Actes de CORIA-TALN 2026
Nantes, France, 29 juin au 3 juillet 2026

Omniprésents et anthropomorphisés : analyse lexico-syntaxique des discours sur les LLM

Clémentine Bleuze[1] Bruno Guillaume[1] Aurélie Névéol[2] Karën Fort[1]
(1) Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France
(2) Université Paris-Saclay, CNRS, Laboratoire Interdisciplinaire des Sciences du Numérique, Orsay, France
clementine.bleuze@univ-lorraine.fr

RÉSUMÉ

Depuis quelques années, les grands modèles de langue (LLM) sont au cœur des discours médiatiques comme scientifiques. Ils sont à cette occasion régulièrement anthropomorphisés, c'est-à-dire, présentés comme ayant des capacités habituellement attribuées aux êtres humains. Dans cet article, nous souhaitons ouvrir une discussion sur les risques de telles pratiques, et appuyer l'importance d'une parole scientifique fiable portée par la communauté du Traitement Automatique des Langues (TAL). Pour illustrer notre propos, nous proposons une méthode de détection de l'anthropomorphisation basée sur des indices lexico-syntaxiques en français, que nous appliquons sur un corpus de 737 publications CORIA-JEP-TALN-RECITAL entre 2020 et 2025. Nos résultats mettent en évidence des formulations potentiellement anthropomorphisantes, bien que non majoritaires, auxquelles il convient de réfléchir collectivement. Quel discours scientifique souhaitons-nous, chercheurs et chercheuses en TAL, porter sur les grands modèles de langue ?

ABSTRACT

Ubiquitous and anthropomorphized : a lexico-syntactic analysis of discourse on LLMs.

In recent years, Large Language Models (LLMs) have been at the center of media discussions as well as scientific discourse. Regularly, cases of anthropomorphization (that is, the depiction of inanimate entities as possessing human-like capabilities) have been reported. In this paper, we wish to open the discussion about the risks of such practices, while stressing the importance of a sound scientific discourse held by the Natural Language Processing (NLP) community. To illustrate our point, we propose a method to detect anthropomorphization using lexical and syntactic clues in French. We apply this method to a corpus of 737 CORIA-JEP-TALN-RECITAL papers published between 2020 and 2025. Our results highlight potentially anthropomorphizing wording, which, although non majority, should be discussed collectively. Which scientific discourse on LLMs do we want to convey as NLP researchers ?

MOTS-CLÉS : TAL pour le TAL, anthropomorphisation, grands modèles de langue.

KEYWORDS: NLP for NLP, anthropomorphization, large language models.

1 Introduction

« La récente émergence des grands modèles de langue (LLM) a révolutionné le domaine du Traitement Automatique des Langues (TAL) ». Ces modèles « atteignent des performances inégalées, établissant

Actes des 33ème Conférence sur le Traitement Automatique des Langues Naturelles.
Volume 1 : articles scientifiques originaux, pages 405–418.
Cette œuvre est mise à disposition sous licence Attribution 4.0 International.

un nouvel état de l'art pour la génération de texte, la compréhension du langage naturel, ou encore la traduction automatique ». Leurs « capacités d'analyse et de raisonnement » leur permettent de « rivaliser, voire de surpasser les humains » dans certaines tâches. Ils sont des « assistants rédactionnels », des « agents intelligents », demain peut-être des « psychologues virtuels », etc.

Tout·e chercheur·euse en TAL reconnaîtra dans ce qui précède des formules familières, évocatrices d'un nombre conséquent de ses lectures académiques récentes. Il semblerait en effet dérisoire de le nier : les LLM sont à l'heure actuelle omniprésents dans la recherche en TAL, comme ils le sont dans les discours médiatiques, politiques, inter-personnels. Cette omniprésence s'accompagne d'un phénomène croissant d'anthropomorphisation (Floridi & Nobre, 2024), c'est-à-dire le fait « [d'] attribuer des capacités et caractéristiques humaines à des systèmes inanimés » (Inie *et al.*, 2026). Notons que ceci peut aussi bien être le fait des locuteur·ices qui produisent des discours sur les LLM, que des concepteur·ices qui implémentent ces modèles dans des systèmes anthropomorphisants « par design » (par ex., leur attribuant un nom, une voix, ou un avatar) (Abercrombie *et al.*, 2023).

Nous souhaitons aborder ici la question de l'omniprésence et de l'anthropomorphisation des LLM, en particulier au sein d'un discours scientifique qui peut, *in fine*, induire en erreur le public sur la nature réelle de ces modèles ainsi que de leurs « capacités » (Shardlow & Przybyła, 2024).

1.1 Conséquences et caractérisation de l'anthropomorphisation

Un certain nombre de travaux se sont penchés ces dernières années sur le phénomène général d'anthropomorphisation des technologies dites d'intelligence artificielle (IA) (Deshpande *et al.*, 2023; Abercrombie *et al.*, 2023; Placani, 2024; Ryazanov *et al.*, 2024; Verwilghen, 2025), et parfois plus spécifiquement des LLM (Shardlow & Przybyła, 2024; Inie *et al.*, 2024; Raza & Meghji, 2025; Inie *et al.*, 2026). Ces études décrivent certaines conséquences d'ordre légal (Deshpande *et al.*, 2023) ou psychologique de cette anthropomorphisation, incluant notamment un faux sentiment d'attachement de la part des utilisateur·ices, des croyances erronées ou une influence sur le niveau de confiance placé dans le modèle (Shardlow & Przybyła, 2024; Inie *et al.*, 2024).

Tandis que DeVrio *et al.* (2025) décrivent des structures anthropomorphisantes dans des textes produits *par* des LLM , nous proposons de nous concentrer ici sur les discours émis par des êtres humains *à propos* de ces mêmes LLM. Inie *et al.* (2024) introduisent une première taxonomie différenciant les énoncés qui attribuent des capacités cognitives, agentives ou communicatives aux LLM, ainsi que ceux reposant sur des métaphores biologiques. Cette taxonomie a depuis été enrichie de nouvelles catégories (par exemple, les analogies avec des fonctions humaines, ou l'évocation d'émotions des modèles) et mise en regard de taxonomies existantes proposant des catégorisations similaires (Inie *et al.*, 2026). Un point commun des différentes définitions proposées est la dimension fréquemment implicite de l'anthropomorphisation [1], qui peut résulter de choix lexico-syntaxiques particuliers : entre autres, Hunger (2023) remarque ainsi que le simple choix d'un verbe d'action comme *créer* ou *produire* peut se révéler anthropomorphisant.

