Convolution et marqueurs multidimensionnels. Description des représentations genrées dans un corpus de films français

Laurent Vanni, Magali Guaresi, Véronique Magri

Convolution et marqueurs multidimensionnels. Description des représentations genrées dans un corpus de films français

Laurent Vanni, Magali Guaresi, Véronique Magri

Université Côte d'Azur, CNRS, BCL

laurent.vanni@univ-cotedazur.fr

magali.guaresi@univ-cotedazur.fr

veronique.magri@univ-cotedazur.fr

Abstract

Convolutional neural networks allow new representations of texts that extend the standard statistical approaches. By combining frequency and context of words as well as allowing multidimensional treatments (graphical form, lemma and part of speech), convolution leads to the extraction of *motifs*, i.e. complex linguistic patterns that are likely to feed interpretation. In this paper, this architecture is tested on movie scripts in order to explore the hypothesis of a gendered differentiation of female and male dialogues.

Keywords: Gender, Movies, Convolution, *Motifs,*Textual Statistic Analysis, Deep learning.

Résumé

Les réseaux convolutionnels appliqués aux textes permettent de nouvelles représentations du texte, prolongeant les résultats de l'analyse de données textuelles. En combinant approche fréquentielle et séquentielle des textes et en permettant un traitement multidimensionnel des mots (forme graphique, lemme et code grammatical), la convolution aboutit à l'extraction de *motifs*, c'est à dire de patrons linguistiques complexes susceptibles d'alimenter l'interprétation. Dans ce papier, ce type de modèle est mis à l'épreuve d'un corpus de scripts cinématographiques dans le but d'explorer l'hypothèse d'une différenciation genrée des dialogues de femmes et d'hommes.

Mots clés : Genre, Cinéma, Convolution, Motifs, Statistique textuelle, Deep learning.

1. Introduction

Les réseaux convolutionnels appliqués aux textes (Collobert et Weston 2008) permettent de dépasser l'approche fréquentielle de la statistique classique, en croisant l'axe paradigmatique (la sélection des unités linguistiques) et l'axe syntagmatique (la combinaison de ces unités dans la linéarité du passage)[1]. Par ailleurs, comme en analyse de données textuelles (ADT), la convolution est capable de considérer plusieurs représentations des mots : forme graphique, catégorie grammaticale et lemme. Ces représentations multidimensionnelles (ou multi-*channels*) permettent, non seulement, d'obtenir de meilleures performances en termes de classification (prédiction), mais aussi d'extraire des motifs au sens de (Mellet et Longrée 2009), susceptibles d'alimenter la description et l'interprétation des textes (Mayaffre et Vanni 2021). À partir d'un corpus de scripts de films français contemporains, cette étude propose une description des dialogues en fonction du sexe des locuteurs/locutrices et du sexe des cinéastes. Nous montrerons que l'approche convolutionnelle multicritère prolonge l'ADT classique en mettant en évidence des patterns linguistiques profonds - combinant les activations des trois *channels* - qui interroge d'une part le deep learning et sa capacité à extraire de l'information linguistique complexe et d'autre part les représentations de genre - compris comme un système

[1] Cette recherche a été financée par l'Agence Nationale de la Recherche (ANR) au titre du projet ANR-21-CE38-0012-01 et du Programme d'Investissement d'Avenir UCA JEDI ANR-15-IDEX-0001.

de bicatégorisation des sexes et des représentations qui y sont associées (Achin et Bargel, 2013) au cinéma.

2. Méthode et corpus

Les réseaux convolutionnels, initialement conçus pour l'analyse d'images, utilisent souvent plusieurs *channels,* c'est-à-dire plusieurs tableaux de données différents, pour représenter une même image en entrée de plusieurs façons. La méthode la plus répandue consiste à utiliser le codage RGB (Red, Blue, Green) pour représenter l'image suivant trois nuances de couleurs. L'idée de cet algorithme est de permettre au modèle d'extraire des contrastes particuliers qui utilisent des couleurs différenciées. Pour les textes, il est possible d'appliquer le même raisonnement pour encoder séparément les mots avec leur forme graphique, leur catégorie grammaticale et leur lemme. Cette approche présente deux avantages majeurs : d'une part, elle autorise le modèle à mixer syntaxe et lexique pour distinguer deux classes différentes dont elle améliore potentiellement la précision, d'autre part, elle permet de restituer des motifs linguistiques profonds basés sur le calcul du *Text Deconvolution Saliency* (TDS) (Vanni et al. 2018), implémenté dans Hyperbase[2]. La capture de ces motifs est une méthode exploratoire fondée sur l'analyse des zones d'activations les plus fortes (pic d'activation du TDS) dans le texte. Ces passages que le deep learning souligne pour prendre une décision (classification) ont vocation à interroger l'analyste sur les données, leur nature et les contrastes qui discriminent chaque texte (Vanni 2021). C'est précisément cette potentielle exploration fine qui nous pousse à utiliser cette méthode dans le cadre d'une étude des représentations genrées dans les films. Celles-ci seront décrites par l'analyse des spécificités (section 3.1), prolongée par l'exploration des motifs profonds qui ouvre des parcours de lecture nouveaux et expose des résultats moins attendus (section 3.2).

