Vers une méthodologie d'annotation des nominations émergentes en corpus Séminaire DéLiCorTal Agata Jackiewicz, LIDILEM, Université Grenoble Alpes Praxiling, CNRS, Université Paul-Valéry Montpellier 3 Grenoble, 9 octobre 2020

Problématique et objectif

- La problématique concerne la systématisation des traces des opérations linguistiques impliquées dans l'activité de nomination, menée dans l'optique de leur analyse outillée.
 - Ce sont des entités émergentes et référentiellement instables qui sont visées, dans la mesure où le travail d'élaboration est alors explicitement marqué dans les discours.
- La méthode s'attache à expliciter et à formaliser une démarche qui permettra de repérer et de qualifier des ensembles de marques et de structures qui matérialisent l'acte de nomination en discours.
 - Nous proposons une architecture de notions pertinentes pour la caractérisation des multiples formes que prend en discours l'élaboration des expressions référentielles susceptibles de faire catégorie.

Cadre scientifique et corpus

Cadre : plusieurs projets de recherche coordonnés

- Projet ANR TALAD, « Analyse et traitement automatique de discours » sur les nominations émergentes
- Thèse de doctorat (M. Pengam), fin. Région Occitanie, sur la notion de radicalisation
- Projet MSH Prose, sur l'engagement écologique (militant, permaculturel)

Principaux corpus

- Corpus projet TALAD (société Reticular) : interviews politiques,
- Corpus thèse M. Pengam « musulmans modérés » (Pengam Jackiewicz 2019),
- Corpus du projet PROSE sur l'écologie et la permaculture,
- Corpus « appropriation culturelle » (Mémoire de master 2, 2019, H. Neff),

Nominations observées

- protectionnisme, patriotisme économique, réfugiés climatiques... (Reticular)
- musulmans modérés, appropriation culturelle, éco-fascisme, écologie intégrale, mobilités douces, révolution verte... (Praxiling)

1. NOTION DE NOMINATION

Notion de nomination

- La nomination est une opération linguistique et cognitive, complexe et multidimensionnelle, indissociable des processus d'appréhension et de catégorisation des réalités.
- Elle possède une dimension discursive et dialogique, car l'expression choisie pour nommer un référent reflète la position que le sujet parlant adopte à son égard.
- Les nominations s'inscrivent dans une dynamique de relations sociales et révèlent des représentations que les locuteurs co-construisent, négocient, imposent et font circuler.
- Quatre facettes: catégorisation, signification, performativité, valorisation

Dubois et Mondada (1995), de Chanay H-C. (2001), Bourdieu (2001), Siblot (2001), Kleiber (2001), Charolles (2002), Chave-Dartoen et al. (2012), Frath (2015)...

Construction des « choses sociales »

Construction et négociation

- Dès le début des années 80, le sociologue Pierre Bourdieu signalait l'importance des « opérations sociales de nomination », et la nécessité d'« examiner la part qui revient aux mots dans la construction des choses sociales » (Bourdieu 2001)
 - « Chaque mot, nous le savons, se présente comme une arène en réduction où s'entrecroisent et luttent les accents sociaux à orientation contradictoire. Le mot s'avère, dans la bouche de l'individu, le produit de l'interaction vivante des forces sociales » (Bakhtine, 1929 : 67).

Engagement d'ordre ontologique

- Donner un nom à un objet (social...), c'est dire « oui, ce phénomène existe », affirmation qui peut avoir ses contradicteurs.
 - En cause : enjeux d'ordre juridique et économique :
 - obsolescence programmée, car si programmée alors délibérée (responsabilité)

Nommer en discours

- Nommer ce qui s'impose à nous ou ce que nous souhaitons faire exister
 - rendre compte des réalités émergentes, nommer une altérité non encore envisagée à un phénomène connu,
 - nommer une création (artefact), une idée ou un concept (acte d'institution),
 - rendre compte des transformations et des évolutions (métamorphoses, fractures, jonctions inédites, syncrétisme), interroger les frontières des entités et des catégories,
 - marquer l'apparition de formes hybrides, mettant à mal des ruptures conceptuelles profondes stabilisées,
 - isoler et nommer un phénomène à l'intérieur d'un ensemble flou ou artificiellement homogène (valorisation d'une partie peu visible),
 - renommer une entité pour être plus en accord avec la connaissance scientifique des faits ou phénomènes naturels, psychologiques ou sociaux,
 - renommer pour recatégoriser, càd, faire évoluer la perception d'un référent (valoriser, dévaloriser, neutraliser, atténuer...).

