



HAL
open science

Littérature et algorithmes : lorsque la critique théâtrale londonienne (2010 – 2020) passe au crible de l'informatique

Mylène Maignant

► To cite this version:

Mylène Maignant. Littérature et algorithmes : lorsque la critique théâtrale londonienne (2010 – 2020) passe au crible de l'informatique. Linguistique. Université Paris sciences et lettres, 2022. Français. NNT : 2022UPSLE065 . tel-04682444

HAL Id: tel-04682444

<https://theses.hal.science/tel-04682444v1>

Submitted on 30 Aug 2024

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.

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL
Préparée à l'École normale supérieure

**Littérature et algorithmes :
lorsque la critique théâtrale londonienne (2010 – 2020)
passe au crible de l'informatique**

Soutenue par

Mylène MIGNANT

Le 13 décembre 2022

Ecole doctorale n° 540

École doctorale Lettres, Arts,
Sciences humaines et sociales

Spécialité

Sciences du langage

Composition du jury :

Ioana, GALLERON Professeure des universités Université Sorbonne Nouvelle – Paris 3	<i>Présidente</i>
Elisabeth, ANGEL-PEREZ Professeure des universités Sorbonne Université	<i>Rapporteure</i>
Clarisse, BARDIOT Professeure des universités Université Rennes 2	<i>Rapporteure</i>
Derek, MILLER Professor Harvard University	<i>Examineur</i>
Thierry, POIBEAU Directeur de recherche au CNRS Paris Sciences et Lettres (Ecole normale supérieure)	<i>Directeur de thèse</i>

UNIVERSITE PARIS SCIENCES ET LETTRES
ECOLE NORMALE SUPERIEURE

THESE DE DOCTORAT

Littérature et algorithmes :
Lorsque la critique théâtrale londonienne (2010 – 2020)
passé au crible de l’informatique

Auteur
Mylène MAIGNANT

Directeur de thèse
Thierry POIBEAU

Laboratoire LATTICE
Ecole doctorale 540 – Transdisciplinaire Lettres / Sciences

Soutenue le 13 décembre 2022

Jury de thèse

Thierry POIBEAU	Ecole normale supérieure	Directeur de thèse
Ioana GALLERON	Université Sorbonne Nouvelle	Présidente
Elisabeth ANGEL-PEREZ	Sorbonne Université	Rapporteure
Clarisse BARDIOT	Université Rennes 2	Rapporteure
Derek MILLER	Harvard University	Examineur

RÉSUMÉ

Fondée sur l'étude d'un corpus de plus de 40 000 critiques théâtrales publiées entre 2010 à 2020, cette thèse a pour finalité d'examiner la réception du théâtre contemporain londonien à travers le regard de deux communautés critiques différentes : les journalistes d'un côté, et les blogueurs de l'autre. A partir de 2010 en Angleterre, une multitude de blogs émerge en ligne. Professionnels du théâtre, étudiants mais aussi simples amateurs publient leurs analyses théâtrales en ligne. La naissance de cette blogosphère remet en question les fondements de la culture journalistique. Le canon de la critique théâtrale londonienne se trouve progressivement décentralisé par ses périphéries. A partir de méthodes informatiques, l'objectif de ce projet de recherche consiste à modéliser les points de convergence et de divergence entre ces deux groupes de spectateurs distincts afin de comprendre leur perception du théâtre contemporain londonien.

MOTS CLÉS

Théâtre, humanités numériques, réception, science des données, critique littéraire, journalisme, blogueurs

ABSTRACT

Based on a corpus of more than 40,000 theatre reviews published between 2010 and 2020, this thesis aims to examine the reception of contemporary London theatre through the eyes of two distinct groups of theatre critics: the journalistic one on the one hand, and the digital one on the other hand. Since 2010 in England, a wave of blogs written by authors coming from various horizons has arisen on the Internet. Students, theatre professionals but also mere amateurs began publishing their own theatre reviews. The birth of this blogosphere challenges the foundations of the journalistic culture. The canon of London theatre criticism is gradually decentralized by its peripheries. By using computational methods, the aim of this research project consists in modeling the differences and similarities between these two groups of spectators in order to understand their perception of London contemporary theatre.

KEYWORDS

Theatre, Digital Humanities, Reception Studies, Data Science, Literary Criticism, Journalism, Blogging

Remerciements

Ayant effectué la quasi-totalité de mes études en lettres, je n'aurais jamais pensé mener un tel projet qui repose autant sur la science des données. Et pourtant... ! Ces trois années de doctorat ont été une très belle opportunité pour plonger dans un monde qui m'était encore inconnu il y a quelques années. Naviguer entre l'informatique et les théories littéraires n'allait pas toujours de soi. Ce voyage quelque peu étonnant m'a cependant permis d'explorer la littérature d'un point de vue ludique. Pour faire fonctionner deux disciplines qui sont rarement associées l'une à l'autre, il faut créer, faire preuve de curiosité, essayer, rater, recommencer. Et je suis heureuse que mes recherches aient pris ce tournant.

*

Je n'aurais pas pu mener à bien ce projet sans l'aide de **Thierry Poibeau**, mon directeur de thèse. Converser dans deux langages différents, l'un technique, l'autre littéraire, n'a pas toujours été facile au début, mais ces échanges m'ont permis de découvrir une autre manière de penser et d'appréhender la connaissance, et je te suis très reconnaissante de cela. Merci aussi pour ta patience et tes relectures attentives. Elles m'ont aidée à grandir intellectuellement.

J'aimerais également remercier **Elisabeth Angel-Perez** qui a été la première professeure à m'enseigner le théâtre contemporain anglais. Vos idées et votre manière d'enseigner ont toujours été une grande source d'inspiration pour moi, et je n'aurais pas effectué cette thèse sans avoir assisté à vos cours. Votre enthousiasme sans faille et votre soutien m'ont été très précieux durant ces années de doctorat. Merci.

Drama Critiques n'aurait jamais pu voir le jour sans les petites mains de l'ombre qui m'ont aidée à mettre en place ce projet. Un immense merci à **Damien**, pour sa bonne humeur, ses encouragements et son entrain. Damien a travaillé à mes côtés afin de m'aider à relire ces quelques 40 000 critiques. Nous avons échangé plusieurs mails tous les jours, pendant six longs mois, afin de créer ce corpus. Grâce à toi, ma solitude dans cette tâche gigantesque n'était plus qu'un vague souvenir. De collègues nous sommes passés à amis, et c'est un lien précieux que j'ai envie de garder pendant longtemps encore.

Gaëtan a également largement contribué à cette recherche. Gaëtan a effectué un stage sous ma direction au laboratoire Lattice pendant six mois. Son expertise en informatique m'a, non seulement permis de progresser rapidement dans ce domaine, mais aussi de donner corps à mes hypothèses littéraires. Merci d'avoir transformé cette trentaine de rendez-vous de travail par Skype en discussions passionnantes et enrichissantes, tant d'un point de vue professionnel que personnel.

Je ne suis pas devenue compétente en informatique du jour au lendemain, et pour progresser plus vite, j'ai pris des cours particuliers. J'ai eu la chance de rencontrer **Théo**. Cela fait presque un an maintenant que ses explications si claires, sa patience et son aide m'accompagnent dans ce projet. Je n'ai pas encore trouvé le moyen de t'exprimer toute ma gratitude. Ton pragmatisme, ta force tranquille et ton humour ont réellement rendu cette fin de doctorat plus douce. Ton intérêt pour mon projet de recherche m'a également confortée dans mes idées. Ce sont donc des remerciements très sincères que je t'adresse, non plus comme une « apprenante », mais comme une amie.

Et puis, évidemment, il y a eu toutes les marques d'affection de celles et ceux qui m'entourent qui m'ont permis de rire lorsque le quotidien devenait compliqué. L'amour, la tendresse et la bienveillance infinies de **Nicolas** m'ont portée pendant de longs mois. Je ne te remercierai jamais assez pour tout ce que tu as fait pour moi et pour tout ce que tu m'as apporté. C'est la plume fine et délicate d'**Anaïs** qui a créé les dessins de Drama Critiques. Cela fait quinze ans que nous sommes amies, et je suis fière que ce soient tes productions qui illustrent mon travail. **Fatima** m'a accueillie pendant des mois chez elle et a toujours fait en sorte que mes conditions de travail soient idéales pour rédiger, notamment grâce à sa cuisine perse. **Charlotte** m'a rassurée en écoutant tous mes doutes et mes questions à la campagne, **Colin** et **Lucence** ont fait un séjour de recherche avec moi sur fond de musique indienne et grégorienne et **Léa** m'a envoyé de belles lettres de soutien. Enfin, **Mathilde** a eu la gentillesse de répondre à mes innombrables appels, **Auphémie** m'a accompagnée dans la rédaction et **Nathan** m'a permis de m'évader en m'emmenant à la mer.

Je tiens également à remercier tous les blogueurs qui ont accepté que je publie leurs critiques et qui ont pris le temps de répondre à mes questions, et plus particulièrement **Laura Kressly**, éditrice de ThePlaysTheThingUK.com, que j'ai eu la chance de rencontrer lors d'un colloque à Reykjavik. **Alison Cook**, éditrice de *Theatre Record*, a également été d'une grande aide dans la diffusion de notre corpus en nous accordant le droit de republier toutes leurs critiques.

*

Cette recherche est dédiée à **Julie** et à **Nicolas**, ainsi qu'à ma mère, **Isabelle**, à mon père, **Philippe**, et à ma sœur, **Manon**. Je ne suis pas sûre qu'un paragraphe soit nécessaire pour exprimer ce que je ressens. Vous le savez déjà, et il me reste toute ma vie pour vous le témoigner.

SOMMAIRE

Introduction

Définition de la question de recherche	10
Finalité de la recherche	12
Apports de la thèse	13
Définition des termes-clés	15
Structure de la thèse	17

I. Etat de l'art

I.A Théâtre, réception et numérique : tour d'horizon	28
--	----

I.A.1 Humanités numériques et études théâtrales

I.A.1.1 Une discipline encore sous-représentée	28
I.A.1.2 Première catégorie de projets : histoire du théâtre	31
I.A.1.3 Deuxième catégorie de projets : étude des textes	34
I.A.1.4 Troisième catégorie de projets : logiciels de visualisation	37

I.A.2 Réception et théâtre anglais

I.A.2.1 Définition des études de la réception	41
I.A.2.2 Hans-Robert Jauss et l'horizon d'attente	42
I.A.2.3 Les études de la réception dans le théâtre	44

I.B Redéfinition de la critique théâtrale au XXI ^e siècle	49
--	----

I.B.1 Critique et numérique : la blogosphère londonienne

I.B.1.1 L'émergence d'un second regard sur le théâtre (1997 – 2010).....	49
I.B.1.2 Naissance d'une véritable communauté numérique (2010 – 2020)	53
I.B.1.3 Centre versus périphéries de la critique théâtrale londonienne.....	56

I.B.2 Positionnement de notre projet

I.B.2.1 Le mariage des études théâtrales et de l'informatique	59
I.B.2.2 Définition de la méthode	60
I.B.2.3 Présentation et limites du corpus	62

II. Constitution et présentation des données

II.A Création du corpus	73
-------------------------------	----

II.A.1 La critique journalistique : *Theatre Record*

II.A.1.1 Quels types d'informations sont mis à disposition des lecteurs ?	73
---	----

II.A.1.2 Quels journaux avons-nous sélectionnés ?	73
---	----

II.A.1.3 Comment avons-nous récupéré le contenu des journaux ?	73
--	----

II.A.2 La critique numérique : les blogs

II.A.2.1 Qu'est-ce qu'un blog ?	82
---------------------------------------	----

II.A.2.2 Quels types d'informations sont mis à disposition des lecteurs ?	83
---	----

II.A.2.3 Comment avons-nous récupéré leur contenu ?	83
---	----

II.B Présentation du contenu du corpus	89
--	----

II.B.1 Structurer les informations

II.B.1.1 Mise en place d'une matrice de données	89
---	----

II.B.1.2 Ajout des métadonnées	91
--------------------------------------	----

II.B.1.3 Application d'outils en Traitement Automatique des Langues (TAL) ...	93
---	----

II.B.2 Présentation générale

II.B.2.1 Deux sous-corpus équilibrés	96
--	----

II.B.2.2 La critique journalistique en quelques chiffres	99
--	----

II.B.2.3 La critique numérique en quelques chiffres	102
---	-----

III. Exploration du corpus

III.A Stylistique computationnelle et « zoning » : affect versus raison	125
---	-----

III.A.1 Les journalistes et les blogueurs écrivent-ils de la même manière ?

III.A.1.1 La critique théâtrale et la notion de style	125
---	-----

III.A.1.2 Etat de l'art de la technique du « zoning »	128
---	-----

III.A.1.3 Définir automatiquement les zones littéraires d'une critique	129
--	-----

III.A.2 Deux façons d'appréhender l'expérience théâtrale

III.A.2.1 Résultats	137
---------------------------	-----

III.A.2.2 Interprétation	140
--------------------------------	-----

III.A.2.3 Limites de l'approche	144
---------------------------------------	-----

III.B Cartographie numérique : centre versus périphéries spatiales	148
III.B.1 Existe-t-il une rupture géographique entre le centre de la critique théâtrale et ses périphéries ?	
III.B.1.1 Analyser le théâtre d'un point de vue spatial	148
III.B.1.2 Que sont les Systèmes d'Information Géographique et comment ont-ils été utilisés dans les études théâtrales ?	149
III.B.1.3 Création d'une carte des lieux fréquentés par la critique journalistique et numérique	153
III.B.2 Deux communautés aux intérêts divergents	
III.B.2.1 Résultats	158
III.B.2.2 Interprétation	164
III.B.2.3 Limites de l'approche	168
III.C Analyse d'opinion : étude des succès sur la scène londonienne	172
III.C.1 Les préférences esthétiques de ces deux groupes de spectateurs sont-elles similaires ?	
III.C.1.1 Quels types de pièces les journalistes et les blogueurs préfèrent-ils ?..	172
III.C.1.2 Etat de l'art de l'analyse d'opinion	173
III.C.1.3 Calcul des pourcentages de positivité et de négativité des critiques ...	174
III.C.2 Des choix esthétiques similaires	
III.C.2.1 Résultats	179
III.C.2.2 Interprétation	181
III.C.2.3 Limites de l'approche	182
IV. Science ouverte : partage et mise en valeurs des données	
IV.A Créer un pont entre les humanités et le numérique	189
IV.A.1 La plateforme DramaCritiques.com	
IV.A.1.1 Identité numérique de la plateforme	189
IV.A.1.2 Construction de la page d'accueil	191
IV.A.2 Un contenu clair, pédagogique et interactif	
IV.A.2.1 Présentation du contenu du site web	195
IV.A.2.2 L'application « Derrière la machine »	200

IV.B Un potentiel d'exploitation pour les communautés littéraires et scientifiques	206
IV.B.1 Ouverture et accès aux données : les Principes FAIR	
IV.B.1.1 Principes 1 et 2 : « Findable » et « Accessible »	206
IV.B.1.2 Principes 3 et 4 : « Interoperable » et « Reusable »	216
IV.B.2 Un exemple de piste de recherche : la question des femmes	
IV.B.2.1 Un moteur de recherche dédié aux dramaturges féminines.....	219
IV.B.2.2 Etude de cas : la place des femmes dans l'industrie théâtrale londonienne (1995 – 2010)	220
Conclusion	225
Annexes	232
Bibliographie	271

Introduction

Vladimir: Moron!
Estragon: Vermin!
Vladimir: Abortion!
Estragon: Morpion!
Vladimir: Sewer-rat!
Estragon: Curate!
Vladimir: Crétin!
Estragon (with finality): Critic!

Waiting for Godot (1954), p. 48

Dans la littérature théâtrale ou romanesque, le critique est souvent représenté comme un parasite. Austère, suffisant et frustré, il est cette figure de l'ombre qui juge et moralise le travail des autres à la nuit tombée. En atteste cette jonglerie verbale entre Vladimir et Estragon, dont l'insulte suprême qui clôturait cet enchaînement d'injures est celle de « Critic ». L'étymologie de ce mot rappelle pourtant que cette profession ne consiste pas à détruire le travail des artistes, mais à l'analyser. « Critiquer » signifie « discerner, disjoindre, trier ». Le critique a ainsi pour dessein de sélectionner certaines caractéristiques d'un événement culturel donné afin de le décortiquer, de le « passer au tamis » comme le rappelle l'étymologie grecque du mot, et d'en donner une analyse. Pour le critique théâtral, il lui faut passer de l'expérience visuelle de la scène à l'écriture. Entre plaisir esthétique immédiat et distance rationnelle, il doit traduire ce qu'il a vu, ressenti et compris le temps d'une représentation en un nombre de mots limités.

L'arrivée d'Internet au début du XX^e siècle remet cependant en question le statut et l'activité de critique. A partir des années 2010 dans le monde anglophone, l'émergence d'une vague de blogueurs redessine les contours de la critique théâtrale journalistique. Etudiants, professionnels du théâtre mais aussi simples amateurs mettent en ligne leur propre site web afin de publier leurs points de vue. Entre journalistes professionnels, rémunérés pour leur activité, et blogueurs amateurs, bénévoles passionnés, les polémiques prolifèrent. Une série d'articles aux titres provocateurs jette la lumière sur l'apparition de cette nouvelle communauté numérique qui vient bousculer le rôle du critique classique : « Is it Curtains for

Critics?¹ », « Blogging and Theatre Criticism: The Debate Continues² », « Are Critics and Bloggers on the Same Side?³ », etc.

Tandis qu'une partie des journalistes dénigre la légitimité des blogueurs, ces derniers, à l'inverse, réclament leur droit de visibilité. Derrière ces crispations se dessine surtout l'opposition entre le canon de la critique théâtrale anglaise et ses périphéries, ou entre une communauté qui fait autorité et une autre qui peine à trouver sa place dans l'espace théâtral public. En 2007, Nicholas Hytner, alors directeur artistique du Royal National Theatre à Londres, qualifia les critiques journalistiques de « dead, white men⁴. » Cette remarque provocatrice résume la position d'un grand nombre de blogueurs. Face à l'élitisme des médias classiques, une nouvelle forme de critique est nécessaire pour assurer une représentation plus juste et plus diverse de la culture théâtrale londonienne.

L'identité de ces plateformes numériques se construit en opposition à la culture journalistique, jugée par la plupart des blogueurs comme discriminante. Le titre de certains blogs affiche explicitement leur désir de différenciation face à ces grandes voix privilégiées de la presse écrite : ViewfromtheCheapSeat.com⁵, AyoungerTheatre.com⁶, NotExactlyBillington.com⁷ ou encore CriticsofColour.tumblr.com⁸ sont autant de dénominations qui soulignent leur appartenance aux marges de la culture théâtrale. Ces nouvelles voix discrètes de l'espace numérique entendent mettre à l'honneur d'autres modes de récits et d'autres types de théâtres qui ne sont pas, ou pas assez, représentés par les journalistes.

¹Jay RAYNER, « Is it Curtains for Critics? », *The Guardian* [En ligne], URL : <https://www.theguardian.com/artanddesign/2008/jul/13/art.comedy>, consulté le 04 mai 2022.

² Karen FRICKE, « Blogging and Theatre Criticism: The Debate Continues », *The Guardian* [En ligne], publié le 28 novembre 2008, URL : <https://www.theguardian.com/stage/theatreblog/2008/nov/28/blogging-critics>, consulté le 04 mai 2022.

³ Lyn GARDNER, « Are Critics and Bloggers on the Same Side? », *The Guardian* [En ligne], publié le 22 août 2013, URL : <https://www.theguardian.com/stage/theatreblog/2013/aug/22/theatre-edinburghfestival>, consulté le 04 mai 2022.

⁴ Ben HOYLE, « Dead White Men in the Critic's Chair Scoring Work of Women Director », *The Guardian* [En ligne], publié le 14 mai 200, URL : <https://www.thetimes.co.uk/article/dead-white-men-in-the-critics-chair-scoring-work-of-women-directors-6jn7025295h>, consulté le 04 mai 2022.

⁵<https://viewfromthecheapseat.com/> [En ligne], consulté le 23 mai 2021.

⁶<https://www.ayoungertheatre.com/> [En ligne], consulté le 04 juillet 2021.

⁷<http://www.notexactlybillington.com/> [En ligne], consulté le 02 mai 2021.

⁸<https://criticsofcolour.tumblr.com/> [En ligne], consulté le 11 mai 2021.

Finalité de la recherche

Cette recherche aura pour dessein d'examiner la relation entre le canon de la critique théâtrale anglaise et ses périphéries. Aucun corpus n'étant directement exploitable pour l'analyse d'une telle problématique, nous avons construit notre propre base de données. Cette dernière se compose de 40 043 critiques théâtrales anglaises qui portent sur la scène londonienne uniquement et qui couvrent les années 2010 à 2020. Nous avons choisi de commencer l'analyse de ces deux formes de réception à partir de 2010 car c'est précisément à ce moment-là que la critique numérique se déploie. Toutes les critiques journalistiques de notre corpus proviennent de la base de données *Theatre Record*⁹, qui fut mise en ligne en janvier 2019. Concernant les critiques numériques, nous avons extrait automatiquement leur contenu grâce à des scripts informatiques.

A partir de moyens techniques, il s'agira de modéliser la relation entre ces deux communautés afin de mieux comprendre leur perception du théâtre contemporain londonien. Toutes deux représentent une forme de réception différente. Toutes deux fonctionnent selon leurs propres codes et leur propre norme. Ces différences affectent nécessairement la manière dont le théâtre est perçu et dont la critique est rédigée. Les questions directrices de cette étude seront donc les suivantes : « Est-il possible d'observer une rupture entre les journalistes et les blogueurs, ou entre le centre de la critique théâtrale londonienne et ses périphéries ? Quels outils informatiques mobiliser pour modéliser cette problématique et comment interpréter les résultats d'un point de vue littéraire ? »

Nos observations littéraires seront fondées sur nos analyses techniques. Ces deux domaines seront dépendants l'un de l'autre. Sans l'informatique, il nous serait impossible d'examiner notre corpus dans son ensemble. Sans l'histoire théâtrale, l'interprétation de nos résultats chiffrés serait limitée. Afin d'étudier ce phénomène dans son ensemble, nous nous sommes focalisés sur trois questions littéraires précises qui correspondent à trois domaines techniques :

1. « Peut-on observer des différences stylistiques d'un corpus à l'autre ? » Posée autrement, cette question revient à étudier la manière dont les critiques journalistiques et numériques écrivent. Pour cela, nous aurons recours à des méthodes en stylistique computationnelle et en Machine Learning.

⁹<https://www.theatrecord.com/>[En ligne], consulté le 11 mai 2021.

2. « Leur répartition spatiale dans Londres confirme-t-elle le clivage « centre / périphéries » ? Afin d'analyser les types de lieux dans lesquels se rendent les journalistes et les blogueurs, nous utiliserons un Système d'Information Géographique pour visualiser leurs déplacements au sein de la capitale.

3. « Les préférences esthétiques de ces critiques varient-elles d'une communauté à l'autre ? » Nous utiliserons l'analyse d'opinion, une méthode algorithmique qui consiste à déterminer automatiquement les opinions d'un auteur dans un texte donné, afin de comprendre quels types de pièces les journalistes et les blogueurs préfèrent.

La finalité de ce projet est donc double : d'une part, analyser la réception du théâtre contemporain londonien à travers le regard de deux publics divergents, et d'autre part, mobiliser diverses techniques informatiques pour apporter des réponses à des questions littéraires. C'est l'articulation de ces deux champs opposés qui nous permettra de construire notre projet.

Apports de la thèse

1. En littérature : études de la réception appliquées au théâtre

Les études de la réception sont encore en construction. Par manque de données appropriées et directement exploitables par la machine, peu d'études sont disponibles sur cette thématique. La recherche que nous proposons permet ainsi de participer à l'élaboration de ce champ en se focalisant sur un exemple précis, celui de la scène londonienne des années 2010 à 2020. De plus, si de nombreux essais ont déjà abordé la question de la critique journalistique, très peu d'entre eux se sont attachés à examiner la critique numérique. Surtout, aucune étude ne propose une mise en regard de ces deux communautés avec un corpus de cette envergure. Cette recherche est donc la première à s'intéresser à la réception du théâtre londonien des années 2010 à 2020 avec autant de données.

2. Dans le domaine des humanités numériques théâtrales

Bien que les humanités numériques soient en pleine expansion depuis plusieurs décennies, les études théâtrales restent néanmoins sous-développées par rapport à d'autres disciplines¹⁰. Ce projet contribue non seulement à la construction de ce champ hybride, mais également à l'enrichissement de méthodologies nouvelles. L'utilisation de la science des données pour l'étude de textes théâtraux ne va pas de soi. Elle requiert une forme de gymnastique linguistique afin de converser dans deux langues différentes, l'une littéraire, l'autre technique. Nous avons détaillé les méthodes que nous avons adoptées et avons mis en libre accès nos scripts de sorte que d'autres chercheurs en humanités numériques puissent soit reproduire nos expériences, soit les adapter à leur corpus.

3. Science ouverte : diffusion et partage des données

La base de données qui a constitué le matériau premier sur lequel nous avons fondé nos analyses est également une des contributions majeures de ce doctorat, tant en littérature qu'en informatique. Il représente le premier corpus en études théâtrales contemporaines à fournir autant de critiques directement exploitables par l'ordinateur. Plus de 36 000 critiques sont en libre accès et sous licence Creative Commons Attribution 4.0 International. Ces données sont aujourd'hui téléchargeables sur différentes plateformes en ligne¹¹. L'intégralité du corpus est autrement soit accessible sur demande, soit mise à disposition des utilisateurs sous forme de moteur de recherche sur notre site DramaCritiques.com¹². La propreté, la richesse des métadonnées ainsi que l'homogénéité des données (50,5% de critiques journalistiques et 49,5% de critiques numériques) offrent de multiples perspectives de recherche originales aux chercheurs, tant dans le domaine de la littérature que de l'informatique.

¹⁰ Clarisse BARDIOT, « Arts de la scène et Culture Analytics », *Revue d'Historiographie du Théâtre* [En ligne], no. 4, 2017. URL : <https://sht.asso.fr/etudes-theatrales-et-humanites-numeriques-une-introduction/>, consulté le 24 janvier 2020.

¹¹ Le corpus est téléchargeable sur :

- **Le dépôt Zenodo :**

<https://zenodo.org/record/6799656#.Y0mXFEzP2so> [En ligne], consulté le 01 juillet 2022.

- **La page dédiée aux téléchargements sur la plateforme DramaCritiques.com :**

<https://dramacritiques.com/fr/telechargements/> [En ligne], consulté le 01 juillet 2022.

¹²<https://dramacritiques.com/fr/base-de-donnees/> [En ligne], consulté le 01 juillet 2022.

4. Dramacritiques.com : création d'un outil numérique à destination de la communauté littéraire et informatique

Enfin, cette recherche a également donné naissance à Dramacritiques.com, un nouvel outil numérique en ligne¹³. Cette plateforme a deux finalités principales. La première consiste à offrir davantage de visibilité aux blogueurs. En mettant en avant leur site et leurs réseaux sociaux, en leur donnant la parole à travers plusieurs entretiens et en mettant à l'honneur leurs critiques au même titre que celles rédigées par les journalistes, nous avons voulu promouvoir leur travail. La seconde finalité de DramaCritiques.com consiste à ouvrir les humanités numériques au plus grand nombre en expliquant les processus algorithmiques qui sous-tendent notre recherche. L'application « Derrière la machine » a été créée dans cette perspective. Elle permet de comprendre les différents procédés techniques utilisés pour rendre un texte lisible par l'ordinateur. Nous avons voulu créer un pont explicite entre les humanités et le numérique : montrer en quoi consiste l'informatique, comment l'utiliser sur des ressources littéraires et, surtout, les limites que présentent l'outil informatique.

