

Humanistes, mais lesquels ? Entre représentations lexicographiques du substantif *humaniste* et usages discursifs : étude pilote

Agnieszka Kaliska^{1*}, Bernadeta Wojciechowska¹

¹ Université Adam Mickiewicz de Poznań, 4 Al. Niepodległości, 61874 Poznań, Pologne

Abstract. Cette étude compare l'évolution sémantique du mot *humaniste* telle qu'elle est reflétée dans les données des dictionnaires avec celle mise en évidence par des exemples de corpus. En s'appuyant sur une approche *corpus-driven* et sur des données provenant de la collection French Web disponible sur la plateforme SketchEngine, elle explore également les implications du déséquilibre des sources dans l'analyse de corpus. L'étude est structurée en deux étapes : (i) un résumé de l'évolution sémantique du mot *humaniste* fondé sur les sens attestés dans les dictionnaires de langue française ; et (ii) une analyse des données de corpus, mettant en évidence un déséquilibre quantitatif entre les sources et son impact potentiel sur les résultats. L'étude se conclut par la formulation de trois hypothèses de recherche fondées sur l'approche *corpus-driven* adoptée, hypothèses qui devront être vérifiées au moyen d'une annotation systématique et d'une interprétation des données.

1 Introduction

Cette étude retrace l'évolution sémantique du mot *humaniste*, telle qu'elle apparaît dans les dictionnaires, et la compare à celle mise en évidence à partir des exemples tirés du corpus. Elle repose sur une approche *corpus-driven* et s'appuie sur les données de la collection *French Web* disponibles sur la plateforme SketchEngine.

Notre réflexion s'ouvre sur un résumé d'une étude consacrée à l'évolution sémantique du mot *humaniste*, fondée sur les acceptions attestées dans les principaux dictionnaires de la langue française [1]. Ce cadrage constitue le point de départ de l'analyse quantitative du corpus, dont l'objectif est d'identifier les valeurs discursives potentielles du mot *humaniste* ainsi que les effets de sens qu'il peut produire en contexte médiatique. L'interprétation qui en découlera constitue une première hypothèse fondée sur des données quantitatives, conformément à une

* Corresponding author: agnieszka.kaliska@amu.edu.pl

approche *corpus-driven*, et pourra ensuite être approfondie au moyen d'une analyse qualitative des données annotées.

Cette analyse s'accompagne également d'une réflexion méthodologique sur les implications d'une surreprésentation de données issues d'une source au sein du corpus.

2 Aperçu critique des acceptions du substantif *humaniste* dans les dictionnaires français

Nous organisons les différentes acceptions du substantif *humaniste* selon trois axes délimités au cours d'une étude basée sur des définitions lexicales d'*humaniste* dans les dictionnaires français [1] : un axe éducatif, dans lequel l'*humaniste* est perçu comme un érudit maîtrisant les langues et littératures classiques ou, plus largement, les sciences humaines ; un axe historique, centré sur les figures majeures de la Renaissance ; et enfin, un axe philosophique, lié à une posture valorisant l'autonomie et la dignité humaines.

2.1 Axe éducatif : humaniste, érudit des humanités classiques

La première mention lexicographique du substantif *humaniste* se trouve dans le *Dictionnaire* de Richelet qui le définit comme « celui qui sait les humanitez, qui sait les Poètes & les Orateurs » [2]. Cette acception se retrouve dans le *Dictionnaire universel* de Furetière ainsi que dans le *Dictionnaire de l'Académie* (dès 1718), qui décrivent l'*humaniste* comme « celui qui sait bien ses humanitez » et « celui qui les enseigne » [3-4].

Au XIX^e siècle, les définitions – « qui étudie, sait, enseigne les humanités » [5] ; « celui qui sait bien les humanités – celui qui les enseigne » [6] ; « Celui qui étudie ou enseigne les humanités. *Un bon humaniste.* » [7] – maintiennent ce lien avec la formation classique, sans pour autant détailler les contenus spécifiques que recouvrent les humanités. Ce n'est que les éditions plus récentes du DAF, avec la formulation « celui qui a une connaissance approfondie de la langue et de la littérature grecque et romaine », et puis avec la formulation « Personne qui a une connaissance approfondie de la langue et de la littérature grecques et latines. *Un grand, un savant humaniste* », qui précisent explicitement « la langue et la littérature grecque et romaine », la dernière édition classant cette acception en deuxième position [4]. Le TLFi reprend cette dimension éducative, il la conçoit toutefois comme une extension d'un sens principal, centré sur la tradition de la Renaissance, en distinguant « spécialiste des lettres grecques et latines » et « élève qui fait ses humanités » [8]. La même dualité est marquée dans le *Grand Larousse* où l'*humaniste* est présenté à la fois comme érudit et élève des sections classiques : « 1. Érudit, lettré versé dans la connaissance des langues et des littératures anciennes : *Un savant humaniste.* [...] 3. Élève qui faisait des études de lettres anciennes (section classique des lycées et collèges) : *J'étais à ma manière un bon petit humaniste* [...] » [9]. La continuité du lien avec la formation scolaire est aussi attestée dans les dictionnaires contemporains, comme *Le Robert* qui définit l'*humaniste* comme « spécialiste des langues et littératures grecques et latines » (deuxième acception) et précise qu'il peut désigner aussi un « lettré de la Renaissance » [10].

Les exemples d'usage lexical intègrent fréquemment des adjectifs valorisants (*bon, excellent, grand, savant*). Une seule exception à cet égard est attestée dans la 5^e édition du DAF où on note la coexistence d'*excellent humaniste* et *médiocre humaniste*. Par ailleurs, l'usage des adjectifs en fonction de l'épithète devant le mot *humaniste* fut recommandé par l'Académie depuis sa 5^e édition [4], on en voit la trace aussi dans *Littré* [11]. Ce dernier est le plus étendu quant aux acceptions relevant de l'axe éducatif et en distingue trois acceptions :

1. Celui qui étudie les humanités dans un lycée, collège.
2. Celui qui sait, qui enseigne les humanités ; en ce sens il ne se dit guère qu'avec une épithète. *Un bon, un savant humaniste. Jeune encore et grand humaniste* [...]
3. En Allemagne, les humanistes, ceux qui font de l'étude des langues classiques la base de l'enseignement ; par opposition à réalistes, ceux qui se livrent à l'industrie.

