
Guide d'annotation ESLO-FLEU

Chloé Tahar, Marie Skrovec et Flora Badin

21 août 2023

Table des matières

1	Introduction	1
1.1	Présentation générale	1
1.2	Méthodologie de l’annotation	2
1.2.1	Références	3
2	D’une situation à l’autre	4
2.1	Liaisons	4
2.1.1	Méthodologie	4
2.1.2	Typologie d’annotation	5
2.1.3	Catégories	5
2.1.3.1	Réalisation de la liaison	5
2.1.3.2	Classification de la liaison	6
2.1.3.3	Enchaînement de la liaison	7
2.1.4	Références	8
2.2	Constructions interrogatives	8
2.2.1	Méthodologie	8
2.2.2	Typologie d’annotation	9
2.2.3	Catégories	9
2.2.3.1	Portée	9
2.2.3.2	Forme	10
2.2.3.3	Position du mot interrogatif	11
2.2.3.4	Inversion	12
2.2.3.5	Fonction	13
2.2.4	Références	14
3	Le parler jeune	15
3.1	Mots du discours	15
3.1.1	Méthodologie	16
3.1.2	Typologie d’annotation	16
3.1.3	Catégories	17
3.1.3.1	Interaction	17
3.1.3.2	Formulation	19
3.1.3.3	Connexion	21
3.1.3.4	Organisation informationnelle	23
3.1.4	Références	25
3.2	Constructions disloquées	26
3.2.1	Typologie d’annotation	27
3.2.2	Catégories	27
3.2.2.1	Dislocation	27
3.2.2.2	Constituant détaché	28

3.2.2.3	Reprise	28
3.2.3	Références	29
4	Récit en interaction	30
4.1	Discours rapporté	30
4.1.1	Typologie d'annotation	31
4.1.2	Quotatif	31
4.1.3	Références	32
4.2	Constructions clivées	33
4.2.1	Méthodologie	34
4.2.2	Typologie d'annotation	34
4.2.3	Clivage	34
4.2.4	Références	35
5	Mémo des requêtes	37
5.1	Principes généraux des requêtes sur TXM	37

Introduction

1.1 Présentation générale

Le projet ESLO-FLEU¹ (ESLO, Fle et Linguistique pour l'Enseignement Universitaire) se donne pour objectif de rendre accessible un grand corpus oral utilisé en linguistique, le corpus des Enquêtes SocioLinguistiques à Orléans², pour la didactique du FLE et l'enseignement de la linguistique en contexte universitaire. Il s'inscrit dans un double contexte scientifique et pédagogique : d'un côté celui de la linguistique actuelle, qui mène des analyses sur des données de corpus enrichies d'annotations, de l'autre, celui des approches de référence en didactique des langues, qui recommandent d'exposer les apprenants à des données authentiques, situées, non fabriquées.

Après une première étape qui a consisté en la création d'un sous-corpus de 150 extraits utilisables en contexte pédagogique de par leur longueur et leur contenu, des modules de formation ont été élaborés par regroupements d'extraits pour former les étudiants aux spécificités du français parlé. Trois de ces regroupements d'extraits ont fait l'objet d'annotations :

- Module 1 : D'une situation à l'autre (liaison, interrogatives)
- Module 2 : Le parler jeune (marqueurs discursifs, dislocations)
- Module 3 : Récit en interaction (discours rapporté, clivées)

Ce document présente les trois modules annotés et détaille les choix d'annotation effectués en proposant des définitions, des exemples ainsi que des références pour chaque phénomène annoté.

Ce travail d'annotation a été guidé par des principes d'ordre linguistique et didactique. Il s'est agi d'adopter une méthodologie linguistique rigoureuse (approche *data-driven* et recherches sur l'état de l'art pour éclairer les analyses) tout en proposant des typologies d'annotation accessibles à des étudiants apprentis-linguistes. Cela a supposé des aménagements terminologiques ponctuels ou un élagage des catégories pour éviter des typologies trop complexes. Toujours selon ce double objectif, le traitement du corpus a été réalisé sous TXM, puis une version html accessible en ligne a été préparée pour les utilisateurs non experts de TXM. Ces données sont disponibles sur le site Ortolang.³

1. <http://eslo.huma-num.fr/index.php/pagelarecherche/projets-de-l-equipe-et-sous-corpus/eslo-fleu>

2. <http://eslo.huma-num.fr/>

3. <https://www.ortolang.fr/market/corpora/eslo-fleu>

1.2 Méthodologie de l'annotation

L'exploration et l'annotation du sous-corpus ESLO-FLEU a été faite avec le logiciel TXM, via une version du corpus au format XML-TEI (écoute de la donnée sonore, métadonnées locuteurs), , avec l'outil d'annotation du logiciel (via des requêtes CQL sur le concordancier). La consultation du corpus annoté est disponible en HTML ou sous TXM. L'annotation sur le logiciel TXM peut se faire selon plusieurs modalités, selon qu'elle cible le mot ou des structures plus grandes appelées séquences de mots :

- **Anotation en propriétés de mot** : Cette modalité d'annotation concerne exclusivement le phénomène de la liaison.
- **Annotation en séquences de mots** : Cette modalité d'annotation concerne tous les phénomènes ciblés, à l'exception de la liaison. Il s'agit d'associer un couple catégorie/valeur à l'ensemble des mots présents dans la séquence.

Chaque objet linguistique à annoter au niveau de la séquence de mots dans le cadre du projet ESLO-FLEU a d'abord été annoté avec la catégorie 'objet', en spécifiant comme valeur le nom de l'objet linguistique qu'il représente (e.g., 'interrogative', 'mot du discours'). Seules les constructions clivées et les constructions disloquées ont été annotées avec la catégorie 'dispositif-syntaxique', selon une même logique d'identification de l'objet linguistique qu'elles représentent. Cette méthodologie a été nécessaire pour un requêtage plus facile des objets linguistiques. Cette structuration aidera par la suite à une facilité de transformation de la donnée.

L'annotation en séquence de mots pose un problème bien connu des annotateurs de corpus d'oral spontané : celui du chevauchement de balises (*markup overlap*). Lors de la transcription, le tour de parole d'un même locuteur est parfois segmenté en plusieurs unités de segmentation, notamment pour rendre compte de certains indices prosodiques (pauses, etc.). Ces unités de segmentation correspondent en quelque sorte aux "phrases" du discours oral. Au niveau du fichier XML, ces unités de segmentation sont identifiées par les balises `<u> . . . </u>`. Voici, à titre d'exemple :

```
<u who="NR390">
...
<!-- segment de discours -->
...
</u>
<u who="NR390">
...
<!-- segment de discours -->
...
</u>
```

Parfois, une séquence de mots à annoter (e.g., une construction interrogative) peut s'étendre sur deux unités de segmentation, comme c'est le cas par exemple en (1), où les balises de l'unité de segmentation sont indiquées par le signe ||.

(1) ESLO2_24H_1250

GEN_client_24H ||euh |[^{INTERROGATIVE}est-ce que vous avez les gants ou
un bonnet qui va avec ou elle est]||
GEN_client_24H ||juste vendue comme ça ?]|||

La syntaxe du langage XML ne permet pas ce type de chevauchement. Comme le remarquent Rühlemann et O'Donnell (2012), la hiérarchie des balises en XML ne permet pas une structure comme $\langle x \rangle \langle y \rangle \langle /x \rangle \langle /y \rangle$, mais permet seulement une structure enchâssée telle que $\langle x \rangle \langle y \rangle \langle /y \rangle \langle /x \rangle$. La solution que nous avons adoptée face à ce problème de chevauchement consiste à segmenter l'annotation de l'interrogative en deux segments. Dans ce cas de figure choisi à titre d'exemple, nous avons choisi d'attribuer deux valeurs à la catégorie 'objet', à savoir 'interrogative' et 'interrogative_suite', comme en (2) :

(2) ESLO2_24H_1250
GEN_client_24H ||euh |[^{INTERROGATIVE}est-ce que vous avez les gants ou
un bonnet qui va avec ou elle est]|||
GEN_client_24H |||[^{INTERROGATIVE_SUITE}juste vendue comme ça ?]|||

Une fois les objets linguistiques identifiés, nous avons procédé à l'affectation de couches d'annotation supplémentaires à partir de typologies d'annotation détaillées par la suite dans le guide d'annotation.

1.2.1 Références

Heiden, Serge et al., (2010), TXM : Une plateforme logicielle open-source pour la textométrie - conception et développement. In *Actes du 10th International Conference on the Statistical Analysis of Textual Data*, Rome, Edizioni Universitarie di Lettere Economia Diritto 2, 1021-1032.

Laboratoire Ligérien de Linguistique - UMR 7270 (LLL) (2023). *ESLO-FLEU : Enquêtes Sociolinguistiques à Orléans - Fle et Linguistique pour l'Enseignement Universitaire* [Corpus]. ORTOLANG (Open Resources and TOols for LANGuage) - www.ortolang.fr, v1, <https://hdl.handle.net/11403/eslo-fleu/v1>.

Rühlemann, Christoph et O'Donnell, Matthew Brook. (2012). "Introducing a corpus of conversational stories. Construction and annotation of the Narrative Corpus" *Corpus Linguistics and Linguistic Theory*, vol. 8, no. 2, pp. 313-350.

D'une situation à l'autre

Les 14 extraits du module "D'une situation à l'autre" ont été sélectionnés dans ESLO2 (période 2008-2021) pour leur représentativité en termes de variation situationnelle (diaphasique) : il s'agissait de former un échantillon des différentes situations de communication des plus formelles aux plus informelles. Les types d'interactions qui y sont représentés sont divers, allant de la conversation familière lors de repas ou à la sortie du cinéma, aux discours académiques en situation de conférence, en passant par des récits en interaction et des discours d'opinion en situation d'entretien. Ce module est également représentatif de la diversité générationnelle (diastatique), dans la mesure où il contient des locuteurs de toutes les tranches démographiques (de 15/25 à + de 65 ans).

Ce module cible l'étude de la **liaison**, en Section 2.1 et de l'**interrogation directe**, en Section 2.2. Il permet de mettre en évidence la variation relative aux types d'interrogatives et à la réalisation ou non des liaisons facultatives en fonction du contexte. Britta Gemmeke, Céline Dugua et Marie Skrovec ont été responsables de l'élaboration de la typologie d'annotation des liaisons et de leur annotation. Pour l'élaboration de la typologie d'annotation des interrogatives directes, sont concernées Dominique Ulma, Marie Skrovec et Chloé Tahar.