1. Par opposition à une anthropomorphisation explicite, par exemple dans un énoncé tel que « Les LLM sont des êtres conscients ».

1.2 Détection de l'anthropomorphisation dans un corpus

Nous recensons dans ce qui suit quelques propositions de méthodes pour détecter automatiquement l'anthropomorphisation dans un corpus. Coll Ardanuy *et al.* (2020) utilisent les propriétés des modèles de langue masqués tels que BERT (Devlin *et al.*, 2019) pour identifier des expressions présentant une « animéité contraire » [2] dans des écrits du 19ème siècle en anglais portant sur les machines. Leur stratégie consiste à masquer les expressions référant à des machines (par ex., « Et pourquoi dirait-on que la [MASK] ne vit pas ? ») et à lister les tokens les plus probables prédits par BERT à cette position (par ex., *personne*, *femme*, etc.). Un score d'animéité moyenne de l'expression masquée est calculé en considérant, pour chacun des tokens prédits, s'il hérite dans *WordNet* (Fellbaum, 1998) du concept d'« entité vivante » (*living thing*) ou non ; ce score est alors mis en regard d'un seuil déterminé expérimentalement, pour conclure s'il s'agit d'un cas d'animéité contraire. C'est sur une approche similaire que repose la métrique ANTHROSCORE (Cheng *et al.*, 2024) mesurant, pour un syntagme nominal masqué donné (en anglais), le quotient des probabilités estimées par un modèle RoBERTa (Liu *et al.*, 2019) de correspondre à un pronom « humain » (*he*, *she*, etc.) *vs.* « non-humain » (*it*, *its*).

Lonke *et al.* (2025) relèvent toutefois qu'ANTHROSCORE s'applique mal aux cas où des indices d'anthropomorphisation nominaux ou adjectivaux se trouvent à l'intérieur du syntagme masqué [3] ; par ailleurs, cette méthode est peu adaptée pour des langues autres que l'anglais [4]. S'appuyant sur une analyse linguistique de l'anthropomorphisation et sur la sémantique des cadres (*frame semantics*), les auteurs proposent d'identifier une diversité de structures syntaxiques qui, couplées à des indices lexicaux inspirés de (DeVrio *et al.*, 2025), peuvent se révéler anthropomorphisantes. Cette méthode leur permet d'extraire des phrases possiblement anthropomorphisantes d'un ensemble de publications tirées de l'ACL Anthology [5] et d'arXiv [6], par la suite annotées manuellement en trois classes : anthropomorphisation avérée, incertaine, ou nulle. ANTHROSET (Lonke *et al.*, 2025), le jeu de données résultant, est présenté comme un corpus d'évaluation pour la détection de l'anthropomorphisation, mais n'a cependant pas à notre connaissance fait l'objet d'analyses qualitatives revenant sur la part d'anthropomorphisation dans les conférences anglophones de TAL. Shardlow *et al.* (2025) annotent quant à eux manuellement 601 résumés d'articles longs publiés à ACL en 2022, dont 74 % comprennent des marques d'anthropomorphisation explicites ou ambigües. Shardlow *et al.* (2025) exploitent ensuite ces annotations pour entraîner un modèle de régression basé sur XLNet (Yang *et al.*, 2019), utilisé en inférence sur l'ACL Anthology, ce qui révèle une tendance à la hausse de l'anthropomorphisation dans toutes les conférences majeures. Cependant, il n'existe à notre connaissance pas de travaux ayant tenté d'identifier automatiquement des cas d'anthropomorphisation des LLM en français, *a fortiori* dans des publications francophones en TAL [7].

2. Suggestion de traduction de *atypical animacy*. En linguistique, l'animéité ou caractère animé est considérée comme une propriété sémantique inhérente des référents (Thuilier *et al.*, 2024).

3. Par exemple dans « les capacités cognitives de ChatGPT ».

4. Notamment, le français ne dispose pas d'un pronom inanimé prototypique équivalent à *it*.

5. https://aclanthology.org.

6. https://arxiv.org/.

7. Le corpus analysé par Shardlow *et al.* (2025) doit en théorie comprendre des articles CORIA-JEP-TALN-RECITAL parmi les conférences regroupées dans la catégorie « Autre », mais cela n'est pas explicité dans l'article. Par ailleurs, l'application d'un modèle entraîné sur de l'anglais à des articles en français risque de ne pas fournir de résultats fiables.

2 Une méthode pour détecter des cas d'anthropomorphisation des LLM en français

Étant donné un corpus d'articles scientifiques en français, nous proposons une méthode de détection des cas potentiellement anthropomorphisants opérant en trois temps :

— identification dans les articles des expressions référant à des LLM, ci-après appelées *mentions* de LLM
— reconnaissance de structures linguistiques incluant ces mentions de LLM à l'aide de relations de dépendances syntaxiques, d'après une méthodologie inspirée de (Lonke *et al.*, 2025)
— estimation du degré d'anthropomorphisation des structures extraites à l'aide d'indices lexicaux et syntactico-sémantiques

2.1 Identification des mentions de LLM

De même que les travaux existants sur le sujet, nous proposons dans un premier temps d'identifier les mentions de LLM présentes dans un article à l'aide de mots-clés caractéristiques incluant à la fois des dénominations génériques (par ex. « modèle de langue ») et des noms de LLM (par ex. « CamemBERT »). Nous nous inspirons d'une liste initiale de mots-clés utilisée par Lonke *et al.* (2025), que nous traduisons en français et étendons pour minimiser le risque de faux négatifs (Table 1) [8]. Nous utilisons des expressions régulières pour extraire les nombreuses variations des noms de LLM (par ex. « ChatGPT » ou « gpt-4o » pour « gpt »).

dénominations génériques	**noms de LLM**
ia ; intelligence artificielle ; llm ; agent {conversationnel, llm} ; modèle {de langue, de langage, de fondation, génératif, auto-régressif, supervisé, auto-supervisé, non-supervisé}	bard ; bert ; claude ; deepseek ; gemini ; gemma ; gpt ; llama ; mistral ; qwen

TABLE 1 – Mots-clés servant à identifier les mentions de LLM dans un texte donné.

