

UNIVERSITÉ DE LILLE

École Doctorale Sciences de l'Homme et de la Société

UMR 8163 CNRS - Savoirs, Textes et Langage

THÈSE présentée par :

Sangwan JEON

soutenue le : **9 janvier 2025**

pour obtenir le grade de :

Docteur en Sciences du langage

**Le discours numérique sur l'immigration en France
entre 2011 et 2022. Une analyse de corpus (Vol. 2)**

**Guide d'annotation et définition opérationnelle des
catégories utilisées**

THÈSE dirigée par :

Mme PIETRANDREA Paola

Professeure, Université de Lille

RAPPORTEUSES :

Mme ESHKOL-TARAVELLA Iris

Professeure, Université Paris Nanterre

Mme GLIKMAN Julie

Professeure, Université de Lorraine

JURY :

M. BERGOUNIOUX Gabriel

Professeur, Université d'Orléans, *président*

Mme CALIENDO Giuditta

Maîtresse de conférences, Université de Lille, *examinatrice*

Mme ESHKOL-TARAVELLA Iris

Professeure, Université Paris Nanterre, *rapporteuse*

Mme GLIKMAN Julie

Professeure, Université de Lorraine, *rapporteuse*

Mme PIETRANDREA Paola

Professeure, Université de Lille, *directrice*

M. WAGENER Albin

Professeur, Université Catholique de Lille, *examineur*

RESUME

Le *Guide d'annotation et définition opérationnelle des catégories utilisées*, second volume de la thèse *Le discours numérique sur l'immigration en France entre 2011 et 2022. Une analyse de corpus*, propose une méthodologie pour l'annotation du corpus diachronique FrMigrTwit, composé de tweets politiques. Ce guide se concentre sur le traitement du **lexique-migr**, regroupant les termes dérivés de la racine latine migr- (p. ex., migrant, immigration, migratoire) et leurs variantes, en structurant l'annotation autour d'unités spécifiques : **MIGR-LEXICON**, pour les termes du lexique-migr ; **MODIFIER** pour leurs modificateurs (adjectifs, déterminants, propositions subordonnées relatives) ; et **NOUN** ; pour les noms modifiés par ces lexèmes. Chaque unité est enrichie de propriétés lexicales, syntaxiques, sémantiques, ainsi que de paramètres relationnels. Le guide explicite les étapes d'annotation dans Analec, notamment la gestion des relations entre unités et l'intégration des métadonnées (date, comptes utilisateurs, retweets). Enfin, cette annotation met en lumière les dynamiques syntaxiques et sémantiques du lexique-migr, permettant une analyse comparative et diachronique du corpus de tweets.

Mots-clés | discours numérique, immigration, corpus Twitter, annotation linguistique, Analec, rôles sémantiques, fonctions syntaxiques, collocations

ABSTRACT

The *Annotation Guidelines and Operational Definitions of the Categories Used*, the second volume of the thesis *Online Discourse on Immigration in France between 2011 and 2022. A Corpus Analysis*, presents a methodology for annotating the diachronic corpus FrMigrTwit, composed of political tweets. This guide focuses on processing the migr-lexicon, encompassing terms derived from the Latin root migr- (e.g., migrant, immigration, migratory) and their variants, structuring the annotation around specific units: MIGR-LEXICON, for migr-lexicon; MODIFIER, for their modifiers (adjectives, determiners, relative subordinate clauses); and NOUN, for nouns modified by these lexemes. Each unit is enriched with lexical, syntactic, semantic properties, as well as relational parameters. The guide details the annotation workflow in Analec, particularly the management of relations between units and the integration of metadata (date, user accounts, retweets). This annotation highlights the syntactic and semantic dynamics of the migr-lexicon, enabling a comparative and diachronic analysis of the tweet corpus.

Keywords | digital discourse, immigration, Twitter corpus, linguistic annotation, Analec, semantic roles, syntactic functions, collocations

Table des matières

Résumé	ii
Abstract	iii
introduction	5
1 L’annotation du lexique-migr	5
1.1 Lexique-migr	5
1.1 Unité d’annotation : MIGR-LEXICON	6
1.1.1. Nom-migr.....	6
1.1.2. Adjectif-migr	8
1.1.3. Hashtag-migr ou Arobase-migr.....	8
1.2 Propriété LEMMA : Lemme du lexique-migr.....	9
1.2.1. Variabilité des lemmes et formes de surface	10
1.2.2. Cas spécifiques liés à Twitter	10
1.3 Propriété MODIFICATION : Paramètres de qualification	12
1.4 Propriété LIST/PAR : Paramètres de construction de liste ou de parallélisme.....	13
1.5 Propriété ARG_VERB : Paramètres d’argument du verbe.....	13
1.6 Propriété FUNC_SYN : Fonction syntaxique du lexique-migr	14
1.7 Propriété ROLE_SEM : Rôle sémantique du lexique-migr	18
1.8 Propriété USERNAME : Nom du compte d’utilisateur Twitter.....	27
1.9 Propriété DATE : Date de mise en ligne	27
1.10 Propriété RT : Paramètres de retweetage.....	27
2 L’annotation des collocations du lexique-migr	28
2.1 Liste et parallélisme	28
1.1 Unité d’annotation : LIST/PARALLELISM	28
1.1.1. Forme de surface de la liste/parallélisme : propriété LIST/PARALLELISM	29
1.1.2. Type de constructions : propriété TYPE	29
1.2 Unité d’annotation : CONJUNCT	30
1.2.1. Propriété #forme# : Forme de surface du conjunct	31
1.2.2. Propriété LEMMA_CONJ : Lemme du conjunct.....	31
1.2.3. Propriété FUNC_SYN_CONJ : Fonction syntaxique du conjunct.....	31
1.2.4. Propriété ROLE_SEM_CONJ : Rôle sémantique du conjunct	32
1.2.5. Propriété POSITION_CONJ : Position du conjunct.....	33
1.3 Mise en relation des unités d’annotation pour l’analyse des liste et parallélisme	34
1.3.1. Mise en relation de la liste/parallélisme au lexique-migr : REL-MIGR-LEX.....	34
1.3.2. Mise en relation des conjuncts au lexique-migr : SCH-MIGR-CONJUNCT et REL-MIGR-CONJ35	
2.2 Modification	38
2.1 Unité d’annotation : MODIFIER	38
2.1.1. Forme de surface du modifieur du lexique-migr : propriété #forme#	38
2.1.2. Propriété LEMMA_MODIF : Lemme du modifieur du lexique-migr.....	39
2.2 Unité d’annotation : NOUN.....	40
2.2.1. Forme de surface du nom modifié par le lexique-migr : propriété #forme#	40
2.2.2. Lemme du nom modifié par le lexique-migr : propriété LEMMA_NOUN	40
2.3 Mise en relation des unités d’annotation pour l’analyse des modifieurs du lexique-migr	41
2.3.1. Mise en relation du lexique-migr au MODIFIER : Relation MIGR-MODIF1/2	41
2.3.2. Mise en relation du lexique-migr au NOUN : Relation MIGR-MODIF1	42
3 Synthèse de la procédure de l’annotation	44
Références bibliographiques	47

INTRODUCTION

Ce guide est structuré en trois sections principales.

La première section détaille les principes et les méthodes d’annotation du lexique-migr. La deuxième section est consacrée à l’annotation des collocations, incluant les modifieurs et les constructions de listes ou de parallélismes. Enfin, la troisième section propose une synthèse de la procédure d’annotation.

1 L’annotation du lexique-migr

Cette section présente les principes et les méthodes d’annotation du lexique-migr. L’objectif est de fournir un cadre précis pour identifier et annoter ses occurrences dans le corpus diachronique de tweets, FrMigrTwit. L’annotation concerne les formes nominales, adjectivales, ainsi que les suites de caractères contenant la racine latine *migr-*, délimitées par des espaces.

1.1 Lexique-migr

Le **lexique-migr** désigne l’ensemble des mots dérivés de la racine latine *migr-* tels que *migrant*, *immigré*, *immigration*, *migration*, *migratoire*, ainsi que leurs variantes et formes dérivées, y compris les hashtags.

L’unité d’annotation du lexique-migr dans Analec est désignée en terminologie anglaise¹ : MIGR-LEXICON. Ce type inclut dix propriétés du lexique-migr couvrant des aspects formels,

¹ Pour la structure d’annotation, j’ai préféré conserver les termes d’origine plutôt que de les traduire en français, s’agissant des étiquettes de la fonction syntaxique et du rôle sémantique empruntés à Jurafsky et Martin (2021), et à Lacheret et ses collègues (2014) respectivement. Et, pour rester cohérent en termes de code métalinguistique dans l’ensemble de la structure, toutes les étiquettes sont en terminologie anglaise.

linguistiques (lexical, syntaxique et sémantique), paramétriques et contextuels (c.-à-d. les métadonnées).

Tableau 1 Propriétés annotées de l'unité MIGR-LEXICON

Type	Propriété	Description
Formel et lexical	Forme de surface (#forme#)	Texte tel qu'il apparaît dans le corpus.
	Lemme (LEMMA)	Entrée lexicale correspondant à la forme canonique de l'unité annotée.
Syntaxique	Fonction syntaxique (FUNC_SYN)	Indique la fonction syntaxique de l'unité annotée.
Sémantique	Rôle sémantique (ROLE_SEM)	Indique le rôle sémantique de l'unité annotée.
Paramétrique	Modification (MODIFICATION)	Indique si le lexique-migr est modifié ou modifiant.
	Liste/parallélisme (LIST/PAR)	Indique si le lexique-migr se trouve dans une construction de liste ou de parallélisme.
	Argument du verbe (ARG_VERB)	Indique si le lexique-migr est un argument d'un verbe au premier plan de l'énoncé.
Métadonnées	Nom du compte (USERNAME)	Nom du twitto politique.
	Date de (re)tweetage (DATE)	Date de publication du (re)tweet.
	Paramètre de retweetage (RT)	Indique si le tweet en un retweet.

1.1 Unité d'annotation : MIGR-LEXICON

L'annotation de la forme de surface du lexique-migr consiste à sélectionner, dans Analec, une séquence contenant la racine migr- délimitée par des espaces. Trois cas de figure sont observés dans ce contexte.