On retrouve en troisième position une acception collective des *humanistes* opposée aux *réalistes* dans le contexte allemand – probablement, reflet des transformations des structures éducatives au XIX^e siècle, notamment la montée des formations techniques [11].

2.2 Axe historique : humaniste, figure emblématique de la Renaissance

L'association explicite de l'humaniste aux érudits de la Renaissance s'impose à partir du XIX^e siècle, bien que les humanistes italiens déjà utilisaient le terme *rinascita* (fr. *renaissance*) pour qualifier leur époque. Le TLFi insiste d'abord sur ce contexte et définit l'humaniste comme « érudit de la Renaissance qui, s'inspirant des auteurs antiques, a exalté la dignité de l'esprit humain ». Cette définition est étayée par des exemples précis, souvent sous forme de citations, notamment : « les hommes marquants de cet âge [le XVe siècle en Italie] sont les humanistes, restaurateurs passionnés des belles-lettres grecques et latines » [8].

L'édition actuelle du DAF confirme cette perspective, présentant l'humaniste comme « penseur des XVe et XVI^e siècles, versé dans l'étude érudite de la philosophie et de l'art antiques », et cite Marsile Ficin, Érasme ou Guillaume Budé [4]. *Larousse* en ligne propose une définition similaire en 3^e position : « nom donné aux savants et littérateurs qui, au XVe et XVI^e siècles, remirent en honneur les chefs-d'œuvre de l'Antiquité classique » [12]. Enfin, *Le Robert* en ligne met en avant trois dimensions essentielles – la période historique, l'activité intellectuelle et l'influence sociale – en définissant l'humaniste comme « lettré de la Renaissance qui se consacrait à l'étude et à la diffusion des auteurs antiques » [10] [voir aussi 15, 16]. Cette définition synthétise érudition, temporalité et rôle social, repris de façon plus détaillée dans le *Grand Larousse*, qui insiste sur la restauration et la transmission des idées antiques par les humanistes italiens et français [9].

2.3 Axe philosophique : humaniste, une posture idéologique

La dimension philosophique du terme se manifeste surtout dans les dictionnaires français en ligne, tels que *Le Robert* et *Larousse*, qui définissent l'humaniste comme un « partisan de l'humanisme », respectivement en première et en deuxième position, autrement que dans les versions imprimées où l'acception « partisan de l'humanisme » suit celle associée à l'éducation dans *Le Petit Robert*, alors qu'elle la précède dans *Larousse* imprimé [10, 12, 13, 14]. Cette définition est également la première dans *Le Dico* (adapté du Wiktionnaire), où elle est illustrée par l'exemple *Érasme était un véritable humaniste* [16]. Dans le TLFi, cette acception, clairement dissociée des dimensions éducative et historique, est formulée ainsi : « Philosophe qui, considérant l'homme comme la mesure de toute chose, cherche à l'épanouir en prônant le développement des facultés proprement humaines » [8]. Cette définition, sans référence à la Renaissance, présente l'humaniste comme un penseur centré sur la valorisation et l'émancipation de l'individu. Les exemples tirés de Massis et Sartre proposés dans cette entrée apportent une nuance importante :

(1) M. Anatole France en est venu à se faire de la vie de l'homme une image mesquine et pitoyable. Cet humaniste est profondément inhumain. Les hauts sentiments de notre nature morale lui sont étrangers. (Massis, *Jugements*, 1923, p. 163)

(2) L’humaniste dit « de gauche » a pour souci principal de garder les valeurs humaines ; il n’est d’aucun parti, parce qu’il ne veut pas trahir l’humain mais ses sympathies vont aux humbles : c’est aux humbles qu’il consacre sa belle culture classique. (Sartre, *La Nausée*, 1938, p. 150).

Massis met en garde contre un humanisme potentiellement égocentrique et déshumanisant, tandis que Sartre associe l’humanisme à un engagement politique de gauche, tourné vers la défense des classes populaires et la transmission de la culture classique. Cette dimension sociale et idéologique dépasse ainsi la simple posture philosophique. L’idée de reconnaissance de l’homme comme valeur essentielle revient également dans le *Grand Larousse*, en quatrième position, et dans l’édition actuelle du DAF où on met en lumière une continuité entre l’héritage humaniste classique et une posture philosophique contemporaine, plus universelle, déliée des références antiques, et centrée sur la dignité et le développement de la personne humaine (troisième position au sein de l’entrée) [4, 9]. Elle fait également l’objet de la définition proposée par *Le Dictionnaire Hachette* : « Personne qui professe un humanisme » [15].

2.4 Bilan

Les frontières entre les axes distingués ne sont guère étanches. L’analyse menée permet toutefois de constater la primauté de l’axe éducatif, non seulement dans la manière même de concevoir le sens du substantif *humaniste* dans les dictionnaires, mais aussi dans la genèse de son évolution sémantique ultérieure. En effet, deux autres axes semblent dériver de l’axe éducatif. Leur émergence est, pourtant, naturellement plus tardive, car déterminée par le développement historique. L’association du substantif *humaniste* aux érudits de la Renaissance devient explicite à partir du XIXe siècle, à mesure que le terme *Renaissance* s’établit dans l’usage comme désignation de la période de l’histoire européenne où s’est manifesté un intérêt renouvelé pour les langues, la culture et la philosophie de l’Antiquité. L’axe historique et l’axe philosophique étant étroitement liés, la dimension philosophique acquiert progressivement une plus grande autonomie et cesse d’être exclusivement associée à la Renaissance. Les résultats de cette première partie de notre étude sont présentés sur la Fig. 1.

