

Le sens de la famille : analyse du vocabulaire de la parenté par les plongements de mots



Ludovic Tanguy, Cécile Fabre,
Nabil Hathout, Lydia-Mai Ho-Dac

CLLE : CNRS & Université de Toulouse



Objectifs de l'étude

Questionner les représentations sémantiques issues des corpus

- Renouveler les études en sémantique lexicale en s'appuyant sur les analyses distributionnelles à bases neuronales
- Stabiliser une méthode
- Produire de nouvelles descriptions exploitant la diversité des corpus

La famille : un champ d'investigation approprié

- Lexique de taille moyenne, fréquent, couvrant la plupart des genres textuels
- Sémantiquement riche, structuré et bénéficiant de descriptions approfondies issues des sciences sociales
- Connecté aux rôles sociaux et donc susceptibles de varier selon les contextes d'énonciation et les époques

Fil directeur

- Y a-t-il adéquation entre les critères habituels de structuration de la famille et les représentations issues des corpus ?
- Ces représentations diffèrent-elles selon les contextes discursifs (genres, époques) ?
- Si oui, comment caractériser ces différences ?

Plan

Matériel

- Choix des termes et des corpus, description componentielle, embeddings

Première analyse des représentations vectorielles

Modélisation des relations familiales

- Croisement des représentations

Annotation des voisins distributionnels

- Annotation humaine
- Annotation automatique

Analyse

- Croisement des représentations

Données et matériel

Choix du lexique

Point de départ : liste de 82 termes (Wikipédia : page “relation de parenté”)

Filtrage par fréquence dans les corpus (minimum 100 occurrences dans chacun)

Rejet de *arrière-petite-cousine*, *arrière-arrière-grand-oncle*, *beau-cousin*, *bru*, *marâtre*...

Cas des termes polysémiques

Neutres (*parent*, *enfant*...) et *femme* non conservés,
filles gardé (au vu des voisinages distributionnels, cf infra)

Liste finale de 25 termes amorces

beau-frère, *beau-père*, *belle-fille*, *belle-mère*, *belle-soeur*, *cousin*, *cousine*, *épouse*, *époux*, *filles*, *fil*, *frère*,
gendre, *grand-mère*, *grand-père*, *mari*, *mère*, *neveu*, *nièce*, *oncle*, *père*, *petit-fils*, *petite-fille*, *soeur*, *tante*

Analyse componentielle

Les liens familiaux peuvent se décomposer en un petit ensemble de relations élémentaires (Augustins 2000, Barry 2018)

Si l'on considère *ego* comme référentiel, chaque terme peut donc être décomposé en traits atomiques :

Terme	genre	ascendance	descendance	germanité	alliance
<i>grand-mère</i>	F	2	0	0	0
<i>belle-soeur</i>	F	0	0	1	1
<i>oncle</i>	M	1	0	1	0,5
<i>cousin</i>	M	1	1	1	0
<i>cousine</i>	F	1	1	1	0

Corpus

Sélection des sources

Double contrainte : variation et taille minimale (dizaines de millions)

5 corpus :

Frantext	Fiction (romans XXe)	30M mots
FrWac	Web tout venant	1600M mots
Le Monde	Presse	200M mots
Revue.org	Articles scientifiques SHS	200M mots
Wikipédia	Encyclopédie	720M mots

Prétraitement

Principes : faciliter le repérage des régularités et limiter l'impact du genre grammatical

Lemmatisation et normalisation

Analyseur Stanza + correction par lexique Morphalou + ligatures

Neutralisation des noms propres (--> "NAM")

Il fut alors recueilli par sa grand-mère paternelle et tutrice, Marie des Moulins, dont la sœur avait donné asile à des maîtres jansénistes , notamment Lancelot, et dont la fille et deux autres de ses sœurs étaient religieuses à Port-Royal. (FrWaC)

*il être alors recueillir par **son** grand-mère **paternel** et tutrice , **NAM** de le **NAM** , dont **le** soeur avoir donner asile à un maître janséniste , notamment **NAM** , et dont **le** fille et deux autre de **son** soeur être **religieux** à **NAM** .*

Plongements de mots (*embeddings*)

Principes

Représentation vectorielle des mots-types basée sur le calcul de cooccurrence (fenêtre)

Méthode neuronale (prédiction des cooccurrences), algorithme Skip-Gram with Negative Sampling (SGNS) implémenté dans *Word2vec* (Mikolov et al. 2013)

Représentation dense dans un espace vectoriel avec des dimensions non interprétables mais permettant des opérations de calcul de similarité (cosinus)

Implémentation

1 représentation par corpus + 1 global (concaténation des embeddings)

Bibliothèque Gensim (fenêtre 5 mots, 100 occurrences minimum, 300 dimensions)

5 runs par corpus pour limiter l'impact des phénomènes aléatoires

Etude des plongements des amorces

Vue d'ensemble

Analyses par modèle :

Observation directe du voisinage distributionnel : plus proches voisins

Structuration interne du lexique de la famille : clustering et projections

Plus proches voisins

Extraction des N mots du corpus les plus proches (cosinus) de chaque amorce, par modèle/corpus (25 m/c : 5 runs x 5 corpus)

Pour toutes les amorces et dans chaque modèle/corpus **présence massive de termes de la famille :**

93% en moyenne au rang 10, 67% au rang 25

Rang / amorce	petite-fille	père	soeur	tante
1	petit-enfant	oncle	frère	oncle
2	cadette	mère	mère	mère
3	parente	grand-père	belle-soeur	cousine
4	nièce	frère	cousine	père
5	petit-fils	tante	nièce	grand-mère
6	cousine	parent	aîné	madame
7	belle-soeur	mari	tante	belle-soeur
8	fiancée	beau-frère	amie	maman
9	gendre	fils	père	grand-père
10	épouse	tuteur	filie	cousin ¹²

Structuration interne

Objectifs :

- faire émerger l'organisation interne du lexique
- identifier des dimensions sous-jacentes
- comparer les modèles/corpus

Deux approches mises en oeuvre :

- clustering hiérarchique
- projection/réduction de dimensions (t-SNE)

Clustering hiérarchique ascendant

Modèle Global, cosinus
critère de Ward

beaucoup de paires
générées (*frère/soeur*) mais
pas seulement
(*père/grand-père*)