Le corpus que nous étudions ici est composé de 86 films français, issus de la base de scripts en ligne *Lecteurs anonymes*[3]. Des travaux ont montré qu'il était possible d'extraire de manière automatique les répliques d'un corpus de scripts et de les attribuer à un sexe (Apoorv et al., 2015). Ici, nous avons utilisé une approche similaire en nous fondant sur un dictionnaire de noms propres et un contrôle manuel. Pour nos expérimentations, le corpus est divisé en quatre sous-corpus regroupant les répliques selon le sexe des personnages et le sexe des réalisateur.rices.

	Occurrences	Vocabulaire	Hapax
Personnages féminins par Réalisatrices	82899	6421	3508
Personnages féminins par Réalisateurs	214503	12821	6953
Personnages masculins par Réalisatrices	49736	5523	3264
Personnages masculins par Réalisateurs	305536	16483	8595

Tableau 1 : Description du corpus « Scripts »

La partition du corpus est de taille inégale, les films de réalisatrices et les dialogues de personnages féminins étant sous-représentés. Mais loin d'être un biais de corpus, cette répartition traduit la réalité de la production cinématographique française. D'après une étude

[2] http://hyperbase.unice.fr
[3]http://lecteursanonymes.org/scenario/

publiée par le Centre national de la cinématographie en 2019, seuls 27% des fims sortis en 2017 ont été réalisés par des femmes et les personnages féminins n'y ont qu'un tiers du temps de parole.

Le modèle entrainé pour cette étude utilise 80% du corpus pour l'entrainement et 20% pour la validation. Les classes choisies correspondent aux quatre parties du corpus décrites dans le tableau 1 et la précision du modèle sur ces classes atteint 85% sur l'ensemble de validation.

3. Etude de cas : des spécificités aux motifs genrés dans un corpus de scripts français

3.1 Spécificités

L'analyse de données textuelles rend compte, par le calcul des spécificités (Lafon 1980), des mots sur-utilisés[4] dans les répliques des personnages féminins comparées à celles des personnages masculins. Qu'ils soient produits par des réalisatrices ou des réalisateurs, les dialogues construisent des protagonistes stéréotypiquement genrés qui n'emploient pas les mêmes mots selon leur sexe. D'un point de vue thématique d'abord : les femmes évoquent, dans les films de réalisatrices, des sujets réputés féminins comme celui de la reproduction (« pilule » [14,52], « enceinte » [10,79], « grossesse » [8,95], etc.)[5], là où les hommes, en particulier dans les films de réalisateurs, abordent de larges thèmes publics, « politiques » [4,6] ou économiques (« entreprise » [3,86], « société » [3,73], « justice » [3,57]), reproduisant ainsi la traditionnelle dichotomie féminin-privé *versus* masculin-public. La différenciation genrée s'opère également pour les personnages masculins des réalisatrices présentés comme compétitifs (« concours » [6,17]) voire sportifs (« ballon » [4,7]), ou encore bricoleurs (« démonter » [4,3]). Les mots caractéristiques des femmes dans les films de réalisateurs montrent, en revanche, des personnages élaborés dans le cadre des relations familiales (« maman » [3,35]) ou affectives (« chéri » [4,35]). Les spécificités lexicales confirment l'existence de différences sensibles entre les discours des femmes et des hommes au cinéma, leur construisant ainsi des places et des rôles différents (Brey 2020).

