



HAL
open science

Etude des stéréotypes genrés dans le théâtre français du XVIe au XIXe siècle à travers des plongements lexicaux

Alexandra Benamar, Cyril Grouin, Meryl Bothua, Anne Vilnat

► To cite this version:

Alexandra Benamar, Cyril Grouin, Meryl Bothua, Anne Vilnat. Etude des stéréotypes genrés dans le théâtre français du XVIe au XIXe siècle à travers des plongements lexicaux. *Traitement Automatique des Langues Naturelles*, 2022, Avignon, France. pp.74-81. hal-03701478

HAL Id: hal-03701478

<https://hal.science/hal-03701478>

Submitted on 24 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Etude des stéréotypes genrés dans le théâtre français du XVI^e au XIX^e siècle à travers des plongements lexicaux

Alexandra Benamar^{1,2}, Cyril Grouin¹, Meryl Bothua², Anne Vilnat¹

(1) Université Paris-Saclay, CNRS, LISN, Orsay, France

(2) EDF Lab R&D, Palaiseau, France

{first name}.{last name}@lisn.upsaclay.fr, {first name}.{last name}@edf.fr

RÉSUMÉ

Les modèles de TAL les plus récents cherchent à capturer au mieux toutes les subtilités de la langue, ce qui implique de récupérer les stéréotypes qui y sont associés. Dans cet article, nous étudions les stéréotypes de genre qui existent dans des modèles Word2Vec. Nous avons constitué un jeu de données composé de pièces de théâtre françaises allant du XVI^e au XIX^e siècle. Nous avons choisi de travailler sur le genre théâtral car il tend à pousser à leur paroxysme certains traits de caractère représentatifs de hiérarchies sociales préexistantes. Nous présentons des expériences dans lesquelles nous parvenons à mettre en avant des stéréotypes de genre en relation avec les rôles et les émotions traditionnellement imputés aux femmes et aux hommes. De plus, nous mettons en avant une sémantique spécifique associée à des personnages féminins et masculins. Cette étude démontre l'intérêt de mettre en évidence des stéréotypes dans des corpus à l'aide de modèles contextuels « classiques ».

ABSTRACT

Studying gender stereotypes in French theater from XVIth to XIXth century through the use of lexical embeddings

The most recent NLP models seek to best capture all the subtleties of the language, which involves recovering its stereotypes. In this article, we study gender stereotypes that exist in Word2Vec models. We have compiled a dataset composed of French plays ranging from the XVIth to the XIXth century. The advantage of theater is that it tends to exaggerate existing social hierarchies. We present experiences in which we highlight gender stereotypes related to the roles and emotions attributed to women and men. In addition, we elaborated forward specific semantics associated with female and male characters. This study demonstrates the interest in using « classic » contextual models to highlight biases in corpora.

MOTS-CLÉS : Plongements lexicaux, Biais, Stéréotypes de genre, Théâtre.

KEYWORDS: Word embeddings, Bias, Gender stereotypes, Theater.

1 Introduction

En apprentissage automatique, les ensembles de données utilisés correspondent à une vision partielle de la réalité. Les données sont susceptibles de contenir des biais, tels que des biais cognitifs humains ou des biais liés à la collecte des données (c'est-à-dire, liés à l'équité de représentation). Les modèles