Définition des termes-clés

Humanités numériques

En 2010, le *Manifeste des digital humanities* rédigé et cosigné par plus de 250 chercheurs et 10 institutions à Paris définissait les humanités numériques comme :

L'ensemble des Sciences humaines et sociales, des Arts et des Lettres. Les digital humanities ne font pas table rase du passé. Elles s'appuient, au contraire, sur l'ensemble des paradigmes, savoir-faire et connaissances propres à ces disciplines, tout en mobilisant les outils et les perspectives singulières du champ du numérique. Les digital humanities désignent une transdiscipline, porteuse des méthodes, des dispositifs et des perspectives heuristiques liés au numérique dans le domaine des sciences humaines et sociales.¹⁴

Notre étude s'insère dans cette transdiscipline qui s'inscrit à l'intersection de deux champs opposés : les études théâtrales et la science des données. Plus précisément, nous nous appuyons sur les écrits théoriques de Franco Moretti qui accordent une place centrale à la

¹³<https://dramacritiques.com/>[En ligne], consulté le 01 juillet 2022.

¹⁴Marin DACOS, « Manifeste des Digital Humanities », *ThatCamp Paris 2010* [En ligne], publié le 26 mars 2011, URL : <http://tcp.hypotheses.org/318>, consulté le 12 janvier 2022.

« lecture distancée » (*distant reading*) pour aborder la littérature sous un angle différent. Très proche de ce que Matthew Jockers nomme les « micro ou macro analyses¹⁵ », en référence à la micro et la macro économie, la lecture distancée repose sur les statistiques, les graphes et la visualisation des données pour explorer des corpus de grande taille qui ne peuvent être abordés par une lecture cursive.

Etudes de la réception

Ce projet de recherche ne porte pas sur les textes du théâtre contemporain londonien, mais sur la manière dont ils sont reçus et analysés par deux groupes de spectateurs différents des années 2010 à 2020. Notre objet d'étude porte par conséquent sur les destinataires du théâtre et sur leur perception singulière de la scène. Cette approche inscrit ainsi notre démarche dans les études de la réception. Nous utiliserons le concept d'« horizon d'attente », tel qu'il est défini par Hans-Robert Jauss dans *Pour une esthétique de la réception*, pour examiner nos textes d'un point de vue théorique.

En proposant une vision historique de la littérature, Jauss prône l'importance de l'histoire des effets d'une œuvre sur le public. C'est en prenant en compte l'interprétation du lectorat durant une période donnée qu'une œuvre prendra tout son sens. Pour comprendre le sens d'une œuvre, il faut reconstituer l'« horizon d'attente » du public, c'est-à-dire ses « attentes concrètes [...] ses intérêts, désirs, besoins et expériences tels qu'ils sont déterminés par la société et la classe à laquelle il appartient aussi bien que par son histoire individuelle¹⁶. » Nous verrons que les deux communautés critiques de notre corpus ont précisément des horizons d'attente divergents, ce qui laisse supposer que leurs analyses du théâtre seront différentes.

Canon / marges, centre / périphéries

Le troisième aspect central à notre recherche porte sur le couple conceptuel « canon / marges » ou « centre / périphéries. » Le terme « canon » vient du grec *kanôn* qui fait

¹⁵ Matthew JOCKERS, *Macroanalysis: Digital Methods & Literary History*, Champaign, University of Illinois Press, 2013.

¹⁶ Hans-Robert JAUSS, *Pour une esthétique de la réception*, Paris, Gallimard, 1978, p. 259.

premièrement référence à un roseau, puis à un instrument de mesure fabriqué à partir du roseau. De manière plus générale, le canon renvoie à la règle de conduite, à la norme, au modèle. C'est autour de son centre que se construisent les périphéries. La définition donnée par Griselda Pollock permet de comprendre en quoi cette notion est adossée à l'idée de domination :

Le canon renvoie aux textes — ou aux objets — que les institutions académiques établissent comme les meilleurs, les plus représentatifs et les plus significatifs dans les domaines de la littérature, de l'histoire de l'art ou de la musique. Dépositaires d'une valeur esthétique transhistorique, les canons des diverses pratiques culturelles établissent non seulement ce qui est incontestablement grand, mais aussi ce qui doit être étudié comme modèle par ceux qui aspirent à l'une de ces pratiques¹⁷.

Pierre Bourdieu l'a démontré dans *La distinction : critique social du jugement* (1979), le centre détermine les fondements d'une culture légitime¹⁸. En construisant leur identité numérique et leurs discours en opposition aux journalistes, les blogueurs représentent ainsi les marges de la critique théâtrale londonienne. Grâce à l'utilisation du numérique, nous verrons quelles stratégies les blogueurs mettent en place afin de contester le discours dominant journalistique.

Structure de la thèse

Chapitre I

Le premier chapitre sera consacré à la présentation des fondements théoriques de cette recherche. Nous avons divisé ce chapitre en deux parties. Dans la première, nous établirons l'état de l'art des différentes disciplines dans lesquelles s'inscrit notre travail. Nous aborderons les humanités numériques théâtrales ainsi que les études de la réception appliquées au théâtre. Nous verrons que notre sujet s'articule à la frontière de ces deux champs encore en construction. La seconde partie portera sur le cœur de notre analyse, c'est-à-dire sur la déconstruction du canon de la critique théâtrale londonienne. Nous

¹⁷ Griselda POLLOCK, « Des canons et des guerres culturelles », *Cahiers du genre* [En ligne], vol. 2, no. 47, 2007, URL : <https://www.cairn.info/revue-cahiers-du-genre-2007-2-page-45.htm>, consulté le 29 août 2022.

¹⁸ Pierre BOURDIEU, *La distinction : critique sociale du jugement*, Paris, Editions de Minuit, 1979.

commencerons par aborder la naissance de la critique numérique, puis nous introduirons notre corpus. Nous détaillerons la nature de nos sources, le choix des journaux et des blogs ainsi que les limites de notre base de données. Ce premier chapitre permettra ainsi de délimiter les contours de notre recherche.

Chapitre II

Après avoir présenté les contours théoriques de notre travail, nous nous focaliserons sur ses aspects techniques. Afin d'étudier la critique théâtrale londonienne dans son ensemble, nous avons créé un corpus de plus de 40 000 critiques théâtrales qui regroupent les productions écrites des journalistes et des blogueurs londoniens des années 2010 à 2020. Comme pour le premier chapitre, nous l'avons divisé en deux parties. Nous nous concentrerons dans un premier temps sur la récupération, la structuration, et le nettoyage des données. Ces tâches constituent un aspect important de notre recherche, puisqu'elles représentent plus d'une année de travail. Dans un second temps, nous exposerons les résultats de ce travail. Nous présenterons nos deux sous-corpus grâce à deux tableaux de bord qui résumeront les caractéristiques principales de notre base de données. Nous finirons par donner une brève présentation des 28 blogs qui constituent notre corpus témoin afin de mettre en avant leur appartenance aux marges de la culture théâtrale.

Chapitre III

Dans le troisième chapitre, nous présenterons les expériences que nous avons menées afin de cartographier les points de convergence et de divergence entre le centre de la critique théâtrale londonienne et ses marges. Ce chapitre sera divisé en trois parties qui correspondront à trois questions littéraires pour lesquelles nous avons mobilisé trois techniques informatiques différentes. La première sera dédiée au style des critiques. Dit autrement, « pouvons-nous observer des différences et / ou des similitudes entre la rédaction des journalistes et des blogueurs ? Et comment interpréter ces résultats ? » En nous appuyant sur la stylistique computationnelle, nous apporterons des premiers éléments de réponse à ces questions. Dans la deuxième partie, nous nous intéresserons à la répartition spatiale des journalistes et des blogueurs. En utilisant un Système d'Information Géographique (SIG), nous verrons comment la cartographie numérique, que nous avons mise au point, nous permet d'analyser

les relations « canon / périphéries » à partir d'une autre perspective. Enfin, la troisième partie portera sur les préférences esthétiques de ces deux communautés. Grâce à l'analyse d'opinion, un outil informatique qui permet de déterminer automatiquement les opinions d'un auteur dans un texte, nous verrons quels types de pièces les auteurs de ces deux communautés mettent à l'honneur.

Chapitre IV

Le dernier chapitre sera consacré aux stratégies que nous avons mises en place afin de contribuer à la Science Ouverte en partageant nos données. La première partie aura pour finalité de présenter la plateforme DramaCritiques.com que nous avons créée. Ce nouvel outil a un objectif principal, celui de participer à la vulgarisation des humanités numériques. En explicitant, en illustrant et en ouvrant nos travaux à l'ensemble de la communauté littéraire et scientifique, nous espérons que notre recherche pourra mener à d'autres analyses. Nous présenterons une partie de ces pistes d'ouverture dans la dernière partie de ce chapitre. Nos expériences réalisées sur la place des femmes dans le théâtre contemporain londonien montrent que notre corpus offre un vaste champ de possibles qui reste à explorer.

Liste des communications relatives à la thèse

20 – 24 juin, 2022. « Drama Critiques »

Colloque international « Digital Humanities in Theatre Research »

Organisé par Nora Probst et Miguel Escobar Varela

Université d'Islande, Reykjavik.

16 – 17 juin, 2022. « La Notion de « succès » des pièces de théâtre jouées sur la scène contemporaine londonienne vue à travers les algorithmes »

Journées d'études « Stylistique outillée et analyse des émotions et sentiments en littérature »

Organisées par Suzanne Mpouli et Dominique Legallois

Université Sorbonne Nouvelle Paris III, Paris.

17 mars, 2022. Présentation de mes recherches doctorales à l'occasion de l'événement « Génération IA 2030 »

Colloque européen des jeunes chercheurs et entrepreneurs en Intelligence Artificielle

Organisé par l'Ambassade de France en Belgique

Digitalcity Brussels, Bruxelles.

10-12 mars, 2022. « Mapping British Theatre Criticism Thanks to Numbers » (en ligne)

Groupe de travail « IFTR Digital Humanities, Data Modeling and /or Multimedia in Digital Theatre Research »

Organisé par Nora Probst et Miguel Escobar Varela

Université de Maryland, Etats-Unis.

18 novembre, 2021. « Algorithmique et scène théâtrale »

Séminaire « Humanités Numériques »

Organisé par Ioana Galleron et Dominique Legallois

Université Sorbonne Nouvelle Paris III, Paris.

9 novembre, 2021. « When Data Science Meets Theatre Studies »

Séminaire DHAI (Digital Humanities and Artificial Intelligence)

Organisé par Scott Trigg

Ecole normale supérieure (Ulm), Paris.

24 - 25 juin, 2021. « Quand la critique théâtrale anglaise passe au crible de l'informatique »

Colloque international « La Réception, perspectives contemporaines »

Organisé par Cécile Beaufils, Anne-Valérie Dulac, Jagna Oltarzewska, Diane Drouin et Maxence Gouleau

Sorbonne Universités, Paris.

18 juin, 2019. « La Critique théâtrale numérique envisagée à l'aune du numérique »

Journée des doctorants du laboratoire LATTICE (Langues, Textes, Traitements Informatiques et Cognition)

Organisée par Mathilde Regnault et Amine Lahouli

Laboratoire Lattice (CNRS, ENS, Paris Sorbonne Nouvelle), Montrouge

Liste des publications relatives à la thèse

Mylène MAIGNANT, Damien PELLE, Gaëtan BRISON, Thierry POIBEAU, « Drama Critiques' Database », [En ligne] *Journal of Open Humanities Data*, Août 2022, disponible sur : <https://openhumanitiesdata.metajnl.com/articles/10.5334/johd.81/>

« Drama Critiques : Comment concilier algorithmes et analyses littéraires ? » [En ligne] Publié le 5 juillet 2022 dans le *Magazine de Paris Sciences et Lettres*
Disponible sur : <https://explore.psl.eu/fr/le-magazine/focus/drama-critiques-comment-concilier-algorithmes-et-analyses-litteraires>

« Dans les coulisses des humanités numériques » [En ligne]
Publié le 11 octobre 2021 sur le site de L'Observatoire des humanités numériques de Paris Sciences et Lettres (PSL)
Disponible sur : <https://odhn.ens.psl.eu/newsroom/dans-les-coulisses-des-humanites-numeriques>

Mylène MAIGNANT, Gaëtan BRISON, Thierry POIBEAU, « Text Zoning of Theater Reviews: How Different are Journalistic from Blogger Reviews? », [En ligne] *Workshop on Natural Language Processing for Digital Humanities*, Décembre 2021, Sicha, India, disponible sur : https://hal.archives-ouvertes.fr/hal-03498270/file/Maignant_text_zoning.pdf

Liste des communications relatives aux humanités numériques

28 mai, 2019. « Mathematics and Computer Science on Samuel Beckett's Stage »
Journée d'études internationale « La Machine dans la littérature et les arts visuels »
Organisée par Omayyah Al-Shabab, Anouk Bottero et Diane Drouin
Sorbonne Universités, Paris.

1-4 mai, 2019. « Coding / Decoding Space and Characters in Samuel Beckett's Drama »
Digital Humanities in Theatre Research, A Working Group of the International Federation for
Theatre Research (IFTR)
Organisé par Doug Reside et Nic Leonhardt
New York Public Library for the Performing Arts, New York.

mars, 2019. « Computing 19th Century French Literature : The OuPoCo Project »
Journée d'études « Plotting Poetry » / « La Génération de poésie par ordinateur »
Organisée par Thierry Poibeau et Valérie Baudouin, avec le soutien du labex Transfers et de
l'IRN Cyclades (International Research Network « Corpora and Computational Linguistics
for Digital Humanities », soutenu par le CNRS).
École normale supérieure, Paris.

Liste des publications relatives aux humanités numériques

Thierry POIBEAU, Mylène MAIGNANT, Frédérique MELANIE-BECQUET, Clément
PLANCQ, Matthieu RAFFARD, et al. « Sonnet Combinatorics with OuPoCo », [En ligne]
*4th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social
Sciences, Humanities and Literature, ACL-SIGHUM*, Décembre 2020, Barcelone, Espagne,
disponible sur : <https://hal.archives-ouvertes.fr/hal-03084603/document>

Thierry POIBEAU, Mylène MAIGNANT, Frédérique MELANIE-BECQUET, Clément
PLANCQ, Matthieu RAFFARD, et al. « OuPoCo, the Combinatorial Poetry Workbench »,
[En ligne]*Colloque Digital Humanities 2020 (DH2020)*, Juillet 2020, Montréal, Canada,
disponible sur : <https://hal.archives-ouvertes.fr/hal-02977154>

Chapitre I

Etat de l'art

Ce premier chapitre a, pour finalité, de présenter les fondements théoriques de notre recherche. Notre problématique s'inscrit dans deux disciplines distinctes : les humanités numériques théâtrales d'une part, et les études de la réception d'autre part. A partir de l'utilisation d'outils techniques, il s'agit d'examiner la réception du théâtre londonien des années 2010 à 2020 à travers le regard de deux communautés de spectateurs différents. Pour comprendre comment se positionne notre travail et ce qu'il apporte dans le champ de la connaissance, nous présenterons ici les différents projets qui ont été réalisés dans ces deux champs de recherche.

Nous avons divisé ce chapitre en deux parties. Dans la première, intitulée « Théâtre, réception et numérique : tour d'horizon », nous nous focaliserons sur l'état de l'art des humanités numériques théâtrales dans un premier temps, et sur les études de la réception appliquées au théâtre dans un second temps. Nous verrons que, dans les deux cas, un nombre restreint d'études a été mené. Les travaux qui incluent une composante technique dans les études théâtrales sont très divers mais restent sous-représentés par rapport à d'autres disciplines, comme l'histoire ou la littérature romanesque. Probablement par manque de données exploitables grâce à des moyens computationnels, les humanités numériques théâtrales sont encore un champ en construction.

Les études de la réception présentent des caractéristiques similaires. Elles apparaissent au début des années 1970 et proposent de jeter la lumière sur le destinataire d'une œuvre pour mieux comprendre son sens. Bien que l'étymologie du mot « théâtre » nous rappelle que la présence du spectateur est nécessaire à une pièce pour qu'elle advienne¹⁹, peu d'auteurs se sont intéressés à la manière dont les œuvres théâtrales étaient reçues par leur public. Surtout, aucune étude ne propose une analyse chiffrée de la scène contemporaine londonienne en examinant les critiques des journalistes et des blogueurs.

Après avoir exposé les deux domaines théoriques dans lesquels s'inscrit notre recherche, nous nous concentrerons sur le cœur de notre projet de thèse, c'est-à-dire sur l'émergence d'une nouvelle forme de réception. Dans cette seconde partie, intitulée « Redéfinition de la critique théâtrale au XXI^e siècle », nous verrons en quoi ce public, par ses attentes différentes, semble offrir un autre regard sur la scène londonienne. La première sous-partie sera dédiée à la présentation de cette nouvelle communauté numérique. L'arrivée d'Internet favorise l'apparition d'une multitude de blogs à partir du début du XXI^e siècle. En 2010, date à laquelle commence notre étude, ce phénomène prit une ampleur véritable. Des

¹⁹ Le mot théâtre vient du grec *theatron*, qui signifie « regarder, contempler. » Le théâtre est donc le lieu « où l'on regarde. »

dizaines de blogueurs commencent à publier leurs analyses en ligne. En remettant en question la culture journalistique, considérée comme élitiste par nombreux de ces blogueurs, ces derniers entendent mettre à l'honneur les marges de la culture théâtrale londonienne.

Enfin, après avoir détaillé les différentes étapes qui ont mené à la naissance de cette nouvelle communauté, nous finirons ce chapitre par présenter les aspects principaux de notre objet d'étude. Nous répondrons plus précisément à ces deux questions :

1. Quelles approches techniques et littéraires avons-nous choisies et comment nos travaux se situent-ils parmi les projets en humanités numériques et en études théâtrales déjà existants ?
2. Quelles sont les caractéristiques et les limites de notre corpus ?

I.A Théâtre, informatique et réception : tour d’horizon

I.A.1 Humanités numériques et théâtre

I.A.1.1 Une discipline encore sous-représentée

Dans son essai *Les humanités numériques, une histoire critique* (2018), Pierre Mounier rappelle que l’expression « humanités numériques » n’apparaît qu’en 2004, dans l’ouvrage collectif *A Companion to Digital Humanities*, dirigé par Susan Schreibman, Ray Siemens et John Unsworth²⁰. L’étude des humanités, grâce à des moyens computationnels, remonte à l’après-guerre lorsque le père jésuite Roberto Busa voulut utiliser les machines d’IBM pour étudier *La Somme théologique* de Thomas d’Aquin. C’est cette demande particulière qui permit aux textes d’être considérés « comme objet[s], préfigurant le concept de *distant reading* proposé par Franco Moretti plusieurs décennies plus tard²¹. » A partir des années 1970, les projets à la frontière de l’informatique et des humanités connurent un succès grandissant. L’usage des outils techniques pour aborder des textes historiques et littéraires, parmi d’autres, devint un formidable moyen pour examiner de larges corpus dans leur ensemble.

Comme l’explique Clarisse Bardirot, les études théâtrales sont cependant restées en marge de ce phénomène²². Les premières tentatives en humanités numériques et en études théâtrales commencent dans les années 1970. En 1973, Joseph Donohue et James Ellis furent partie des premiers universitaires à utiliser le numérique pour leurs recherches théâtrales. Intitulé *The London Stage 1800-1900: A Documentary Record and Calendar of Performances*²³, leur projet avait pour dessein de réunir et d’organiser toutes les informations relatives aux spectacles (titres des pièces produites, nombre de spectateurs, prix du ticket d’entrée, etc.) joués au XIX^e siècle à Londres. Un site sommaire fut créé afin que les utilisateurs puissent avoir accès aux quelques pages qui réunissaient les premiers matériaux collectés par les deux auteurs. Comme le montre la capture d’écran en figure 1, l’utilisation du

²⁰Susan SCHREIBMAN, Ray SIEMENS et John UNSWORTH, *A Companion to Digital Humanities*, Oxford, Blackwell, 2004.

²¹ Pierre MOUNIER, *Les humanités numériques, une histoire critique*, Paris, Éditions de la Maison des sciences de l’homme, 2018, p. 24.

²² BARDIROT, *op. cit.*

²³ Joseph DONOHUE, « The London Stage 1800-1900: A Data Base for a Calendar of Performances on the Nineteenth-Century London Stage », *Computers and the Humanities*, Vol. 9, no. 4, 1975.

numérique était alors encore restreinte. Les fonctionnalités proposées étaient rudimentaires. Les informations, qui étaient alors récupérées manuellement, étaient simplement mises à disposition sous forme de listes ou de tableaux en noir et blanc :

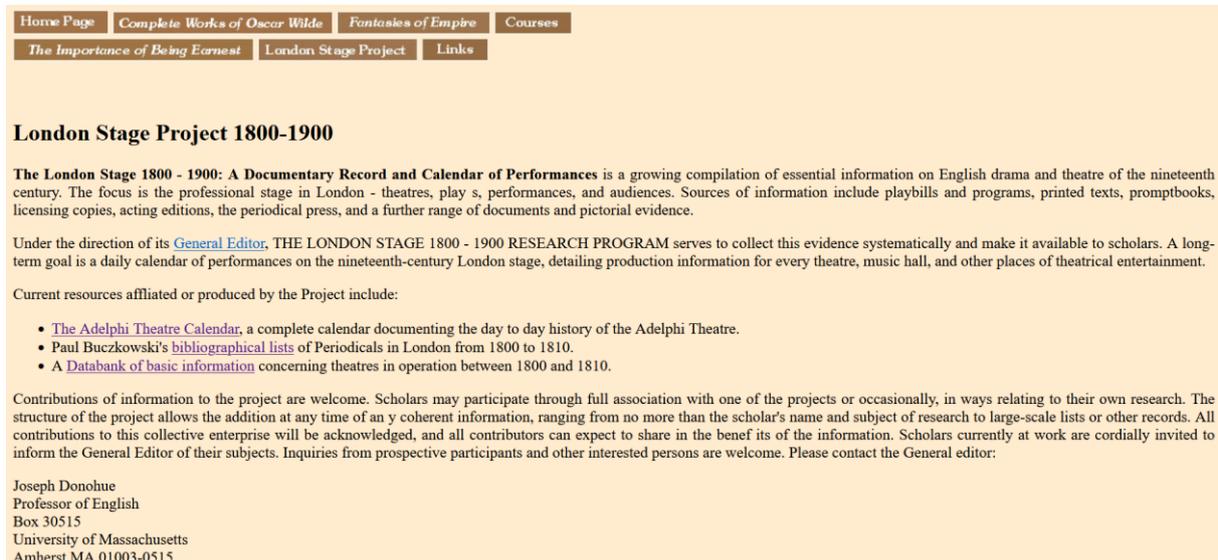


Fig.1. Capture d'écran de la page d'accueil du projet *The London Stage 1800-1900: A Documentary Record and Calendar of Performances*.

Il faudra attendre plusieurs décennies pour que les études théâtrales s'insèrent pleinement dans le champ des humanités numériques. Alors que de nombreux projets numériques dans l'ensemble des humanités voient le jour à partir de la fin des années 1990, ce n'est qu'à partir de 2010 qu'une « prise de conscience collective et institutionnelle se structure autour de plusieurs initiatives dans les études théâtrales²⁴ ». Bardirot le rappelle, deux événements en particulier marquent le début de ce nouveau consensus. Tandis que le 20 et 21 juin 2013 est organisé le premier THATcamp à New York (« Performing Arts: Presenting and Archiving Performance in the Digital Age »), la première *unconference* « DH and Theatre Research », organisée par l'American Society of Theatre Research apparaît quelques mois plus tard. La même année, le premier article sur les études théâtrales classiques et sur les humanités numériques est publié par Paul Spence en langue espagnole²⁵.

C'est en 2014 que le premier groupe officiel de recherche en humanités numériques et en théâtre est créé par la FIRT (Fédération Internationale pour la Recherche Théâtrale). Ce groupe se réunit aujourd'hui tous les ans pendant plusieurs jours afin de permettre à ses

²⁴BARDIROT, *op. cit.* p. 2.

²⁵Paul SPENCE, « Teatro clásico y humanidades digitales: el cruce entre método, proceso y nuevastecnologías », *Teatro de Palabras*, no. 7, 2013, p. 9 – 38.

participants de se fédérer autour de thématiques communes. En dehors de cette conférence, il propose également un atelier de réflexion annuel qui permet à chacun de ses membres de présenter ses travaux de recherche inachevés mais en cours de réalisation. Ces temps de partage permettent aux chercheurs d'affiner leurs hypothèses, tant dans le domaine des études théâtrales qu'en informatique, et d'aborder les problématiques rencontrées. Les études théâtrales n'en sont donc qu'à leurs débuts dans le vaste champ des humanités numériques. La pluralité des outils techniques et la profusion des sujets ne favorisent pas son unité.

Trois catégories de projets

Ces projets hybrides, à l'intersection de deux disciplines différentes, peuvent être divisés en plusieurs catégories. Pour Nic Leonhardt, ce sont sept catégories qui peuvent être identifiées et réparties entre les projets qui portent sur les archives, sur les acteurs ou sur les dramaturges, sur l'architecture des théâtres, sur la danse, sur l'étude des réseaux, et enfin sur la recherche expérimentale ou éducative²⁶. Deborah Caplan propose, quant à elle, une répartition en quatre catégories : les projets qui reposent sur les archives et sur la numérisation, ceux sur les environnements théâtraux, ceux sur la visualisation des données et enfin ceux qui se focalisent sur la création de bases de données²⁷. Sarah Bay-Cheng réduit ces catégories à trois, en identifiant d'abord la collection des données, puis leur analyse, et enfin leur partage²⁸.

Afin de mieux cerner les contours théoriques et pratiques de ce champ dans lequel s'inscrit cette recherche, nous nous appuyons sur la structure et les arguments de l'article « Arts de la scène et *culture analytics*²⁹ » rédigé par Clarisse Bardiot. Comme Bay-Cheng, Bardiot identifie trois types de projets en humanités numériques et en théâtre. La classification qu'elle propose est non seulement particulièrement claire, mais aussi pertinente pour comprendre dans quels sillages théoriques s'inscrit ce travail.

Selon Bardiot, se positionnent en premier lieu les chercheurs qui s'intéressent à l'histoire théâtrale dans sa globalité. Une longue ou une courte période historique est ainsi prise en considération et est analysée grâce à différentes techniques de modélisation informatique. Viennent ensuite les projets qui se focalisent sur un objet d'étude plus concis.

²⁶ Cité par Miguel VARELA ESCOBAR, *Theater as Data*, Michigan, University of Michigan Press, 2021, p. 2.

²⁷ Debra CAPLAN, « Notes from the Frontier: Digital Scholarship and the Future of Theatre Studies », *Theatre Journal*, vol. 67, no. 2, 2015, p. 347–59.

²⁸ Sarah BAY-CHENG, 2017. « Digital Historiography and Performance », *Theatre Journal*, vol. 68, no. 4, 2017, p. 507–27.

²⁹ BARDIOT, *op. cit.*

Ces derniers examinent le texte des pièces de théâtre exclusivement. Enfin, la dernière catégorie que Bardiot identifie concerne les outils de visualisation des données théâtrales. Pour mieux comprendre en quoi consistent ces trois catégories, nous avons sélectionné deux exemples pour chacune d'entre elles. Cette recontextualisation nous permettra de saisir comment se situe notre propre recherche par rapport aux travaux déjà effectués dans ce domaine et ce qu'elle peut apporter à ce champ encore en construction.

I.A.1.1 Première catégorie : L'histoire du théâtre

Exemple 1 : Visualizing Broadway

La première catégorie concerne les projets qui examinent l'histoire théâtrale dans sa globalité. Une des recherches emblématiques de cette approche est intitulée « Visualizing Broadway » et a été menée par Derek Miller. En novembre 2000, la League of American Theatres and Producers et l'Etat de New York ont impulsé l'Internet Broadway Database³⁰. Cette base de données gratuite recense un grand nombre d'informations sur toutes les représentations théâtrales qui ont eu lieu à Broadway depuis 1915 jusqu'à nos jours.