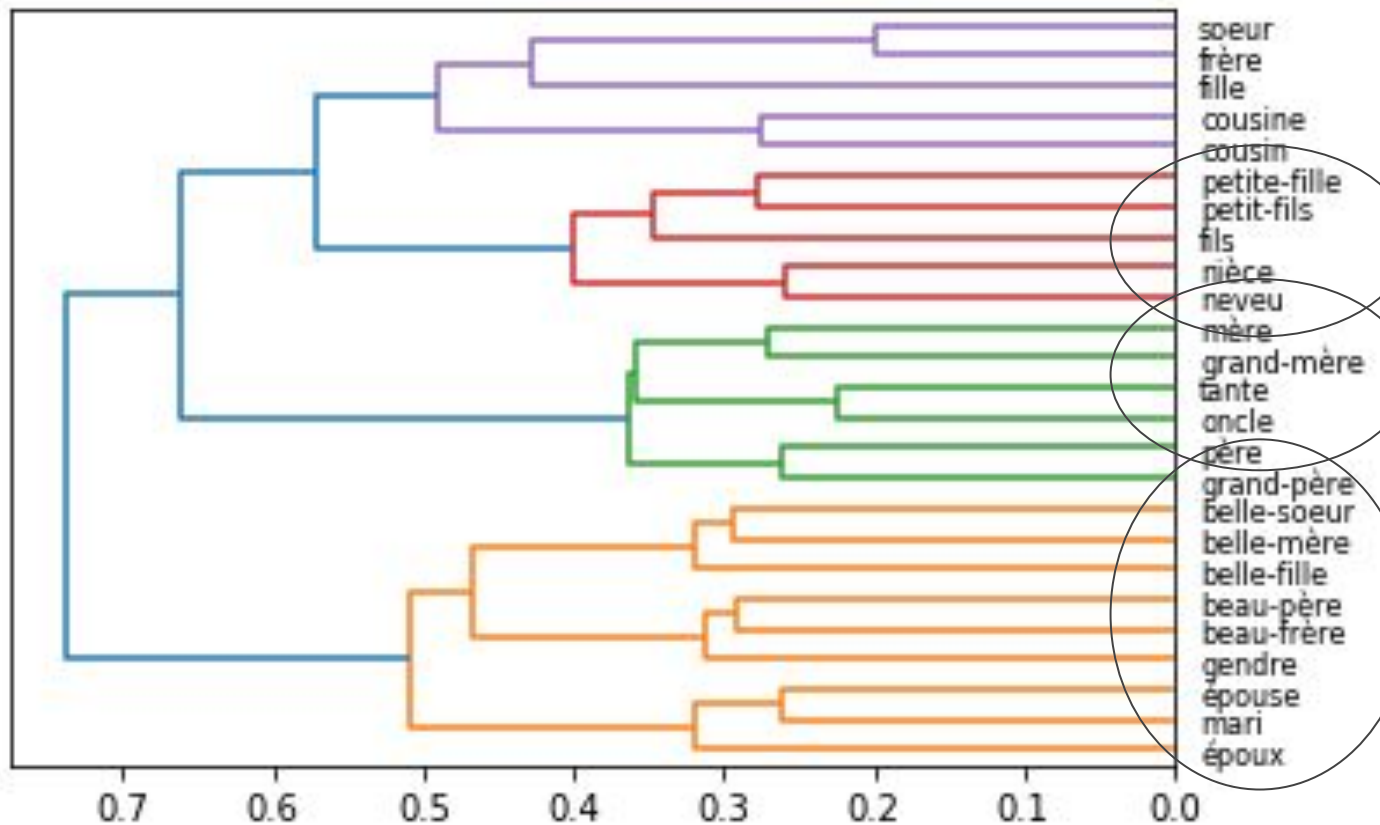
→ plusieurs dimensions
de rapprochement

alliance

ascendance

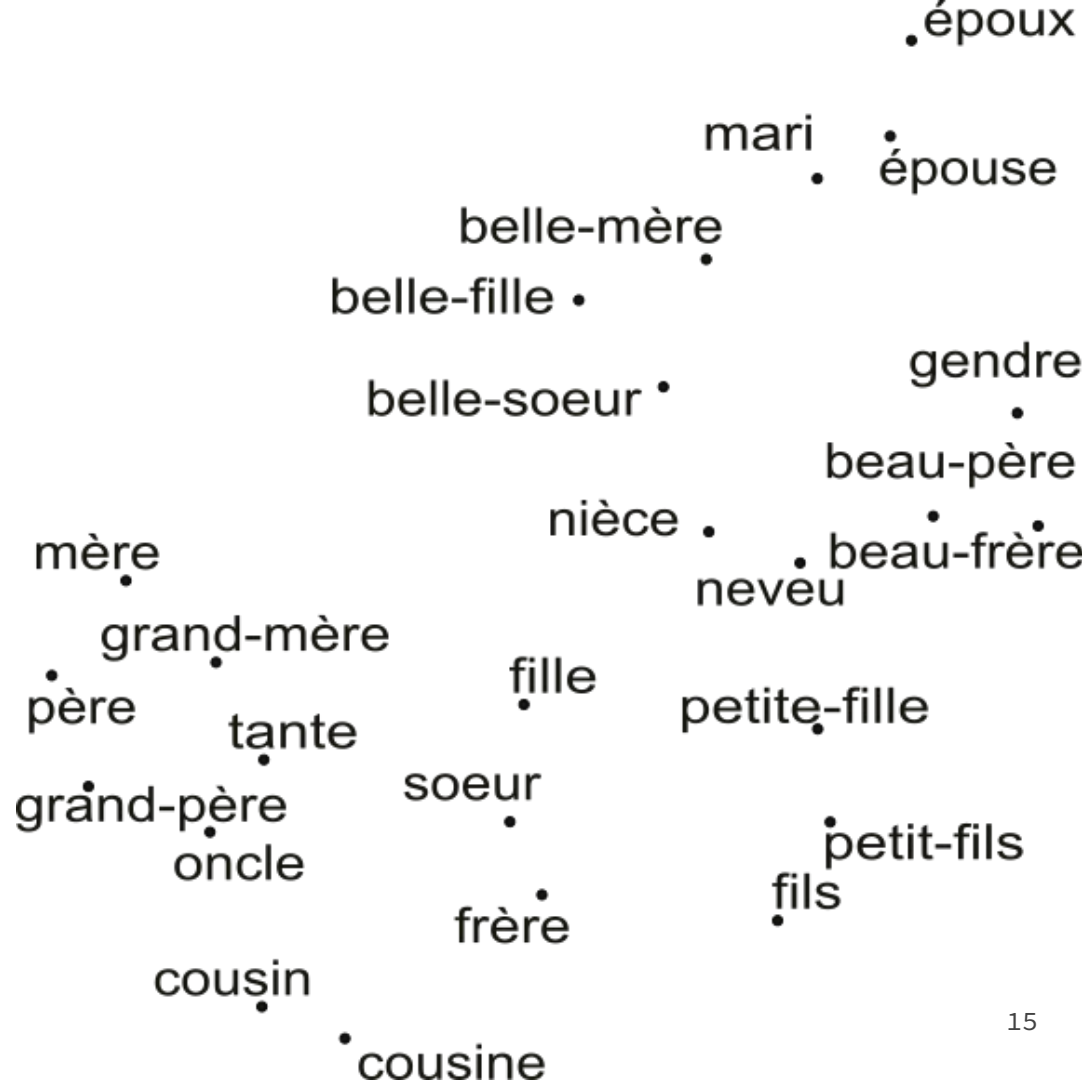
descendance

Quelles priorités entre ces
dimensions ?



Projection t-SNE

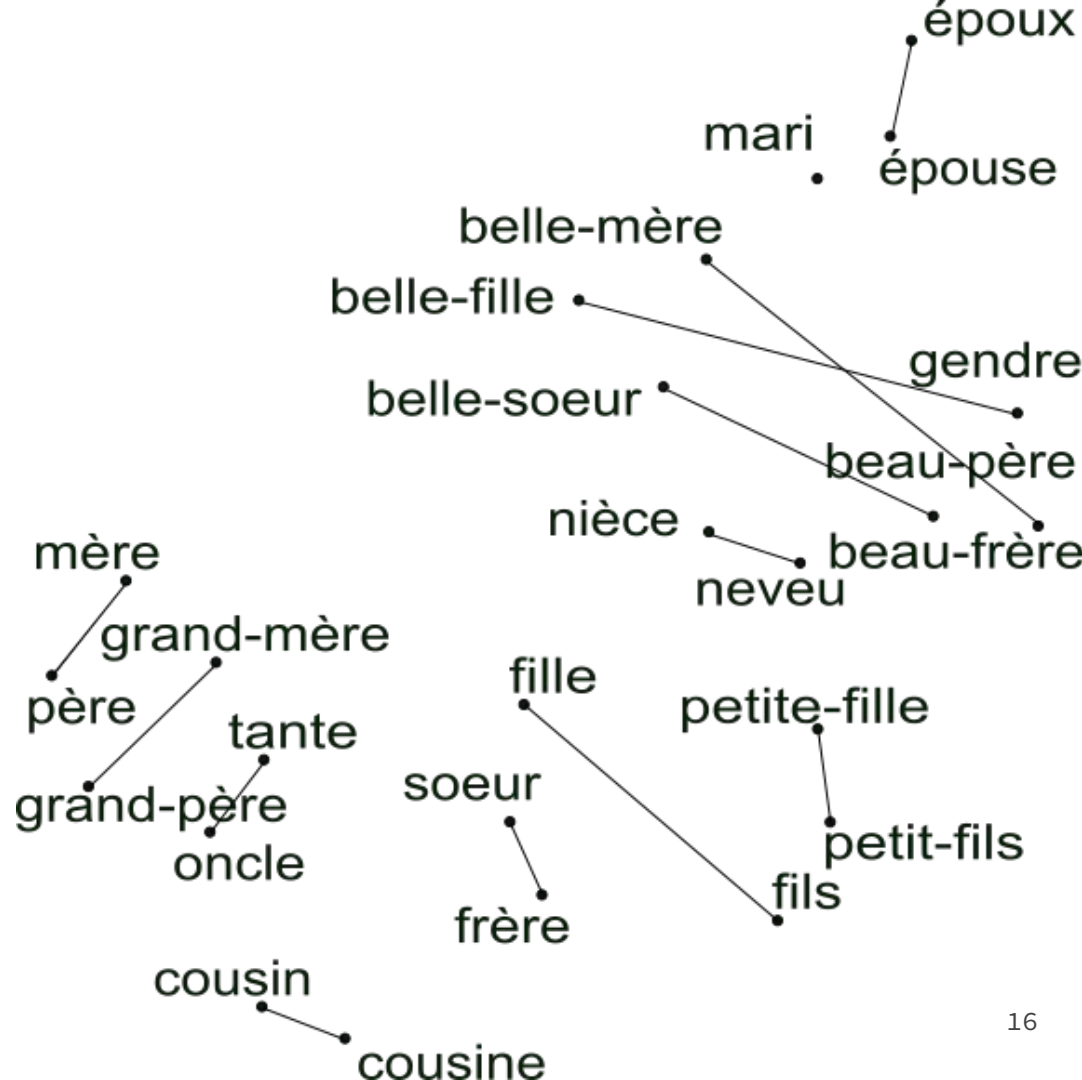
Modèle global, perplexité = 5



Projection t-SNE

Modèle global, perplexité = 5

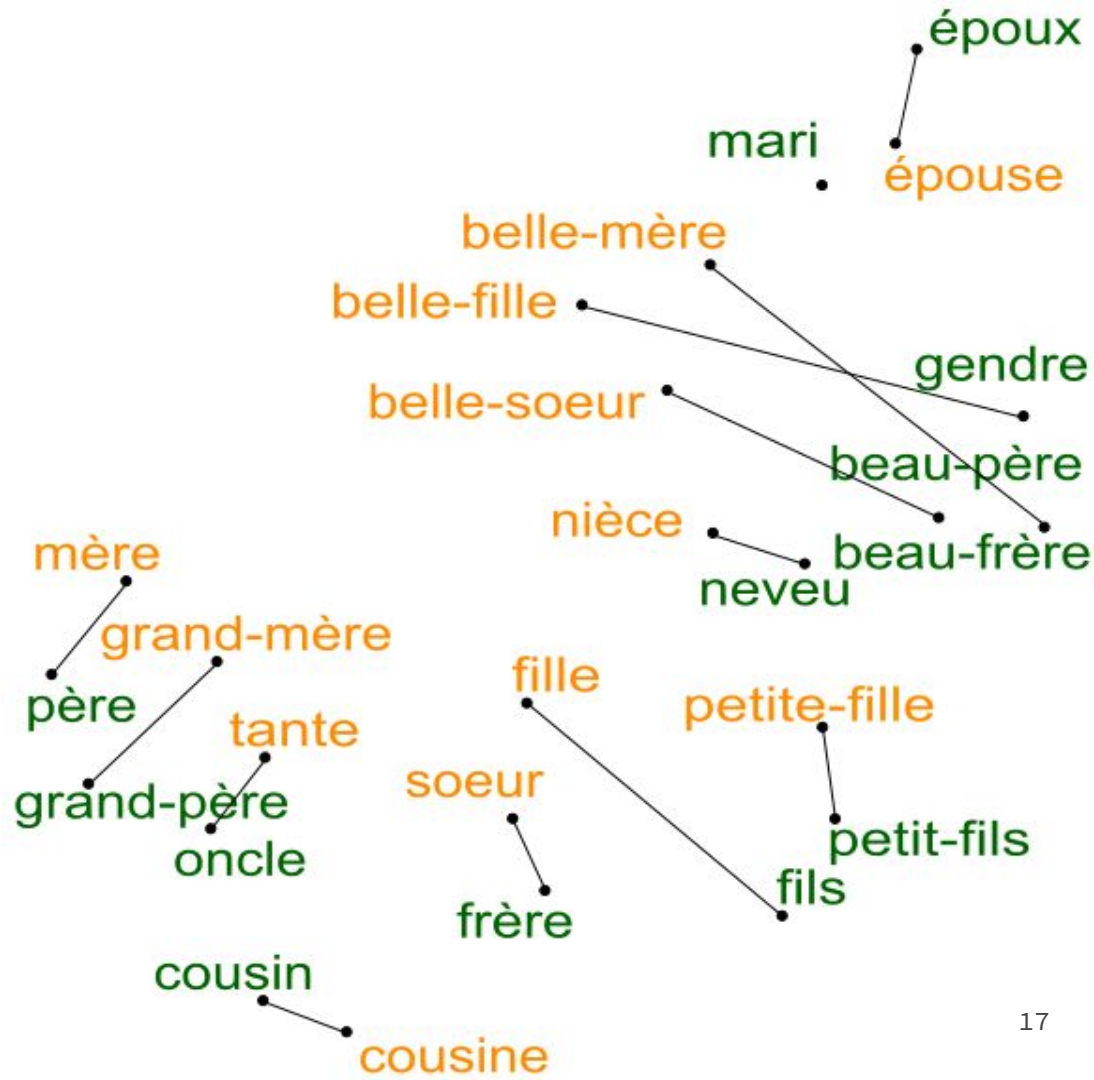
degré de proximité variable
des paires genrées



