3 camps\ veut la paix\ (0.8) et cette paix semble impossible\ (0.6)  
 4 où est/ le bien commun/

Cette catégorie couvre aussi les cas d'attributs verbaux :

(22) DPU_MED / 00:34:41.345 - 00:34:47.652 [ID 40019]

1 SAN là:/ (.) on a l'air de dire il y a un pouvoir législatif  
 2 un pouvoir exécutif un pouvoir judiciaire .h et un pouvoir  
 3 médiatique\

(23) REU_AC2 / 00:23:47.831 - 00:23:57.731 [ID 20017]

1 BOR [i` doit i` doit bien rigoler le gars (0.7) et (.) u- il a l'air de  
 2 se faire des °putes° en or hein/ là-dessus (1.3) `fin d` gagner  
 3 beaucoup plus °euh° (sa vie) avec [°le chanvre légal que  
 4 illégal/°]

(24) REU_CL1 / 01:05:30.279 - 01:05:36.279 [ID 20019]

1 OLR i` peut jouer ce rôle-là si eddy veut partir/  
 2 (0.1)  
 3 SEB mh (0.4) °s'il veut°  
 4 YVA MAIS i` i` [xx]  
 5 OLR [eddy avait] l'air d` vouloir partir\ h

(25) DPU_UDC / 00:31:12.578 - 00:31:35.618 [ID 20003]

1 CH1 le troisième truc c'est que si vous créez l'événement par la  
 2 provocation: en en mettant des choses très très .h: puissantes sur heu  
 3 (.) l'opinion publique vous marchez aussi dans les médias ça y a y a  
 4 des phénomènes mécaniques de ce type-là\ (0.4) et et les médias ne  
 5 peuvent pas pas en parler: (.) ils sont pris dans: dans ce cycle-là\  
 6 ça me paraît être les (.) les tr- il y a le quatrième truc c'est  
 7 aussi: le climat international/ (.) on est pas dans un phénomène  
 8 suisse on est dans un phénomène européen/

Dans tous ces cas, la construction est attributive, « personnelle », en distinction des cas de constructions impersonnelles (« il semble (que) » ; « il paraît (que) »), qui sont traités comme « prédicat verbal à complément » (voir ci-dessous).

- c. Construction à verbe modal (« constructions characterized by an inflectionally constrained verbal predicate that takes an aspectually incomplete infinitive as a complement and that attributes a truth value to it »). Pour rappel (voir ci-dessus 1.2.4.), c'est à l'exception des cas où le verbe à l'infinitif sur lequel porte le modal est lui-même un lemme épistémique/évidentiel.

(26) DPU_REC / 00:51:06.161 - 00:51:50.799 [ID 41096]

1 NA2 y a (.) une partie de ça qui sert euh qui sont les recours y a  
2 une petite partie\ (0.5) euh donc je pense que ça se chiffre en  
3 quelques dizaines de millions de francs\ (0.5) au total\ (0.3)  
4 peut-être\ (0.7) °peut-être° cent millions de euh à à vue de nez  
5 ça doit être quelque chose comme ça\ ce qui est tout à fait  
6 supportable pour la suisse/ en tout cas je pense que

- d. Prédicat verbal à complément (« verbal predicates that take a sentential predicate as a complement and that attribute a truth value to it »). Cela comprend aussi les cas de constructions impersonnelles mobilisant « sembler/paraître » (ex. « il semble que ce soit le cas »), qui sont à distinguer des constructions attributives à verbe copule (voir ci-dessus). Attention : le « que » de subordination peut être absent (situation de complétives asyndétiques) (Ex. « je pense il est arrivé »), cas qui n'est donc pas assimilé aux parenthétiques (voir ci-dessous), qui ne peuvent quant à eux intervenir qu'au milieu ou en fin de proposition. Finalement, cela comprend également les cas de discours rapportés directs (Ex. « on dit oh oui on veut garantir par le financement public une certaine indépendance « recherche » »).

(27) DPU_MED / 00:20:32.394 - 00:20:41.534 [ID 23701]

1 SAN alors on pourrait penser qu' une femme est aussi un être normal:/  
2 et qu'après tout euh si elle correspond:/ à ce que certains  
3 souhaitent et bien il est absolument légitime qu'elle soit élue:\

(28) DTV_VEH / 00:51:41.386 - 00:51:57.248 (lemme = évident) [ID 41155]

1 PET le véhicule a plus de chances de se retourner et heu: (.) plus  
2 de risques de décès pour ses occupants.h et il a plus de  
3 risques pour les piétons ou: en cas de collision ou les vélos  
4 ou les motards.h donc juste pour cette raison ça nous paraît  
5 évident que ce type de véhicules pas tous (.) mais ceux qui  
6 sont dangereux: ou trop dangereux doivent être retirés de la  
7 vente/

(29) DPU_PRE / 00:24:36.995 - 00:24:44.370 [ID 22068]

1 NA1 heu j` crois (0.5) mon s- mon (0.5) ma proposition a pas  
2 été très bien comprise\ (.) on est justement pour une  
3 plus forte démocratisation des études\ .h

(30) DTV_VEH / 00:04:43.909 - 00:04:54.473 [ID 41401]

1 PON [alors] il n'em- pas bush pour prendre cet exemple-là n-  
2 à supposer qu'il soit capable de penser quelque chose (.)  
3 n'en pense pas le premier mot il a perdu les élections: et i`  
4 faut qu'il cocole un peu sa majorité démo[cratique ça n'a  
5 rien à voir]

(31) REU_BM3 / 00:25:02.759 - 00:25:14.922 [ID 20049]

1 CHR alors faut noter sur les plans\ (0.4) sur les plans parce que euh::  
2 (1.6)

3 SAL i g p à °euh:° on contrôle mais i` me semble avoir écrit: n c s  
 4 ou i g p à choix\  
 5 CHR mh[mh\  
 6 SAL °[de l'ar]chitecte°

(32) REU_BM3 / 00:43:53.490 - 00:44:07.500 [ID 20050]

1 CHR par contre on n'a pa- pour la façade on n'a pas discuté du  
 2 calepinage\  
 3 SAL ah ouais\  
 4 (0.7)  
 5 CHR donc le calepina:ge/ (0.4) toi t` as r`gardé c` qui propo:s[e]  
 6 SAL [.tsk]  
 7 .h ALORS j'ai regardé\ j'ai PAS encore rapproché °d` jusqu'en::  
 8 ° haut mais il me semble c'est TRÈS TRÈS proche à ce que: ça c'est  
 9 ce que on avait proposé nous

- e. Syntagme adverbial (« a syntactically non obligatory predicates that syntactically depend on another predicate (whether a verb, an adjective, an adverb or an entire clause) and that attribute a truth value to it »)

(33) DTV_VEH / 00:51:57.987 - 00:52:07.351 [ID 40229]

1 PET .h (.) c'est surtout: pour ces: ces raisons-là que les  
 2 familles comme vous dites achètent ce type de véhicules/  
 3 mais je pense [que ces familles sont mal informées] sur la&  
 4 PON [pour écraser les piétons:]  
 5 PET &sécurité de ces véhicules\ (.) visiblement\  
 6

(34) REU_AC2 / 00:48:24.000 - 00:48:34.837 [ID 26198]

1 BOR .tsk et les SEULS qui ont réussi à rebondir un peu c'est la radio\  
 2 et encore ils rebondissent pas bien haut hein/=  
 3 ROG =ouais  
 4 (1.0)  
 5 BOR et peut-être que dans un an ils s`ront plus là du tout\ °donc  
 6 euh c'est un peu:° (1.4)°un peu gli:ssant\  
 7

- f. Parenthétique (« predicates syntactically independent of the scope to which they attribute a truth value »). Intitulé « marqueur discursif » dans le projet MODAL, nous préférons utiliser la terminologie de « parenthétique » ici. Attention : un marqueur qui admettrait une subordination avec un « que » avant sa portée (ex. « je pense (que) P ») est traité comme un prédicat verbal à complément (avec absence du « que ») et non comme un parenthétique (Thompson & Mulac, 1991). Un parenthétique peut être produit par un autre locuteur que celui de la portée, notamment dans le cas d'une réponse à une question (voir troisième exemple ci-dessous).

(35) DPU_PRE / 00:39:50.525 - 00:40:13.936 [ID 40064]

1 P16 enfin au niveau de leurs comptes en fait le fait qu'ils font des  
 2 sorties en boîte enfin rien qu` ça [j` veux dire]&  
 3 NA1 [mhm/]  
 4 P16 &ça fait que y a un retour au niveau de la: dynamique dynam-  
 5 dynamique économique de la région\ (.) donc heu::: (0.3) bon:::  
 6 ils font pas que heu ponctionner de l'argent à l'état il me  
 7 semble/ h::

(36) DPU_UDC / 00:16:24.525 - 00:16:44.162 [ID 40014]  
 1 CH1 et puis un journal qui dans ses éditos\ à part quelques (.) tirs  
 2 d'indignation mais assez rares\ était (0.4) comme l'est toujours le  
 3 temps relativement heu nuancé\ (0.4) .h (0.5) et puis alors **me**  
 4 **semble-t-il** après heu on a vu plutôt euh des journaux assez  
 5 **tranquilles assez assez complaisants assez neutres/**

(37) REU_AM1 / 00:29:27.535 - 00:29:29.418 [ID 24188]  
 1 CAR **il lui montrait quand** (.) t`avais dit/  
 2 BOR **je sais pas**

- g. Proposition complète (« syntactically and semantically complete utterances that attribute a truth value to a scope. The reference to the scope is made through a co-reference relation that links an element of the utterance and (an element of) the scope. »).