En analysant cette masse de données gigantesque grâce à diverses méthodes informatiques et en la croisant avec les données de la Playbill Vault³¹, Miller relit un siècle de l'histoire des grands spectacles de Broadway à travers un prisme socio-économique³². En se concentrant non plus sur les succès de Broadway mais sur l'ensemble des spectacles, Miller jette ainsi la lumière sur l'importance de toutes les professions théâtrales qui sont rarement mentionnées, comme les scénographes, les costumiers ou les régisseurs.

Son site Internet, VisualizingBroadway.com³³, permet aux utilisateurs de découvrir ses autres travaux de recherche en construction. Ses résultats, parfois accompagnés d'explications à la fois littéraires, sociologiques et techniques, montrent l'originalité et la richesse qu'offre l'utilisation du numérique. La visualisation présentée en figure 2 est extraite de son site VisualizingBroadway.com, et représente les personnes les plus influentes de Broadway entre 1940 et 1960 :

³⁰<https://www.ibdb.com/>[En ligne], consulté le 23 mai 2021.

³¹<https://www.playbill.com/vault>[En ligne], consulté le 23 mai 2021.

³² Derek MILLER, « Average Broadway », *Theatre Journal*, vol. 68, no. 4, 2016.

³³<https://visualizingbroadway.com/index.html>[En ligne], consulté le 23 mai 2021.

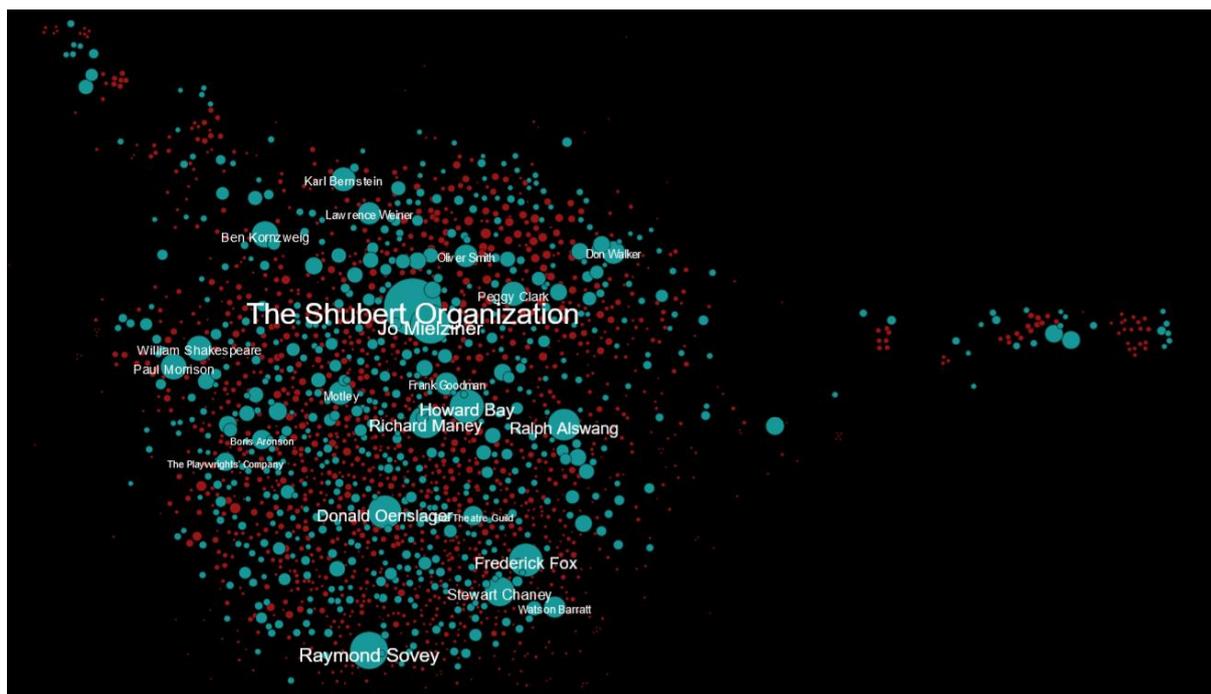


Fig.2. Capture d'écran issue de la catégorie « Words in Progress » du site VisualizingBroadway.com créé par Derek Miller.

Exemple 2 : Ibsen Stage

Un autre projet qui s'inscrit également dans une perspective numérique à grande échelle est l'IbsenStage³⁴. Le Centre Ibsen de l'Université d'Oslo ainsi que la bibliothèque nationale de Norvège et l'Australian Live Performance Database AusStage³⁵ ont collaboré afin de produire une base de données gratuite de plus de 23 000 informations sur la réception des pièces d'Henrik Ibsen dans le monde, de 1850 à nos jours. Les utilisateurs peuvent aujourd'hui avoir accès à ces données grâce à une interface web qui est régulièrement mise à jour par ses créateurs.

La plateforme IbsenStage ne propose pas d'analyses littéraires, sociologiques ou historiques de ces données. Comme l'Internet Broadway Database, elle a été conçue comme un espace numérique de recueil d'informations destiné à quiconque serait intéressé par l'histoire des pièces d'Ibsen. En revanche, dix modules ont été intégrés au site Internet afin que ses utilisateurs puissent obtenir certaines visualisations des données et puissent en tirer des analyses. Il est également possible d'effectuer des recherches avancées grâce à la sélection d'un grand nombre de caractéristiques sur les personnes qui ont participé aux productions

³⁴<https://ibsenstage.hf.uio.no/>[En ligne], consulté le 23 mai 2021.

³⁵<https://www.ausstage.edu.au/pages/browse/>[En ligne], consulté le 23 mai 2021.

(acteurs, metteurs en scène, musiciens, etc.), sur le genre des adaptations (théâtre pour enfants, cabarets, films, etc.), sur les langues dans lesquelles les pièces ont été traduites, etc.). L'image ci-dessous représente la page d'accueil de leur site web avec les différents onglets qui permettent d'accéder aux dix catégories de la plateforme :

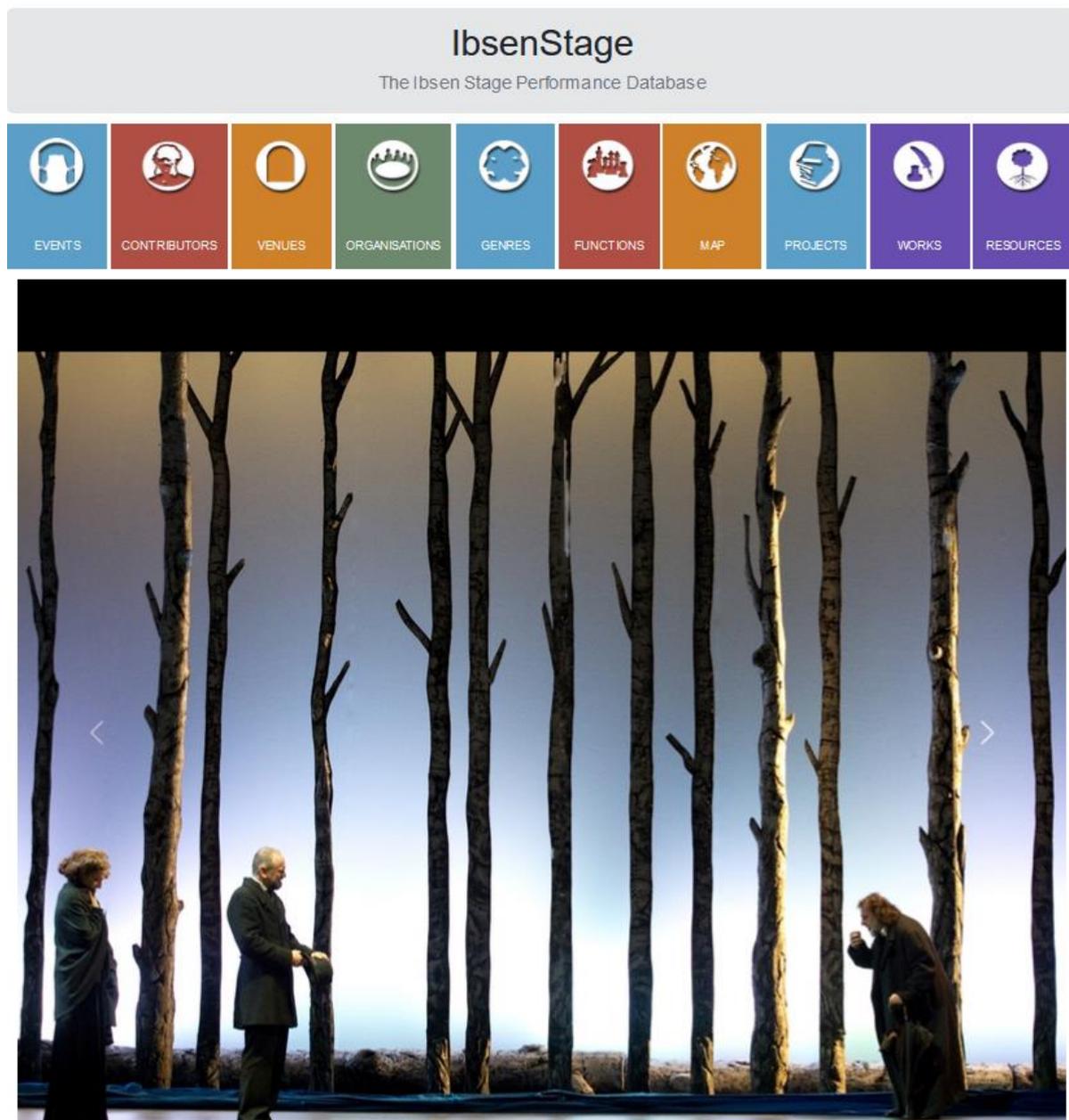


Fig.3. Capture d'écran de la page d'accueil du site IbsenStage.

La richesse d'IbsenStage a donné lieu à deux études quantitatives sur le théâtre d'Ibsen. En 2016, Julie Holledge, Jonathan Bollen, Frode Helland et Joanne Tompkins ont

travaillé sur le contexte social, politique et économique de *A Doll's House* (1879)³⁶ afin de comprendre pourquoi et comment cette pièce était devenue l'un des plus grands succès d'Ibsen. En reposant leur étude sur deux approches complémentaires, l'une portant sur la transmission de la pièce, l'autre sur ses adaptations dans le monde, ces quatre auteurs ont pu représenter la circulation de cette pièce à travers les déplacements de plus de 16 000 artistes dans le monde. Cette expérience leur a notamment permis de démontrer, entre autres, comment *A Doll's House* s'était progressivement inscrite dans le canon théâtral.

Deux ans plus tard, Jens-Morten Hanssen s'est également appuyé sur la base de données IbsenStage afin d'étudier les représentations théâtrales d'Ibsen sur la scène allemande de 1876 à 1918³⁷. Ses modélisations lui permettent par exemple d'observer que le premier grand succès d'Ibsen fut la pièce *Pillars of Society* (1877), alors que sa très célèbre pièce *A Doll's House* (1879) reçut une critique plutôt négative à la même époque sur la scène allemande.

Ces exemples illustrent la manière dont l'informatique permet d'aborder des thématiques théâtrales sous un angle différent. Il ne s'agit plus tant d'étudier les textes en eux-mêmes mais plutôt leur rayonnement à travers leur circulation dans l'histoire. Les outils numériques servent ainsi de tremplins à la construction de nouveaux savoirs en complétant les études littéraires dites « classiques ». Sans l'informatique et ces bases de données, il serait impossible d'avoir une vue d'ensemble précise de ces phénomènes culturels.

I.A.1.3 Deuxième catégorie : l'étude des textes

Exemple 1 : Les pièces de William Shakespeare

Tandis que la première catégorie se focalise sur l'histoire du théâtre, la deuxième catégorie inclut des projets qui se focalisent sur les pièces de théâtre en elles-mêmes, c'est-à-dire sur les textes littéraires. L'exemple le plus connu est celui qui porte sur la paternité des pièces de William Shakespeare. La stylométrie, c'est-à-dire l'utilisation d'outils statistiques pour décrire les propriétés stylistiques d'un texte, a été employée à de nombreuses reprises pour déterminer si certaines de ses pièces avaient été écrites par lui seul ou étaient le fruit

³⁶Julie HOLLEDGE, Jonathan BOLLEN, Frode HELLAND, Joanne Elizabeth TOMPKINS, *A Global Doll's House: Ibsen and Distant Visions*, Londres, Palgrave Macmillan, 2016.

³⁷ Jens-Morten HANSSSEN, *Ibsen on the German Stage 1876 – 1918, A Quantitative Study*, Narr FranckeAttempo Verlag, Tübingen, 2018.

d'une collaboration de plusieurs auteurs. Hugh Craig et Arthur Kinney dans *Shakespeare, Computers, and the Mystery of Authorship*³⁸, Brian Vickers dans son article « Shakespeare and Authorship Studies in the Twenty First Century³⁹ » ou bien encore Jackson MacDonald dans « Authorship and the Evidence of Stylometrics⁴⁰ » explorent cette thématique à travers l'utilisation de diverses méthodes numériques.

D'autres chercheurs ont utilisé ces méthodes computationnelles pour analyser plus finement le style de Shakespeare sur l'ensemble de sa carrière. C'est le cas de Jonathan Hope et de Michael Witmore qui, avec DocuScope, un logiciel d'aide à la visualisation des données, ont pu observer trois périodes qui correspondent à trois groupes de caractéristiques linguistiques particulières. Les deux chercheurs émettent l'hypothèse que « le style adopté par Shakespeare à la fin de sa carrière commence bien plus tôt que ne l'affirme la tradition⁴¹. » A l'aide d'outils stylistiques, Hope et Witmore démontrent comment le style de Shakespeare à la fin des années 1590 « indique le passage d'une dénotation du monde réel vers l'expression de la subjectivité de l'interlocuteur⁴². »

Exemple 2 : Samuel Beckett Digital Manuscript Project

Le Samuel Beckett Digital Manuscript Project propose, lui aussi, une approche computationnelle des archives des œuvres de Beckett⁴³. En numérisant tous les brouillons, les notes de lecture et les traductions de Beckett, Mark Nixon et Dirk Van Hulle ont non seulement contribué à la préservation de documents fragiles et disséminés dans divers endroits (musées, bibliothèques, collections privées, etc.), mais aussi à construire une compréhension plus globale de l'héritage beckettien. Grâce à une inscription payante à la plateforme Beckettarchive.org, les utilisateurs peuvent désormais avoir accès à l'ensemble des documents qui ont nourri les réflexions de Beckett (pages cornées, passages soulignés, ouvrages de sa bibliothèque personnelle, etc.).

³⁸ Hugh CRAIG et Arthur KINNEY, *Shakespeare, Computers, and the Mystery of Authorship*, Cambridge, Cambridge University Press, 2009.

³⁹ Brian VICKERS, « Shakespeare and Authorship Studies in the Twenty-First Century », *Shakespeare Quarterly*, vol. 62, no. 1, 2011, p. 106 – 42.

⁴⁰ Jackson P. MACDONALD, « Authorship and the Evidence of Stylometrics », *Shakespeare beyond Doubt: Evidence, Argument, Controversy*, Cambridge, Cambridge University Press, 2013, p. 100 – 110.

⁴¹ Jonathan HOPE et Michael WITMORE, « Quantification and the language of later Shakespeare », *Actes des congrès de la Société française Shakespeare* [En ligne], vol. 31, 2014. URL : <https://journals.openedition.org/shakespeare/2830>, consulté le 11 mai 2021.

⁴²*Ibid.*

⁴³<https://www.beckettarchive.org/> [En ligne], consulté le 11 mai 2021.

Un important travail fondé sur la TEI a également été réalisé. La TEI (Text Encoding Initiative) vise à uniformiser le codage de documents pour faciliter leurs échanges et leurs publications à un niveau international. Pour Lou Bernard, l'un des fondateurs de la TEI, son objectif principal consiste à « fournir des recommandations pour la création et la gestion sous forme numérique de tout type de données créées et utilisées par les chercheurs en Sciences humaines et sociales⁴⁴. » La TEI représente aujourd'hui une part importante des travaux en humanités numériques et permet, entre autres, d'enrichir et de rendre apparentes la structure et la signification d'un texte.

En encodant les manuscrits rédigés par Beckett, Nixon et Van Hulle jettent ainsi la lumière sur la génétique de son œuvre. Toutes les ratures, les ajouts et les suppressions de mots ainsi que les déplacements de certaines phrases ou de certains paragraphes par Beckett ont été minutieusement encodés. Ce travail a donné lieu à plusieurs publications qui ont permis de révéler de nouveaux éléments dans la construction de son œuvre. L'image ci-dessous (Fig. 4) est une capture d'écran des deux types d'informations mises à disposition des utilisateurs. Tandis qu'à gauche se situe une version scannée de la première page du script de *Krapp's Last Tape* (1958), à droite se trouve sa version encodée en TEI :

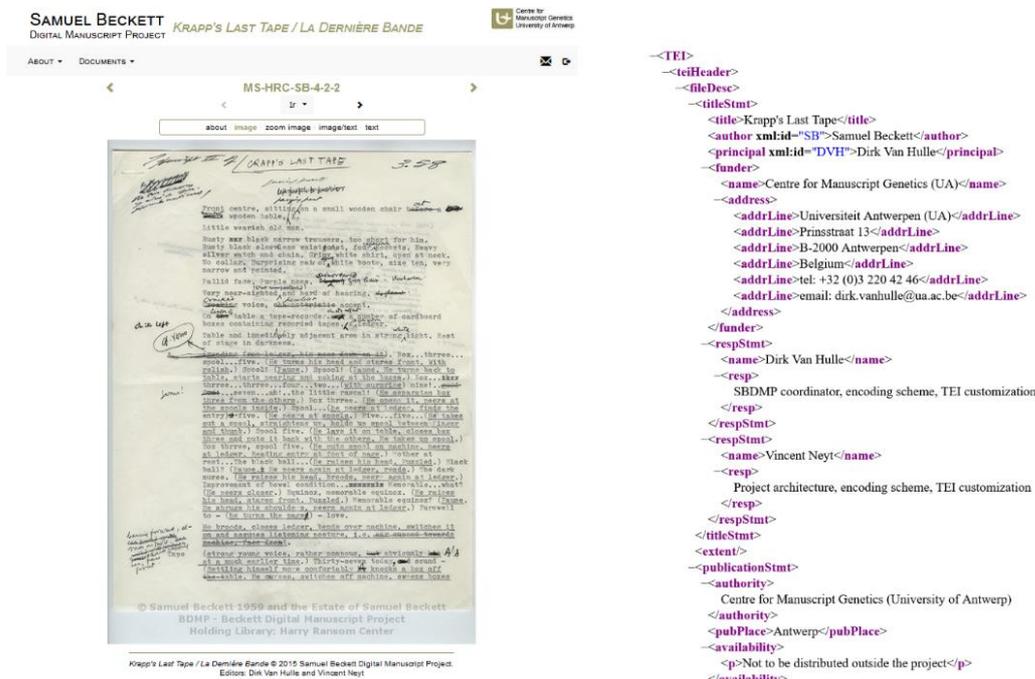


Fig.4. Version scannée de la première page de *Krapp's Last Tape* (gauche) et sa version encodée en TEI (droite).

⁴⁴Lou BURNARD, « Introduction », *Qu'est-ce que la Text Encoding Initiative ?* [En ligne], OpenEdition Press, coll. « Encyclopédie numérique », 2015, URL : <http://books.openedition.org/oepp/1297> consulté le 08 mai 2022.

I.A.1.4 Troisième catégorie : Logiciels de visualisation

Exemple 1 : ReKall

La troisième catégorie des projets en humanités numériques et en études théâtrales identifiée par Bardirot concerne les travaux qui portent sur la visualisation des scripts des pièces et sur les représentations scéniques. ReKall⁴⁵, fondé par Bardirot, est un environnement open source emblématique de cette catégorie. Sa version professionnelle commence à prendre forme en 2013 et est mise à disposition du grand public à la fin du mois de septembre 2015. Ce logiciel permet à des artistes, à des professeurs, à des étudiants mais aussi à des institutions culturelles d'analyser, de documenter, d'archiver et de partager des documents visuels. Le principe repose sur la création de « capsules » qui consistent à ajouter des commentaires, des liens Internet ou des documents aux vidéos. Les utilisateurs peuvent ensuite publier leur travail en ligne. La capture d'écran (Fig. 5) ci-dessous est extraite du site RemoRekall.com, et illustre l'interface de la plateforme :

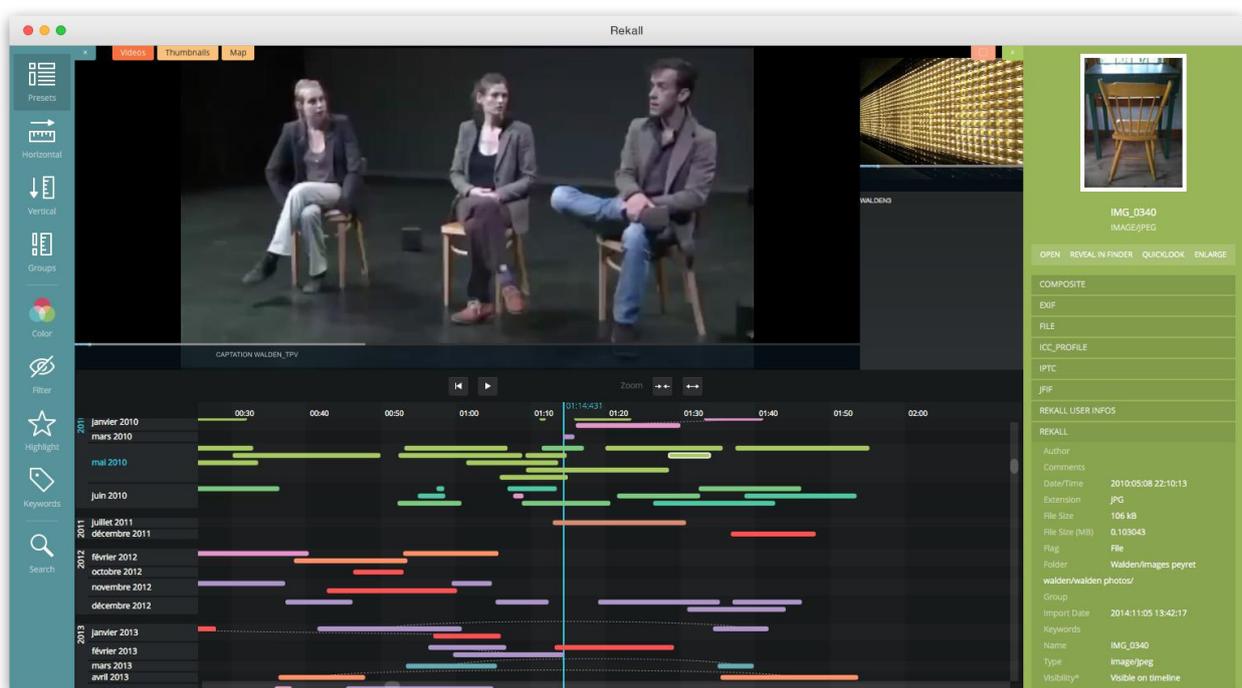


Fig.5. Capture d'écran du logiciel ReKall.

⁴⁵<https://memorekall.com/fr/> consulté le 02 mai 2021.

Exemple 2 : SET (Simulated Environment for Theatre)

Le SET (Simulated Environment for Theatre)⁴⁶ est un autre projet qui s'insère dans cette troisième catégorie. La première version du SET fut mise en place par une équipe de chercheurs pluridisciplinaires venant de différentes universités américaines en 2010. Elle permet aux utilisateurs de travailler simultanément sur leur texte et sur la mise en scène de la pièce choisie. Comme le montre la capture d'écran ci-dessous (Fig. 6), un encadré dans lequel est affiché le texte de la pièce se situe en bas de la page, accompagné d'une section « Notes & SDs » qui permet d'ajouter des commentaires. Cet outil numérique invite ainsi ses utilisateurs à réfléchir sur l'aspect visuel des spectacles vus, mais aussi à créer et à imaginer leur propre scénographie.

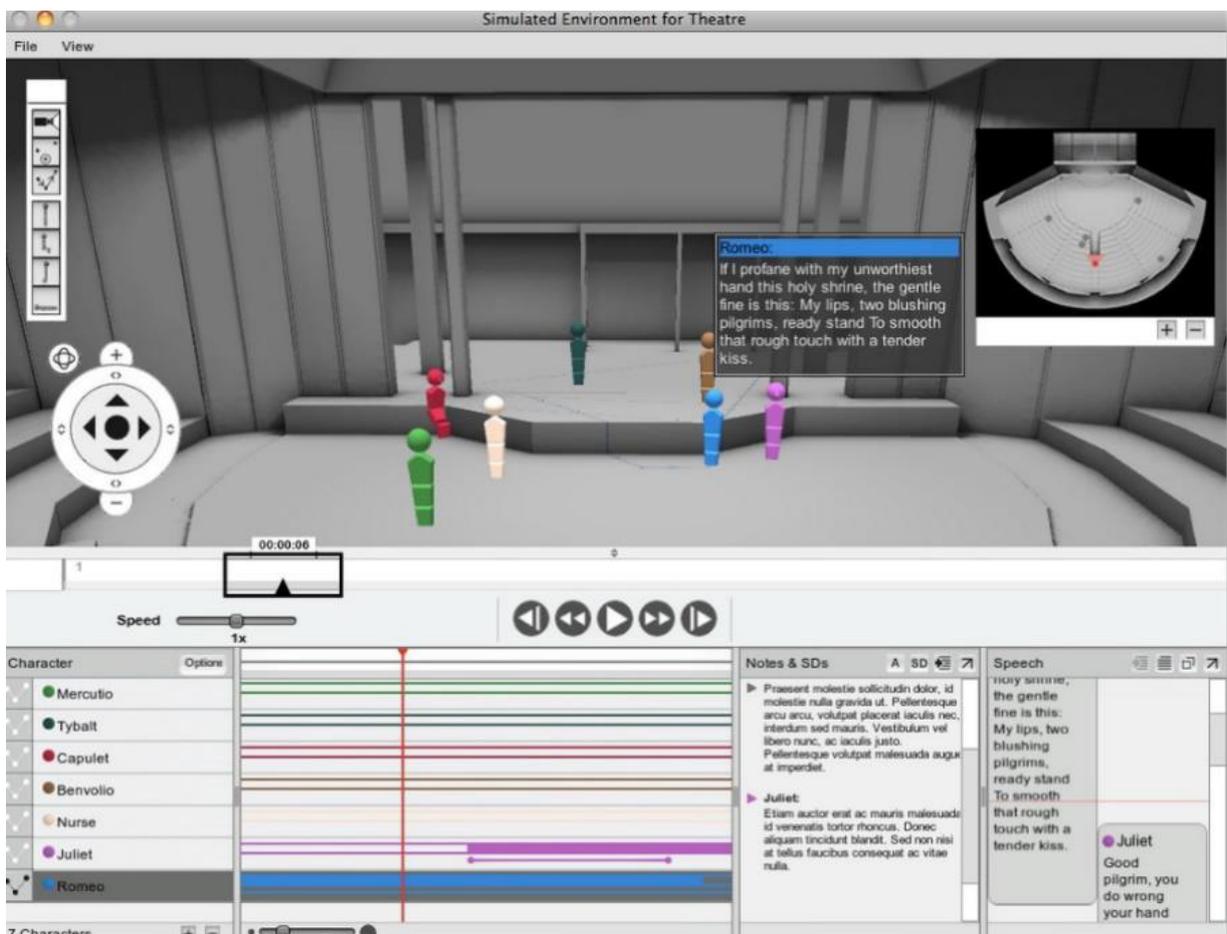


Fig.6. Capture d'écran du logiciel Simulated Environment for Theatre.