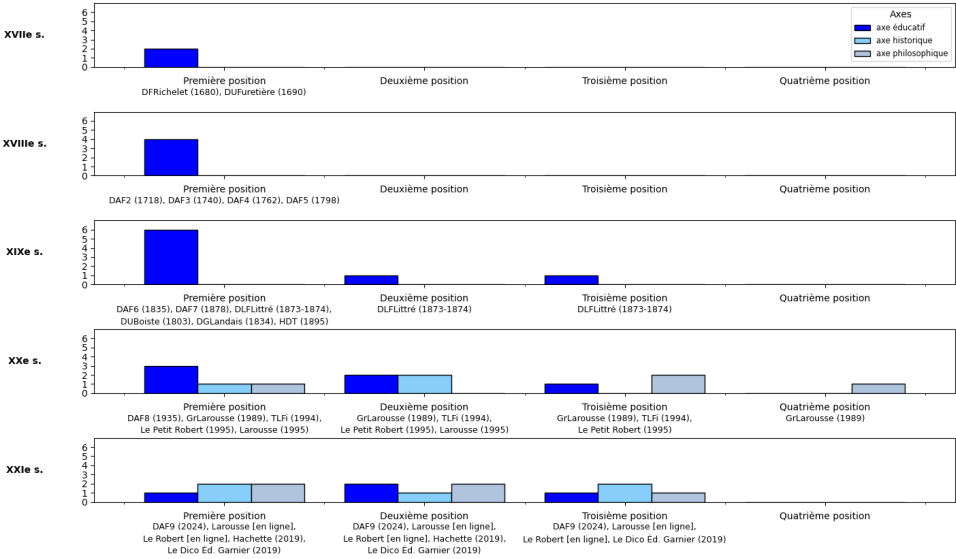


Fig. 1. Répartition des axes au sein des entrées analysées.

Les échelles placées à gauche indiquent le nombre des dictionnaires dans lesquels apparaissent des acceptions relevant des axes éducatif, historique et philosophique, selon leurs positions dans les entrées analysées. Les numérotations d'origine (1, 2...), les balisages (A, B... ; a, b...) et les marqueurs tels que *Philos.* ou *Hist.* ont été normalisés pour être pris en compte dans les calculs.

3 Choix méthodologiques et corpus

Notre analyse est menée dans l'optique *corpus-driven*, approche en linguistique de corpus permettant de formuler des hypothèses avant de passer à l'analyse qualitative des données collectées [17, 18]. Dans cette approche inductive, l'usage est privilégié, de sorte que toute hypothèse formulée ne saurait être validée que si le corpus analysé répond au principe de référentialité. Dans notre cas, l'analyse porte sur les emplois du substantif *humaniste* dans l'espace web. Bien que circonscrit, un tel corpus se caractérise par une diffusion massive, liée à l'essor des technologies numériques. Il présente toutefois un certain nombre de limites : les textes sont des échantillons bruts issus des sites web, collectés automatiquement, *crawled* ou *scraped* tels quels, sans attention portée à leur qualité. C'est pourquoi, le caractère disparate et éphémère des données, le risque élevé d'erreurs et de bruit sans valeur ajoutée, ainsi que la présence de contenus dupliqués conduisent certains auteurs à qualifier le web d'espace « anarchique » (cf. « The Web is anarchic », [19]).

Voulant donc faire un premier pas vers le complément de l'analyse fondée exclusivement sur des données lexicographiques relatives au substantif *humaniste*, nous avons lancé une recherche sur la plateforme SketchEngine en restreignant le corpus aux données issues de sites web de neuf titres de la presse numérique française, à savoir :

- Quotidiens : *Le Monde* (sensibilité de centre-gauche), *Le Parisien* et *Le Figaro* (marqués à droite), *Libération* (marqué à gauche), *L'Humanité* (journal communiste) ;
- Hebdomadaires d'actualité : *L'Express* et *Le Point* (sensibilité de centre-droite, libéraux), *Nouvelobs.com* (marqué à gauche) ;
- Mensuel : *Le Monde diplomatique* (sensibilité de gauche critique).

Rappelons que, créée initialement dans le souci de soutenir l'enseignement des langues étrangères, la linguistique de corpus a pour objectif d'analyser l'emploi des mots dans des textes authentiques. Elle vise ainsi à identifier des collocations et à en dégager des généralisations quant au fonctionnement des unités lexicales [18]. C'est à ce point qu'intervient la notion de représentativité, pourtant illusoire [20, 21].

La notion de représentativité parfaite d'un corpus relève de l'utopie, car il est impossible de capturer la totalité des variations de la langue, toutes proportions gardées, dans un seul et même ensemble de données. Il n'en demeure pas moins que la constitution d'un corpus fondée sur le principe d'équilibre entre genres textuels, registres, etc., en fonction de l'objectif visé, constitue une démarche nécessaire afin d'assurer, *dans la mesure du possible*, la représentativité des données et, par là même, leur efficience heuristique [21].

Dans cette perspective, bien que les titres retenus pour notre étude ne constituent qu'un échantillon de médias numériques français, celui-ci a été composé dans le souci de maintenir un équilibre entre des publications appartenant à deux sensibilités opposées (ou à des proximités) et sélectionnées de manière à assurer, dans la mesure du possible, un équilibre également du point de vue du nombre de lecteurs, ces paramètres étant pris en compte sur la base des données actuelles disponibles sur le site : [22].

Notre requête a été poursuivie sur quatre sous-corpus de la collection française *French Web (frTenTen)* : 2012, 2017, 2020 et 2023 [23], ce qui nous a permis d'obtenir, outre les résultats regroupés ensuite par titres, une vue d'ensemble sur une diachronie courte. Bien que la requête effectuée sur le corpus ait été paramétrée de manière à ne retenir que le

substantif *humaniste* (lemme ; voir Fig. 2), à l'exclusion donc des emplois adjectivaux, une vérification manuelle des données s'est avérée nécessaire. À titre d'illustration, ont été éliminées quelques dizaines de concordances dans lesquelles *humaniste* était employé comme épithète (p. ex. : *cause humaniste*, *démarche humaniste*, *dimension humaniste*, *discours humaniste*, *éthique humaniste*, *idée humaniste*, *mode humaniste*, *loi humaniste*, *posture humaniste*, *principe humaniste*, *projet humaniste*, *propos humaniste*, *raison humaniste*, *sens humaniste*, *valeur humaniste*).