(38) DPU_UDC / 00:32:55.873 - 00:33:05.943 [ID 41143 ; 28474]  
 1 CH1 alors oui/ on va c'est une tendance heu systémique pour mille  
 2 raisons: (0.3) heu: **y a une tendance de personnalisation: de mise**  
 3 **en spectacle de la politique** **ça c'est évident** (.) **ça ça me paraît**  
 4 **tout à fait clai:r**

(39) REU_CL4 / 01:00:39.871 - 01:00:40.314 [ID 26370]  
 1 PAU .h **[c'est] peut-être moi qui signe la lettre pour euh t a r j'en sais**  
 2 **RIEN** mais [faut que cje soit personnali[té/] .h personnalisé:/ .h

(40) DPU_ETU / 01:44:06.250 - 01:44:11.900 [ID 21212]  
 1 BAC le premier/ éventuellement que **l'université peut un petit peu vous**  
 2 **préparez pou:r** mais: **j'en doute un petit peu**

(41) REU_AC2 / 00:39:47.822 - 00:40:02.271 [ID 21143]  
 1 ROG ouais/ donc pour moi l'ÉCHEC c'est d'avoir (trainer; trader) aussi  
 2 long:temps: dans un [truc aussi:: (.) ARCHAïque\]  
 3 BOR [un truc aussi ARCHAÏQUE que ça]  
 4 (0.3)  
 5 °ça c'est° p[as faux]  
 6 CAR [mh ouais]  
 7 (0.2)  
 8 ROG **c'est ça l'écheC**  
 9 (0.7)  
 10 **vous v- v- v- v- vous construisez comme y A CENT ANS/**  
 11 CAR oua[is]  
 12 ANN [m]hm[h]  
 13 BOR [a]h on construit [là complètement]  
 14 ANN **[ça c'est sûr\]**  
 15 (0.2)

### 3.3 Informations relatives au contexte interactionnel

Les *informations relatives au contexte interactionnel* permettent de documenter la position séquentielle du marqueur, relativement à l'unité de construction du tour (UCT) au sein de laquelle le marqueur émerge (3.3.1) et relativement au tour de parole (TDP) comprenant l'UCT en question (3.3.2), ainsi que la fonction interactionnelle prise par celle-ci (3.3.3)

La combinaison de ces trois informations permet ainsi de documenter aussi bien le cas d'un « peut-être » apparaissant seul comme réponse à une question (totalité de l'UCT, UCT correspondant à la totalité du TDP, fonction réactive de l'UCT) que d'un « peut-être » apparaissant au milieu d'une UCT émergeant au milieu d'un long tour multi-UCT (et ayant possiblement une fonction interactionnelle indéterminée).

### 3.3.1 Position du marqueur dans l'UCT (col. S)

La *position du marqueur dans l'unité de construction de tour (UCT)* implique de regarder le rapport temporel entre le marqueur et l'UCT au sein de laquelle le marqueur émerge. L'UCT consiste en une suite linguistique manifestant une complétude au plan syntaxique et prosodique et accomplissant, au niveau pragmatique, une action *accountable* (Clayman, 2013; Schegloff, 2007). En pratique, la segmentation des tours de parole (TDP) en UCT, et en particulier la segmentation des longs tours de parole multi-UCT, est une problématique relativement complexe et qui fait débat dans la littérature (pour une discussion, voir Jacquin, 2014, p. 79-92). Pour ce projet, on admet qu'il y a UCT dès qu'il y a complétude syntaxique et intonative et que la suite linguistique ainsi découpée est interprétable pragmatiquement. Il peut arriver que différentes ressources viennent s'ajouter pour projeter une complétude du tour au-delà de l'UCT en cours et ainsi projeter la construction d'un long tour multi-UCT (Selting, 2000, p. 504-511), mais cela n'implique pour autant de supprimer la frontière initiale entre UCT.

Ainsi, dans l'exemple qui suit, la pertinence pour une transition de locuteurs entre « il a perdu les élections » et « et il faut qu'il cocole... » est certes réduite (notamment par l'allongement prosodique de « élections »), mais il y a bien à cet endroit une rupture entre deux composants complets du TDP de PON, ce qui n'est pas le cas entre le premier « alors » et la suite, ou entre « quelque chose » et « n'en pense pas... » ou encore entre « le premier mot » et « il a perdu ». Aucun de ces points intermédiaires ne manifeste une complétude syntaxique et prosodique accomplissant une action langagière identifiable.

NB : ci-dessous, les couleurs (vert et bleu) renvoient à la segmentation des TDP en UCT ; elles n'ont aucun rapport avec l'annotation de marqueurs épistémiques/évidentiels.

(42) DTV_VEH / 00:04:43.299 - 00:04:57.015

1 ZEN pensent comme [ça]  
 2 PON [alors] il n'em- pas bush pour prendre cet exemple-  
 3 là n- à supposer qu'il soit capable de penser quelque chose (.)  
 4 n'en pense pas le premier mot il a perdu les élections: (.) et i  
 5 faut qu'il cocole un peu sa majorité démo[cratique ça n'a rien à  
 6 voir]  
 7 ZEN [alors y a une communauté  
 8 scienti]fique y a monsieur blair [y a monsieur chirac et y a]

(43) DTV_VEH / 00:04:43.299 - 00:04:57.015

1 BOR j- je saute du coq à l'âne\ j'ai j'ai un copain qui était à new  
 2 york et qui est allé dans cette i- dans cette euh ce (1.0) euh::  
 3 (.) ça s'appelle sleep no more (0.3) et en fait on on entre dans  
 4 un::: (0.3) y a des énigmes à rés- c'est c'est une immense  
 5 maison: (.) .h: et puis on entre pis on on fait soi-même son  
 6 chemin:  
 7 CAR ah mais j'en ai entendu parler au[ssi]

Les petits mots et autres marqueurs discursifs (tels que « alors », « ouais », etc.) peuvent être traités comme des UCT à part entière lorsque la réalisation prosodique et l'environnement pragmatique vont dans ce sens. Ainsi, on distinguera le cas de « ouais ça me semble logique » ci-dessous (ligne 6), produit d'un seul tenant et consistant en une seule UCT, du cas où la locutrice aurait énoncé « ouais\ (.) il me semble »).

(44) REU_BL1 / 00:18:20.936 - 00:18:30.522

1 ALE [et] puis [à l'intérieur]&  
2 CHA [à l'intérieur\  
3 ALE &ben [c'est]&  
4 CHA [°ouais°]  
5 ALE &no[ir\ le même qu` y a un peu]&  
6 FRA [ouais ça m` semble logique\  
7 ALE &quand on rentre [dans la s]&  
8 FRA [mhmh\]

La segmentation des TDP en UCT venant d'être précisée, 6 positions de marqueurs au sein de l'UCT sont distinguées :

a. Début : le marqueur est au tout début de l'UCT.

(45) REU_CL4 / 00:21:18.523 - 00:21:26.067 [ID 40221]

1 YVA [remis en question/]  
2 OLR [on on sait pas c` qu'i` s'est pa[ssé depuis/]  
3 WIL [<((souffle)) H::>]  
4 (1.0)  
5 WIL [°(c` n'est que) euh euh rencontre°]  
6 SEB [mais c'était deux y a deux ou] trois ans déjà ça non/  
7 WIL °ouais ouais°  
8 OLR **apparemment** **il s` passe pas grand-CHOSE/** °donc°  
9 WIL non\  
non\  
non\

(46) DTV_DEM / 00:13:24.562 - 00:14:24.003 [ID 40221]

1 LOR [non attendez revenons au complot dont on parlait]  
2 BEN on est en train de parler: .h de ce mot: .h qui  
3 fait irruption dans la politique suisse et moi je dis de  
4 ne pas être nAif\ .h **la gauche aujourd'hui dit**  
5 **que** .h **mertz et couchepin complotent contre les**  
6 **assurances sociales/** .h veulent ca:sser la sécurité  
7 **sociale en-** (.) aujourd'hui (.) c'est le thème que  
8 vous avez xx aujourd'hui alors vous êtes mal placé  
9 j'entends pour faire ce genre de chose .h la deuxième  
10 [chose (.) la deu:xi]ème&

b. Début* : le marqueur est au début de l'UCT mais précédé de 1 maximum 2 mot(s) (par ex. « heu », « mhm », « ben », « mais »).

(47) DPU_PRE / 00:24:36.995 - 00:24:44.370 [ID 22068]

1 NA1 heu **j` crois** (0.5) **mon s- mon (0.5) ma proposition a pas**  
2 **été très bien comprise** (.) on est justement pour une  
3 plus forte démocratisation des études\ .h

(48) DPU_ETU / 01:32:13.260 - 01:32:21.550 [ID 23071]

1 TOP et heu j` pense que on on de- enfin on peut pas obliger les gens à  
2 choisir des facultés par rapport aux débouchés parce que si on a  
3 pas envie d'étudier la matière qu'on étudie/

(49) DTV_SM1 / 00:18:03.194 - 00:18:04.271 [ID 22151]

1 LEV .h ouais mais j` crois que c'est faux\

(50) REU_CL4 / 00:02:28.744 - 00:02:49.688 [ID 20034]

1 OLR alors moi j` préfère être honnête ça m'est égal mais j'ai  
2 l'important j- j` pense de regrouper les les les données on  
3 pourrait l` mettre là d'dans/ [en f- e]n annexe ou à la fin de&  
4 WIL [oui/]  
5 OLR &ce: document mais (0.7) j` pense faudrait& l` faire et ÇA pour moi  
6 c'était un point c'est-à-dire que: (0.8) ces concurrents  
7 (1.4)  
8 OLR j` sais pas c` que tu en penses/ mais moi j'ai l'impression de MA:L  
9 connaître euh:: °enfin c'est pas une impression en plus° ((rire))  
10 <((en riant)) mal connaître le milieu>

### c. Milieu : le marqueur est au milieu de l'UCT

(51) DTV_DEM / 00:04:15.483 - 00:04:43.617 [ID 20079; 40238]

1 MAI j` trouve qu'on a pas l` droit d` dire que: les politiciens:  
2 sont tous pourris: sont tous: montrent leurs vrais visages .h (.)  
3 mais i` faut simplement essayer d` comprendre ce qui s` passe\ .h  
4 c` qui se passe c'est qu` y a un parti: qui semble manifestement  
5 aux abois\ (. ) parce qu` y a un certain nombre de choses qui sont  
6 dites dans un RAPport Officiel: (. ) de la commission d`  
7 ge[stion (. ) qui]&  
8 LOR [on en parlera un peu plus loin]  
9 MAI &qui manifestement: leur faire très peur:

(52) DTV_REC / 00:43:22.595 - 00:43:33.789 [ID 40178]

1 LAT heu attendez je veux dire (0.7) on peut évidemment\ (0.4) euh  
2 s'asseoir à une table (0.2) et discuter/ (. ) mais: (. ) la suite  
3 c'est quoi/ la suite c'est: (0.4) on veut: (. ) un plan  
4 d'aménagement global (0.4) de nendaz/

### d. Fin* : le marqueur est en fin d'UCT mais suivi de 1 maximum 2 mot(s).

(53) DTV_REC / 00:10:04.364 - 00:10:28.425 [ID 41160]

1 NA3 .h ce qui nous gêne/ (. ) c'est que souvent ces associations  
2 (. ) jouent sur les deux tableaux\ (. ) et c'est le cas pour  
3 le musée cantonal des beaux-arts/ puisque franz weber:  
4 a lancé le référendum/ (. ) dem[ande au peuple vaudois de  
5 voter/ .h et]  
6 LOR [heu peuh peuh peuh je suis pas  
7 si sûr] pour franz weber (. ) j'ai pas [lu ça hein/]  
8 NA3 [il est initi]ant il est  
9 ns les [initi: dans les dans les référendaires/]

### e. Fin : le marqueur est en toute fin d'UCT

(54) DTV_SM2 / 00:59:54.743 - 01:00:10.659 [ID 42315]