Projection t-SNE

Modèle global, perplexité = 5

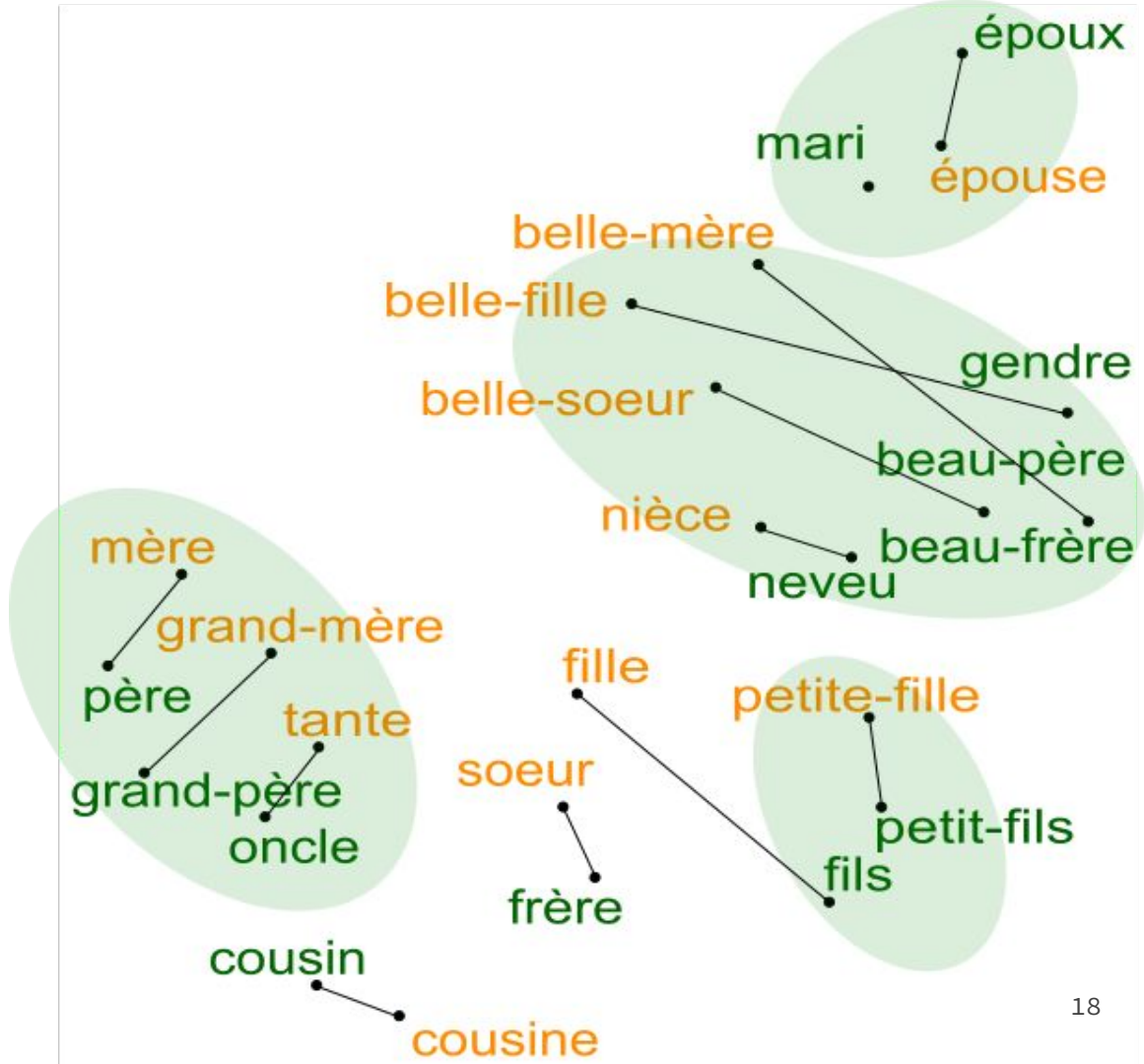
Pas de regroupement évident selon la distinction
féminin masculin



Projection t-SNE

Modèle global, perplexité = 5

des regroupements selon
l'alliance et l'ascendance



Projection t-SNE par corpus

Wikipédia, marquage des genres

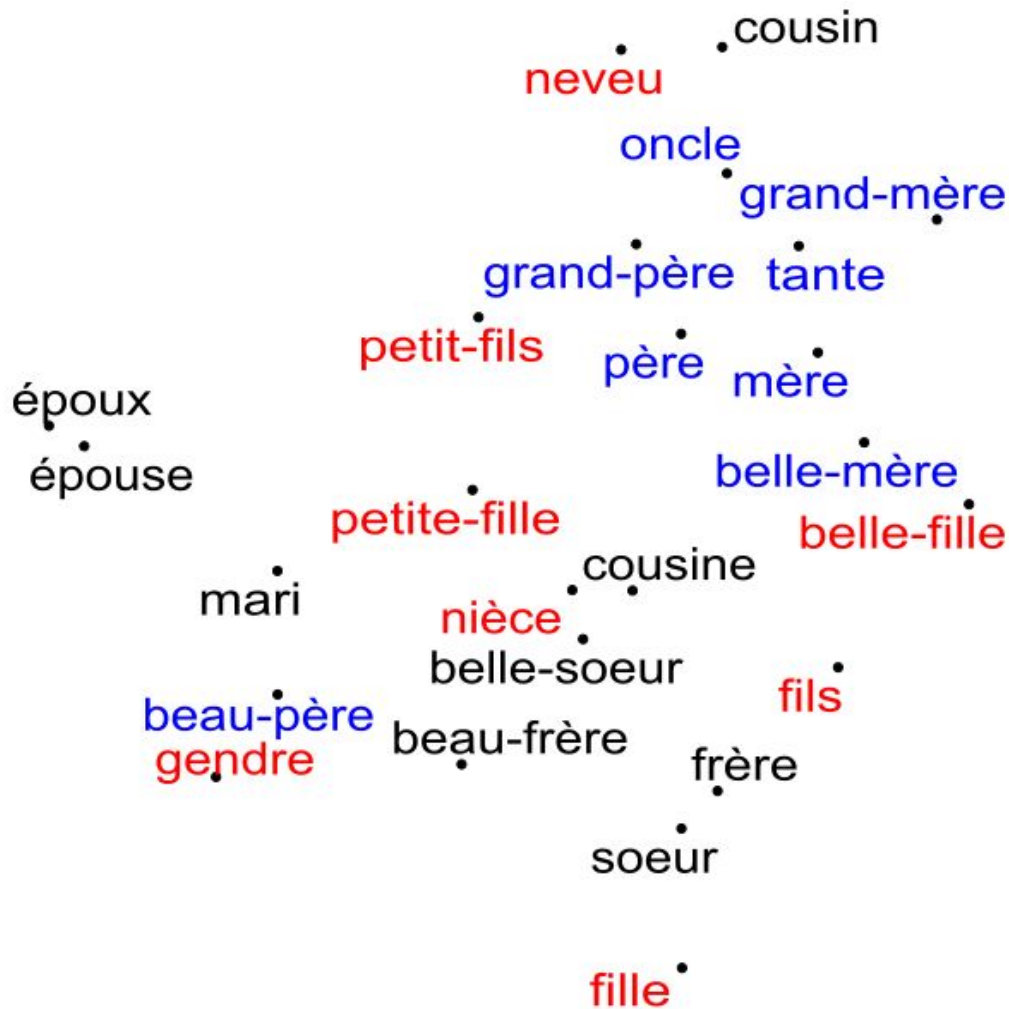
démarcation nette selon le
genre dans ce corpus



Projection t-SNE par corpus

Frantext, marquage des
générations (**ascendants**
descendants)

démarcation plutôt liée à
l'ascendance (hormis *beau-père*)



Croisement avec l'analyse composentielle de la famille

Analyse différentielle

Pour chaque paire d'amorces, calcul des différences sur chaque trait :

père :

genre=M, ascendant=1, descendant=0, germanité=0, alliance=0

grand-mère :

genre=F, ascendant=2, descendant=0, germanité=0, alliance=0

(père, grand-mère) :

genre=1, ascendant=1, descendant=0, germanité=0, alliance=0

Croisement plongements / traits

Pour chaque paire d'amorces (N=300), croiser la décomposition en traits différentiels et la similarité distributionnelle (par corpus)

Régression linéaire multiple : VD=similarité, VI=différences de traits

Corrélation : congruence des deux représentations

(R^2 ajusté du modèle linéaire)

Analyse par trait : part de la dimension de parenté dans la relation distributionnelle

(perte de R^2 suite à l'ablation de la VI)

R² du modèle linéaire et part de chaque trait

Modèle	R ² ajusté	genre	ascendance	descendance	germanité	alliance
Frantext	0,238	29%	35%	19%	17%	20%
FrWaC	0,109	26%	17%	27%	9%	40%
Le Monde	0,375	18%	25%	24%	9%	41%
Revue.org	0,297	33%	16%	25%	15%	30%
Wikipédia	0,363	50%	34%	18%	6%	10%
Global	0,350	32%	26%	22%	11%	28%

Confirmation : les plongements de mots semblent suivre les propriétés organisant les familles, avec des différences entre les corpus

Annotation des voisins humaine et LLM

Annotation des voisins humaine

Annotation humaine des voisins

Objectif : représenter les membres de la famille selon le type de voisins distributionnels qui les entourent dans les différents corpus

- 4 niveaux d'annotation :
 - Lien avec la famille
 - Catégorie/facette utilisée pour désigner l'individu (si voisin désignant un individu)
 - Thématiques générales associées
 - Polarité (connotation positive / négative)
- Observer les variations entre amorces, entre corpus, selon les traits componentiels

Protocole d'annotation des voisins

1. **Sélection des unités à annoter :**
 - a. Extraction des 200 plus proches voisins des 22 amorces obtenus pour modèles/corpus (25)
> **550 m/c i.e listes de voisins (22x25) de 200 mots**
 - b. Sélection des voisins fréquemment extraits : 53% apparaissent dans moins de 40 m/c, "aîné" dans 550 m/c, "bibou" dans 11 m/c
Seuil fixé à une fréquence de 10 i.e. voisins apparaissant dans au moins 10 m/c :
> **N=1173 voisins à annoter**
2. Première **annotation exploratoire** pour dégager des classes de voisins. Idée première :
 - a. Distinguer les voisins impliquant une relation familiale des autres
 - b. Proposer des catégories/thématiques
3. Proposition d'un **premier schéma d'annotation**
4. **(Multi-)Annotation** d'une sous partie des voisins
5. Calcul de l'accord inter-annotateur et première adjudication
6. **Stabilisation des classes**
7. Nettoyage et **mono-annotation** du reste
8. **Adjudication**

Premier schéma d'annotation selon 3 niveaux

Test sur 884 HF pour les catégories et 460 F2 et HF pour les thématiques

niveau	nb.	excl.	Description et tags
type	3	oui	Lien avec la famille F1 (variante d'amorce) : <i>tata, frangin, mère</i> F2 (autre lien familial) : <i>divorce, maternité, illégitime</i> HF (hors famille) : <i>amie, roi, vieille, assassiner, chambellan, lâchement, M, †</i>
catégorie	5	oui	Catégorie/facette utilisée pour désigner un individu HF métier/occupation : <i>gouvernante, comédienne, bedeau, prostituée</i> titre/fonction : <i>tutrice, archiduchesse, tsar, pacha, ex-ministre</i> Caractéristique : <i>marmot, éploré, play-boy, agée, décéder, inconsolable</i> lien/relation : <i>copine, chéri, colocataire, amitié, dévouer, ex-petit</i> <i>NoCat</i>
thématique	??	oui	Thématiques générales associées à F2 ou HF - Proposer une thématique par voisin

Analyse de la première multi-annotation

Fiabilité du schéma d'annotation, accord inter-annot.