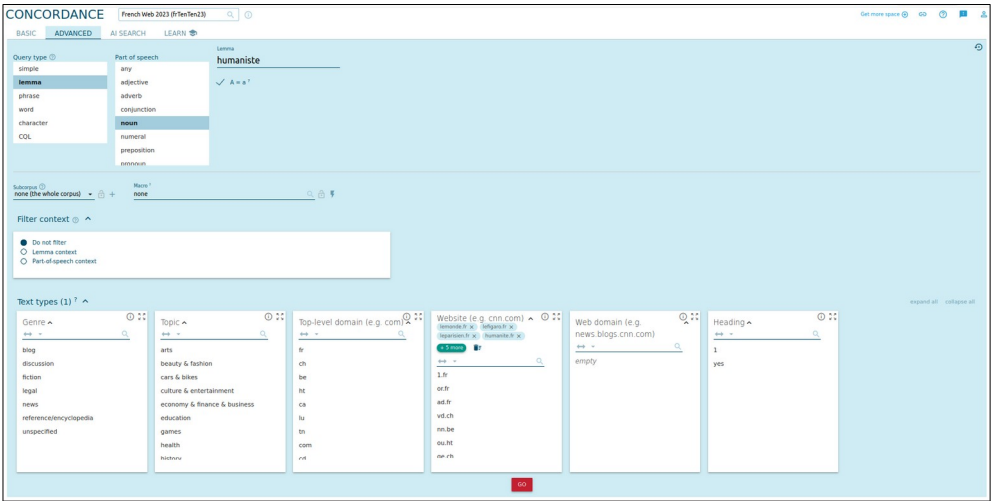


Fig. 2. Paramètres de la requête [SketchEngine, dernier accès : le 2 Janv. 2026].

3.1 Enjeux de la représentativité des données

Les statistiques du corpus (des concordances) sont détaillées dans la Table 1. Les chiffres obtenus dévoilent un déséquilibre important entre les données venant des médias de sensibilités opposées, principalement en raison de la surreprésentation des exemples provenant du site du quotidien *Le Figaro*.

Table. 1. Statistiques des concordances par titres calculées à l'aide d'*Hyperbase* [24].

Sous-corpus	Occurrences	Lexèmes	Hapax	Occurrences d' <i>humaniste</i>
humanite.fr	1875	715	103	38
lefigaro.fr	104619	12517	4784	2150
lemonde.fr	16789	4220	1266	346
leparisien.fr	1400	592	87	27
lepoint.fr	12308	3308	787	253
lexpress.fr	8763	2565	554	183
liberation.fr	3804	1265	185	81
monde-diplomatique.fr	1166	484	52	26

nouvelobs.com	11899	3162	814	249
Total des occurrences: 162623				
Total des occurrences du substantif <i>humaniste</i> (lemme) : 3339				

La répartition des données, selon la source et l’année, telle que visualisée sur la carte de chaleur (voir Fig. 3), met en évidence la prépondérance d’une source particulière.

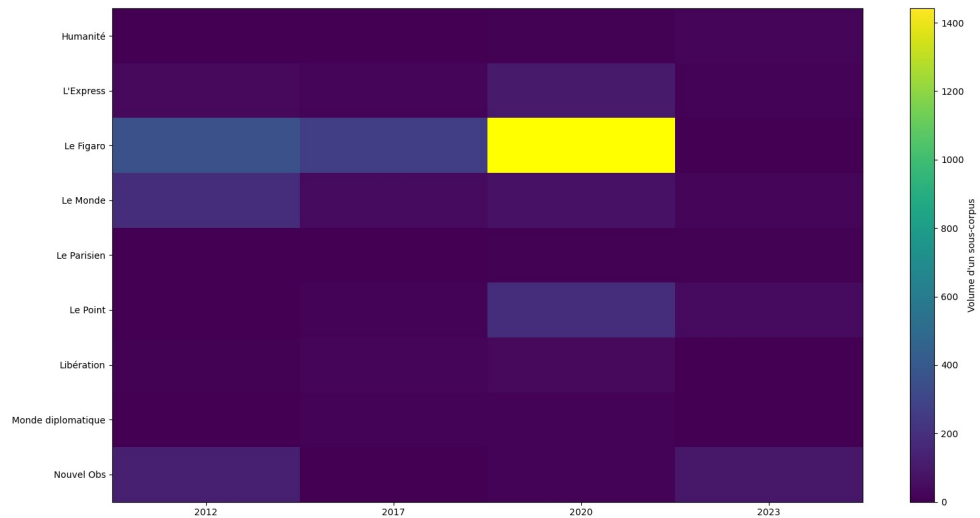


Fig. 3. Répartition des données selon la source et l’année.

Cette surreprésentation, qui ne peut être constatée qu’a *posteriori*, s’explique néanmoins par une taille brute (*token coverage*) prépondérante des données issues du site lefigaro.fr qui sont stockées dans le corpus *French Web (frTenTen)* sur la plateforme SketchEngine. Il s’agit en particulier de la sous-partie relative au corpus de l’année 2020 qui constitue 60 % de toutes les données en provenance du site *Le Figaro*, soit 40,9 % de l’ensemble sur lequel nous avons poursuivi notre requête (voir Fig. 4).

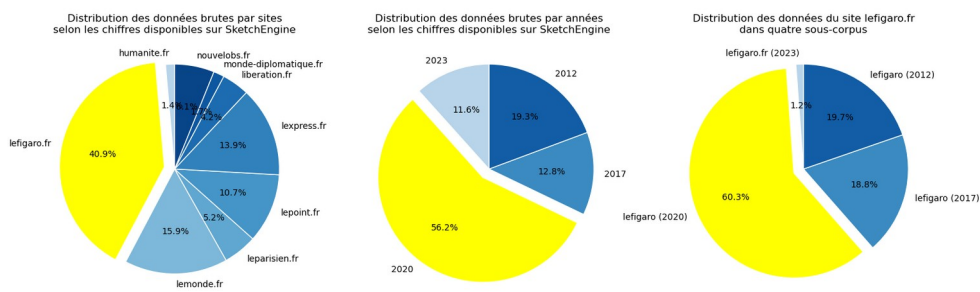


Fig. 4. Visualisation faite à l’aide de Python 3.10 à partir des chiffres disponibles sur SketchEngine, onglets : *Corpus Info* : *Text Type Analysis (Token Coverage)* [dernier accès : le 28 Mars 2026].