1 MA2 [là vous êtes euh:::]  
 2 CNI [non la la raison numéro un] pour la pau[vreté c'est]&  
 3 MA2 [((petit rire))]  
 4 CNI t- être euh chômage d'avoir pas d'emploi ou pas  
 5 suffisamment d'emploi  
 6 MA2 pas du tout c'est être famille monoparentale c'est être  
 7 divorcé c'est avoir un bas salaire c'est le rapport deux  
 8 mille [dix sorti par monsieur burkhalter qui le dit]  
 9 BER [.h: et d'être une femme non/ et d'être une femme]  
 10 (0.3) euh ça [aide\ (.) à être euh à être pauvre [si on peut  
 11 s'exp]rimer ain]si

#### f. Totalité : le marqueur consiste en la totalité de l'UCT

(55) REU_AM1 / 00:45:08.733 - 00:45:12.435 [ID 23300]  
 1 BOR [ouI:: ben C'EST UNE CARTE DE VOEUX DE NOËL] C'EST  
 2 [UNE CARTE DE JOYEU]SES [FÊ:TES]  
 3 ROG [ça::] ça i- il m` la balance à travers la  
 4 figure [c'est sûr\  
 5 BOR [j` pense pas\  
 6 ROG c'est sûr\  
 7

### 3.3.2 Position de l'UCT dans le TDP (col. T)

La position de l'UCT dans le tour de parole (TDP) implique de regarder le rapport temporel entre l'UCT au sein de laquelle le marqueur émerge et le TDP contenant l'UCT. La segmentation de la parole en TDP peut être plus ou moins problématique, notamment dans le cas de certains silences ou certains chevauchements, et compte tenu des différences de contextes interactionnels et donc de dispositifs de *turn-taking* documentés par le projet. On prendra ici le TDP comme une occasion de parler pour un locuteur, située entre deux autres occasions prises en charge par un ou deux autre(s) locuteur(s). Ainsi,

- si le locuteur continue de parler après un silence, quel qu'en soit sa durée, on considérera ici qu'il continue son TDP.

(56) REU_AC1 / 00:22:39.500 - 00:23:28.372  
 1 BOR j- je saute du coq à l'âne j'ai j'ai un copain qu` était à new  
 2 york et qui est allé dans cette i- dans cette heu ce (1.0)  
 3 heu:: (.) ça s'appelle sleep no more (0.3) et en fait on on  
 4 entre dans un::: (0.3) y a des énigmes à rés- c'est une  
 5 immense maison: (.) .h: et puis on entre pis on on fait  
 6 soi-même son chemin:(.) mais tous les gens qui sont là: ont  
 7 un rôle sont des acteurs  
 8 (0.2)  
 9 CAR ah mais j'en ai entendu parler au[ssi]

- si le locuteur continue de parler après un continuateur tel que « mhm » l'encourageant à poursuivre, on considérera qu'il continue son TDP.

(57) DPU_PRE / 00:39:50.525 - 00:40:13.936  
 1 P16 enfin au niveau de leurs comptes en fait le fait qu'ils font des  
 2 sorties en boîte enfin rien qu` ça [j` veux dire]&  
 3 NA1 [mhm/]  
 4 P16 &ça fait que y a un retour au niveau de la: dynamique dynam-



5           dynamique économique de la région\ (.) donc heu::: (0.3) bon:::  
6           ils font pas que heu ponctionner de l'argent à l'état il me  
7           semble/ h::

- si le locuteur est **chevauché** par un autre participant mais qu'il maintient la trajectoire syntaxique/prosodique/pragmatique de son tour, on considérera qu'il continue son tour de parole.

(58) DTV_VEH / 00:51:57.987 - 00:52:07.351

1 PET       .h (.) c'est surtout: pour ces: ces raisons-là que les  
2           familles comme vous dites achètent ce type de véhicules/  
3           mais je pense [que ces familles sont mal informées] sur la&  
4 PON           **[pour écraser les piétons:]**  
5 PET       &sécurité de ces véhicules\ (.) visiblement\

En cas de doute sur la segmentation en TDP, l'annotateur consigne le cas, qui est discuté au sein de l'équipe.

4 possibilités de relation entre le TDP et l'UCT où intervient le marqueur sont distinguées :

a. Début : l'UCT est la première d'un TDP multi-UCT.

(59) DPU_MED / 00:51:00.201 - 00:51:27.424 [ID 22061]

1 P11       à fo:rce (0.5) de donner l'information/ (.) on provoque un  
2           sentiment (.) qui est totalement dirigé chez le (0.7) chez le  
3           lecteur\ (.) et la simple transmission de faits (.) permet  
4           d'influencer sur les idées je pense que faut pas: (.) scinder les  
5           deux\  
6           (0.4)  
7 SAN       **ça je crois que** **vous avez tout à fait raison** MAIS on se trouve  
8           devant le même problème aussi sur le plan (0.5) je dirais d'un  
9           enseignement scientifique/ on cherchera à avoir l'air très  
10          objectif/ (.) mais on est quand même plutôt persuadé en général  
11          par les éléments qu'on a soi-même trouvés: que par celles qui sont  
12          parvenues à l'esprit d'autrui/ et par conséquent/

b. Milieu : l'UCT est au milieu d'un TDP multi-UCT.

(60) DTV_DEM 00:18:27.631 - 00:18:35.889 [ID 21202]

1 LOR       °j` vous ai posé une° question euh on essaie de faire notre boulot  
2           honnêtement **sans doute que** **vous aussi/** est-ce que pour vous il y a  
3           oui ou non (0.5) complot/ ou est-ce que finalement on exagère/

c. Fin : l'UCT est à la fin d'un TDP multi-UCT.

(61) DPU_ETU 00:55:57.110 - 00:56:30.170 [ID 20208]

1 SDZ       est-ce qu'il faut vraiment/ (.) que tout le monde/ (.) aille à  
2           l'uni\ (1.0) est-ce qu'il faut vraiment continuer/ (0.9) à mettre  
3           cette pression/ (0.5) sur les: sur les étudiants heu::: (0.8) j'ai  
4           des enfants qui sont encore trop petits pour y chi- pour euh mais  
5           ils vont arriver en cinquième et je pense que ça va commencer/ (.)  
6           on va commencer à s'en soucier\ est-ce que/ est-ce qu'il y a PAS/  
7           un un moment (.) une nécessité de se poser la question de resserrer  
8           le (0.5) le processus de sélection/ à l'université qui n'a jamais  
9           été aussi populaire qu'elle ne l'est (0.6) actuellement (.) heu en

10 suisse\ (.) pas forcément (.) accessible (.) mais en tout cas::

#### d. Totalité : l'UCT et le TDP sont co-extensifs (TDP à une seule UCT)

(62) DPU_ETU / 00:44:20.846 - 00:45:26.527 [ID 41355; 41356]  
1 DUM [euh moi] j` suis pas tout à fait d'accord qu` ça n'existe pas  
2 plus près heu en université par exemple heu saint-gall est f- est  
3 financé en partie par de:s par des (.) entreprises privées bon ça  
4 reste (0.7) le domaine économique mais j` pense que c'est pa::s et  
5 et contrairement à ce qu'on dit y a pas une influence directe de  
6 l'entreprise sur euh .h [sur les]&  
7 SAN [si si]  
8 DUM &c` qui est enseigné dans dans dans les cours\ NON j` suis  
9 [désolée\  
10 SAN [C'EST RECONNU c'est] c'est prouvé\  
[si si]

### 3.3.3 Rôle de l'UCT dans la séquence (col. U)

Le rôle de l'UCT dans la séquence renvoie à la fonction interactionnelle prise par l'UCT au sein de laquelle le marqueur apparaît.

On ne distingue ici que trois grandes possibilités :

- a. (clairement) initiatif : l'UCT relève clairement d'une première partie de paire adjacente (de type question-réponse, ordre-obtempération, statement-confirmation, invitation-acceptation, etc).

(63) REU_BL2 / 00:45:59.882 - 00:46:07.832 [ID 24265]  
1 SOP nous on voulait se voir:/ j` sais pas si t` as eu : l` temps de:  
2 regarder une fois sur ces trois parcelles là de:: orlatti:/  
3 FLO euh::  
4 FLO j` crois on avait parlé d` deMAIN

- b. (clairement) réactif-évaluatif : l'UCT relève clairement, au niveau de l'interaction, d'une seconde partie de paire adjacente (de type question-réponse, ordre-obtempération, statement-confirmation, invitation-acceptation, etc) ou d'une évaluation du TDP précédent. Attention : sont exclus les cas où le locuteur se positionne par rapport à des propos qu'il rapporte.

(64) REU_AC2 / 00:47:19.355 - 00:47:25.211 [ID 22205]  
1 ANN et pis uve[o\ ils ont] pas fait un flop au début/  
2 ROG [et pis] (0.6) .h  
3 (0.3)  
4 ANN °avant de° fonctionner/  
5 (1.5)  
6 CAR j` crois pas\  
[et pis]

- c. Indéterminé : aucun des deux rôles précédents ne peut clairement être attribué à l'UCT, ou l'UCT fait un peu des deux. Le cas échéant, c'est l'analyse qualitative de collections qui pourra affiner la caractérisation du rôle séquentiel.

(65) DPU_ETU / 00:55:52.645 - 00:56:35.816 [ID 20208]  
1 SDZ on va chan[ger d` thème\  
2 BAC [(j` vous l'offre)]

3 (0.3)  
4 SDZ [se poser] la question:/ de l'accès\  
5 BON [mhm]  
6 (1.0)  
7 SDZ est-ce qu'il faut vraiment/ (.) que tout le monde/ (.) aille à l'uni\  
8 (0.2) est-ce qu'il faut vraiment continuer/ (.) à mettre cette  
9 pression/ (0.5) sur les: sur les étudiants heu:: (0.8) j'ai  
10 des enfants qui sont encore trop petits pour y chi- pour euh mais ils  
11 vont arriver en cinquième et je pense que ça va commencer/ (.) on va  
12 commencer à s'en soucier\ est-ce que/ est-ce qu'il y a PAS/ un un  
13 moment (.) une nécessité de se poser la question de resserrer le (0.5)  
14 le processus de sélection/ à l'université qui n'a jamais été aussi  
15 populaire qu'elle ne l'est (0.6) actuellement (.) heu en suisse\  
16 (0.3) pas forcément (.) accessible (.) mais en tout cas:: (1.5)  
17 MA1 ben si vous permettez que je commence/ parce que là je suis (.) prête  
18 à entrer dans l'arène je [veux dire]

### 3.4 Informations relatives au locuteur du marqueur

#### 3.4.1 Locuteur du marqueur (col. V)

Le *locuteur* renvoie au participant ayant prononcé le marqueur. L'information a été introduite par le copier-coller effectué depuis les résultats de la recherche effectuée dans ELAN (étape 1.2.4)

Ex. ROG

#### 3.4.2 Place interactionnelle du locuteur du marqueur (col. W)

La *place interactionnelle du locuteur du marqueur* renvoie au rôle / à la fonction du locuteur du marqueur au sein de l'événement (voir Jacquin, 2014, p. 294-308). L'information est automatiquement introduite lors de la saisie du locuteur (col. V)

Ex. BOR → participant à la réunion

Ex. LOR → modérateur

#### 3.4.3 Statut socio-professionnel du locuteur du marqueur (col. X)

Le *statut socio-professionnel du locuteur du marqueur* renvoie aux caractéristiques identitaires du locuteur du marqueur qui ont justifié sa participation à l'événement à la place interactionnelle mentionnée (voir Jacquin, 2014, p. 283-294). L'information est automatiquement introduite lors de la saisie du locuteur (col. V)

Ex. BOR → directeur

Ex. LOR → journaliste à la RTS

## 3.5 Informations relatives à l’/aux allocutaire(s) du marqueur

### 3.5.1 Allocutaire du marqueur (col. Y)

Par allocutaire, on entend la personne à laquelle ou les personnes auxquelles le locuteur adresse son énoncé. Cet adressage s’évalue en premier lieu au niveau énonciatif, c’est-à-dire au niveau du formatage linguistique et multimodal de l’énoncé, indépendamment du comportement interactionnel des interlocuteurs.