⁴⁶ Jennifer ROBERTS-SMITH, « Visualizing Theatrical Text: From Watching the Script to the Simulated Environment for Theatre (SET) », *Digital Humanities Quarterly* [En ligne], vol. 7, no. 3, 2013, URL : <http://www.digitalhumanities.org/dhq/vol/7/3/000166/000166.html>, consulté le 02 mai 2021.

Transition

Nous venons de présenter les trois catégories de projets en humanités numériques et en études théâtrales que Bardiot a identifiées. La présentation des six exemples dans ce domaine nous a permis de constater que l'utilisation du numérique permet d'aborder des questions théâtrales sous un angle nouveau. Que cela soit à grande ou à petite échelle, l'informatique aide le chercheur à obtenir une vue d'ensemble de son corpus et à en extraire des éléments difficilement repérables par l'humain. Bien que ces projets soient très variés, leur sous-représentation par rapport à d'autres disciplines témoigne néanmoins d'une nécessité d'explorer des voies nouvelles pour construire ce champ qui en est encore à ses débuts.

Notre recherche s'inscrit dans la première catégorie définie par Bardiot, c'est-à-dire celle qui concerne les projets qui portent sur l'histoire théâtrale. À l'aide d'outils techniques, la finalité de cette thèse est d'explorer la réception du théâtre londonien de 2010 à 2020 à travers le regard de deux communautés critiques différentes. Notre objet d'étude ne porte donc pas sur les textes de théâtre en eux-mêmes, mais plutôt sur la circulation des discours qui le construisent.

Comme pour *Visualizing Broadway* ou *Ibsen Stage*, nous nous intéressons donc à l'histoire du théâtre. Le numérique nous servira à analyser notre corpus dans son ensemble, afin d'en extraire des *patterns*, ou des schémas répétitifs, qui nous aideront à comprendre ce phénomène culturel à grande échelle. L'approche numérique n'ayant encore jamais été exploitée pour l'étude de ce sujet, les résultats apportés seront nécessairement novateurs.

Le deuxième aspect central de notre recherche porte sur les études de la réception. Notre travail se concentrera sur la manière dont le théâtre londonien est *reçu* par les journalistes et les blogueurs. Ces derniers constituent deux groupes de spectateurs divergents, ce qui nous a amenés à faire l'hypothèse que leur perception du théâtre sera différente. Au delà de l'aspect numérique, ce projet s'inscrit donc également dans un autre champ théorique précis qu'il convient de délimiter pour comprendre les concepts littéraires auxquels il fait référence.

Dans la partie suivante, nous nous attacherons à présenter les caractéristiques principales des études de la réception. Ces dernières proposent d'étudier l'importance de la subjectivité du destinataire d'une œuvre. Pour comprendre un objet culturel, il faut replacer les attentes de son public dans leur contexte historique, social et politique. Nous exposerons dans la partie qui suit un des concepts-clés des études de la réception. Ce que Hans-Robert

Jauss nomme l'« horizon d'attente » est précisément la multitude des façons de « réceptionner » une œuvre.

Nous verrons en quoi ce concept nous permettra de penser la relation entre les deux formes de réception que nous étudions (les journalistes versus les blogueurs). Ces notions littéraires abordées, nous finirons cette partie par décrire les études qui ont été menées sur le théâtre et sur leur public. Nous pourrions conclure qu'un nombre restreint de travaux a été réalisé jusqu'ici, et que notre recherche contribue ainsi à la construction de ce champ.

I.A.2 Réception et théâtre anglais contemporain

I.A.2.1 Définition des études de la réception

Les études de la réception apparaissent au début des années 1970. D'origine française et latine, le terme « réception » est revenu par l'allemand (Rezeption), et est traduit dans de nombreuses langues européennes (« receptié » en néerlandais, « recepcja » en polonais, etc.)⁴⁷. Le Dictionnaire de la critique littéraire le définit comme l'acte de

perception d'une œuvre par le public. (...) Etudier la réception d'un texte, c'est accepter que la lecture d'une œuvre est toujours une réception qui dépend du lieu et de l'époque où elle prend place⁴⁸.

Ces études proposent une approche relationnelle de la littérature au texte en mettant la sensibilité du récepteur en son centre. Ce n'est plus tant le contenu des objets culturels qui est examiné, mais plutôt la circulation des discours qu'ils engendrent parmi le public. Le sens d'une œuvre, qu'elle soit théâtrale, cinématographique, ou littéraire, est créé par le corps social qui la reçoit. Les études de la réception se focalisent ainsi sur l'expérience vécue par le destinataire. C'est l'interprétation donnée par ce dernier qui permettra au sens d'une œuvre d'advenir.

Cette dynamique inversée du rapport de l'œuvre au public n'a pu advenir que parce qu'un contexte politique et culturel particulier le permettait. A partir des années 1970, une série de ruptures théoriques favorisa la mise en place d'une nouvelle sensibilité qui allait peu à peu accorder davantage d'importance à l'individu. Le groupe Tel Quel, qui remit en question les notions d'auteur, d'œuvre et de création, fut un des premiers à ouvrir la voie à ces changements. Roland Barthes, dans *La Mort de l'auteur* (1967), mit par exemple en avant que le sens d'une œuvre ne devait non plus être compris en fonction de l'auteur, mais plutôt grâce à son intertextualité.

Les essais sociologiques de Pierre Bourdieu permirent également de mettre en évidence que la culture pouvait être un instrument de domination. Les pratiques culturelles étant définies par les positions sociales, l'existence d'une culture dite « légitime » devint par conséquent un concept relatif. Enfin, un nombre de publications philosophiques et

⁴⁷ Yves CHEVREL, « Les études de réception », *Précis de littérature comparée*, Paris, Presses Universitaires de France, 1989, p. 177.

⁴⁸ Joëlle GARDE-TAMINE, Marie-Claude HUBERT, *Dictionnaire de critique littéraire*, Paris, Edition Armand Colin, 2002, p.174.

ethnographiques, notamment par de Michel de Certeau et Jean-François Lyotard, jetèrent la lumière sur la multiplicité des points de vue. Pour de Certeau, l'activité de lecture est semblable à celle du braconnage. Le lecteur, comme le braconnier, met en place des tactiques, joue et résiste au texte, de sorte que « le texte n'a de signification que par ses lecteurs⁴⁹. » En remplaçant l'unité par la pluralité, ces changements théoriques favorisèrent ainsi la prise en compte de la multiplicité des points de vue.

I.A.2.2 Hans-Robert Jauss et l'horizon d'attente

C'est Hans-Robert Jauss (1921 – 1997), historien et théoricien allemand de la seconde moitié du XX^e siècle, qui fut un des premiers à théoriser les études de la réception. Le terme « réception » est souvent associé à son essai *Pour une esthétique de la réception* qui parut en 1972. Aux côtés de Wolfgang Iser, il fut à l'origine de l'Ecole de Constance, un groupe de recherche littéraire qui mit le destinataire au centre de ses théories⁵⁰.

Proposant une vision historique de la littérature, Jauss prône l'importance de l'histoire des effets d'une œuvre sur le public. C'est en prenant en compte l'interprétation du lectorat durant une période donnée qu'une œuvre prendra tout son sens. « Qu'il la rejette, l'oublie, la réhabilite ou l'encense, le destinataire détermine donc sa postérité⁵¹. » La notion d'œuvre est donc redéfinie. Ce n'est pas parce qu'une pièce de théâtre existe qu'elle peut être qualifiée d'œuvre. C'est parce qu'elle est lue, et dans le contexte du théâtre, vue par un public, qu'elle devient « œuvre. »

Le terme-clef de sa théorie repose sur ce que Jauss appelle l'« horizon d'attente », déjà en germe dans les réflexions d'Husserl, de Mannheim et de Popper⁵². Ce concept fait référence aux diverses attentes d'un public. Celles-ci sont façonnées par les connaissances préalables du lecteur ou du spectateur et par la hiérarchie des valeurs esthétiques d'une époque donnée. Pour comprendre le sens d'une œuvre, il faut alors

⁴⁹ Michel DE CERTEAU, *L'Invention du quotidien. T. 1. Arts de faire*, Paris, Gallimard, 1980, p. 241.

⁵⁰ Isabelle KALINOWSKI, « Hans-Robert Jauss et l'esthétique de la réception », *Revue germanique internationale* [En ligne], vol. 8, 1997, URL : <https://journals.openedition.org/rgi/649>, consulté le 15 septembre 2015.

⁵¹ Elsa MARPEAU, François-René MARTIN, « RÉCEPTION, art et littérature », *Encyclopædia Universalis*[En ligne], URL : <https://www.universalis.fr/encyclopedie/reception-art-et-litterature/>, consulté le 12 septembre 2021.

⁵² Gisèle SAPIRO, « LITTÉRATURE - Sociologie de la littérature », *Encyclopædia Universalis*[En ligne], disponible sur : <https://www.universalis.fr/encyclopedie/litterature-sociologie-de-la-litterature/>, consulté le 15 septembre 2021.

s'intéresser aux interprétations qu'elle a reçues et qui sont influencées par les conditions sociales, esthétiques et cognitives d'un individu. C'est cet ensemble qui forme « l'horizon d'attente » d'un spectateur, c'est-à-dire un

système de références objectivement formulable qui, pour chaque œuvre au moment de l'histoire où elle apparaît, résulte de trois facteurs principaux : l'expérience préalable que le public a du genre dont elle relève, la forme et la thématique d'œuvres antérieures dont elle présuppose la connaissance, et l'opposition entre langage poétique et langue pratique, monde imaginaire et réalité quotidienne⁵³.

Il y a donc pour Jauss un déjà-là littéraire dans chacune des œuvres qui est partagé par un ensemble d'individus. Dit autrement, le bagage culturel ainsi que le vécu personnel (« réalité quotidienne ») affecteront le jugement qu'un individu portera sur une œuvre donnée.

L'œuvre littéraire n'a qu'une autonomie relative. Elle doit être analysée dans un rapport dialectique avec la société. Plus précisément, ce rapport consiste dans la production, la consommation et la communication de l'œuvre lors d'une période définie, au sein de la praxis historique globale⁵⁴.

Avec *Pour une Esthétique de la réception*, Jauss dessine les contours d'un nouveau paradigme. Prendre en compte le milieu culturel d'un public devient une composante essentielle pour saisir le sens d'une œuvre qu'il lui donnera.

Ce concept d'« horizon d'attente » est pertinent pour répondre à notre problématique. Nous étudions deux publics de spectateurs aux « horizons d'attente » divergents, ce qui laisse sous-entendre que leur vision du théâtre sera différente. Bien que nous n'ayons pas de chiffres précis, une partie importante des journalistes a effectué des études de lettres dans les grandes institutions britanniques. La rémunération qu'ils perçoivent pour leur activité de critique leur assure également une relative aisance financière ainsi que des places de premier choix dans les théâtres de Londres.

A l'autre bout du spectre se situent les blogueurs. Leurs parcours académiques sont différents. Par manque de temps et d'argent, leur activité bénévole ne leur permet pas toujours d'assister aux grands succès joués sur la scène contemporaine londonienne. Pour la grande majorité, ces personnes ont un travail à côté. Surtout, leur désir d'offrir un nouveau regard sur le théâtre londonien les différencie des chroniqueurs des grands médias journalistiques. Dans un souci de représenter les marges de la culture théâtrale, ces blogueurs revendiquent leur mise en valeur des théâtres périphériques.

⁵³ JAUSS, *op. cit.*, p. 54.

⁵⁴*Ibid.*, p. 269.

Le concept d'« horizon d'attente », tel qu'il est défini par Hans-Robert Jauss, nous permet ainsi d'apporter un éclairage théorique à notre objet d'étude. Ces deux groupes de spectateurs aux horizons d'attente différents, c'est-à-dire avec des « attentes concrètes [...] intérêts, désirs, besoins et expériences tels qu'ils sont déterminés par la société et la classe à laquelle il[s] apparten[nent] aussi bien que par [leur] histoire individuelle⁵⁵ » proposeront sûrement deux manières différentes de concevoir le théâtre londonien.

I.A.2.3 Les études de la réception dans le théâtre

Contrairement aux études cinématographiques, les horizons d'attente du public théâtral ont été peu explorés jusqu'à présent. Probablement par manque de données directement exploitables, l'étude de la réception du théâtre reste donc largement incomplète. Une des premières à avoir analysé le rôle du spectateur d'un point de vue théorique est Susan Bennett, dans sa thèse publiée en 1988 à l'Université de McMaster au Canada. La publication tirée de sa thèse, *Theatre Audiences*(1990)⁵⁶, fait aujourd'hui référence dans ce champ d'études. La raison pour laquelle Bennett a commencé à explorer la réception du théâtre est similaire à celle qui nous a amenés à conduire notre étude. Dans la préface de sa thèse, elle explique comment l'émergence de nouvelles formes théâtrales moins conventionnelles attiraient déjà, dans les années 1980, des spectateurs d'un genre nouveau :

Beyond the lacunae of the theorists, my interest has been stimulated by the diverse practices of 'alternative' theatre. The development in the last thirty years of a flourishing alternative sector, offering drama outside, and often in opposition to, the conventional product of the "mainstream", a pre-scripted play performed in a recognized theatre space, has expanded the nature of theatre, and thus audience, open to study⁵⁷.

Ce que Bennett appelle « les théâtres alternatifs », c'est-à-dire les théâtres marginaux qui apparaissent dans les années 1960, fait écho à la culture théâtrale périphérique à laquelle les blogueurs s'intéressent. En examinant des pièces qui ont lieu dans les rues de Londres, dans les associations de quartier ou parfois même dans les bowlings, les critiques numériques offrent non seulement un autre regard sur le théâtre londonien, mais ils constituent également un nouveau public qui semble prêter attention à d'autres types de spectacles.

⁵⁵ JAUSS, *op. cit.*, p. 259.

⁵⁶Susan BENNETT, *Theatre Audiences*, Londres et New York, Routledge, 1990.

⁵⁷ Susan BENNETT, *The Role of the Theatre Audience*, Thèse de doctorat, Université de MacMaster, 1988, p. vi.

Bien que l'importance du rôle du spectateur était déjà prise en compte dans le théâtre grec, ce n'est que dans la seconde moitié du XX^e siècle que quelques études académiques ont commencé à théoriser la réception dans les études théâtrales. En 1968, dans *The Empty Space*, Peter Brook insistait sur la nécessité de la présence du spectateur pour que le théâtre advienne : « A man walks across this empty space whilst someone else is watching him, and this is all that is needed for an act of theatre to be engaged⁵⁸. » La même année, Jerzy Grotowski proposa dans *Towards A Poor Theatre* que le théâtre était précisément ce qui « a lieu entre le spectateur et l'acteur » (« *takes place between spectator and actor*⁵⁹ »). Quelques années plus tard, David Bain, dans *Actors and Audience* (1987)⁶⁰, examina en profondeur les liens entre le public et le théâtre grec. Enfin, plus récemment, l'essai d'Helen Freshwater *Theatre and Audience* (2009)⁶¹ offre un panorama des études de la réception dans le théâtre.

Bennett rappelle que c'est en réaction au théâtre naturaliste du début du XX^e siècle que l'intérêt porté au rôle du spectateur connut un succès grandissant. Jean Jullien, dramaturge, théoricien et critique français, résume explicitement la position du spectateur durant cette période dans *Le Théâtre vivant* (1892) en demandant à ses acteurs de jouer « comme s'ils étaient à la maison, en ignorant les émotions suscitées parmi le public » (« *as if at home, ignoring the emotions they arouse in the public*⁶². »). Le rôle du spectateur au début du XX^e siècle se caractérise alors par la passivité. Relégué au second plan, il n'est qu'un simple élément dans le décor théâtral. Les pratiques théâtrales qui suivirent ce mouvement prirent le contrepied de cette tendance. Le spectateur devint peu à peu un élément central au théâtre, comme dans les pratiques de Filippo Marinetti (1876 – 1944), fondateur du mouvement futuriste. En utilisant de la poudre à gratter ou à éternuer, en mettant de la colle sur certains des sièges du public ou en vendant le même ticket à deux spectateurs différents, Marinetti mit au centre des pratiques théâtrales l'importance du rôle du spectateur⁶³.

A partir du milieu du XX^e siècle, les travaux théoriques sur les études théâtrales et la réception prennent une véritable ampleur. Deux approches peuvent être observées : celles, bien moins nombreuses, qui reposent sur des analyses quantitatives, et celles qui adoptent un point de vue psychanalytique. Concernant la première approche, les travaux de Jim Davis et

⁵⁸ Peter BROOKE, *The Empty Space, A Book About the Theatre: Deadly, Holy, Rough, Immediate*, New York, Scribner, 1968, p.9.

⁵⁹ Jerzy GROTOWSKI, *Towards a Poor Theatre*, Londres, Methuen, 1968, p.32.

⁶⁰ David BAIN, *Actors and Audience: A Study of Asides and Related Conventions in Greek Drama*, Oxford, Oxford University Press, 1987.

⁶¹ Helen FRESHWATER, *Theatre and Audience*, Basingstoke, Palgrave Macmillan, 2009.

⁶² BENNETT, *op. cit.*, p. 6.

⁶³ *Ibid.*

de Victor Emeljanow sont particulièrement proches de notre projet⁶⁴. En se concentrant sur sept théâtres emblématiques de Londres de 1840 à 1880 (Surrey Theatre, Royal Victoria Theatre, Whitechapel Pavilion, Britannia Theatre, Sadler's Wells et le Queen's Theatre, Drury Lane Theatre), les deux auteurs donnent une analyse détaillée des différents types de spectateurs qui ont assisté aux représentations théâtrales de ces lieux à cette époque. Leur étude repose sur des calculs statistiques tirés d'un nombre de sources variées. Rapports de police, programmes de théâtre, autobiographies et journaux intimes, compte-rendus journalistiques, politiques managériales des théâtres de l'époque, parmi d'autres, sont finement passés en revue afin de cartographier les attentes des spectateurs du XIXe siècle.

L'étude de Dan Olsen propose également une approche quantitative de la réception du théâtre mais en adoptant une échelle différente⁶⁵. Afin de mieux comprendre ce qui attire l'attention des spectateurs dans une représentation théâtrale, les deux auteurs ont rédigé des questionnaires précis pour un même public qui assiste à huit pièces venant de quatre pays différents. En ayant recours à des méthodes inspirées de la sociologie, Cabral et Olsen parviennent ainsi à déterminer les différentes caractéristiques qui intéressent le public.

La seconde approche qui étudie la réception du théâtre repose sur des théories psychanalytiques. Anne Ubersfeld, dans *Le Plaisir du spectateur* (1980), utilise par exemple des concepts freudiens pour explorer la relation entre le public et la scène. Reprenant la théorie de Freud selon laquelle l'identité de l'enfant est conditionnée par un plaisir inatteignable, Ubersfeld dresse un parallèle entre le plaisir de la scène et celui du spectateur. L'expérience esthétique fonctionne selon une dynamique similaire : le caractère éphémère des représentations théâtrales et l'inconstance du désir du spectateur impliquent une frustration permanente :

The object flees from the eye and the touch of the one that desires it: not only the actor flees from us, but all the beauty shimmering on the surface of the stage... And the flight of our desire is no less frustrating than the flight of the object: the desire of the spectator travels from object to object⁶⁶.

Dans *The Conditions of Reception in the Theatre: Psychological and Psychoanalytical Approaches* (2001)⁶⁷, Patrice Pavis effectue un résumé de tous les outils issus de la

⁶⁴ Jim DAVIS, Victor Emeljanow, *Reflecting the Audience: London Theatregoing, 1840 – 1880*, Hertfordshire, University of Hertfordshire Press, 2001.

⁶⁵ Dan OLSEN, « A Quantitative Research Method in Reception Theory in International Theatre: Case Studies: 'Transit', Brasil 'Motherland', Jamaica 'Sandhom', 'FIF' and 'Visit to Turkey', Denmark, 'Vida de Los Perros' and 'El Meón', Bolivia, '5PSA – o Filho', Brazil. », *Caribbean Quarterly* [En ligne], no. 1/2, vol. 53, 2007, p. 181 – 93, URL : <http://www.jstor.org/stable/40654983>, consulté le 13 octobre 2021.

⁶⁶ Anne UBERSFELD, « The Pleasure of the Spectator », Trans. Pierre Bouillaguet et Charles Jose, *Modern Drama*, vol. 25, no. 1, 1982, p. 138.

psychanalyse pour examiner les réactions du spectateur. Enfin, plus récemment, en 2015, Maria Grazia Turri a terminé son doctorat à l'Université d'Exeter en proposant une thèse qui explore l'engagement émotionnel inconscient du spectateur⁶⁸.

Malgré ces exemples décrits, peu d'études sont aujourd'hui disponibles sur les publics des salles de théâtre, et encore moins sur les publics marginaux. Comme le note Ric Knowles dans *Reading the Material Theatre* (2004), « how audiences produce meaning in negotiation with the particular, local theatrical event, fully contextualised... has only rarely been analysed or modelled in any detail⁶⁹. » Concernant la critique théâtrale, une littérature conséquente a été rédigée, mais pas du point de vue des études de la réception. Plusieurs essais examinent le rôle du critique, son évolution au cours de l'histoire et les effets qu'il a pu provoquer sur une pièce de théâtre⁷⁰. Aucun en revanche, à notre connaissance, ne considère la critique théâtrale comme un public à part entière.

Surtout, très peu se sont encore intéressés à la critique numérique londonienne. Outre les articles qui sont parus en ligne dans les blogs ou dans les journaux britanniques, seulement deux essais ont été publiés à ce sujet. Le premier, *Theatre Criticism, Changing Landscape* (2016), offre un panorama de l'évolution de la critique au XXI^e siècle dans le monde⁷¹. Une série de chapitres présente tout à tour la critique allemande, russe, américaine, lettonienne, grecque et britannique. Un chapitre seulement expose les enjeux de la critique anglaise au XXI^e siècle. Le second essai qui porte sur cette thématique est celui de Megan Vaughan, blogueuse londonienne, dont la thèse réalisée à l'Université de Royal Holloway a été publiée récemment sous le titre de *Theatre Blogging, The Emergence of a Critical Culture* (2020)⁷². A nouveau, si ces deux essais révèlent des informations précieuses à propos de la blogosphère anglaise, aucun ne propose une analyse chiffrée de ce public. Surtout, aucun ne les envisage sous l'angle théorique de la réception.

⁶⁷ Patrice PAVIS, « The Conditions of Reception in the Theatre: Psychological and Psychoanalytical Approaches », *New Approaches to Theatre Studies and Performance Analysis: Papers Presented at the Colston Symposium, Bristol, 21 mars 1997*, Berlin, New York, Max Niemeyer Verlag, p. 13 – 30.

⁶⁸ Maria GRAZIA TURRI, *A Psychoanalytic Perspective on Theories of Spectator-Character and Actor Character Identification in the Theatre*, Thèse de doctorat, Université d'Exeter, 2015.

⁶⁹ Rick KNOWLES, *Reading the Material Theatre*, Cambridge, Cambridge University Press, 2004, p.17.

⁷⁰ Irving WARDLE, *Theatre Criticism*, Londres, Routledge, 1992.

Charles W. MEISTER, *Dramatic Criticism: A History*, Londres, McFarland and Company, 1985.

Noel CARROLL, *On Criticism*, Londres, Routledge, 2008.

⁷¹ Duška RADOSAVLJEVIC, *Theatre Criticism: Changing Landscapes*, Londres, Bloomsbury, 2016.

⁷² Meghan VAUGHAN, *Theatre Blogging: The Emergence of a Critical Culture*, Londres, Bloomsbury, 2020.

Transition

Cette deuxième sous-partie avait pour finalité de présenter les études de la réception, autre aspect théorique central à notre recherche. A partir des années 1970, ce champ met au jour l'importance du rôle du récepteur dans l'actualisation d'un objet culturel. C'est grâce au regard ou à la lecture du destinataire qu'une œuvre adviendra. Pour saisir le sens de cette dernière, il faut alors restituer l'« horizon d'attente » de son public, c'est-à-dire ses attentes esthétiques qui dépendent de son contexte social, politique et culturel. Cette conception historique de la littérature permet ainsi de reconnaître l'importance de la pluralité des points de vue.

Cette dialectique inversée du rapport de l'œuvre à son destinataire a été peu explorée au sein des études théâtrales. Excepté l'essai de Susan Bennett, *Theatre Audiences*, qui fait aujourd'hui référence dans ce domaine, peu d'études ont été menées pour analyser les publics des salles de théâtres. Deux tendances peuvent néanmoins être observées parmi ces travaux : les rares recherches qui reposent sur des analyses quantitatives, et celles qui choisissent une approche fondée sur la psychanalyse. Notre projet s'inscrit dans la première catégorie. A partir de la création d'un corpus de plus de 40 000 critiques théâtrales regroupant les écrits de deux groupes de spectateurs distincts, nous explorerons leur vision du théâtre londonien des années 2010 à 2020.

Les deux principaux aspects théoriques de notre recherche étant maintenant présentés, nous nous concentrerons dans la partie suivante sur le cœur de ce projet de thèse. A partir des années 2010, une nouvelle forme de réception émerge grâce à Internet. Les journalistes ne sont désormais plus les seuls à publier leurs critiques sur le théâtre londonien. Des dizaines de blogueurs commencent à mettre en ligne leurs propres critiques. Ces nouvelles voix marginales de l'espace numérique bousculent les conventions journalistiques. Leur identité se construit en opposition aux médias traditionnels, jugés élitistes par nombre de ces blogueurs. Ce sont ainsi deux publics aux horizons d'attente opposés qui se font face. Pour mieux comprendre en quoi cette nouvelle communauté propose un regard différent sur la scène contemporaine londonienne, nous aborderons sa naissance dans la partie qui suit.

I.B Redéfinition de la critique théâtrale au XXI^e siècle : enjeux littéraires

I.B.1 Critique et numérique : la blogosphère londonienne

I.B.1.1 L'émergence d'un second regard sur le théâtre londonien (1997 – 2010)

Les premières tentatives de critiques théâtrales britanniques commencent dans les années 1690, avec la revue *The Gentleman's Journal*, soit soixante-dix ans après l'avènement du premier journal britannique. En 1734, Aaron Hill et William Popple créent *The Prompter*, le premier journal consacré au théâtre anglais. C'est en 1770 que la critique théâtrale devient progressivement partie intégrante des journaux, et à partir du début du XIX^e siècle que l'activité de critique est reconnue comme profession à part entière⁷³. Le développement et le rayonnement de la presse écrite au XIX^e siècle permet ensuite à la critique de s'établir pleinement dans les journaux. Les grands noms de la presse écrite commencent à apparaître, *The Times* (1785), *The Observer* (1791), *The Guardian* (1821), pour n'en citer que quelques-uns, et laissent progressivement davantage d'espace à la critique théâtrale.

A la fin du XX^e siècle, l'arrivée d'Internet redéfinit le paysage de la critique théâtrale classique. De 1990 à 2010, l'apparition de blogs permet à des auteurs d'horizons différents de publier leurs critiques théâtrales en ligne. C'est à partir de 2010 que ce mouvement prend une ampleur véritable, notamment grâce à la popularité de plateformes comme WordPress.com, LiveJournal.com ou Blogger.com. L'activité de critique s'élargit, elle n'est désormais plus seulement réservée à une fraction de la population, mais à un ensemble d'individus. Une diversité de voix indépendantes commence à fleurir dans le paysage de la critique théâtrale britannique. Etudiants, professionnels du théâtre, mais aussi simples amateurs s'emparent de leur clavier pour rédiger une histoire du théâtre anglais différente de celle proposée par les grands journaux de la presse anglaise.