Mesure : Alpha de krippendorf (Marzi et al. 2024), supporte des cas d'annotation manquante :

Fusion des types et des catégories, 7 tags exclusifs

F1	F2	métier/occupation	titre/fonction	caractéristique	lien/relation	NoCat
----	----	-------------------	----------------	-----------------	---------------	-------

- 0,761 pour les 303 voisins HF annotés par au moins 2 annotateurs (sur 884)
- 0,795 pour les 39 voisins annotés à 4 annotateurs

Désaccords et décisions pour le Type

- **F1 vs. F2** : on peut dire "mon F1"
Ex : *mes enfants, ma concubine* = F1 vs. ~~mon~~ *marié* = F2
- **F2 vs. HF** : si au moins un des sens implique un lien familial, alors F2
Ex : *maîtresse, naissance, famille, marraine* = F2 vs. *chérie, bien-aimée* = HF

Analyse de la première multi-annotation

Fiabilité du schéma d'annotation, accord inter-annot.

	métier/occ.	titre/fonction	caractéristique	lien/relation	NoCat	
métier/occupation	27	0	1	1	0	29
titre/fonction	7	31	2	2	1	43
caractéristique	1	4	31	5	6	47
lien/relation	3	2	1	14	8	28
NoCat	4	1	5	3	33	46
	42	38	40	25	48	193

Désaccords et décisions pour les catégories

- **métier vs. titre** : tout ce qui est religieux est un titre
- **lien/relation** : test du "X de" (*chef de, tuteur de*)
- **caractéristique** (nécessairement d'un individu) : *richissime, barbu, galanterie*
- **NoCat** : terme qui ne s'applique pas à l'individu : *trône, affectueusement, fugue, hériter*

Analyse de la première multi-annotation

Définition des thématiques

Thématiques

- Alpha de krippendorf = 0.82 sur 133 voisins multi-annotés (sur 460)

Décisions :

- Identification de 14 thématiques : notable/noble, sentiment/amour, religion, enfance, âge, crime/délit, foyer, mort, art, sexe, armée, argent, héritage
- Possibilité de proposer de nouvelles thématiques (éducation, santé, procréation)
- Permettre plusieurs thématiques par voisin
- Indiquer le degré de certitude
- Ajout de l'annotation de la connotation positive / négative

Stabilisation du schéma d'annotation

niveau	nb.	excl.	Description et tags
type	3	oui	Lien avec la famille F1 : relation familiale avec ego (on peut dire "mon F1") F2 : un des sens implique une relation familiale (hors F1) HF : hors famille
catégorie	5	oui	Catégorie/facette utilisée pour désigner un individu HF <ul style="list-style-type: none">- métier/occupation, titre/fonction, caractéristique d'un individu, lien/relation (X de)- <i>NoCat</i> : ne s'applique pas à un individu
thématique	16	non	Thématiques générales associées à F2 ou HF <ul style="list-style-type: none">- notable/noble, sentiment/amour, religion, enfance, âge*, crime/délit, foyer, mort, art, éducation*, sexe, armée, argent, santé, héritage, procréation*- <i>Pas de thématique</i>
polarité	3	oui	Connotation associée à F2 ou HF <ul style="list-style-type: none">- pos, neg- <i>na</i>

Définitions par l'exemple

noblesse/...	aristocrate (titre), ministre (titre)
sentiment/...	chagriner ("na"), antipathique (caractéristique), choupette (lien), lesbien (caractéristique)
religion	coreligionnaire (titre), évêque (titre), église ("na"), Jesus ("na")
enfance	prématuré (caractéristique), peluche ("na"), pouponner ("na"), nourrice (métier)
âge	11an ("na"), veillard (caractéristique), anniv ("na")
crime/délit	assassin (lien), cambrioleur (métier), proxénète (métier)
foyer	bonne (métier), maisonnée ("na")
mort	deuil ("na"), décédé (caractéristique)
art	accordéoniste (métier), romancière (métier), musique ("na")
éducation	<i>Nouvelle thématique</i>
sexe	prostituée (métier), viol ("na"), souteneur (métier), chaste (caractéristique), castrat (métier), eunuque (caractéristique)
armée	capo(titre), régiment("na")
argent	banquier (métier), dot ("na"), milliardaire (caractéristique)
santé	<i>Nouvelle thématique</i>
héritage	légataire (relation), exécuteur (relation), testament ("na")
procréation	<i>Nouvelle thématique</i>

Annotation complète (300 voisins par annotateur)

Mot	notable/noble	religion	sentiment/amour	crime/délit	enfance	art	mort	foyer	âge	armée	sexe	argent	hérita	santé	éducation	procréation	polarité	com
défunt							X											
lévite		X																
bénir		X																pos
pacha	X																	
général-comte	X									X								
nonne		X																
millionnaire												X						
athée		X																
hériter													X					
ravisneur				X														
archidiacre	X	X																
assassiner				X			X											neg
bénédiction		X																pos
conteuse						X												
crucifier		X					X											neg
..																		

- Annotation sans consultation de la fréquence, du type ou de la catégorie
- X sur fond jaune si aucune hésitation
- X sur fond blanc si hésitation
- ? si gros doute

Distribution des types, catégories et thématiques

Types et catégories	#	%	Top 3 (fq m/c)
F1	139	12	aîné, neveu, nièce (550)
F2	56	5	maîtresse, parente, amant (481)
HF	978	83	amie, comtesse, compagne (484)
total	1173		
caractéristique	264	27	prénommer, fillette, éploré (219)
titre/fonction	211	22	comtesse, princesse, comte (449)
NoCat	201	21	épouser, tendrement, assassiner (148)
métier/occupation	153	16	gouvernante, servante, précepteur (366)
lien/relation	149	15	amie, compagne, bien-amiée (381)
total (HF)	978		

Distribution des thématiques

63% des F2+HF avec au moins 1 thématique (#651)

- 52% 1 thématique
- 9% 2 thématiques
- 2% 3 thématiques
- 1 avec 4 thématiques
 - pension (33 m/c) : enfance, foyer, argent, éducation

18% des F2+HF avec une polarité (#184)

- 60% neg, souffre-douleur (138 m/c)
- 40% pos, chère (155 m/c)

Thématiques	#	%	Top 1 (fq m/c)
notable/noble	154	15	maîtresse (525)
sentiment/amour	107	10	amie (536)
religion	101	10	marraine (358)
enfance	70	7	maîtresse (525)
âge	57	6	fillette (287)
crime/délit	47	5	géôlier (194)
foyer	43	4	gouvernante (462)
mort	33	3	orpheline (251)
art	28	3	comédienne (91)
éducation	27	3	maîtresse (525)
sexe	27	3	souteneur (131)
armée	25	2	général-baron (103)
argent	22	2	douairière (218)
santé	21	2	docteur (78)
héritage	16	2	héritier (282)
procréation	12	1	bébé (147)
sans thématique	522	45	parente (513)
total (F2 + HF)	1034		

Annotation des voisins LLM

Annotation linguistique fondée sur les LLM

Nous souhaitons compléter l'annotation des voisins des termes familiaux du français (mots qui apparaissent dans le voisinage du terme de moins de 10 runs)

Nous souhaitons réaliser la même étude sur l'anglais

Motivations :

- améliorer la précision de la caractérisation des termes familiaux du français. Passer de 1173 voisins à 3112 voisins
- confirmer la généralité de la méthode. Peut-on caractériser les termes familiaux de l'anglais à partir de leurs voisins distributionnels , en utilisant les mêmes catégories et les mêmes thématiques ?
- réaliser une étude en diachronie en utilisant le corpus COHA (Corpus of historical American English, 1810-2000, 400 millions de mots) avec l'objectif d'observer l'évolution du sens des termes familiaux.

Mise au point de la méthode sur les données du français

Nous disposons pour le français de données annotées manuellement

La méthode consiste à réaliser l'annotation au moyen de LLM

Il existe 2 types d'études fondées sur les LLM

1. les LLM sont l'objet d'étude. L'objectif est de tester ou de mesurer la qualité de leurs connaissances
2. les LLM sont des outils. L'objectif est de réaliser une tâche le mieux possible, en combinant par exemple les réponses de LLM différents.

Nous avons choisi d'utiliser des petits LLM (coût, souplesse, autonomie, comparaison)

Les annotations comportent des biais induits par l'alignement (= l'adaptation des LLM aux attentes des utilisateurs)

Les biais sont de nature sociale, **culturelle** et pas que linguistique.

LLM testés (ollama)

LLM	paramètres	taille (GB)
nous-hermes2	10.7b	6.1
hermes 3	8.0b	4.7
llama 3	8.0b	4.7
llama 3.1	8.0b	4.9
llama 3.2	3.0b	2.0
mistral-small	22.0b	12.0
mixtral	8 x 7.0b	26.0

Des prompts en anglais et en français

Les modèles sont **multilingues**.

- prompts en français (= dans la même langue que les données)
- prompts en anglais (possiblement plus “intelligibles” aux modèles)

Méthode “few-shots” avec présentation d'exemples d'annotation pour chacune des étiquettes (catégories et thématiques)

Tâches

Nous avons créé des prompts différents pour chaque tâche :

1. annotation séparée des catégories F1 et F2
F1 = termes familiaux directement ou indirectement en relation avec EGO
F2 = termes qui impliquent une relation familiale avec EGO, à l'exception des termes F1
2. annotation des termes F1 et F2 au moyen de la même étiquette "famille"
3. annotation des thématiques (hors voisins F1 et F2)

Exemple de prompt

Je suis un lexicographe français.

Je travaille sur les noms des membres de la famille en français comme "père", "oncle", "cousin", "petite-fille".

Je voudrais que vous m'aidiez à caractériser un ensemble de voisins distributionnels des noms.

Vous devez annoter un ensemble de mots du français, en utilisant un ensemble de catégories prédéfinies dont la liste est donnée plus loin.

Chaque mot peut appartenir à une seule catégorie.

Les sorties doivent être présentées de la manière suivante : le mot suivi d'un deux point (:) suivi de la catégorie.

Si aucune catégorie ne convient, vous devez indiquer "néant".

Vous ne devez indiquer aucune catégorie qui ne soit pas dans la liste.

Vous ne devez ajouter aucun commentaire ni aucune information supplémentaire.

Par exemple, si le mot est "mère", vous devez produire :

mère : famille

Exemple de prompt (suite)

Voici la liste des catégories :

- * "famille"
- * "caractéristique"
- * "titre"
- * "fonction"
- * "métier"
- * "occupation"
- * "lien"
- * "relation"
- * "nocat"

Exemple de prompt (suite)

"famille" est la catégorie des mots que l'on peut utiliser pour désigner un membre de la famille et des mots qui impliquent l'existence d'un lien familial. Exemples : "père", "frère", "veuf", "meuf", "aîné", "ex-femme", "génitrice", "veuve", "meuf", "héritage", "adultère", "maternité"

"caractéristique" est la catégorie de mots comme : "richissime", "barbu", "breton", "fainéant"

"titre" est la catégorie de mots comme : "pape", "comte", "député", "vice-président"

"fonction" est la catégorie de mots comme : "brigadier", "sacrificateur"

"métier" est la catégorie de mots comme : cuisinière, accordéoniste, conseiller

"occupation" est la catégorie de mots comme : cambrioleur, tueur, gigolo

"lien" est la catégorie de mots comme : "amant", "chéri"

"relation" est la catégorie de mots comme : "voisin", "assassin", "concurrent", "copropriétaire", "collège"

"nocat" est la catégorie des mots qui ne se réfèrent pas à des personnes comme : "deuil", "église", "11 ans", "1912", "chagriner"

Voici les mots que vous devez annoter, un par ligne :

Exemple de réponse

ex-président : politique/héritage

ancêtre : héritage/famille

ange : religion/sentiment/amour

secourir : santé/procréation

instigation : crime/délit

mère-grand : famille/enfance

fermière : professionnel/foyer

kidnappeur : crime/délit

compagne : sentimental/amour

grand- : noblesse/notabilité

ex-maire : politique/héritage

régente : administration/héritage

mere : famille/enfance

filleule : enfance/famille

vizir : politique/héritage

maternité : procréation/santé

prénommer : culture/littérature

Répéter les interrogations pour réduire la variabilité

Les réponses des modèles varient d'un run à l'autre pour un même modèle et un même prompt

Nous avons réalisé **10 runs** pour chaque modèle et chaque prompt

Nous avons réalisé un post-traitement pour extraire les réponses

Exemples de réponse

gérante : femme/équipe, noble
impératrice : notable/noble, religion
prodigue : noble
curé : prêtre
ex-collègue : confrère
cousine : famille
descendance : héritage
ouaille : enfance
epouse : foyer/mariage, religion
concierge : foyer
morganatique : noble
franc-maçon : société
rondouillard : prêtre

compère : ami
ménage : foyer
officiellement : administration
assassinat : crime/délit
espionne : armée
natif : famille/peuple
passagère : voyage
prophète : religion
télégraphier : communication

Exemples de sorties

notaire : justice (pas dans la liste, mais je l'ai ajoutée pour cette réponse)

couvent : lieu de culte (removed, not in list) -> nocat

32. bénédictin : métier

['colleague': notable, éducation]

['enterrement': nocat, 'ourson': caractéristique, 'demoiselle': famille, 'rondouillard': caractéristique, 'roturier': lien, 'vieille': caractéristique, 'neveu': famille, 'putain': lien, 'fugueur': occupation, 'ennemie': relation, 'bréviaire': nocat, 'milliardaire': caractéristique, 'prénommer': nocat, 'lettré': caractéristique, 'lycéenne': famille, [...]