Lorsqu’une seule personne physique est adressée comme allocutaire spécifique par la présence d’une deixis personnelle et/ou par une direction du regard franche (col. AC), l’annotateur le choisit dans la liste (par ex. DUM).

Lorsqu’un sous-groupe est clairement adressé par la présence d’une deixis personnelle et/ou par une direction du regard franche (col. AC), l’annotateur le choisit dans la liste (par ex. sous_groupe_DTV_VEH). Il peut s’agir d’un sous-groupe de deux invités externes à une réunion professionnelle (ANT et CHA dans REU_BL1), du sous-groupe constitué par les débattants invités d’un débat, etc. Le public sur place d’un débat public a par contre un identifiant particulier (par ex. 01 vaut pour le public du débat DPU_BIE), tout comme, les téléspectateurs d’un débat télévisé, adressés via le regard-caméra (par ex. CT5 vaut pour la caméra-téléspectateurs de DTV_SM1).

Lorsque le locuteur s’adresse à toutes les personnes présentes sans distinction et sans précision, ou ne semble pas avoir d’allocutaire déterminé, l’annotateur choisit l’option « tous » correspondant à l’événement (Ex. « tous REU_AM1 »).

Ainsi, la présence d’une deixis personnelle et/ou d’une direction du regard franche sont de puissants critères pour définir qui est l’allocutaire du marqueur en question. Cependant, lorsqu’il n’y a pas de deixis personnelle et que le regard est indéterminable/non clairement orienté vers un type d’allocutaire, la présence ou l’absence d’un geste de pointage peuvent être déterminantes. La combinaison de ces trois critères permet différentes possibilités représentées dans le tableau ci-dessous.

Tableau 1 : Les différents cas de sélection du ou des allocutaire(s)

Cas	Deixis verbale	Regard	Pointage	Décision
1	oui	oui	oui	une personne spécifique, un sous-groupe, ou tous est adressé comme allocutaire
2	oui	oui	non	une personne spécifique, un sous-groupe, ou tous est adressé comme allocutaire
3	oui	non	oui	une personne spécifique, un sous-groupe, ou tous est adressé comme allocutaire
4	oui	non	non	une personne spécifique, un sous-groupe, ou tous est adressé comme allocutaire
5	non	oui	oui	cette possibilité s’évalue contextuellement, sur la base des indices disponibles.
6	non	oui	non	cette possibilité s’évalue contextuellement, sur la base des indices disponibles.

7	non	non	oui	aucun allocataire clairement adressé (que ce soit une personne, un sous-groupe, ou tous) ; le pointage seul ne suffit pas => tous
8	non	non	non	aucun allocataire clairement adressé (que ce soit une personne, un sous-groupe, ou tous) => tous

De manière à exemplifier la complexité dont peut relever les cas 5 et 6, considérons l'exemple suivant :

```
(66) REU_BM1 / 01:02:27.459 - 01:02:58.261 [ID 40004]
1 MIC      alors moi j` trouve bizarre comment t- t- c'est c'est bien/
2          tu fais (d`) l'exercice\ (0.5) mais j` trouve bizarre dans: dans::
3          °mm° dans le DEgré: de pri- de priorité: .h: j- j`dirais moi
4          déjà: la première chose c'est que: on a dû:: mettre du LOGEMENT
5          °là dedans°\ (0.6) la première contrainte qu'on a:: qui à mon avis
6          est la plus complexe (0.5) c'est l` logement:\
7 DAP      °mm-mm°
8          (0.4)
9 MIC      il m` semble\
10         (0.6)
11 LOU     oui: et non parce que dans les DEUX cas le logement il::: fon-
12         il fonctionnerait la même chose\
13         (.)
14 MIC     OUI/ mais [mais] mais::: mais:: (.) pour moi y a quand même
15 LOU     [fin:::]
```

En 9, MIC conclut son développement avec « il me semble », qui porte sur la dernière partie de son développement, en bleu. Ce marqueur est accompagné ici d'un clair regard en direction de LOU, qui prend la parole en 11 en prenant ses distances par rapport à la demande de confirmation produite par MIC. Si LOU a été désigné personnellement par de la deixis verbale (« tu » en 2), ce n'est plus le cas dans la suite, où MIC produit par ailleurs différents balayages visuels, notamment sur DAP, qui confirme ainsi en 7. Il y a donc une tension ici entre le fait de considérer que la demande confirmation hétéro-sélectionne LOU comme allocataire, ou plutôt vise plus généralement à obtenir l'assise non seulement de LOU mais d'autres participants à l'interaction. Dans ce cas, l'annotateur fait un choix fondé sur une pondération des indices.

### 3.5.2 Place interactionnelle de l'allocataire du marqueur (col. Z)

La *place interactionnelle de l'allocataire du marqueur* renvoie au rôle / à la fonction de l'allocataire du marqueur au sein de l'événement (voir *supra* sur la place interactionnelle du locuteur). L'information est automatiquement introduite lors de la saisie de l'allocataire (col. Y).

### 3.5.3 Statut socio-professionnel de l'allocataire du marqueur (col. AA)

Le *statut socio-professionnel de l'allocataire du marqueur* renvoie aux caractéristiques identitaires de l'allocataire du marqueur qui ont justifié sa participation à l'événement à la place interactionnelle mentionnée (voir *supra* sur le statut du locuteur). L'information est automatiquement introduite lors de la saisie de l'allocataire (col. Y).

## 3.6. Informations relatives à la mimo-gestualité

### 3.6.1 Direction du regard (relatif au marqueur)

La *direction du regard (relatif au marqueur)* implique, sur la base des perspectives documentées par les différentes caméras, d'établir, avant (3.6.1.1), pendant (3.6.1.2) et après (3.6.1.3) le marqueur, ce que regarde le locuteur. Les choix ci-dessous constituent une simplification du modèle présenté par Rossano (2013), dans la mesure où certaines distinctions proposées par l'auteur sont trop fines ou trop soumises à l'interprétation au vu des données considérées ici.

5 possibilités sont distinguées :

- a. regard sur l'allocutaire ou les allocutaires → allocutaire(s). Il s'agit de l'allocutaire identifié dans la colonne Y.

L'exemple ci-dessous représente un cas simple où le regard est clairement orienté vers l'allocutaire, identifié ici par la deixis verbale.

```
(67) REU_AC1 / 00:35:25.015 - 00:35:27.148 [ID 20168]
1 JEA <((à isaline)) euh j` t'envoie c'est su- t` as #surement raison
fig                                         #fig.1
2      mais j` me souviens plus
```



Figure 1

Dans les deux exemples suivants, considérés comme des cas complexes, les regards sont également annotés comme étant orientés vers l'allocutaire. En effet, malgré une disposition particulière des caméras, il est possible de reconstituer vers qui le regard est orienté. Dans l'exemple (68), la combinaison des deux images permet ainsi de déterminer que le regard du locuteur est dirigé vers l'allocutaire du marqueur, tandis que dans l'exemple (69), c'est le plan d'ensemble suivant directement le marqueur (fig. 3) qui permet d'affirmer que la direction du regard en fig. 1 et fig. 2 (gros plan) est dirigé vers l'allocutaire PON ; absence de mouvement de la tête, regard fixe et présence d'une deixis verbale dans le TDP. Soulignons que c'est la somme de ces différents éléments qui nous permet d'identifier l'orientation du regard ; dans des cas plus ambigus (par exemple, absence de plan d'ensemble dans un contexte proche de gros plans), l'annotateur choisira « indéterminable ».

```
(68) DPU_BIE / 00:39:41.860 - 00:39:55.450 [ID 20010]
1 P01 est-ce que (.) c'est finalement cet intérêt général que moi/ j'ai
2 l'impression\# d'avoir traité comme finalement bien public/ (0.8) ne
fig.                                         #fig.1
```





Figure 1

3 peut pas être comparé (.) sinon par le fait que le bien commun/ (0.4)  
4 peut s'extendre (.) sur le:: euh c` qu'il y a sur le marché/

(69) DTV_VEH / 00 :02 :44.186 - 00 :02 :57.263 [ID 40185]

1 ZEN c'est vrai ça:/ euh cette déclaration que vous avez faites heu c'est  
2 une plaisanterie/ évidemment\  
fig #fig. 1 #fig. 2  
3 PON #.H: c'est une plaisanterie mais qui reflète probablement une  
#fig. 3



Figure 1



Figure 2



Figure 3

4 certaine irritation face à cette nouvelle forme de politiquement  
5 (correct 0.05).h (0.09) dont on aura l'occasion de se dire quelques  
6 mots tout à l'heure\

b. regard sur un tiers → tiers. Ce cas est identifiable au fait que l'énoncé est clairement orienté vers un allocutaire *précis* mais que ponctuellement un autre participant (un tiers) est désigné du regard.

Dans l'extrait ci-dessous, la locutrice répond à la question d'une personne du public. Le regard de la locutrice est orienté vers son allocutrice, identifiée par la deixis verbale (fig. 1) et ensuite vers le public qui est donc, dans ce cas, le « tiers » (fig. 2).

(70) DPU_MED / 00:41:26.626 - 00:41:37.346 [ID 40117]

1 SAN donc si vous voulez:/ c- ENCORE une fois on retom:be SU:R c- cette  
2 EXCELLent:e chose et pire des choses qu'on a xx xx et qui dépend  
3 #évide #mment de la qualité de celui qui l'utilise\  
fig #fig.1 #fig.2



Figure 1



Figure 2

Dans l'extrait ci-dessous, le locuteur formule une (longue) question qui est adressée à l'un de deux débattants. Le regard du locuteur s'oriente ponctuellement vers l'autre débattant (fig. 1) pour s'orienter ensuite à nouveau vers son allocutaire (fig. 2).