1.1.1 Nom-migr

Lorsque le lexique-migr constitue une unité nominale, le déterminant du nom-migr est inclus dans la portée annotée, comme indiqué dans l'exemple ci-dessous, représenté entre crochets

[{déterminants défini, indéfini, démonstratif ou possessif} + (Adj) + nom-migr]_{MIGR-LEXICON}

La forme de surface de l'élément annoté est présentée en gras italique dans les exemples suivants :

ex (1) [...] qui accueille, nourrit oriente ***les migrants*** [...]

ex (2) Face aux hausses d'impôts, à la baisse du pouvoir d'achat des Français, à ***une immigration*** qui continue, à une insécurité croissante [...]

ex (3) Plus de repas pour ***les #migrants*** et une petite prime pour les forces de l'ordre, [...]

ex (4) Il n'y a rien dans ***ce #PJLAsileImmigration*** sur les relations que nous devons avoir [...]

ex (5) [...] "aucun port français" ne doit accueillir l'Aquarius et ***ses migrants*** ?

ex (6) ***La #LoiAsileImmigration*** malgré les paroles fermes affichées par @gerardcollomb ne va pas répondre aux attentes des Français [...]

Si le nom-migr est imbriqué dans un syntagme prépositionnel, et que son déterminant défini au masculin fusionne avec la préposition *à* ou *de*, la fusion de la préposition et du déterminant défini est annotée avec le nom-migr. Voici quelques exemples illustrant cette situation :

ex (7) l'accueil des migrants

ex (8) la chasse aux migrants

ex (9) [...] la distribution de repas aux #migrants [...]

1.1.2. *Adjectif-migr*

Lorsque le lexique-migr a pour fonction de qualifier un nom, seule la forme adjectivale du lexique-migr est annotée.

ex (10) Nous sommes confrontés à une submersion **migratoire** incontrôlée et ininterrompue. [...]

ex (11) La loi **immigration** sera une petite loi administrative, [...]

Dans les exemples ci-dessous, le lexique-migr apparaît sous forme de mots composés. Si les éléments constitutifs sont reliés par un symbole conventionnellement utilisé, comme un tiret « - » ou une barre oblique « / », et qu'il n'y a pas d'espace entre ces éléments, l'ensemble est annoté comme un seul mot composé, comme illustré dans l'ex (12) et l'ex (13).

ex (12) [...] Ce matin, dans L'Humanité, ma contribution à un débat sur la loi **Asile/Immigration** : [...]

ex (13) Mon intervention lors du débat sur la loi **Asile-Immigration** @AssembléeNat [...]

En revanche, si les éléments constitutifs du mot composé sont délimités par des espaces, ils sont annotés séparément.

ex (14) Ce dimanche, mobilisation de notre Groupe @Republicains_An dans le débat sur pjl #Asile #**Immigration**.

1.1.3. *Hashtag-migr ou Arobase-migr*

Les séquences précédées d'un hashtag (#) ou d'une arobase (@), appelées respectivement hashtags et mentions, lorsqu'elles remplissent une fonction nominale ou adjectivale, elles sont

annotées selon les règles établies pour les noms et adjectifs du lexique-migr (cf. ex (3), ex (4), ex (6), ex (9) et ex (14)).

Cependant, si ces séquences ne participent pas directement à la formation de la phrase, comme lorsqu'elles se trouvent en position initiale ou finale d'un tweet, seule la suite de caractères précédée du symbole (# ou @) est annotée. Par exemple :

ex (15) 968819768563257344; #*Migrants*. Ce matin, dans L'Humanité, ma contribution à un débat sur la loi Asile/Immigration: "Le migrant en ennemi de l'interieur" @EELV @ImmigrationEELV @EELVIdF @senateursCRCE @pamannoni @CedricHerrou @SophieBorner @gerardcollomb @Aurelientache <https://t.co/Gtlpr4nl5g> ; 2018-02-28T12:06:47.000Z; @EstherBenbassa; 18; 0; 21; 0

1.2 Propriété LEMMA : Lemme du lexique-migr

Le lemme du lexique-migr (qu'il s'agisse de noms, d'adjectifs, hashtags ou mentions avec arobase) correspond à l'entrée lexicale canonique, telle que défini en lexicographie. Cela signifie que le lemme est systématiquement présenté sous sa forme masculine et singulière. Par convention, les lemmes de notre schéma sont toujours écrits en minuscule.

Exemples de lemmes du lexique-migr :

- anti-migrant
- asile-immigration
- immigrant
- immigration
- immigrationniste
- immigré
- migrant
- migration

- migratoire
- pro-immigration
- pro-migrant

1.2.1. Variabilité des lemmes et formes de surface

Les formes de surface du lexique-migr peuvent varier selon :

- La casse : majuscules ou minuscules.
- Les symboles typographiques : tirets (« - ») ou barres obliques (« / »).
- Les pratiques d'écriture en ligne : utilisation de symboles hypertexte tels que le croisillon (« # ») ou l'arobase (« @ »).

Exemple : Toutes les formes suivantes sont associées au même lemme canonique : **asile-immigration.**

- la loi #Asile/#immigration
- la loi #Asile/Immigration
- la loi #Asile/#immigration

1.2.2. Cas spécifiques liés à Twitter

La pratique technolangière sur Twitter introduit des particularités : Lorsque les mots sont soudés après un croisillon (« # ») ou une arobase (« @ ») sans espaces intermédiaires, ces combinaisons forment un unique lemme. Par défaut, tous les caractères sont en minuscule, sauf pour les noms propres.

Exemples de lemmes intégrant des symboles Twitter :

Hashtags :

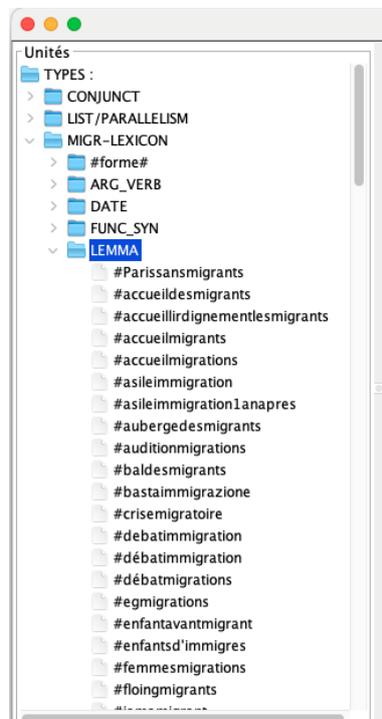
- #accueilmigrants
- #asileetimmigration
- #loiasileimmigration
- #pouruneautrepolitiquemigratoire
- #tousmigrants

Mentions :

- @aubergemigrants
- @immigrationeelv
- @infomigrants

La Figure 1 montre un extrait de la liste des lemmes du lexique-migr, comprenant des termes standardisés en minuscules, tels que *migrant* et *immigration*, ainsi que des hashtags comme *#asileimmigration*.

Figure 1 Visualisation de la liste des lemmes du lexique-migr dans la structure d'annotation



1.3 Propriété MODIFICATION : Paramètres de qualification

Cette propriété évalue si le lexique-migr est impliqué dans une relation de modification.

Trois valeurs possibles sont attribuées :

- MODIFIED : Le lexique-migr est modifié par un qualificatif (par exemple, un adjectif ou une proposition subordonnée relative).
- MODIFYING : Le lexique-migr modifie un autre nom.
- NO : Le lexique-migr n'est ni modifié ni modifiant.

Exemple :

- 1492842459382042627; Nous assistons à *une immigration* de peuplement que les Français n'ont pas voulue. Je leur rendrai la parole sur ce sujet fondamental, en les consultant par referendum sur l'arrêt de l'immigration. #MarinePrésidente \n \n ???? <https://t.co/G74OTZDHjV> <https://t.co/rjpIf9TP2D>; 2022-02-13T12:45:52.000Z; @MLP_officiel; 229; 67; 600; 10
→ **MODIFIED** (*immigration* est modifié par « de peuplement » et par « que les français n'ont pas voulue »)
- 1492966509928398855; ?? Si ne voulons nous préserver des attaques islamistes, avoir une influence sur les flux *migratoires*, il faut travailler avec la Syrie et la Libye. #Cdanslair #Elysée2022 <https://t.co/L5crJCKgcV>; 2022-02-13T20:58:48.000Z; @MLP_officiel; 220; 46; 472; 10
→ **MODIFYING** (*migratoire* qualifie « flux »)
- 1493297043485478921; Pour nos dirigeants, *l'immigration* n'est pas un problème : c'est un projet, qui se poursuit malgré l'opposition majoritaire des Français. \n \n Dès mon élection, je lancerai le grand chantier de l'arrêt de *l'immigration* ???? <https://t.co/G74OTZDHjV> <https://t.co/on4znTIwag>; 2022-02-14T18:52:13.000Z; @MLP_officiel; 270; 92; 683; 15
→ **NO**

1.4 Propriété LIST/PAR : Paramètres de construction de liste ou de parallélisme

Cette propriété identifie si le lexique-migr figure dans une construction de liste ou de parallélisme. Les valeurs attribuées sont :

- LIST : Le lexique-migr est un membre constitutif (c.-à-d. *conjunct*) d'une liste.
- PARALLELISM : Le lexique-migr est un membre constitutif (c.-à-d. *conjunct*) d'un parallélisme.
- NO : Le lexique-migr ne fait pas partie d'une liste ou d'un parallélisme.

Exemple :

- 180395595868028928; *Je propose une brigade spécialisée dans la lutte contre les filières clandestines. {#Immigration | #Hollande2012 | #DPDA}; 2012-03-15T20:50:36.000Z; @fhollande; 22; 11; 3; 0 → LIST*
- 180611091641024513; *Gang des postiches 2/2: Sarkozy prétend défendre la laïcité et les Français pour mieux distiller {l'islamophobie et le rejet des immigrés}; 2012-03-16T11:06:54.000Z; @yjadot; 1; 0; 0; 0 → PARALLELISM*
- 1040652935426924544; *.@NicolasBay_ : "#Salvini obtient déjà des résultats très importants dans la lutte contre l'immigration clandestine." @franceinfo; 2018-09-14T17:26:08.000Z; @RNational_off; 20; 2; 32; 0 → NO*

1.5 Propriété ARG_VERB : Paramètres d'argument du verbe

Cette propriété évalue le degré d'actualisation du lexique-migr dans une structure verbale.

Deux valeurs principales sont attribuées : YES ou NO.

YES : Le lexique-migr est l'argument principal ou secondaire d'un verbe.

- Niveau principal : Le lexique-migr est un complément direct ou indirect du verbe principal de l'énoncé.

ex (16) Nous accueillons beaucoup trop de *#migrants* en France

- Niveau secondaire : Le lexique est un argument dans une proposition subordonnées ou sur le plan secondaire dans une proposition introduite par un complémenteur comme dans les exemples ci-dessous.

ex (17) La première humanité consiste [à dire *aux #migrants* qui veulent venir en Europe] [...]

ex (18) [...] plutôt que d'organiser le blocage et la réquisition de bâtiments [pour y installer *des migrants*,] [...]