"m'sieur" : nocat

["m'sieur": contraction of "me" + "sieur")

Mesures

Annotation par vote majoritaire (méthode “ensemble”; majorité relative) sur 1 à 10 runs

Nous utilisons l’exactitude pour l’annotation des catégories

Nous utilisons le rappel, la précision et la mesure F1 pour l’annotation des catégories (vote majoritaire, majorité absolue)

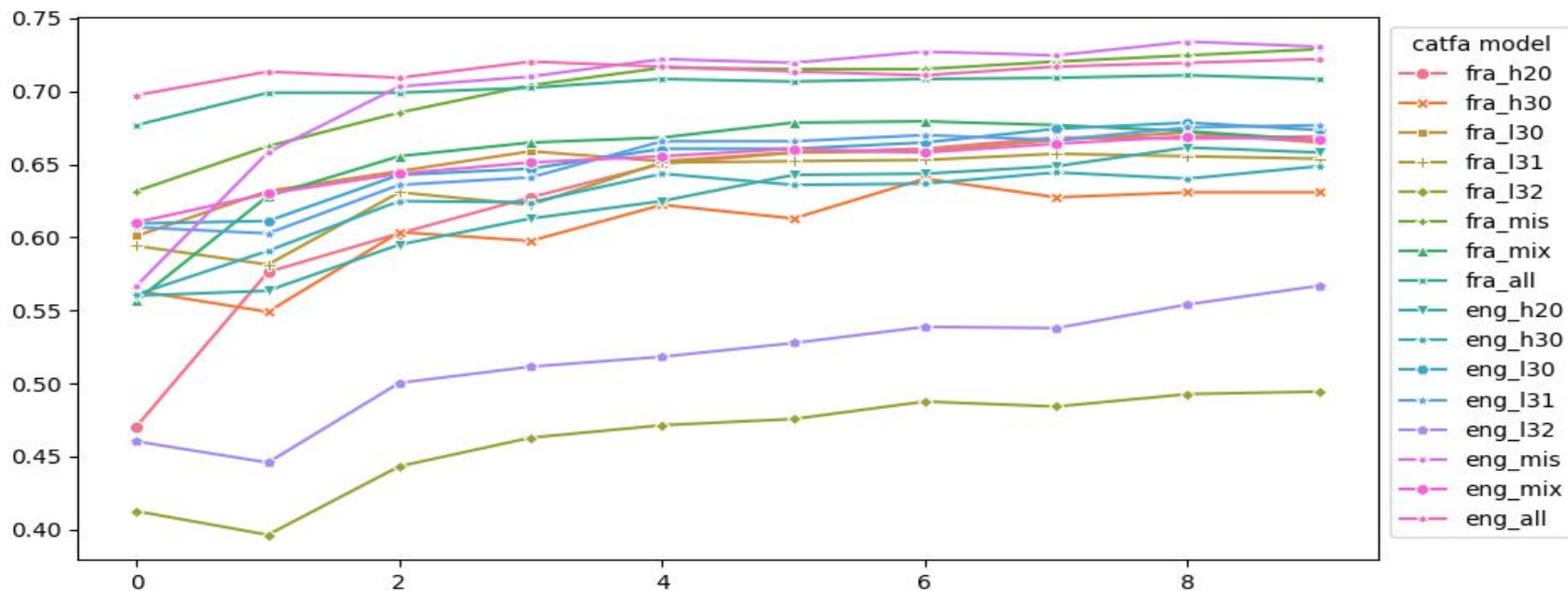
Nous utilisons le rappel, la précision et la mesure F1 pour l’annotation des thématiques.

- Les mesures sont calculées séparément pour chaque thématique
- F1 moyenne (= macro-average sur les LLM)

Exactitude pour l'annotation de catégories (famille)

0	0.471	0.564	0.601	0.594	0.413	0.632	0.557	0.677	0.560	0.561	0.610	0.607	0.460	0.567	0.610	0.697
1	0.576	0.549	0.632	0.581	0.396	0.662	0.628	0.699	0.564	0.591	0.611	0.603	0.446	0.658	0.630	0.714
2	0.603	0.604	0.645	0.631	0.443	0.685	0.656	0.699	0.595	0.625	0.643	0.636	0.500	0.703	0.644	0.709
3	0.627	0.598	0.659	0.622	0.463	0.704	0.665	0.702	0.613	0.624	0.647	0.641	0.512	0.710	0.651	0.720
4	0.650	0.622	0.652	0.651	0.471	0.716	0.668	0.708	0.625	0.644	0.661	0.666	0.518	0.722	0.656	0.717
5	0.658	0.613	0.658	0.652	0.476	0.715	0.679	0.707	0.643	0.636	0.661	0.666	0.528	0.720	0.661	0.714
6	0.661	0.640	0.659	0.653	0.488	0.715	0.679	0.708	0.644	0.637	0.665	0.670	0.539	0.727	0.658	0.711
7	0.668	0.627	0.667	0.657	0.484	0.720	0.677	0.709	0.649	0.645	0.674	0.667	0.538	0.725	0.664	0.717
8	0.668	0.631	0.672	0.656	0.493	0.725	0.673	0.711	0.662	0.640	0.679	0.675	0.554	0.734	0.669	0.720
9	0.669	0.631	0.665	0.654	0.494	0.729	0.667	0.708	0.658	0.649	0.673	0.677	0.567	0.731	0.667	0.722
	fra_h20	fra_h30	fra_l30	fra_l31	fra_l32	fra_mis	fra_mix	fra_all	eng_h20	eng_h30	eng_l30	eng_l31	eng_l32	eng_mis	eng_mix	eng_all

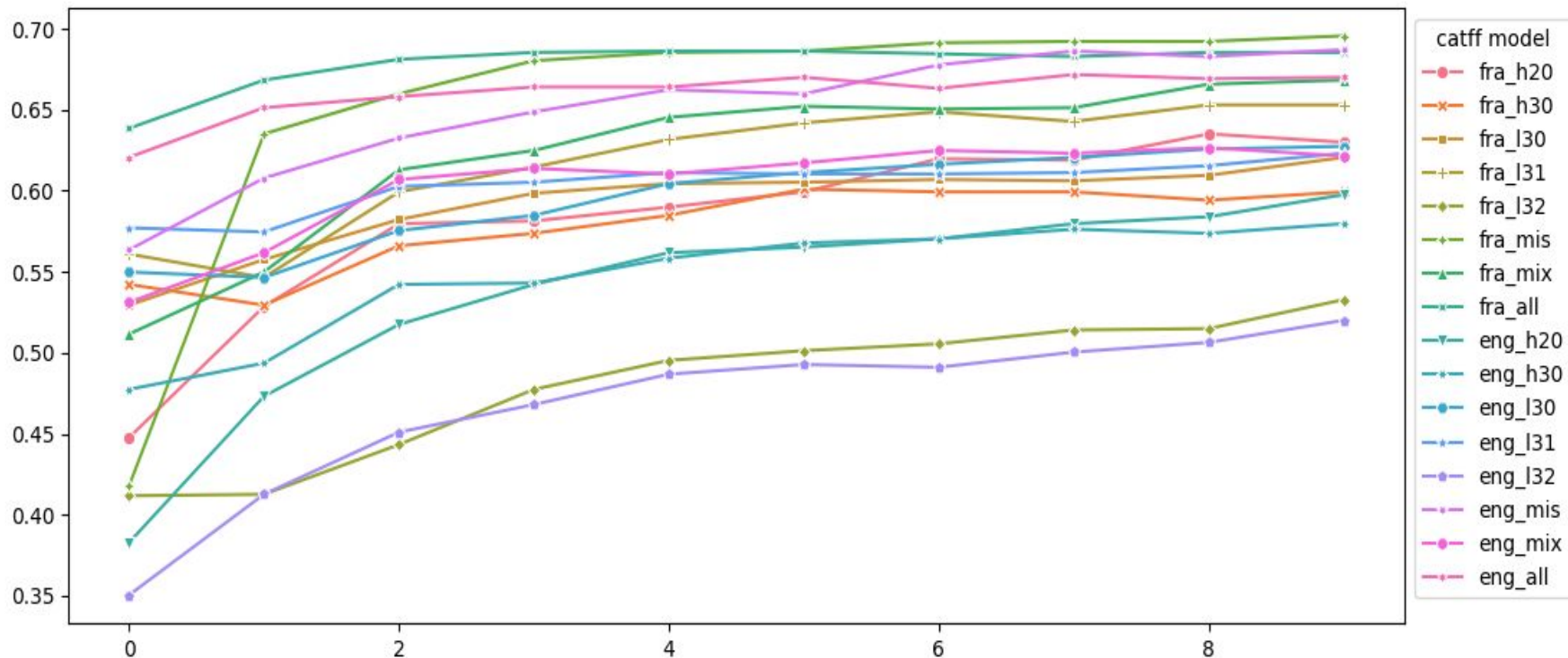
Exactitude pour l'annotation de catégories (famille)



Exactitude pour l'annotation de catégories (F1, F2)

0	0.448	0.542	0.529	0.561	0.412	0.418	0.512	0.639	0.383	0.477	0.550	0.577	0.350	0.564	0.531	0.621
1	0.529	0.529	0.558	0.546	0.413	0.635	0.550	0.668	0.473	0.494	0.546	0.575	0.413	0.608	0.562	0.651
2	0.580	0.566	0.582	0.599	0.443	0.660	0.613	0.681	0.517	0.542	0.575	0.603	0.451	0.633	0.607	0.658
3	0.581	0.574	0.598	0.615	0.477	0.680	0.625	0.685	0.542	0.543	0.585	0.605	0.468	0.649	0.614	0.664
4	0.590	0.585	0.604	0.632	0.495	0.685	0.645	0.686	0.562	0.558	0.604	0.611	0.487	0.662	0.610	0.664
5	0.599	0.601	0.605	0.642	0.501	0.686	0.652	0.686	0.565	0.568	0.611	0.610	0.493	0.660	0.617	0.670
6	0.620	0.599	0.607	0.649	0.506	0.691	0.650	0.685	0.570	0.570	0.616	0.610	0.491	0.678	0.625	0.663
7	0.619	0.599	0.606	0.643	0.514	0.692	0.651	0.683	0.580	0.576	0.621	0.611	0.500	0.686	0.623	0.672
8	0.635	0.594	0.610	0.653	0.515	0.692	0.666	0.685	0.584	0.574	0.626	0.616	0.506	0.683	0.627	0.669
9	0.630	0.599	0.621	0.653	0.533	0.696	0.668	0.685	0.598	0.580	0.627	0.623	0.520	0.687	0.621	0.670
	fra_h20	fra_h30	fra_l30	fra_l31	fra_l32	fra_mis	fra_mix	fra_all	eng_h20	eng_h30	eng_l30	eng_l31	eng_l32	eng_mis	eng_mix	eng_all