2.2 Reconnaissance de structures syntaxiques

Suivant la méthodologie de Lonke *et al.* (2025), nous proposons d'extraire automatiquement une série de structures syntaxiques potentiellement anthropomorphisantes. Nous utilisons un modèle spaCy [9] du français performant pour l'annotation en dépendances syntaxiques au format *Universal Dependencies* (UD) (de Marneffe *et al.*, 2021). Nous appelons ci-après *entité-LLM* la tête syntaxique d'une mention de LLM, dans le cas où cette mention comprend plus d'un mot : ainsi *modèles* est l'entité-LLM de la mention « les grands modèles de langue ». Nous récapitulons en Table 2 les structures syntaxiques considérées ainsi que les relations de dépendances employées pour les identifier. Bien que Lonke *et al.* (2025) considèrent une sixième structure de comparaison entre entité-LLM et une entité référant à des humains (par ex., « l'IA peut être aussi créative que les humains »), nous la trouvons plus complexe à implémenter et la laissons de côté pour l'instant.

8. Inversement, cette approche peut retourner des faux positifs, et nous reconnaissons qu'*intelligence artificielle* n'est pas un synonyme de *LLM*. Cependant, précisément parce que nous nous situons dans le contexte de conférences liées au TAL, nous nous attendons à ce que ces occurrences génériques soient plus ou moins substituables à des occurrences de *LLM*.
9. https://spacy.io/models/fr#fr_dep_news_trf.

structure syntaxique	**description / exemple**	**caractérisation syntaxique UD**
entité-LLM sujet verbal	entité-LLM en position de sujet verbal (« le modèle **encode** des propriétés sémantiques »)	nsubj($e_{\{\text{POS=VERB}\}}$, e_{LLM})
entité-LLM objet/complément verbal	entité-LLM en position d'objet ou de complément verbal (« nous **entraînons** un LLM », « nous nous **basons** sur RoBERTa »)	obj($e_{\{\text{POS=VERB}\}}$, e_{LLM}) *ou* iobj($e_{\{\text{POS=VERB}\}}$, e_{LLM}) *ou* obl :arg($e_{\{\text{POS=VERB}\}}$, e_{LLM})
modificateur adjectival/attribut du sujet de l'entité-LLM	modificateur qualifiant directement l'entité-LLM comme adjectif ou attribut du sujet (« une IA **empathique** », « BERT est **meilleur** »)	amod(e_{LLM}, $e_{\{\text{POS=ADJ}\}}$) *ou* nsubj($e_{\{\text{POS=ADJ}\}}$, e_{LLM})
modificateur nominal de l'entité-LLM	structure nominale doublée susceptible de décrire le rôle ou la fonction de l'entité-LLM (« le modèle **professeur** »)	nmod(e_{LLM}, $e_{\{\text{POS=NOUN}\}}$)
entité-LLM complément du nom	structure dans laquelle l'entité-LLM fait partie d'un complément du nom, exprimant possiblement une relation d'appartenance (« les **capacités** du modèle »)	nmod($e_{\{\text{POS=NOUN}\}}$, e_{LLM}) *et* case(e_{LLM}, e')

TABLE 2 – Structures syntaxiques ciblées par notre analyse, légèrement modifiées d'après (Lonke *et al.*, 2025). La caractérisation UD est indiquée au format *relation_de_dépendance(tête, dépendant)*. e_{LLM} désigne une entité-LLM ; e et e' désignent des entités quelconques.

2.3 Agrégation d'indices lexicaux et sémantico-syntaxiques

Les structures syntaxiques décrites précédemment ne constituent pas en elles-mêmes des indices concluants d'anthropomorphisation. Pour chacune d'entre elle, nous vérifions donc si les verbes, adjectifs, et noms impliqués figurent dans une liste pré-définie d'indices lexicaux (Table 5).

Pour affiner l'analyse des structures verbales, nous exploitons également des traits sémantiques de verbes du français contenus dans le dictionnaire DICOVALENCE (van den Eynde & Mertens, 2010; van den Eynde *et al.*, 2017), inventoriant les cadres valenciels de plus de 3 700 verbes du français. DICOVALENCE s'appuie sur « l'Approche Pronominale » en syntaxe (van Den Eynde & Blanche-Benveniste, 1978), qui précise pour chaque position syntaxique dans la valence d'un verbe donné le paradigme de pronoms associé, déterminant ainsi certaines restrictions de sélection. Les traits sémantiques « humain », « concret non humain » et « abstrait » nous intéressent ici particulièrement : le verbe *voir* comporte par exemple dans DICOVALENCE quatre cadres de valence (donc, quatre entrées distinctes), dont chacune comporte la restriction d'admettre en position sujet un nom/pronom « humain ». Par conséquent, accoler le verbe *voir* à un sujet non-humain constituerait un indice fort d'anthropomorphisation.

Nous proposons ainsi de modéliser l'« humanitude » du cadre de valence d'un verbe vis-à-vis d'une position syntaxique donnée comme une variable binaire valant 1 si cette position admet le trait « humain » mais pas le trait « non humain » [10], et 0 sinon. Au niveau du verbe, le ratio du nombre d'entrées « humaines » vis à vis d'une position syntaxique donnée, sur l'ensemble de ses entrées,

10. En théorie, une position pouvant accepter à la fois des référents humains et non-humains possèderait les deux traits.

constitue alors un « score d'humanitude » compris entre 0 et 1. Cette approche permet d'attribuer à « voir » (possédant 4 entrées dans DICOVALENCE) un score d'humanitude de 1 pour la position sujet, mais de 0,25 pour la position objet/complément [11].