NO : Le lexique-migr n'est pas un argument du verbe, ou il est imbriqué dans une structure complexe (par exemple, un syntagme nominal).

ex (19) Sur [le sujet *migratoire*] il faut avoir de la détermination, [...]

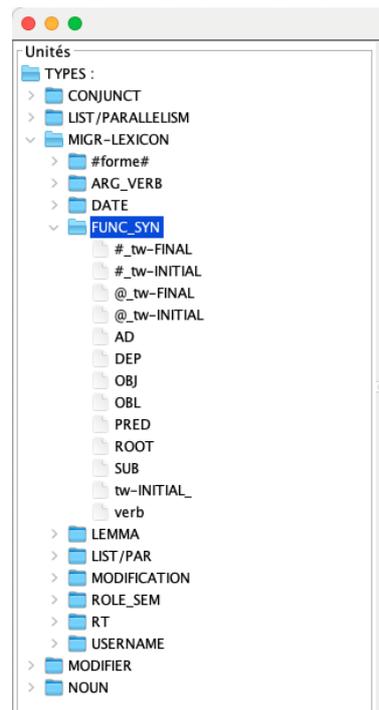
ex (20) La première humanité c'est d'en finir avec [cette folle politique d'accueil de *migrants*]. [...]

ex (21) [La submersion programmée de nos pays par *les #migrants*] est une atteinte à notre intégrité. [...]

1.6 Propriété FUNC_SYN : Fonction syntaxique du lexique-migr

La propriété **FUNC_SYN** identifie la fonction syntaxique occupée par le lexique-migr dans la structure de l'énoncé. Elle permet de catégoriser les relations entre les unités annotées en fonction de leur fonction syntaxique. Les catégories syntaxiques sont empruntées au projet Rhapsodie (Lacheret et al., 2014) et adaptées pour annoter les occurrences du lexique-migr dans un contexte discursif. La Figure 2 visualise la liste des fonctions syntaxiques dans Analec.

Figure 2 Visualisation de la liste des fonctions syntaxiques dans la structure d'annotation



Présentation des fonctions syntaxiques

Le tableau suivant recense les fonctions syntaxiques attribuées au lexique-migr. Ces fonctions permettent de décrire précisément les fonctions syntaxiques que le lexique-migr peut occuper au sein des énoncés. Chaque fonction est illustrée par des exemples extraits du corpus.

Tableau 2 Inventaire des fonctions syntaxiques

Abréviations et définitions des fonctions syntaxiques du projet Rhapsodie² (Lacheret et al. 2014)

1. ROOT • « unité gouvernée par aucune autre »

ex (22) « #Macron a promis l'ouverture du regroupement familial aux migrants « mineurs » !
#Migrants qui, au passage, se déclarent tous « mineurs », même s'ils ont une canne de
 vieillard ! » #LeCrest #TournéeMLP <https://t.co/MutaresADH>; 2018-02-04T13:59:23.000Z;
 @MLP_officiel

ex (23) 84 **#migrants** installés à l'ancien hôtel F1 d'Appoigny, un mur vient d'être construit
 par l'hôtel voisin. Ils veulent supprimer nos frontières mais construisent des murs dans notre
 pays pour protéger les nôtres. \n Demandons le retour des frontières
 nationales!#ExpulserCestProtéger

2. SUB • « sujet grammatical »

ex (24) Il suffit qu'**un migrant** de 45 ans se déclare mineur, pour qu'il soit inexpulsable de
 notre pays... [...]

3. OBJ • « complément d'objet direct »

ex (25) Il faut arrêter toute **immigration**. On est complet !

4. OBL • « complément actanciel oblique, incluant les compléments d'objet indirect »

ex (26) L'Europe a besoin de votre fermeté pour la protection des frontières, pour lutter contre
l'immigration illégale, contre le fondamentalisme islamiste, [...]

5. AD • « adjectif au verbe »

ex (27) À **@AubergeMigrants** pour soutenir leur travail auprès des migrants. [...]

² Sauf la catégorie JUNC qui sert à annoter la « jonction entre les conjoints et le joncteur, en suivant l'analyse asymétrique de la coordination de Mel'čuk (1988) ». Puisque le lexique-migr fait partie des termes à contenu conceptuel, cette étiquette qui concerne les termes à contenu procédural (conjonctions de coordination en particulier) ne sera pas incluse dans la liste des fonctions syntaxiques de notre schéma d'annotation.

6. **PRED** • « tous les éléments qui forment un prédicat complexe avec un verbe (adjectif attribut, forme verbale complexe...) »

ex (28) [...] La vraie vie, c'est *l'immigration* qui n'est pas maîtrisée. \ [...]

ex (29) Il y a 250 millions de *migrants* climatiques potentiels dans les années à venir, voilà l'enjeu. [...]

7. **DEP** • « tous les dépendants de formes non verbales »

- Les syntagmes nominaux introduits par le signe de ponctuation deux-points « : » sont dépendants de la phrase précédente.

ex (30) "Mon ambition est de reconquérir les terrains que nous avons abandonnés aux censeurs du politiquement correct : la nation, *l'immigration*, les frontières, les classes moyennes, l'école, l'écologie, la culture, le social" @laurentwauquiez #CNLR

ex (31) "Nous nous opposons à la politique d'Emmanuel #Macron, qui organise le grand déclassé des classes moyennes et populaires françaises, au bénéfice des gagnants : ses amis, et *les migrants* !" #ConfFN

- Complément prépositionnel :

ex (32) La première humanité c'est d'en finir avec cette folle politique d'accueil de *migrants*. [...]

- Lexique-migr ayant pour fonction de modifieur

ex (33) [...] Nous sommes actuellement dans une situation de véritable submersion *migratoire*.

Positions syntaxiques liées à la structure formelle et sémiotique du tweet
--

1. **#_tw-FINAL** • Hashtag-migr en position finale de tweet
2. **#_tw-INITIAL** • Hashtag-migr en position initiale de tweet

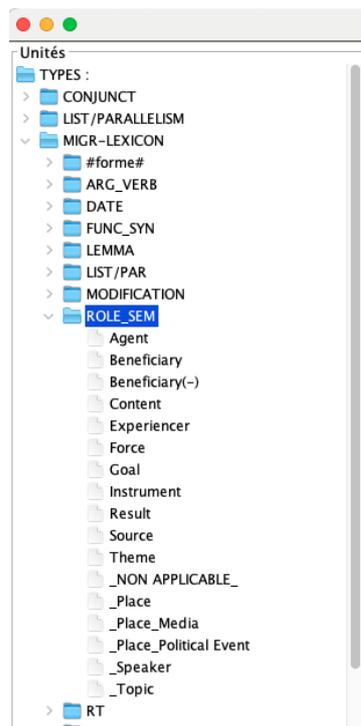
3. **@_tw-FINAL** • Arobase-migr en position finale de tweet
4. **@_tw-INITIAL** • Arobase-migr en position initiale de tweet
5. **tw-INITIAL** • Lexique-migr en position initiale de tweet

1.7 Propriété ROLE_SEM : Rôle sémantique du lexique-migr

Une partie des caractéristiques attribuées à la catégorie des migrants et au phénomène de l’(im)migration peut être analysée grâce à l’annotation des rôles sémantiques associés au lexique-migr. Le choix des catégories de rôles sémantiques repose sur des objectifs spécifiques au domaine de recherche et varie selon les approches. Dans le cadre de notre annotation, nous avons retenu dix rôles principaux : Agent, Expérienceur, Force, Thème, Résultat, Contenu, Instrument, Bénéficiaire, Source et Destination. Les définitions, présentées dans le tableau ci-dessous, sont adaptées et réajustées à partir des travaux de Jurafsky et Martin (2021), Tellier (2004) et Saeed (2016).

Cas particulier : Lorsque le lexique-migr est intégré dans des structures syntaxiques complexes, telles que des syntagmes nominaux [N1 de N2], [N1 à N2], [N1 pour N2], [N1 contre N2], ou [N1 entre N2 et N3], l’annotation repose sur l’analyse détaillée de la micro-structure syntaxique et de la relation sémantique argumentale entre N1 et N2. Cette approche s’appuie notamment sur les principes établis par Bartning (1996).

Figure 3 Visualisation de la liste des rôles sémantiques dans la structure d'annotation



Présentation des rôles sémantiques

Le tableau suivant propose un inventaire des rôles sémantiques utilisés pour annoter le lexique-migr dans le corpus. Ces rôles permettent de décrire les relations sémantiques associées aux entités et phénomènes mentionnés dans les tweets.

Tableau 3 Inventaire des rôles sémantiques

Liste de 10 rôles sémantiques Jurafsky & Martin 2021 (2 nd Edition)
--

- 1. Agent** • Entité [+animé], instigateur volontaire qui est à l'origine du procès ou de l'action.

The waiter spilled the soup. (Jurafsky & Martin 2021)

- ex (34) Il suffit qu'*un migrant* de 45 ans se déclare mineur, [...]
- ex (35) Pourquoi *les #migrants* qui fuient la guerre d'Irak et de Syrie viendraient-ils en Europe ? [...]
- ex (36) [...] les hyper-riches se gavent et *les migrants* ont la belle vie ! [...]
- ex (37) [...] Mais favoriser le passage d'*un #migrant* est un délit. [...]
- ex (38) [...] Ma réaction à la profanation de la basilique de #SaintDenis par *migrants* et extrême-gauche !
- ex (39) [...] Pour ceux qui utilisent ce genre de terme, l'arrivée *des migrants* vient sonner le tocsin
- ex (40) [...] on a vu des scènes de guerre civile et d'affrontements entre *migrants* armés. [...]
- ex (41) *L'immigration* continue : il y a une hausse des titres de séjour de 14%, une augmentation du [...]

- 2. Expériencer** • Entité [+animé] « qui ressent un sentiment, une émotion, ou qui se trouve dans un état psychologique » (Tellier 2004).

John has a headache. (Jurafsky & Martin 2021)

- ex (42) La passivité de l'Europe en matière d'immigration ne fait qu'entretenir les filières de passeurs qui font le malheur *des migrants*.

ex (43) @dupontaignan à raison de dire que le véritable humanisme est de financer un grand plan d'investissement dans le 1/3 monde pour combattre les causes et non subir les conséquences de la tragédie *migratoire* !

ex (44) [...] C'est douloureux pour tout le monde : les français et *les migrants* »

- 3. Force**
- « Entité involontaire qui est à l'origine de l'action. » (The non-volitional causer of the event [Ma traduction]) (Jurafsky & Martin 2021)
- The wind* blows debris from the mall into our yards. (ibid.)
- L'action ou le résultat de cette action causé par l'entité peuvent ne pas être explicités dans la phrase comme dans l'ex (45).

ex (45) Notre pays a besoin de frontières pour se protéger contre *l'immigration* massive et incontrôlée, l'Union européenne le refuse ! [...]

ex (46) Les gens n'en peuvent plus de l'insécurité, d'*une immigration* incontrôlée, d'une mondialisation qui tue notre agriculture et nos petites entreprises. [...]

ex (47) flux de N2, flots de N2, afflux de N2³

ex (48) [...] 50 djihadistes tunisiens seraient arrivés en Italie dans le flux de *#migrants*. [...]

ex (49) On subit des flux de *migrants* clandestins ininterrompus.

ex (50) [...] C'est nous qui avons tort de ne pas refuser ces afflux de *#migrants* !

ex (51) [...] Depuis les premiers mois d'Emmanuel Macron on assiste à une explosion de *l'immigration* en France. [...]