En Angleterre, au Canada, aux Etats-Unis mais aussi en Nouvelle-Zélande, ce mouvement s'étend progressivement. Tout spectateur passionné par le théâtre et sachant utiliser Internet peut désormais ouvrir son blog afin d'y publier ses propres analyses. En Angleterre, cette restructuration de la critique théâtrale se divise en trois phases. Nous nous

⁷³Mark FISHER, *How to Write about Theatre*, Londres, Bloomsbury, 2015, p. 12 – 16.

appuyons ici sur le chapitre intitulé « A Brief History of Online Criticism in England⁷⁴ », rédigé par Andrew Haydon, afin d'exposer les différentes étapes de l'émergence de cette communauté. Nous avons repris la chronologie qu'Haydon propose, tout en complétant ses arguments par nos propres analyses.

Première phase (1997 – 2007)

La première phase débute en 1997 avec la création de deux sites encore en ligne aujourd'hui (2021) : BritishTheatreGuide.info⁷⁵ et WhatsOnStage.com⁷⁶. BritishTheatreGuide.info fut mis en place par Peter Lathan, alors enseignant dans le nord-est de l'Angleterre à Sunderland. Afin d'alimenter la base de données de la New York-based General Internet, il réunit plusieurs matériaux relatifs aux études théâtrales pour d'autres enseignants. Ce sont ses premiers matériaux qui donneront naissance au premier site officiel de critique théâtrale en ligne. A l'inverse, WhatsOnStage.com fut impulsé par le magazine *EMAP* en décembre 1996, et plus particulièrement par Terry Paddock qui fut nommée éditrice à mi-temps de cette nouvelle plateforme en ligne.

Bien que l'aspect numérique de ces deux sites fut novateur pour la fin des années 1990, leur contenu restait néanmoins similaire à ce que proposait la presse écrite à cette époque. Tous deux continuaient de respecter les normes éditoriales proposées par les magazines au format papier : un éditeur était à la tête de ces plateformes, et plusieurs critiques professionnels publiaient régulièrement du contenu.

Deuxième phase (2007 – 2010)

C'est à partir de 2006 qu'une deuxième vague de blogs émerge et propose des critiques d'un genre différent. Deux blogueurs anonymes, Andrew et Phil, fondent leur site intitulé WestEndWhingers.wordpress.com⁷⁷ (*Les Pleurnichards du West End*). Leur première critique qui porte sur la pièce *Resurrection Blues* (2002) d'Arthur Miller, jouée au théâtre The Old Vic en mars 2006, donne le ton de cette nouvelle plateforme qui entend réécrire les codes de la critique classique :

⁷⁴ Andrew HAYDON, « A Brief History of Online Criticism in England », *Theatre Criticism: Changing Landscapes*, Londres, Bloomsbury, 2016, p. 135 – 151.

⁷⁵<https://www.britishtheatreguide.info/>, [En ligne], consulté le 22 novembre 2021.

⁷⁶<https://www.whatsonstage.com/>, [En ligne], consulté le 22 novembre 2021.

⁷⁷<https://westendwhingers.wordpress.com/>, [En ligne], consulté le 22 novembre 2021.

It was unwatchable. Everyone was terrible. The set was terrible. The costumes were terrible. It cost a fortune. Why did we return after the interval? Well, we've always wanted to boo at the end of a show, but we've never seen anything bad enough to warrant it. This was our One Big Chance and Phil seized it with aplomb⁷⁸.

Rédigés avec un humour satirique mordant, leurs écrits marquent un tournant dans la critique théâtrale traditionnelle. Phil et Andrew se moquent ouvertement de certaines pièces et du style de la critique journalistique. Leur analyse de la pièce *Death and The King's Horseman* (1975) par exemple se présente sous la forme d'un mail frauduleux signé par la « No. 1 Nigerian Scam Agency⁷⁹. »

Leur système de notation ridiculise également les schémas de la critique classique. A la place des traditionnelles étoiles qui sont attribuées aux spectacles, Phil et Andrew leur discernent des « puddings du Yorkshire », des verres de whiskey ou le plus souvent des verres de vin. Comme le rappelle le jeu de mots dans la présentation de leur blog, leur rôle consiste principalement à aider le spectateur à choisir entre le « Merlot et le Marlowe (« Andrew is on sabbatical but Phil is soldiering on to help you decide between the Merlot and the Marlowe and generally putting London's West End theatre to rights. »). Ces verres de vin sont accompagnés de commentaires caustiques. Voici l'illustration de leur système de notation :



Fig. 7 : Système de notation du site
WestEndWhingers.wordpress.com

⁷⁸ Andrew et Phil, « Resurrection Blues : Well it Made us Cross », *The West End Whingers* [En ligne], publié le 8 mars 2006, URL : <https://westendwhingers.wordpress.com/2006/03/08/resurrection-blues-well-it-made-us-cross/>, consulté le 22 novembre 2021.

⁷⁹ Andrew et Phil, « Death and the King's Horseman, National Theatre », *The West End Whingers* [En ligne], publié le 8 avril 2009, URL <https://westendwhingers.wordpress.com/2009/04/08/review-death-and-the-kings-horseman-national-theatre/>, [En ligne], consulté le 22 novembre 2021.

Ainsi, pour deux verres de vin sur cinq par exemple, le spectacle est qualifié de « Légèrement bouchonné ou vinaigré (*Slightly corked or vinegar*). » Phil et Andrew ont également créé un « Temple de la renommée des spectacles qui ont obtenu 0 étoile » (*Zero Stars Hall of Fame*⁸⁰). A l'inverse des discours policés publiés dans les grands organes de presse, Phil et Andrew tirent parti de l'absence de règles établies dans l'espace numérique pour renverser et réécrire les codes de la critique journalistique.

WestEndWhingers.wordpress.com est ainsi le premier blog à affirmer un style qui va à l'encontre de la critique classique. Plusieurs blogs sont également créés cette même année. Le dramaturge anglais Chris Goode commença son blog beescope.blogspot.com⁸¹, Matt Trueman « Carousel of Fantasies » (qui n'est plus en ligne aujourd'hui), Andy Field TheArcadesProject.blogspot.com⁸² et Fin Kennedy FinKennedy.blogspot.com⁸³. Réalisés pour la majorité grâce à des plateformes comme Blogger.com, Blogspot.com ou Wordpress.com, ces blogs sont administrés par l'auteur lui-même et proposent des contenus variés. La critique telle qu'elle est ordinairement conçue est ici aussi repensée. Il ne s'agit plus d'une succession de textes courts avec une organisation stricte et rigoureuse des arguments, mais plutôt d'un ensemble de discussions et d'interviews sur le théâtre londonien et sur ses périphéries.

L'émergence de cette nouvelle communauté numérique donne naissance à un grand nombre de polémiques sur le statut de la critique en tant qu'objet littéraire. En 2007, le titre provocateur de l'essai de Rónán McDonalds incarne le point de vue de nombreux auteurs qui rédigent dans les grands journaux publics. *The Death of the Critic*⁸⁴ remet en question la montée des *Cultural Studies* et leur décentralisation du canon qui, selon McDonalds, aboutissent à un relativisme culturel. Le débat dans la sphère théâtrale se cristallise autour de deux positions : les critiques professionnels d'un côté, qui sont rémunérés pour leur activité et qui publient dans les grands noms de la presse anglaise (*The Times*, *The Guardian*, *The Independent*, etc.), et les critiques amateurs de l'autre, qui écrivent par passion et bénévolement dans leur blog.

⁸⁰ <https://westendwhingers.wordpress.com/zero-stars-hall-of-fame/>, [En ligne], consulté le 22 novembre 2021.

⁸¹ <http://beescope.blogspot.com/>, [En ligne], consulté le 22 novembre 2021.

⁸² <http://thearcadesproject.blogspot.com/>, [En ligne], consulté le 22 novembre 2021.

⁸³ Accessible sur autorisation seulement.

⁸⁴ Ronan MCDONALD, *The Death of the Critic*, Londres, Bloomsbury, 2007.

I.B.1.2 Naissance d'une véritable culture numérique (2010 à aujourd'hui)

La troisième vague de blogs qui émerge en 2010 est celle à laquelle nous nous intéressons dans ce travail. Elle s'inscrit dans la continuité du mouvement que nous venons de décrire, excepté qu'elle est d'une plus grande ampleur. La création de deux plateformes numériques, ExeuntMagazine.com et AYoungerTheatre.com, initient ce troisième mouvement. La première voit le jour à la fin de l'année 2010 grâce à trois personnes : Daniel B. Yates, critique et écrivain, Natasha Tripney, aujourd'hui critique dans le journal *The Stage*, et Diana Damian, alors doctorante à l'Université de Royal Holloway. L'identité d'ExeuntMagazine.com se construit explicitement en opposition aux discours de la presse écrite :

Exeunt believes in making beautifully written, experimental, fierce and longform writing about theatre available for free. We give critics the freedom to expand beyond the tight deadlines and brutal word counts of more established media outlets. To give theatre and performance the consideration and thought it deserves.⁸⁵

AyoungerTheatre.com affirme des positions similaires. Destinée à des individus âgés entre 16 et 35 ans seulement, cette plateforme encourage ouvertement la participation de jeunes auteurs qui font partie des minorités⁸⁶ et qui s'intéressent aux scènes périphériques du théâtre londonien :

At the heart of our work is a platform for those who are often unheard. We champion the emerging generation. We celebrate excellence and pride ourselves on professionalism, while also leaving room for risk, failure and learning.⁸⁷

En 2011 est créée une autre plateforme qui présente des caractéristiques similaires à celles que nous venons de présenter. EverythingTheatre.com se veut être un espace dans lequel l'activité de critique n'est plus considérée comme un métier réservé à une fraction privilégiée de la population, mais comme une activité que chaque individu peut pratiquer. Ses contributeurs sont autant de passionnés qui travaillent dans le domaine du droit, de

⁸⁵<http://exeuntmagazine.com/about-exeunt/>, [En ligne], consulté le 24 novembre 2021.

⁸⁶<https://www.ayoungertheatre.com/jobs/>, [En ligne], consulté le 24 novembre 2021.

Voici le texte rédigé dans leur section « Join the Team, Work with us » :

« We are also keen to ensure that our writing team is representative of the society that we live in and so, we particularly welcome applications from people of colour, people with disabilities and the LGBTQ+ community. Please note that if for any reason you are unable to send an application using the forms below, we are very happy to receive video applications too. We are expanding the company past 'traditional' forms of criticism but more than anything, wish to be part of a more open society that doesn't discriminate against those with disabilities or learning difficulties. »

⁸⁷<https://www.ayoungertheatre.com/about/>, [En ligne], consulté le 24 novembre 2021.

l'événementiel, du marketing, de l'éducation que du théâtre. Les quatre principes mis en valeur par la couleur rouge dans leur page de présentation soulignent leur désir d'étendre la critique au plus grand nombre

OUR AIMS:

1. To share our **infectious enthusiasm** for theatre with the person on the street.
2. To provide **simple, unpretentious and easy-to-understand** coverage of the theatre scene which does not assume any level of theatrical knowledge.
3. To remain **upbeat and constructive** in everything we do; we're certainly not here to put people down!
4. To provide **all theatre companies** – both big and small – with an even chance of having their work reviewed and shared with as wide an audience as possible.⁸⁸

D'autres blogs administrés par une seule personne voient également le jour et portent les marques de ce désir de distinction face à la critique journalistique. Nick, éditeur de PartiallyObstructed-View.Blogspot.com, écrit dans son profil :

Like most people, I usually pay for my tickets (so I can tell you what the view's like from the cheap seats - theatre shouldn't just be for people who can mortgage their house for a premium ticket) and believe that entitles me to an opinion.⁸⁹

Avant d'ajouter en note de bas de page « not everyone agrees with me about this. » Maryam Philpott a, quant à elle, intitulé son blog « Cultural Capital », une référence explicite à l'expression employée par Pierre Bourdieu pour désigner l'ensemble des ressources culturelles d'une personne qui sont inégalement réparties. Michael Davis, créateur du blog Breaking-The-Fourth-Wall.com, est décrit comme un « passionné par les théâtres marginaux et par faire de la publicité aux spectacles qui ne reçoivent pas nécessairement l'attention du grand public⁹⁰. »

Que cela soient des plateformes collectives comme ExeuntMagazine.com, AyoungerTheatre.com ou EverythingTheatre.com, ou des blogs administrés par une seule personne, tous prônent l'indépendance de leur écriture et la nécessité de rendre audibles des voix qui ne sont pas assez entendues dans l'espace théâtral public. La naissance de ces blogs se crée ainsi en opposition à la culture journalistique et constitue une nouvelle forme de réception avec des spectateurs aux horizons d'attente opposés. La frise chronologique

⁸⁸ Description de la plateforme donnée en 2020. Les éditeurs de la plateforme EveryThingTheatre.com ont depuis réécrit leur page de présentation.

⁸⁹<http://partially-obstructed-view.blogspot.com/p/about-me.html> [En ligne], consulté le 24 novembre 2021.

⁹⁰<https://mytheatremates.com/about-us/authors/> [En ligne], consulté le 24 novembre 2021.

présentée à la page suivante (Figure 8) résume les différentes étapes de l'émergence de cette communauté numérique.

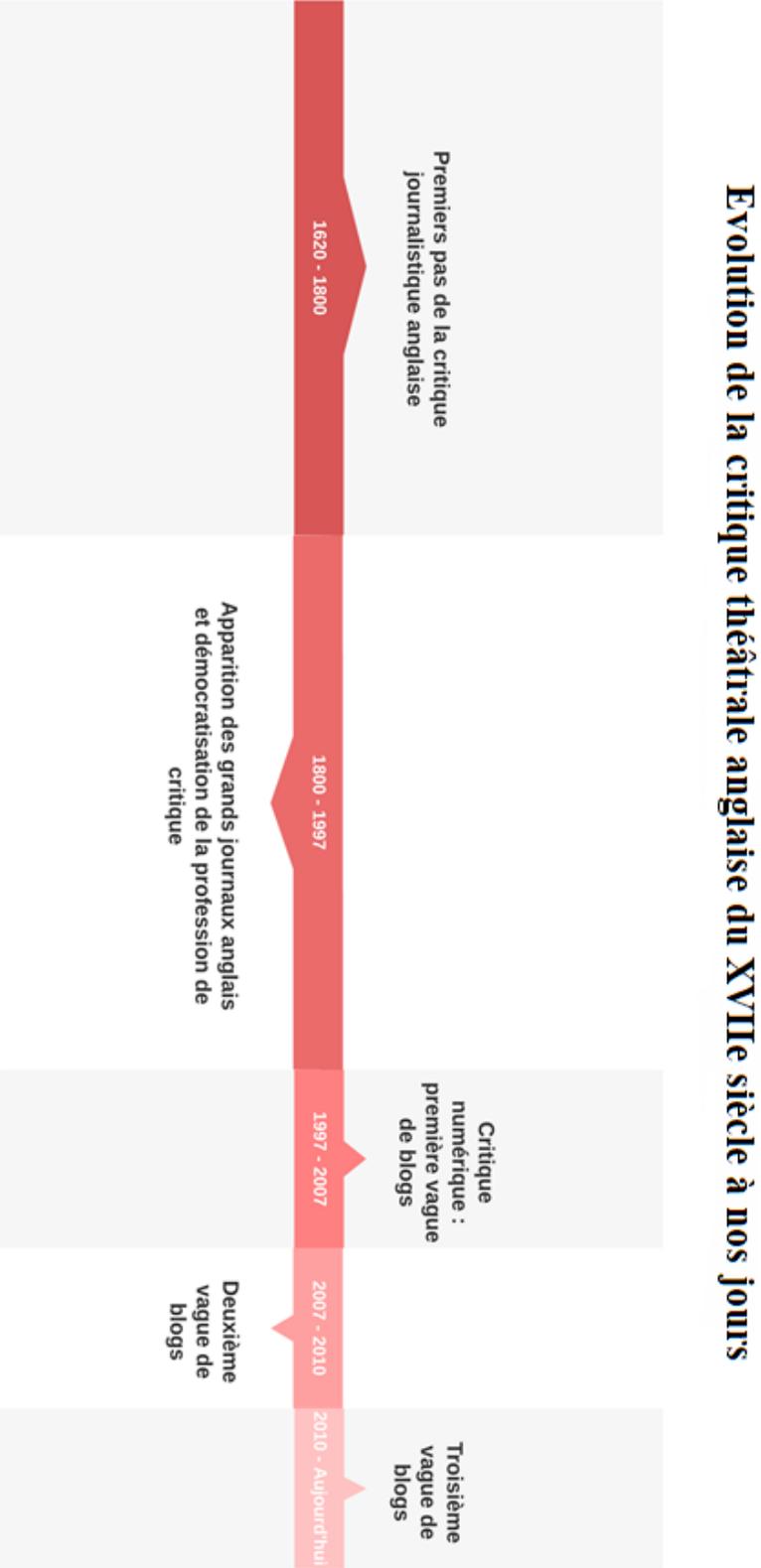


Fig. 8 : Frise chronologique de l'évolution de la critique théâtrale anglaise

I.B.1.3 Centre versus périphéries de la critique théâtrale londonienne

L'émergence de cette nouvelle communauté numérique met en lumière l'opposition « centre versus périphéries », ou « canon versus marges. » Le nom de certains blogs porte explicitement les marques de son appartenance aux marges de la culture théâtrale londonienne. Créé en 2014, NotExactlyBillington.blogspot.com affirme par exemple son désir de s'écarter des écrits d'une des plus grandes figures de la critique journalistique du XXI^e siècle. Nous pouvons également citer ViewfromtheCheapSeat.com et PartiallyObstructedView.blogspot.com dont les noms mettent en évidence les différences de budget entre un critique professionnel et un critique amateur. D'autres, comme Criticsof Colour.tumblr.com, s'attachent à donner uniquement la parole à une minorité de la population qui est sous-représentée dans le canon de la critique théâtrale anglaise.

L'opposition conceptuelle « centre / périphérie » émerge dans les années 1900 en géographie et en économie⁹¹. C'est dans les années 1960 que ces termes deviennent populaires avec la décolonisation et la remise en question de l'impérialisme. Le terme « canon » quant à lui vient du grec *kanôn* qui fait premièrement référence à un roseau, puis à un instrument de mesure fabriqué à partir du roseau. De manière plus générale, le canon renvoie à la règle de conduite, à la norme, au modèle. En astrologie ou en chronologie, il désigne une liste ou un catalogue. Les tables astrologiques de Ptolémée au II^e siècle avant Jésus-Christ., appelées « canons », permettaient par exemple de déterminer des points de référence pour le changement des saisons.

Le canon désigne également « le centre » autour duquel se forment les périphéries, ou les marges. Comme le rappelle le programme de la journée d'études qui a eu lieu en 2013 sur les littératures périphériques, « le centre domine par son prestige, la masse de son lectorat, la densité de ses instances de reconnaissance, sa capacité à organiser la production littéraire, et impose ainsi la norme⁹². » La périphérie se construit en opposition à ce centre. Elle est définie par ses relations internes et par ses relations avec l'extérieur. Elle se « définit par un éloignement qui peut être spatial, social, national, idéologique, sociolinguistique, intellectuel ou politique⁹³. »

⁹¹ Nelly BLANCHARD, Thomas MANNAIG, *Des littératures périphériques*, Rennes, Presses universitaires de Rennes, 2014, p. 11.

⁹² « Connexions et ruptures dans les littératures périphériques », Appel à contribution, *Calenda*, Publié le mardi 10 juillet 2012, URL : <https://calenda.org/209224> consulté le 19 septembre 2022.

⁹³*Ibid.*

Cette conception binaire témoigne surtout des rapports de force qu'il existe entre ces deux communautés. Le centre domine les périphéries par la culture « légitime » qu'il impose. En se définissant eux-mêmes comme une minorité et en prônant l'inclusion d'une diversité culturelle, les blogueurs entendent ainsi déconstruire la norme construite par la critique journalistique. L'objet de cette recherche consiste précisément à mesurer l'écart entre le centre de la critique journalistique et ses périphéries.

Transition

Nous venons de présenter les caractéristiques principales de la seconde forme de réception qui apparaît au tournant du XXI^e siècle. Trois mouvements peuvent être identifiés dans l'émergence de cette communauté numérique. Le premier apparaît au début des années 1990. Bien que l'apparition de deux nouvelles plateformes numériques marque le début d'une nouvelle ère dans la critique théâtrale, leur contenu reste néanmoins très proche des médias journalistiques traditionnels. C'est à partir de 2007 qu'un réel changement s'opère. Les blogs reflètent alors davantage la personnalité de leur éditeur. Leurs contenus sont plus longs et de nature hybride. Enfin, c'est en 2010 que ce phénomène prend une ampleur véritable.

Des dizaines de plateformes et de blogs personnels voient le jour. Ces nouveaux espaces numériques se construisent en réaction au journalisme. Il s'agit pour ces spectateurs de faire entendre leurs voix et de rendre visibles les marges de la culture théâtrale londonienne, encore trop discrètes dans l'espace public selon eux. La naissance de cette communauté remet en question la légitimité des grands noms de la critique théâtrale anglaise et de la culture dominante qu'ils mettent à l'honneur. L'opposition entre les journalistes et les blogueurs fait ainsi écho à la dichotomie « canon versus périphéries », ou « centre versus marges ».

Ce clivage est-il cependant si net ? Les journalistes et les blogueurs prônent-ils deux types de culture théâtrale différente ? La dernière sous-partie de ce chapitre aura pour finalité de présenter la méthodologie que nous avons choisie afin de répondre à ces questions qui constituent la problématique de notre recherche. Comme les parties précédentes, elle sera divisée en deux sous-parties. Dans la première, nous nous focaliserons sur l'aspect numérique de ce travail. Nous avons adopté une approche computationnelle pour explorer la restructuration de la critique londonienne des années 2010 à 2020. Nous aborderons ainsi les enjeux d'une telle démarche située à l'intersection de l'informatique et des lettres. Dans la seconde sous-partie, nous exposerons les caractéristiques principales du corpus que nous avons créé ainsi que ses limites.

I.B.2 Positionnement de notre projet

I.B.2.1 Le mariage des études théâtrales et de l'informatique

Notre recherche se situe dans la première catégorie de projets que Bardiot a identifiée et que nous avons présentée au début de ce chapitre⁹⁴, c'est-à-dire dans la catégorie qui porte sur l'analyse de l'histoire théâtrale. Il s'agit, à partir de la construction d'une base de données réunissant blogs et journaux, de cartographier une partie de l'histoire des discours sur le théâtre anglais grâce à l'utilisation d'outils techniques. Comme Visualizing Broadway et Ibsen Stage, nous nous intéressons à la manière dont l'histoire théâtrale se construit à travers la circulation de ses discours. Comme dans ces deux projets, nos analyses reposent sur ce que Franco Moretti nomme le *distant reading*, ou « la lecture à distance » :

The reality of the text undergoes a process of deliberate reduction and abstraction. 'Distant reading', I have once called this type of approach; where distance is however not an obstacle, but a *specific form of knowledge*: fewer elements, hence a sharper sense of their overall interconnection. Shapes, relations, structures. Forms. Models⁹⁵.

La lecture distante permet de rendre les thématiques ou les motifs récurrents dans un corpus apparents. Les algorithmes nous aideront ainsi à révéler les points de connexion entre les journalistes et les blogueurs, ou entre le centre de la critique théâtrale et ses périphéries. Loin de remplacer l'humain cependant, les traitements informatiques nous permettront d'organiser, de trier et d'analyser cette masse de données. La lecture distante ne remplace pas la lecture proche. Comme l'écrit Bardiot, il ne s'agit pas de « sacrifier les méthodologies classiques sur l'autel du progrès informatique⁹⁶ ». Au contraire, il se forme une « boucle rétroactive⁹⁷ » entre l'utilisation des algorithmes et l'étude du texte telle qu'elle est traditionnellement conçue en littérature. Tandis que l'informatique nous permet d'observer des tendances générales dans notre corpus, les études de textes et les lectures théoriques dans le domaine des humanités nous apportent des éléments d'explications plus nuancés.

C'est précisément cette dialectique entre technicité et littérature qui est au centre de ce doctorat. Théories littéraires et statistiques sont associées afin de modéliser une partie de l'histoire théâtrale pour apporter des réponses chiffrées à des questions littéraires. Les deux

⁹⁴Voir p. 31 – 39.

⁹⁵ Franco MORETTI, *Graphs, Maps, Trees, Abstract Models for Literary History*, Londres, Verso, 2005, p.1.

⁹⁶Bardiot, *op. cit.*, p. 5.

⁹⁷*Ibid.*

disciplines sont dépendantes l'une de l'autre. S'il nous était impossible d'examiner plus de 40 000 critiques littéraires sans la machine, les résultats ne produiraient de sens que grâce aux théories littéraires. Ce sont donc des allers-retours constants entre l'informatique et les lettres, une dynamique centrale aux humanités numériques, qui nous ont permis de construire ce projet de thèse. A l'inverse de Vizualizing Broadway et d'Ibsen Stage en revanche, notre corpus n'était ni complètement défini, ni directement exploitable par la machine. Notre projet s'apparente ainsi à d'autres initiatives qui ont uniquement pour objectif de préparer et de partager des corpus à l'ensemble de la communauté académique.

I.B.2.2 Définition de la méthode

La méthode que nous avons adoptée pour mener cette recherche repose sur une succession d'étapes propre aux projets en science des données. L'analyse d'un grand nombre d'informations requiert plusieurs protocoles distincts afin de garantir le bon déroulement des expériences. En anglais, ce protocole à suivre est appelé « Data science life cycle » par les ingénieurs, ou « Le cycle de vie des données » en français. Nous avons suivi ce modèle en l'adaptant à nos données littéraires. Il comporte sept catégories que le schéma ci-dessous illustre :

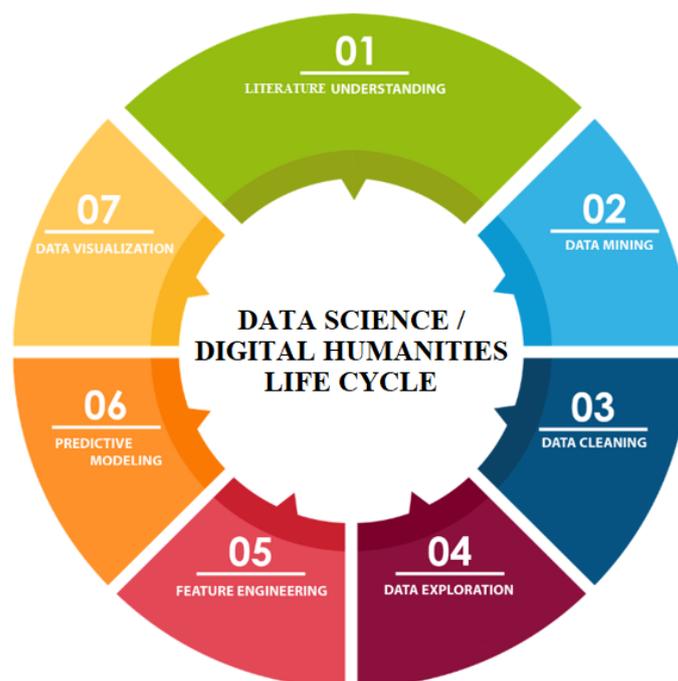


Fig. 8 : Représentation du cycle de la vie des données.