Quelques propriétés

- Type des entités, condensation du sens, schémas actanciels, rapports sémantiques, réseaux terminologiques...
 - homicide, féminicide, uxoricide, (homme ou femme, femme en tant que femme, femme en tant qu'épouse adultère) [cible+cause],
 - réfugiés politiques, réfugiés climatiques, réfugiés médicaux [cause ou motif]
 - pré-radicalisation, radicalisation, anti-radicalisation, déradicalisation...
 [étape d'un processus / degré d'achèvement]
 - changement climatique, dérèglement climatique, urgence climatique, crise climatique, péril climatique [intensité d'un processus + degré de gravité]
 - harcèlement scolaire ou bullying, harcèlement professionnel ou mobbing, cyber harcèlement, harcèlement de rue [circonstance]; harcèlement sexuel, harcèlement moral, stalking [mode opératoire]

- Formation d'expressions qui modifient les contours des catégories établies
 - Le concept de réfugiés environnementaux ou écologiques s'est progressivement forgé dans un contexte de prise de conscience des dégradations de la planète, mais son émergence a bouleversé les catégories classiques du droit des migrations.
 - « Réfugiés » : une catégorie très particulière de migrants que le droit international a entérinée par la Convention de Genève de 1951 (protection des individus contre diverses formes de violences et d'exclusion liées à l'appartenance « ethnique » ou aux convictions politiques ou religieuses)
 - L'environnement rendu inhabitable n'en fait pas partie
 - problème sémantique (variantes et nuances)
 - Réfugiés|migrants + climatiques|écologiques|environnementaux ;
 éco-réfugié...
 - problème politique et juridique
 - La « reconnaissance » de la catégorie doit se traduire en obligations réelles du droit international

- Création de termes (en apparence) oxymoriques
 - Les nouvelles pratiques du végétal (en agroécologie, permaculture, architecture) font apparaître des formes d'hybridité entre domestiqué et sauvage, exploité et intact, entre maîtrise et laisser-faire ...

forêt-jardin, agriculture sauvage, agriculture réensauvagée, agriculture invisible, farming the woods, spontané dirigé, jachère productive ou encore ville sauvage

- [3] De nombreux « peuples de la forêt », grâce à leur interaction millénaire avec le végétal, créent aux alentours des villages des systèmes d'agro-forêts, apparemment <u>sauvages</u> mais en réalité extrêmement <u>domestiqués</u>, ce qu'on appelle « l'agriculture invisible ». (Revue Silence)
- [4] Si «ensauvager la ville» signifie rendre la ville plus humaine, alors ce paradoxe nous offre l'occasion de nous interroger sur la nature profonde de la ville. Ainsi en même temps qu'envisager «le sauvage pour dépayser la ville», envisageons aussi à la bonne échelle cette autre dimension urbaine : «la ville jardin pour cultiver l'humain». (Libération, Ville sauvage, 30/07/2020)

- Signification et valorisation étroitement liées
- appellations renvoyant littéralement ou de manière métaphorique, souvent avec une intention injonctive, à des <u>réalités désirables</u>, ou celles qui sont essentielles mais menacées
 - expressions désignent des modèles pour les villes de demain.
 - « Villes respirables en 5 ans », « Nantes ville comestible », « Imaginez, une ville comestible, respirable, solidaire et inclusive » ;
 - ville durable, ville cyclable, ville vélo-friendly, ville éco-intelligente, ville sauvage...
- expressions péjoratives mobilisant les figures des <u>repoussoirs</u> (absolus) présents dans l'imaginaire collectif d'une société à un moment donné: *Khmers, talibans, djihadistes, ayatollahs,* fascistes... + XX).
 - [5] Je me suis toujours battu contre les Khmers rouges, je ne plierai pas aujourd'hui devant les Khmers verts! (Gérard COLLOMB, lelab.europe1.fr, 22/11/2011)