- 4. Theme**
- Entité [± animé], « qui subit l'action ou qui est l'objet de l'action » (Tellier 2004)

³ Le concept principal des N flot, flux, afflux est le mouvement d'un *liquide*. On élimine le rôle sémantique d'agent du fait de l'absence de volonté. Remarquons la métaphore à la catastrophe naturelle par analogie à l'inondation. L'inondation est définie selon le TLFi : « Submersion des terres par les eaux en crue d'un cours d'eau, d'un lac ou par celles de la mer à la suite d'un raz de marée ou de la rupture de digues ».

- De ce fait, le sujet et l'attribut du sujet du verbe d'existence *être* seront annotés comme Thème.

ex (52) [...] la répartition *des #migrants*

ex (53) [...] il faut couper les pompes aspirantes de *l'immigration*

5. Result • « Le résultat ultime du procès » (The end product of an event [Ma traduction]) (Jurafsky & Martin 2021)

The city built a *regulation-size baseball diamond*... (ibid.)

ex (54) Le niveau de l'immigration dans notre pays est spectaculaire : 410 000 *immigrés* légaux par an, un nombre de clandestins que l'on ne veut pas connaître, 50 000 "mineurs isolés", 100 000 demandeurs d'asile..." #QuestionsPol

6. Content • « La proposition ou le contenu du procès exprimé par les verbes de parole » (The proposition or content of a propositional event [Ma traduction]) (Jurafsky & Martin 2021)

Mona asked "*You met Mary Ann at a supermarket?*" (ibid.)

- La proposition ou le message du procès implicite impliqué dans l'acte même de tweeter.

ex (55) 1038849644619620353; RT @GNationVar: ?? @GNation_off se mobilise contre l'arrivée de 72 migrants à #Châteaudouble, petit village de 477 habitants ! \n \n Malgré l'agitation de 5 gauchistes déjà prêts à accueillir ces clandestins, les habitants, eux, ne veulent pas de cette colonisation ! \n \n #StopMigrants <https://t.co/ah8AmEK7CU>; 2018-09-09T18:00:30.000Z; @RNational_off; 460; 0; 0; 0

→ Glose : Je tweete que

[@GNation_off se mobilise contre l'arrivée de 72 migrants à #Châteaudouble, petit village de 477 habitants ! \n \n Malgré l'agitation de 5 gauchistes déjà prêts à accueillir ces clandestins, les habitants, eux, ne veulent pas de cette colonisation !]

et par là, je dis

#StopMigrants

7. **Instrument** • « Instrument utilisé dans le procès ou dans l'accomplissement du procès »
(An instrument used in an event [Ma traduction]) (Jurafsky & Martin 2021)
He poached catfish, stunning them *with a shocking device*... (ibid.)

ex (56) Attirer la jeunesse d'Afrique avec le miroir aux alouettes de *l'immigration* en France ne sert pas ces pays.

ex (57) [...] Nous avons un problème démographique que nous ne voulons pas résoudre par *l'immigration* mais par des politiques familiales. » Viktor Orban

ex (58) Les « nouvelles ruralités" : ce lobby qui veut repeupler nos campagnes avec *des migrants* » [...]

ex (59) La submersion programmée de nos pays par *les #migrants* est une atteinte à notre intégrité. [...]

8. **Beneficiary** • « Le bénéficiaire [ou le maléficiaire] du procès » (The beneficiary of an event [Ma traduction]) (Jurafsky & Martin 2021)
« Whenever Ann Callahan makes hotel reservations *for her boss*... » (ibid.)
• Beneficiary(-) : valeur pour maléficiaire

ex (60) Nous accueillons beaucoup trop de *#migrants* en France.

ex (61) [...] hausse des aides à l'accueil *des #migrants* ! [...]

ex (62) [...] la distribution par l'Etat de repas *aux #migrants* [...]

- ex (63) Quand la police sert de taxi *aux #migrants*
- ex (64) #Macron a promis l'ouverture du regroupement familial *aux migrants* « mineurs » !
[...]
- ex (65) Je n'ai pas vu dans loi de M. Collomb la suppression de la présomption de minorité pour *les #migrants*. En revanche, il prévoit pour eux le regroupement familial.
- ex (66) [...] L'Europe a besoin de cette force pour lutter contre [*l'immigration*]_{beneficiary(-)}, le fondamentalisme, et préserver nos valeurs de civilisation.
- ex (67) [...] ce sont elles qui [...] mettent la vie de [*ces migrants*]_{beneficiary(-)} en danger. [...]

9. Source • « Provenance de l'objet de l'action de transfert » (The origin of the object of a transfer event [Ma traduction]) (Jurafsky & Martin 2021)

• « Entité à partir de laquelle quelque chose se déplace, aussi bien littéralement [...] que métaphoriquement [comme dans les exemples suivants] » (the entity from which something moves, either literally [...] or metaphorically [Ma traduction]) (Saeed 2016)

The plane came back *from Kinshasa*.

6.20 We got the idea *from a French magazine*.

- ex (68) Marin a été agressé par des racailles issues de *l'immigration* massive, ça on n'en a pas beaucoup parlé dans les médias. [...]

10. Goal • « Destination de l'objet de l'action de transfert » (The destination of an object of a transfer event [Ma traduction]) (Jurafsky & Martin 2021)

• « Entité pour laquelle quelque chose se déplace, aussi bien littéralement [...] que métaphoriquement [comme dans les exemples suivants] » (the entity toward which something moves, either literally [...] or metaphorically [Ma traduction]) (Saeed 2016)

Sheila handed her license *to the policeman*.

Pat told the joke *to his friends*.

ex (69) La première humanité consiste à dire **aux #migrants** qui veulent venir en Europe que
[...]

ex (70) 262 000 titres de séjour délivrés en 2017 : en marche vers toujours plus d'**immigration**
\n Mon communiqué de presse : <https://t.co/QHxoTHEa7n> \n #Calais

ex (71) Nous devons couper toutes les pompes aspirantes de l'immigration et dire aux
candidats à **l'immigration** : [...]

ex (72) Notre action sera guidée par 3 grands principes : \n 1?? Pas d'appel d'air pour **les migrants** \n 2?? Expulsion des déboutés du droit d'asile et des clandestins [...]

ex (73) @gerardcollomb nous explique dans Le Parisien qu'il va élargir le regroupement
familial aux parents des "réfugiés mineurs" ! Il va ouvrir une nouvelle pompe à **immigration**.

ex (74) « En janvier 2018, quelques mois après l'élection de Macron, les Français voient la
réalité de ce quinquennat, le président s'inscrit dans la continuité de ses prédécesseurs :
immigration, fédéralisme européen, avantages aux plus riches .. »

Fonctions sémantico-pragmatiques des hashtags et mentions (@)

11. **_Place** Cf. 21.2.4

12. **_Place_Media** Cf. 21.2.4

13. **_Place_Political Event** Cf. 21.2.4

14. **_Speaker** Cf. 21.2.4

15. **_Topic** Valeur hybride comprenant deux niveaux d'analyse

- D'un point de vue de la structure formelle de Twitter, il s'agit des hashtags (ou des SN dépourvus du déterminant) qui sont placés en initiale ou en finale de tweet comme dans les ex (75) et ex (77).
- D'un point de vue de la structure sémantique du syntagme, de la proposition ou de la phrase, il s'agit des thèmes sur lequel porte un débat, un texte, une conférence, une émission, un livre, ou encore du thème à propos duquel une personnalité prend telle ou telle position, émet tel propos, comme dans les ex (78) et ex (79).

ex (75) Effrayant ! Sommes nous encore dans un état de droit? Comment le maire et le Prefet de #Nantes peuvent laisser faire une chose pareille ? Que fait le ministre de l'intérieur ?

#migrants

ex (76) [...] plutôt que de regarder la situation *des #migrants* en Méditerranée [...]

ex (77) J'étais hier l'invitée de #MardiPol sur @France24_fr et @RFI.
<https://t.co/bzUOAC34Oc> **#Migrants** #Aquarius

ex (78) Ce soir, réunion publique de soutien à @BuchetonLuc notre candidat @DLF_Officiel dans la #circo4504 à Cepoy. \n Baisse du pouvoir d'achat, passage à 80km/h, *migrants*... : les Français ouvrent les yeux et attendent une alternative crédible et sérieuse à la politique de #Macron !

ex (79) RT ! Fiscalité, *migrants*, chômage, Union européenne : comme ses prédécesseurs @laurentwauquiez tient un discours de fermeté. Dommage pour lui, son bilan est là pour rappeler qui il est vraiment. #LeVraiWauquiez #LEmissionPolitique

ex (80) [...] Pendant ce temps, les chiffres de *l'immigration* continuent de monter [...]

ex (81) C'est cette loi qui a concentré les difficultés liées à *l'immigration*. Entre 20% et 30% des demandes de logements sociaux sont faites par des étrangers. Encore une fois, les organismes HLM ne donnent pas les chiffres précis, pour ne pas choquer les Français. \n #Chateaudun; 2022-01-07T20:39:26.000Z; @ZemmourEric

1.8 Propriété USERNAME : Nom du compte d'utilisateur Twitter

Cette propriété correspond aux noms d'utilisateur Twitter, précédés du symbole arobase (@). La casse des caractères est respectée. Cette propriété permet d'identifier l'auteur du tweet ou du retweet.

1.9 Propriété DATE : Date de mise en ligne

Les valeurs attribuées à cette propriété indiquent la date exacte de publication du tweet ou du retweet par l'utilisateur. Les dates sont formatées de manière standard : **AAAA-MM-JJ** (par exemple, 2011-01-01).

1.10 Propriété RT : Paramètres de retweetage

Cette propriété distingue les tweets originaux des retweets. La valeur par défaut est **No** si l'utilisateur est l'auteur d'origine du contenu. Si le tweet est un retweet, la propriété est annotée avec le sigle **RT**, suivi du nom d'utilisateur de l'auteur initial.