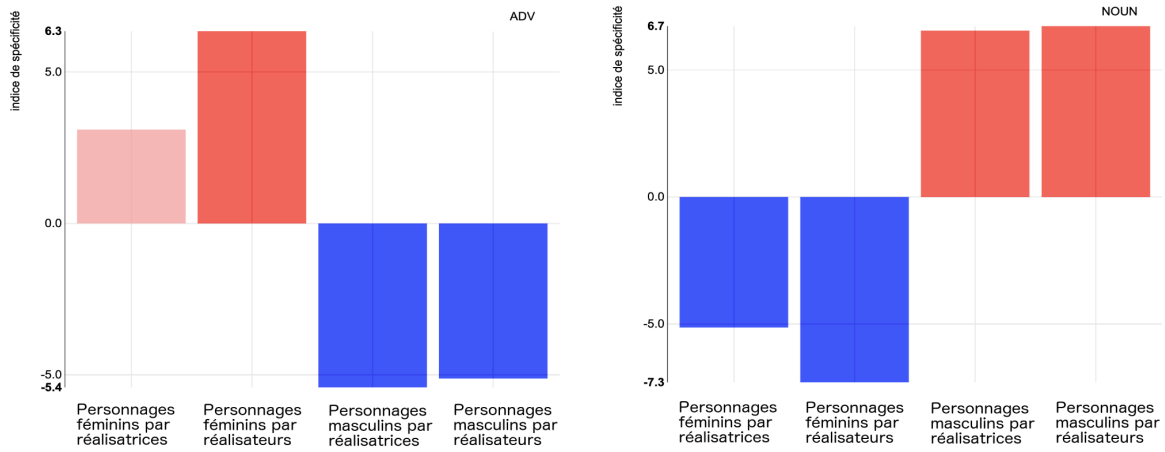

Figure 2 : Sur-utilisation des adverbes et sous-utilisation des noms dans les répliques de personnages féminins

[4] Les principaux mots caractéristiques sont repris dans la suite du développement avec leur indice de spécificité entre crochet.

[5] Ces spécificités lexicales témoignent de la sur-représentation dans le corpus des réalisatrices de comédies dramatiques ou de drames consacrés aux questions du contrôle des naissances. Sur le sujet, lire aussi Lécossais 2017.

Pour dépasser la simple analyse thématique, la distribution statistique des catégories grammaticales suggère des différences syntaxiques entre les répliques prononcées par les femmes et par les hommes (figure 2). Les premières sur-utilisent les adverbes et les pronoms alors que les hommes emploient davantage de noms, laissant penser à deux économies discursives distinctes (discours verbal et modalisé des femmes *versus* discours nominal et substantiel des hommes).

Si des méthodes statistiques poussées nous autoriseraient une analyse plus fine, utilisant par exemple les segments répétées (Salem, 1986) ou les (poly)cooccurrences spécifiques (Vanni, 2016), l'extraction des motifs profonds par le *deep learning* offre un parcours interprétatif complémentaire qui croise à la fois le lexique et la syntaxe dans une même analyse.

3.2 De l'urne au réseau : les motifs genrés profonds

Nous l'avons vu, le deep learning offre une lecture nouvelle des données. À première vue, les zones d'activation (TDS) confirment les analyses préliminaires faites avec l'ADT classique.