2.1 Liaisons

La liaison est un phénomène typique de la phonologie du français qui se caractérise par la production d'un son consonantique à la frontière entre deux mots. Il s'agit d'un cas particulier d'enchaînement, cette consonne n'étant pas prononcée lorsque le mot qui la porte est produit seul (ex : "premier" /pʁɛmjɛ/ entraîne une liaison et "car" /kaʁ/ un enchaînement) et elle peut apparaître quand le mot qui suit est à voyelle initiale (ex : "premier étage" /pʁɛmjɛʁetaʒ/). Les liaisons sont le plus souvent réalisées en /n/, /z/, /t/ et plus rarement en /p/ et /ʁ/. Les contextes de liaison sont généralement classés en catégoriques (toujours réalisées), variables (les formes réalisées et non réalisées sont possibles) et impossibles.

2.1.1 Méthodologie

On isole les contextes de liaison potentiels de manière automatique, plus précisément les mots porteurs d'une potentielle consonne de liaison lorsqu'ils sont suivis d'un mot commençant par une voyelle, avec la requête donnée ci-dessous.

```
a:[_modulefleu_desc="D'une situation à l'autre"&word=".*s|.*x|
.*z|.*d|.*t|.*g|.*k|.*n|.*r|.*p"%cd&word!=".*ar|.*ir|.*or|.*ur|
```

```
bag|basket|ben|bifteck|cet|cher|direct|donc|Eltit|et|foot|hein|
hiver|hm|huit|hum|index|internet|jadis|Jennifer|master|NPERS|
ouais|post-it|sept|set|sud|tennis|test|zut"%cd] [_u_who=a.u_who
&_u_start=a.u_start&word="e.*|a.*|h.*|-(a|e|i|o|u|y).*|
o.*|u.*|y.*|i.*"%cd&word!="hein|hm|hum|ouais|oui"%cd]
```

Notez que les liaisons après *et* (sinon impossibles) sont exclues par la requête. Cette requête permet également que les mots qui se suivent soient prononcés par le même locuteur (a : et `_u_who=a.u_who`). On a également fait en sorte que les mots soient dans le même segment de parole avec `_u_start=a.u_start`. Le `%cd` permet de ne pas faire attention à la casse.

2.1.2 Typologie d'annotation

Ensuite, on annote ces occurrences en propriété de mots.

TABLE 2.1 – Jeu d'étiquettes adopté sur le premier mot de la liaison

Catégories	Valeurs
realisation_liaison	N, Z, T, R, P ABS 0
classification_liaison	catégorique impossible variable ajout 0
enchainement_liaison	AVEC SANS 0

2.1.3 Catégories

2.1.3.1 Réalisation de la liaison

N, Z, T, R, P Pour les liaisons réalisées en /n/, /z/, /t/, /ʁ/, /p/ (même en cas d'ajout)

- /n/ : 'bien aimé' /bjɛ̃neme/

(1) ESLO2_CINE_1187

GD954 j'aurais *bien aimé* un film plus euh il est très très beau hein mais j'aurais
bien aimé un film tu vois un peu plus
 MK912 qui change les idées

- /z/ : 'vous avez' /vuzave/

(2) ESLO2_24H_1250

GEN_client_24H j'ai vu une écharpe ici
 NR390 oui

GEN_client_24H euh est -ce que *vous avez* les gants ou un bonnet qui va avec ou elle est juste vendue comme ça ?

- /t/ : 'tout en' /tutã/

(3) ESLO2_REPAS_1263

GJ845 elle a une veste euh camouflage et les manches pas en cuir mais *tout en* paillettes or

- /ʁ/ : il n'y a pas de liaison en /ʁ/ réalisée dans le corpus, mais des contextes potentiels indiqués ici :

(4) ESLO2_24H_1253

NR390 je vais *jouer au* jeu du pied (/ʒueʁo/)

(5) ESLO2_24H_1250

NR390 ça permet de *cibler un* peu (/sibleʁœ/)

- /p/ : il n'y a pas de liaison en /p/ réalisée dans le corpus, des exemples autres : 'J'ai beaucoup appris durant ce stage' (/bokupapʁi/)

ABS Pour l'absence de liaison (= liaison possible non réalisée) : 'c'est absolument' /seapsolymã/

(6) ESLO2_ENT_1013

ZD13 donc au lieu d'éduquer on fait de la répression voilà

ch_PP6 hm hm

ZD13 donc moi là-dessus *c'est absolument* pas mon point de vue

0 Pour tous les cas hors classification (quand il ne s'agit pas d'une liaison possible) : 'le traitement automatique' /lətʁetmãotomatik/

(7) ESLO2_CONF_1242

ch_GB9 sur le *traitement automatique* des langues il y a des aspects sur lesquels nous ne travaillerons pas ici

2.1.3.2 Classification de la liaison

Catégorique Pour les liaisons qui sont toujours réalisées, c'est-à-dire dans les contextes catégoriques suivants :

- **déterminant + nom** : ex., 'un ami' /œnami/, 'les habitants' /lezabitã/
- **déterminant + adjectif** : 'les autres voisins' /lezotʁəvwazẽ/, 'mon ancien collègue' /mõnãsjẽkolɛg/
- **pronom clitique + verbe** : 'on entend' /õnãtã/, 'ils ont' /ilzõ/
- **verbe + pronom clitique** : 'dit-elle' /ditɛl/
- **quelques expressions figées** : 'tout à fait' /tutafɛ/

Impossible Pour les liaisons qui ne sont jamais réalisées, c'est-à-dire dans les contextes impossibles suivants :

- **après un nom au singulier** : 'génération automatique' /ʒeneʁasjɔtomatik/
- **devant un h aspiré** : des haches /deɑʃ/
- **devant les numéraux** : (huit, onze, huitième, onzième)

Variable Pour les liaisons qui peuvent être réalisées ou pas, c'est-à-dire dans tous les autres contextes de liaison.

- (8) ESLO2_CINE_1187
GD954 Je *fais* un stage dans la protection de l'enfance (/fɛ(z)œ/)

Ajout : Pour des consonnes de liaison produites dans des contextes qui ne sont des contextes potentiels de liaison, (absence de consonne de liaison dans la graphie : à repérer à l'oreille), ex., ajout d'un /t/

- (9) ESLO2_CONF_1243
CPelage une forme de parité visant à réparer cette *injustice historique* (/ɛ̃zɥstɪstɪstɔvɪk/)

0 Pour tous les cas hors classification :

- pour les cas où, en raison d'une frontière syntaxique entre deux constituants, la liaison paraît impossible. Ex : frontière de deux phrases / de deux constituants, dislocations, hésitations, énoncés inachevés, vocatif (appels / interpellations), coordonnants
- (10) ESLO2_REPAS_1270
TG634FIE bah tu *m'étonnes* en même temps comment veux-tu
- s'il y a une pause entre les deux mots et que la liaison n'a pas été réalisée
 - pour tout autre cas où le contexte de liaison paraît impossible. Ex : devant une onomatopée ou une interjection, dans le cas de répétitions
- (11) ESLO2_24H_1253
NR390 dans *les euh* genre euh
- lorsque les annotateurs ne parviennent pas à décider à l'écoute si la liaison est réalisée et/ou comment

2.1.3.3 Enchaînement de la liaison

Avec Pour les liaisons qui sont réalisées avec enchaînement (par défaut)

Sans Pour les liaisons qui sont réalisées sans enchaînement (le # marque une pause)

- (12) ESLO2_DISC_1239
JLBernard une enfant dont l'existence ne dura que dix-neuf années mais qui *fut* une figure emblématique de l'histoire de France (/fyt#ɥnfigyʁ/)

0 Pour tous les cas hors classification et les liaisons absentes

2.1.4 Références

Chevrot, Jean-Pierre; Fayol, Michel et Laks, Bernard (2005). La liaison : de la phonologie à la cognition. *Langages*, 158 : 3-7.

Chevrot, Jean-Pierre; Dugua, Céline; Harnois-Delpiano, Mylène; Siccardi, Anne et Spinelli, Elsa (2013). Liaison acquisition : debates, critical issues, future research. *Language Sciences*, 39 : 83-94.

De Jong, Daan (1994). La sociophonologie de la liaison orléanaise. In C. Lyche (dir.), *French Generative Phonology : Retrospectives and Perspectives*. Salford : ESRI : 95-129.

Delattre, Pierre (1996). *Studies in french and comparative phonetics : selected papers in French and English*. The Hague, London, Paris : Mouton & Co.

Dugua, Céline et Baude, Olivier (2017). La liaison à Orléans, corpus et changement linguistique : une première étude exploratoire. *Journal of French Language Studies*, 27 : 41-54.

Durand, Jacques; Laks, Bernard; Calderone, Basilio et Tchobanov, Atanas (2011). Que savons-nous de la liaison aujourd'hui ? *Langue française*, 169 : 103-135.

Encrevé, Pierre (1988). *La liaison avec et sans enchaînement, phonologie tridimensionnelle et usage du français*. Paris : Éditions Seuil.

Harnois-Delpiano, Mylène; Cavalla, Cristelle et Chevrot, Jean-Pierre (2012). L'acquisition de la liaison en L2 : étude longitudinale chez des apprenants coréens de FLE et comparaison avec enfants francophones natifs 3ème Congrès Mondial de Linguistique Française : 1575 - 1589.

Wauquier, Sophie (2009). Acquisition de la liaison en L1 et L2 : stratégies phonologiques ou lexicales ? *AILE... LIA*, 2 : 93-130.

2.2 Constructions interrogatives

Les constructions interrogatives - ou questions - sont un phénomène à l'interface entre syntaxe, sémantique, pragmatique et prosodie (voir Dayal, 2016). Généralement caractérisées par l'inversion sujet-verbe dans la littérature, les constructions interrogatives en contexte de conversation spontanée conservent souvent un ordre SVO, et sont marquées par une intonation montante, qui signale qu'il ne s'agit pas d'une assertion. Les constructions interrogatives ont pour fonction pragmatique prototypique d'adresser au destinataire une requête d'information. Cependant, les constructions interrogatives en contexte de conversation spontanée peuvent assumer une gamme variée de fonctions pragmatiques. Les interrogatives ont été annotées au fil de l'eau, en parcourant la transcription.

2.2.1 Méthodologie

L'annotation des constructions interrogatives ne prend pas en compte les interrogatives indirectes, car elles sont rares dans le module.