Nous proposons d'agréger ces indices lexicaux et sémantico-syntaxiques en considérant qu'une structure syntaxique donnée est anthropomorphisante si elle implique un verbe, un adjectif ou un nom considéré comme un indice lexical d'anthropomorphisation ; ou si elle implique un verbe dont le score d'humanitude calculé à partir de DICOVALENCE est supérieur ou égal à 0,5 [12]. L'anthropomorphisation d'un article entier est alors estimée par le nombre de structures anthropomorphisantes qu'il contient.

3 Étude de cas sur le corpus CORIA-JEP-TALN-RECITAL

À titre d'illustration, nous appliquons notre méthode à un corpus de 737 articles représentatifs de la production francophone en TAL. Nous considérons l'intégralité des articles recensés dans l'ACL Anthology entre 2020 et 2025 au sein des conférences CORIA, JEP, TALN et RECITAL [13]. Nous publions librement le code permettant de reproduire ou d'adapter cette étude exploratoire [14].

3.1 Évolution diachronique des mentions de LLM

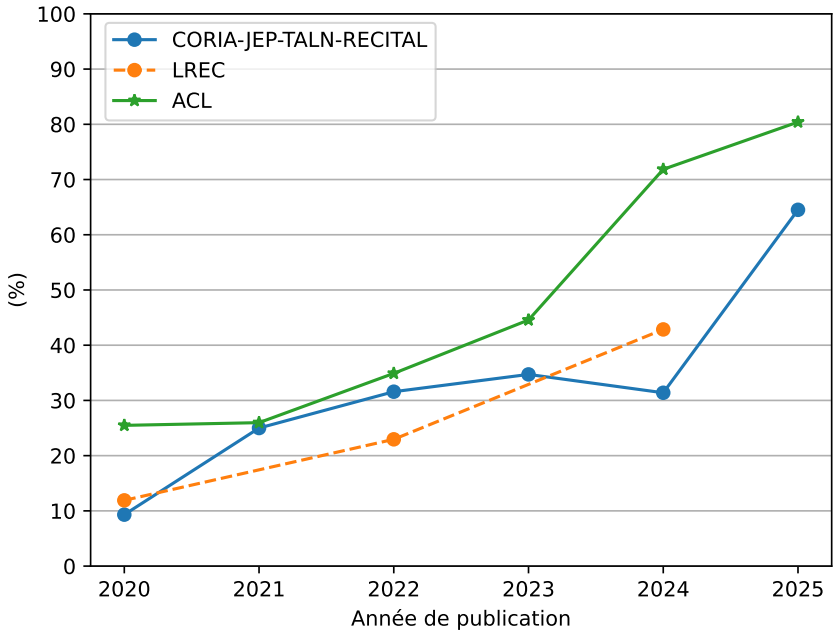

FIGURE 1 – Évolution de la part annuelle de publications comportant au moins une mention de LLM dans leur titre et/ou résumé pour différentes conférences représentatives du domaine du TAL.

11. Pour « voir quelqu'un·e ». Nous considérons ensemble les positions d'objet direct, indirect, et de complément verbal.
12. Ce seuil est arbitraire, et pourrait être modifié au besoin.
13. Ces articles sont indexés dans l'ACL Anthology à l'adresse https://aclanthology.org/venues/jeptalnrecital/, et incluent également les articles d'ateliers joints comme DEFT (DÉfi Fouille de Textes). Bien que CORIA et JEP ne soient pas des conférences de « pur » TAL, ceci nous permet d'obtenir un corpus de volume plus conséquent, et d'adopter un point de vue plus large sur les productions de la communauté. L'expérience est cependant tout-à-fait reproductible à grain plus fin.
14. https://github.com/ClementineBleuze/LLM-anthropomorphisation.

Dans un premier temps, nous mesurons la part annuelle de publications CORIA-JEP-TALN-RECITAL comprenant une ou plusieurs mentions de LLM dans leur titre et/ou résumé. Nous observons une hausse significative [15] entre 2020 et 2025 (voir Figure 1), avec une présence des LLM qui passe de 9,32 % à 64,52 % de l'ensemble des publications. Par ailleurs, cette part a plus que doublé depuis 2024, où elle couvrait 31,39 % des publications. À titre de comparaison, nous mesurons également l'évolution de ce taux pour les conférences majeures internationales LREC et ACL [16] qui comptent respectivement 4 761 et 15 003 publications entre 2020 et 2025. Nous constatons que les LLM sont de plus en plus présents dans l'ensemble des trois conférences, bien que la tendance soit plus marquée à ACL (80,41 % des publications de 2025), et semble plus modérée à LREC (42,86 % des publications en 2024). Il est à noter que cette présence pourrait être sous-estimée par notre méthode qui ne considère que les titres et résumés des articles, et non leur texte intégral.

3.2 Anthropomorphisation des LLM

Dans ce qui suit, nous considérons uniquement les 258 articles CORIA-JEP-TALN-RECITAL faisant mention de LLM pour identifier des cas possiblement anthropomorphisants [17].

3.2.1 Entités-LLM en position de sujet et d'objet/complément verbal

Nous comptons 113 entités-LLM placées en position de sujet verbal, réparties dans 81 articles. Sur les 66 verbes relevés, 35 (53,0 %) ont un score d'humanitude (pour la position de sujet) supérieur ou égal à 0,5 (calculé d'après leurs traits dans DICOVALENCE), et 4 (6,1 %) sont présents dans notre liste d'indices lexicaux (*connaître*, *distinguer*, *capturer*, *différencier* ; cf. Table 5). De même, 178 entités-LLM sont placées en position d'objet ou de complément verbal dans 122 articles. Sur les 65 verbes relevés, 8 (12,3 %) ont un score d'humanitude (pour la position d'objet/complément) supérieur ou égal à 0,5 ; deux (3,1 %) sont présents dans notre liste d'indices lexicaux (*identifier*, *demander*). Quelques exemples sont présentés en Table 3.