La première, « **Literature Understanding** » correspond à la problématique de la recherche. Pour les ingénieurs, cette catégorie s'intitule « **Business Understanding** ». Il s'agit ici de délimiter le sujet et de comprendre ce que l'on veut démontrer grâce aux outils numériques. Que cela concerne une question littéraire (« Peut-on observer des divergences d'un corpus à l'autre ? »), comme dans les humanités numériques, ou un besoin commercial (« Les utilisateurs ont-ils aimé ce produit ? »), comme dans une entreprise, la démarche reste la même. La question doit ainsi correspondre à un ou à plusieurs domaines techniques. Dans notre cas, nous avons posé trois questions spécifiques afin de mieux comprendre la relation entre le canon de la critique théâtrale londonienne et ses périphéries. La première porte sur le style littéraire des critiques, la deuxième sur leurs préférences esthétiques et la troisième sur leur répartition dans Londres. A ces trois questions correspondent trois champs techniques : la stylistique computationnelle, l'analyse d'opinion et les Systèmes d'Information Géographique.

La deuxième étape, le « **Data Mining** », consiste à récupérer les données nécessaires pour répondre à la problématique posée. Pour construire notre corpus, nous nous sommes appuyés sur deux types de sources. La première porte sur la critique journalistique et constitue notre premier sous-corpus. Nous avons réutilisé les archives numérisées du journal *Theatre Record* qui republie toutes les critiques parues dans les grands journaux de la presse anglaise depuis 1981. La seconde nous a permis de créer notre corpus témoin, c'est-à-dire celui fondé sur la critique numérique. Grâce à une technique appelée le *web scraping*, nous avons pu extraire automatiquement le contenu des 28 blogs les plus populaires de la toile.

La troisième étape, le « **Data Cleaning** », a pour finalité de préparer les données aux traitements algorithmiques. Cette étape chronophage implique à la fois un travail automatique et manuel. Pour analyser une masse d'informations à l'aide d'outils techniques, il faut que les données soient organisées selon une certaine structure. Il faut également retirer tous les éléments susceptibles de fausser les résultats. Un corpus qui comporte beaucoup d'erreurs, qu'elles soient de langue ou de structure, induira nécessairement un biais dans les traitements algorithmiques. Cette partie représente plus d'un an de travail dans le cadre de notre projet.

Une fois le corpus prêt, nous pouvons commencer à analyser son contenu. Cette quatrième étape est intitulée la « **Data Exploration** » dans le « cycle de vie des données. » Le terme « exploration » est justement choisi. Dans un premier temps, nous établissons un état de l'art des différentes méthodes informatiques qui nous permettront d'aborder notre problématique. Il s'agit ensuite d'essayer plusieurs modèles et de comparer leurs performances. Travailler en informatique requiert de nombreux essais avant de parvenir à un

résultat satisfaisant. Certaines techniques se révèlent parfois insuffisantes pour analyser un problème donné, et il faut alors reconsidérer le choix des outils.

La nécessité de retravailler sur les modèles, de les affiner et de les adapter en fonction de nos données et des résultats escomptés correspond à la cinquième étape qui est appelée le « **Feature Engineering**. » Cette partie consiste à améliorer les modèles informatiques pour obtenir des scores de performance plus précis. C'est probablement cette tâche qui requiert le plus de compétences techniques puisqu'il faut comprendre comment l'algorithme fonctionne pour l'améliorer.

La sixième étape est facultative et dépend du type de l'expérience menée. Le « **Predictive Modeling** », ou le « Machine Learning », correspond à l'automatisation d'une tâche. Une fois que le ou les modèle(s) ont été choisis et paramétrés (ce qui correspond à l'étape précédente), ils peuvent être implémentés afin d'analyser le corpus dans son entièreté.

La dernière étape, la « **Data Visualization** », permet de mettre en valeur les résultats obtenus. Bien que cette partie puisse sembler être la plus simple, elle fait référence à tout un domaine particulier de l'informatique qui nécessite une réelle réflexion afin de rendre les résultats pertinents d'un point de vue visuel. Le choix du type de graphes, les éléments que l'on décide de montrer ou non, l'utilisation des couleurs et des échelles de valeur sont autant de caractéristiques différentes auxquelles le chercheur doit réfléchir en amont.

Le caractère cyclique de cette méthode implique des allers-retours constants entre ses différentes étapes. Nous pouvons avoir atteint l'étape 7, c'est-à-dire la « Data Visualization », et nous rendre compte que le nettoyage des données n'est pas suffisant, ce qui implique de revenir à l'étape 2, soit au « Data Cleaning. » C'est une composante essentielle à de nombreux projets en informatique et qui ajoute beaucoup de temps aux expériences : le va-et-vient récurrent entre les différentes étapes techniques.

I.B.2.3 Présentation du corpus

Sous-corpus I : la critique journalistique

La mise en place de cette méthodologie fut nécessaire pour construire notre base de données. Pour analyser le théâtre londonien à travers le regard de deux communautés divergentes, nous avons rassemblé et organisé plus de 40 000 critiques littéraires. Ces dernières sont structurées en deux sous-corpus. Le premier, qui représente le centre la critique

théâtrale londonienne et qui est fondé sur les journaux, a été créé à partir de la base de données *TheatreRecord*⁹⁸. *Theatre Record* est originellement un magazine papier édité par Ian Herbert qui fut critique dans le journal *The Stage* et co-éditeur de sa section internationale pendant de nombreuses années. Publié toutes les deux semaines en Angleterre depuis janvier 1981, il retranscrit une grande partie des critiques théâtrales parues dans les journaux portant sur la scène londonienne et régionale anglaises.

C'est l'interruption momentanée du magazine *Plays and Players* qui poussa Herbert à créer *Theatre Record*. *Plays and Players* était à l'époque le seul magazine qui retranscrivait une grande partie des informations liées aux spectacles anglais. Afin de pallier la disparition de ce magazine, Herbert demanda la permission à tous les critiques journalistiques de réunir et de retranscrire leurs publications. En février 1981, le premier numéro de *London Theatre Record* vit ainsi le jour. Douze spectacles furent mentionnés. Seuls le *Times* et le *Sunday Times* n'y figuraient pas encore. Il fallut attendre les années 1990 pour que ces derniers soient inclus⁹⁹.

La revue fut ensuite enrichie par d'autres rubriques. Les pièces jouées dans les régions furent progressivement ajoutées au journal, ainsi qu'une section sur le festival d'Edinburgh et plusieurs photographies en noir et blanc sur les représentations théâtrales. A partir de 1991, *London Theatre Record* devint *Theatre Record*, et plusieurs CD-ROM réunissant l'intégrale de ses archives furent mis à disposition de ses utilisateurs afin de faciliter l'accès à ses multiples ressources¹⁰⁰.

A partir de janvier 2019, la version papier devint disponible en ligne grâce à Julian Oddy. Tous les journaux publiés depuis 1981 furent numérisés en format PDF et mis à disposition des utilisateurs grâce à une inscription payante à leur plateforme. Bien qu'aujourd'hui une grande partie de ces critiques soit disponible en ligne sur les sites web des journaux, il est néanmoins difficile d'avoir accès soit aux archives, soit à la totalité des articles. *Theatre Record* est ainsi la première et la seule base de données qui offre un tel vivier d'informations portant sur la scène contemporaine anglaise.

L'originalité de cette base de données nous a conduits à approfondir le sujet de la critique théâtrale londonienne. Non seulement sa récente numérisation garantissait des résultats novateurs, mais son contenu était également particulièrement intéressant dans le

⁹⁸<https://www.theatrecord.com/>, consulté le 12 janvier 2012.

⁹⁹ Nick SMURTHWAITE, « How Theatre Record's Survival is a Beacon of Hope while Arts Critics are Axed », *The Stage* [En ligne], publié le 23 juin 2017, URL : <https://www.thestage.co.uk/obituaries--archive/archive/how-theatre-records-survival-is-a-beacon-of-hope-while-arts-critics-are-axed>, consulté le 12 janvier 2022.

¹⁰⁰*Ibid.*

contexte des humanités numériques théâtrales et des études de la réception. Pour des raisons liées aux droits d'auteurs, peu de corpus sont disponibles pour l'analyse du théâtre contemporain. De plus, *Theatre Record* nous permettait de mettre en regard une autre forme de réception très peu analysée jusqu'ici d'un point de vue théorique, et encore moins d'un point de vue computationnel. En construisant un second corpus sur la critique numérique qui émerge à partir des années 2010, *Theatre Record* nous permettait ainsi d'examiner le théâtre ultra contemporain londonien à travers le regard de deux communautés de spectateurs différents.

Sous-corpus II : la critique numérique

Notre second corpus, qui repose sur les marges de la critique théâtrale, est fondé sur les 28 blogs anglais les plus influents d'Internet en 2020. Le terme « blog » est le diminutif de « weblog », qui est la contraction des termes « Web » et « Log », qui signifie « connexion. ». « De façon très synthétique, un blog est un type de site composé essentiellement d'actualités (ou billets) publiées au fil de l'eau et apparaissant selon un ordre chronologique (les plus récentes en haut de page), le plus souvent enrichies de liens hypertextes externes¹⁰¹. » C'est à partir des années 2000 que ce format devint populaire, en particulier grâce à l'apparition de la plateforme Blogger qui permit à de nombreux individus de publier du contenu en ligne sans avoir de connaissances en programmation informatique.

Comme nous l'avons expliqué au cours de ce chapitre, nous nous sommes intéressés à la troisième vague de blogs, c'est-à-dire aux sites qui apparaissent à partir des années 2010. Aucune étude n'ayant encore analysé la blogosphère dans sa totalité, nous avons dû effectuer de longues recherches pour obtenir un panorama des différents blogueurs qui écrivaient sur le théâtre anglais. Pour garantir une certaine partialité dans le choix des blogs, nous nous sommes appuyés sur deux sites qui référençaient les plateformes les plus populaires de la toile en 2020. Le premier concerne le classement des 10 blogs anglais les plus influents de l'année 2020 établi par Vuelio.com. Le second, MyTheatreMates.com, répertorie une cinquantaine de blogs britanniques. MyTheatreMates.com fut créé en 2014 par les critiques Mark Shenton et Terri Paddock. Après plus de dix ans en tant que critiques professionnels au *Sunday Express* pour Mark Shenton et dix-sept ans à *WhatsOnStage* pour Terri Paddock, leurs postes furent supprimés en décembre 2013. A la suite de ce licenciement, Shenton et Paddock décidèrent de

¹⁰¹Cyril FIEVET, Emily TURRETTINI, *Blog Story*, Paris, Eyrolles, 2004.

mettre en place une plateforme gratuite sur laquelle d'autres auteurs indépendants non rémunérés pouvaient publier leurs critiques. Tous les blogueurs qui contribuent à MyTheatreMates.com respectent ces quatre critères :

- Ils ont leur propre site web.
- Ils publient un contenu relatif au théâtre au moins une fois toutes les deux semaines.
- Ils sont en mesure de fournir trois références professionnelles artistiques (des artistes interviewés, des places de théâtre offertes par un producteur afin d'écrire la critique d'un spectacle, etc.).
- Ils ont un compte Twitter et sont actifs dessus.

Au moment de la création de ce second corpus (septembre 2020 à décembre 2021), 52 auteurs étaient membres de MyTheatreMates.com. Nous avons choisi les blogs qui présentaient un nombre de critiques important (au moins 100) ainsi que ceux qui se focalisaient sur la scène londonienne principalement. Enfin, les blogs que nous avons inclus dans notre second sous-corpus peuvent être catégorisés selon deux types. Le premier concerne les plateformes qui regroupent de multiples auteurs, tandis que le second porte sur les blogs qui sont administrés par une seule personne. Les 28 sites web que nous avons réunis partagent néanmoins trois caractéristiques communes :

- Ils ne possèdent pas de version papier
- Leur contenu est entièrement gratuit
- Les auteurs des critiques ne sont pas rémunérés pour leurs publications

Limites du corpus

La première limite de notre corpus est liée à l'analyse du profil des critiques. Bien que ces deux corpus représentent deux types de spectateurs différents, nous n'avons pas de chiffres précis sur l'âge, les revenus moyens et le bagage culturel de la plupart de ces individus. D'une part, il nous était impossible d'effectuer des recherches sur l'ensemble des auteurs des critiques de notre corpus. Cela aurait représenté une tâche bien trop chronophage. D'autre part, même en envisageant de récupérer ces informations automatiquement, pour gagner du temps, ces éléments n'étaient pas toujours disponibles en ligne. Lorsque ces informations étaient disponibles cependant, notamment pour les blogueurs, nous les avons

intégrées à nos analyses.

La seconde limite de notre corpus est commune à de nombreux projets en science des données. Elle concerne l'objectivité des sources et des analyses du corpus. Bien que dans une certaine mesure, l'examen de la littérature à grande échelle assure une forme de partialité des résultats, la constitution et l'analyse d'une base de données ne sont jamais neutres. C'est ce que Lisa Gitelman met en évidence par le titre de son essai « Raw Data » is an Oxymoron¹⁰². Malgré leur étymologie latine, les données ne sont pas « données. » Loin d'être des masses d'informations neutres qui permettraient d'apporter une vérité absolue grâce à leur traitement informatique, elles résultent d'un long processus de collecte, d'organisation et d'interprétation qui est mené par un chercheur. En ce sens, aussi large un corpus soit-il, il sera toujours le reflet d'une subjectivité humaine, et donc susceptible d'être biaisé.

Comme nous l'avons expliqué dans cette partie, nous avons essayé d'être le plus impartial possible dans le choix des blogs en fondant notre sélection sur des sites qui répertoriaient les plateformes les plus influentes d'Internet en 2020. Nous aurions pu néanmoins intégrer d'autres blogs à notre corpus, ce qui aurait probablement légèrement modifié nos résultats. De la même manière, bien que *Theatre Record* republie une très grande majorité des critiques journalistiques, certaines sont laissées de côté par manque de place dans leur magazine. C'est pourquoi notre étude apporte des réponses sur la réception du théâtre londonien de manière *globale*, et non *totale*.

Enfin, au-delà des limites propres à notre corpus, les conclusions que nous avons tirées grâce à l'analyse des données sont, elles aussi, à considérer avec une distance critique. A nouveau, l'étude des chiffres repose sur des interprétations humaines. La présence de big data n'implique pas la production de connaissances neutres et absolues. Il s'agit seulement de considérer un objet culturel à plus grande échelle afin d'en proposer une analyse générale.

¹⁰² Lisa GITEMAN, *Raw Data is an Oxymoron*, Cambridge, The MIT Press, 2013.

Remarques conclusives sur le chapitre I

Nous avons vu tout au long de ce premier chapitre dans quels sillages théoriques et pratiques s'inscrit cette recherche. Le premier aspect que nous avons mis en évidence est celui de son appartenance aux humanités numériques théâtrales. Bien qu'il y ait une diversité de projets dans ce domaine, l'étude du théâtre grâce à l'utilisation d'outils techniques reste en marge par rapport à d'autres disciplines comme l'histoire ou la littérature romanesque. Notre étude contribue ainsi à la construction d'un domaine qui n'en est encore qu'à ses débuts.

Le deuxième aspect que nous avons abordé concerne les études de la réception. Apparu au début des années 1970, ce champ encore en construction propose une approche sensible de la littérature. Il s'agit désormais de considérer les objets culturels à travers « l'horizon d'attente » de leurs destinataires. Bien que certains essais en études théâtrales se soient intéressés à cette question, très peu d'entre eux se sont focalisés sur la critique théâtrale londonienne, et aucun à travers le prisme du numérique.

Ce constat nous a ensuite amenés à présenter le cœur de notre projet : l'émergence d'une nouvelle forme de réception au début du XXI^e siècle. La naissance d'une blogosphère à Londres remet en question la légitimité des journalistes. Toutes ces nouvelles voix de l'espace numérique s'élèvent contre la culture journalistique jugée trop élitiste par nombreux de ces blogueurs. Le canon de la critique théâtrale londonienne se trouve ainsi progressivement décentralisé par ses marges qui réclament le droit à la visibilité.

Ce clivage est-il cependant si net ? Bien que ces deux communautés de spectateurs aient des horizons d'attente divergents, n'est-il pas possible d'observer des points de convergence entre eux ? Cette question, centrale à nos travaux, nous a amenés dans une dernière partie à présenter le matériau sur lequel nous allons fonder nos analyses. En créant un corpus de plus de 40 000 critiques théâtrales écrites par les journalistes et les blogueurs londoniens entre 2010 et 2020, nous allons répondre à cette problématique grâce à des moyens computationnels.

Chapitre II

Constitution et présentation des données

Ce deuxième chapitre aura ainsi pour finalité de présenter les enjeux techniques de la création de notre corpus. Ce travail représente plus d'une année de réflexion, d'ajouts et de suppressions de contenu, de traitements informatiques et de relecture. La structuration et le nettoyage des données représentent environ 80 % d'un travail en science des données¹⁰³. Si les sources numériques sont de plus en plus importantes aujourd'hui, cela signifie également que le travail de nettoyage sera proportionnellement plus conséquent. Sans organisation rigoureuse des données, les résultats resteront imprécis et sujets à davantage d'erreurs. L'ordinateur est une machine puissante mais pas encore capable d'ordonner un jeu de données de manière autonome.

C'est donc au chercheur d'anticiper les questions auxquelles il veut répondre afin de construire une architecture de corpus appropriée aux problématiques posées. La nature des questions affectera non seulement le choix des outils informatiques, mais aussi la structuration des données. Il est donc nécessaire de réfléchir en amont aux enjeux techniques que pose une problématique donnée. Ce point névralgique des humanités numériques faisait par exemple partie intégrante du dernier numéro de la *Revue d'Historiographie du Théâtre*. Chacun des contributeurs avait pour mission de répondre à un questionnaire précis sur la conversion de leurs données et sur les problèmes techniques rencontrés¹⁰⁴.

Afin de présenter les processus qui nous ont menés à la mise en place de notre corpus, nous avons divisé ce chapitre en deux parties. Dans la première, intitulée « Création du corpus », nous nous concentrerons sur l'aspect technique de ce travail. Nous aborderons la base de données *Theatre Record*, et plus précisément les protocoles que nous avons mis en place pour rendre ce corpus numérisé en format PDF exploitable par les algorithmes. Nous adopterons ensuite la même démarche pour notre second sous-corpus. Après avoir donné la

¹⁰³Tamrapani DASU, Theodore JOHNSON, *Exploratory Data Mining and Data Cleaning*, New Jersey, Wiley, 2003, p. ix.

¹⁰⁴Marine ROUSSILLON, Christophe SCHUWEY, « Questionnaire – Écrire l'histoire des spectacles avec des bases de données », *Revue d'Historiographie du Théâtre* [En ligne], no. 5, 2017, URL : <https://sht.asso.fr/les-confidants-du-theatre-classique-essai-denquete-quantitative/>, consulté le 25 janvier 2021.

Voici deux des questions du questionnaire :

(a) Comment avez-vous constitué le corpus de votre base de données ? S'agit-il d'un fonds cohérent, déjà constitué par d'autres opérations d'archivage inscrites dans une histoire propre ? Ou bien a-t-il été constitué dans la perspective de la base de données, et dans ce cas selon quels critères ? Dans quelle mesure la constitution de ce corpus opère-t-elle des recompositions de corpus existants ?

(b) Comment les documents (écrits, images...) du corpus ont-ils été convertis en données ? Comment avez-vous choisi les données recensées par la base ? Quels obstacles avez-vous rencontrés ? Quels sont les gains et les pertes réalisés dans cette opération ? Du fonds d'archives manuscrites ou imprimées, du corpus de documents à la base de données, quels sont les effets du changement de support ? Comment les avez-vous pris en compte, voire comment avez-vous essayé de les minimiser ?

définition d'un blog, nous expliquerons quelle technique informatique nous avons utilisée pour récupérer automatiquement le contenu de ces sites Internet.

Dans la seconde partie, « Présentation du contenu du corpus », nous commencerons par introduire les métadonnées que nous avons ajoutées ainsi que les outils en Traitement Automatique des Langues que nous avons appliqués à notre base de données. Nous détaillerons ensuite le contenu de nos deux sous-corpus en nous appuyant sur deux tableaux de bord qui représentent leurs caractéristiques principales sous forme de chiffres (nombre de médias, de critiques rédigées, de pièces examinées, etc.). Nous commencerons par la critique journalistique avant d'introduire la critique numérique. Pour cette dernière, moins connue du grand public, nous avons rédigé une brève description de chacun des blogs. Cette seconde partie du chapitre II permettra ainsi aux lecteurs de comprendre en quoi ces deux jeux de données incarnent deux communautés différentes : le canon de la critique théâtrale anglaise d'un côté, et ses périphéries de l'autre.

II.A Création du corpus

II.A.1 La critique journalistique : *Theatre Record*

II.A.1.1 Quels types d'informations sont mis à disposition des lecteurs ?

Les numéros publiés par *Theatre Record* sont organisés de la même manière. Pour chacun des spectacles mentionnés, un nombre variable de critiques ainsi qu'une série d'informations relatives à la production sont donnés. Ainsi, tous les corps de métier liés à la sphère théâtrale sont listés pour chacune des pièces mentionnées : metteurs en scène, costumiers, acteurs, éclairagistes, décorateurs, scénographes, etc. Le lieu dans lequel a été jouée la pièce de théâtre (The Royal Court Theatre, The National Theatre, The Old Vic, etc.) ainsi que les dates d'ouverture et de fermeture du spectacle sont également indiqués. Les journaux dans lesquels sont parues les critiques sont, pour la plupart, connus de tous : *The Times*, *The Guardian*, *The Independent*, *The Sun*, *The Observer*, etc.

II.A.1.2 Quels types de journaux avons-nous sélectionnés ?

Sur les 84 journaux mis à disposition par *Theatre Record*, nous en avons retenu 23 au total. Cette sélection s'est effectuée selon trois critères. Premièrement, nous avons retiré les critiques issues des plateformes numériques (*Theartsdesk.com*, *LondonTheatre.co.uk*, *BritishTheatreGuide.info*, etc.) car nous voulions obtenir deux sous-corpus de nature distincte. Enfin, nous avons supprimé les journaux qui comportaient un nombre limité de critiques (25) et que nous avons jugés peu connus du grand public (*The Hampstead & Highgate Express*, *The Hackney Gazette*, etc.).

II.A.1.3 Comment avons-nous récupéré le contenu des journaux ?

Les journaux numérisés par *Theatre Record* sont présentés chronologiquement sur leur plateforme en ligne. De 1981 à aujourd'hui, entre 20 et 25 fichiers par année, de 60 pages chacun environ et en format PDF, réunissent l'ensemble des critiques publiées chaque année

dans les plus grands journaux anglais. Voici un aperçu de l'année 2010 par exemple :

Theatre Record

Reviews Archive What's On News About Contact My Account Logout

ARCHIVE

All issues of *Theatre Record* since 1981, with production details, cast lists and reviews for more than 52,000 productions.
Older issues may take slightly longer to load.

Search archive by title, author, theatre, director, cast or creatives

VOLUME 30 - ISSUES

1 - 2 Pages 1 - 96	3 Pages 97 - 152	4 Pages 153 - 212
5 Pages 213 - 276	6 Pages 277 - 340	7 Pages 341 - 396
8 Pages 397 - 456	9 Pages 457 - 516	10 Pages 517 - 572
11 Pages 573 - 632	12 Pages 633 - 692	13 Pages 693 - 744
14 Pages 745 - 804	15 Pages 805 - 868	16 - 17 Pages 869 - 916
18 Pages 917 - 1012	19 Pages 1013 - 1092	20 Pages 1093 - 1156
21 Pages 1157 - 1212	22 Pages 1213 - 1266	23 Pages 1269 - 1324
24 Pages 1325 - 1368	25 - 26 Pages 1369 - 1472	

LATEST NEWS

Julian Oddy
It is with great sadness to tell you that the previous editor of Theatre Record, Julian Oddy passed away...

See more on the [News Page](#).

AFFILIATES

[doollee.com](#)
Doollee

Search the largest collection of Plays & Musicals in the world.
[Find Your Script](#)
stageplays.com
[Stageplays](#)

LATEST REVIEWS

- [Good](#)
British Theatre Guide
- [Jekyll & Hyde](#)
The Times
- [La Performance](#)
The Times

© 2022 Theatre Record Ltd

Fig. 9 : Capture d'écran de l'année 2010 sur le site TheatreRecord.com.

Pour accéder au fichier PDF, il suffit seulement de l'ouvrir dans une autre page web et de le télécharger. Afin de faciliter le traitement des données pour la suite du travail, la première étape consiste à convertir ces fichiers PDF en fichiers texte. Il est possible de travailler à partir d'un format PDF directement avec un langage de programmation (comme R ou Python), mais les possibilités d'analyses sont restreintes. Plusieurs solutions ont été envisagées avant d'en trouver une satisfaisante. Le principal obstacle à ce travail est lié à la disposition du texte dans les fichiers PDF. Si nous prenons le troisième fichier de l'année 2010 par exemple :

EVENING STANDARD

3.2.10

Sarah Frater

Combine the often gruesome authenticity of 999 calls with their life-and-death urgency and you have a potent source of drama, albeit one at a permanent panic pitch.

The baffled child whose mother is dying, the terrified husband whose wife is choking, and the luckless neighbour who finds the man next door has cut off his arm with a chainsaw — Gate co-director Carrie Cracknell has taken the transcripts from real-life emergency calls and presented them on the stage of the Notting Hill micro-theatre.

I say present, because for all its ability to unsettle the audience and reveal our longing to keep life intact, the 45-minute Breathing Irregular is more a recital than a piece of theatre. Cracknell will surely argue that she intentionally under-metamorphosed the material for the stage to keep it raw but the result often feels like a read-through.

Jane Mason's dance sequences are more conspicuously performed yet they are collaged rather than integrated with the speech (a conflation dance-and musical-theatre achieved decades ago), plus they lack a vocabulary that illuminates the stories in movement form.

In more adept hands, the dance would resolve the stories, or tell you

limited opportunities of the feeble drama, limp direction and shapeless text. One wonders to what purpose Breathing Irregular was conceived; it simply regurgitates the words of the calls with hammy shouting but adds no insight and yields no surprises. The real-life emergencies must have been unbearable drama for those involved but, recreated here together, they add up to a whole lot of diddly.

GUARDIAN

5.2.10

Lyn Gardner

BREATHING IRREGULAR

GATE

2–28 February 2010

BREATHING IRREGULAR

A new piece

Performed by

Termitope Ajose-Cutting Bryony Hannah Brendan Hughes

Éva Magyar Mary Erskine

Director Carrie Cracknell

Decor Holly Waddington

Lighting Lucy Carter

Sound Edward Lewis

Choreography Jane Mason

Music Tom Mills

it will surely encourage us all to sign up to that first-aid course we never got round to. Despite a terrific quartet of performers, too often the work is merely illustrative of the transcripts. While there are moments that make you hold your breath, I often felt uneasy — like a voyeur peering into somebody else's disaster.

THE TIMES

5.2.10

Donald Hutera

A few years ago Carrie Cracknell and Natalie Abrahami, the co-

characters, situations and states with a breathless, heroic fluidity. Hannah is especially truthful, demonstrating her range first as a reassuring phone operator advising an anxious caller on how to keep his seriously ailing wife alive, and then switching gears as a frightened but plucky five-year-old boy whose mother has inexplicably collapsed and won't respond.

The show's structure is fragmentary, but the shards accumulate into a multifaceted portrait of raw fear and trembling suspense. All the sobering, simulated life-or-death emotion might be difficult to sustain or tolerate were the piece longer than an hour. It actually lasts less.

Breathing Irregular rarely gripped me, and the ending is more ambiguous than necessary, but the performance is delivered with an artful power. Interestingly, while some of the incidents we hear and glimpse remain unresolved on stage, in real life none ended in despair or defeat. These are tales of survival.

OBSERVER

7.2.10

Susannah Clapp

A couple of years ago the *Guardian* published transcripts of telephone calls made by members of the public to the emergency services. Carrie Cracknell has used some of these to

Fig. 10 : Capture d'écran d'un des fichiers PDF de l'année 2010 mis à disposition par TheatreRecord.com.