2.2.2 Typologie d'annotation

TABLE 2.2 – Repérage de l'objet

Catégories	Valeurs
objet	interrogative interrogative_suite

TABLE 2.3 – Jeu d'étiquettes adopté

Catégories	Valeurs
objet	interrogative interrogative_suite
portee	alternative partielle totale
forme	verbale averbale
motqu-	ex-situ in-situ 0
inversion	simple complexe sans est-ce-que 0
fonction	phatique pédagogique requête_action requête_confirmation requête_information rhétorique écho

2.2.3 Catégories

2.2.3.1 Portée

L'interrogative est soit totale, soit partielle soit alternative, selon le constituant sur lequel elle porte. Selon le constituant ciblé par l'interrogation (par exemple, la phrase dans son intégralité ou bien le mot interrogatif), l'interrogation n'offre pas à l'interlocuteur le même type de réponses possibles.

Totale L'interrogative ne contient pas de mot interrogatif et porte sur la phrase entière. Le typage interrogatif de la phrase est déterminé par la syntaxe (est-ce que, inversion sujet-verbe) ou par la prosodie, telle qu'une intonation montante. L'interrogation totale est une interrogative qui appelle une réponse en "oui" ou en "non" (dans la littérature anglophone, on parle de "yes-no question" ou "polar question").

CD322 alors elle elle habite Tavers maintenant avant elle était à euh Sandillon
vous connaissez Sandillon ?
 FT715 oui
 CD322 oui c'est pas loin

Partielle L'interrogative contient un mot interrogatif (ou plusieurs). Beaucoup de ces mots interrogatifs commencent par *qu-* (e.g., *quoi, qui, quand*). Ils peuvent être désignés par le label "mot qu-" (correspondant au label anglophone "wh-word", d'où la notion de "wh-question"). L'interrogative partielle appelle une réponse ouverte, en fournissant à l'interlocuteur la possibilité d'identifier la valeur de l'entité correspondant à la prédication contenue dans la question.

- (14) ESLO2_REPAS_1270
 TG634FIE hé *vous faites quoi*
 pour le nouvel an alors vous ?
 TG634 eh ben on le fait avec euh les NPERS et les NPERS

Alternative L'interrogative présente deux interrogatives totales coordonnées par le connecteur disjonctif *ou*. L'interrogation alternative présente plus d'une option au destinataire et appelle à une réponse partielle, destinée à valider l'un des constituants de l'interrogative.

- (15) ESLO2_24H_1250
 GEN_client_24H euh *est-ce que vous avez les gants ou un bonnet qui va avec ou elle est juste vendue comme ça ?*

2.2.3.2 Forme

On distingue les interrogatives verbales des interrogatives averbales.

Verbale L'interrogative est de forme canonique, c'est-à-dire qu'elle contient un prédicat verbal.

- (16) ESLO2_ENT_1013
 ZD13 euh on tombe dans un système très très répressif
 qui nous prive quand même de beaucoup de nos libertés
 ch_PP6 ouais
 tu parles des caméras ?

Averbale L'interrogative est de forme non-canonique (ou "elliptique") : elle ne contient pas de prédicat verbal. Dans ce cas de figure, elle contient un mot interrogatif (averbale partielle), voir (17), ou n'en contient pas (averbale totale), voir (18) (voir Lefeuvre et Rossi-Gensane, 2015).

- (17) ESLO2_REPAS_1270
 TG634FIL1 je vais l'inviter au nouvel an
 TG634 *qui ça ?*
 (18) ESLO2_REPAS_1270
 TG634FIL2 *le machin de douze mètres ?*

2.2.3.3 Position du mot interrogatif

Un mot interrogatif dont la fonction grammaticale est celle d'un complément du verbe peut occuper différentes positions syntaxiques par rapport au prédicat verbal. Cette catégorie n'est donc pertinente qu'en ce qui concerne les interrogatives partielles de forme verbale.

Ex-situ Le mot interrogatif est antéposé au verbe, c'est-à-dire qu'il n'est pas localisé dans la position syntaxique qu'occupe canoniquement le complément du verbe dans une phrase déclarative. Dans le cadre théorique de la grammaire générative, le mot interrogatif est antéposé au verbe via un mouvement de montée en tête de phrase.

- (19) ESLO2_CINE_1191
NX422 donc *quel film* vous venez de voir?

In-situ Le mot interrogatif est postposé au verbe : il est localisé dans la position syntaxique qu'occupe canoniquement du verbe dans une phrase déclarative. La montée en tête de phrase du mot interrogatif n'a pas eu lieu.

- (20) ESLO2_REPAS_1270
TG634FIE Marie part en vacances avec lui avec ses parents et tout
TG634FIL2 frères de cœur
TG634 tu le vois quand pou- parce qu'il est en terminale tu le vois *quand* toi?

0 On attribue l'étiquette 0 pour non pertinent dans 3 cas de figure :

- lorsque l'interrogative ne comporte pas de mot interrogatif
 - dans le cas des interrogatives caractérisées par l'ellipse du prédicat verbal (interrogatives averbales), ce qui implique que le positionnement du mot interrogatif par rapport au prédicat verbal est non pertinent
 - dans certains "contextes catégoriques" identifiés par Coveney (1995) où les interrogatives partielles ex-situ ne peuvent pas être construites autrement, car la variante in-situ correspondante serait agrammaticale. Ces contextes catégoriques sont les suivants :
 - QUEL + copule + NP
- (21) ESLO2_CONF_1241
JGuarrigues alors *quels* étaient mes questionnements dans cette recherche?
- (22) *Alors mes questionnements étaient *quels* dans cette recherche?
- Le mot interrogatif est POURQUOI
- (23) *Pourquoi* elle ne vient pas?
- (24) *Elle ne vient pas *pourquoi*?
- Mot interrogatif + EST-CE QUE
- (25) ESLO2_CINE_1215
CD322FEM si c'est pas indiscret *qu'est-ce que* vous comptez faire plus tard?

FT715 professeur des écoles

(26) *Est-ce que vous comptez faire quoi plus tard ?

2.2.3.4 Inversion

Dans les phrases déclaratives du français contemporain, l'ordre canonique de la phrase est SVO (Sujet-Verbe-Objet). Dans les phrases interrogatives, l'ordre d'apparition du sujet (clitique ou nominal) et du verbe peut être inversé, suivant un ordre VSO (Verbe-Sujet-Objet). Notez que l'inversion du sujet et du verbe est principalement attestée dans les registres les plus formels. D'une manière générale, l'inversion est rare dans l'oral spontané, voir aussi Coveney (2020). Aussi est-elle rare dans les extraits du corpus ESLO-FLEU rattachés au module "D'une situation à l'autre". Notez également l'inversion ou la non-inversion du sujet et du verbe n'est possible qu'à la condition que l'interrogative soit de forme verbale.

Simple En cas d'inversion simple, le sujet clitique est placé après le verbe, selon la structure (Q-)V-Cl.

(27) ESLO2_REPAS_1270

TG634 mon sapin artificiel j'en s'en je suis je crois à cinq ou six hivers
il tient toujours

TG634FIE bah tu m'étonnes en même temps comment *veux-tu* qu'il meurt ?

Complexe En cas d'inversion complexe, le sujet nominal est placé avant le verbe et est redoublé par un sujet clitique placé après le verbe, selon la structure (Q-)GN-V-Cl.

(28) ESLO2_CONF_1240

BGratuze *ce verre est-il* vraiment

à B- euh de Bernard Perrot

ou est-il plutôt euh une production différente ?

Sans Dans les cas où l'inversion est possible mais non réalisée, l'interrogative reçoit l'étiquette "sans inversion".

(29) ESLO2_CINE_1187

NW940 moi tout le monde au taf me dit *tu as vu* Polisse tu as vu Polisse

Est-ce-que L'inversion du sujet et du verbe est impossible lorsque l'interrogative est construite avec la locution "est-ce que", suivi du sujet et du verbe. Il s'agit d'un contexte catégorique.

(30) ESLO2_REPAS_1270

TG634 Antoine *qu'est-ce qu'il a* d'extraordinaire qu'est-ce qu'il fait ?

0 On relève d'autres cas où l'inversion est agrammaticale :

- Le mot interrogatif assume la fonction de sujet grammatical :

(31) ESLO2_24H_1253

NR390AMI *qui* a pris le mot anglais ?

- Mot interrogatif + copule :

- (32) ESLO2_CONF_1241
JGuarrigues alors *quels étaient mes questionnements*
dans cette recherche ?

De plus, les interrogatives averbales ne présentant pas de verbe ne peuvent pas faire l'objet d'une inversion. On peut relever aussi le cas des non questions comme "ça va ?", où l'inversion n'est pas possible.

2.2.3.5 Fonction

Les phrases interrogatives peuvent assumer une diversité de fonctions pragmatiques, que nous ne prétendons pas identifier de manière exhaustive.

Phatique L'interrogative sert à établir un contact avec l'interlocuteur.

- (33) ESLO2_REPAS_1270
TG634FIE c'est Alex
qui voulait aller couper un sapin au Parc Pasteur pour le mettre
dans le dans la déco
de classe
TG634 ça va pas le faire
tous les sapins qui étaient là c'était nos sapins
quand j'étais petite de Noël qu'on replantait tout le temps dans le jardin
TG634FIL1 et euh
TG634 *tu entends ?*

Pédagogique L'interrogative sert à introduire un nouveau thème dans le discours : le locuteur connaît déjà la réponse à sa question, mais l'interlocuteur ne la connaît pas encore. Les interrogatives pédagogiques permettent au locuteur de répondre lui-même à sa question par la suite du discours. Également appelées "interrogatives introductives de thème" (voir Coveney, 2020), les interrogatives pédagogiques sont fréquentes en contexte de prise de parole académique.

- (34) ESLO2_CONF_1241
JGuarrigues alors *quels étaient mes questionnements*
dans cette recherche ?

Requête_action L'interrogative sert à adresser une requête à l'action du destinataire.

- (35) ESLO2_REPAS_1270
TG634 oui *tu peux te lever aller chercher de l'eau j'ai soif s'il te plaît* Marc

Requête_confirmation L'interrogative sert à solliciter la confirmation de l'interlocuteur. Cette fonction pragmatique est associée de manière privilégiée aux questions négatives orientées positivement (dans la littérature anglophone, "high negation questions"), voir aussi Krifka (2017).

- (36) ESLO2_REPAS_1263

GJ845MER *tu voulais pas essayer de trouver la petite veste euh
camouflage avec les petits strass que portait Karine Lemarchand euh dans son
émission ?*

Requête_information L'interrogative sert à adresser au destinataire une requête d'information. C'est la fonction canonique associée aux phrases interrogatives.

- (37) ESLO2_CINE_1215
CD322FEM *si c'est pas indiscret qu'est ce que vous comptez faire plus tard ?*
FT715 professeur des écoles

Rhétorique Les interrogatives rhétoriques sont des questions qui présupposent que la réponse est impliquée par le contexte (voir Biezma et Rawlins, 2017).