- « *Musulman modéré* » : une catégorie référentiellement et sémantiquement instable.
 - Socialement souhaitable et à construire pour les uns, contestée par d'autres, cette catégorie ne renvoie pas à une représentation unifiée de la figure du musulman, vu, tour à tour, comme
 - non pratiquant ou pratiquant convenablement,
 - typique ou atypique en tant que musulman,
 - à l'opposé des islamistes « qui passent à l'acte » ou terroriste en puissance,
 - éloigné des extrêmes ou placé en opposition au musulman aux pratiques rigoristes,
 - être réel (pouvant être incarné par une personne existante) ou concept abstrait (pure invention de l'esprit).

(Pengam et Jackiewicz 2019).

Critères de stabilisation

- Plusieurs critères permettent d'apprécier le degré (+/avancé) de stabilisation d'une expression :
 - sociétal :
 - · entrée dans Wikipédia,
 - emploi récurrent comme mot-dièse (sur Twitter, Facebook) #outragesexiste,
 - extralinguistique :
 - reconnaissance juridique, par le droit local et international,
 - lexicographique :
 - inscription dans un dictionnaire en tant que mot-vedette, ou
 - à minima au titre d'une illustration,
 - discursif :
 - emplois détournés, emplois métaphoriques : obsolescence programmée des employés, obsolescence déprogrammée.

2. ARCHITECTURE NOTIONNELLE

Architecture notionnelle

Facettes (complexité)

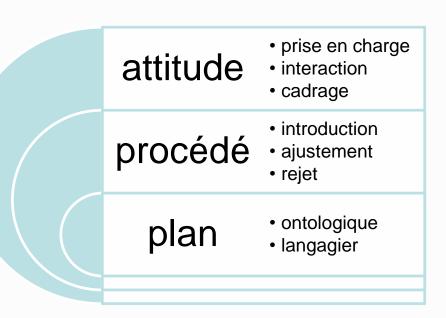
catégorisation signification

performativité valorisation

Repérage (instabilité)

Repéré

 Opérations discursives (enchevêtrement)



Facettes

[6a] les musulmans modérés osent de moins en moins, par solidarité religieuse, désapprouver leurs <u>coreligionnaires</u> intégristes

Catégorisation

Découpage structuration

Performativité

Actes de langage

[6c] La <u>préfecture</u> défend ces mesures « préventives » destinées à « <u>dessiner</u> un **islam modéré** ». [6b] je trouve - à titre personnel - maladroit et même navrant d'accoler <u>l'épithète</u> "modérés" à musulmans. Ce sont les islamistes radicaux et extrémistes qu'il faut spécifier toujours, jamais le fait d'être musulman « ordinaire ». La <u>sémantique</u> a son importance...

Signification

Vigilance épistémique et épilinguistique

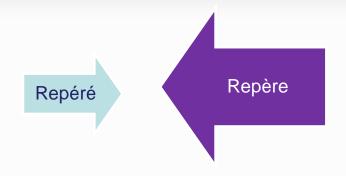
Valorisation

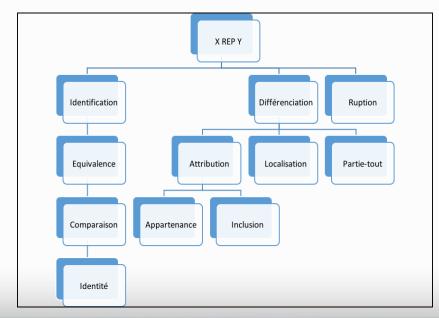
Désidérabilité, préférences, normes

[6d] ... la réaffirmation de notre <u>soutien</u> à **ces musulmans modérés**, <u>éclairés</u>; Il faut <u>aider l'élite</u> **musulmane modérée**

Repérage

- L'appréhension et la désignation des phénomènes émergents ou instables (faiblement autonomes) s'opèrent par rapprochement (assimilation, différentiation...) avec ce qui est mieux connu ou plus stable (« déjà-là »).
- Les différents modes de repérage qui mobilisent un réseau de relateurs statiques et dynamiques - sont d'autant plus intéressants, qu'ils révèlent l'état des connaissances et des représentations qu'ont les locuteurs à propos des objets de leur discours.