Comme le montre la figure 10, nous pouvons remarquer que chaque critique se succède sous forme de colonnes. Ce type de format présente des difficultés pour le traitement informatique car cela demande au programme de respecter la succession des colonnes, et non un traitement linéaire de gauche à droite. La réussite de la séparation des colonnes dépend de plusieurs facteurs : la qualité de la numérisation mais aussi la simplicité de la mise en page et l'absence d'éléments parasites (comme ceux qu'induit une page trop transparente par exemple). Pour éviter le mélange des colonnes, il est possible d'utiliser un outil comme ScanTailor¹⁰⁵ qui permet de découper les images en détectant les colonnes avant d'appliquer un outil de transcription dessus. L'inconvénient est que cela peut prendre beaucoup de temps pour un gros corpus.

Après avoir passé en revue trois techniques différentes¹⁰⁶, nous avons finalement utilisé ABBYY FineReader¹⁰⁷. FineReader est un logiciel de reconnaissance optique de

¹⁰⁵ <https://scantailor.org/> [En ligne], consulté le 09 novembre 2021.

¹⁰⁶ Ces trois techniques sont détaillées en annexes p. 234 – 235.

¹⁰⁷ <https://pdf.abbyy.com/fr/> [En ligne], consulté le 09 novembre 2021.

caractères qui permet de convertir des fichiers PDF dans un vaste champ de formats (.txt, .png, .csv, etc.) et ce, dans plus de dix langues différentes. La conversion des fichiers PDF vers les fichiers texte prend environ dix minutes par fichier. Puisque nous avons 228 fichiers PDF au total, cette tâche représente environ 38 heures de traitement informatique.

Les colonnes sont prises en compte, ce qui permet d'assurer une certaine cohérence du contenu des textes retranscrits. De plus, les textes sont convertis sous forme de paragraphes qui se succèdent les uns aux autres. Cela retire donc le problème des colonnes qui, même sous format texte, pose des difficultés. Voici un aperçu du fichier texte une fois converti par FineReader :

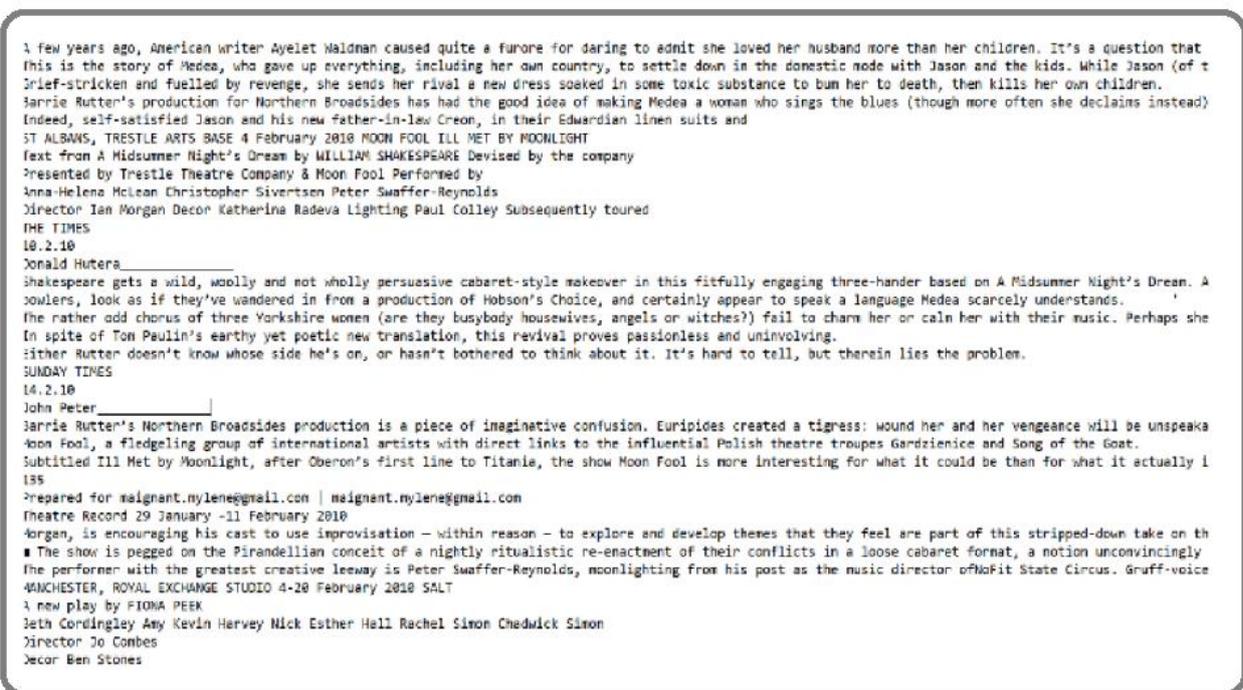


Fig. 11 : Capture d'écran d'un exemple de fichier PDF converti par ABBY FineReader.

L'étape suivante consiste à séparer les métadonnées du corps du texte. Au début de chaque fichier PDF se trouve une série d'informations relatives au numéro publié. En dehors des notes de l'éditeur qui ne nous sont pas utiles pour la suite de nos travaux, d'autres informations précieuses ont été conservées dans un fichier texte séparé afin de créer les métadonnées de chaque fichier. Ainsi, au début de chaque PDF se trouvent :

1. La liste des pièces présentes dans le fichier en question
2. Le nom du dramaturge
3. Le théâtre dans lequel a été jouée la pièce

4. La date d'ouverture et de fermeture du spectacle

Pour le premier PDF de l'année 2010 par exemple (sachant que chaque année est constituée de 25 PDF environ), voici comment se présente son architecture :

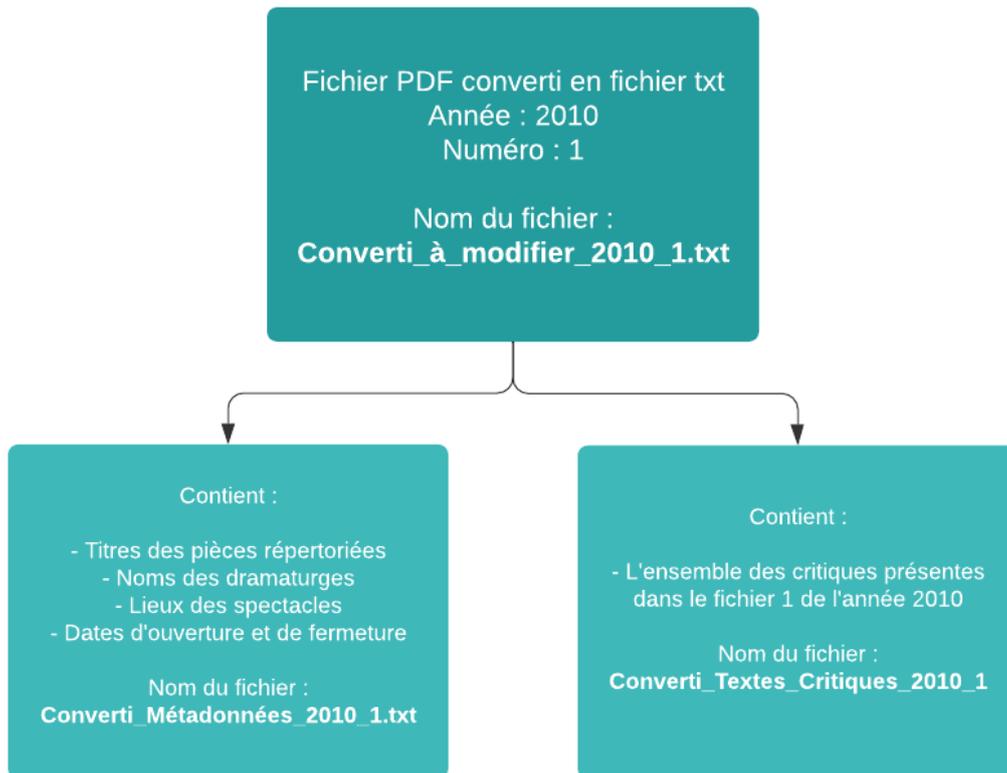


Fig. 12 : Schéma de l'exemple de l'architecture d'un fichier.

Nous rédigeons ensuite un script avec Python afin de nettoyer ces données. La librairie « Re¹⁰⁸ » est particulièrement utile pour cette étape car elle permet d'utiliser les expressions régulières. Les expressions régulières permettent de chercher des chaînes de caractères dans un texte et de les modifier, de les remplacer ou de les supprimer. Ainsi, nous supprimons tous les numéros de pages, les mentions de notre adresse e-mail (*Theatre Record* étant une base de données payante, la mention de l'e-mail du lecteur est inscrite en bas de chaque page), les informations relatives aux photographies... Tout ce qui est à corriger et repérable automatiquement est nettoyé grâce au script.

Toutes les erreurs ne sont cependant pas détectables automatiquement. Deux types en

¹⁰⁸<https://docs.python.org/3/library/re.html> [En ligne], consulté le 09 novembre 2021.

particulier qui requièrent une relecture humaine. La première concerne les titres des pièces ainsi que la liste des professions théâtrales qui lui sont associées (metteurs en scène, acteurs, chorégraphes, etc.). Ces informations se trouvent soit au milieu, soit sur les côtés d'une page, avec une taille et une police de caractère différentes ainsi qu'une mise en page centrée. La figure 11 ci-dessous en rend compte :

Theatre Record 29 January - 11 February 2010

Lope de Vega was all these hundreds of years ago. You are reminded that this is not real life but a play. The ending can change - the characters can be given different lovers. They are, after all, actors - and actors can improvise or interact with the audience.

This is a version by David Johnston, who a few years ago translated the superb *The Dog in the Manger*, also by de Vega. Here Johnston has transformed the original script, "reallocation lines, combining characters and playing with form".

And yet the play is not gripping throughout. Perhaps the reason for that is, at times, the pace is simply too frenetic - the characters do not pause enough for the audience to be sufficiently engaged with their fate. The fact that so many of the cast swoon in love is a central point, but it nonetheless gives a touch of senseness to the action.

One memorable moment has the keeper and a couple of the inmates collecting in the streets for the sayings. Valencia is justifiably proud of its madhouse but is demised if it is going to pay for it. More of that would have been welcome.

Yet there is a lot to like. All the cast put in good performances. I particularly enjoyed Lawrence Fuller as the asylum keeper, cleverly showing that the one character who may be mad is not an inmate, but the person designated to guard the insane.

Last year saw a version of Calderon's *Life is a Dream*. Now we have another of the writers of the Spanish golden age. It's a reminder that at the time Marlowe,

Shakespeare and Jonson were writing, Spain had its own great. Keep those revivals coming.

TIME OUT LONDON
18.2.10
Sam Marlowe

Floriano has murdered in the name of love; to avoid detection, he has himself committed. Effilia has fled her patriarchal home with a servant who betrays her; stripped of her jewels and clothes and abandoned, she too is flung into the madhouse. There, they meet and fall in love, both fringing insanity. Is true love nothing more than lunacy? And in a crazy world, is there illusion, delusion, or as solid as any other construct we choose to term reality?

These are the questions tossed into the air and juggled with wit and skill by Lope de Vega in this Golden Age drama, jauntily translated by David Johnston. Simon Evans's production, transferring from the White Bear, is energetic, stylistically designed by Kate Matthews and acted with verve. While it delights in mechanical games and in some joyously grotesque performances, it neglects the work's darker potential.

There's a horror in the play's motifs of incarceration, forced marriage and marital torment that ought to supply a disturbing counterpoint to its comedy. Kathryn Beaumont and William Bechtelmann are charming as the lovers, and Evans's staging is great fun, but it should be more than that.

Panaphobia

LOST
28 January - 14 February 2010
PANAPHOBIA
A new play by CERI ASHCROFT & MIRANDA KEELING

Universal Citizens

Chi-Lin Niu	Roke
Miranda Keeling	Melgold
Austin Spangler	Mr Wills
Carl Ashcroft	Lisa
Jerrish Yakubu	Mrs Irene Lowkey
Geoff O'Brien	Tom
Michael Yakubu	Daisy
Lauren Olson	
Harry Myers	Voice of Dad

Director: Ray Downing
Kate McGregor
Lighting: Mike Swan
Sound: Austin Spangler

TIME OUT LONDON
4.2.10
Robert Shore

When fearless young Rose is sent on a mission to deliver supplies to her sick aunt at the top of a tower block, a power cut brings her into contact with the neurotic neighbours - is there a Big Bad Wolf among them?

This Universal Citizens reworking of the "Little Red Riding Hood" story has been expanded from a 30-minute devised piece. It's a worthy exercise, but there isn't really

The 14th Tale

COTTESLOE
10 February - 14 March 2010
THE 14TH TALE
Written and performed by INJIA ELLAMS

Director: Thierry Lawson
Lighting: Michael Nabarro

OBSERVER
14.2.10
Kate Kellaway

Injia Ellams is a self-diagnosed bad boy from a line of troublemakers. The 14th Tale, a short, semi-autobiographical piece, begins in Nigeria where he seldom escaped whipping from a "mercenary of mum". We pursue him through a "dark afternoon" in a Nigerian

particular tag must have stood out as much for its artistry as its exotism.

Delivered in a voice as sonorous as a steel drum, Ellams's poetry is smooth enough to melt in the ear, yet packed full of snap, crackle and pop. Accompanied by a fluid, simple movement score, his words conjure vivid cartoonish images imbued with atmosphere and detail.

The latest in 'a long line of troublemakers', "The 14th Tale" recalls a string of mischievous antics and sets them up as farrago. The revenge taken on a school bully using a tube of toothpaste and a pack of pins. The cheating ex-girlfriend's shower-head unscrewed and filled with paint. And, of course, the doozing with his doodle.

We've all got stories of misdeeds from our past, but Ellams's

enough material to fill the new hour-long format: there's a lot of fairly vacuous sitting (and walking) around on stage. A bigger problem is that the story is crying out for a gritty, site-specific treatment: there's no real stakes of fear when the lights suddenly go out in this well-sophistered and distinctly unvigilant auditorium. In addition, the sophistication (projection screen, complicated sound design etc) of Ray Downing and Kate McGregor's production only serves to emphasize how undernourished the character drawing is.

my image, tried to stay close, but still trouble came." On stage, he presents himself as a nervous kid with bloody trousers, pecky glasses and an impish grin. His writing is slight but sinuous. And he moves like a careful cat. But isn't performing at the National as compromising for a troublemaker as getting a gold star from teacher?

TIME OUT LONDON

SUNDAY TIMES
21.2.10
Louis Wise

In this one-hour, one-man show, Injia Ellams deals on his troubled teenage youth and his heart-breaking family past. (The two are one and the same: "I'm from a long line of troublemakers.") He delivers a pleasant and fluid performance, explaining the bits of mischief he has got up to in Nigeria, then London, then Dublin, growing up (or not) along the way. It's a delicate work, touching on heavy themes with a light touch, and its impact rests entirely with the

Fig. 13 : Capture d'écran d'un exemple de fichier PDF mis à disposition par Theatre Record.

Le logiciel FineReader retranscrivant toutes les données les unes à la suite des autres de manière linéaire, comme vu dans la capture d'écran présentée sur la figure 11, ces informations se trouvent ainsi mêlées aux textes des critiques. Ces informations étant très aléatoires (chaque pièce est associée à une liste de noms précis qui changent d'une pièce à

l'autre), il nous est donc difficile de reconnaître un quelconque motif linguistique récurrent que nous pourrions nettoyer automatiquement. Il nous faut alors relire ces textes en diagonale afin de repérer les titres de pièces suivis des autres noms et prénoms des différentes professions afin de les supprimer nous-mêmes.

Le second élément que nous ne pouvons corriger automatiquement concerne le texte des critiques. À nouveau, la disposition des textes dans les fichiers PDF se révèle parfois problématique. Nous pouvons le voir dans la capture d'écran présentée par la figure 13 : trois pièces sont ici passées en revue. La première est contenue dans le premier rectangle en haut, à gauche, la deuxième en haut, à droite (*Panaphobia*, une seule critique) et la troisième (*The 14th Tale*, trois critiques) dans le bas de la page. FineReader retranscrivant les colonnes les unes à la suite des autres, il considère ainsi que la critique pour *Panaphobia* continue jusqu'au bas de la page. En réalité, ce morceau de texte appartient à la pièce *The 14th Tale*, publiée dans le *Time Out London*. Ces erreurs sont néanmoins en partie repérables dans le fichier texte de sortie. Toutes les critiques qui ne se finissent pas par un point signalent un problème de ce genre. Dans ce cas-là, il faut donc retourner sur le fichier PDF d'origine afin de récupérer la suite cohérente du texte.

Les étapes que nous venons de décrire concerne un fichier PDF seulement. Nous avons dû répéter ce processus sur 228 fichiers, ce qui représente à peu près 1 150 heures de travail qui ont été réparties entre deux chercheurs. Bien que ces problématiques semblent éloignées de notre question initiale, à savoir « Quels sont les points de convergence et de divergence entre le centre de la critique théâtrale londonienne et ses périphéries ? », elles sont pourtant nécessaires pour y répondre. L'informatique permet d'aborder des questions littéraires classiques sous un angle nouveau, mais elle impose également ses propres contraintes. Ces dernières, souvent discrètes lors de la présentation des résultats, requièrent néanmoins beaucoup de temps et de réflexion. Des tâches qui pourraient sembler insignifiantes au départ, comme la conversion de fichiers PDF en fichiers texte, représentent en réalité de véritables enjeux dans la réalisation d'un projet.

Afin de rendre compte de toutes ces étapes qui nous ont menés à la création de notre premier sous-corpus, nous les avons résumées dans le schéma suivant (Fig. 14) :

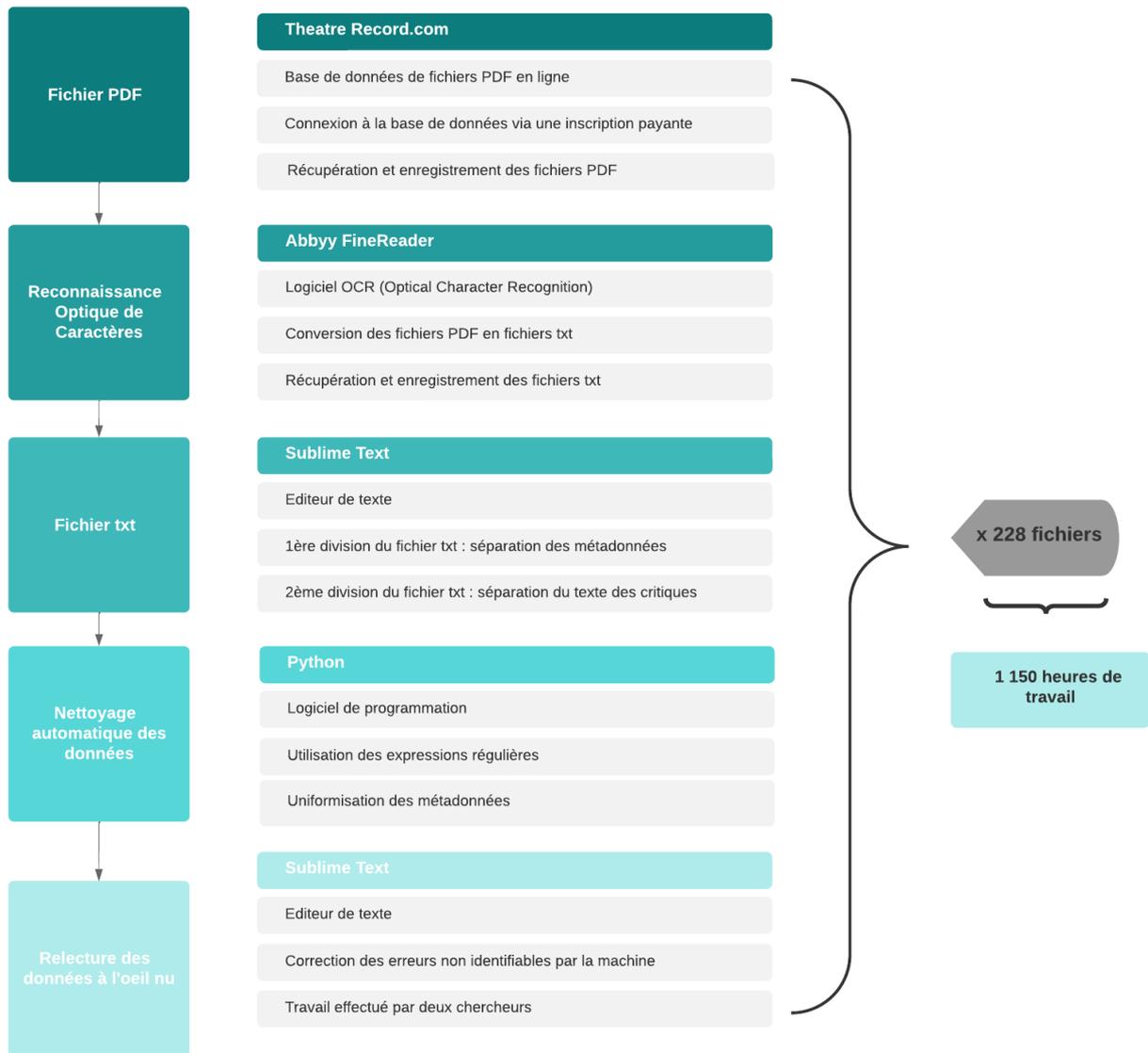


Fig. 14 : Schéma récapitulatif des différentes étapes nécessaires pour la création du sous-corpus I.

Transition

Pour résumer cette première étape, nous sommes partis de 228 fichiers PDF que nous avons téléchargés puis convertis en fichier texte grâce au logiciel FineReader. Nous avons ensuite divisé ces fichiers de sortie en deux types de sous-fichiers : ceux qui contiennent les métadonnées d'un côté, et ceux qui sont constitués du texte des critiques de l'autre. Nous avons ensuite rédigé un script avec Python pour effectuer un premier nettoyage à l'aide de la librairie « Re ». Ce premier travail restant insuffisant pour obtenir des données propres, nous avons dû engager un travail de relecture effectué par deux chercheurs.

A nouveau, ces problématiques, très techniques et pourtant éloignées de notre objet d'étude sur le théâtre, sont pourtant une des composantes majeures de notre projet, et de nombreux travaux dans les humanités numériques en général. Si ces données correctement nettoyées et structurées représentent une tâche chronophage et laborieuse, elles sont cependant les garantes de résultats cohérents. Au total, ce sont donc plus de 21 717 critiques sous licence (Creative Commons Attribution 4.0 International) que nous avons répertoriées pour ce premier sous-corpus, et qui sont désormais en accès libre sur Zenodo¹⁰⁹ et sur notre site Internet¹¹⁰.

Dans la partie suivante, nous nous attacherons à présenter le travail effectué pour la construction de notre corpus témoin, c'est-à-dire celui fondé sur les critiques numériques. Contrairement à ce premier corpus, nous n'avons pas de liste prédéfinie des blogs disponibles sur Internet. Après avoir passé en revue la grande majorité d'entre eux, nous nous sommes appuyés sur deux sources Internet qui répertorient les blogs jugés les plus populaires en 2020, date à laquelle nous avons commencé ces recherches. Après avoir donné une définition du terme « blog », nous détaillerons les procédés informatiques que nous avons implémentés afin de récupérer le contenu de ces sites Internet.

¹⁰⁹<https://zenodo.org/record/6799656#.Y0hSDkzP2so> [En ligne], consulté le 28 août 2022.

¹¹⁰<https://dramacritiques.com/fr/telechargements/> [En ligne], consulté le 28 août 2022.

II.A.2 La critique numérique : les blogs

II.A.2.1 Qu'est-ce qu'un blog ?

Vaughan rappelle que le terme « blog » est le diminutif de « weblog » qui fut inventé par le blogueur américain Jorn Barger en décembre 1997. Un an plus tard, un autre blogueur, Jesse James Garrett, fit la liste de tous les « weblogs » disponibles en ligne. Les auteurs y listaient principalement les différents liens Internet qu'ils jugeaient intéressants et y ajoutaient leurs commentaires. Comme l'explique Megan Vaughan dans *Theatre Blogging*, c'est à partir de l'automne 1999 que le terme « blog » devint populaire. Non seulement il fut intégré à *Jargon Scout*, un catalogue qui répertoriait les nouveaux termes liés à Internet, mais également à une plateforme intitulée « Blogger ». Mis en place par Pyra Labs, ce nouvel outil permettait aux utilisateurs de publier du contenu en ligne sans avoir de connaissances en programmation informatique¹¹¹ (Vaughan, p. 4-5).

De plus en plus d'individus commencèrent ainsi à ouvrir leur blog à partir du début des années 2000. La définition de ce dernier évolua également. Grâce à la multiplication de plateformes qui facilitaient la mise en ligne de contenus, les blogueurs pouvaient désormais réécrire certains de leurs textes ou ajouter des entrées au cours du temps. En 2003, Jill Walker Rettberg définit le blog comme un « continuum from confessional, online diaries to logs tracking specific topics or activities¹¹² ». C'est à partir de ce moment que les blogueurs commencèrent à exprimer davantage leurs opinions politiques, notamment aux Etats-Unis. Il ne s'agissait plus seulement de narrer la quotidienneté de leurs activités et de relayer des informations trouvées sur d'autres sites Internet. Les blogueurs publiaient désormais de réels billets engagés. En 2005, la blogosphère connut un véritable succès : le blog devint un moyen de communication aisément manipulable qui permit à un nombre grandissant d'utilisateurs de s'exprimer.

Les blogs de critiques théâtrales ont suivi une chronologie similaire. A partir de 1997, *BritishTheatreGuide.com* et *WhatsOnStage.com*, les deux premières plateformes spécialisées en théâtre, apparaissent sur Internet. En 2007, soit deux ans après le premier succès de la blogosphère, de nombreux étudiants, professeurs et amateurs créent leur site. Enfin, à partir de 2010, la blogosphère théâtrale londonienne prend son essor : des dizaines d'individus

¹¹¹ VAUGHAN, *op. cit.*, p. 4 – 5.

¹¹² Walker RETTBERG, « Final Version of weblog definition », *Jill/txt* [En ligne], publié le 28 juin 2003, URL : [final version of weblog definition – jill/txt \(jilltxt.net\)](http://final-version-of-weblog-definition-jill/txt-jilltxt.net), consulté le 09 novembre 2021.

prennent la parole dans l'espace numérique en ouvrant leur propre blog. C'est précisément ce tournant qui est au centre de notre étude.

II.A.2.2 Quels types d'informations sont mis à disposition des utilisateurs dans les blogs ?

Comme pour le journal *Theatre Record*, un nombre variable d'informations relatives à la production examinée par le critique est donné. Le lieu dans lequel a été jouée la pièce de théâtre (The Royal Court Theatre, The National Theatre, The Old Vic, etc.), le nom du dramaturge ainsi que la date de publication de la critique sont nécessairement indiqués. Par ailleurs sont précisées, en fonction des blogs, les dates d'ouverture et de fermeture du spectacle avec un lien vers une billetterie en ligne ou la mention du metteur en scène. Dans les plateformes qui regroupent plusieurs contributeurs, un portrait des critiques est parfois donné à la fin de chacune de leur publication.

II.A.2.3 Comment avons-nous récupéré ces informations ?

Afin de créer notre second corpus fondé sur les blogs, nous avons utilisé une technique appelée le web scraping. Le web scraping permet d'extraire le contenu d'une ou de plusieurs pages Internet automatiquement en créant un utilisateur factice qui se connectera à chacune des pages et qui aspirera le texte voulu. Au lieu de récupérer les informations manuellement en répétant l'opération « copier / coller » dans un fichier texte, nous pouvons faire appel soit à un Web scraper en ligne, c'est-à-dire un logiciel qui aspire les données de manière autonome, soit rédiger un script nous-mêmes à partir du langage de programmation de notre choix.