- (38) ESLO2_REPAS_1270
TG634 *mon sapin artificiel j'en s'en je suis je crois à cinq ou six hivers
il tient toujours*
TG634FIE *bah tu m'étonnes en même temps comment veux-tu qu'il meurt ?*

Écho L'interrogative écho sert à demander à l'interlocuteur de répéter ce qu'il a dit en cas de mauvaise compréhension.

- (39) ESLO2_24H_1253
NR390 *sinon tu nous fais des doggy bags hein n'hésite pas*
NR390AMIE2 *chaud cacao*
NR390AMI *des quoi ?*

2.2.4 Références

- Beyssade, Claire (2006). La structure de l'information dans les questions : quelques remarques sur la diversité des formes interrogatives en français. *Linx*, 55 : 173-193.
- Biezma, Maria et Rawlins, Kyle (2017). Rhetorical questions : Severing questioning from asking, *Proceedings of SALT 27* : 302-322
- Coveney, Aidan (2020). L'interrogation directe. *Encyclopédie grammaticale du français*.
- Coveney, Aidan (1995). The use of the QU-final interrogative structure in spoken French. *Journal of French Language Studies*, 5 : 143-171.
- Dayal, Veneeta (2016). *Questions*. Oxford University Press.
- Druetta, Ruggero (2021). La structure alternative en français : un type interrogatif à part entière. *Langue française*, 4, 212 : 91-105.
- Krifka, Manfred (2017). Negated polarity questions as denegations of assertions, in L. Chung-min ; K. Ferenc et M. Krifka (dir.), *Contrastiveness in information structure : alternatives and scalar implicatures* : 359-398. Berlin : Springer
- Lefeuvre, Florence et Rossi-Gensane, Nathalie (2015). Interrogation, in P. Larrivée & F. Lefeuvre (dir.) *Projet Fracov*.
- Quillard, Virginie (2001). La diversité des formes interrogatives : comment l'interpréter. *Language et société*, 1(95) : 57-72.

Le parler jeune

Les 20 extraits du module "Le parler jeune" ont été sélectionnés dans ESLO2 (période 2008-2021) selon l'âge des locuteurs, afin d'illustrer les pratiques contemporaines des locuteurs jeunes dans des situations informelles ou entre pairs. Il implique majoritairement des locuteurs de la tranche d'âge des 15/25 ans dans des échanges qui se caractérisent par une interactivité très élevée et une absence de préparation. Ce module, issu d'une diversité de modules d'ESLO2 (24heures, Cinéma, Entretiens, Entretiens jeunes, Repas), contient des conversations familières sur différents sujets (sujets humoristiques, relatifs à la vie quotidienne ou aux études universitaires), collectées en situations informelles (entre autres autour de jeux vidéo, de jeux de plateau ou devant la télévision), ainsi que des interactions en situation d'entretien sur différents sujets (loisirs, sorties, réseaux sociaux, fêtes orléanaises, etc.).

Dans la mesure où les locuteurs jeunes font un usage important des marqueurs discursifs, ce module permet une sensibilisation à ce phénomène, présenté sous l'étiquette des "**mots du discours**", en Section 3.1 en ciblant leurs fonctions pragmatiques dans les interactions spontanées. Le phénomène des **dislocations** (constructions détachées) y est également représenté en Section 3.2. Le groupe de travail responsable de l'élaboration de la typologie d'annotation des mots du discours est constitué de Britta Thörle, Marie Skrovec, Chloé Tahar, Britta Gemmeke et Christian Koch. Le groupe de travail en charge des constructions détachées est composé de Marie Skrovec, Britta Thörle et Chloé Tahar.

3.1 Mots du discours

Ce module est construit pour familiariser les apprenants avec un ensemble d'objets actuellement décrits en linguistique sous l'appellation "marqueurs discursifs". Devant certaines difficultés de catégorisation liées à la nature-même de ces marqueurs (formes polysémiques, ensemble non homogène d'un point de vue syntaxique, etc.), et pour répondre à une nécessaire adaptation terminologique pour un public non expert, le terme "mot du discours" (MD), renvoyant aux *Gesprächswörter* de la tradition germanophone (cf. Koch & Oesterreicher 1990), a semblé pertinent pour désigner une diversité de phénomènes incluant entre autres interjections, connecteurs et marqueurs de l'interaction.

On y dénombre 2125 occurrences pour 48 types de MD. L'annotation a été menée sur 23 MD (981 occurrences) pour spécifier ces formes d'un point de vue sémantico-pragmatique, selon 4 grands ensembles fonctionnels détaillés ci-après.

3.1.1 Méthodologie

Annotation manuelle des mots du discours via un inventaire des formes attestées dans le sous-corpus et une requête ciblée sur chacun d’entre eux, sous forme de séquences même pour les formes simples pour garantir l’homogénéité.

3.1.2 Typologie d’annotation

TABLE 3.1 – Repérage de l’objet

Catégorie	Valeurs
objet	mot du discours
	mot du discours_suite

TABLE 3.2 – Jeu d’étiquettes adopté

Catégories	Valeurs
interaction	connivence
	désaccord
	emphase
	interjection
	ouverture_tour
	ratification
	sollicitation_approbation
	surprise
formulation	approximation
	atténuation
	clôture_liste
	correction
	hésitation
	illustration
	ponctuant
	proposition_formulation
connexion	précision
	reformulation
	additif
	conclusif
	consécutif
information	contrastif
	résultatatif
	activation_topic
	clôture_topic
	développement_topic
	parenthèse_topic
	réactivation_topic
	résolution_topic

3.1.3 Catégories

3.1.3.1 Interaction

Ce domaine fonctionnel concerne les marqueurs de structure de l'interaction et de co-construction des tours de parole (e.g., marqueurs d'ouverture et de fermeture de tours), les marqueurs de gestion intersubjective, qui remplissent des fonctions interactives (e.g., marqueurs d'expressivité ou d'emphase), et les marqueurs de gestion des connaissances partagées¹ (e.g., marqueurs de ratification).

Connivence Les marqueurs de connivence participent à la gestion des connaissances partagées. Leur utilisation sert à indiquer que ce à quoi le locuteur réfère est déjà connu de l'interlocuteur. C'est le cas du marqueur "là" dans l'exemple (1), qui signale que l'émission de télévision que le locuteur a l'intention de désigner est déjà connue de l'interlocuteur.

- (1) ESLO2_ENT_1001
 BV1AMI ouais c'est TF1
 ch_OB1 hé ah ouais ouais ouais non non
 BV1 c'est TF1 ah non c'est quoi sur euh M6 là

Désaccord Les marqueurs de désaccord participent à la gestion intersubjective. Ils sont utilisés pour réfuter ou se mettre en opposition avec ce qu'a dit précédemment un interlocuteur.

- (2) ESLO2_ENT_1001
 BV1AMI ah *mais* ça c'est pas Confessions Intimes c'est Pascal le Grand Frère
 BV1 *mais* non non c'était Super ma- Nanny là

Emphase Les marqueurs d'emphase sont utilisés pour focaliser de manière contrastive un constituant, c'est-à-dire pour exprimer que le locuteur assume que l'interlocuteur ne considèrera pas le contenu du constituant comme normal ou probable (d'après Zimmerman, 2017), mais comme inattendu.

- (3) ESLO2_ENTJEUN_1234
 WT075 là la nouvelle elle a trop une pauvre tête
 WZ853 c'est vrai ?
 la pauvre ça se fait pas de dire ça
 WT075 bah tant pis
 pour elle
 en même temps euh J-
 WZ853 je sais pas
 je l'ai jamais vue
 WT075 Jeanne d'Arc *quoi*
 enfin
 WZ853 ouais
 c'est clair faut en vouloir pour faire Jeanne d'Arc

1. La notion de connaissances partagées (ou *common ground* dans la littérature anglophone, d'après la notion de Stalnaker (1974)) désigne l'ensemble des propositions que les participants de la conversation acceptent mutuellement de tenir pour vraies, c'est à dire l'état informationnel partagé par les participants de la conversation.

Interjection Les interjections participent de la gestion intersubjective. Elles sont utilisées pour exprimer l'attitude du locuteur (état cognitif ou émotionnel) par rapport au contenu de la proposition.

(4) ESLO2_24H_1249

NR390AMIE on a acheté un micro sapin pour la micro petite
NR390 *oh* c'est mignon

Ouverture_tour Les marqueurs d'ouverture de tour participent à la stratégie de gestion de la co-construction des tours de parole. Leur utilisation sert à signaler à l'interlocuteur que le locuteur s'empare du tour de parole.

(5) ESLO2_24H_1253

NR390AMI et donc tu peux prend- piocher des objets si tu veux
ça coûte une minute
NR390AMIE bah allez c'est parti
NR390 mais ça sert à rien là votre euh de
NR390AMIE3 *mais* j'en m- j'en pioche combien

Ratification Les marqueurs de ratification participent à la gestion des connaissances communes aux interlocuteurs. Leur utilisation sert à signaler que le locuteur confirme l'ajout de la proposition aux connaissances communes.

(6) ESLO2_ENTJEUN_1228

WT646 est-ce que tu joues à des jeux vidéo ?
GF429 ouais *en effet* je suis un gros geek aussi

Sollicitation_approbation Les marqueurs de sollicitation de l'approbation participent à la gestion des connaissances communes. Leur utilisation sert à exprimer que le locuteur assume que la proposition est vraie mais demande à l'interlocuteur de confirmer l'ajout de cette proposition aux connaissances communes.

(7) ESLO2_CINE_1187

NX422 et vous avez réussi vos partiels après ?
parce que
nous euh du coup on se dit que si on ne va pas en cours euh
MK912 comme il ne donne pas les cours euh
LG356 franchement on a réussi
après on n'est pas un exemple mais euh
mais tu peux quand même euh franchement on a eu notre licence à euh
NW940 mais oui mais parce que euh
nous c'est bah je sais que c'est toujours la moyenne de tout *en fait* ?
MK912 oui oui c'est ça
NX422 ouais

Surprise Les marqueurs de surprise (ou « miratifs ») participent de la gestion intersubjective. Ils sont utilisés pour exprimer l'attitude de surprise du locuteur, généralement causée par la découverte soudaine de quelque chose de nouveau.

(8) ESLO2_24H_1249

NR390AMIE tu fais quoi alors pendant les vacances là ?