Repérage

Relation de ruption

 « X est en rupture /est incompatible avec Y ». Relation non réflexive et symétrique.

« Musulman modéré »

- Les locuteurs tentent de préciser le sens de musulman modéré, au moyen de marqueurs d'opposition et de tours injonctifs qui posent l'existence d'une autre catégorie, plus prégnante, vue comme incompatible, dont il faut se désolidariser.
- Cette catégorie porte des désignations multiples :
 musulmans radicaux, musulmans extrémistes, islam fanatique, islam radical, islam rigoriste, fanatiques, djihadistes, terroristes, fondamentalistes, extrémistes, islamistes radicaux et extrémistes.

[7] Me Blanchart décrit son client comme « un **musulman modéré**, <u>aux</u> <u>pratiques religieuses convenables</u>. Je n'ai pas rencontré un **fanatique**, une **poudrière**. (La Voix du Nord, le 19/11/2015)

Plans

Ontologique

- Modes d'existence et d'organisation des référents
- Description et élaboration de ce qui « est » ou « devrait être »
- Catégorisation, représentation des connaissances (dichotomies, arbres, réseaux...)
- Linguistique référentielle, logique, descriptive

[8] Il faut laisser émerger des modèles de **musulmans modérés** mais pas en inventant une religion modérée. (lesechos.fr, le 30/04/2015)

Langagier

- Mise en mots et leur négociation pour stabiliser la dénomination des états des choses
- Performativité (acte de baptême), valorisation
- Désignation, schémas actanciels, relations, prise en charge, métadiscours...
- Sémantique discursive, énonciation...

[9] Plutôt que le <u>terme</u> de « **radicalisation** », dont on voit qu'il obscurcit autant qu'il éclaire, j'ai proposé dès mars dernier dans une tribune au « Monde » d'utiliser le <u>mot</u> de « **conversion** ».

Procédés d'élaboration

	Plan ontologique	Plan langagier
Introduction	Annoncer (attester) l'existence de R, Faire exister R	Introduire le terme T pour désigner et fixer R
Ajustement	Spécifier R repérer (situer) R par rapport à un référent mieux établi S : Identification, différentiation, ruption entre le répéré R et le repère S)	Ajuster l'emploi du terme T au référent R et au contexte : Discuter l'adéquation du terme au contenu R à représenter, « déplier » la signification, signaler une non coïncidence -> élaboration, glose, illustration, reformulation
Rejet	Rejeter R Infirmer l'existence de R, Rejeter R au profit de R' (c'est R' et non R), Dire la difficulté de concevoir R	Rejeter le terme T : abandonner le terme T Remplacer T par T' (renommer) Dire la difficulté d'employer T (stratégies d'évitement)

Elaboration

- Introduction (création, réappropriation)
 - [10] Glenn Albrecht a créé en 2003 le néologisme solastalgie pour désigner une forme de détresse psychique ou existentielle causée par des changements environnementaux.
 - [11] J'en suis à dire que l'État, un État fort c'est un État quand il est actionnaire de Thales, quand il est actionnaire de la SNCF, quand il actionnaire d'Alstom, il s'organise pour protéger nos intérêts, ça, c'est ce que j'appelle le patriotisme économique. (corpus Reticular)