2 L'annotation des collocations du lexique-migr

Cette section détaille les procédures d'annotation des collocations associées au lexique-migr. Ces collocations comprennent les constructions de liste et de parallélisme, ainsi que les modifications associées au lexique-migr ou réalisées par celui-ci.

2.1 Liste et parallélisme

Deux unités d'annotation sont créées pour traiter les constructions de liste et de parallélisme :

- CONJUNCT : élément coordonné ou proposition mise en parallèle avec le lexique-migr.
- LIST/PARALLELISM : annotation de la liste ou du parallélisme

1.1 Unité d'annotation : LIST/PARALLELISM

Cette unité est utilisée pour annoter les constructions de liste et de parallélisme dans lesquelles le lexique-migr est impliqué. Elle comprend deux catégories principales de propriétés :

Propriétés propres à l'unité LIST/PARALLELISM :

- Forme de surface (#forme#) : le texte tel qu'il apparaît dans le corpus.
- TYPE : indique si la construction est une liste ou un parallélisme.

Propriétés mises en relation avec l'unité MIGR-LEXICON :

- Forme de surface du lexique-migr impliqué (#forme#-MigrLex).
- Lemme du lexique-migr (LEMMA-MigrLex).
- Nom qualifié par le lexique-migr, si applicable (LEMMA_NOUN-1).

- Fonction syntaxique du lexique-migr (FUNC_SYN-MigrLex)
- Rôle sémantique du lexique-migr (ROLE_SEM-MigrLex)
- ARG_VERB-MigrLex
- Paramètre de retweetage (RT-1)
- Nom du twitto (USERNAME-1)
- Date de publication (DATE-1)

1.1.1. Forme de surface de la liste/parallélisme : propriété LIST/PARALELLISM

Le point de départ de la portée de la forme de surface correspond au nœud qui domine les *conjuncts*.

ex (82) Il est temps de mettre un terme à la confusion [entre les migrants économiques que nous n'avons pas les moyens d'accueillir dignement et les demandeurs d'asile que nous devons accueillir]_{LIST/PARALELLISM}.

1.1.2. Type de constructions : propriété TYPE

Cette propriété a deux valeurs : LIST et PARALLELISM. La valeur LIST est attribuée lorsque le lexique-migr constitue l'élément coordonné et qu'il est la tête de cet élément coordonné.

ex (83) Il est temps de mettre un terme à la confusion [entre [*les migrants*]_{tête.MIGR-LEX} que nous n'avons pas les moyens d'accueillir dignement]_{conjunct1} et [les demandeurs d'asile que nous devons accueillir]_{conjunct2}]_{LIST}

ex (84) Penser tous les jours de l'année qui vient à ceux qui sont dehors, [*#sansabris*]_{conjunct1}, [*#migrants*]_{conjunct2.MIGR-LEX}, [*#mineursnonaccompagnés*]_{conjunct3}]_{LIST}

ex (85) [En matière [*d'immigration*]_{conjunct1.MIGR-LEX}, et [de droit d'asile]_{conjunct2}] LIST,
 Macron va plus loin que la droite

La valeur PARALLELISM est attribuée lorsque le lexique-migr constitue un des éléments coordonnés de la construction de liste, mais qu'il n'en est pas la tête. Autrement dit, c'est le cas où le nœud dominant du syntagme du lexique-migr est la tête du *conjunct*. Nous considérons ces types de cas comme pseudo-parallélisme et l'annotons comme PARALLELISM (p. ex (86)) L'unité PARALLELISM peut donc s'observer aussi bien au niveau de la proposition libre ou dépendante qu'au niveau du syntagme (pseudo-parallélisme).

ex (86) [[Matraquage fiscal]_{conjunct1.CONJ}, [loi Travail au service des grands groupes]_{conjunct2.CONJ}, [[*laxisme*] [*migratoire*]_{MIGR-LEX}]_{conjunct3} et [inertie sur l'insécurité]_{conjunct4.CONJ}]_{PARALLELISM} : la communication #Macron a ses limites

ex (87) #Macron prétend à la fermeté mais [*baisse de 7 millions d'euros les moyens de lutte contre l'immigration clandestine*]_{conjunct1} et [augmente le budget de l'AME]_{conjunct2.CONJ}]_{PARALLELISM}.

ex (88) [[La vraie vie, c'est l'augmentation de l'essence.]_{conjunct1.CONJ} \n [*La vraie vie, c'est l'immigration*]_{MIGR-LEX} *qui n'est pas maîtrisée.*]_{conjunct2} \n [La vraie vie c'est l'insécurité qui monte]_{conjunct3.CONJ}]_{PARALLELISM}.

1.2 Unité d'annotation : CONJUNCT

Cette unité est dédiée à l'annotation des éléments coordonnés (*conjuncts*) ou des propositions mises en parallèle avec le lexique-migr.

1.2.1. Propriété #forme# : Forme de surface du conjunct

La forme de surface correspond à l'ensemble du constituant coordonné.

ex (89) Il est temps de mettre un terme à la confusion entre [*les migrants*]MIGR-LEXICON économiques que nous n'avons pas les moyens d'accueillir dignement et [*les demandeurs d'asile que nous devons accueillir*]CONJUNCT.

1.2.2. Propriété LEMMA_CONJ : Lemme du conjunct

Les lemmes des *conjuncts* sont standardisés en minuscule, sauf pour les noms propres (personnes, pays, lieux), comme *Merkel* ou *Marine Le Pen*. Les abréviations ou noms d'organismes, tels que *NDDL* ou *sncf*, suivent également ce format en minuscules. Lorsque la séquence est composée de plusieurs mots non séparés par des espaces (par exemple, *@partisocialiste*), le symbole est inclus dans le lemme.

1.2.3. Propriété FUNC_SYN_CONJ : Fonction syntaxique du conjunct

Cette propriété attribue une fonction syntaxique au *conjunct*, en utilisant les mêmes valeurs que celles définies dans le Tableau 2 (p.16).

Voici quelques exemples de la fonction syntaxique PRED :

ex (90) "La vraie vie, c'est *l'augmentation* de l'essence. \n La vraie vie, c'est l'immigration qui n'est pas maîtrisée. \n La vraie vie c'est *l'insécurité* qui monte. \n Je veux qu'on règle les problèmes des Français et qu'on ne les cache pas." #Mardipol

ex (91) Les Républicains sont englués dans leurs contradictions. Ils *votent au Parlement européen en faveur de tous les textes pour toujours plus de fédéralisme*, soutiennent les quotas imposés de migrants...

1.2.4. Propriété *ROLE_SEM_CONJ* : Rôle sémantique du conjunct

Le paradigme des rôles sémantiques suit celui défini pour MIGR-LEXICON dans le Tableau 3 Inventaire des rôles sémantiques, tout en intégrant des fonctions sémantico-pragmatiques spécifiques aux hashtags et mentions (@) :

Fonctions sémantico-pragmatiques des hashtags et mentions (@)	
_Place	<ul style="list-style-type: none"> • Hashtag ou mention faisant référence à un lieu autre que le média ou le lieu politique
_Place_Media	<ul style="list-style-type: none"> • Hashtag ou mention faisant référence à dese médias. Lorsqu'il s'agit du nom d'une émission audiovisuelle, du nom d'un(e) journaliste, le hashtag ou la mention ont souvent pour fonction de contextualiser la production du tweet, soit contextualiser le lieu où l'énoncé du tweet est produit (chaîne audiovisuelle, nom de l'émission ou du présentateur), soit s'y référer pour commenter le propos émis lors de l'émission, soit évoquer la participation des figures politiques à l'émission audiovisuelle. • Cette valeur sert également à annoter la fonction sémantico-pragmatique du nom de la presse écrite hashtagué (ou sous forme de mention) comme source de l'énoncé du tweet.
_Place_Political Event	<ul style="list-style-type: none"> • Hashtag ou mention faisant référence à des événements politiques.
_Speaker	<ul style="list-style-type: none"> • Mention (@xxx) faisant référence au détenteur du compte Twitter

Autres valeurs

PARALLELISM Lorsque le *conjunct* n'est pas un syntagme nominal, la valeur _PARALLELISM_ est attribuée. Si, néanmoins, les éléments subordonnés au *conjunct* sont analysables comme occupant de façon plus ou moins symétrique la même position syntaxique que le lexique-migr (du moins la même position que le micro syntagme du lexique-migr), les *conjuncts* subordonnés reçoivent les valeurs de rôle sémantique (ex (93), ex (94)).

ex (92) Le gouvernement va stopper sa chasse aux migrants, [offrir des solutions pérennes aux SDF, [protéger les salariés], [augmenter le SMIC], [ramener la part du nucléaire à 50% d'ici 2025], [augmenter les APL], [traquer les fraudeurs fiscaux], [mener une politique décente].

ex (93) L'UE, c'est les problèmes [migratoires]_{MIGR-LEX} [la concurrence déloyale]_{CONJ+1 :PRED}, [les traités de libre-échange]_{CONJ+2:PRED} qui nous font importer des produits ne respectant pas nos normes sanitaires et environnementales !

ex (94) Avec l'UE, [les immigrés]_{MIGR-LEX} arrivent, [les emplois s'en vont]_{CONJ+1:SUB}!

1.2.5. Propriété *POSITION_CONJ* : Position du *conjunct*

Cette propriété indique la position et la distance du *conjunct* par rapport au lexique-migr.

- L1 : cotexte gauche (*Left*) immédiat du lexique-migr.
- L2 : cotexte gauche à un intervalle d'un *conjunct*.
- R1 : cotexte droit (*Right*) immédiat du lexique-migr.
- R2 : cotexte droit à un intervalle d'un *conjunct*.

1.3 Mise en relation des unités d'annotation pour l'analyse des liste et parallélisme

La relation entre MIGR-LEXICON et LIST/PARALLELISM est essentielle pour analyser les constructions de listes et de parallélismes impliquant le lexique-migr.

1.3.1. Mise en relation de la liste/parallélisme au lexique-migr : REL-MIGR-LEX

La relation MIGR-LIST/PAR est annotée dans la Gestion des relations du logiciel.

- Élément 1 : Unité LIST/PARALLELISM.
- Élément 2 : Unité MIGR-LEXICON.

Règles spécifiques :

1. Lorsque plusieurs éléments de la liste appartiennent au lexique-migr, seul le premier élément sera relié à LIST/PARALLELISM. Cette contrainte découle des limitations du logiciel, qui n'autorise chaque élément à être en relation directe qu'avec un seul autre élément.
2. Pour compenser cette limitation, une autre relation, MIGR-CONJ, permet d'établir des connexions entre les différents éléments coordonnés.

Note importante : Cette relation ne possède pas de propriétés supplémentaires.

Astuce pratique : Dans la vue de gestion des relations, le logiciel permet d'afficher directement le lemme de l'unité MIGR-LEXICON associée à LIST/PARALLELISM. Cela facilite l'identification des connexions dans l'interface.