Avant même d’entamer l’analyse proprement dite, nous sommes donc confrontés à une observation méthodologique, quelque peu déconcertante, découlant de la structuration interne d’une ressource web, pourtant très populaire parmi les linguistes. Cette observation

Le présent chapitre est composé de deux sections. La première est consacrée à une analyse strictement quantitative. Elle repose sur le calcul des spécificités réalisé, selon la double démarche décrite ci-dessus, sur l'ensemble du corpus, puis sur le corpus réduit. Cette première section a pour objectif de nous orienter vers la formulation d'hypothèses sur la représentation discursive de l'*humaniste* dans le corpus web. La deuxième section amorce une première étape de l'analyse qualitative des données.

En utilisant des fonctionnalités d'*Hyperbase*, nous avons accédé aux co-occurents les plus *spécifiques* des mots-formes pivots *humaniste* et *humanistes* afin d'essayer de dégager par leur biais des connotations qui leur sont associées dans le discours (voir Fig. 5 et 6). Par *connotation* nous entendons une valeur secondaire, un sentiment positif ou négatif (p. ex. une ironie), qu'un mot peut suggérer ou éveiller chez le lecteur, selon le contexte.

[illegible][illegible]

Les adjectifs épithètes, ou substantifs homophones (p. ex. *républicain*), ainsi dévoilés, suggèrent un écart quant à l'usage du lexème *humaniste* dans le corpus par rapport aux exemples de dictionnaires. Celui-là semble mobilisé prioritairement pour désigner une identité politique (p. ex. *démocrate*, *républicain*, *progressiste*) ou idéologique (p. ex. *engagé*, *intransigeant*), autrement que dans les dictionnaires où ses dimensions éducative et

professionnelle (*celui qui enseigne ou apprend les humanités*) sont systématiquement mises en avant.

En outre, parmi les co-occurents les plus spécifiques, figure également l’adjectif axiologique *grand* à polarité positive. Or, de nombreux adjectifs appartenant à la modalité aléthique – également détectés (p. ex. *authentique, véritable, vrai* vs *autoproclamé, faux, prétendu, soi-disant*) – auraient, peut-être, pour effet de mettre en cause la sincérité ou la véracité des valeurs assimilées aux attitudes humanistes. Ainsi est-il possible que les énoncés tels que *grand humaniste* aient une portée ironique.

4.1.1. Calcul opéré sur l'ensemble du corpus

Nous présentons ci-dessous les données relatives aux dix co-occurents les plus spécifiques (signes de ponctuation exclus), parmi lesquels figurent certains adjectifs déjà évoqués (p. ex. *grand, engagé, faux*). Il est intéressant de les situer par rapport aux autres co-occurents les plus spécifiques, entre autres certains déterminants comme *cet* et *ces* qui figurent également sur cette liste – voir Table 2.

Table. 2. Co-occurents les plus spécifiques dans le corpus entier calculés à l’aide d’*Hyperbase*.

	humaniste				humanistes			
	Indice de spéci-ficité	Probabilité	Fré-quence corpus	Mot	Indice de spéci-ficité	Probabilité	Fré-quence corpus	Mot
1	24,38	2,36E-102	2088	un	16.46	9.88e-52	2672	les
2	14,79	5,02E-43	302	grand	16.12	6.71e-50	2201	des
3	7,31	5,36E-13	254	homme	13.3	8.14e-36	188	grands
4	7,03	3,28E-12	19	engagé	11.3	4.09e-27	411	ces
5	6,38	2,08E-10	125	cet	9.49	3.66e-20	524	sont
6	6,29	3,71E-10	5	tape	8.23	7.28e-16	93	faux
7	5,7	1.15e-8	20	démocrate	7.92	7.04e-15	23	pacotille
8	5,69	1.21e-8	18	progressiste	7.53	1.13e-13	237	Les
9	5,39	6.44e-8	8	crétin	7.31	5.08e-13	35	complet
10	5.38	6.76e-8	2464	l'	7.25	7.99e-13	8	lapin

Il est clair que les co-occurents détectés sont des chaînes de caractères séparées les unes des autres par des espaces blancs, la notion de *mot* n’étant qu’une construction théorique, et que le *mot* ne pourrait donc être détecté par une machine [17]. C’est uniquement par un retour au texte que nous sommes capables de constater quelle est la fonction de certaines formes que l’intuition juge d’obscures, comme dans le cas de la forme *tape* (voir Table 2) qui revient dans trois sous-corpus :

(3) Votre parti d’humaniste, on s’en tape. Dans humaniste il y a humain, dans humain il y a homme, dans homme il y a famille, dans famille il y a identité, dans identité, il y a nation. (lefigaro.fr, 2012, 2017 et 2020)

Il en est de même de la forme *lapin* qui s’avère être collocatif dans *en peau de lapin*, p. ex. :

(4) Durant les années 40 à Paris il y avait beaucoup d’humanistes en peau de lapin qui vont se mettre à plat ventre devant l’Autre. (lefigaro.fr, 2020)

4.1.2. Examen de l'impact induit par une surreprésentation des données

Le calcul des spécificités appliqué à l’ensemble du corpus nous permet de voir de plus près les contrastes entre différents sous-corpus. Ainsi, trois formes qui ont été détectées comme étant parmi les plus spécifiques du corpus – *ces*, *faux* et *grands* – s’avèrent prédominantes seulement dans une partie de celui-ci et biaisent alors les résultats globaux (voir Fig. 7).

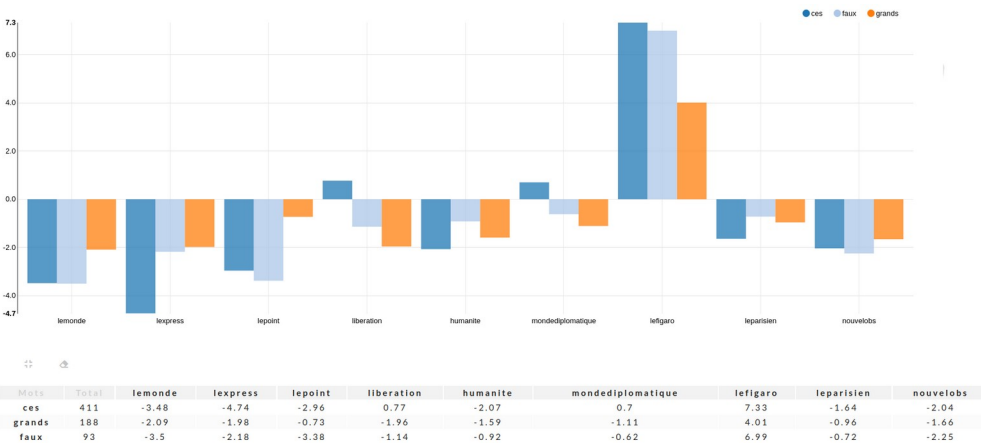


Fig. 7. Spécificités des chaînes *ces*, *faux* et *grands* calculées à l’aide d’Hyperbase.