(71) DPU_BIE / 00:40:16.640 - 00:40:21.180 [ID 40099]

1 P01 #après évi#demment euh il manque tout le côté du processus (.) pis  
fig #fig.1 #fig.2



Figure 1



Figure 2

- c. regard clairement orienté vers un objet (ordinateur, feuille, maquette, tableau, etc.) → objet. Ce cas est identifiable au fait que le regard est clairement orienté vers un objet sémantiquement et/ou pragmatiquement pertinent au moment de l'énonciation et au sein de l'espace interactionnel coconstruit par les participants.

(72) DPU_UDC / 00:16:35.670 - 00:16:44.130 [ID 40014]

1 CH1 et puis alors me semble-t-il a#près heu on a #vu plutôt euh des  
fig #fig. 1 #fig. 2

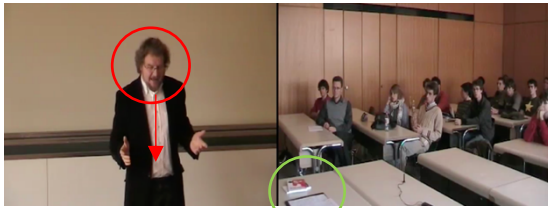


Figure 1



Figure 2

2 journaux assez tranquilles assez assez complaisants assez  
3 neutres/

Dans l'exemple ci-dessus, l'image de gauche nous permet de définir que le regard est orienté vers le bas, tandis que l'image de droite nous indique l'objet regardé. Dans la fig. 1, bien que la tête du locuteur n'apparaisse pas dans l'image de droite, le regard est annoté comme orienté vers un objet (ce qui est confirmé par la figure 2).

- d. regard moyenne distance, au loin, en bas, en haut ou indéterminé → autre.

Dans l'extrait ci-dessous, le regard du locuteur (NA1) est orienté vers le haut.

(73) DPU_REC / 00:22:31.050 - 00:22:38.460 [ID 20002]

1 NA1 euh (0.7) c` qui paraît# logique c'est qu'on commence par aller  
fig #fig.1  
2 voir l'état en lui disant bon on a besoin d'un terrain:/ on a b`soin:  
3 (.) on apporte des emplois/ euh quel comment est-ce qu'on peut faire/



Figure 1



Dans l'extrait ci-dessous, le regard du locuteur semble être dirigé vers la table sans pour autant fixer un objet précis.

(74) REU_CL1 / 00:12:34.296 - 00:12:37.965 [ID 20135]

1 YVA .h oui ils étaient intéressés mais que **probablement** # **ils**  
 fig #fig.1  
 2 **devaient nous mettre en concurrence** hein/

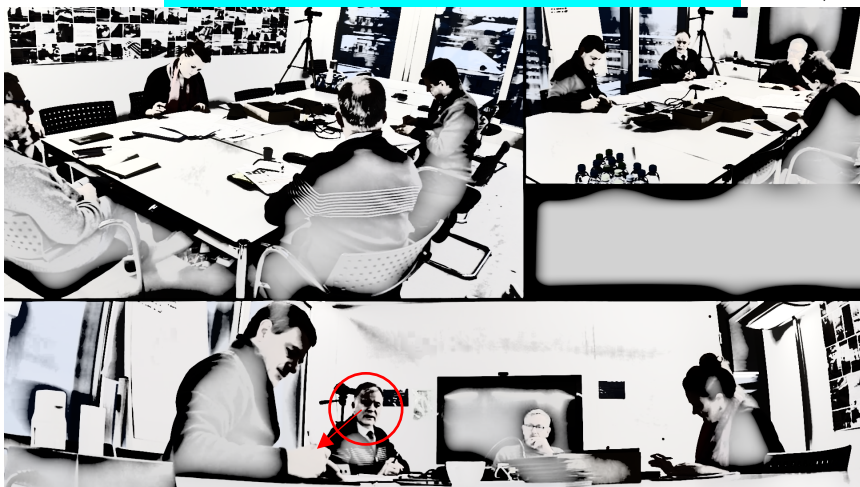


Figure 1

e. yeux fermés → yeux fermés.

(75) DTV_SM1 / 00:32:50.889 - 00:33:02.811 [ID 40180]

1 LEV et et pour répondre à michel zendali peut-être qu'au départ on  
 2 disait qu'on: (0.4) on était censé euh et qu'on devrait défendre/  
 3 (0.3) les y compris les: les salariés étrang#ers **évidemment** **on l'**  
 fig #fig.1



Figure 1

4 **fait** la moitié des [membres d'uni]a c'est [des personnes qui sont  
 5 d'origine étran]gè[re]

Les cas où le locuteur se cache les yeux sont également annotés comme *yeux fermés*.

(76) REU_BL4 / 00:03:50.730 - 00:04:02.065 [ID 20132]

1 CHR il a les PLANches qui s` développent/(0.8) il a euh:: mmh:: le:  
 2 le: **le: projet euh::: FENÊTRE bleue auss[i/]**  
 3 LUC [mhmh]  
 4 (0.8)  
 5 CHR **qui va qui va** #**probablement** **se:: se débloquent/**

fig #fig.1



Figure 1

- f. direction du regard indéterminable parce que les prises de vues ne permettent pas de l'établir avec certitude → indéterminable.

Dans le cas ci-dessous, il est possible de déterminer la direction du regard avant et pendant la production du marqueur, car l'allocutrice est visible sur l'une des caméras et en l'occurrence, sur l'autre caméra, le regard du locuteur est orienté vers l'allocutrice (fig. 1 et 2). Toutefois, après la production du marqueur, le visage du locuteur se trouve hors caméra et donc la direction du regard, à ce moment-là, est indéterminable (fig. 3). Il est impossible d'affirmer s'il y a un changement de regard ou si le regard est continu vers son allocutrice.

(77) REU_BL4 / 00:01:35.593 - 00:01:39.401 [ID 20131]

1 CHR <((en entrant dans la pièce et à quelqu'un dans le couloir)) on va  
2 pas pouvoir faire une séance d'une heure #proba#blement\# ça ça va  
fig #fig.1#fig.2 #fig.3  
3 durer une minute j` pense\>



Figure 1

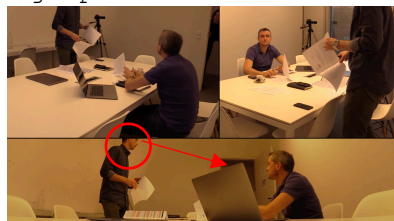


Figure 2

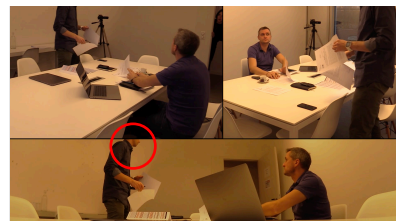


Figure 3

Dans l'extrait ci-dessous, le regard est catégorisé comme indéterminable, car il n'est pas possible de définir l'orientation du regard du locuteur (regard vers le bas (autre) ? Regard sur des feuilles (objet symbolique) ?) ; aucun indice – qui serait présent sur un plan d'ensemble – ne nous permet d'identifier ce regard.

(78) DTV_SM2 / 00:35:52.350 - 00:36:09.350 [ID 40002]

1 BER euh:: sergio rossi nous attend vous êtes euh vous êtes l'expert  
2 euh sergio rossi ce soir euh professeur d'économie .H: euh: quatre  
3 mille francs pour l'ensemble du pays sur l'ensemble du pays pour  
4 tous les cantons (0.2) est-ce que ça vous #semble/(0.4) vous (.)  
fig #fig.1



Figure 1

5 qui enseignez à l'université de fribourg un bon moyen finalement de  
6 lutter contre la pauvreté en suisse

### 3.6.1.1 Regard avant le marqueur (col. AB)

Il s'agit ici de choisir, parmi les options présentées ci-dessus, la direction du regard qui précède immédiatement la réalisation du marqueur (col. N).

### 3.6.1.2 Regard pendant le marqueur (col. AC)

Il s'agit ici de choisir, parmi les options présentées ci-dessus, la direction du regard pendant la réalisation du marqueur (col. N). S'il s'agit d'un marqueur relativement long et que des changements de direction de regard interviennent pendant son émission au point qu'aucun regard « majoritaire » ne se dégage, le cas est consigné et discuté en équipe.

### 3.6.1.3 Regard après le marqueur (col. AD)

Il s'agit ici de choisir, parmi les options présentées ci-dessus, la direction du regard qui suit immédiatement la réalisation du marqueur (col. N).

## 3.6.2 Gestualité (relative au marqueur)

### 3.6.2.1 Présence ou non de geste(s) accompagnant le marqueur (col. AE)

Pour l'annotation de la gestualité, il s'agit d'abord d'établir s'il y a présence ou non d'un geste pertinent accompagnant le marqueur (dans le cas où il est impossible de le déterminer, par exemple si le locuteur est invisible, choisir l'option correspondante dans la liste déroulante).

En se fondant sur la littérature existante concernant les gestes coverbaux (notamment Ferré, 2010; Schegloff, 1984; ter Bekke et al., 2020), trois critères sont retenus pour établir si le lemme du marqueur considéré doit bien être considéré comme un « *lexical affiliate* » (Schegloff, 1984) d'un geste observé :

- a. Critère temporel : le *stroke* du geste observé doit apparaître dans une fenêtre de - 2.0 secondes à +1.0 seconde par rapport au lemme du marqueur.
- b. Critère typologique : le geste doit être de type représentationnel, ce qui inclut les gestes iconiques, métaphoriques, déictiques ainsi que les emblèmes (pour une introduction à cette typologie, voir Kipp, 2003; McNeill, 1992, 2005). Les déictiques peuvent être réalisés par n'importe quelle partie du corps (par ex. pointage du menton).
- c. Critère sémantique : tous les gestes déictiques apparaissant dans la fenêtre (critère a) sont annotés, mais les emblèmes, les gestes iconiques et les gestes métaphoriques ne sont annotés qu'à partir du moment où ils ont un lien sémantique avec le marqueur (col. M) ("for co-speech gestures to play a role in semantic predictive processing, it is crucial that the gestures [...] contain semantic information that is related to speech", in ter Bekke et al., 2020, p. 1). En d'autres termes, s'il y a un geste iconique, métaphorique ou emblématique saillant dans la fenêtre temporelle examinée mais que le geste a un rapport avec un élément de la portée et non du marqueur, il n'est pas pris en compte.

### 3.6.2.2 Type de geste (col. AF)

La typologie des gestes utilisée pour l'annotation est celle de Kipp (2003; sur la base de McNeill, 1992), qui présente l'avantage de proposer un dictionnaire illustré des principaux types de gestes observés et annotés sur corpus. La répartition de certains types de gestes entre les catégories métaphoriques, iconiques et emblématiques est parfois discutable, notamment pour les gestes tournant autour de la préhension, mais cela n'empêche pas d'utiliser ce dictionnaire à ce stade.