d'apprentissage automatique sont fortement susceptibles de créer, *a posteriori*, des discriminations entre différentes populations (par exemple, des modèles de classement de CVs tendent à accroître des biais femme/homme). La question de l'analyse de biais dans des systèmes de traitement automatique des langues (TAL) a émergé ces dernières années. Parmi ces travaux, on retrouve deux grandes approches : l'évaluation des biais dans des espaces de plongements lexicaux (Bolukbasi *et al.*, 2016; Caliskan *et al.*, 2017; Gonen & Goldberg, 2019) ainsi que des travaux sur des biais relatifs à des tâches comme la traduction automatique (Wisniewski *et al.*, 2021), l'analyse de sentiments (Thelwall, 2018), ou encore la résolution de coréférences (Alfaro, 2019). Bien que ces travaux constituent un enjeu crucial dans l'analyse et la critique des systèmes de TAL et dans la mise en évidence des comportements négatifs que peuvent avoir ces systèmes, aucun consensus sur le concept de biais n'existe en TAL, et les définitions sont parfois très éloignées de ce qui est présenté en sociologie. La définition du terme constitue donc un « parti pris » dès le début de l'étude : quels types de comportements sont préjudiciables, de quelle manière, à qui et pourquoi ? Par exemple, on distingue deux types de préjugés dans ces travaux (Blodgett *et al.*, 2020) :

1. Les stéréotypes qui propagent une généralisation négative sur des groupes sociaux particuliers (tels que stéréotypes de genre, stéréotypes ethniques, d'âge, etc.).
2. Les différences de performances de systèmes de TAL pour différents groupes sociaux.

Nous nous intéressons à la première catégorie de préjugés, et plus précisément aux stéréotypes de genre. Parce que le concept de « biais » est flou dans la littérature (Blodgett *et al.*, 2020), nous utiliserons ici le terme « stéréotype » défini en sociologie comme « *une image préconçue, une représentation simplifiée d'un individu ou d'un groupe humain. Il repose sur une croyance partagée relative aux attributs physiques, moraux et/ou comportementaux, censés caractériser ce ou ces individus.* ». Dans cet article, nous étudions les stéréotypes de genre, sans recherche de réutilisation de ces vecteurs pour une tâche spécifique, dans un corpus de théâtre français du XVI^e au XIX^e siècle, qui présente l'avantage d'amplifier les stéréotypes existant dans le monde réel à des fins de critique sociale ou de comédie. Au vu de la spécificité des genres du théâtre, nous étudions tout particulièrement les stéréotypes existant dans deux genres : comédie et tragédie. Concrètement, la question centrale de ce travail est de déterminer si les stéréotypes de genre attendus dans les pièces de théâtre se retrouvent dans les modèles contextuels. Pour cela, nous nous intéressons d'abord à la sémantique des termes associés à « femme » et « homme » et notamment aux émotions qui les caractérisent. Puis, nous nous intéressons à quelques personnages types du théâtre (e.g., *valet, servante, maître*, etc.) afin d'étudier les spécificités de genre et de rôles captées par ces modèles.

2 Les stéréotypes dans le théâtre

Le théâtre se répartit en deux genres majoritaires : la tragédie et la comédie. Tandis que le premier, souvent illustré par Racine ou Corneille, est placé au sommet de la hiérarchie (avec son exigence de vraisemblance et ses convenances), la comédie française, représentée par Molière, est moins codifiée et valorisée. Durant le siècle des Lumières, le théâtre subit une révolution : « *la comédie, sous ses apparences de gaieté, de légèreté et de fantaisie, dénonce les injustices sociales, renverse les hiérarchies et tend vers la critique ou la satire* » (Marcandier, 2011). Au contraire, la tragédie laisse place au drame bourgeois visant à « [...] consacrer leurs écrits à des situations réalistes, proches du public bourgeois, à des intrigues reposant sur des conflits familiaux ». Étant donné que les deux genres sont construits pour des publics différents, les stéréotypes présents dans ces oeuvres le sont également. Gruffat (2012) étudie la représentation des héros classiques dans la tragédie et