NR390 euh bah
 je fais je fais Noël avec mes parents et
 ma soeur et le lendemain on on se fait un resto avec des amis
 NR390AMIE hm
 NR390 et je pars à Londres
 pour le premier de l'an avec mes potes
 NR390AMIE *mais* oui c'est vrai le jour de l'an

3.1.3.2 Formulation

Ce domaine fonctionnel concerne le travail de formulation au sens large en tant que parole "en train de se faire" et regroupe les procédures de retouche liées à l'immédiateté de la production orale (hésitations, corrections, précisions, reformulations) et de modalisation des contenus discursifs (approximation, atténuation, proposition de formulation, illustration)

Approximation Les marqueurs d'approximation sont utilisés pour signaler que le locuteur hésite quant à une quantification ou à l'identification d'un référent comme appartenant à une certaine catégorie (voir Inkova, 2023).

- (9) ESLO2_CINE_1187
 LG356 et en première année vous étiez combien ?
 CH578 on était ouais quatre-vingt-dix
 MK912 euh quatre-vingt-dix
 quelque chose comme ça
 LG356 ah bah nous
 c'était encore pire parce que en première année on était *genre* cent et quelques

Atténuation Les marqueurs d'atténuation sont utilisés pour mitiger, c'est-à-dire minimiser l'engagement du locuteur vis-à-vis de son assertion ("softening"), en signalant par exemple une possible sur-assertivité. (Beeching 2005 : 159).

- (10) ESLO2_ENTJEUN_1229
 WT646 et donc ça prend une *quand même* une bonne partie de ton temps ?
 WR641 hm oui

Clôture_liste Les clôtureurs (d'après le label de Kahane & Pietrandrea, 2012) de liste ont pour fonction de "signaler l'achèvement d'une énumération ou d'une liste" (voir Béguelin & Corminboeuf, 2017) tout en suggérant que "la liste entreprise ne reflète pas exhaustivement sa référence" (Johnsen, 2017 : 237), invitant ainsi l'interlocuteur à "inférer d'autres membres potentiels" (Johnsen, 2017 : 237).

- (11) ESLO2_ENT_1003
 KC3 j'aime bien les films indiens aussi
 Bollywood *tout ça*

Correction Les marqueurs de correction sont utilisés pour corriger l'énoncé qui précède, c'est-à-dire proposer un énoncé alternatif.

- (12) ESLO2_ENT_1038

AJ38 euh à partir de treize ans j'ai commencé les cours de chant
 ch_PP6 d'accord
 AJ38 *enfin* non douze ans
 douze treize ans ouais je sais plus douze treize ans j'ai commencé les
 cours de chant

Hésitation Les marqueurs d'hésitation sont utilisés dans le travail de gestion de la progression discursive comme des "remplisseurs" (dans la littérature anglophone, des *fillers*), qui servent à signaler l'hésitation.

- (13) ESLO2_ENT_1001
 BV1AMI bah nous on adore
 les reportages les trucs comme ça
 BV1 ouais M6 surtout
 ch_OB1 ouais après M6 c'est quoi c'est Capital *euh*
 c'est quoi le truc de reportage après *euh*
 BV1 Enquête Exclusive

Illustration Les marqueurs d'illustration sont utilisés pour donner un exemple de référent associé à une catégorie.

- (14) ESLO2_ENTJEUN_1234
 WT075 tu vas euh des fois aux aux évènements culturels ?
 genre euh la fête de Jeanne d'Arc
 marché de Noël
 festival de Loire

Ponctuant Les ponctuants sont utilisés dans le travail de gestion de la progression discursive pour la délimitation des unités discursives. Ils n'ont pas de contenu sémantique.

- (15) ESLO2_ENTJEUN_1232
 WT646 est-ce que par exemple tu fais de la musique ?
 on dit souvent sport musique euh est-ce que tu fais de la musique ?
 MM428 bah
 la musique c'est ma vie *en fait*
 WT646 c'est ta vie ?
 MM428 je la
 ouais je la conçois euh
 ma vie sans musique je la conçois pas *en fait*

Proposition_formulation Les marqueurs de proposition de formulation servent à signaler que le locuteur cherche une formulation adéquate et présente l'énoncé comme une formulation possible.

- (16) ESLO2_24H_1251
 NR390collègue2 ils ont eu quoi vingt-et-un vingt-deux et bah là ils sont en
 divorce
 même pas un an

NR390 ouais donc en fait euh
 ça sert à rien de précipiter les choses mais tu vois Laëtitia par contre je
 pense que ça
 NR390collègue2 hm
 NR390 que c'est *voilà* réfléchi sérieux
 et qu'ils sont vraiment bien ensemble

Précision Les marqueurs de précision sont utilisés pour apporter une reformulation explicite de l'énoncé qui précède.

(17) ESLO2_ENT_1038

AJ38 moi je chante pour le plaisir je chante pas euh
 pour euh pour moi c'est pas un sport quoi c'est euh
 c'est un plaisir donc euh moi tu me feras pas chanter une chanson que
 je n'aime pas enfin
 je peux pas
enfin je peux mais ce sera pour moi ce sera médiocre parce que euh
 je serais pas dedans

Reformulation Les marqueurs de reformulation sont utilisés pour apporter une reformulation de l'énoncé précédent lors d'une "activité discursive de paraphrasage" où "deux énoncés sont produits et enchaînés de telle manière qu'ils doivent et peuvent être compris de la même manière" (Gülich et Kotschi, 1982).

(18) ESLO2_ENTJEUN_1235

CC423 j'ai fait Jérusalem
 avant deux mille euh
 avant deux mille un
bref avant les attentats du onze septembre

3.1.3.3 Connexion

Cet ensemble fonctionnel regroupe les connecteurs discursifs, catégorie qui regroupe les conjonctions ainsi que d'autres catégories grammaticales qui servent à exprimer un sens logico-argumentatif (e.g., conséquence, résultat, contraste). On adoptera une définition large de la notion de "connecteurs", comme regroupant l'ensemble des outils argumentatifs qui permettent de mettre en relation des constituants discursifs (voir Rossari, 1999). Cette définition implique que les connecteurs ne mettent pas seulement en relation deux propositions, mais peuvent s'appliquer à des propositions implicites ou inférables dans le contexte.

Additif Les marqueurs additifs connectent deux segments discursifs, sur le mode de la coordination.

(19) ESLO2_ENTJEUN_1232

MM428 euh je joue de la guitare de la batterie euh
 euh je chantais pour une chorale chrétienne
 et je rappais
 euh voilà
 euh *après* euh
 euh j'aime aussi euh composer

Conclusif Les marqueurs conclusifs expriment une relation de conséquence entre des prémisses et une conclusion (voir Rossari et Jayez, 1997).

(20) ESLO2_ENT_1038

AJ38 j’aime bien ces deux bars-là et euh
 Mac Ewans des fois parce que euh la terrasse elle euh
 elle est de passage et en fait des fois quand je suis dans rue de Bour-
 gogne euh
 j’ai beaucoup de personnes euh d’amis qui vont au Mac Ewans
 et ils sont sur la terrasse et moi je passe devant
 et *donc* je me fais engrainer

Consécutif Les marqueurs de consécution expriment une relation de cause à effet, ou de moyen à fin entre deux constituants discursifs (voir Zenone, 1983). Dans l’exemple (21), le connecteur “du coup” permet d’établir une relation de consécution entre une situation (“ça va être le week-end”) qui a certains effets (“va y avoir pas mal de monde”).

(21) ESLO2_ENTJEUN_1234

WT075 tu vas euh des fois aux aux évènements culturels?
 genre euh la fête de Jeanne d’Arc
 marché de Noël
 festival de Loire
 WZ853 ouais
 j’aime bien mais après y a beaucoup de monde
 et euh quand je peux y aller euh bah c-
 ça va être le week-end
du coup va y avoir
 pas mal de monde

Contrastif Les connecteurs contrastifs expriment une relation d’opposition entre les deux segments discursifs qu’ils connectent entre eux.

(22) ESLO2_ENT_1001

ch_OB1 Confessions Inti- non mais Confessions Intimes c’est fascinant moi
 ça me fascine
 quand je regarde euh je regarde jusqu’au bout hein
 BV1 ouais moi ça ça cartonne
 je tolère pas du tout les gens qu’il y a dedans *mais*
 ça me fait rire

Résultatif Les connecteurs résultatifs expriment que le segment discursif qu’ils introduisent est un résultat, par exemple, en (23), le résultat d’une action non-langagière.

(23) ESLO2_24H_1253

NR390 mais pourquoi moi j’ai pas eu le droit de prendre un objet?
 NR390AMIE3 parce que euh
 NR390AMI parce que c’est là les objets
 NR390AMIE3 je prends ça
 hop

et ça c'est
 une
 un extincteur
 enfin il me semble
 NR390 comme ça
 NR390AMIE3 ils feraient pas des cartes aussi petites ouais c'est ça
voilà je me suis grillé (...)
 deux minutes

3.1.3.4 Organisation informationnelle

Les marqueurs d'organisation informationnelle sont des marqueurs qui servent à la structuration des topics conversationnels dans la structure de la conversation.

Activation_topic Les marqueurs d'activation d'un nouveau topic servent à introduire un nouveau topic.

- (24) ESLO2_24H_1249
 NR390AMIE tu fais quoi *alors* pendant les vacances là ?
 NR390 euh bah
 je fais je fais Noël avec mes parents et
 ma soeur et le lendemain on on se fait un resto avec des amis

Clôture_topic Les marqueurs de clôture de topic servent à indiquer la fin d'un topic.

- (25) ESLO2_ENTJEUN_1232
 MM428 ma vie sans musique je la conçois pas en fait
 c'est euh
 c'est depuis petit [= !ptit /instantaneous/N] c'est mon dada
 c'est euh
 cette crème qu'on m'a mis à la naissance
 euh je joue de la guitare de la batterie
 je chantais pour une chorale chrétienne
 et je rappais
 euh *voilà*
 euh
 après euh
 euh j'aime aussi euh composer

Développement_topic Les marqueurs de développement topical servent à introduire le développement d'un topic.

- (26) ESLO2_ENT_1038
 AJ38 j'aime bien ces deux bars-là et euh
 Mac Ewans des fois parce que euh la terrasse elle euh
 elle est de passage et *en fait* des fois quand je suis dans rue de Bourgogne
 euh
 j'ai beaucoup de personnes euh d'amis qui vont au Mac Ewans
 et ils sont sur la terrasse et moi je passe devant
 et donc je me fais engrainer

Parenthèse_topic Les marqueurs de parenthèse topicale servent à signaler un décrochement du topic conversationnel, pour apporter par exemple une explication que le locuteur juge nécessaire pour la compréhension du topic en cours.