Dans le cas où deux gestes (ou plus) doivent être annotés, les gestes sont indiqués dans la case en commençant par les éventuels gestes déictiques et en les séparant par un point-virgule (ex. Deictic.Space;Iconic.Write).

S'agissant des gestes déictiques, une adaptation du modèle de Kipp a été nécessaire, (i) en remplaçant le cas de Deictic.Space par celui de Deictic.Indeterminable pour les situations où le référent du geste de pointage est invisible sur les prises de vues (que ce soit pour des raisons de caméras ou de mode d'accès au référent), et (ii) en ajoutant le cas de Deictic.Object pour le pointage vers un objet visible sur les prises de vue :

- ~~Deictic.Space~~ => Deictic.Indeterminable

(79) DTV_VEH / 00:07:37.906 - 00:07:44.078 [ID 20009]

1 NOR si on fait pas ces deux choses c'est notre prospérité[é qui est]&  
2 PON [mais]  
3 NOR &en jeu c'est notre confort qui est en jeu\ (0.2) #et **contrairement**  
fig #fig.1



Figure 1

4 NOR **à c` que (.)** **a l'air de** **[dire monsieur]**

(80) DTV_REC / 00:44:21.540 - 00:44:26.685 [ID 40228]

1 CHE le droit de recours aussi# sur les murs [comme on vous l'a expliqué&  
fig #fig.1



Figure 1

2 NA3 [pas pour les décisions populaires]  
3 &mais **visiblement** **vous n'entendez PAS/ ce qu'on vous explique\]**

Attention, dans l'exemple ci-dessous, le geste déictique en fig. 1 est annoté comme un Deictic.Addressee (Kipp, 2003, p. 220). En effet, bien que le référent du geste de pointage soit invisible sur cette prise de vue, le plan d'ensemble qui suit nous permet de situer le référent dudit geste. Dans des cas plus ambigus (par exemple, absence de plan d'ensemble dans un contexte proche de gros plans), l'annotateur choisira indéterminable, et ce même s'il y a présence d'une deixis verbale.

(81) DTV_VEH / 00:57:57:477 - 00:58:05.061 [ID 20012]

1 ZEN [on a- on arrive au terme on arrive au terme [madame rebetez juste  
2 cette question à]&  
3 ND3 [non j` crois qu` y a  
4 un problème de génération]  
5 ZEN #vous/ euh vous avez l'impression que c'est une question de  
fig #fig.1



Figure 1



Figure 2

6 généra#tion/ (0.4) ou d` siècle comme dit monsieur rossi/ (0.3) on  
fig #fig.2  
7 a changé d` siècle/

- Deictic.Object

(82) REU_BM3 / 00:15:23.682 - 00:15:26.142 [ID 21175]

1 ROB [ouais c'est ça/ c'est s#ûr] qu` si on fait le TOUR  
fig #fig.1  
2 y en aura plein\ [c'est clair\]



Figure 1

## 4. Annotation de la portée

### 4.1 Informations relatives à la portée elle-même

#### 4.1.1 Identification de la portée (col. AG)

L'identification de la portée renvoie au code de la portée correspondant à celui du marqueur. L'information a déjà été introduite à l'étape 2.2.1.2.



Ex. 56791

#### 4.1.2 Timing début (col. AH) et fin (col. AI) de la portée

Le *timing début* et le *timing fin* de la portée renvoient aux coordonnées temporelles de la portée d'un marqueur au sein de l'événement. Il s'introduit en millisecondes à partir des informations stockées dans ELAN à l'étape 2.2.1.1 ci-dessus).

Ex. la portée ayant l'ID 56791 a comme timing début 698916 et comme timing fin 702781.

#### 4.1.3 Réalisation de la portée (col. AJ)

La *réalisation de la portée* indique la portée telle qu'elle a été effectivement et fidèlement transcrite dans ELAN. Il suffit dès lors de copier-coller la portion de transcription dans ELAN correspondant à la portion identifiée comme portée (col. AG).

Dans le cas où le marqueur apparaît à l'intérieur de la portée (voir 5.1. ci-dessous : Direction), le marqueur est mis entre signes {...}. Ex. « elle est {peut-être} passée la semaine dernière ».

### 4.2 Informations relatives au locuteur de la portée

#### 4.2.1 Locuteur de la portée (col. AK)

Le *locuteur de la portée* renvoie au participant ayant prononcé la portée du marqueur.

Ex. BOR

#### 4.2.2 Place interactionnelle du locuteur de la portée (col. AL)

La place interactionnelle du locuteur de la portée renvoie au rôle / à la fonction du locuteur de la portée au sein de l'événement (voir *supra*). L'information est automatiquement introduite lors de la saisie du locuteur de la portée (col. AK).

Ex. BOR → participant à la réunion

Ex. LOR → modérateur

#### 4.2.3 Statut socio-professionnel du locuteur de la portée (col. AM)

Le statut socio-professionnel du locuteur de la portée renvoie aux caractéristiques identitaires du locuteur de la portée qui ont justifié sa participation à l'événement à la place interactionnelle mentionnée (voir *supra*). L'information est automatiquement introduite lors de la saisie du locuteur (col. AK).

Ex. BOR → directeur

Ex. LOR → journaliste à la RTS

## 4.3 Informations relatives à l'/aux allocutaire(s) de la portée

### 4.3.1 Allocutaire de la portée (col. AN)

Lorsqu'une seule personne physique est clairement identifiable comme allocutaire par la présence d'une deixis personnelle et/ou par une direction du regard franche (col. AC), l'annotateur le choisit dans la liste (par ex. DUM)⁸.

Lorsqu'un sous-groupe est clairement identifiable par la présence d'une deixis personnelle et/ou par une direction du regard franche (col. AC), l'annotateur le choisit dans la liste (par ex. sous_groupe_DTV_VEH). Il peut s'agir d'un sous-groupe de deux invités externes à une réunion professionnelle (ANT et CHA dans REU_BL1), du sous-groupe constitué par les débattants invités d'un débat, etc. Le public sur place d'un débat public a par contre un identifiant particulier (par ex. 01 vaut pour le public du débat DPU_BIE), tout comme, les téléspectateurs d'un débat télévisé, adressés via le regard-caméra (par ex. CT5 vaut pour la caméra-téléspectateurs de DTV-SM1).

Lorsque le locuteur s'adresse à toutes les personnes présentes sans distinction et sans précision, ou ne semble pas avoir d'allocutaire déterminé, l'annotateur choisit l'option « tous » correspondant à l'événement (Ex. « tous REU_AM1 »).

### 4.3.2 Place interactionnelle de l'allocutaire de la portée (col. AO)

La place interactionnelle de l'allocutaire de la portée renvoie au rôle / à la fonction de l'allocutaire de la portée au sein de l'événement (voir *supra*). L'information est automatiquement introduite lors de la saisie de l'allocutaire (col. AN).

### 4.3.3 Statut socio-professionnel de l'allocutaire de la portée (col. AP)

Le statut socio-professionnel de l'allocutaire de la portée renvoie aux caractéristiques identitaires de l'allocutaire de la portée qui ont justifié sa participation à l'événement à la place interactionnelle mentionnée (voir *supra*). L'information est automatiquement introduite lors de la saisie de l'allocutaire (col. AN).

## 5. Annotation de la relation

### 5.1 Direction (col. AQ)

En reprenant les catégories du projet MODAL, la *direction* renvoie à la relation temporelle entre le marqueur et la portée. La direction est identifiée sur la base de ce qui est réellement réalisé, même dans le cas de portée tronquée (ex. « ça me paraît » ; vous avez l'air » ; c'était je crois »). On distingue 3 possibilités⁹ :

- marqueur>portée : le marqueur précède la portée

⁸ Se référer également au Tableau 1 en section 3.5.1 pour une discussion systématique des possibilités d'identification du ou des allocutaire(s).

⁹ Le projet MODAL identifie une quatrième possibilité, qui correspond au cas où le marqueur et sa portée sont co-extensifs. Cette situation est impossible dans le cadre du présent guide, qui repose radicalement sur l'idée qu'un marqueur doit avoir, même en cas de chevauchement partiel, une portée distincte de lui.

(83) DPU_ETU / 00:07:34.469 - 00:07:37.139 [ID 24009]

1 SAN j` **sais pas si** je vais vous présenter heu son voisin

- portée>marqueur : la portée précède le marqueur

(84) REU_BM3 / 00:23:45.437 - 00:23:47.895 [ID 20170]

1 SAL i` faudrait l'ajouter ici aussi\ **°surement°**

- à l'intérieur : le marqueur est à l'intérieur de la portée

(85) REU_AM2 / 00:20:11.376 - 00:20:16.647 [ID 20169]

1 ROG marquA:::ge intéri[eur ascensEU:R morges ç]a c'est un VIEUX trUc/  
2 CAR [°j` sais pas c` qu'elle a fait (bon bref)°]  
3 CAR OUI **ça s`ra** **surement** jamais

## 5.2 Source (col. AR)

La *source* renvoie au rapport entre les origines énonciatives respectives du marqueur et de la portée. Deux possibilités¹⁰ sont calculées automatiquement à partir des informations introduites dans les colonnes renvoyant aux locuteurs respectifs du marqueur (V) et de la portée (AK) :

- ML (même locuteur) : le marqueur et la portée sont émis par le même locuteur

(86) DTV_VEH / 00:51:29.823 - 00:52:03.218 [ID 41155]

1 PET le véhicule a plus de chances de se retourner et heu: (.) plus  
2 de risques de décès pour ses occupants.h et il a plus de  
3 risques pour les piétons ou: en cas de collision ou les vélos  
4 ou les motards.h donc juste pour cette raison **ça nous paraît**  
5 **évident que** ce type de véhicules pas tous (.) mais ceux qui  
6 sont dangereux: ou trop dangereux doivent être retirés de la  
7 vente/

- AL (autre locuteur) : le marqueur et la portée sont émis par deux locuteurs différents

(87) REU_AC1 / 00:22:39.500 - 00:23:28.372 [ID 43501]

1 BOR j- je saute du coq à l'âne j'ai j'ai un copain qu` était à new  
2 york et qui est allé dans cette i- dans cette heu ce (1.0)  
3 heu:: (.) **ça s'appelle sleep no more (0.3) et en fait on on**  
4 **entre dans un::: (0.3) y a des énigmes à rés- c'est une**  
5 **immense maison: (.) .h: et puis on entre pis on on fait**  
6 **soi-même son chemin:(.) mais tous les gens qui sont là: ont**  
7 **un rôle sont des acteurs**  
8 (0.2)  
9 CAR ah mais **j'en ai entendu parler au[ssi]**

¹⁰ Le projet MODAL identifie une troisième possibilité, correspondant au cas suivant : « the marker and the scope are uttered by the same speaker but in two different speech turns » (Pietrandrea, 2022). Cette option nous semblant théoriquement et analytiquement difficile à mettre en place et pas forcément pertinente, nous avons fait le choix de simplifier en ne conservant que l'option ML (le marqueur et la portée sont émis par le même locuteur).