démontre que le héros tragique amoureux était le produit d’une époque (il parle ici d’un changement brutal des hommes, chez Racine comme chez Pradon, où le caractère guerrier laisse place à une faiblesse et à une transformation profonde des héros amoureux). Le théâtre constitue, comme toutes les oeuvres littéraires et artistiques, un miroir de la société de l’époque à laquelle il renvoie. Ces écrits nécessitent d’être replacés dans leurs contextes culturels et sociaux afin de comprendre leurs impacts et d’analyser leur contenu. Selon la metteuse en scène Myriam Marzouki, la lourdeur des stéréotypes de genre est plus marquée dans le répertoire classique que dans le répertoire contemporain. Selon elle, « *les personnages de femmes dans le théâtre classique sont épouses, filles, mères ou servantes, leur rapport au monde étant toujours médiatisé par leur lien avec un homme.* »¹. Dans cette étude, l’objectif est de déterminer dans quelle mesure des modèles de traitement automatique des langues (TAL) mettent en avant les stéréotypes femmes/hommes présents dans les pièces de théâtre.

3 Jeu de données

Le corpus de pièces de théâtre que nous avons constitué² se compose de 817 pièces de théâtre dont 554 comédies et 263 tragédies. Elles ont été extraites au format XML et nettoyées « à la main » : suppression des en-têtes, du contenu relatif au à l’oeuvre plutôt qu’au contenu des oeuvres (dates, liste des personnages, actes et scènes) et des noms dans les tours de paroles. Les documents ont ensuite été segmentés en phrases à l’aide d’expressions régulières. Le vocabulaire contient 79 183 mots pour une moyenne de 7900 mots/pièce. Le tableau 1 présente la répartition des pièces par mouvement littéraire.

	Baroque 1580-1660	Classicisme 1660-1690	Lumières 1690-1789	Romantisme 1789-1860	Symbolisme 1850-1900
#Documents	144	147	404	77	45
#Mots	33 874	29 691	46 986	21 760	17 658
Moyenne #caractères/document	38 539	36120	35 282	34 269	37 908
Moyenne #mots/document	9 448	8 857	8 670	8 476	9 265

TABLE 1 – Répartition des pièces de théâtre du corpus dans les mouvements littéraires

4 Modèles Utilisés

Word2Vec (Mikolov *et al.*, 2013) est un modèle vectoriel permettant de représenter des mots dans leur contexte d’utilisation (une phrase, par exemple). Son utilisation repose sur l’apprentissage de représentations vectorielles de mots dans des corpus suffisamment grands. Nous réutilisons ce type de modèle pour représenter directement les mots de notre corpus, dans le contexte des phrases. Afin de comparer notre modèle à des modèles existants, appris sur des données plus volumineuses et représentant des stéréotypes différents, nous utilisons également des modèles pré-entraînés sur Wikipédia et le Web (Fauconnier, 2015). Nous nous limitons à l’entraînement Skip-Gram de Word2Vec, plus

1. Rapport d’information n° 704 (2012-2013) de Mme Brigitte Gonthier-Maurin, fait au nom de la délégation aux droits des femmes du Sénat. « La place des femmes dans l’art et la culture : le temps est venu de passer aux actes », 2013.

2. <http://www.theatre-classique.fr/pages/programmes/PageEdition.php>, licence CREATIVE COMMONS BY-NC-ND.

	Modèles existants			Entraînement	
	Wiki	Web	Comédie+Tragédie	Comédie	Tragédie
femme	filles, maîtresse, prostituée, fiancée, servante	filles, mari, jeune, prostituée, adolescente	mari, belle-soeur, vindicative, accoucha, veuve	mari, veuve, feue, épousé, vertueuse	mari, veuve, ressource, belle-mère, comtesse
homme	diplomate, politicien, agriculteur	femme, humaine, galop, humanité, nationalisme, sabre	garçon, hébété, désintéressé, aventurier, dissipateur	rustre, coureur, moine, joueur, bravoure	méchamment, fou, franc, président, raisonnable

TABLE 2 – Cinq plus proches voisins des mots « femme » et « homme » en utilisant des modèles Word2Vec. La distance cosinus est utilisée pour récupérer les mots les plus similaires

propice à la représentation de la sémantique que l'apprentissage CBOW qui est utile pour des calculs de similarité syntaxique. Tous les modèles génèrent des plongements lexicaux de dimension 1 000.