L’hypothèse initiale qui s’en dégage suggère moins une évolution sémantique stabilisée du mot *humaniste* qu’un déplacement possible de ses connotations dans les usages attestés en ligne vers des emplois polémiques ou idéologiquement marqués. Cette hypothèse doit être envisagée avec prudence, dans la mesure où elle repose sur un corpus déséquilibré et ne constitue pas le schéma dominant à l’échelle de l’ensemble des données.

Il semble donc méthodologiquement relevant de confronter cette hypothèse avec les résultats issus d’une analyse analogue menée sur un corpus remanié, c’est-à-dire expurgé de sa sous-partie dominante. La Table 3 contient dix co-occurents les plus spécifiques après l’élimination de la sous-partie dominante du corpus.

Table 3. Co-occurents les plus spécifiques dans le corpus restreint calculés à l’aide d’Hyperbase.

	humaniste				humanistes			
	Indice de spéci-ficité	Probabilité	Fré-quence corpus	Mot	Indice de spéci-ficité	Probabilité	Fré-quence corpus	Mot
1	15.51	1,09E-46	893	un	11,06	4,08E-26	712	des

2	10,84	3,06E-25	114	grand	7,3	5,48E-13	838	les
3	5,7	1,14E-8	932	l'	6,15	8,42E-10	18	progressistes
4	5,12	2,67E-7	6	réglée	5,55	2,65E-8	4	taille
5	5,1	2,98E-7	8	crétin	5,55	2,65E-8	4	pleurent
6	5	4,79E-7	10	engagé	5,46	4,34E-8	134	sont
7	5	4,79E-7	10	démocrate	5,29	1,11E-7	5	différence
8	4,49	0,0000061	60	cet	4,77	0,00000152	20	Renaissance
9	4,41	0,0000088	8	progressiste	4,67	0,00000251	42	grands
10	4,24	0,0000192	14	enfin	4,62	0,00000325	7	croient

Les principales différences par rapport aux listes obtenues précédemment résident dans la disparition ou le déplacement de certains co-occurents du substantif *humanistes* au pluriel : le déterminant *ces* ne figure plus parmi les dix co-occurents les plus spécifiques (il se situe, toutefois, en onzième position !), alors que l’adjectif *faux* disparaît complètement de la liste. L’adjectif *grands* se déplace vers une position inférieure, désormais précédée par *Renaissance*. L’allusion aux grandes figures de la Renaissance, que ce co-occurent jusqu’à absent laisse présager, modifie l’hypothèse initiale ou, en tout cas, la remodele. La polarisation précédemment présumée – entre, d’un côté, *grands (et vrais) humanistes* et, de l’autre, *grands mais faux humanistes* – ne disparaît pas nécessairement, mais semble dépendre de configurations discursives spécifiques.

4.2 Un retour au texte : les démonstratifs *cet* et *ces*

Le retour au texte constitue une démarche de nature qualitative. Il permet de confronter à l’épreuve des faits ce que les chiffres ne font que suggérer. L’une des manières d’appréhender les nuances véhiculées par les mots consiste à les étiqueter de façon systématique. Une telle approche permet également de gérer les ambiguïtés découlant de la flexibilité sémantique naturelle des mots. Les emplois annotés pourraient ensuite servir de référence (*gold standard*) pour la validation automatique de l’analyse qualitative. À l’ère de l’essor des modèles d’IA, une telle perspective apparaît comme prometteuse. Ci-dessous, nous amorçons une première étape de l’analyse qualitative : l’examen des valeurs véhiculées par les démonstratifs *cet* et *ces*. Ces démonstratifs, en présupposant un savoir partagé avec le lecteur, constituent un point d’observation intéressant pour analyser les valeurs associées au mot *humaniste*.

S’agissant des fonctions discursives assumées par les démonstratifs, la reprise anaphorique coréférentielle apparaît, généralement, comme la plus caractéristique. Elle consiste à reprendre le contenu d’un énoncé antérieur comme dans l’exemple ci-dessous :

(5) Avisé, Léon implante également de ce côté-ci de l’eau l’ostréiculture. Dans la foulée, cet humaniste fonde une société de protection des parqueurs. (lexpress.fr, 2012)

Parfois, toutefois, la reprise prend la forme d’une cataphore. Dans les emplois cataphoriques de *CE N*, une entité est annoncée avant d’être identifiée :

(6) Comment rester insensible au terrible sort de cet humaniste vénitien et futur cardinal Pietro Bembo (1470 – 1547). (lepoint.fr, 2020)

Dans le discours médiatique, la fonction expressive semble prédominante. En nous basant sur les données du corpus, nous avons délimité trois types de structures syntaxiques par lesquels s'opère l'actualisation de la valeur expressive du démonstratif, encore que la coexistence de différentes valeurs et de différentes formes puisse donner lieu à des structures mixtes. L'expressivité de la forme n'exclut donc pas qu'elle assume simultanément une fonction de reprise anaphorique, en se référant à une entité évoquée dans l'article commenté ou un autre commentaire. Comparons :