### 5.3 Polarité (col. AS)

La *polarité* renvoie au type de validation que le marqueur propose relativement au contenu propositionnel de la portée. Cette validation n'est pas binaire mais scalaire, ce qui signifie qu'elle admet des degrés intermédiaires et que chaque polarité doit être vue comme un pôle vers lequel le marqueur peut tendre.

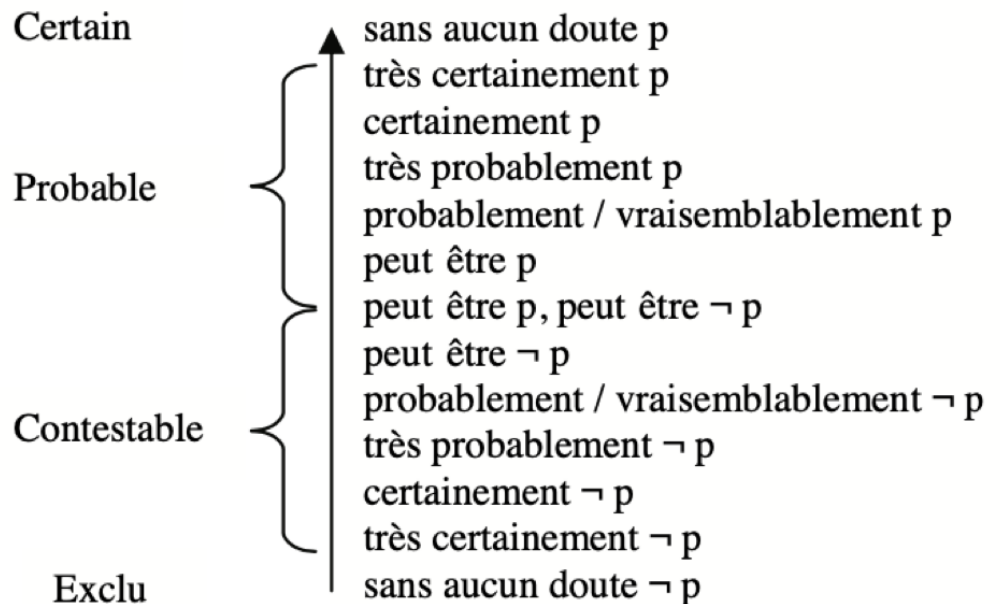


Figure 3 : De l'exclu au certain (Gosselin, 2010, p. 86)

Cette validation est évaluée hors contexte, c'est-à-dire qu'on s'en tient strictement à la relation entre le marqueur et sa portée, sans autres informations contextuelles. En effet, il est possible qu'une relation entre un marqueur et sa portée soit neutre au niveau de la polarité (ce qu'on souhaite bien annoter à ce niveau), mais que prise dans un contexte plus large, on comprenne que ce neutre bascule vers le positif ou le négatif. Reprenant la typologie du projet MODAL, trois possibilités sont ainsi distinguées :

- a. positive : « the modal construction puts forward the propositional content of the scope as (mostly) true » (Pietrandrea, 2022). Cela regroupe les cas de validation stricte (Ex. « je pense que tu es guéri ») comme les cas tendant vers la validation (Ex. « il est probablement venu », « il va sans doute venir », « ça doit être quelque chose comme soixante mille »).

(88) REU_BL2 / 00:50:14.759 - 00:50:17.049 [ID 20130]

1 SOP on s` verra (.) **probablement** pas:: beaucoup cette semaine

(89) DTV_SM1 / 0:39:14.517 - 00:39:20.241 [ID 23210]

1 LEV on vient de voter au parlement treize milliards (.) pour  
 2 l'agriculture sur quatre ans\ j` l'ai soutenu **j` pense que** c'est  
 3 **juste**  
 4 WEI °°ouais°°

(90) DTV_SM1 / 00:28:14.239 - 00:28:29.019 [ID 20153]

1 JAN euh:: BRUT à mettre peut-être en relation avec les bénéfices de la

2 migros f- les derniers chiffres disponibles c'est deux mille cinq/  
 3 c'est six cents nonante-neuf millions\  
 4 (0.20)  
 5 STE alors **certainement pas migros valais:/** éventuellement ((petit  
 6 rire))) l'ensemble de la [migros mais de migros suis]se et [dix  
 7 co]opératives région[ales .h]

(91) REU_CL1 / 00:48:06.877 - 00:48:12.117 [ID 20015]

1 YVA mais::: (0.6) pfff: **j'ai l'impression qu'il va rester en HAUT**  
 2 selon MOI mais[:]  
 3 PAU [ça fait] pas tant d` (boulot) qu` ça finalement/

(92) REU_CL4 / 00:21:18.523 - 00:21:26.067 [ID 40221]

1 SEB [mais c'était deux y a deux ou] trois ans déjà ça non/  
 2 WIL °ouais ouais°  
 3 OLR **apparemment il s` passe pas grand-CHOSE/** °donc°  
 4 WIL non\

(93) DPU_UDC / 00:15:42.324 - 00:16:09.342 [ID 42042; 21502; 40088]

1 CHE il souvent il fallait lire entre les lignes heu tout  
 2 n'était pas extrêmement clair difficile d'en déduire un  
 3 positionnement général du journal c'est pas la nature du  
 4 journal d'ailleurs/ (0.4) et **son rédacteur en chef**  
 5 **(0.2) heu::: roth disait à juste titre me semble-t-il qu'au**  
 6 **fond/ (0.14) c'est heu (.) le temps a beaucoup**  
 7 **fonctionné dans cette campagne dans l` structure**  
 8 **d'accueil de positions an- de de de positions de (0.3) de**  
 9 **personnes qui: se sont EXPRIMÉES à l'intérieur des**  
 10 **colonnes du temps (.) sans que ce soit l'avis du temps**

(94) DPU_ETU / 00:44:20.846 - 00:45:26.527 [ID 41355; 41356]

1 DUM [euh moi] j` suis pas tout à fait d'accord qu` ça n'existe pas plus  
 2 près heu en université par exemple heu saint-gall est f- est financé  
 3 en partie par de:s par des (.) entreprises privées bon ça reste (0.7)  
 4 le domaine économique mais j` pense que c'est pa::s et et  
 5 contrairement à ce qu'on dit y a pas une influence directe de  
 6 l'entreprise sur euh .h [sur les] c` qui est enseigné dans dans &  
 7 SAN **[si si]**  
 8 DUM & dans les cours\ NON j` suis [désolée\  
 9 SAN **[C'EST RECONNU c'est] c'est prouvé**

Dans ce dernier exemple, si on admet que « c'est reconnu » et « c'est prouvé » portent sur « si si » (= oui c'est le cas, il y a une influence directe de l'entreprise sur ce qui est enseigné dans les cours), alors les deux marqueurs ont une polarité positive, même si, par rebond, ils sont négativement polarisés par rapport au contenu « y a pas une influence... ».

- b. négative : « the modal construction puts forward the propositional content of the scope as (mostly) false » (Pietrandrea, 2022). Cela regroupe les cas d'invalidation stricte (Ex. « je ne pense pas que tu sois guéri » => invalidation du contenu propositionnel [tu es guéri]) comme les cas tendant vers l'invalidation (« je suis pas sûr que »).

(95) REU_AC1 / 00:18:16.837 - 00:18:26.424 [ID 23263]

1 CAR [((ri)r(es))]  
 2 JEA [non: eh moi] je m'excuse mais je vais directement au but parce que

3 j'ai eu une idée (0.5) j` pense pas que ça pourrait être la seule  
4 idée mais sérieusement j'y crois beaucoup à cette idée

(96) REU_AC1 / 00:15:24.369 - 00:15:27.733 [ID 21105]

1 JEA ça c'est ça c'est votre avis/  
2 BOR ouais  
3 CAR oui  
4 JEA [j` suis pas sûr qu` c` soit l'avis de tout le m[onde]  
5 CAR [ah ben] ouais

(97) DTV_SM1 / 00:15:48.244 - 00:15:54.592 [ID 21217]

1 WEI OU euh je doute je doute que: qu'avec ce salaire si c'est une  
2 personne seule elle bénéficie de de l'aide sociale\

(98) DPU_ETU / 01:27:50.940 - 01:28:01.450 [ID 20214]

1 FRI heu: donc moi moi j` suis pour garder/ heu:: une pa:rt/ (.) heu  
2 disons un peu de: (0.4) pas forcément d` rêve/ mais disons de de  
3 POTENTIALITÉ/: euh dans dans en tout cas dans les universités\

- c. neutre : « the modal construction puts forward the truth-value of the propositional content of the scope as not decidable » (Pietrandrea, 2022). Cela regroupe les cas où le marqueur ne se positionne pas quant à la vérité de la portée – qu'il ne fait que prendre en compte (voir par ex. Coltier et al., 2009) – ou que le positionnement est simplement indécidable. Comme pour le cas de la polarité positive ou négative, cela s'évalue strictement au niveau du rapport entre le marqueur et sa portée, en faisant abstraction du contexte. Ainsi, pour les discours rapportés, seront jugés neutres les cas où le locuteur rapporte des propos grâce à un marqueur qui ne contient pas par ailleurs de marquage d'accord ou de désaccord quant au contenu rapporté (distinguer « il a dit que » de « vu qu'il a dit que », « comme il dit », « il a révélé que », etc.).