Ajustement / rejet

- [12] « Femmes de réconfort (ian-fu?) » est l'euphémisme employé au Japon à propos des victimes, souvent mineures, du système d'esclavage sexuel de masse organisé à travers l'Asie par et pour l'armée et la marine impériales japonaises, en particulier durant la Seconde Guerre mondiale. L'emploi de ce terme est fortement contesté par les organisations qui exigent du gouvernement japonais des excuses formelles et des réparations, et préfèrent le terme non édulcoré d'« esclaves sexuelles ». (Wikipédia)

Attitude	Caractéristiques et modalités	
Prise en charge	positionnement subjectif (monoglossique) du locuteur par rapport au référent ou à la nomination employée	
Interaction	 positionnement intersubjectif (hétéroglossique): interaction dialogale ou dialogique avec d'autres qui réalisent des actes de nomination contraction: accord (alignement), désaccord (refus, opposition), expansion (ouverture à d'autres points de vue): imputation, prise en compte (reprise distanciée du point de vue de l'autre), consonance: accord implicite, dissonance: désaccord implicite. 	
Cadrage	expression des cadres de validité : circonstances des événements, des procès et des états de faits dénotés dans le discours (temporelles, spatiales, univers de connaissance)	

- Prise en charge ou positionnement monoglossique
 - [13] Je fais partie de ceux qui soutiennent le principe d'appropriation culturelle depuis le début.
 - [14] Nous pensons qu'il existe des musulmans modérés (non violents, partageant nos valeurs, partisans d'une pratique privée de la religion) mais que l'islam lui-même n'est pas modéré. (vigilanceislam.com)

- Interaction
 - Contraction : désaccord
 - [15] Bravo l'appropriation culturelle et les néologismes exotisants. Pitoyable.
 - Expansion (ouverture à d'autres points de vue):
 imputation
 - [16] Le concept d'appropriation culturelle désigne, si l'on s'en tient à la définition qu'en donne le sociologue Eric Fassin, tout « emprunt culturel s'inscrivant dans un contexte de domination auquel on s'aveugle ». (Le Monde, 18 avril 2019)
 - Prise en compte (reprise neutre ou reprise distanciée du point de vue de l'autre)
 - [17] La notion d' « appropriation culturelle » déchire le monde littéraire.

- Cadrage spatial, temporel, médiatif, univers de connaissance...
 - [18] L'expression « harcèlement de rue », traduction de l'anglais street harassement, est utilisée en Europe et en Amérique du Nord pour désigner des pratiques de harcèlement sexuel, subies principalement par des femmes dans l'espace public (lieux publics ou transports publics) de la part d'inconnus de sexe masculin.
 - [19] Comme l'explique la professeure de droit américaine, Olufunmilayo Arewa, "l'emprunt devient de l'appropriation à partir du moment où il renforce les rapports de domination historiques, ou prive les pays africains d'opportunités de tirer profit de leur patrimoine culturel".

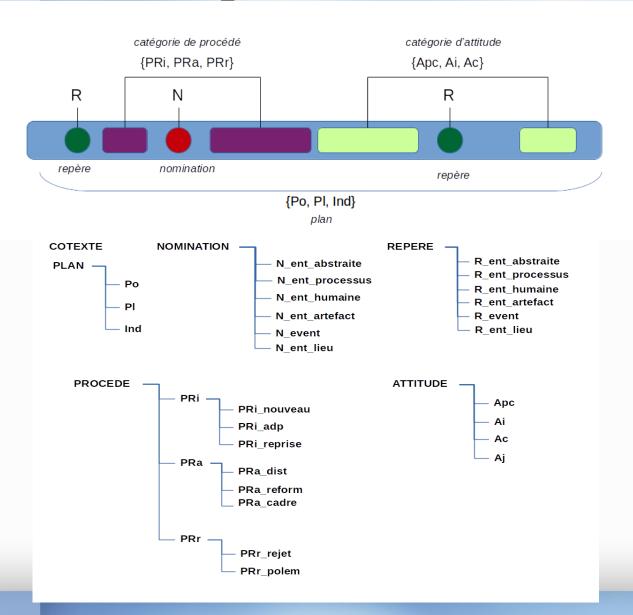
Cadrage

- imbriqué et contrastif :
- [20] Non, je crois que c'est très important, je vais vous dire pourquoi, parce que la France doit retrouver de la croissance pour retrouver le plein emploi, la croissance elle est où? Elle est dans le monde et on ne pourra pas imaginer, c'est pour ça que nous sommes toujours un peu inquiets aujourd'hui du protectionnisme latent, je dirais, de Donald Trump. La France ne s'en sortira pas dans le protectionnisme et donc elle doit absolument grandir dans son économie pour être forte en Europe, de façon à ce que l'Europe soit forte et soit éventuellement autonome.