5 Expériences

5.1 Stéréotypes de genre

Voisinage des termes « femme » et « homme » Dans cette section, nous nous intéressons aux stéréotypes de genre mis en avant par les modèles Word2Vec. Afin de déterminer quels stéréotypes sont spécifiques aux pièces de théâtre, on compare des modèles appris sur trois sources de données : deux sources généralistes (Web et Wikipedia) et les pièces de théâtre. Les plus proches voisins sont les termes ayant la similarité cosinus la plus grande avec le terme source. Le tableau 2 présente une vision synthétique des cinq termes les plus similaires à « femme » et « homme ». Tout d'abord, les modèles renvoient des résultats plutôt similaires lorsque l'on s'intéresse au terme « femme », alors qu'ils sont bien plus variés pour le terme « homme ». Une faible richesse lexicale signifie que l'utilisation du terme « femme » sur le Web, sur Wikipédia et dans les pièces de théâtre est très similaire, et donc très stéréotypée. On observe que les termes associés à « femme » sont principalement axés sur le rapport social qu'elles entretiennent par rapport aux hommes (e.g., *épousée, fille, maîtresse, veuve, fiancée*), sur leur âge et leur physique (e.g., *jeune, adolescente, beauté*) ou sur leur sexualité (*prostituée, maîtresse, vertueuse*). Les termes les plus similaires à « femme » renvoient principalement à leur rôle dans le contexte familial (e.g., *mère, tante, julie, épouse, mari*). Au contraire, on observe une variété lexicale plus importante entre les modèles pour le terme « homme ». Sur Wikipédia, on observe une prépondérance de termes qui renvoient aux métiers (e.g., *diplomate, politicien, avocat*). En revanche, deux champs lexicaux apparaissent sur le Web : le champ lexical relatif à la guerre (*galop, sabre*) et celui de « l'Homme » (*humanité, humaine*). Le deuxième est dû au passage en minuscule lors de l'apprentissage. Sur les pièces de théâtre, on parvient à obtenir des termes typiques du genre (en comparaison avec les autres modèles). Par exemple, dans la comédie, on observe différents personnages types (tel que *coureur, joueur, rustre, bravoure*). Les personnages étant traités différemment en tragédie, le vocabulaire identifié est plutôt lié à des personnages odieux au statut social important. On remarque ici que, contrairement au terme « femme », il n'y a pas de termes liés au champ lexical de la famille dans le voisinage du mot « homme ».

	Homme	Femme	Indéterminé
Émotion	Fierté, colère	Joie, soulagement, espoir, surprise, compassion, tristesse, honte, intérêt	Jalousie, dégoût, mépris, peur

TABLE 3 – Tendence majoritaire attribuée aux émotions (Raymondie & Steiner, 2020)

Expression des émotions Dans cette partie, on cherche à déterminer quelles émotions sont les plus utilisées dans le contexte de « femme » et « homme ». Nous reprenons les émotions présentées par Raymondie & Steiner (2020) qui étudie le « genre des émotions » dans le contexte du travail. Le tableau 3 présente la tendance d’attribution majoritaire observée dans l’article pour quatorze émotions. La Figure 1 présente la proximité calculée entre les émotions et les termes « homme » et « femme ». Les émotions sont associées aux femmes et aux hommes très différemment en fonction du corpus. On note un accord entre les trois modèles sur quatre émotions (surprise, peur, soulagement et honte), qui sont toutes attribuées aux femmes, avec une légère différence par rapport aux hommes, ce qui est corroboré avec l’étude d’origine. Dans les pièces de théâtre, trois émotions sont principalement associées aux hommes (fierté, colère et compassion). Bien que les deux premières soient attendues et soutiennent l’hypothèse que les stéréotypes sont renforcés dans le théâtre, la présence de la troisième est surprenante. Le terme « compassion » dans le corpus de théâtre est principalement utilisé dans trois contextes : un personnage parle d’un homme sans compassion³, un personnage demande à un homme de faire preuve de compassion⁴ ou un personnage (souvent féminin) éprouve de la compassion pour un homme⁵. Cette émotion n’est donc pas associée aux hommes parce qu’ils en font preuve mais plutôt parce qu’il s’agit d’une problématique récurrente dans ces œuvres. De façon générale, les stéréotypes de genre associés aux émotions dans le théâtre correspondent à ceux qui sont attendus. On note qu’avec le modèle appris sur des données Web, les émotions « compassion » et « mépris » sont associées aux hommes, bien que contradictoires. On suppose que ces émotions ne sont pas uniquement employées en tant que qualificatif mais aussi comme requête (i.e., dans des contextes où on implorerait la compassion). De plus, étant donné que les hommes étaient souvent associés à « l’Homme » dans ce modèle, il se peut que ces adjectifs sont utilisés pour parler de l’humanité.