(99) REU_CL4 / 00:21:18.523 - 00:21:26.067 [ID 24777]

1 YVA [remis en question/]  
2 OLR [on on sait pas c` qu'i`] s'est pa[ssé depuis/]  
3 WIL [<((souffle)) H::>]

(100) DTV_REC / 00:56:42.099 - 00:56:51.993 [ID 45001]

1 NID on est (.) vraiment au cœur d'un débat qui apparaît pour ce  
2 qu'il est/ c'est-à-dire un débat idéologique\ qui a d'un côté  
2 les organisations de défense de la nature qui auraient le  
3 monopole/.h: de la conscience écologique comme le parti  
4 socialiste français celui du cœur peut-être\

(101) DTV_DEM / 00:13:24.562 - 00:14:24.003 [ID 42248]

1 BEN on est en train de parler: .h de ce mot: .h qui fait irruption dans  
2 la politique suisse et moi je dis de ne pas être nAïf\ .h la  
3 gauche aujourd'hui dit que .h mertz et couchepin complotent contre  
4 les assurances sociales/ .h veulent ca:sser la sécurité sociale  
5 en- (.) aujourd'hui (.) c'est le thème que vous avez xx aujourd'hui  
6 alors vous êtes mal placé j'entends pour faire ce genre de chose .h  
7 la deuxième [chose (.) la deu:xi]ème&

(102) REU_AC1 / 00:22:39.500 - 00:23:28.372 [ID 40231]

1 CAR ah mais j'en ai entendu parler au[ssi]  
2 BOR [c'est] il paraît que c'est

3 [exceptionnel\]  
 4 CAR [oui c'est le] truc [hyper branché]

(103) REU_AC1 / 00:34:32.187 - 00:34:35.670 [ID 41300]  
 1 JEA [j` j` suis désolé] là je [dois pas ici on]&  
 2 BOR [j'ai ouais ouais]  
 3 JEA &doit pas couper [mais là]&  
 4 BOR [ouais ouais\]  
 5 &là c'[est écrit] pas de fun  
 6 BOR [ouais ouais/]

Le cas ci-dessous articule deux polarités différentes :

(104) REU_AC1 / 00:18:16.837 - 00:18:26.424 [ID 23263]  
 1 JEA on: eh moi] je m'excuse mais je vais directement au but parce que  
 2 j'ai eu une idée (0.5) j` pense pas que ça pourrait être la seule  
 3 idée mais sérieusement j'y crois beaucoup à cette idée en fait  
 4 c'est le mot bulle

La relation entre « je pense pas que » et « ça pourrait être la seule idée » est polarisée négativement, tandis qu'à l'intérieur, la relation entre « ça pourrait » et « être la seule idée » est neutre. On a donc ici un locuteur qui tend vers l'invalidation d'une éventualité.

## 6. Résultats des scores d'accords inter-annotateurs

Le présent guide a été soumis à trois tests d'accords inter-annotateurs successifs, chacun ayant mené à des modifications plus ou moins substantielles avant d'opérer le test suivant, de manière à atteindre des scores d'accord allant de « substantiel » (Kappa de Cohen entre 0.6 and 0.8) à « presque parfait » (entre 0.8 and 1.0) pour chaque variable (Cohen, 1960; Landis & Koch, 1977). Le Kappa de Cohen a été adapté pour les variables continues, notamment tout ce qui relève de l'annotation des segments identifiés comme « marqueur » et « portée » (Bavaud, 2021).

1. Le premier test, effectué en février 2021, a été consacré spécifiquement à la question du tri des faux positifs (voir *supra*, 1.2). Quatre annotatrices et annotateurs experts ont ainsi annoté 50 tokens en choisissant pour chacun si les critères retenus exigeaient ou non de le retenir pour l'analyse. Un score d'accord a été attribué à chaque binôme, pour un total de 6 scores. L'ensemble des 6 scores a dépassé un Kappa de Cohen de 0.6 (de l'accord « substantiel » à l'accord « presque parfait »). Le score d'accord entre les deux annotatrices concernées par les deux tests suivants ainsi que par l'annotation complète des marqueurs retenus par le projet a été de 0.808 (seuil de l'accord « presque parfait »). La discussion qui a suivi cette phase a permis d'explicitier encore davantage quelques critères de sélection des tokens à annoter.
2. Le second test, effectué en mars 2021, a permis d'évaluer le guide une première fois. 50 tokens ont été annotés par deux annotatrices expertes. Les discussions qui ont suivi ont mené à une révision du guide en conséquence, avec un accent particulier mis sur les variables pour lesquelles le Kappa de Cohen ne dépassait pas un score d'accord satisfaisant (fixé à 0.6) : Position de l'UCT dans le TDP (voir *supra*, 3.3.2), Rôle de l'UCT dans la séquence (voir *supra*, 3.3.3), allocutaire(s) du marqueur (voir *supra*, 3.5.1), Gestualité (voir *supra*, 3.6.2).
3. Un troisième et dernier test été effectué en avril 2021, par l'annotation de 30 tokens supplémentaires, à nouveau par les deux annotatrices expertes. Les résultats ont été très

significativement améliorés pour les variables qui s'étaient révélées les plus problématiques à l'issue du test précédent. Les discussions sur les résultats de ce test ont mené à quelques modifications plus superficielles apportées au présent guide, qui a ensuite été publié dans sa version 1 (en 2022).

## 7. Mise à disposition de la base de données

La base de données constituée des marqueurs épistémiques et évidentiels identifiés, extraits et annotés au travers du présent guide (près de 4000 tokens) est consultable et interrogeable dans une instance fournie par le DaSCH (Jacquin, Keck, et al., 2024). Celle-ci permet des recherches multicritères ; par exemple, il est possible de filtrer les données en ne retenant que les occurrences de « peut-être » en position réactive et accompagnées d'un haussement d'épaules, ou les « je sais pas » sans complémentation syntaxique et émergeant en fin de tour de parole. Grâce au système de permaliens pointant vers CLAPI/POSEPI (voir *supra*, 3.2.1), il est possible de consulter chaque occurrence en contexte, c'est-à-dire au sein du segment transcrit et associé aux données audiovisuelles.

## Références

- Bavaud, F. (2021). *Similarité entre intervalles* [Notes].
- Clayman, S. E. (2013). Turn-Constructional Units and the Transition-Relevance Place. In J. Sidnell & T. Stivers (Éds.), *The Handbook of Conversation Analysis* (p. 150-166). Wiley-Blackwell.
- Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), 37-46. <https://doi.org/10.1177/001316446002000104>
- Coltier, D., Dendale, P., & De Brabanter, P. (2009). La notion de prise en charge : Mise en perspective. *Langue française*, 162, 3-27.
- ELAN (Version 6.6) [Computer software] (6.6). (2023). [Logiciel]. Max Planck Institute for Psycholinguistics, The Language Archive. <https://archive.mpi.nl/tla/elan>
- Ferré, G. (2010). *Timing Relationships between Speech and Co-Verbal Gestures in Spontaneous French*. W6, 86-91. <https://hal.archives-ouvertes.fr/hal-00485797>
- Gosselin, L. (2010). *Les modalités du français. La validation des représentations*. Rodopi.
- Jacquin, J. (2014). *Débattre. L'argumentation et l'identité au coeur d'une pratique verbale*. De Boeck.
- Jacquin, J., Etienne, C., Keck, A., Petitjean, C., Robin, C., Roh, S., & Stern, G. (2024). *Corpus POSEPI* [jeu de données]. CLAPI, ENS-Lyon (France). <http://clapi.icar.cnrs.fr/Posepi>
- Jacquin, J., Keck, A., Robin, C., Roh, S., & Rivoal, M. (2024). *Marqueurs épistémiques et évidentiels du français identifiés et annotés dans un corpus d'interactions relevant de débats politiques et de réunions d'entreprise* ([Projet FNS 188924, POSEPI]) [jeu de données]. DaSCH. <https://ark.dasch.swiss/ark:/72163/1/0120>
- Keck, A. (en préparation). *Le degré de certitude en français. Étude systématique de la modalité épistémique dans un corpus d'interactions sociales*. Université de Lausanne.
- Kipp, M. (2003). *Gesture generation by imitation : From human behavior to computer character animation* [Universität des Saarlandes]. <http://dx.doi.org/10.22028/D291-25852>
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159-174. <https://doi.org/10.2307/2529310>
- McNeill, D. (1992). *Hand and Mind : What Gestures Reveal About Thought*. University of Chicago Press.
- McNeill, D. (2005). *Gesture & Thought*. University of Chicago Press.
- Mondada, L. (2019). *Conventions for multimodal transcription*. <https://www.lorenzamondada.net/multimodal-transcription>
- Nuyts, J. (2001). *Epistemic Modality, Language, and Conceptualization*. John Benjamins.
- 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. <https://doi.org/10.1016/j.pragma.2016.06.014>
- Pekarek Doehler, S. (2019). At the Interface of Grammar and the Body : Chais pas ("dunno") as a Resource for Dealing with Lack of Recipient Response. *Research on Language and Social Interaction*, 52(4), 365-387. <https://doi.org/10.1080/08351813.2019.1657276>
- Pietrandrea, P. (2018). Epistemic constructions at work. A corpus study on spoken Italian dialogues. *Journal of Pragmatics*, 128, 171-191. <https://doi.org/10.1016/j.pragma.2017.10.006>
- Pietrandrea, P. (2022). *Modal. Modèles de l'annotation de la modalité à l'oral*. <https://modal.msh-vdl.fr>
- Pietrandrea, P., & Cervoni, V. (2016). *Modal – annotation guidelines. Version 1.0*. [http://paolapietrandrea.altervista.org/guidelines_decision_trees_for_French_with_examples.pdf](http://paolapietrandrea.altervista.org/guidelines_decision_trees_for_French_with_examples.pdf)
- Robin, C. (en préparation). *Marquer la source de l'information : Approches interactionnelle, énonciative et multimodale de l'évidentialité en français*. Université de Lausanne.
- Rossano, F. (2013). Gaze in conversation. In J. Sidnell & T. Stivers (Éds.), *The Handbook of Conversation Analysis* (p. 308-329). Wiley-Blackwell.
- Schegloff, E. A. (1984). On some gestures' relation to talk. In J. M. Atkinson & J. Heritage (Éds.), *Structures of Social Action : Studies in Conversation Analysis* (p. 266-296). Cambridge University Press/Maisons des Sciences de l'Homme.
- Schegloff, E. A. (2007). *Sequence organization in interaction : A primer in conversation analysis / Volume 1*. Cambridge University Press.
- Selting, M. (2000). The construction of units in conversational talk. *Language in Society*, 29(4), 477-517.
- ter Bekke, M., Drijvers, L., & Holler, J. (2020). *The predictive potential of hand gestures during*

- conversation : An investigation of the timing of gestures in relation to speech* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/b5zq7>
- Thompson, S. A., & Mulac, A. J. (1991). A Quantitative Perspective on the Grammaticization of Epistemic Parentheticals in English. In E. C. Traugott & B. Heine (Éds.), *Approaches to Grammaticalization* (Vol. 2, p. 313-329). John Benjamins.
- Willems, D. (2011). Les degrés d'intégration syntaxique de la modalité épistémique. Le cas de sembler et paraître. In G. Corminboeuf & M.-J. Béguelin (Éds.), *Du système linguistique aux actions langagières. Mélanges en l'honneur d'Alain Berrendonner* (p. 61-72). De Boeck Supérieur. <https://www.cairn.info/du-systeme-linguistique-aux-actions-langagieres--9782801116470-page-61.htm?contenu=resume>

## Remerciements

Le présent guide a été rédigé dans le cadre du projet *Prendre une position épistémique dans l'interaction. Les marqueurs du savoir, du non-savoir et du doute en français* (POSEPI) financé par le Fonds National Suisse de la recherche [subside 100012_188924].

Les auteurs tiennent également à remercier Marion Rivoal, Carole Etienne, Johanna Miecznikowski, Elena Battaglia, Christian Geddo et Paola Pietrandrea pour leurs précieux commentaires à différents stades de rédaction.