3. EXPÉRIENCES D'ANNOTATION

Premiers résultats et discussion

Schéma général d'annotation



Corpus TALAD

- Composition
 - 3166 interviews avec 561 personnalités politiques, entre 10/06/2016 et 4/12/2017, soit 10 millions de mots
- Annotations protectionnisme, patriotisme, isolationnisme, moralisation, Europe, violences obstétricales, réfugiés climatiques

Nominations	836		
Plans	772 (92%)		
Elaborations	1339 (intro 42, rejet 65, ajust 1200)		
Attitudes	1253 (cadrage 656, PEC 436)		
Repère	654 (78%)		Total S: 5891
	REL association	487	
	REL ruption	124	
	REL localisation	64	
	REL Identité	69	Total R: 934

Au total, 6825 items d'informations ajoutés ; 8 pour chaque nomination caractérisée

Annotation des segments

[20] [Depuis le XIXe siècle]_Ac, [les droites dures]_Ac invoquent la [«défense du territoire »]_R_ent_abstraite pour justifier le rejet de l'étranger et la sauvegarde des « lois naturelles». Des idées diffusées [en France]_Ac, R_ent_lieu notamment par [la nouvelle droite, et certains catholiques traditionalistes] _R_ent_humaine qui prônent [une][« »]_PRa_dist [écologie intégrale]_N_ent_abstraite [»]_PRa_dist

Annotation en relations (repérage)

- [« écologie intégrale »]_N_ent_abstraite -rs_localisation--> [en France]_R_ent_lieu
- [« écologie intégrale »]_N_ent_abstraite -rs_association--> [la nouvelle droite, et certains catholiques traditionalistes]_R_ent_humaine
- [« écologie intégrale »]_N_ent_abstraite -rs_association- [«défense du territoire »]_R_ent_abstraite

Premiers résultats

Régularités d'expression

- Collecte des marqueurs prototypiques des opérations explicitées par le modèle,
- Identification des schémas d'expression récurrents

Tendances observées

- Le procédé d'ajustement et l'attitude de cadrage constituent les cas majoritairement observés
- activité de nomination appelle celle de désambiguïsation
 - recours à des constructions explicatives, cadratives, des formes de relativisation (un certain protectionnisme).

Tendances à analyser

- Croiser les formes d'ajustement et les types de relations de repérage
- Observer les spécificités des cotextes des nominations en fonction du type sémantico-logique de l'entité désignée (entité humaine, processus social, événement...)

Structures

Schémas d'expression (structures) prototypiques

[21] Je pense qu'il faut faire face au **protectionnisme** agressif de Trump.

["je pense" +"il faut" + "faire face à" [Ai]],	attitude interaction (dissonance, voire désaccord)
["protectionnisme"[N_ent_abstraite]],	nomination entité abstraite
["agressif"[PRa_spécification]]	procédé d'ajustement spécification
["de Trump"[PRa_cadre]].	procédé d'ajustement cadre