Exactitude pour l'annotation de catégories (F1, F2)



Matrice de confusion pour fra_f1f2_mis_9

actual/LLM	caractéristique	f1	f2	lien/relation	métier/occupation	nocat	titre/fonction
caractéristique	166	8	1	4	20	54	11
f1	1	94	15	23	0	5	1
f2	10	6	13	11	0	10	6
lien/relation	9	2	4	72	19	21	22
métier/occupation	2	1	0	1	135	0	14
nocat	11	2	4	4	7	165	8
titre/fonction	5	3	0	2	15	12	174

Annotation des catégories

Il est important de réaliser plusieurs runs

La distinction F1-F2 est difficile à capter

Les gros modèles sont les plus performants

L'anglais n'améliorent pas toujours les performances

Annotation catégorie

Si l'on rejette les réponses des modèles données dans de 50% des runs, ainsi que les réponses "néant",

une partie des voisins perdent leur annotation

on calcule une précision, un rappel et une mesure F1 pour l'annotation de catégories (micro-moyenne)

Rappel pour l'annotation de catégories (famille)

0	0.471	0.564	0.601	0.594	0.413	0.632	0.557	0.580	0.560	0.561	0.610	0.607	0.460	0.567	0.610	0.592
1	0.576	0.549	0.632	0.581	0.396	0.662	0.628	0.608	0.564	0.591	0.611	0.603	0.446	0.658	0.630	0.621
2	0.521	0.563	0.604	0.608	0.411	0.616	0.596	0.577	0.541	0.596	0.611	0.608	0.437	0.637	0.602	0.590
3	0.594	0.582	0.641	0.616	0.451	0.689	0.634	0.596	0.596	0.621	0.639	0.636	0.490	0.691	0.637	0.608
4	0.554	0.560	0.598	0.606	0.384	0.654	0.580	0.583	0.548	0.595	0.623	0.621	0.421	0.662	0.604	0.595
5	0.609	0.579	0.631	0.630	0.431	0.672	0.624	0.578	0.606	0.621	0.649	0.645	0.475	0.693	0.631	0.608
6	0.540	0.563	0.598	0.607	0.397	0.642	0.582	0.574	0.562	0.590	0.624	0.621	0.420	0.667	0.606	0.604
7	0.604	0.583	0.628	0.634	0.424	0.676	0.617	0.579	0.602	0.614	0.651	0.640	0.462	0.693	0.627	0.614
8	0.560	0.559	0.607	0.611	0.388	0.633	0.593	0.572	0.559	0.596	0.633	0.619	0.429	0.673	0.615	0.606
9	0.596	0.579	0.620	0.634	0.412	0.665	0.607	0.577	0.598	0.610	0.650	0.637	0.456	0.691	0.628	0.611
	fra_h20_recall	fra_h30_recall	fra_l30_recall	fra_l31_recall	fra_l32_recall	fra_mis_recall	fra_mix_recall	fra_all_recall	eng_h20_recall	eng_h30_recall	eng_l30_recall	eng_l31_recall	eng_l32_recall	eng_mis_recall	eng_mix_recall	eng_all_recall

Nombre de réponse

	réponses
fra_h20	935
fra_h30	982
fra_l30	1025
fra_l31	1044
fra_l32	814
fra_mis	998
fra_mix	980
fra_all	818
eng_h20	951
eng_h30	1049
eng_l30	1068
eng_l31	1039
eng_l32	792
eng_mis	1050
eng_mix	1033
eng_all	863

Précision pour l'annotation de catégories (famille)

0	0.594	0.602	0.635	0.635	0.454	0.682	0.657	0.778	0.599	0.590	0.647	0.635	0.512	0.689	0.667	0.800
1	0.596	0.561	0.638	0.587	0.404	0.671	0.652	0.778	0.573	0.597	0.618	0.608	0.461	0.690	0.647	0.798
2	0.694	0.673	0.695	0.684	0.559	0.735	0.720	0.811	0.670	0.679	0.689	0.704	0.614	0.762	0.710	0.824
3	0.665	0.627	0.678	0.643	0.530	0.725	0.697	0.815	0.653	0.642	0.662	0.665	0.572	0.738	0.682	0.813
4	0.735	0.687	0.720	0.712	0.597	0.762	0.732	0.829	0.712	0.703	0.719	0.729	0.667	0.781	0.729	0.833
5	0.716	0.652	0.688	0.679	0.559	0.756	0.723	0.826	0.698	0.668	0.687	0.703	0.615	0.751	0.702	0.821
6	0.760	0.709	0.718	0.730	0.622	0.783	0.744	0.829	0.740	0.697	0.727	0.740	0.673	0.784	0.732	0.834
7	0.745	0.678	0.707	0.698	0.584	0.763	0.725	0.829	0.723	0.679	0.709	0.705	0.644	0.764	0.711	0.829
8	0.774	0.718	0.733	0.733	0.624	0.793	0.754	0.837	0.754	0.706	0.732	0.744	0.687	0.792	0.736	0.833
9	0.748	0.691	0.709	0.713	0.593	0.782	0.727	0.828	0.738	0.682	0.714	0.719	0.676	0.772	0.713	0.831
	fra_h20_precision	fra_h30_precision	fra_l30_precision	fra_l31_precision	fra_l32_precision	fra_mis_precision	fra_mix_precision	fra_all_precision	eng_h20_precision	eng_h30_precision	eng_l30_precision	eng_l31_precision	eng_l32_precision	eng_mis_precision	eng_mix_precision	eng_all_precision

Mesure F1 pour l'annotation de catégories (famille)

0	0.525	0.582	0.617	0.614	0.433	0.656	0.603	0.664	0.579	0.575	0.628	0.620	0.485	0.622	0.637	0.681
1	0.586	0.555	0.635	0.584	0.400	0.667	0.640	0.682	0.568	0.594	0.615	0.606	0.453	0.674	0.638	0.698
2	0.595	0.613	0.647	0.644	0.474	0.670	0.652	0.674	0.599	0.635	0.648	0.652	0.511	0.694	0.652	0.688
3	0.628	0.604	0.659	0.630	0.487	0.707	0.664	0.688	0.623	0.631	0.650	0.650	0.528	0.714	0.658	0.696
4	0.632	0.617	0.654	0.655	0.467	0.704	0.647	0.685	0.619	0.645	0.668	0.671	0.516	0.717	0.660	0.694
5	0.658	0.613	0.658	0.653	0.487	0.712	0.670	0.680	0.649	0.644	0.668	0.673	0.536	0.721	0.665	0.699
6	0.631	0.627	0.653	0.663	0.485	0.705	0.653	0.678	0.639	0.639	0.672	0.676	0.517	0.721	0.663	0.700
7	0.667	0.627	0.665	0.665	0.491	0.717	0.667	0.682	0.657	0.645	0.679	0.671	0.538	0.727	0.667	0.706
8	0.650	0.629	0.664	0.667	0.478	0.704	0.664	0.679	0.642	0.646	0.679	0.676	0.528	0.728	0.670	0.702
9	0.663	0.630	0.662	0.671	0.486	0.719	0.661	0.680	0.661	0.644	0.681	0.675	0.545	0.730	0.668	0.704
	fra_h20_f1 --	fra_h30_f1 --	fra_l30_f1 --	fra_l31_f1 --	fra_l32_f1 --	fra_mis_f1 --	fra_mix_f1 --	fra_all_f1 --	eng_h20_f1 --	eng_h30_f1 --	eng_l30_f1 --	eng_l31_f1 --	eng_l32_f1 --	eng_mis_f1 --	eng_mix_f1 --	eng_all_f1 --

Rappel, précision, F1 pour l'annotation des catégories

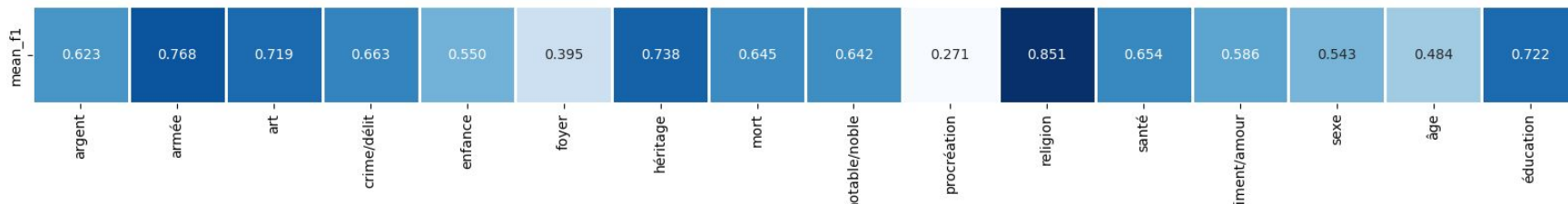
Les prompts en anglais donnent de meilleurs résultats

Selon que l'on souhaite privilégier la précision, le rappel ou la mesure F1, le modèle / le cumul de modèles à utiliser n'est pas le même

Les modèles permettent un dégrossissage de l'annotation, mais une révision manuelle reste nécessaire

Les résultats obtenus ne tiennent pas compte de la variabilité des annotations des humains, du fait de la variabilité dans l'interprétation des labels des catégories.

Thématique. F1 moyenne (sur l'ensemble des LLM)



Les performances sont variables.

Certaines thématiques sont mal captées par les modèles

Il faudrait revoir le jeu de thématiques

Une solution serait de laisser les modèles proposer leurs propres thématiques et de sélectionner un jeu de thématiques à partir de leurs réponses.

Support des thématiques

notable/noble	154
religion	101
sentiment/amour	107
crime/délit	47
enfance	70
art	28
mort	33
foyer	43
âge	57
armée	25
sexe	27
argent	22
héritage	16
santé	21
éducation	27
procréation	12

Thématiques. Discussion

Envisager une annotation manuelle qui consiste à réviser l'annotation automatique. Quel serait l'accord inter-annotateurs pour la tâche de révision ?