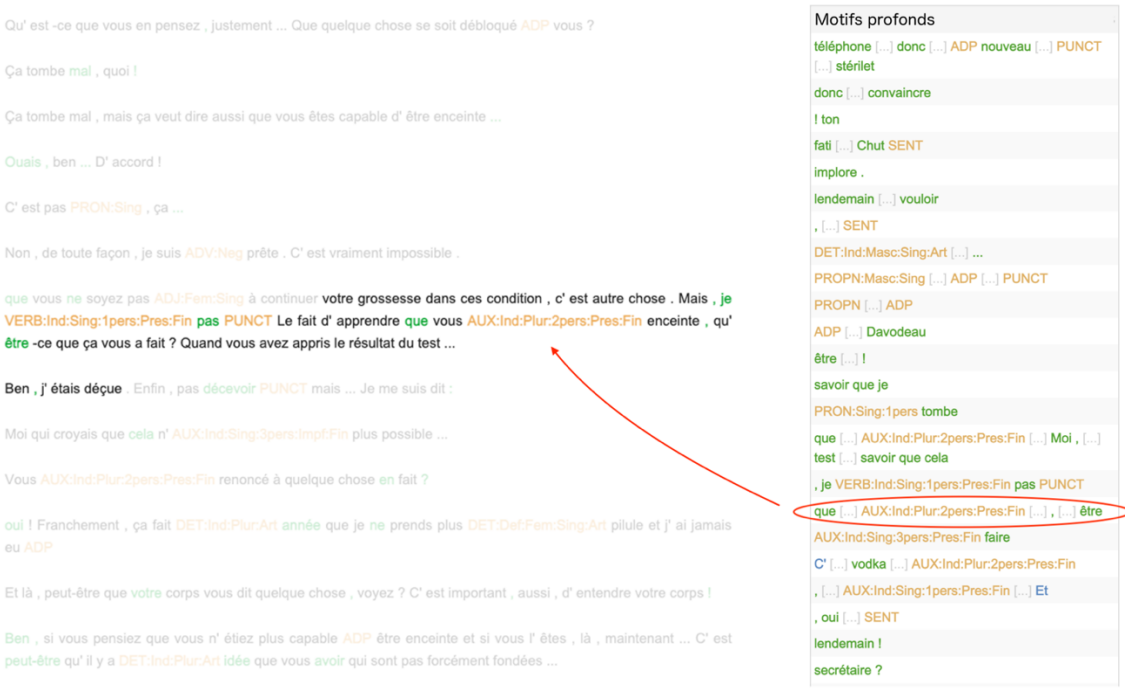

*Figure 3 : Sur-utilisation et activation du motif « que *** AUX *** LEM:être »*[6] *dans les répliques de femmes chez les réalisatrices*

Le lexique occupe une place importante dans les marqueurs qui permettent de discriminer chacune des classes. Ainsi, nous retrouvons, pour la classe « Personnages féminins par réalisatrices » (Figure 3), le mot « stérilet » qui marque la plus forte activation du réseau. D'autres unités comme l'adverbe « donc » ou le point d'exclamation, présents dans les passages-clés, sont aussi des marqueurs attestés par le calcul de spécificités. Mais la plus-value de la méthode ne se situe pas dans ces premiers marqueurs détectés. Un peu plus loin dans la liste des motifs (20 premiers résultats), un premier schéma linguistique semble proposer un véritable *motif* (Figure 3). Composé du subordonnant « que » suivi immédiatement ou non d'un auxiliaire[7] conjugué à la deuxième personne de pluriel, et plus loin, du verbe « être », ce passage peut, en effet, être considéré comme un *motif profond* (Vanni, 2021) en tant que suite de marqueurs non contigus et multidimensionnels qui marque une singularité dans les dialogues

[6] Les « *** » signifient, dans le langage des expressions régulières, n'importe quelle(s) unité(s) linguistique(s).
[7] Notons que l'étiquette AUX proposée par Hyperbase correspond aux verbes « être » ou « avoir » qu'ils soient employés en tant qu'auxiliaire ou non.

de femmes chez les réalisatrices et que la statistique classique peine à détecter automatiquement.

« Que » est un terme polysémique qui recouvre en langue plusieurs natures grammaticales. Dans le corpus, sous la multiplicité des réalisations grammaticales, des traits communs émergent. Dans l'exemple figurant sur la figure 3, le motif renvoie à une interrogation adressée par une conseillère du Planning familial à une jeune femme. Celle-ci lui demande : « Le fait d'apprendre que vous étiez enceinte, qu'est-ce que ça vous a fait ?[8] ». Outre le lexique repéré par la statistique comme « féminin », la structure « que + auxiliaire + *** + être », repéré par l'algorithme de deep learning comme caractéristique des prises de parole féminines dans les films de réalisatrices, montre une tendance à la mise au premier plan de la perception affective ou évaluative des faits.

Dans la majorité des motifs tirés du corpus, ces processus de retours sur les faits se concrétisent par un retour sur un « dire », au travers de formes de la parole rapportée ou de la reformulation :

> « Donc, vous me disiez que vous avez vu le médecin hier, en urgence… C'était tard non ?[9] »
>
> « ça veut dire que vous êtes capable d'être mère vous aussi[10] »

La complétive introduite par « que » est placée sous la rection du verbe de parole « dit », soit pour reprendre les propos d'un tiers dans une forme de discours indirect (premiers exemple), soit pour reformuler et utiliser la formule « vouloir dire » qui introduit une glose explicative (deuxième exemple). Les tournures emphatiques sont encore plus remarquables quand elles thématisent le prédicat, c'est-à-dire l'évaluation du fait, en le plaçant en début d'énoncé. L'emphase repose dans l'extrait suivant sur la redondance du sujet exprimé sous la forme développée de la complétive.

> « ça ne m'étonne pas que vous soyez dehors, vous aimez bien ce temps-là. C'est magnifique votre site[11] »

Les répliques des femmes dans ce contexte paraissent manifester une dissociation marquée entre le *dictum* (ce qui est dit) et le *modus* (l'attitude exprimée par l'auteur.e de l'énoncé) en faveur de la mise en évidence du second placé volontiers au premier plan dans la phrase, en position frontale. L'énoncé manifeste un doublage modalisateur persistant.