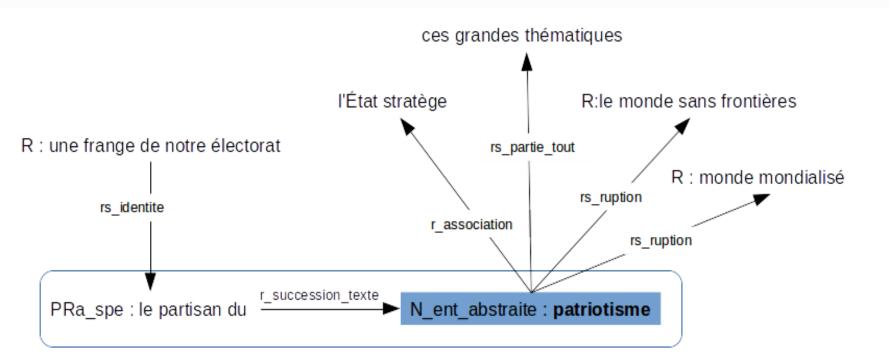
attitude (interaction)

nomination

ajustement (spécificatif)

ajustement (cadratif)

Repères : exemple de graphe local



Des réseaux de cooccurrences peuvent être construits à partir des sous-graphes locaux pour caractériser le texte (corpus) dans sa globalité, car ces graphes permettent de détailler la sémantique des cooccurrences en lien avec les nominations à l'étude.

Influence du genre discursif

- Le genre discursif apparaît comme paramètre explicatif de présence ou d'absence de certains procédés (éléments co-textuels)
- Le cas des relations de repérage
 - Le discours institutionnel ne fait pas usage de la ruption pour définir le processus de *radicalisation* et discriminer ce qu'il n'est pas, ceci notamment en raison des procédés de lissage discursifs qui le structurent (Oger et OllivierYaniv 2006).
 - Le discours militant pourrait à l'inverse faire apparaître de façon plus significative ce type de relation, dans des négations polémiques.
 - [22] Non, un féminicide n'est pas un « différend de couple » !
 (https://legroupef.fr/nonun-feminicide-nest-pas-un-different-de-couple/)

Difficultés

- Maîtrise élevée des concepts d'analyse de discours et de linguistique
- Connaissance (poussée) des domaines dont relèvent les nominations émergentes et des débats que celles-ci suscitent
 - Expressions à la frontière des vocabulaires courants et des terminologies spécialisées (et des systèmes de connaissances associées)
 - anthropocène, service écosystémique, limites planétaires...
 - semences paysannes, variétés population....
 - ...ni l'écologie, ni la radicalisation violente, ni l'appropriation culturelle, ni le protectionnisme ne vont de soi...
- Cela confirme l'intrication du linguistique et de l'extralinguistique, et de fait... l'intérêt du modèle construit.

Travaux en cours et à venir

- Finalisation et formalisation dans le gestionnaire de thésaurus OpenTheso, au format SKOS - d'une ressource terminologique (400 termes linguistiques) relative aux phénomènes de la référence et de la nomination
- Articulation des deux ensembles de connaissances linguistiques dans une plateforme commune.

Champs d'application en SHS

- Contribution linguistique aux études SHS des processus à l'oeuvre dans la catégorisation des réalités sociales
- Problématiques
 - observer les figements des catégories et leurs usages prescriptifs,
 - étudier des processus d'instrumentalisation des catégories à des fins socio-politiques (hiérarchisation, marginalisation, discrimination, valorisation...),
 - savoir discriminer des catégorisations socialement discutables (à caractère ethnicisant, raciste...),
 - étudier les ressorts de la revalorisation des catégories initialement stigmatisées...

Quelques références

- Jackiewicz, A., Bebeshina, N., Cassier, M., Frontini, F.,
 Halftermmever, A., Longhi, J., Luxardo, G. et Nouvel, D. (2019).
 Vers une ontologie de la nomination et de la référence dédiée à l'annotation des textes. *ToTH 2019*, 6-7 juin 2019, Chambéry.
- Pengam, M. et Jackiewicz, A. (2019). Sens et emplois de l'expression « Musulmans Modérés » dans les discours médiatiques. Open Library of Humanities. DOI: https://doi.org/10.16995/olh.431
- Jackiewicz, A et Pengam, M. (2020). Un modèle pour l'étude des nominations émergentes. Notion de repérage pour saisir les modalités d'ajustement sémantique et discursif, CMLF 2020, https://doi.org/10.1051/shsconf/20207812004