5.2 Voisinage de « personnages types »

Nous nous intéressons à la représentation de termes traditionnellement attribués à des personnages types au théâtre. L’objectif de cette étude est de déterminer si les modèles vectoriels parviennent à capturer les stéréotypes associés à ces rôles. Étant donné que les personnages types sont des rôles prépondérants et très caricaturaux, et qu’ils sont associés à des traits de caractères forts, on s’attend à retrouver de fortes différences dans la sémantique qui leur est associée avec les modèles vectoriels. De plus, les personnages au théâtre contiennent, à l’image de la vie réelle, des stéréotypes de genre assez forts qui orientent leurs caractéristiques. C’est pourquoi, nous nous attachons à distinguer les personnages types féminins et masculins dans cette étude. Cela nous permet d’étudier deux types de préjugés : ceux associés à des classes sociales (représentées par les statuts des personnages) et ceux

3. *« Il est toujours insensible et froid, l’homme qui se refuse à tes feux : son coeur qui s’isole se durcit, il n’est plus disposé à la compassion ni à la pitié. »* - « Les Tombeaux de Vérone », Louis-Sébastien Mercier (1782)

4. *« Ami, qui que tu sois, si ton âme sensible à la compassion peut se rendre accessible, un jeune gentilhomme implore ton secours ; »* - « Clitandre », Corneille (1631)

5. *« Petit bonhomme, veuille le ciel que vous ne vous trompiez pas, et que ce soit mon semblable que j’embrasse dans une créature pourtant si méconnaissable ! Vous me pénétrez de compassion pour vous. »* - « L’île de la raison », Marivaux (1727)

Proximité entre des émotions et les termes « femme » et « homme »

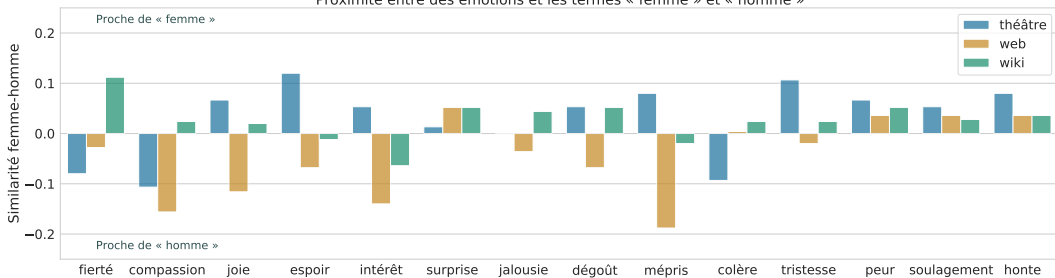


FIGURE 1 – Similarité cosinus entre des émotions et les termes « femme » et « homme ». La similarité est calculée entre le vecteur d’une émotion $v_{\text{émo}}$ et les vecteurs de « femme » v_f et « homme » v_h avec la formule : $\text{proximité} = \cos(v_{\text{émo}}, v_f) - \cos(v_{\text{émo}}, v_h)$. Une proximité de 1 signifie que l’émotion est exclusivement associée à « femme » et une proximité de -1, qu’elle est associée à « homme »