4. Conclusion

Cette étude montre, sur le plan méthodologique, que les réseaux convolutionnels offrent aujourd'hui de nouvelles représentations des textes. En combinant l'approche fréquentielle et séquentielle et en considérant ensemble les activations des trois *channels* (lexique, lemme, code grammatical), le modèle permet de mettre au jour des motifs linguistiques profonds susceptibles de prolonger les résultats en statistique textuelle sur les représentations discursives de genre au cinéma. Ces motifs présentent une double plus-value : d'une part, ils sont des unités contextuelles *de facto* porteuses de sens et d'autre part, ils combinent différents niveaux de granularité de la textualité pour être de nouveaux observables complexes du genre.

Ces premières analyses, à mettre à l'épreuve d'autres données filmiques, suggèrent la prégnance des stéréotypes genrés dans l'élaboration des personnages féminins et masculins.

[8] *Les Bureaux de Dieu*, 2008.
[9] *Les Bureaux de Dieu*, 2008.
[10] *Les Bureaux de Dieu*, 2008.
[11] *Roxane*, 2019. Dans cet exemple, le verbe « être » est au subjonctif présent. L'embedding (word2vec) utilisé par le modèle fait apparaître cette catégorie dans les 10 plus proches voisins (cosine distance) de l'auxiliaire à l'indicatif. Nous faisons donc l'hypothèse, ici, que cette variation fait partie des motifs soulignés par le modèle. Cette hypothèse sera validée dans des travaux ultérieurs.

Outre les spécialisations lexicales, qui cantonnent les femmes au privé et à la reproduction, les outils de l'analyse de données textuelles et du deep learning montrent des économies discursives sexuées différentes : aux hommes, le discours nominal, référentiel ou informatif ; aux femmes le discours adverbialisé et modalisé. La répartition de la parole entre les sexes dans les films est régie par des rapports de pouvoir, proches de ceux observés dans les conversations dans nos sociétés (Monnet 1998). Les répliques féminines sont moins fréquentes, plus courtes et souvent sous forme de questions ou de demande de précisions sur ce qui vient d'être dit. Comme l'exemplifie le motif étudié dans la dernière section, les répliques féminines se démarquent par une mise au premier plan de l'énonciation avant le contenu même de l'énoncé. Ce marquage plus fort de la subjectivité des énonciatrices dans les scénarios traduit moins une emprise directe et active sur le monde qu'un retour, évaluateur ou affectif, par rapport aux évènements qui sont évoqués et rapportés dans les dialogues.

Références

Achin C. et Bargel L. (2013). *Dictionnaire Genre et Science politique*, Paris, Presses de science politique.

Apoorv et al. (2015). Keys Female Characters in Film Have More To Talk About Beside Men : Automating Betchdel Test, *Human Language Technologies: The 2015 Annual Conference of the North American Chapter of the ACL*, 830–840, Denver, Colorado.

Brey I. (2020). *Le Regard féminin. Une révolution à l'écran*, Paris, Edition de l'Olivier.

Collobert, R. et J. Weston (2008). A Unified Architecture for Natural Language Processing : Deep Neural Networks with Multitask Learning. *Proceedings of the 25th International Conference on Machine Learning*. ICML '08. Helsinki, Finland : ACM, 160–167.

Lafon P. (1980). Sur la variabilité de la fréquence des formes dans un corpus. *Mots,* 1: 127–165.

Lécossais S. (2017). Les mots de la grossesse. *Études de communication*, 48 : 155-176.

Mayaffre D. et Vanni L. (2021). *L'intelligence artificielle des textes. Des algorithmes à l'interprétation*, Champion, Paris, 2021.

Mellet S. et Longrée D. (2009). Syntactical motifs and textual structures. *Belgian Journal of Linguistics*. T. 23 : 161–173.

Monnet C. (1998), La répartition des tâches entre les femmes et les hommes dans le travail de conversation », *Nouvelles Questions Féministes*, vol. 19 (n°1) : 9-34.

Salem A. (1986). Segments répétés et analyse statistique des données textuelles. *Histoire & Mesure*, vol 1 (n°2) : 5-28.

Vanni L. (2021). *De l'analyse statistique de données textuelles aux réseaux de neurones artificiels. Vers des motifs linguistiques profonds*. Thèse de doctorat, Université Côte d'Azur.

Vanni L., Ducoffe M., Mayaffre D., Precioso F., Longrée D., et al. (2018). Text Deconvolution Saliency (TDS) : a deep tool box for linguistic analysis. *In : 56th Annual Meeting of the Association for Computational Linguistics (ACL)*, p. 548–557, Melbourne, Australia.

Vanni L. et Mittmann A. (2016). Cooccurrences spécifiques et représentations graphiques, le nouveau "Thème" d'Hyperbase. *12es Journées internationales d'Analyse statistique des Données Textuelles*. T. 1 : 295–305.