- (27) ESLO2_24H_1253
 NR390AMIE3 donc là qu'est-ce que je fais déjà ?
 ah je je fais une action
 bah je peux pas rester où je suis je suis
 sur la place du capitaine ah je peux sortir
 NR390AMI alors
 euh ouais
 NR390 bah oui c'est le but non ?
 NR390AMI donc tu peux prendre
 NR390AMIE mon scaphandre
 NR390AMI euh des objets
 dans euh *en fait*
 là y a des objets
 tu les caches là

Réactivation_topic Les marqueurs de réactivation de topic servent à réactiver un topic qui a été introduit dans le discours qui précède et momentanément suspendu.

- (28) ESLO2_ENT_1047
 ch_MP10 je ne les connais pas du
 tout les boîtes de nuit d'Orléans
 LD47 ah j'en ai fait qu'une seule pour le moment depuis que je suis à Orléans
 par contre
 ça j'en ai fait qu'une seule c'est le
 LD47 c'est le gardel's
 donc qui en plus qui est pas loin de mon boulot donc c'est
 bah vous avez deux salles vous avez une salle plutôt pour les jeunes
 puis une salle un peu pl- pour les personnes un peu plus âgées donc
 euh c'est une boîte
 ouais c'est sympathique c'est pas
 ch_MP10 c'est quel g- quel genre de boîte en fait ? c-
 c'est y a des thèmes
 c'est des soirées thématiques ou des des
 LD47 euh dans la boîte euh jeune ça va être plutôt de la alors là
 en général euh je me trompe souvent sur le le nom des styles de musique
 mais
 c'est de la de la trance de la techno
 du truc assez rythmé avec beaucoup de basses quoi
 et dans la salle euh donc pour les personnes un peu plus âgées
 en général c'est une trentaine d'années
 euh ça va être plus de la musique euh contemporaine quoi
 ce qu'on entend à la radio sur Vibration sur Forum sur euh

Résolution_topic Certains marqueurs ont pour fonction d'indiquer la résolution d'un topic. Cette fonction est identifiée par Pelletier & Ranson (2020 : 8) pour le marqueur "bon", qui peut servir à "marquer la progression du discours" en introduisant un "résultat". Cette opération correspond à la "résolution du thème en cours".

- (29) ESLO2_CINE_1187
 LG356 en plus non les cours qui sont intéressants
 genre tout ça tu y vas
 mais alors nous
 NW940 c'était boycott direct
 NX422 ah ouais ?
 LG356 c'était boycott
 on se rejoignait devant la & fac bon
 on va faire les magasins
 NW940 c'était huit
 heures huit heures je me rappelle
 c'était huit heures neuf heures le vendredi

3.1.4 Références

- Abouda, Lotfi et Skrovec, Marie, (2016), Du mouvement au figement : pragmatification de la forme on va dire. Étude micro-diachronique sur un corpus oral, *Language Design*, Special Issue 2016, 121-145.
- Aijmer Karin, Simon-Vandenberghe Anne-Marie, (2009), Pragmatic markers, in J.-O. Östman et J. Verschueren (éds), *Handbook of Pragmatics*, Amsterdam : Benjamins, 1-29.
- Auchlin, Antoine, (1981), Réflexions sur les marqueurs de structuration de la conversation, *Études de linguistique appliquée*, 44 : 88-103.
- Béguelin, Marie-José et Corminboeuf, Gilles (2017), *Ou comme ça, machin et autres marqueurs d'indétermination dans les listes*, Discours 20
- Dostie Gaétane, (2004), Pragmatification et marqueurs discursifs. *Analyse sémantique et traitement lexicographique*, Bruxelles, De Boeck-Duculot.
- Dostie, Gaétane, Pusch, Claus D., (2007), Présentation. Les marqueurs discursifs. Sens et variation, *Langue française*, 154 : 3-12.
- Fisher, Karin (éd.), 2006, *Approaches to discourse particles*, Studies in Pragmatics, Elsevier.
- Fraser, Bruce, 1999, What are discourse markers? *Journal of pragmatics*, 31 : 931-952.
- Gülich, Elisabeth et Kotschi, Thomas (1982) Les marqueurs de la reformulation paraphrastique. *Cahiers de Linguistique française*, 5. Genève : Univ. de Genève : 305-351.
- Inkova, Olga (2023). Approximation et structures sémantiques apparentées. *Langages*, 229
- Johnsen, Laure Anne, (2017). *L'expression de la sous-détermination dans les procédés de référence en français contemporain*. Thèse de doctorat. Université de Neuchâtel.
- Kahane, Sylvain et Pietrandrea, Paola. (2012). La typologie des entassements en français. In F. Neveu, V. Muni Toke, P. Blumenthal, T. Klingler, P. Ligas, S. Prévost et S. Teston-Bonnard (éd.), *Actes du 3e congrès mondial de Linguistique française – CMLF 2012* Vol. 1 : 1809-1828.
- Koch, Peter and Oesterreicher, Wulf. (1990). *Gesprochene Sprache in der Romania Französisch, Italienisch, Spanisch*. 31, Romanistische Arbeitshefte
- Maurer, Bruno, (1998), Utilisation des ponctuations et apprentissage de la compétence de communication. In : M. Souchon (éd.) *Actes du Xe colloque international : Pratiques discursives et acquisition des langues étrangères*, Besançon 1996, Presses universitaires franc-comptoises, 183-192.

Mosegaard-Hansen, Maj-Britt, (1996), Some common discourse particles in spoken French. In : M.-B. M. Hansen, G. Skytte (éds), *Le discours : cohérence et connexion*, Copenhagen, Museum Tusculanum Press : 105-149.

Pfander, Stefan et Skrovec, Marie (2010) : "Donc, entre grammaire et discours. Pour une reprise de la recherche des universaux de la langue parlée à partir de nouveaux corpus", in : Drescher, Martina & Neumann-Holzschuh, Ingrid (eds.) : *Syntaxe de l'oral dans les variétés non-hexagonales du français*. Section 5 du Congrès des Franco-romanistes, Augsburg, sept. 2009, 183-196.

Rossari, Corinne (1999) Les relations de discours avec ou sans connecteurs. *Cahiers de Linguistique française* 21 : 181-192

Rossari, Corinne (1990) Projet pour une typologie des opérations de reformulation. *Cahiers de Linguistique française*, Université de Genève, 11 : 345-359

Rossari, Corinne et Jayez, Jacques (1997) Connecteurs de conséquence et portée sémantique. *Cahiers de Linguistique française*, Université de Genève, 19 : 233-265.

Repp, Sophie (2013). Common Ground Management : Modal Particles, Illocutionary Negation and Verum. In Daniel Gutzmann & Hans-Martin Gartner (eds.), *Beyond Expressives : Explorations in Use-Conditional Meaning*. Brill

Schiffrin, Deborah, 1987, *Discourse markers*. Cambridge University Press.

Skrovec, Marie ; Kanaan-Caillol, Loyal et Akihiro, Hisae, 2022, Le marqueur après à l'oral : une approche micro-diachronique, variationniste et interactionnelle. *Langages*, 117-131.

Stalnaker, Robert (1974). Pragmatic Presuppositions. In *Context and Content*. Oxford University Press. pp. 47-62.

Thörle, Britta, 2016, Turn openings in L2 French : An interactional approach to discourse marker acquisition. In : Borreguero Zuloaga, Margarita, Thörle, Britta (éds.) : *Discourse Markers in Second Language Acquisition. Studies on Italian and French as L2*. Special issue of *Language, Interaction and Acquisition* 7 :1, 117-144

Thörle, Britta, 2019, La particule additive aussi dans les détachements à gauche en français parlé. In : *Bulletin de la Société Néophilologique* II/CXX, 365-387.

Thörle, Britta, 2020, "Et le chien aussi i' regarde" : Particule additive aussi et structure informationnelle en français L2, *Language, Interaction and Acquisition*, Vol. 11-2, 298-324.

Zimmermann, Malte (2008). Contrastive focus and emphasis. *Acta Linguistica Hungarica*, Vol : 55, Issue 3-4. 347-360

Zenone, Anna (1982). La consécution sans contradiction : donc, par conséquent, alors, ainsi, aussi (première partie). *Cahiers de Linguistique française* 4, Genève, 107-141.

3.2 Constructions disloquées

Les constructions dites "disloquées" (ou "détachées") sont des constructions syntaxiques qui se caractérisent par "la présence d'un syntagme à la périphérie d'une structure phrastique, au sein de laquelle celui-ci conserve généralement une trace sous une forme pronominale" (Raickovic & Skrovec 2020 : 1). Ce syntagme adjoint à la construction verbale se situe généralement dans une position non-canonique par rapport à l'ordre SVO et peut entretenir une relation de co-référentialité plus ou moins exacte avec l'élément de reprise pronominale au sein du noyau prédicatif. Dans la continuité des travaux de Blasco-Dulbecco (2000), les configurations suivantes ont été annotées comme "dislocations" :

- un argument du verbe est exprimé deux fois, par un constituant détaché en périphérie et par un pronom défini qui le représente dans la phrase,
- le syntagme détaché est repris par un pronom clitique indéfini comme "ça" ou "ce",
- le syntagme est considéré comme détaché dans la mesure où il se situe dans une position non-canonique mais il n'y a pas de reprise pronominale. Dans ce dernier cas, l'élément disloqué peut être ou non un argument du verbe.

3.2.1 Typologie d'annotation

TABLE 3.3 – Repérage de l'objet

Catégorie	Valeurs
dispositif-syntaxique	dislocation dislocation_suite

TABLE 3.4 – Jeu d'étiquettes adopté

Catégories	Valeurs
dislocation	droite gauche
constituant-detache	SN SV SAdv pronom_tonique
reprise	ce-ça il-y-a pronom topique_suspendu

3.2.2 Catégories

3.2.2.1 Dislocation

Le site d'accueil du constituant détaché peut être à gauche, en début de phrase ou à droite, en fin de phrase.

Droite

- (30) ESLO2_24H_1253
NR390AMI elles sont là les *les vodkas*

Gauche

- (31) ESLO2_24H_1251
NR390 *la nana du Pareil Au Même* elle est mariée divorcée elle a vingt-huit ans

3.2.2.2 Constituant détaché

Le constituant détaché n'est pas forcément un syntagme nominal ou un pronom tonique. Il peut également s'agir d'un syntagme verbal, d'une proposition ou d'un syntagme adverbial.

SN

- (32) ESLO2_ENTJEUN_1232
MM428 *la musique* c'est ma vie en fait

SV

- (33) ESLO2_ENT_1003
KC3 *écrire euh des histoires*
et cetera ça me plaît bien

SAdv

- (34) ESLO2_ENT_1038
AJ38 *où j'étais euh* c'est pas comme le conservatoire

Pronom-tonique

- (35) ESLO2_24H_1253
NR390 mais pourquoi *moi* j'ai pas eu le droit de prendre un objet ?