Envisager de caractériser les termes de la famille en utilisant les thématiques proposées par les modèles

Envisager d'avoir une gradation dans l'appartenance d'un voisin à une thématique (en fonction du nombre de runs dans lesquels elle a été proposées)

Analyse de la famille : Caractérisation sémantique des amorces

Analyse Composantes Principales (ACP)

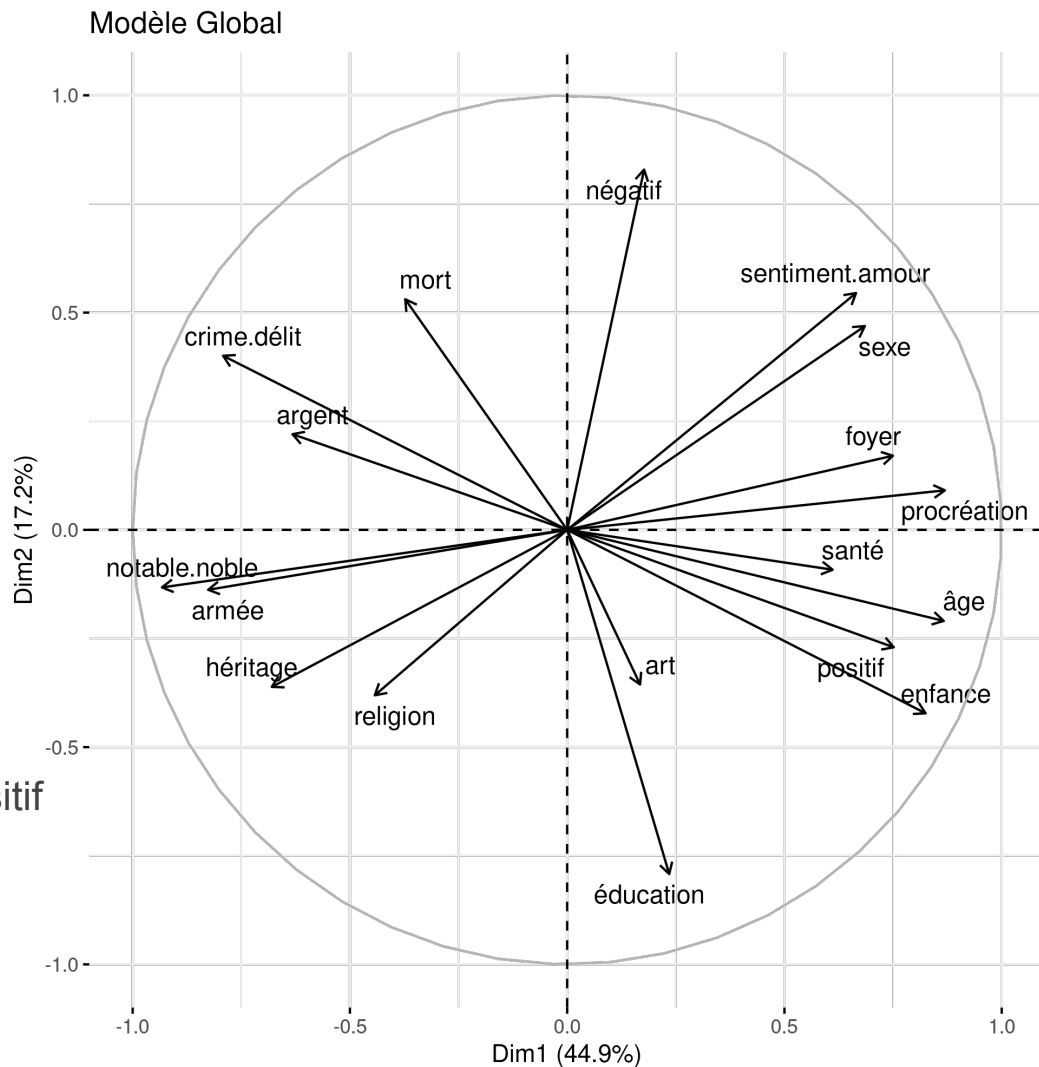
- Par corpus et pour le modèle global
 - Regroupement des informations annotées pour chaque amorce :
 - Poids cumulé des types, catégories, thématiques et polarité des voisins
 - Pondéré par le nombre de modèles
- Analyse des annotations humaines adjudiquées et LLM
 - ACP thématiques (non exclusives) et polarité
 - ACP types et catégories (exclusives)
 - Projection des amorces
 - Projection des traits de décomposition

Thématiques et polarité

Dimension 1 :

- armée, noble, héritage, crime, argent, religion, mort

- Âge, procréation, enfance, sentiment, sexe, foyer, santé, positif

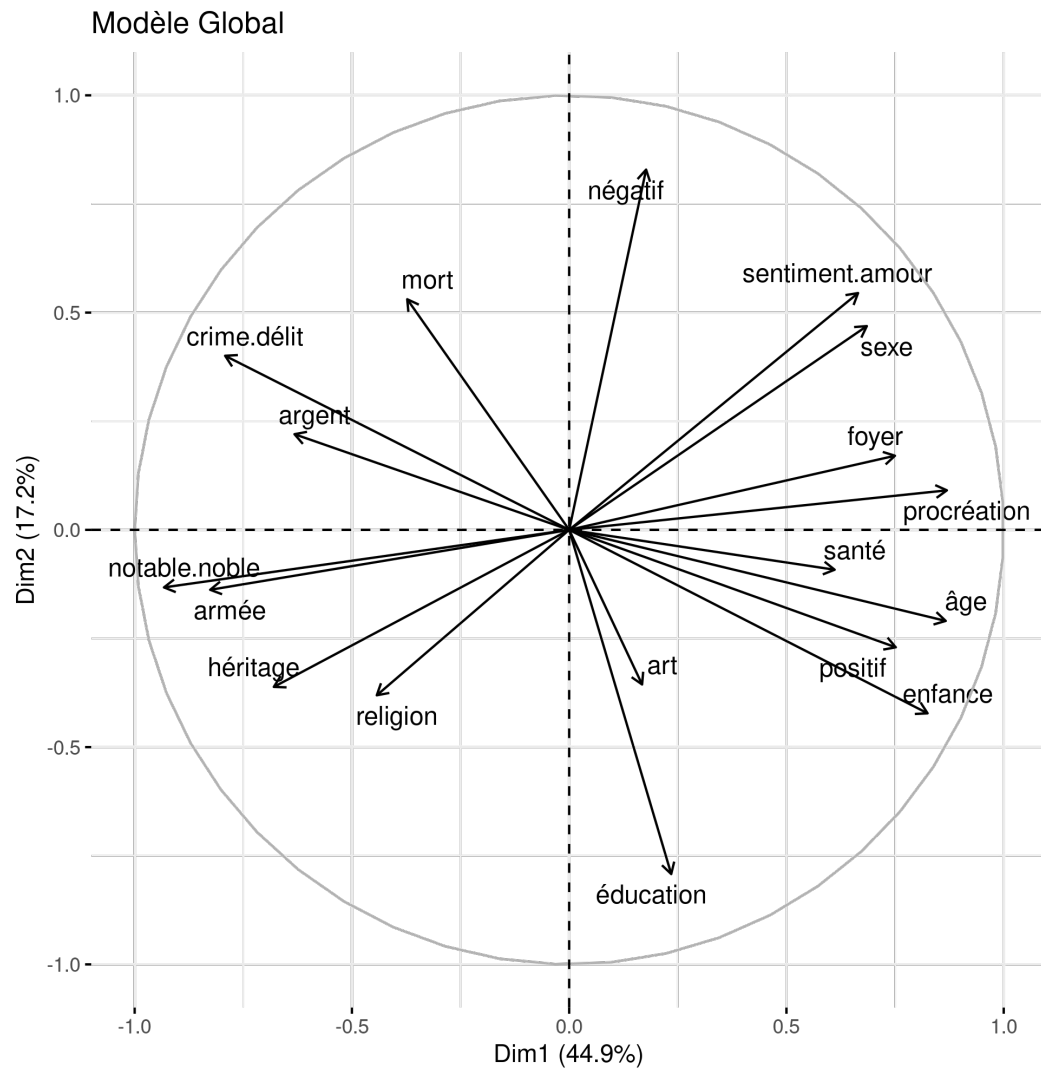


Thématiques et polarité

Dimension 2 :

- Négatif, sentiment, mort, sexe, crime

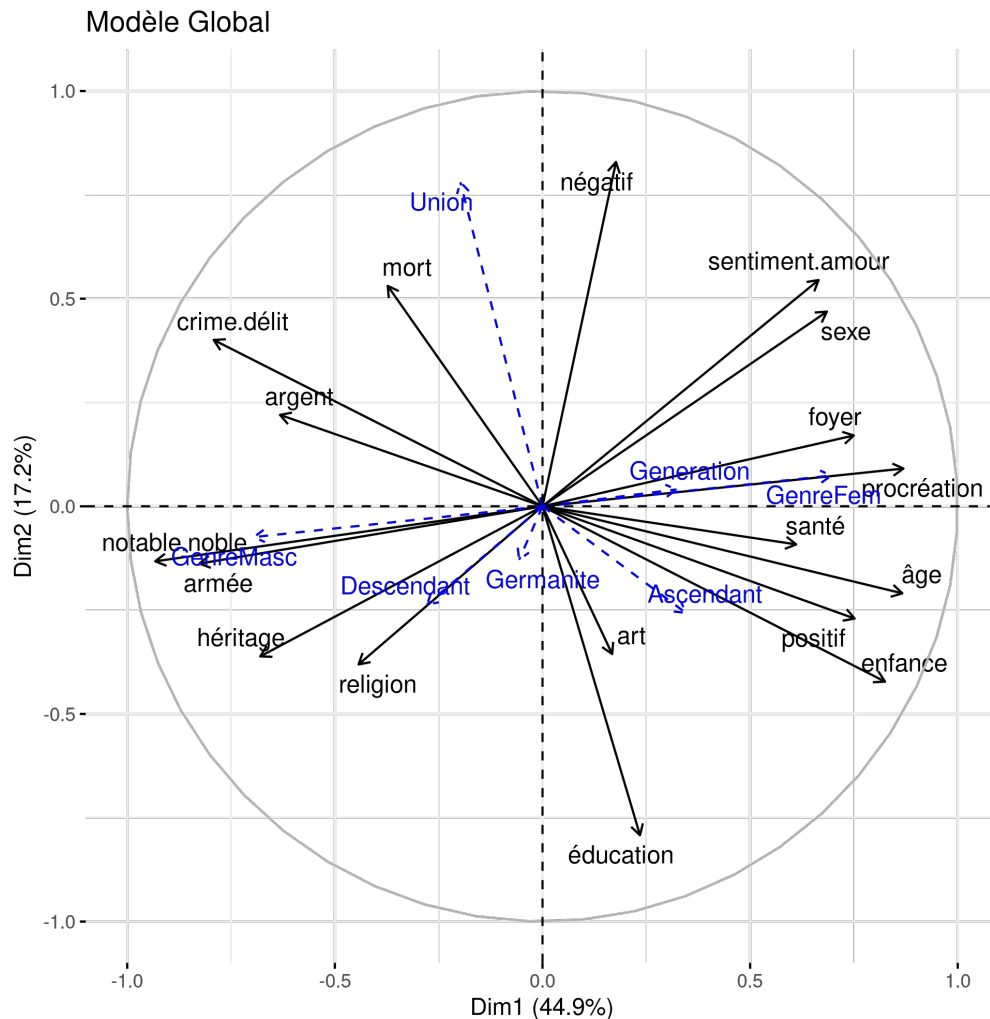
- Éducation, religion, enfance, art, positif



Thématiques et polarité

Dimension 1 : masculin vs. féminin

Dimension 2 : union vs ascendance



Types et catégories

Dimension 1 :