- *CE N Adj*, p. ex.:
 - (7) [L]e grand Hugo, cet humaniste, ce républicain forcené... (lemonde.fr, 2012)
- *CE N qui*, p. ex.:
 - (8) Qui sait si la Palestine n'est pas la raison de vivre de ces humanistes qui ainsi se sentent utiles, intéressants et importants ? (lefigaro.fr, 2012 et 2020)
 - (9) [C]es humanistes qui ont des châteaux grâce aux subventions des contribuables...ben non...ça suffit... (lefigaro.fr, 2017)
 - (10) Car, à n'en pas douter, ces humanistes qui à force de pratiquer le grand écart doivent en avoir des douleurs insupportables à l'entre-jambe... (lefigaro.fr, 2020)
- *CE N Prép N*, p. ex.:
 - (11) Lire une citation de Montaigne donne envie de se ressourcer chez ces humanistes de la Renaissance (Rabelais, La Boétie – fort ami avec Montaigne – Erasme, lequel ne manquait pas d'humour... (lemonde.fr, 2012)
 - (12) Mais ce qui est le marqueur de ces humanistes d'opérette semble vouloir nous faire croire que les islamistes ne sont pas plus dangereux que les catholiques. (lefigaro.fr, 2012 et 2020)
 - (13) La suffisance dont font preuve ces humanistes de paille n'en arrêtent pas l'importation. (lefigaro.fr, 2020)

Dans tous les exemples analysés sauf le premier, les pronoms démonstratifs assument des fonctions discursives marquées. Ils servent d'abord à construire une catégorie référentielle évaluée, en désignant « ces humanistes » comme un groupe social identifiable et présenté comme objet de jugement. L'emploi démonstratif contribue ainsi à mettre à distance les référents et à renforcer la dimension polémique ou ironique du propos. Par ailleurs, ces formes démonstratives fonctionnent comme un mécanisme d'inclusion du lecteur dans une présupposition partagée : elles suggèrent que l'énonciateur et le co-énonciateur savent déjà de qui il est question, et qu'ils partagent une même représentation stéréotypée de ce groupe. Ainsi, les démonstratifs contribuent à la mise en scène d'une relation polémique avec le référent, à la construction d'un effet d'évidence partagée avec le lecteur, et au renforcement de la valeur axiologique des énoncés.

Le prédéterminant *TOUT* peut renforcer la valeur expressive du démonstratif. Ce fait n'est pourtant observable que dans les syntagmes au pluriel, p. ex. :

- (14) [I]gnoble imposture de tous ces humanistes courtisans (lefigaro.fr, 2017)
- (15) [T]ous ces humanistes qui rejettent en bloc 20 % de la population (lefigaro.fr, 2020)
- (16) [T]ous ces humanistes de pacotille à gants blancs... (lefigaro.fr, 2020)

Les éléments postposés au *N* dans trois types de structures délimités ci-haut peuvent assumer une fonction évaluative, porteuse d'un jugement de valeur (p. ex. *humanistes de paille*), ou une fonction restrictive qui délimite la classe des objets dénotés en précisant lesquels sont visés (p. ex. *humanistes de gauche*, *humanistes qui ont des châteaux*). Cette distinction est opératoire notamment à travers l'application du test interrogatif *Lesquels ?* qui s'applique naturellement aux syntagmes restrictifs mais pas aux syntagmes évaluatifs :

Humanistes, lesquels ? – Ces humanistes qui ont des châteaux.

*Humanistes, lesquels ? – Ces humanistes de paille.

En l'absence d'un contexte élargi, la fonction précise du démonstratif demeure pourtant indéterminée. Elle peut découler tant du *co-texte* (verbal) que du *contexte* (discursif) d'un énoncé. Il semble, en effet, qu'une référence à la *Deixis am Phantasma* [25], selon laquelle le démonstratif mobiliserait une connaissance partagée implicite, relevant d'un arrière-plan discursif commun, soit possible pour interpréter une partie d'emplois.

D'après les données recueillies, il apparaît qu'il s'agit d'un stéréotype négatif que suggèrent de nombreux adjectifs, mais aussi de nombreuses locutions adjectivales, p. ex. *de pacotille, en peau de lapin, de salon, de bazar, d'opérette, de circonstance, à géométrie variable, de façade*. Certains qualificatifs prennent une position éloignée par rapport au pivot, p. ex. : *humaniste chef de gondole des droits de l'homme demi-saison* – si l'interprétation de la fonction syntaxique peut manquer d'univocité, la valeur attribuée à l'ensemble du syntagme ne prête pas à confusion. La polarité négative caractérise également les structures où sont employés les marqueurs de totalisation comme *de tout poil* et *de tout bord*, indicateurs d'un ensemble qui n'est pas homogène, alors qu'on s'attendrait plutôt à une communauté de valeurs partagées.

Mais le corpus contient également des exemples de propos élogieux, p. ex. :

(17) Ce qu'est initialement Albert Jacquard : Un savant penché sur les mystères de la génétique. L'humaniste est aussi, selon l'étymologie, un homme que le passé passionne afin de questionner le présent. (lemonde.fr, 2012)

(18) Michel-François Delannoy, maire de Tourcoing et numéro deux de la communauté urbaine. L'humaniste exigeant et conséquent. À 92 ans, il continuait d'écrire... (lepoint.fr, 2023)

5 Remarques conclusives

Rappelons qu'une première recommandation d'employer le substantif *humaniste* avec un épithète a apparu dans le *Dictionnaire de l'Académie* (éd. 5, 1798) et, depuis, elle revenait régulièrement, mais seulement jusqu'à la 7^e édition (1878), en spécifiant tout de même que l'évaluation ainsi exprimée concerne : *celui qui sait, qui enseigne les humanités*. Différents éditeurs reprenaient des exemples analogues pour illustrer cette règle, p. ex. : *un excellent humaniste* et *un médiocre humaniste* (DAF, 1798, 1835, 1878), *un bon humaniste* (Littré), *un savant humaniste* (Littré, DAF éd. 9), *un grand humaniste* (DAF éd. 9). Ces exemples témoignant du respect et de l'estime dont auraient joui les humanistes depuis l'époque de la Renaissance constituent à la fois un cadre culturel et une perception traditionnelle, déterminée par le passé et les succès des humanistes dont les noms reviennent parfois dans les dictionnaires, p. ex. : *Marsile Ficin, Érasme, Guillaume Budé, Robert Estienne* (cf. DAF, éd. 9).

Notre analyse, quelque préliminaire qu'elle soit, réduite par ailleurs à un échantillon de données web, nous oriente vers les hypothèses suivantes, à vérifier lors d'une annotation qualitative systématique d'un plus grand corpus :

Hypothèse 1 : Une valeur nouvelle, perceptible à travers les données du corpus web, tend à se stabiliser dans l'usage, à savoir : *humaniste* en tant qu'élément de l'ethos d'un positionnement politique.