1.3.2. Mise en relation des conjuncts au lexique-migr : SCH-MIGR-CONJUNCT et REL-MIGR-CONJ

Cette étape vise à établir un schéma d'annotation permettant de relier les *conjuncts* identifiés aux unités du lexique-migr. La procédure s'effectue en deux phases : la création et l'attribution de propriétés au schéma, suivies de la mise en relation entre le schéma et le lexique-migr.

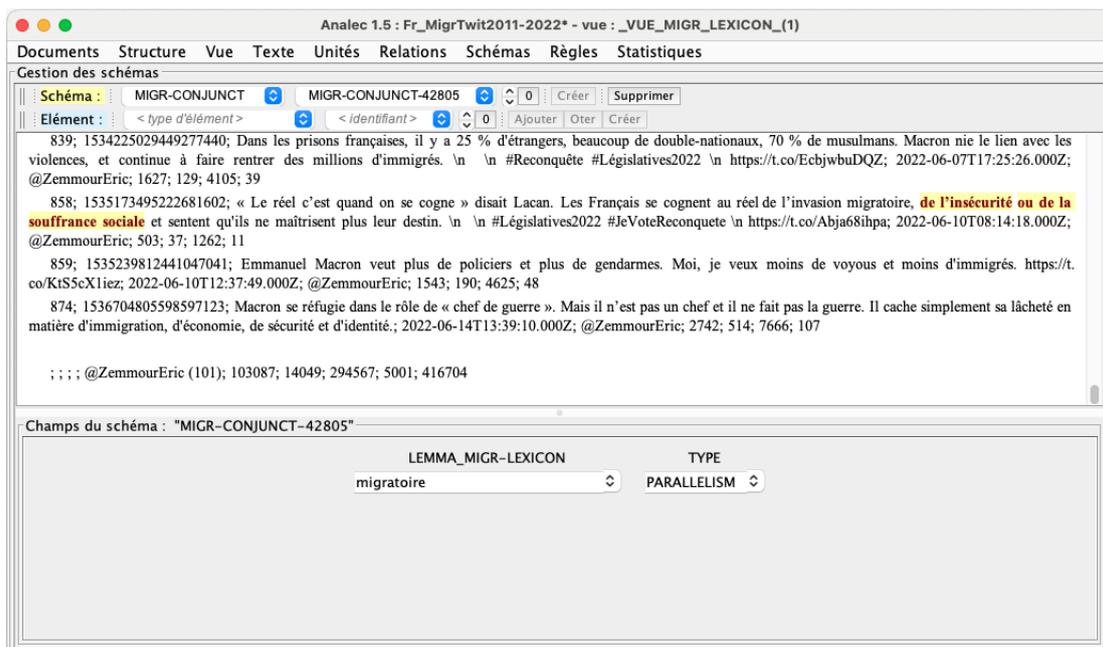
1. Phase 1 : Création du schéma SCH-MIGR-CONJUNCT

Dans cette phase, tous les conjuncts sont regroupés dans un schéma unique au sein de la Gestion des Schémas. L'annotateur attribue au schéma deux propriétés principales :

- **TYPE** : Identifie la nature du schéma, avec deux valeurs possibles : LIST ou PARALLELISM.
- **LEMMA_MIGR-LEXICON** : Correspond au lemme du lexique-migr auquel les *conjuncts* sont associés. Cette valeur est sélectionnée depuis l'onglet correspondant.

La Figure 4 visualise la création du schéma des *conjuncts*.

Figure 4 Visualisation du schéma des conjuncts dans Analec (SCH-MIGR-CONJ)



2. Phase 2 : Mise en relation avec MIGR-LEXICON

Le schéma entier SCH-MIGR-CONJUNCT (Élément 1) est ensuite relié à l'unité MIGR-LEXICON (Élément 2) dans la Gestion des Relations.

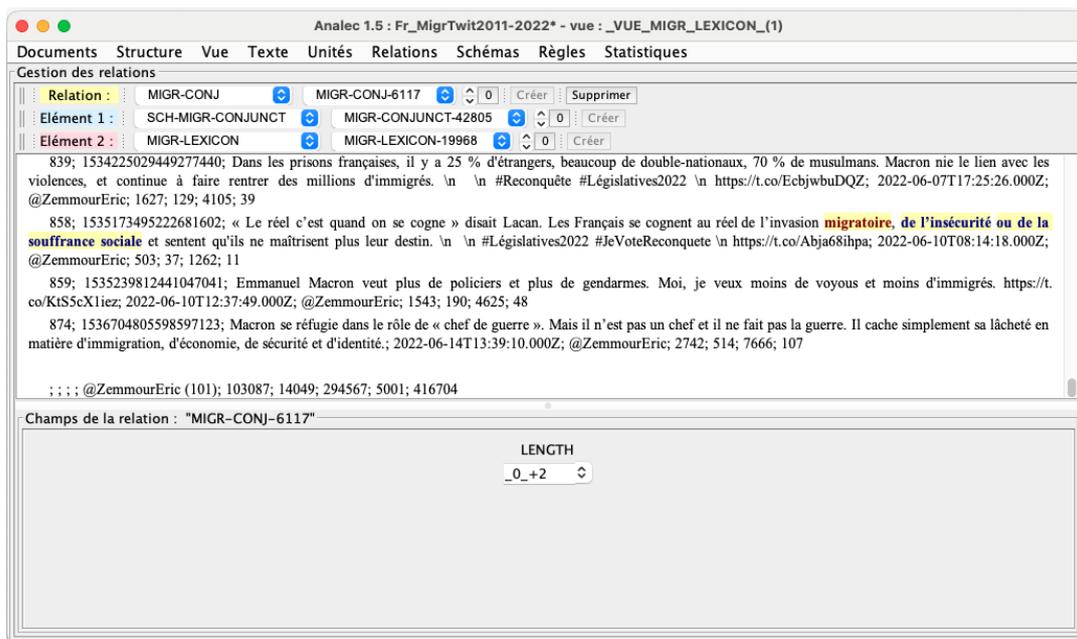
Propriété LENGTH : Longueur de la construction

Lors de la création de la relation MIGR-CONJ, l'annotateur spécifie la longueur de la construction de liste ou de parallélisme via la propriété LENGTH. Les valeurs possibles incluent des caractères numériques et spéciaux :

- **0** : Correspond au lexique-migr.
- **1** : Représente un conjunct situé immédiatement à gauche (-1) ou à droite (+1).
- **2** : Indique un conjunct situé à un intervalle plus éloigné à gauche (-2) ou à droite (+2).
- **Tiret bas (_)** : Utilisé pour marquer la frontière entre les éléments coordonnés.

La Figure 5 visualise la création de la relation entre le SCH-MIGR-CONJ et l'unité MIGR-LEXICON.

Figure 5 Visualisation de la relation reliant le SCH-MIGR-CONJ à l'unité MIGR-LEXICON



Facilitation visuelle : La manipulation de l'interface permet d'afficher dans la Gestion des unités CONJUNCT le lemme du lexique-migr associé aux *conjuncts* identifiés. Cela simplifie la navigation et la cohérence des annotations.

Cas particuliers : Doublons dans les relations

Pour des exemples comme ex (86), où le lexique-migr est simultanément un conjunct de liste et de parallélisme, une duplication de l'unité MIGR-LEXICON est nécessaire. Cela permet de relier correctement une unité unique à plusieurs constructions LIST/PARALLELISM.

2.2 Modification

Cette section décrit les procédures d'annotation des adjectifs ou noms qualificatifs associés au lexique-migr, ainsi que des noms modifiés par celui-ci.

2.1 Unité d'annotation : MODIFIER

L'unité **MODIFIER** est utilisée pour annoter les éléments qui modifient le lexique-migr, notamment :

- Adjectifs qualificatifs,
- Déterminants numériques et quantitatifs,
- Syntagmes prépositionnels introduits par *de* (e.g. *immigration de masse*) ou *en* (e.g. *migrants en danger de mort*),
- Participes présents et passés,
- Propositions subordonnées relatives.

2.1.1. Forme de surface du modifieur du lexique-migr : propriété #forme#

La portée de l'annotation inclut la séquence entière constituant le modifieur.

ex (95) La première humanité consiste à dire aux #migrants [*qui veulent venir en Europe*]MODIFIER que nous n'avons plus les moyens de les accueillir.

ex (96) Nous ne pouvons pas accueillir [*l'ensemble des*]MODIFIER migrations [*qui continueront à augmenter de façon exponentielle*]MODIFIER !

ex (97) L'asile était un beau principe, complètement dévoyé pour devenir une filière d'immigration [*hors contrôle*]MODIFIER. [...]

ex (98) Il est temps de mettre un terme à la confusion entre les migrants [*économiques*]MODIFIER [*que nous n'avons pas les moyens d'accueillir dignement*]MODIFIER et les demandeurs d'asile que nous devons accueillir.

2.1.2. Propriété LEMMA_MODIF : Lemme du modifieur du lexique-migr

Les formes lemmatiques des syntagmes prépositionnelles sont précédées de la préposition *de/en*. Les lemmes des modifieurs suivent les règles suivantes :

- Adjectifs qualificatifs, déterminants quantitatifs et numéraux : lemmes au masculin, singulier.
- Déterminants numéraux : écrits sous forme de chiffres arabes.
- Syntagmes prépositionnels : précédés de *de* ou *en*.

Voici ci-dessous des exemples de lemmes :

LEMMA_MODIF	#forme#
% de	90% des migrants qui viennent en France sont des hommes, qui ne fuient pas la guerre mais viennent pour des raisons économiques.
qui	90% des migrants qui viennent en France sont des hommes, qui ne fuient pas la guerre mais viennent pour des raisons économiques.
que	Il est temps de mettre un terme à la confusion entre les migrants économiques que nous n'avons pas les moyens d'accueillir dignement et les demandeurs d'asile que nous devons accueillir.
_Qt.Num	Près de 500 000 migrants sont entrés [...] [...] Pourquoi Salvini ne pourrait pas expulser 500 000 migrants ? [...]
_Qt	[...] s'ils veulent bien accueillir plusieurs centaines de migrants clandestins ?
clandestin	[...] s'ils veulent bien accueillir plusieurs centaines de migrants clandestins ?
de l'aquarius	[...] pour l'accueil des migrants de l'#Aquarius [...]
en masse	[...] va-t-on continuer avec cette Union européenne qui nous impose la concurrence déloyale, les migrants en masse , les sacrifices contre nos retraités et nos services publics ?
aidé	[...] Les #migrants aidés de leurs complices détournent le droit d'asile.

bousculant	Les "responsables" ne sont pas ceux qui manifestent mais ceux qui ont laissé entrer des centaines de milliers de #migrants <i>bousculant et remettant en cause l'identité même de nos pays européens et donc notre mode de vie.</i> [...]
lisant du Victor Hugo	[...] On est loin des gentils #migrants <i>lisant du Victor Hugo !</i> [...]