Ces résultats semblent en accord avec l'observation que la fonction sujet attire les référents humains en français (Thuilier *et al.*, 2024), ce qui la rend plus propice à l'anthropomorphisation des LLM que les positions d'objet ou de complément. Outre les quelques exemples déjà cités, nous listons ci-après l'ensemble des verbes ayant un score d'humanitude de 1,0 relevés dans le corpus : *afficher*, *affirmer*, *connaître*, *exceller*, *exploiter*, *favoriser*, *hériter*, *introduire*, *négliger*, *obtenir*, *recommander*, *respecter*, *surpasser*, *traiter*, *évaluer* pour la position de sujet ; et *permettre* pour la position d'objet/complément.

3.2.2 Adjectifs, attributs du sujets, et modificateurs nominaux

Nous relevons de nombreux adjectifs en relation de modificateur adjectival ou d'attribut du sujet avec des entités-LLM : 358 mentions réparties dans 160 articles couvrent 100 adjectifs distincts. Sans

15. Validée par un test de Cochran-Armitage avec $p < .001$.

16. Nous utilisons cette fois un modèle spaCy de l'anglais (https://spacy.io/models/en#en_core_web_sm) ainsi que les codes événements "acl" et "lrec" de l'ACL Anthology.

17. Les structures extraites ont été parcourues manuellement pour vérifier leur pertinence en contexte. Nous avons noté quelques erreurs marginales provenant du modèle (mauvaise lemmatisation, attribution de la mauvaise dépendance syntaxique), dont nous pourrons fournir une description plus systématique dans des travaux ultérieurs.

verbe	position de l'entité-LLM	#occurrences (#articles)	score hum.	indice lex.	exemple tiré du corpus
pouvoir	sujet	14 (13)	0,5	non	« Les LLMs **peuvent**-ils répondre à des QCM ? »
obtenir	sujet	8 (8)	1,0	non	« Le LLM **obtient** des résultats impressionnants »
montrer	sujet	5 (5)	0,67	non	« Les modèles CamemBERT **montrent** une plus grande exactitude »
baser	objet/compl.	21 (16)	0,0	non	« un entraînement **basé** sur le modèle de langage GPT-2 »
entraîner	objet/compl.	11 (11)	0,43	non	« nous **entraînons** des modèles BERT à partir de zéro »
demander	objet/compl.	1 (1)	0,57	oui	« en **demandant** aux LLMs de répondre à ces questions »

TABLE 3 – Exemples de verbes fréquents et/ou anthropomorphisants relevés dans le corpus.

surprise, les adjectifs les plus fréquents font en réalité partie des mentions de LLM : *grand*, *génératif*, *conversationnel*, etc. En revanche, nous retrouvons 6 adjectifs (6,0 %) faisant partie de nos indices lexicaux : *artistique*, *capable*, *confiant*, *fiable*, *français*, et *équitable*.

Le cas de l'adjectif *français* a attiré notre attention : il est différent de dire « un modèle du français » (*français* est alors un modificateur nominal de *modèle*) et « un modèle français » (modificateur adjectival tel que ceux que nous listons ici). Dans « CamemBERT-bio : Un modèle de langue français savoureux », *français* évoque à la fois l'origine du modèle CamemBERT-bio (Touchent *et al.*, 2023) (au sens plus large d'un « produit français ») et la langue traitée par le modèle. Il y a alors jeu de mot sur un double sens, qui n'est pas anthropomorphisant dans la même mesure que « les biais constatés dans les modèles français », où un modèle semble être doté d'une nationalité. Cet exemple nous a permis de remarquer dans le corpus des usages similaires : « le modèle *québécois* », ainsi que « le modèle *natif* » opposé à des « modèles *étrangers* ». Bien que ces derniers adjectifs ne figurent pas parmi nos indices lexicaux, il nous semble qu'ils peuvent constituer des indices d'anthropomorphisation.

La liste des modificateurs nominaux identifiés (51 occurrences, dont 32 uniques, dans 31 articles) ne contient pas d'indices pertinents d'anthropomorphisation. Il s'agit de noms ou de spécifications de modèles (*gpt-2*, *bio*, *14b*, etc.), à l'exception de *détecteur* exprimant une fonction, qui n'est ici selon nous pas anthropomorphisante. Il est également possible que la construction de modification nominale soit moins utilisée en français qu'en anglais, ce qui en fait ici une structure moins productive.

3.2.3 Compléments du nom

Enfin, nous identifions 336 structures où un nom se voit modifié par une entité-LLM selon une relation de complément du nom, dans 168 articles. Sur les 144 noms uniques relevés, 13 (9,0 %) constituent des indices d'anthropomorphisation : *aide*, *analyse*, *capacité*, *comportement*, *compétence*, *confiance*, *connaissance*, *décision*, *hallucination*, *intelligence*, *préférence*, *raisonnement* et *réponse*. Nous remarquons cependant qu'*intelligence* constitue un faux positif modifiant l'entité *IA* dans « l'intelligence artificielle (IA) », et non pas dans une formulation du type « l'intelligence du modèle ». D'autres cas relèvent de relations qui ne traduisent pas de relation d'appartenance à l'entité-LLM,

comme la formulation fréquente de « l'*utilisation* d'un modèle ». Quelques exemples fréquents et/ou anthropomorphisants sont listés en Table 4.

nom	**#occurrences (#articles)**	**indice lex.**	**exemple tiré du corpus**
performance	24 (24)	non	« améliorer les **performances** des modèles »
capacité	17 (17)	oui	« nous évaluons la **capacité** lexicale de quatre LLM standards »
comportement	3 (3)	oui	« des indicateurs interprétables sur le **comportement** des modèles »
raisonnement	1 (1)	oui	« le **raisonnement** moral des modèles en français »

TABLE 4 – Exemples de noms fréquents et/ou anthropomorphisants relevés dans le corpus.