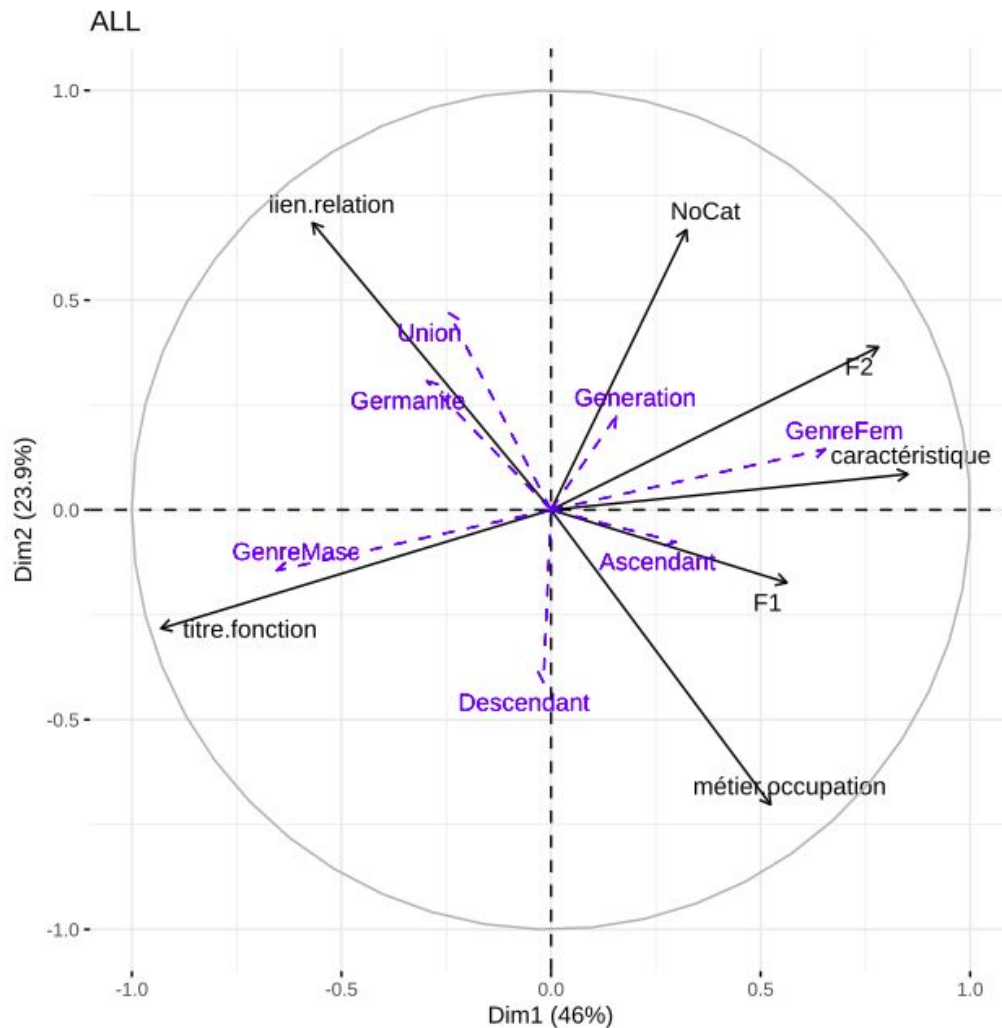
masculin vs. féminin

titre.fonction vs. caractéristique +
famille (F1 + F2)

Dimension 2 :

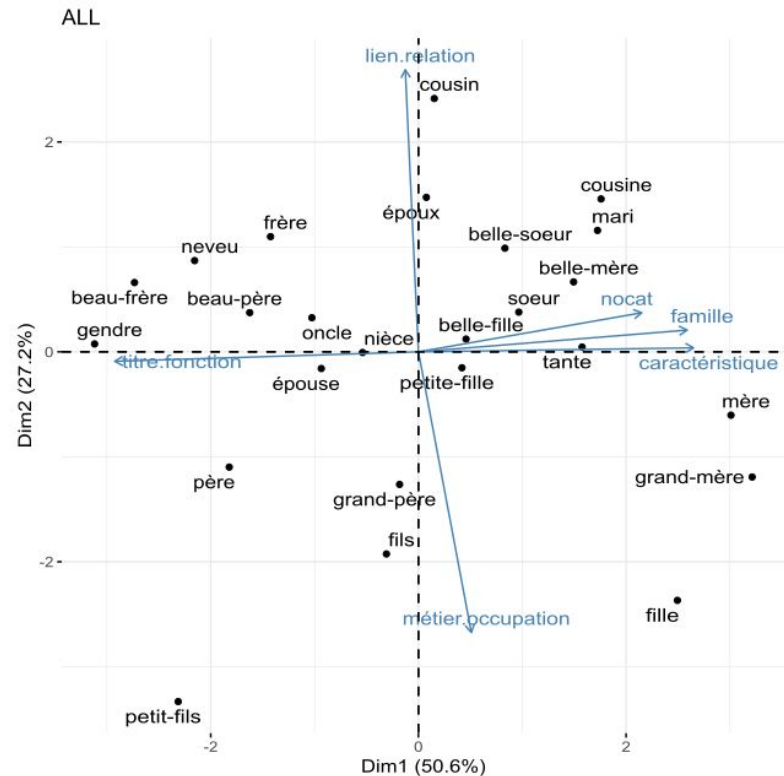
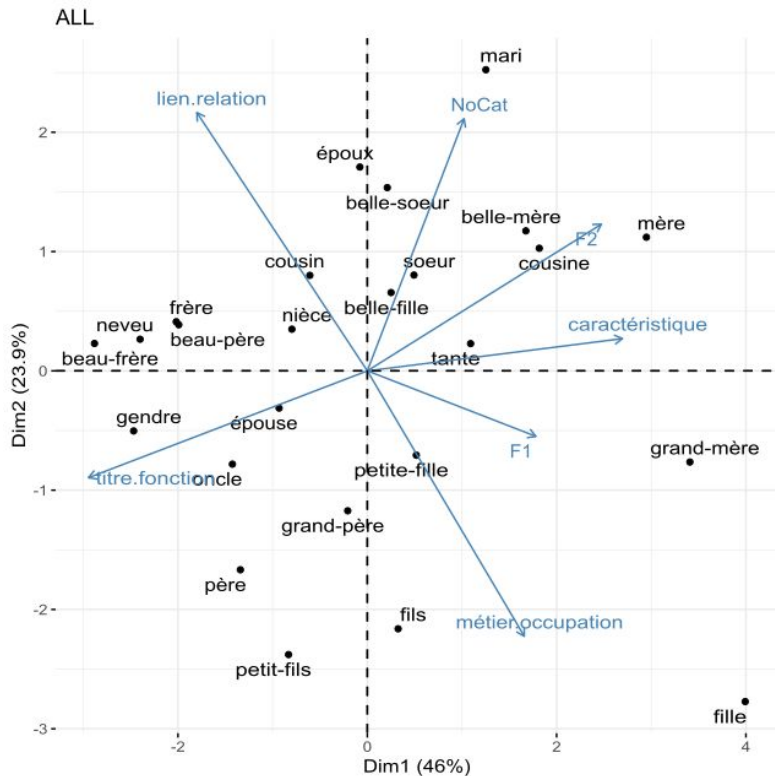
union vs ascendance

Lien relation et non individus vs.
métier.occupation



Types et catégories : annotations humaines et LLM

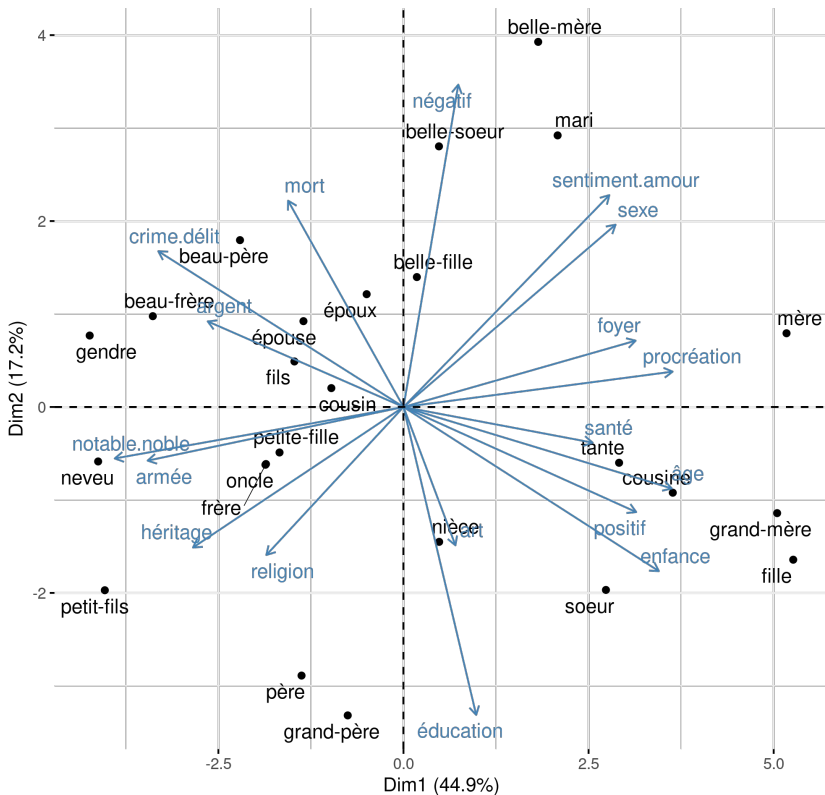
Dim. 1 : masculin vs. féminin; Dim. 2 : union vs ascendance



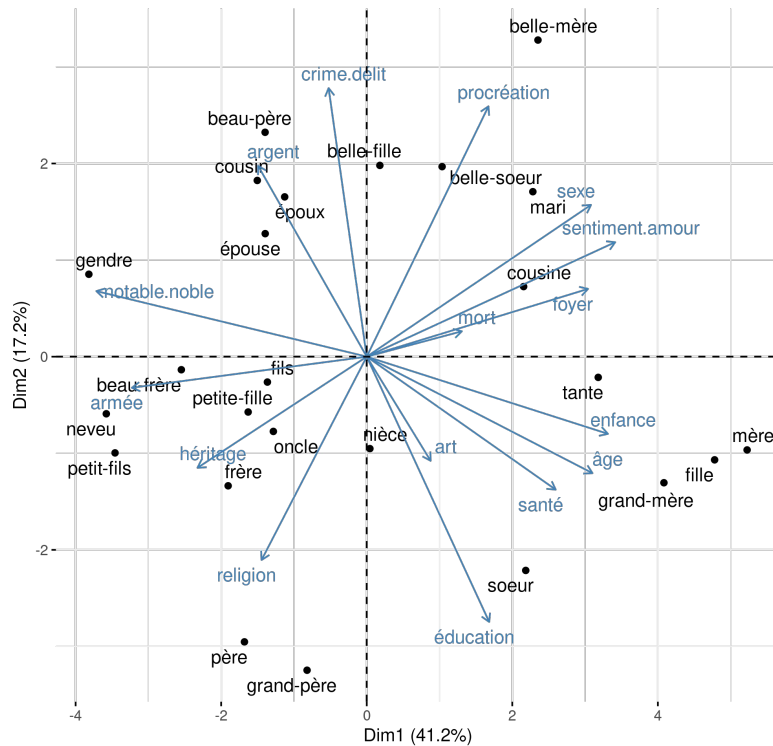
Thématiques : annotations humaines et LLM

Dim. 1 : masculin vs. féminin; Dim. 2 : union vs ascendance

Modèle Global



Modèle Global - LLM



Fin et suite

Bilan

Des dimensions anthropologiques que l'on retrouve en grande partie dans la similarité distributionnelle

Avec des importances relatives différentes

Et de grandes variations entre les corpus

Enrichissement de la description via des analyses thématiques

Un objet facilement transposable

Autres langues et corpus (en cours : diachronie en anglais, COHA 1850-2000)

Une porte d'entrée vers l'interdisciplinarité (anthropologie, histoire...)

- Marzi, G., Balzano, M., & Marchiori, D. (2024). K-Alpha Calculator—Krippendorff's Alpha Calculator: A User-Friendly Tool for Computing Krippendorff's Alpha Inter-Rater Reliability Coefficient. *MethodsX*, 12, 102545.
<https://doi.org/10.1016/j.mex.2023.102545>