2.2 Unité d'annotation : NOUN

L'unité NOUN est utilisée pour annoter les noms modifiés par le lexique-migr.

2.2.1. Forme de surface du nom modifié par le lexique-migr : propriété #forme#

Cette propriété capture la forme de surface du nom modifié par le lexique-migr.

ex (99) Le nombre record de demandes d'asile en France en 2017 prouve que *la pression* migratoire dans notre pays continue d'exploser.

ex (100) [...] les centres existent partout en France qu'elle vienne à l'ONF à Velaine en Haye 100 *jeunes hommes* migrants

ex (101) On assiste à une comédie bien orchestrée sur le projet de *loi* immigration. Rien ne va changer réellement.

2.2.2. Lemme du nom modifié par le lexique-migr : propriété LEMMA_NOUN

Les lemmes des noms modifiés sont indiqués au masculin, singulier.

2.3 Mise en relation des unités d'annotation pour l'analyse des modificateurs du lexique-migr

2.3.1. Mise en relation du lexique-migr au MODIFIER : Relation MIGR-MODIF1/2

Cette relation relie MIGR-LEXICON au MODIFIER. La direction de la relation dépend de la position du modificateur. La direction de la jonction entre les deux éléments dépend de la position de l'adjectif modifiant le lexique-migr. Si le modificateur se trouve dans le cotexte gauche immédiat du lexique-migr, l'élément 1 correspond au modificateur et l'élément 2 au lexique-migr. En revanche, lorsque le modificateur se trouve dans le cotexte droit immédiat, l'élément 1 correspond au lexique-migr et l'élément 2 au modificateur. Ce deuxième cas est illustré dans la Figure 6, dans son contexte d'interface d'annotation.

	MODIFIER	MIGR-LEXICON	MODIFIER	MODIFIER
REL-MIGR-MODIF_1	Elément1	Elément2		
REL-MIGR-MODIF_2		Elément1	Elément2	
		Element1		Element2

Figure 6 Visualisation de la relation entre MIGR-LEXICON (élément 1) et MODIFIER (élément 2)

Analec 1.5 : Fr_MigrTwit2011-2022 - vue : _VUE_MIGR_LEXICON_(1)

Documents Structure Vue Texte Unités Relations Schémas Règles Statistiques

Gestion des relations

Relation	Elément 1	Elément 2
MIGR-MODIF_1	MIGR-LEXICON	MODIFIER

L11 (MigrTw : 21)

; data__id; data__text; data__created_at; author__username; RT; reply; like; quote

1; 31002218702311424; Intervention de police à la Cité Nationale de l'Histoire de l'Immigration <http://bit.ly/dWUJiu>; 2011-01-28T14:54:39.000Z; @partisocialiste; 3; 0; 0; 0

2; 35809112277131264; Sarkozy a pourtant défendu l'immigration zéro... #pdef http://www.dailymotion.com/video/x2sg7z_sarkozy-et-immigration-zero_politics; 2011-02-10T21:15:32.000Z; @partisocialiste; 8; 0; 0; 0

3; 43733771631075328; RT @eurodeputesse: «Jusqu'à présent, l'UE comptait sur Kadhafi pour maintenir en rétention les migrants» H.Flautre Libé: <http://ht.ly/485kL>; 2011-03-04T18:05:18.000Z; @EELV; 6; 0; 1; 0

4; 56335463228768256; Le PS dénonce l'attitude de l'Italie et de la France à l'égard des migrants <http://bit.ly/dP5msm>; 2011-04-08T12:39:55.000Z; @partisocialiste; 3; 0; 0; 0

5; 57821652863565824; Dialogue de sourds en Europe sur l'immigration <http://bit.ly/eZOVp7>; 2011-04-12T15:05:30.000Z; @partisocialiste; 0; 0; 0; 0

Champs de la relation : "MIGR-MODIF_1-1"

2.3.2. Mise en relation du lexique-migr au NOUN : Relation MIGR-MODIFI

Pour les lexiques-migr ayant une fonction d'adjectif, cette relation relie MIGR-LEXICON (Élément 2) au NOUN qu'il modifie (Élément 1). La direction reste constante.

Facilitation visuelle :

La fonctionnalité d'Analec « Vue » permet d'afficher dans l'unité MIGR-LEXICON la propriété LEMMA_NOUN. L'interface de l'unité MIGR-LEXICON visualisée ci-dessous illustre ainsi la richesse de l'annotation à dimensions multiples intégrée.

Figure 7 Visualisation de l'annotation de l'unité migr-lexicon dans Analec (1)

The screenshot shows the Analec 1.5 interface with the following details:

- Window title: Analec 1.5 : forme allégée pour présentation* - vue : _VUE_MIGR_LEXICON_(1)
- Menu bar: Documents, Structure, Vue, Texte, Unités, Relations, Schémas, Règles, Statistiques
- Unit selection: MIGR-LEXICON, MIGR-LEXICON-653
- Text area: A list of units with their metadata, including #forme#, LEMMA, ROLE_SEM, FUNC_SYN, ARC_VERB, LIST/PAR, REL-LIST/PAR, and LENGTH-1.
- Annotation fields: LEMMA_MODIF_L1, LEMMA_MODIF_R1, LEMMA_MODIF_R2, LEMMA_MODIF_R3, LEMMA_MODIF_R4.
- LEMMA_NOUN-1 field: A blue dashed box highlights the field containing the word 'flux'.
- RT, USERNAME, DATE fields: RT is 'No', USERNAME is 'L_@EstherBenbassa', DATE is '2015-07-30'.

Cet atout unique du logiciel en fait un outil d'analyse efficace et pertinent, offrant aux chercheurs la possibilité de réaliser des annotations pluriniveau et d'explorer de manière plus approfondie et structurée les unités annotées. La Figure 8 illustre une annotation pluridimensionnelle de l'unité MIGR-LEXICON-18983 qui apparaît à la fois comme modifiée et intégrée dans une construction de liste. Cette annotation met en évidence les relations

syntaxiques entre les unités annotées et permet de cibler des facteurs d'analyse spécifiques, en fonction des questions de recherche.

Figure 8 Visualisation de l'annotation de l'unité migr-lexicon dans Analec (2)

Analec 1.5 : forme allégée pour présentation* - vue : _VUE_MIGR_LEXICON_(1)

Documents Structure Vue Texte Unités Relations Schémas Règles Statistiques

Gestion des unités

Unité : MIGR-LEXICON MIGR-LEXICON-18983 0 Créer Supprimer Rectifier une borne

https://t.co/z1IkcS2H9; 2022-02-20T20:16:08.000Z; @J_Bardella; 111; 39; 338; 4

333; 1495509959093723154; Arrêter l'immigration et expulser les clandestins réglera en très grande partie l'insécurité en France. https://t.co/TT6WbjhRvD; 2022-02-20T21:25:34.000Z; @J_Bardella; 186; 64; 657; 6

344; 1498325185186914312; Accueillir des familles ukrainiennes le temps du conflit redonnerait tout son sens à un droit d'asile transformé en nouvelle autoroute migratoire. 'n 'n Parce que l'Europe est une réalité géographique et civilisationnelle, c'est en son sein que la solidarité doit s'exprimer. #Ukraine; 2022-02-28T15:52:16.000Z; @J_Bardella; 208; 94; 814; 10

359; 1499712502598361089; Emmanuel Macron est l'homme de la submersion migratoire : 1.5 million d'étrangers supplémentaires sont entrés durant son mandat par la seule voie légale. Marine Le Pen stoppera l'immigration et rendra aux Français la maîtrise de leur destin. 'n 'n #MarinePrésidente #SanaLui @SanaLui. https://t.co/BXnPHS9qj; 2022-03-04T11:44:58.000Z; @J_Bardella; 99; 34; 303; 3

387; 1502234103693062146; Le laxisme migratoire n'épargne pas l'Outre-mer. Nous remettons de l'ordre partout, nous appliquerons les lois, nous supprimerons le droit du sol et nous instaurerons la priorité nationale. https://t.co/KcK8QA0ho; 2022-03-11T10:44:55.000Z; @J_Bardella; 66; 44; 184; 0

410; 1503463652745175051; RT @MLP_officiel: ?? Présidente de la République, je vous rendrai votre argent, en augmentant votre pouvoir d'achat et en sauvant notre système de protection sociale ; et je vous rendrai votre pays, en vous protégeant de l'immigration anarchique et massive et de l'insécurité. #FaceALaGuerreTF1 https://t.co/VWfoeGYYyz; 2022-03-14T20:10:42.000Z; @J_Bardella; 717; 0; 0; 0

433; 1504545606093742084; Je le dis aux Français : la politique est partout dans votre quotidien. Le racket fiscal, l'insécurité, l'immigration incontrôlée que vous subissez, ce sont les conséquences de décisions d'Emmanuel Macron. Ne vous faites pas voler cette élection : les 10 et 24 avril, VOTEZ ! https://t.co/U3kYDk8AHg; 2022-03-17T19:50:00.000Z; @J_Bardella; 86; 17; 204; 1

453; 1505162832173797380; Il n'y a rien à commémorer le 19 mars, date de début de massacres impardonnables. Si l'« apaisement » avec l'Algérie qu'Emmanuel Macron appelle de ses vœux signifie naïveté migratoire, courbettes face aux dirigeants anti-français, repentance permanente de la France, c'est non !; 2022-03-19T12:42:38.000Z; @J_Bardella; 229; 37; 804; 5

459; 1505549722383110144; Les filières migratoires irakiennes ou syriennes étaient composées à 90% d'hommes. En Ukraine, nous sommes confrontés à des familles qui fuient la guerre et dont les hommes sont restés combattre. La France peut en accueillir le temps du conflit, dans le respect du droit d'asile. https://t.co/iXpx5GFUIH; 2022-03-20T14:20:00.000Z; @J_Bardella; 62; 22; 213; 1

498; 1506641919199248392; Les LR ont un problème de crédibilité et de sincérité. Leur bilan, c'est la suppression de postes de policiers, plus d'immigration que sous la gauche, la fin de la double peine qui nous interdit d'expulser les criminels étrangers. Nous en payons encore le prix ! 'n 'n #GrandDébatVA https://t.co/IJKKqUX2O8; 2022-03-23T14:40:00.000Z; @J_Bardella; 132; 5; 305; 0

499; 1506651985809977347; Marine Le Pen n'est pas dans le slogan, mais dans l'efficacité. Sa candidature offre des réponses clés en main, concrètes, sur l'immigration, l'islamisme, le pouvoir d'achat. Sa candidature est celle de la solidité, aussi bien dans la vision que dans sa