3.2.4 Niveau global d'anthropomorphisation

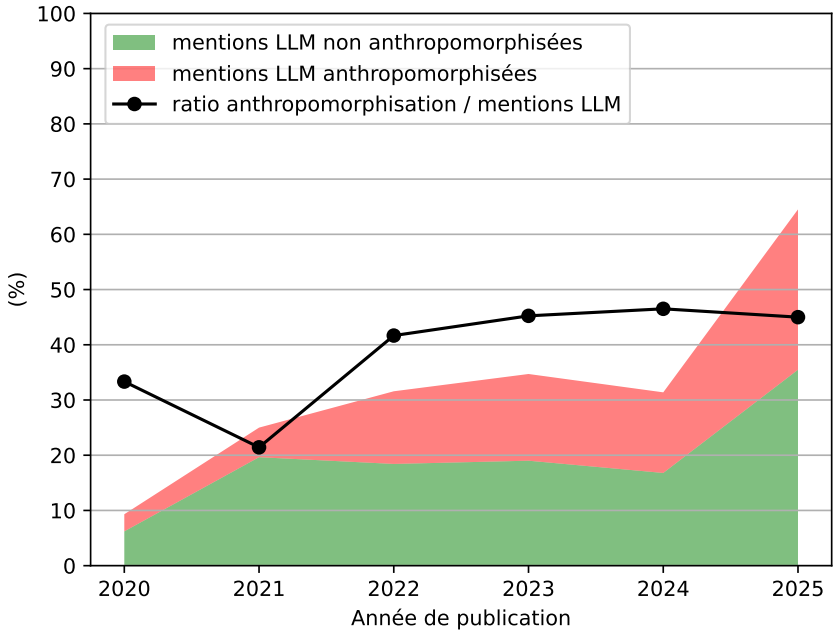

FIGURE 2 – Part annuelle des publications comportant des mentions de LLM non anthropomorphisées *vs.* comprenant des mentions de LLM anthropomorphisées au sein des conférences CORIA-JEP-TALN-RECITAL.

Sur les 258 articles CORIA-JEP-TALN-RECITAL faisant mention de LLM, nous trouvons que 147 (57,0 %) ont un score global d'anthropomorphisation égal à 0, correspondant à l'absence de structures anthropomorphisantes (cf. Section 2.3). Parmi les 111 (43,0 %) publications restantes, le score global d'anthropomorphisation moyen est de 1,45, avec un écart-type de 0,78 et un score maximal de 6, atteint par un seul article. La distribution temporelle entre articles comprenant des mentions anthropomorphisées (en considérant un seuil de score global de 1) et non anthropomorphisées est présentée en Figure 2, et ne présente pas de variations significatives d'une année à l'autre [18].

18. Un test exact de Fisher (préféré en cas de faibles effectifs) retourne $p \approx .604$.

4 Discussion

Nous avons proposé une méthode pour détecter des cas d'anthropomorphisation des grands modèles de langue dans des textes en français, que nous avons appliquée à un corpus d'articles CORIA-JEP-TALN-RECITAL publiés entre 2020 et 2025. Insistons sur le fait que cette étude exploratoire se veut à la fois une illustration pratique des intérêts de notre approche, et une occasion d'ouvrir le dialogue sur les questions conjointes de l'omniprésence et de l'anthropomorphisation des LLM dans notre discipline. Nous pensons en effet qu'une approche « TAL pour le TAL » (Francopoulo *et al.*, 2015) est souhaitable pour rendre compte des dynamiques qui animent actuellement le domaine. Il ne s'agit en aucun cas de pointer le doigt sur la communauté francophone ; au contraire, si la présence des LLM s'y renforce (par ailleurs dans une moindre mesure qu'à ACL), l'anthropomorphisation y demeure stable dans le temps, et malgré tout relativement modérée à la lecture des exemples du corpus. À l'heure où la parole scientifique est remise en question, voire attaquée directement (notamment aux États-Unis), il est fondamental que nous continuions à porter un discours précis qui rend compte fidèlement de la réalité de nos objets de recherche.

La question persiste malgré tout de s'accorder sur ce qui constitue ou non des cas d'anthropomorphisation avérée, et les travaux d'Inie *et al.* (2026) portent à ce sujet une vision plus radicale que celle de Lonke *et al.* (2025). Est-ce à dire que l'anthropomorphisation se cache dans toutes nos expressions ? La ligne à tracer entre métaphores « standard » pour les initiés du TAL, et expressions franchement anthropomorphisantes peut en effet diviser : les LLM *apprennent*-ils quoi que ce soit ? Quel rapport entre l'organisation fonctionnelle de notre cerveau et les *neurones* de GPT-5 ? S'il peut sembler difficile de se séparer de certaines expressions bel et bien ancrées dans la communauté, Inie *et al.* (2026) avancent que cet effort permet à la fois de mieux armer son propre regard de lecteur contre les expressions anthropomorphisantes, et de partager des connaissances scientifiques de manière plus précise. *A minima*, l'emploi de guillemets permet de marquer une distance symbolique avec des termes aussi flous que « l'IA ». Pour notre part, nous souhaitons attirer l'attention sur des termes plus récemment ancrés (donc dés-ancrables), selon nous exagérément psychologisants et mal définis, tels qu'*hallucinations* ou *alignement moral* (ou *culturel*, de *valeurs*, etc.) des LLM : ces termes ont-ils émergé de la communauté du TAL ? Ou bien s'agit-il d'expressions poussées par les grandes entreprises de la tech, par ailleurs de plus en plus représentées dans nos conférences majeures (Abdalla *et al.*, 2023) ? Au sujet du terme d'*alignement* des LLM, notons que Kirk *et al.* (2023) le qualifient de « signifiant vide » (*empty signifier*) (Laclau, 1996) ; cependant que ce vide sert précisément les intérêts d'acteurs promouvant une « universalité » des LLM vouée à demeurer abstraite. Pourtant, les principes initialement mobilisés derrière ce concept semblent sans appel : nous souhaitons évidemment empêcher que quiconque soit discriminé ou lésé du fait d'utilisations inadaptées de modèles de langue.

Ainsi, au-delà d'une définition univoque de l'anthropomorphisation (sans doute difficilement atteignable), on comprend que c'est de principes directeurs partagés et discutés collectivement dont nous aurions le plus besoin. Quel type de technologies souhaitons-nous voir se développer en TAL dans les prochaines années, pour effectuer quelles tâches, pour traiter de quelles questions linguistiques ? Inévitablement, déplorer l'omniprésence des LLM revient encore à parler des LLM. Aussi, nous souhaiterions plaider pour une ré-appropriation par les chercheurs en TAL à la fois des discours sur nos objets de recherche, et de la liberté de choisir si, et comment les étudier. Ce faisant, nous n'appauvrissons en rien les possibilités scientifiques, au contraire, nous remettons en lumière le « L » de « TAL » [19], et de « LLM ».