3.2.2.3 Reprise

Le constituant disloqué peut être représenté dans la phrase par différents moyens grammaticaux (ce-ça, il y a, pronom). On parle de "topique suspendu" (dans la littérature anglophone, "hanging topic") en cas d'absence de reprise pronominale.

Ce-ça

- (36) ESLO2_ENT_1001
ch_OB1 y a plein de séries avec des trucs d'amour quand même
BV1 ah ouais les Frères Scott
mais ça c'est dans le lit avec le PC portable

Il-y-a

- (37) ESLO2_ENT_1047
LD47 bah le bowling bah de toute façon à
je crois que si je me trompe pas à Orléans *y en a* deux

Pronom

- (38) ESLO2_ENTJEUN_1234
WZ853 la statue de Jeanne d'Arc *elle* est éclairée en bleu

Topique-suspendu

- (39) ESLO2_ENT_1001
 ch_OB1 donc ce type d'émissions-là à la télé les infos les trucs comme ça euh
 non?
 BV1 ouais si *info* j'aime bien regarder ouais
- (40) ESLO2_24H_1249
 NR390 si c'est pour jouer aux petits soldats *moi* c'est non

3.2.3 Références

Berrendonner, Alain, (2021), Constructions disloquées, in *Encyclopédie Grammaticale du Français*

Blasco-Dulbecco, Mylène, (2000), *Les dislocations en français contemporain, Etudes syntaxiques*, Honoré Champion : Paris.

De Fornel, Michel, 1988, Constructions disloquées, mouvement thématique et organisation préférentielle dans la conversation, *Langue française*, 78, 101-123.

Horlacher, Anne-Sylvie et Müller, Gabriele (2005) : L'implication de la dislocation à droite dans l'organisation interactionnelle, *Travaux neuchâtelois de linguistique*, 2005, 41, 127-145

Horvath, Marton, 2018, *Le français parlé informel. Stratégies de topicalisation*. Amsterdam : John Benjamins.

Grutschus, Anke et Schwab, Sandra, 2020, Le redoublement du sujet en français parlé : motivations rythmiques. In : *Studia linguistica romanica* 4, 129-149.

Pekarek Doehler, Simona, (2001), Dislocation à gauche et organisation interactionnelle. *Marges linguistiques*, 2, 177-194.

Pekarek Doehler, Simona ; De Stefani, Elwys et Horlacher, Anne-Sylvie, (2015), *Time and Emergence in Grammar. Dislocation, topicalization and hanging topic in French talk-in-interaction*. Amsterdam : John Benjamins.

Pekarek Doehler, Simona et Stoenica, Ioana-Maria (2012) : Émergence, temporalité et grammaire-dans-l'interaction : disloquée à gauche et nominativus pendens en français contemporain, *Langue française*, 111-127.

Raickovic, Luka et Skrovec, Marie 2020, Syntaxe en interaction et variation diaphasique : l'exemple des dislocations dans ESLO2, SHS Web Conf., 78 (2020) 14004.

Récit en interaction

Les 13 extraits de ce module ont été sélectionnés dans les corpus ESLO1 et ESLO2, de manière à fournir une collection de récits non préparés et élaborés au fil de l'interaction. Ces récits sont généralement produits spontanément par le locuteur témoin en situation d'entretien ou en situation de repas. Plusieurs extraits issus de ce module portent sur le récit d'une expérience vécue à la première personne, telles que des expériences difficiles ("trouble telling", Jefferson, 1988), des expériences professionnelles ou relevant de la vie quotidienne qui sortent de l'ordinaire, c'est-à-dire qui transgressent une norme implicite concernant le déroulement normal des choses. Ces récits peuvent être destinés à exprimer une plainte, à faire rire l'interlocuteur, ou tout simplement à retracer la chronologie d'événements personnels. D'autres extraits portent sur le récit d'expériences vécues par un tiers, dont le locuteur raconte l'histoire (parfois dans le but de solliciter l'empathie); ou encore de récits à partir d'un support externe, comme les récits de films ("story retelling", Norrick, 1998). Ce module permet de sensibiliser au phénomène du récit en interaction, tel que décrit par la linguistique interactionnelle (Gulich et Quasthoff, 1986; Vincent et Bres, 2001; Ochs et Capps, 2002) et donne l'occasion de cibler plus spécifiquement le **discours rapporté** (Morel, 1996), en Section 4.1, et certains dispositifs syntaxiques spécifiques comme les **constructions clivées** (Rouquier, 2018), en Section 4.2. Le groupe de travail responsable de l'élaboration de la typologie d'annotation du discours rapporté est constitué de Anke Grutschus, Marie Skrovec, Chloé Tahar, Christian Koch et Britta Thorle. Celui en charge des constructions clivées, de Marie Skrovec, Britta Thorle et Chloé Tahar.

4.1 Discours rapporté

Avec Rendulic et Abouda, on identifie le discours rapporté comme relevant de la "représentation d'une interaction autre", plus précisément comme la reconstitution d'une "interaction qui a eu lieu dans le passé", ou la représentation d'une interaction "présentée comme imaginée" (Rendulic & Abouda 2015, Abouda & Rendulic 2017).

Le discours indirect étant un phénomène marginal dans le module, l'objet annoté est restreint au discours direct, défini par Authier-Revuz comme le phénomène où "l'énonciateur rapporte un autre acte d'énonciation *e*, en faisant usage de ses mots à lui dans la description qu'il fait de la situation d'énonciation de *e* (...) c'est-à-dire dans ce qu'on appelle le syntagme introducteur, mais il fait mention des mots du message qu'il rapporte" (Authier-Revuz, 1992 : 40).

On cherche ainsi à repérer les séquences de discours rapporté (au discours direct) ainsi que les segments introducteurs, les "quotatifs". Pour ces derniers, on observe deux configurations : (a) présence de verbes introducteurs, (b) présence d'autres segments introducteurs (Cf. discours direct libre, Authier-Revuz, 1992 : 41).

4.1.1 Typologie d'annotation

Une fois la séquence de discours rapporté identifiée, on identifie comme une séquence différente le quotatif.

TABLE 4.1 – Repérage de l'objet

Catégorie	Valeurs
objet	discours-rapporté discours-rapporté_suite

TABLE 4.2 – Jeu d'étiquettes adopté

Catégories	Valeurs
quotatif	Vdire Vdire_suite Vfaire Vfiction autre autre_suite

4.1.2 Quotatif

Parmi les quotatifs, on relève plusieurs types de verbes : des verbes de dire courants (ou *verba dicendi*, définis comme «introduceurs d'une activité de parole, dont "dire", Anscombre, 2015 : 104), des verbes de fiction, et d'autres verbes introduceurs d'activité de parole. On ne trouve en revanche aucune occurrence de verbes de pensée avec fonction de quotatif.

Vdire verbes du dire comme *dire, demander*, les plus fréquents dans le module.

- (1) ESLO1_ENT_010
JI306 ce monsieur cuisinier me vient me voir
et puis *me dit* euh tiens Madame euh X a téléphoné

Vfaire verbe *faire* comme introduceur d'activité de parole.

- (2) ESLO2_REPAS_1265
DC040 j'ai regardé ma pote *je lui fais* ben je veux pas monter sur scène hein

Vfiction verbes de fiction comme introduceur d'activité de parole (un seul cas dans le module).

- (3) ESLO2_REPAS_1265
MH568 juste pour le faire chier j'ai dit à Charlotte j'ai trop envie de dire au
vigile qu'y en a un qui fume
faire ma grosse balance mais
PF437 *tu imagines*
monsieur

autre On observe deux cas de figure :

- Verbe ou syntagme nominal sémantiquement proche d'un verbe du dire, déclencheur de parole.
- (4) ESLO2_ENT_1041
BC41 *les auditeurs appelaient beaucoup* mais il est génial c'est qui machin tout ça
- Absence de verbe d'activité de parole mais présence d'éléments contextuels qui construisent la représentation d'un cadre interactionnel ("setting interactionnel") et introduisent une séquence de "représentation d'une interaction autre".
- (5) ESLO2_ENT_1041
BC41 *je rencontre un couple* et euh
voilà est -ce que vous allez bien machin tout ça impeccable
il fait beau aujourd'hui alors moi j'aimerais savoir euh
comment vous vous êtes rencontrés et tout
- (6) ESLO1_INTPERS_403
403PERS *je vois des parents qui viennent*
euh voilà j' ai mon fils qui a vingt-deux ans n' est -ce pas ?
il est en deuxième année de faculté
alors euh je voudrais savoir si

4.1.3 Références

- Abouda, Lotfi et Rendulic, Nina (2017), Séquence d'introduction de discours représenté : faire ou dire ?, *Discours*, 21.
- Anscombre, Jean-Claude (2015). Verbes d'activité de parole, verbes de parole et verbes de dire : des catégories linguistiques ?, *Langue française*, 186, 103-122.
- Anderson, Hanne Leth (2002), Le choix entre discours direct et discours indirect en français parlé : facteurs syntaxiques (et pragmatiques). *Faits de Langues*, 19.
- Authier-Revuz, Jacqueline, (1992), Repères dans le champ du discours rapporté, *Information grammaticale*, n° 55 : 8-42.
- Authier-Revuz, Jacqueline, (2004), La représentation du discours autre : un champ multiplement hétérogène. In : J.-M. Lopez-Munoz, S. Marnette, L. Rosier (éds), *Le discours rapporté dans tous ses états : question de frontières*, Paris, L'Harmattan : 35-53.
- Diwersy, Sascha et Grutschus, Anke, (2014), Écoute(z) en tant que marqueurs de discours rapporté. In : Weidenbusch, Waltraud (éd.) : *Diskursmarker, Konnektoren, Modalwörter*. Tübingen : Narr, 55-67.
- Fauré, Laurent et Verine, Bertrand, (2004), Authentifier un discours autre en y mettant du sien : les vocalisations ah et oh en frontière de discours rapporté direct à l'oral. In : Juan López Muñoz, Sophie Marnette, Laurence Rosier (éds). *Le discours rapporté dans tous ses états*. Paris : L'Harmattan, 317-327.
- Gülich, Elizabeth and Quasthoff, Uta (1986). Story-telling in Conversation. Cognitive and Interactive Aspects, *Poetics*, 15 : 217-241.
- Jefferson, Gail, (1988). On the Sequential Organization of Troubles-Talk in Ordinary Conversation, *Social Problems*, 35 : 418-441.
- Norrick, Neal, (1998). Retelling Stories in Spontaneous Conversation, *Discourse Processes*, 25 : 75-97

Morel, Mary-Annick, (1996), *Le discours rapporté direct dans l'oral spontané*, *Cahiers du français contemporain* 3 :77-89.