Champs de l'unité : "MIGR-LEXICON-18983"

#forme#	LEMMA	ROLE_SEM	FUNC_SYN	ARG_VERB	LIST/PAR	REL-LIST/PAR	LENGTH
immigration	immigration	_Topic	ROOT	no	_NONE_	LIST	-2_0_

MODIFICATION

MODIFIED

LEMMA_MODIF_L1

LEMMA_MODIF_R1

LEMMA_MODIF_R2

LEMMA_MODIF-R3

LEMMA_MODIF-R4

LEMMA_NOUN-1

RT

USERNAME

DATE

No

R_@_J_Bardella

2022-03-17

3 Synthèse de la procédure de l'annotation

L'annotation s'effectue en deux niveaux. Cinq types d'unités sont annotés :

- Lexème-migr (nom, adjectif, adverbe et verbe) (MIGR-LEXICON)
- Modifieur du nom-migr (MODIFIER)
- Nom modifié par l'adjectif-migr (NOUN)
- Liste/parallélisme contenant un lexème-migr (LIST/PARALLELISM)
- *Conjunct* de la liste/parallélisme-migr (CONJUNCT)

Afin d'illustrer la procédure de l'annotation dans Analec, prenons pour exemple l'annotation de l'unité MIGR-LEXICON. Cette opération consiste à annoter manuellement une séquence de caractères contenant la racine latine *migr-* en faisant glisser le curseur sur la séquence en question. Une fois l'unité d'annotation identifiée avec sa forme de surface (*#forme#*), l'annotateur lui attribue les différentes propriétés linguistiques – lexicale, syntaxique et sémantique.

- Lorsque le lexème-migr est modifié par l'adjectif. L'étape suivante consiste à annoter cet adjectif comme unité MODIFIER. Le lemme du modifieur sera annoté par la suite.
- Si le lexème-migr est un membre constitutif de la liste, l'annotateur sera amené à recourir à deux unités : LIST/PARALLELISME et CONJUNCT. Pour l'unité LIST/PARALLELISME, le type sera précisé par l'annotateur (list ou parallélisme). Pour l'unité CONJUNCT, l'élément coordonné dans une construction de liste/parallélisme, l'annotateur précisera le lemme, la fonction syntaxique, le rôle sémantique ainsi que la position par rapport au lexique-migr.
- Lorsque le lexème-migr a pour fonction de modifier un autre nom (p.ex. submersion migratoire), l'annotateur recourt à l'unité NOUN. Le lemme sera choisi par l'annotateur.

Le deuxième niveau d'annotation consiste à relier les unités annotées les unes aux autres en fonction de la relation syntaxique. Pour créer une relation syntaxique entre le modifiant et le modifié, l'annotateur est amené à sélectionner la fonctionnalité Relations du logiciel. Le schéma

d'annotation pour cette fonctionnalité Relations est déployé dans la deuxième colonne de la structure des annotations. En fonction de la position syntaxique du modifieur par rapport au lexème-migr, l'annotateur choisira le type de relation MIGR-MODIF.

La fonctionnalité Relations permet de créer une relation unilatérale entre deux éléments. Par exemple, une unité A peut être reliée à une unité B et l'annotateur peut indiquer l'ordre d'apparition des unités dans le texte en choisissant la direction soit de gauche à droite, soit de droite à gauche. Si le lexème-migr est modifié par deux adjectifs dans ses cotextes gauche et droit immédiats, l'annotateur choisira la relation MIGR-MODIF_1. Lorsqu'un deuxième adjectif se trouve dans le cotexte droit du lexème-migr, l'annotateur choisira la relation MIGR-MODIF_2.

Pour annoter la relation entre le lexème-migr et la construction de liste/parallélisme, l'annotateur choisira la relation MIGR-LIST/PAR. Nous avons également annoté la relation entre les membres constitutifs de la liste (CONJUNCT) et le lexème-migr. Pour réaliser cette annotation de la relation MIGR-CONJ, il nous a fallu recourir au préalable à la fonctionnalité Schémas. Car plusieurs unités CONJUNCT sont susceptibles d'être mises en relation avec une unité MIGR-LEXICON. La fonctionnalité Schémas permet justement de relier les unités qui doivent se comporter comme une seule unité et les éléments constitutifs du schéma peuvent recevoir les mêmes propriétés en même temps. Dans la troisième colonne de la structure des annotations se trouve le schéma MIGR-CONJUNCT. Une fois les unités CONJUNCT reliées sous forme de schéma l'annotateur peut relier ce schéma à l'unité MIGR-LEXICON en se rendant à nouveau dans l'interface des Relations du logiciel.

L'unité MIGR-LEXICON est considérée comme noyau des relations de construction de liste/parallélisme et de modification et les unités de *conjunct* (sous forme de schéma), de modifieur et de nom comme satellites :

- (1) Construction de Liste/parallélisme : il s'agit de relier l'unité MIGR-LEXICON à l'unité LIST/PARALLELISME (Relation MIGR-LIST/PAR)
- (2) Construction de Liste/parallélisme : : il s'agit de relier l'unité MIGR-LEXICON au schéma constitué des CONJUNCT (Relation MIGR-CONJ)
- (1) Modification du nom-migr : il s'agit de relier l'unité MIGR-LEXICON à l'unité de MODIFIER. La place de l'adjectif par rapport à l'unité MIGR-LEXICON est prise en compte. (MIGR-MODIF_1, 2, 3, 4 et 5)
- (2) Modification du nom par l'adjectif-migr : il s'agit de relier l'unité MIGR-LEXICON à l'unité NOUN. (MIGR-MODIF_1)

REFERENCES BIBLIOGRAPHIQUES

Méthodes générales de linguistique de corpus et annotation du corpus

- Bergounioux, G. (Ed.). (2016). *Linguistique de corpus. Une étude de cas*. Honoré Champion.
- Bergounioux, G., Jacobson, M., & Pietrandrea, P. (2017). L'annotation des corpus oraux. *HAL Archives Ouvertes*. <https://halshs.archives-ouvertes.fr/halshs-03082419>
- Eshkol-Taravella, I., Baude, O., Marel, D., Hriba, L., Dugua, C., & Tellier, I. (2011). Un grand corpus oral disponible : le corpus d'Orléans 1968–2012. *Traitement Automatique des Langues*, 52(3), 17–46.
- Lacheret, A., Kahane, S., Beliao, J., Dister, A., Gerdes, K., Goldman, J.-P., Obin, N., Pietrandrea, P., & Tchobanov, A. (2014). Rhapsodie : un Treebank annoté pour l'étude de l'interface syntaxe-prosodie en français parlé. *SHS Web of Conferences*. <https://doi.org/10.1051/shsconf/20140801305>
- Lacheret, A., Kahane, S., & Pietrandrea, P. (2015). *Rhapsodie: A Prosodic-Syntactic Treebank for Spoken French*. Amsterdam/Philadelphia: John Benjamins Publishing Company. <https://doi.org/10.1075/scl.89>

Analec dans les projets d'annotation

- Landragin, F. (2019). *Manuel d'annotation du projet DEMOCRAT (version 2.4.3)*. Projet ANR DEMOCRAT.
- Landragin, F., Poibeau, T., & Victorri, B. (2012). ANALEC: A new tool for the dynamic annotation of textual data. *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2012)* (pp. 357–362). Istanbul, Turkey: European Language Resources Association.
- Landragin, F. (2016). *Description, modélisation et détection automatique des chaînes de référence en français*. Présentation générale du projet DEMOCRAT.
- Landragin, F., & Schnedecker, C. (2014). Les chaînes de référence. *Langages*, 195, 3–9.
- Pietrandrea, P., & Cervoni, V. (2016). *Modal – annotation guidelines: Version 1.0*.
- Quignard, M., Le Mené, M., & Landragin, F. (2021). *Élaboration du corpus DEMOCRAT : Procédures d'annotation et d'évaluation*. *Langages*, 224(4), 11-24.

Schnedecker, C., Glikman, J., & Landragin, F. (2017). Les chaînes de référence en corpus. *Langue française*, 195, 3–14. Armand Colin.

Constructions de liste et de parallélisme

Kahane, S., & Pietrandrea, P. (2012). La typologie des entassements en français. *Actes du CMLF*, 1809-0828.

Kahane, S., Pietrandrea, P., & Gerdes, K. (2018). The annotation of list structures. In A. Lacheret, S. Kahane, & P. Pietrandrea (Eds.), *Rhapsodie: A Prosodic-Syntactic Treebank for Spoken French* (pp. 171–186). Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/scl.89>

Masini, F., Mauri, C., & Pietrandrea, P. (2018). List constructions: Towards a unified account. *Italian Journal of Linguistics*, 30(1), 49–94. <https://doi.org/10.26346/1120-2726-116>

Pietrandrea, P., & Battaglia, E. (2022). Migrants and the EU: The diachronic construction of ad hoc categories in French far-right discourse. *Journal of Pragmatics*, 192, 139–157. <https://doi.org/10.1016/j.pragma.2022.02.012>

Modifieurs

Rijkhoff, J. (2014). Modification as a propositional act. *Studies in Functional and Structural Linguistics (SFSL)*. <https://doi.org/10.1075/sfsl.68.06rij>

Rijkhoff, J. (2008). Descriptive and discourse-referential modifiers in a layered model of the noun phrase. *Linguistics*, 46(4), 789–829. <https://doi.org/10.1515/LING.2008.026>

Bases théoriques des étiquettes annotées : rôles sémantiques et fonctions syntaxiques

Jurafsky, D., & Martin, J. H. (2019). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* (3rd ed.). Pearson.

Lacheret, A., Kahane, S., Beliao, J., Dister, A., Gerdes, K., Goldman, J.-P., Obin, N., Pietrandrea, P., & Tchobanov, A. (2014). Rhapsodie : un Treebank annoté pour l'étude de l'interface syntaxe-prosodie en français parlé. *SHS Web of Conferences*. <https://doi.org/10.1051/shsconf/20140801305>

Lacheret, A., Kahane, S., & Pietrandrea, P. (2015). *Rhapsodie: A Prosodic-Syntactic Treebank for Spoken French*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
<https://doi.org/10.1075/scl.89>

Saeed, J. I. (2016). *Semantics* (4th ed.). John Wiley & Sons, Inc.

Tellier, C. (2004). *Éléments de syntaxe du français : Méthodes d'analyse*. Armand Colin.

Constructions nominales complexes

Bartning, I. (1996). Éléments pour une typologie des SN complexes en *de* en français. *Langue Française*, 109, 29–43. <https://doi.org/10.3406/lfr.1996.5332>