19. Pour reprendre une formulation de Véronique Hoste, présentatrice invitée à TALN en 2024.

Remerciements

Ce travail a bénéficié du soutien du projet ANR InExtenso (ANR-23-IAS1-0004).

Références

ABDALLA M., WAHLE J. P., RUAS T., NÉVÉOL A., DUCEL F., MOHAMMAD S. & FORT K. (2023). The elephant in the room : Analyzing the presence of big tech in natural language processing research. In A. ROGERS, J. BOYD-GRABER & N. OKAZAKI, Éds., *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1 : Long Papers)*, p. 13141–13160, Toronto, Canada : Association for Computational Linguistics. DOI : 10.18653/v1/2023.acl-long.734.

ABERCROMBIE G., CERCAS CURRY A., DINKAR T., RIESER V. & TALAT Z. (2023). Mirages. on anthropomorphism in dialogue systems. In H. BOUAMOR, J. PINO & K. BALI, Éds., *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, p. 4776–4790, Singapore : Association for Computational Linguistics. DOI : 10.18653/v1/2023.emnlp-main.290.

CHENG M., GLIGORIC K., PICCARDI T. & JURAFSKY D. (2024). Anthroscore : A computational linguistic measure of anthropomorphism. In Y. GRAHAM & M. PURVER, Éds., *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1 : Long Papers)*, p. 807–825, St. Julian's, Malta : Association for Computational Linguistics. DOI : 10.18653/v1/2024.eacl-long.49.

COLL ARDANUY M., NANNI F., BEELEN K., HOSSEINI K., AHNERT R., LAWRENCE J., MCDONOUGH K., TOLFO G., WILSON D. C. & MCGILLIVRAY B. (2020). Living machines : A study of atypical animacy. In D. SCOTT, N. BEL & C. ZONG, Éds., *Proceedings of the 28th International Conference on Computational Linguistics*, p. 4534–4545, Barcelona, Spain (Online) : International Committee on Computational Linguistics. DOI : 10.18653/v1/2020.coling-main.400.

DE MARNEFFE M.-C., MANNING C. D., NIVRE J. & ZEMAN D. (2021). Universal dependencies. *Computational Linguistics*, **47**(2), 255–308. DOI : 10.1162/coli_a_00402.

DESHPANDE A., RAJPUROHIT T., NARASIMHAN K. & KALYAN A. (2023). Anthropomorphization of ai : Opportunities and risks. In *Proceedings of the Natural Legal Language Processing Workshop 2023*, p. 1–7, Singapore : Association for Computational Linguistics. DOI : 10.18653/v1/2023.nllp-1.1.

DEVLIN J., CHANG M.-W., LEE K. & TOUTANOVA K. (2019). Bert : Pre-training of deep bidirectional transformers for language understanding. In J. BURSTEIN, C. DORAN & T. SOLORIO, Éds., *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies, Volume 1 (Long and Short Papers)*, p. 4171–4186, Minneapolis, Minnesota : Association for Computational Linguistics. DOI : 10.18653/v1/N19-1423.

DEVRIO A., CHENG M., EGEDE L., OLTEANU A. & BLODGETT S. L. (2025). A taxonomy of linguistic expressions that contribute to anthropomorphism of language technologies. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, p. 1–18, New York, NY, USA : Association for Computing Machinery. DOI : 10.1145/3706598.3714038.

FELLBAUM C. (1998). *WordNet : An Electronic Lexical Database*. MIT Press.

FLORIDI L. & NOBRE A. C. (2024). Anthropomorphising machines and computerising minds : The crosswiring of languages between artificial intelligence and brain cognitive sciences. *Minds and Machines*, **34**(1), 5. DOI : 10.1007/s11023-024-09670-4.

FRANCOPOULO G., MARIANI J. & PAROUBEK P. (2015). Nlp4nlp : The cobbler's children won't go unshod. *D-Lib Magazine*, **21**(11/12). DOI : 10.1045/november2015-francopoulo.

HUNGER F. (2023). *Unhype Artificial 'Intelligence' ! A proposal to replace the deceiving terminology of AI*. DOI : 10.5281/zenodo.7524493.

INIE N., DRUGA S., ZUKERMAN P. & BENDER E. M. (2024). From "ai" to probabilistic automation : How does anthropomorphization of technical systems descriptions influence trust ? In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '24, p. 2322–2347, New York, NY, USA : Association for Computing Machinery. DOI : 10.1145/3630106.3659040.

INIE N., ZUKERMAN P. & BENDER E. M. (2026). De-anthropomorphizing "ai" : From wishful mnemonics to accurate nomenclature. *First Monday*. DOI : 10.5210/fm.v31i2.14366.

KIRK H. R., VIDGEN B., RÖTTGER P. & HALE S. A. (2023). The empty signifier problem : Towards clearer paradigms for operationalising "alignment" in large language models. In *37th Conference on Neural Information Processing Systems (NeurIPS)*. DOI : 10.48550/arXiv.2310.02457.

LACLAU E. (1996). *Emancipation(s)*. Verso. Google-Books-ID : QFwiHJP8mdUC.

LIU Y., OTT M., GOYAL N., DU J., JOSHI M., CHEN D., LEVY O., LEWIS M., ZETTLEMOYER L. & STOYANOV V. (2019). Roberta : A robustly optimized bert pretraining approach. arXiv :1907.11692 [cs], DOI : 10.48550/arXiv.1907.11692.

LONKE D., BLOEM J. & SOMMERAUER P. (2025). Anthroset : a challenge dataset for anthropomorphic language detection. In *Proceedings of the First Interdisciplinary Workshop on Observations of Misunderstood, Misguided and Malicious Use of Language Models*, p. 27–39 : Incoma Ltd. Shoumen, BULGARIA. DOI : 10.26615/978-954-452-101-1-003.