Hypothèse 2 : Cette nouvelle valeur favorise l'émergence d'une plus grande variété de connotations, y compris négatives, mises en œuvre par *humaniste*, lorsque celui-ci est employé pour désigner un positionnement politique.

Hypothèse 3 : L'image évaluative liée à des contextes polémiques ou idéologiques, est pourtant *sectorielle*. Une telle image peut découler d'une sensibilité d'un site et/ou de sensibilités de ses lecteurs.

Une autre question qui se pose suite à ce travail reste en lien avec la surreprésentation des données provenant d'un site avec pour corollaire de nombreux emplois évaluatifs, souvent négatifs. Un plus faible nombre d'emplois évaluatifs issus d'autres sites résulte-t-il de lacunes dans les données, qui auraient existé sans avoir pu être intégrées au corpus ? Perplexes au vu des résultats obtenus, nous en venons, à la lumière de l'étude effectuée, à poser à nouveau, des questions déjà connues : Qu'est-ce qu'un corpus représentatif et comment le constituer ? [26] Les périodiques choisis ont-ils permis de constituer un échantillon fiable, conforme au principe de *corporalité* du corpus ? [21] Une analyse incluant davantage de titres aurait-elle produit des résultats sensiblement différents ? Dans quelle mesure le recours à SketchEngine comme source de données permet-il de garantir la représentativité du corpus ? Ces incertitudes ne conduisent pas seulement à circonscrire la portée des résultats, mais incitent à une réflexion plus générale sur les effets structurants des choix de corpus, qui orientent nécessairement les valeurs et les régularités mises au jour.

Références

1. B. Wojciechowska, A. Kaliska, Enseignant, élève, philosophe... ami du peuple. Termes *humaniste* et *humanista* selon les dictionnaires français et polonais (du XVIIIe s. au présent), *Studia Neophilologica*, 2026 [à paraître].
2. P. Richelet, Le Dictionnaire françois contenant les mots et les choses. (J.-H. Widerhold J.-H. Widerhold, Genève, 1680).
3. A. Furetière, Dictionnaire Universel (A. & R. Leers, La Haye, 1690).
4. Dictionnaire de l'Académie française (2e éd. 1718, 3e éd. 1740, 4e éd. 1762, 5e éd. 1798, 6e éd. 1835, 7e éd. 1878, 8e éd. 1935, et la plus récente édition électronique terminée en 2024). URL : <https://www.dictionnaire-academie.fr/>
5. P.-C.-V. Boiste, Le Dictionnaire universel de la langue française, avec le latin, et manuel d'orthographe et de néologie (Desray, Paris, 1803).
6. N. Landais, Dictionnaire général et grammatical des dictionnaires français (Au Bureau Central, Paris, 1834).
7. A. Hatzfeld, A. Darmsteter, avec le concours d'A. Thomas, Dictionnaire général de la langue française du commencement du 17e siècle jusqu'à nos jours (C. Delagrave, Paris, 1895).
8. Le Trésor de la langue français informatisé (TLFi), terminé en 1994. ATILF - CNRS & Université de Lorraine.
9. L. Guilbert, R. Lagane, G. Niobey, Grand Larousse de la langue française (Larousse, Paris, 1989).
10. Le Robert Dico en ligne, URL : <https://dictionnaire.lerobert.com>
11. E. Littré, Dictionnaire de la langue française (L. Hachette, Paris, 1873-1874). Electronic version created by François Gannaz, URL : <http://www.littre.org>.
12. Dictionnaire de français Larousse, URL : <https://www.larousse.fr/dictionnaires/francais-monolingue>
13. J. Rey-Debove, A. Rey, Le Nouveau Petit Robert (Dictionnaires Le Robert, Paris, 1995).

14. C. Kannas et al. (dir.), Dictionnaire de la langue française Larousse (Larousse, Paris, 1995).
15. J.-P. Mével, B. Gaillard, Dictionnaire Hachette 2019 : noms propres et noms communs interclassés (Hachette, Paris, 2019).
16. Le Dico, adapté du Wiktionnaire (Éd. Garnier, Paris, 2019).
17. E. Tognini-Bonelli, Corpus Linguistics at Work (Benjamin, Amsterdam, 2001).
18. W. Teubert, La linguistique de corpus : une alternative [version abrégée], Semen, 27 (2009), DOI : 10.4000/semen.8914.
19. A. Kilgarriff, G. Grefenstette, Introduction to the Special Issue on Web as Corpus, Computational Linguistics, 29 (3), (Association for Computational Linguistics, 2003).
20. P. Charaudeau, Dis-moi quel est ton corpus, je te dirai quelle est ta problématique, Corpus, 8 (2009), 37-66. DOI : <https://doi.org/10.4000/corpus.1674>
21. D. Mayaffre, Corpus et web-corpus. Réflexion sur la corporalité numérique, Cahiers de praxématique, 54-55 (2010), 233-248. DOI : <https://doi.org/10.4000/praxematique.1170>
22. Orientation des médias français, URL: <https://www.integrersciencespo.fr/orientations-politiques-de-la-presse-etrangere-et-francaise#quotidiens-francais>
23. M. Jakubíček, A. Kilgarriff, V. Kovář, P. Rychlý, V. Suchomel, The TenTen corpus family. In 7th International Corpus Linguistics Conference CL (July 2013), 125-127.
24. L. Vanni, Hyperbase Web. (Hyper)Bases, Corpus, Langage. Corpus, 25 (2024), DOI : 10.4000/corpus.8770
25. G. Kleiber, Démonstratifs et expressivité : les démonstratifs – « titres », C. Chauvin, M. Kauffer (éds.), La fonction expressive. Écart et expressivité. Vol. 3 (Presses universitaires de Franche-Comté, Besançon, 2013). DOI : 10.4000/books.pufc.42725.
26. J.-P. Honoré, N'oublie pas le cœur du débutant. Remarques empiriques sur la formation des corpus en AD, Mots. Les langages du politique, 69 (2002), DOI : 10.4000/mots.10633.