	Comédie	Tragédie
roi	prince, successeur, monarque, empereur, ambassadeurs	prince, nom, couronne, choix, droits
reine	souveraine, héritière, prisonnière, orgueilleuse, ingrate	faveur, honteux, princesse, hymen, jaloux
valet	coquin, vilain, fou, fripon, drôle	fripon, faquin, impertinent, domestique, arlequin
servante	demoiselle, andouille, soubrette, impertinente, gueuse	commère, minette, chonchon, chienne, hymen
maître	brave, scélérat, instruit, successeur, imposteur	valet, laquet, riche, province, méchant
enfant	bâtard, riche, jeune, mari, femme	veuve, gendre, comte, nièce, orphelin

TABLE 4 – Cinq plus proches voisins (distance cosinus) de personnages types (modèles pré-appris et entraînés sur le corpus de théâtre) pour les genres comédie et tragédie

associés au genre des personnages. Étant donné que les personnages ne sont pas traités de la même façon en comédie et en tragédie (on s’attend à des stéréotypes plus accrus en comédie), on analyse le vocabulaire associé aux personnages pour chaque genre.

Zaragoza (2006) évoque plusieurs personnages types du théâtre. Pour cette étude, nous avons sélectionné des antagonistes pouvant être analysés par paire : roi et reine, valet et servante (Gunny, 1978), et maître (par opposition au valet) (Maija, 2012). Nous avons ajouté le rôle de l’enfant afin de comparer le vocabulaire qui lui est attribué par rapport aux personnages adultes. Les résultats de l’étude sont présentés dans le tableau 4.

Dans la comédie, le valet est un personnage clé, parfois qualifié de « maître du jeu » (Da, 2009). La sémantique retrouvée dans les modèles est très spécifique de ce qui est attendu du personnage (*coquin, vilain, fou, fripon, drôle*). Dans la tragédie, Naugrette (2003) illustre le rôle de valet avec Ruy Blas : « *il contrevient à la loi, se révolte devant le sort qui lui est fait, et en meurt.* » Le modèle Word2Vec renvoie à la première définition du valet, similaire à la comédie (*fripon, faquin, impertinent*) allant jusqu’à relier le terme « arlequin » au rôle du valet. De même, la sémantique associée à la servante renvoie à son rôle emblématique de personne intrigante et commère dans la comédie alors qu’elle est plus sexualisée dans la tragédie, tout comme la reine (le terme *hymen* apparaît pour ces deux rôles). Nous observons que le vocabulaire du « roi » et de la « reine » demeurent spécifiques à leur genre. Bien que le roi apparaisse comme un dirigeant de haut rang, la reine est réduite aux aspects négatifs de sa personnalité (e.g., *orgueilleuse, ingrate*). Enfin, le maître est associé au statut de bourgeois riche et instruit avec des stéréotypes typiques de cette classe sociale (e.g., *méchant, imposteur*). La

sémantique liée à l'enfant est très proche de celle de la femme, avec des termes du champ lexical de la famille utilisés dans le même contexte que les femmes⁶.