PLACANI A. (2024). Anthropomorphism in ai : hype and fallacy. *AI and Ethics*, **4**(3), 691–698. DOI : 10.1007/s43681-024-00419-4.

RAZA M. O. & MEGHJI A. F. (2025). Anthropomorphizing ai : A multi-label analysis of public discourse on social media. In P. PRZYBYŁA, M. SHARDLOW, C. COLOMBATTO & N. INIE, Éds., *Proceedings of Interdisciplinary Workshop on Observations of Misunderstood, Misguided and Malicious Use of Language Models*, p. 64–73, Varna, Bulgaria : INCOMA Ltd., Shoumen, Bulgaria.

RYAZANOV I., ÖHMAN C. & BJÖRKLUND J. (2024). How chatgpt changed the media's narratives on ai : A semi-automated narrative analysis through frame semantics. *Minds and Machines*, **35**(1), 2. DOI : 10.1007/s11023-024-09705-w.

SHARDLOW M. & PRZYBYŁA P. (2024). Deanthropomorphising nlp : Can a language model be conscious ? *PLOS ONE*, **19**(12), e0307521. DOI : 10.1371/journal.pone.0307521.

SHARDLOW M., WILLIAMS A., ROADHOUSE C., VENTIROZOS F. & PRZYBYŁA P. (2025). Learn, achieve, predict, propose, forget, suffer : Analysing and classifying anthropomorphisms of llms. In *Proceedings of the First Interdisciplinary Workshop on Observations of Misunderstood, Misguided and Malicious Use of Language Models*, p. 86–94 : Incoma Ltd. Shoumen, BULGARIA. DOI : 10.26615/978-954-452-101-1-010.

THUILIER J., DA CUNHA Y., FAGHIRI P. & PINEAU L. (2024). Prototypicalité des arguments et animéité dans le domaine postverbal du français. In *Langues et Langage à la croisée des disciplines*, Paris, France : GDR LLcD et Sorbonne Université.

TOUCHENT R., ROMARY L. & DE LA CLERGERIE E. (2023). CamemBERT-bio : Un modèle de langue français savoureux et meilleur pour la santé. In C. SERVAN & A. VILNAT, Éds., *Actes de CORIA-TALN 2023. Actes de la 30e Conférence sur le Traitement Automatique des Langues Naturelles (TALN), volume 1 : travaux de recherche originaux – articles longs*, p. 323–334, Paris, France : ATALA.

VAN DEN EYNDE K. & BLANCHE-BENVENISTE C. (1978). Syntaxe et mécanismes descriptifs : Présentation de l'approche pronominale. *Cahiers de lexicologie*, **1**(32), 3–27. DOI : 10.15122/ISBN.978-2-8124-4285-8.P.0005.

VAN DEN EYNDE K. & MERTENS P. (2010). Le dictionnaire de valence DICOVALENCE : manuel d'utilisation version 2.0.

VAN DEN EYNDE K., MERTENS P. & EGGERMONT C. (2017). Dicovalence. DOI : https ://doi.org/10.82270/dicovalence/v1.

VERWILGHEN L. (2025). *When Algorithms 'Think' : A Computational Analysis of Anthropomorphism in AI News Coverage*. Master's thesis.

YANG Z., DAI Z., YANG Y., CARBONELL J., SALAKHUTDINOV R. & LE Q. V. (2019). Xlnet : generalized autoregressive pretraining for language understanding. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, p. 5753–5763, Red Hook, NY, USA : Curran Associates Inc. DOI : 10.5555/3454287.3454804.

Annexes

partie du discours	indices lexicaux
verbes	accomplir, adversaire, aider, analyser, apprendre, assister, attendre, **capturer**, choisir, classer, collaborer, communiquer, **compatir**, comprendre, conclure, **confondre**, connaître, conseiller, contribuer, coordonner, croire, **créer**, demander, différencier, dire, discuter, distinguer, décider, déduire, déterminer, élaborer, enfreindre, enseigner, entendre, espérer, expliquer, **halluciner**, identifier, imaginer, **imiter**, **incarner**, inférer, **inquiéter**, insulter, interpréter, justifier, manipuler, mentir, mémoriser, **oublier**, parler, partenaire, penser, percevoir, planifier, prédire, préférer, **préoccuper**, **prévoir**, raisonner, reconnaître, retenir, **répondre**, résoudre, réussir, rêver, savoir, sentir, **simuler**, souvenir, stratégiser, suggérer, suivre, superviser, sélectionner, tricher, tromper, voir
adjectifs	aligné, amical, artistique, attentif, attentionné, autonome, aveugle, capable, communicatif, compétent, confiant, confus, conscient, conseiller, créatif, drôle, désorienté, **empathique**, **équitable**, **étranger**, expert, **fatigué**, **fiable**, **français**, honnête, indépendant, intelligent, malhonnête, malveillant, manipulateur, **natif**, **relecteur**, responsable, réactif, sensible, sincère, sourd, **thérapeutique**, trompeur, vulnérable
noms	action, aide, **analyse**, art, artiste, assistance, assistant, capacité, compagnon, comportement, **compréhension**, compétence, confiance, **confusion**, connaissance, conscience, contresens, coopération, coordination, **création**, **culture**, **docteur**, difficulté, décision, délibération, désir, écriture, **éditeur**, employé, employée, esprit, étudiant, **explication**, feedback, force, **hallucination**, humour, idée, **imitation**, incompréhension, **inquiétude**, **intelligence**, **intention**, interprétation, ironie, **lien**, malentendu, manager, manipulation, maîtrise, mensonge, **morale**, **médecin**, **neurone**, **oubli**, personnalité, perspicacité, poète, poésie, professeur, **préférence**, **préoccupation**, **prévision**, **psychologue**, raisonnement, réaction, **réponse**, savoir, **sentiment**, **simulation**, souvenir, **suggestion**, tendance, **thérapeute**, tuteur, **valeur**

TABLE 5 – Indices lexicaux traduits en français d'après la taxonomie de Lonke *et al.* (2025). En gras : ajouts ultérieurs suivant des exemples relevés par Inie *et al.* (2026), ainsi que des observations personnelles.