Ngamounskika, Édouard, (2014), *Le discours rapporté dans l'oral spontané. L'exemple du français parlé en République du Congo*. Bordeaux : Presses Universitaires de Bordeaux.

Ochs, Elinor and Capps, Lisa. (2002). *Living narrative : creating lives in everyday storytelling*, Harvard University Press.

Rendulic, Nina, (2015), *Le discours représenté dans les interactions orales. De l'étude des structures en contexte vers la construction de l'image des relations interlocutives*. Thèse de doctorat, Linguistique. Université d'Orléans.

Rendulic, Nina et Abouda, Lotfi. (2013) *La mise en scène dans la représentation des interactions autres. Aux marges du discours. Personnes, temps lieux, objets*, *Actes du Xe Congrès International de Linguistique Française*, pp.287-298.

Rosier, Laurence, (2008), *Le Discours rapporté*, Paris, Ophrys.

Verine, Bertrand, (2005), *Comment les interjections vocaliques modalisent l'insertion argumentative des discours rapportés directs non-véridiques à l'oral*. In : Juan Lopez Munoz, Sophie Marnette, Laurence Rosier (éds), *Dans la jungle des discours : genres de discours et discours rapporté*, Cadiz : Servicio de Publicaciones de la Universidad, 497-506.

Vincent, Diane et Dubois, Sylvie, (1997), *Le Discours rapporté au quotidien*, Québec, Nuit blanche.

4.2 Constructions clivées

Les constructions clivées sont des constructions bi-phrastiques dont le premier segment correspond à un constituant focalisé, généralement le sujet, tandis que la partie prédicative est introduite sous la forme d'une subordonnée de type relative (également connue sous le nom de "pseudo-relative"). La clivée se repère parce qu'au niveau sémantique, un ordre SVO y est substituable (même si la fonction pragmatique diffère).

Dans les clivées prototypiques, le constituant Sujet est focalisé au moyen d'une construction en "c'est" - "C'est X + QU + Verbe".

On choisit ici de considérer également comme constructions clivées les constructions où le sujet est focalisé au moyen d'une construction en "il y a", dans la lignée de Lambrecht (2001), Verwimp & Lahousse (2016) et De Cesare (2017).

Notez que dans les clivées en "c'est", l'élément clivé peut appartenir à de nombreuses les catégories grammaticales ; ce n'est pas nécessairement un syntagme nominal.

Il est établi que certaines occurrences de la forme "C'est X + QU + Verbe" ne doivent pas être confondues avec les clivées : contrairement aux clivées, elles ne peuvent pas être ramenées à une structure SVO équivalente, et le pronom "ce" y est référentiel (voir De Cesare, 2017 : 538 pour des exemples).

4.2.1 Méthodologie

On identifie au fil de l'eau toutes les occurrences de dispositifs syntaxiques de forme "c'est/il y a /j'ai X + QU + Verbe", en parcourant la transcription. Puis, on annote le segment pertinent en spécifiant son type, en se basant sur les catégories définies dans la littérature (notamment De Cesare, 2017).

4.2.2 Typologie d'annotation

TABLE 4.3 – Repérage de l'objet

Catégories	Valeurs
dispositif_syntaxique	clivage
	clivage_suite

TABLE 4.4 – Jeu d'étiquettes adopté

Catégories	Valeurs
clivage	clivée_focus
	clivée_présentative
	clivée_interrogative
	clivée_circonstancielle
	pseudo_clivée
	0

4.2.3 Clivage

clivée_focus pour les cas où l'élément clivé identifie la valeur d'une variable définie par la clause pseudo-relative. Par exemple, en (7), il y a une valeur indéterminée pour une variable ("quelqu'un était malade") que spécifie l'élément clivé ("c'était Mathieu"), à l'exclusion d'autres. L'élément clivé joue le rôle informationnel de focus (pointage de l'information nouvelle parmi plusieurs options).

- (7) ESLO2_ENT_1041
- ch_LA11 donc le le le collègue que tu avais remplacé parce qu'il était
malade ah oui d'accord hm hm
- BC41 bah *c'était Mathieu en fait qui était malade* du coup

- (8) ESLO2_ENT_1041
- BC41 bon là il a il a il est en arrêt euh maladie euh
parce qu'il ne peut plus
il a été arrêté par euh et c'est pas lui hein
c'est euh la la la médecine du travail qui l'arrête

clivée_présentative pour les cas où l'élément clivé présente une nouvelle entité ou un nouvel événement, sans identifier la valeur d'une variable (voir Verwimp & Lahousse, 2016).

- (9) ESLO2_ENT_1029

ch_NS3 c'était plusieurs fois par an on sortait du collège on allait dans un vieux cinéma
 et là y a quelqu'un qui venait y a l'animateur du ciné-club qui disait aujourd'hui euh
 euh je vais vous passer le euh hm une histoire de baleine et puis il nous passait Moby Dick

clivée_interrogative pour les cas où l'élément clivé est un constituant interrogatif.

- (10) ESLO2_ENT_1329
 ch_OB1 et alors *c'est qui qui qui gueulait le plus ?*

clivée_circonstancielle pour les cas où l'élément clivé appartient à la catégorie des syntagmes adverbiaux à signification circonstancielle.

- (11) ESLO1_ENT_010
 JI306 alors j'ai donc ce monsieur allait chez ces gens-là et j'ai donc été chez ces gens-là plusieurs fois
 servir des déjeuners et des soupers
c'est donc là que j'ai appris que la jeune fille se mariait

pseudo_clivée pour les cas où l'élément clivé (et sa subordonnée) est le sujet du verbe "être" ("ce qui pousse les gens à se déchirer est que ...").

- (12) ESLO2_ENT_1041
 BC41 j'ai l'impression qu'on prend pas le temps de connaître et ce qui est *c'est ce qui pousse les gens à se détester à se déchirer à*

0 on observe un segment introducteur "c'est" suivi d'une relative mais il ne s'agit pas d'un cas de clivage. La subordonnée est une relative restrictive, nécessaire à l'identification du référent indéfini : l'ordre SVO n'est pas substituable, le pronom "ce" est référentiel (co-référent de "mon père").

- (13) ESLO2_ENT_1041
 BC41 il pouvait pas renvoyer euh mon père parce que c'est ce qu'on appelle un bon soldat
c'est quelqu'un qui n'a jamais été malade

4.2.4 Références

Apotheloz, Denis et Roubaud, Marie-Noëlle, (2015), Constructions pseudo-clivées. In : *Encyclopédie Grammaticale du Français*.

Apothéloz, Denis, (2012), Pseudo-clivées et constructions apparentées. In : Groupe de Fribourg, *Grammaire de la période*. Berne : Peter Lang, 207-232.

De Cesare, Anna-Maria (2017) Clefts constructions, in : Dufter, Andreas and Stark, Elisabeth. *Manual of Romance Morphosyntax and Syntax*, Berlin, Boston : De Gruyter, 2017

De Stefani, Elwis, (2008), De la malléabilité des structures syntaxiques dans l'interaction orale : le cas des constructions clivées. In : Durand J. Habert B., Laks B. (éds.) : *Discours, pragmatique et interaction*, Congrès Mondial de Linguistique Française 2008, 703-720.

Lambrecht, Knud (2001), A framework for the analysis of cleft constructions, *Linguistics* 39, 463–516.

Lambrecht, Knud (2004), Un système pour l'analyse de la structure informationnelle des phrases. L'exemple des constructions clivées, in : Jocelyne Fernandez-Vest/Shirley Carter-Thomas (edd.), *Structure informationnelle et particules énonciatives*, Paris, L'Harmattan, 22–61.

Roubaud, Marie-Noëlle, (2000), *Les constructions pseudo-clivées en français contemporain*, Honoré Champion : Paris.

Roubaud, Marie-Noëlle et Sabio, Frédéric, (2015), Les clivées en c'est là où et c'est là que : structure et usages en français moderne. *Repères- Dorif : autour du français : langues, cultures et plurilinguisme*, 6.

Rouquier, Magalie, (2018), Les constructions clivées, in *Encyclopédie grammaticale du français*.

Verwimp, Lyan et Lahousse, Karen, (2016), Definite Il y a-clefts in spoken French. In : *Journal of French Language Studies* 27/3 : 1-28.

Mémo des requêtes

5.1 Principes généraux des requêtes sur TXM

Les requêtes données ci-après sont données à titre d'exemple et peuvent être modifiées selon vos besoins.

Requête par module Pour obtenir tous les extraits d'un module (e.g., D'une situation à l'autre) :

```
[_.modulefleu_desc="D'une situation à l'autre"]
expand to modulefleu
```

Requête par objet linguistique Pour obtenir toutes les occurrences d'un objet linguistique (e.g., interrogatives). Notez que l'expression régulière `.*` permet d'obtenir n'importe quelle suite de caractères, elle permet de faire une requête qui récupère à la fois les objets ayant la valeur `interrogative` et ceux ayant la valeur `interrogative_suite` dans l'exemple suivant :

```
[_.objet_ref="interrogative.*"] expand to objet
```

Pour mémoire, les constructions clivées et disloquées sont annotées comme des dispositifs syntaxiques :

```
[_.dispositif-syntaxique_ref="clivage.*"]
expand to dispositif-syntaxique
```

```
[_.dispositif-syntaxique_ref="dislocation.*"]
expand to dispositif-syntaxique
```

Requête par catégorie-valeur Pour obtenir toutes les occurrences d'un objet linguistique hors liaison relevant d'une certaine catégorie et ayant une certaine valeur (e.g., tous les mots du discours dont la fonction informationnelle est celle d'activer un nouveau topic) :

```
[_.objet_ref="mot du discours.*"& _.information_ref=
"activation_topic"] expand to objet
```

Ou :

```
[_.information_ref="activation_topic"] expand to information
```


Si vous souhaitez obtenir dans le concordancier l'unité de segmentation dans son intégralité (expand to u) :

```
[_.objet_ref="mot du discours.*"& _.information_ref="activation_topic"] expand to u
```

Requête par catégorie-valeur Pour obtenir toutes les occurrences de l'objet linguistique liaison relevant d'une certaine catégorie et ayant une certaine valeur :

```
[realisation_liaison="N"]
```

Pour connaître tous les couples catégorie/valeur du corpus les méthodes sous TXM ne sont pas forcément intuitives, le plus simple est donc de vous référer aux tableaux de ce document.