6 Conclusion

Dans cet article, nous avons analysé les stéréotypes de genre dans un corpus de pièces de théâtre en utilisant des modèles de plongements lexicaux. Nous avons comparé des modèles Word2Vec appris sur des corpus externes (Wikipédia ou Web) et des modèles entraînés sur nos données. Nous avons mis en évidence des stéréotypes de genre dans les pièces de théâtre, ce qui montre que des modèles de plongements lexicaux classiques sont suffisants pour révéler des stéréotypes. Nous observons ainsi des différences systématiques entre la sémantique utilisée pour qualifier les femmes et les hommes. Alors que les hommes sont perçus différemment sur Wikipédia, le Web et dans les pièces de théâtre, les femmes sont qualifiées de façon très similaire dans ces corpus et souvent réduites au champ lexical de la famille et à leur sexualité. Nous relevons que les enfants, dans les pièces de théâtre, sont souvent mentionnés lorsqu'il s'agit de femmes et sont d'ailleurs très proches dans l'espace vectoriel. Au contraire, les hommes possèdent un champ lexical différent et plus varié (en fonction du genre de pièces). Nous relevons également que les stéréotypes liés aux émotions « genrées » peuvent être retrouvés facilement dans ce corpus. Enfin, nous avons observé que certains rôles de personnages-types au théâtre contiennent une sémantique attendue (les rôles de valet et servante sont très représentatifs de leurs attributs dans les comédies). Dans de futurs travaux, nous souhaitons appliquer cette étude à un corpus de courriels clients. L'objectif serait de mettre en évidence des stéréotypes de genre fréquents dans des conversations clientèles.

Références

- ALFARO F. (2019). Study and experimentation of gender bias in co-reference resolution. Mémoire de master, Universitat Politècnica de Catalunya.
- BLODGETT S. L., BAROCAS S., DAUMÉ III H. & WALLACH H. (2020). Language (technology) is power : A critical survey of " bias" in nlp. *arXiv preprint arXiv :2005.14050*.
- BOLUKBASI T., CHANG K.-W., ZOU J. Y., SALIGRAMA V. & KALAI A. T. (2016). Man is to computer programmer as woman is to homemaker? debiasing word embeddings. *Advances in neural information processing systems*, **29**.
- CALISKAN A., BRYSON J. J. & NARAYANAN A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, **356**(6334), 183–186.
- DA F. (2009). En quoi le valet est-il maître du jeu au théâtre? *Publications Pimido*.
- FAUCONNIER J.-P. (2015). French word embeddings.
- GONEN H. & GOLDBERG Y. (2019). Lipstick on a pig : Debiasing methods cover up systematic gender biases in word embeddings but do not remove them. *arXiv preprint arXiv :1903.03862*.
- GRUFFAT S. (2012). La représentation du héros amoureux dans les tragédies classiques : pour une conception évolutive du moi? *Litteratures classiques*, (1), 143–160.

6. « Vous seul nous arrachant à de nouvelles flammes nous avez fait laisser nos enfants et nos femmes. » - « Iphigénie », Jean Racine (1794).

GUNNY A. (1978). Emelina, jean : Les valets et les servantes dans le théâtre comique en france de 1610 à 1700. *Zeitschrift für französische Sprache und Literatur*, **88**, 265–266.

MAIJA N. (2012). La relation maître-valet dans le théâtre de comédie. *Publications Pimido*.

MARCANDIER C. (2011). Le théâtre.

MIKOLOV T., CHEN K., CORRADO G. & DEAN J. (2013). Efficient estimation of word representations in vector space. In Y. BENGIO & Y. LECUN, Éd., *1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings*.

NAUGRETTE F. (2003). Le devenir des emplois comiques et tragiques dans le théâtre de hugo. *Littératures classiques*, **48**(1), 215–225.

RAYMONDIE R. A. & STEINER D. D. (2020). Stéréotypes de genre concernant l'expression des émotions : pensez subordonné–pensez femme ? *Carrières, leadership et conflits*, p. 203.

THELWALL M. (2018). Gender bias in sentiment analysis. *Online Information Review*.

WISNIEWSKI G., ZHOU L., BALLIER N. & YVON F. (2021). Biais de genre dans un système de traduction automatique neuronale : une étude préliminaire (gender bias in neural translation : a preliminary study). In *Actes de la 28e Conférence sur le Traitement Automatique des Langues Naturelles. Volume 1 : conférence principale*, p. 11–25.

ZARAGOZA G. (2006). *Le personnage de théâtre*. Armand Colin.