Bien que certains logiciels soient efficaces pour effectuer ce type de tâche sans passer par la programmation, comme l'extension « Webscraper¹¹³ » de Google Chrome, nous avons décidé de rédiger le script nous-mêmes avec Python. Si cette méthode prend davantage de temps, elle permet néanmoins d'obtenir une meilleure vision de l'architecture du site sur lequel se trouvent les données et, ainsi, d'obtenir des résultats plus précis. Pour réaliser ces scripts, nous nous sommes inspirés de la méthode proposée dans le manuel *Web Scraping with*

¹¹³<https://chrome.google.com/webstore/detail/web-scraper-free-web-scr/jnhgnonknehpejjnehehlkklplmbmhn>
[En ligne], consulté le 09 novembre 2021.

*Python: Collecting Data from the Modern Web*¹¹⁴ ainsi que des cours en ligne de Xavier Dupré¹¹⁵, professeur à l'École Nationale de la Statistique et de l'Administration Économique (ENSAE). Pour rendre compte du travail qui a été effectué sur la récupération du contenu des 28 blogs, nous avons choisi d'explicitier le script utilisé pour aspirer les données de la plateforme WestEndWilma.com¹¹⁶. Pour les autres blogs et plateformes, le code est sensiblement similaire. La structure reste identique, excepté le nom des balises qui varient d'un blog à l'autre.

Ce travail se divise en deux étapes. La première consiste à récupérer les liens où se situent les pages. Pour cela, il faut, de manière automatique, se rendre sur chacune des pages et enregistrer leur lien URL dans une liste. Dans un second temps, il s'agit de récupérer le texte de chacune de ces pages. Ces deux étapes sont explicitées ci-dessous :

Partie 1 du web scraping :

Nous commençons par l'analyse d'une seule page. La capture d'écran que montre la figure 15 est tirée de la première page du blog WestEndWilma.com. Nous avons indiqué par un rectangle bleu la partie de la page web qui nous intéresse et que nous allons devoir récupérer automatiquement. Cette dernière contient les quatre liens qui renvoient chacun à une critique théâtrale.

¹¹⁴Ryan MITCHELL, *Web Scraping with Python: Collecting Data from the Modern Web*. Sebastopol, O'Reilly Media, 2015.

¹¹⁵ Xavier DUPRE, « Python pour un data scientist / Economiste, Web scraping », *Python dans tous ses états*[En ligne], URL : http://www.xavierdupre.fr/app/ensae_teaching_cs/helpsphinx/notebooks/TD2A_Eco_Web_Scraping.html, consulté le 09 novembre 2021.

¹¹⁶ Le script commenté de l'extraction des données de la plateforme WestEndWilma.com est disponible en annexes p. 236 – 239.



LOVE, LOVE, LOVE at the Lyric Hammersmith 'promised so much more than it delivered'

March 26, 2020 // Plays, Reviews // Comments are off

Written by Mike Bartlett, Love, Love, Love serves up the lives of Sandra and Kenneth in three acts. We begin in 1967 when they meet [Read more](#)



Irish Play ON BLUEBERRY HILL opens at Trafalgar Studios exploring forgiveness, survival and love

March 12, 2020 // Plays, Reviews // Comments are off

Sebastian Barry's 2017 production of On Blueberry Hill has already been seen by audiences in Ireland, Paris and New York and is now making its [Read more](#)

Receive the latest news and special offers by email

Email Address

Please tick the box below to confirm you would like to receive emails from West End Wilma with theatre news updates, competitions and special offers.

Email

Subscribe



March 26, 2020.

LOVE, LOVE, LOVE at the Lyric Hammersmith 'promised so much more than it delivered'



March 12, 2020.

Irish Play ON BLUEBERRY HILL opens at Trafalgar Studios exploring forgiveness, survival and love



February 24, 2020.

La Cage aux Folles (Park Theatre)



February 20, 2020.

THE UPSTART CROW at London's Gielgud Theatre is 'a rollicking ride' *****



CHECK OUT THE RED CARPET ACTION FROM THE WEST END WILMA AWARDS 2019.

Fig. 15 : Capture d'écran de la première page du site WestEndWilma.com.

Un site Web est un ensemble d'informations codées en langage *HTML* qui permet de décrire à la fois le contenu et la forme d'une page *Internet*. Pour extraire les données voulues, nous devons parcourir l'arborescence de la page HTML et sélectionner les balises correspondantes. Les balises structurent le contenu d'une page HTML. Elles sont signifiées par les symboles « <> » qui indiquent le début ou la fin d'une partie. Pour notre travail, il faut repérer les balises dans lesquelles sont contenues les informations que nous désirons extraire. C'est dans l'outil « Menu », puis dans l'onglet « Inspecteur » et enfin dans la section « Développement web » que les informations nécessaires pourront être trouvées :

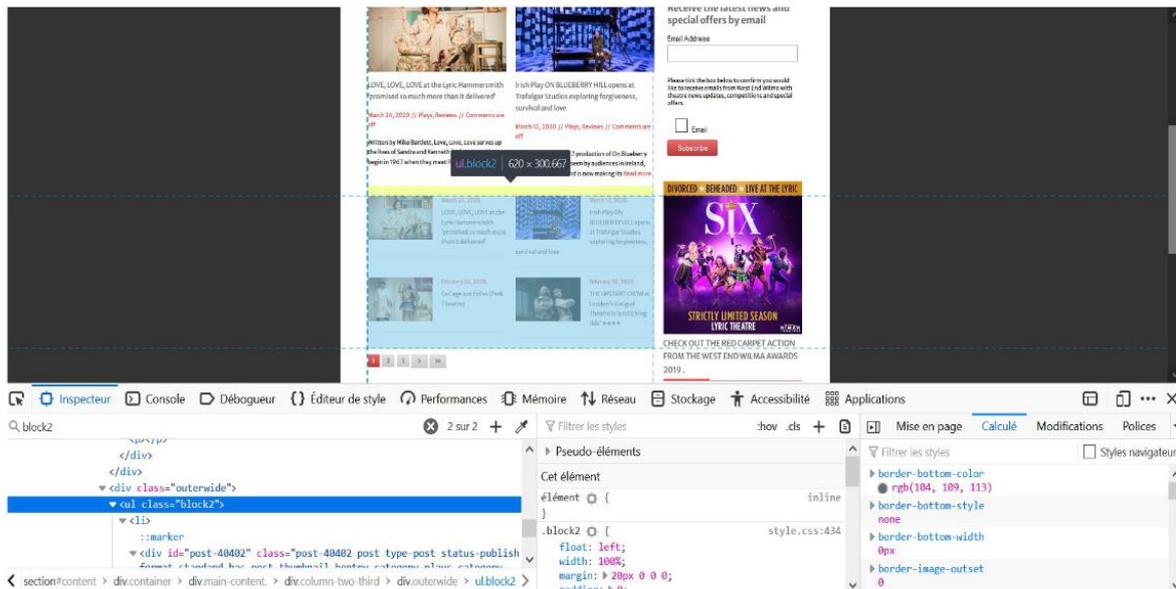


Fig. 16 : Capture d'écran de l'ouverture de l'outil « Développement Web ».

Dans ce cas précis, comme le montre la figure 15 ci-dessus, nous pouvons remarquer que le bloc qui nous intéresse correspond à une liste (ul) de class « block2 ». Les listes et les blocks sont des éléments de langage HTML. Le principe de cette étape repose sur la récupération d'éléments précis qui sont invisibles à l'écran mais disponibles dans le code source. Dans cette liste se trouve tout ce dont nous avons besoin. Chaque page du blog est construite ainsi : quatre liens vers quatre critiques sont donnés à chaque page. Nous enregistrons d'abord ces données, puis nous allons chercher les liens qui conduisent aux autres pages (263 au total à ce jour) afin de réitérer la même opération. Nous créons ensuite une fonction pour effectuer ces étapes automatiquement et stocker tous les liens de toutes les pages.

Partie 2 du web scraping :

Il s'agit désormais d'aller récupérer le texte des pages et de l'enregistrer dans un fichier texte. La démarche à suivre est sensiblement similaire à ce que nous venons de décrire. Il faut d'abord repérer les balises dans lesquelles sont contenus le texte des critiques ainsi que les métadonnées que nous voulons extraire (date, titre, auteur de la critique, etc.). Ces

informations extraites, nous les envoyons ensuite dans un fichier texte indépendant, de sorte qu'à la fin du processus, nous ayons autant de fichiers textes que de critiques.

Nous avons répété cette opération pour les autres blogs. Selon leur structure, il était plus ou moins aisé de récupérer les informations voulues. En fonction du volume de données à extraire du site web, l'algorithme tournait plus ou moins lentement, et plus ou moins longtemps. Pour la plateforme AyoungerTheatre.com par exemple, le programme a dû tourner plus de dix heures. Il nous a également fallu passer davantage de temps sur certaines plateformes car le code HTML des pages avait évolué au fur et à mesure des années. La mise en page des textes publiés au début des années 2010 différait de celle des textes de 2020. Dans ce cas-là, il fallait repérer sur quelle page intervenait le changement, et changer la structure du code sur le script informatique. Enfin, comme pour le premier sous-corpus, un long travail de nettoyage des données a été effectué, tant manuel qu'automatique. Au total, ce sont environ 250 heures qui ont été requises pour construire ce corpus témoin. Le schéma ci-dessous récapitule les étapes que nous venons de décrire :

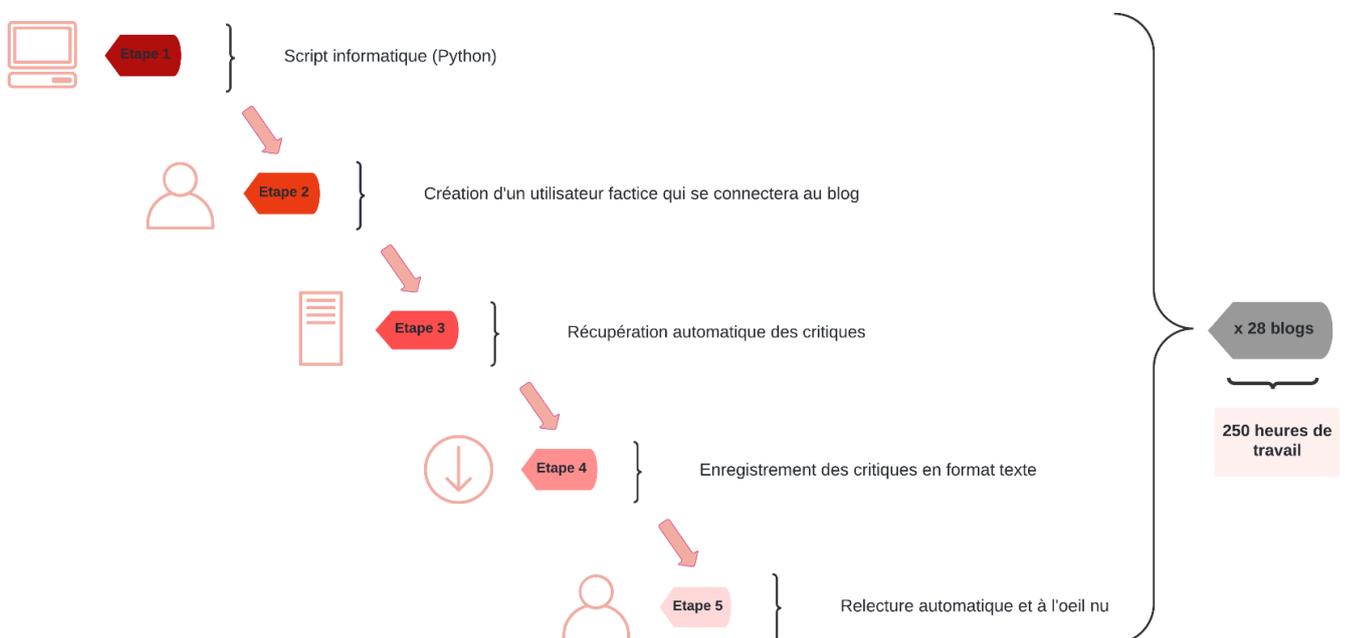


Fig. 17 : Récapitulatif des étapes nécessaires pour créer le second sous-corpus.

Transition

Cette partie nous a permis de présenter la construction de notre second sous-corpus fondé sur la critique numérique. Nous avons sélectionné les 28 blogs de notre base de données en fonction de deux sources : Vuelio.com et MyTheatreMates.com. Nous avons ensuite utilisé le web scraping, une technique de récupération automatique des données, afin d'extraire le contenu de chacun de ces blogs. Nous avons répertorié 21 326 critiques au total, dont les trois quarts sont aujourd'hui en accès libre (15 049 critiques) sur Zenodo et sur notre site Internet, sous la licence Creative Commons Attribution 4.0 International. Nous avons envoyé une demande d'autorisation de publication à chacun des 28 blogueurs. Seulement 21 ont répondu, ce qui explique que la totalité du corpus ne peut être mise en ligne.

Nous allons maintenant présenter le contenu de notre corpus. Nous avons divisé cette partie en deux sous-parties. Dans la première, nous nous concentrerons sur les dernières étapes techniques qui nous ont permis de finaliser la préparation de notre corpus pour les analyses computationnelles à venir. L'application d'algorithmes requiert une organisation rigoureuse des informations. Toutes ces critiques collectées doivent désormais être structurées sous forme de matrice de données, ou de tableau, afin de les analyser. Nous présenterons également les métadonnées que nous avons ajoutées ainsi que les outils en Traitement Automatique des Langues que nous avons appliqués.

II.B Présentation du contenu du corpus

II.B.1 Structurer les informations

II.B.1.1 Mise en place d'une matrice de données

Maintenant que nous avons récupéré et converti les données dans un format plus facilement manipulable (le format texte), nous devons les structurer. Pour pouvoir les étudier, c'est-à-dire pour pouvoir les analyser grâce à des algorithmes, nous devons les organiser selon une certaine structure qui sera reconnaissable par l'ordinateur et rapidement manipulable par l'humain grâce à l'informatique. L'article d'Hadley Wickham¹¹⁷ intitulé « Tidy Data » est particulièrement éclairant dans ce contexte puisqu'il détaille l'architecture d'un jeu de données structuré. Pour cela, ce dernier doit comporter les trois caractéristiques suivantes :

1. Chaque variable est représentée sous forme de colonne et contient des valeurs
2. Chaque observation constitue ce qui est communément appelée une ligne
3. Chaque unité d'observation forme un tableau

Ce format, appelé « matrice de données », facilitera l'analyse des textes pour la suite de notre travail. Nous avons décidé de structurer nos données en format csv. Nous aurions cependant pu utiliser les formats JSON ou XML, deux autres manières répandues dans le domaine de la science des données pour stocker des informations. Tandis que le format JSON permet l'enregistrement de données plus complexes, le format XML permet de définir précisément les données à l'aide de balises. Dans la mesure où nous voulions ouvrir ce corpus au plus grand nombre, nous avons opté pour le csv, un format plus aisément compréhensible et maîtrisable que celui du JSON ou du XML.

La figure 18 illustre une partie de notre corpus en montrant les différentes critiques rédigées pour la représentation de la pièce *The Caretaker*, en 2016, d'Harold Pinter. Il suffit désormais d'appliquer les algorithmes sur la colonne « Review » et d'effectuer ensuite des calculs comparatifs entre les différentes variables.

¹¹⁷Hadley WICKHAM, « Tidy Data », *Journal of Statistical Software*, vol. 59, no. 10, 2014, p. 4.

Corpus	Newspaper / Blog	Creation	Publication	Reviewer	Gender	Title	Playwright	Rating	Review	Theatre
Corpus I	Evening Standard	1859	07.04.2016	Henry Hitchings	Male	THE CARETAKER	Harold Pinter	None	Timothy Spal	Old Vic Theatre
Corpus I	Guardian	1821	08.04.2016	Michael Billington	Male	THE CARETAKER	Harold Pinter	None	The plays of	Old Vic Theatre
Corpus I	Times	1785	08.04.2016	Ann Treneman	Female	THE CARETAKER	Harold Pinter	None	It's an exotic	Old Vic Theatre
Corpus I	Daily Telegraph	1855	08.04.2016	Dominic Cavendish	Male	THE CARETAKER	Harold Pinter	None	The last time	Old Vic Theatre
Corpus I	Daily Mail	1896	08.04.2016	Patrick Marnion	Male	THE CARETAKER	Harold Pinter	None	A slimline Ti	Old Vic Theatre
Corpus I	Spectator	1828	09.04.2016	Lloyd Evans	Male	THE CARETAKER	Harold Pinter	None	Eager crowds	Old Vic Theatre
Corpus II	Rev Stan Blog	2007	09.04.2016	Rev Stan	Female	THE CARETAKER	Harold Pinter	None	Timothy Spal	Old Vic Theatre
Corpus I	Observer	1791	10.04.2016	Susannah Clapp	Female	THE CARETAKER	Harold Pinter	None	Here is an ide	Old Vic Theatre
Corpus I	Sunday Times	1821	10.04.2016	Maxie Szalwinska	Female	THE CARETAKER	Harold Pinter	None	The role of D	Old Vic Theatre
Corpus I	Sunday Express	1918	10.04.2016	Michael Arditi	Male	THE CARETAKER	Harold Pinter	None	WHEN Harol	Old Vic Theatre
Corpus I	Financial Times	1888	11.04.2016	Sarah Hemming	Female	THE CARETAKER	Harold Pinter	None	If only it wou	Old Vic Theatre
Corpus I	Time Out	1968	12.04.2016	Andrzej Lukowski	Male	THE CARETAKER	Harold Pinter	None	Returning to	Old Vic Theatre
Corpus II	Kate in Brockley Blog	2011	12.04.2016	Kate Gould	Female	THE CARETAKER	Harold Pinter	4	THE minute	Old Vic Theatre
Corpus I	Jewish Chronicle	1841	14.04.2016	John Nathan	Male	THE CARETAKER	Harold Pinter	None	The plays of	Old Vic Theatre
Corpus I	Stage	1880	14.04.2016	Natasha Tripney	Female	THE CARETAKER	Harold Pinter	None	This is comic	Old Vic Theatre
Corpus I	Daily Express	1900	15.04.2016	Neil Norman	Male	THE CARETAKER	Harold Pinter	None	IT is said that	Old Vic Theatre
Corpus I	Mail on Sunday	1982	17.04.2016	Georgina Brown	Female	THE CARETAKER	Harold Pinter	None	Comparisons	Old Vic Theatre
Corpus II	Execunt Blog	2010	18.04.2016	Tim Bam	Male	THE CARETAKER	Harold Pinter	None	Spall can't st	Old Vic Theatre

Fig. 18 : Exemple d'une partie de notre corpus une fois structuré.

II.B.1.2 Ajout des métadonnées

Une métadonnée est une donnée qui sert à décrire ou à compléter une autre donnée principale. Elle permet de mieux comprendre le contexte d'un corpus et d'enrichir les analyses statistiques en ajoutant d'autres types d'informations. Dans notre cas, un nombre de métadonnées était déjà inclus dans les critiques que nous avons récupérées. Nous les avons donc conservées et les avons intégrées à notre matrice de données. Afin d'approfondir certaines de nos analyses, nous avons ajouté d'autres types d'informations. La liste ci-dessous répertorie les métadonnées de notre corpus en détaillant celles qui étaient déjà présentes au moment de la récupération des données, et celles que nous avons ajoutées :

Métadonnées déjà incluses dans nos sources au moment de leur récupération :

- Le type de corpus (journalistique / numérique)
- Le nom du journal / du blog
- La date de publication de la critique
- Le nom et le prénom du critique
- Le titre de la pièce
- Le nom et le prénom du dramaturge (lorsque ces informations étaient disponibles)
- La note attribuée par le critique (lorsque ces informations étaient disponibles)
- Le lieu où a été jouée la pièce

Métadonnées non incluses dans nos sources que nous avons ajoutées nous-mêmes, manuellement ou automatiquement :

- La date de création du journal / du blog

Cette information a été ajoutée afin de pouvoir hiérarchiser les journaux et les blogs de notre corpus en fonction de leur date de création.

- Le genre de l'auteur de la critique

Nous avons déterminé automatiquement le genre des auteurs des critiques grâce à la librairie « GenderizeR » avec R. Cette information permet ensuite d'effectuer des analyses en fonction du sexe des critiques.

- Les coordonnées géographiques des théâtres

Les coordonnées géographiques des théâtres ont été récupérées automatiquement grâce à des techniques de web scraping. Elles sont utiles dans le cadre de l'utilisation des Systèmes d'Informations Géographiques. Dit autrement, elles permettent de générer des cartes automatiquement.

- Le lien URL du lieu où a été jouée la pièce

Nous avons récupéré l'adresse web des 446 lieux de représentation mentionnés dans notre corpus à la main. Ces métadonnées permettent non seulement de créer une carte géographique plus détaillée en ajoutant le visuel de l'architecture des différents lieux, mais aussi d'apporter davantage d'informations quant à la politique culturelle de ces lieux.

- Le type de lieu dans lequel a été jouée la pièce¹¹⁸

Nous avons classé ces 446 lieux en 24 catégories. Dans la mesure où les théâtres ne représentent qu'une partie des endroits visités par les critiques, nous avons répertorié le reste de ces lieux en différentes catégories (musées, institutions culturelles, restaurants, etc.). Cela permet d'analyser plus finement la répartition de ces deux communautés dans Londres.

- Les capacités d'accueil du lieu où a été jouée la pièce

Enfin, la dernière métadonnée que nous avons ajoutée à notre corpus concerne le nombre de places par lieu de représentation. Nous avons également ajouté à la main cette information lorsqu'elle était disponible sur le site web de l'endroit en question. Cette métadonnée permet de mieux comprendre quels types d'endroits sont fréquentés par les journalistes et les blogueurs.

¹¹⁸ La liste de toutes les catégories de lieux est disponible en annexes p. 240 – 250.

II.B.1.3 Application d'outils en Traitement Automatique des Langues (TAL)

Les métadonnées ajoutées, l'étape suivante consiste à appliquer des outils en Traitement Automatique des Langues (TAL) afin de rendre les textes intelligibles pour l'ordinateur. Bien qu'une machine soit puissante, elle ne peut analyser un jeu de données de manière autonome. C'est donc à l'humain, et en l'occurrence au chercheur, d'appliquer une série d'algorithmes aux textes afin de les rendre manipulables d'un point de vue computationnel. Nous avons utilisé la librairie SpaCy¹¹⁹ avec Python, qui est une des librairies les plus utilisées par la communauté scientifique pour réaliser ce genre de tâche. Cinq étapes sont nécessaires pour rendre notre corpus analysable par la machine. Nous décrivons chacune d'entre elles ci-dessous :

Etape 1 : La tokenisation

La tokenisation consiste à découper un texte en « tokens », c'est-à-dire en différentes unités qui ont une signification sémantique (mots, phrases, paragraphes, etc.).

Etape 2 : La lemmatisation

La lemmatisation permet de normaliser chacune des unités linguistiques en leur attribuant leur forme fléchie, c'est-à-dire en leur attribuant leur lemme. Cette étape facilite le rapprochement des mots et permettra ainsi d'affiner nos analyses computationnelles. Pour le verbe « to perform » par exemple, ses différentes formes (« performed », « performs », « performing ») seront automatiquement remplacées par le lemme « perform ». Lorsque nous effectuerons des analyses sur le style des auteurs par exemple, l'ordinateur repérera le nombre de fois que le verbe « perform » apparaîtra, plutôt que le nombre d'occurrences de « performed », « performs » ou « performing. »

Etape 3 : Etiquetage morpho-syntaxique

La troisième étape apporte une autre donnée informative aux unités linguistiques. L'étiquetage morpho-syntaxique attribue à chaque « token » sa classe grammaticale (adverbe,

¹¹⁹<https://spacy.io/>[En ligne], consulté le 09 novembre 2021.

verbe, nom...) ainsi que son genre et son nombre. Cela nous permettra par exemple, grâce à l'aide d'un script, de faire une distinction entre le verbe « to play », qui fait référence à l'action de jouer, et le nom « a play », qui désigne une pièce de théâtre. Cette désambiguïté lexicale nous permettra de conduire des analyses plus fines sur notre corpus.

Etape 4 : Analyse en dépendance

L'analyse en dépendance a pour finalité d'annoter automatiquement les relations syntaxiques entre les mots dans une phrase. Cela permet notamment de comprendre le fonctionnement interne d'une phrase et d'avoir un aperçu de la manière dont sont organisés les discours d'un point de vue syntaxique.

Etape 5 : Reconnaissance des entités nommées

Une entité nommée est une expression linguistique qui permet d'affecter une catégorie, comme les personnes, les lieux ou les organisations, aux noms. Le « Royal Court Theatre » sera par exemple automatiquement annoté comme un « lieu ». « Michael Billington » ou « Lyn Gardner » seront, quant à eux, labellisés comme « personne ».

Transition

Cette partie nous a permis de présenter les dernières étapes nécessaires pour examiner notre corpus grâce aux algorithmes. Les analyses computationnelles requièrent un long travail d'organisation des données. Si l'ordinateur est une machine puissante, il n'est cependant pas capable de comprendre un texte de manière autonome. C'est pourquoi nous devons structurer les informations en une matrice de données, ou en un tableau, avant d'appliquer des outils en Traitement Automatique des Langues qui permettront de rendre nos critiques intelligibles pour l'ordinateur.

Dans la seconde sous-partie suivante, nous nous focaliserons sur la présentation de nos deux sous-corpus. Nous commencerons par aborder notre base de données dans sa globalité (nombres de mots, répartition des critiques dans le temps, top 10 des journaux et blogs les plus représentés, etc.). Nous aborderons ensuite nos deux sous-corpus individuellement. Après avoir présenté les principales caractéristiques de la critique journalistique, nous effectuerons le même travail avec la critique numérique. Cette partie nous permettra de voir en quoi nos deux sous-corpus représentent deux formes de réception distinctes.

II.B.2 Présentation du corpus

II.B.2.1 Deux sous-corpus équilibrés

Après plus de 1 400 heures de travail, ce sont très précisément 43 043 critiques qui ont été réunies et qui forment le matériau premier sur lequel nous avons fondé nos analyses. Ces critiques se répartissent de manière presque égale entre les deux corpus, puisque le premier sous-corpus est composé de 21 717 critiques, soit 50,5% du corpus total, tandis que le second est constitué de 21 326 critiques, soit 49,5% du corpus total. Le tableau ci-dessous répertorie les caractéristiques principales que nous venons de présenter :

	Corpus I	Corpus II	Corpus I et II
Nombre de critiques	21 717	21 326	40 043
Nombres de mots	7 689 704	11 416 428	19 106 132

Fig. 19 : Caractéristiques principales du corpus.

Cette base de données restitue les critiques portant sur 9 020 pièces différentes qui ont toutes été jouées à Londres entre 2010 et 2020. Ces critiques ont été rédigées par 1 069 auteurs différents, journalistes et blogueurs inclus. Au total, ce sont 51 médias, journaux et blogs compris, qui sont représentés. La répartition du genre des critiques est relativement équilibrée, même si nous pouvons observer un léger avantage pour les critiques de sexe masculin. Nous pouvons en effet compter 54,1% d'hommes contre 45% de femmes. Le 0,9% restant n'a pas pu être déterminé car les prénoms étaient soit d'origine étrangère, soit mixtes.

Pour les cinq premiers auteurs les plus influents de ce corpus, quatre d'entre eux sont des blogueurs (Nick, Ian, Libby Purves et Laura Kressly). Michael Billington, critique de longue date dans le journal *The Guardian* et figure reconnue de la critique théâtrale anglaise, se place en troisième position parmi ceux qui ont publié le plus d'articles. La tendance s'inverse pour les cinq derniers auteurs les plus prolifiques de ce corpus. Tandis que quatre d'entre eux sont des journalistes (Henry Hitchings, Paul Taylor, Andrzej Lukowski et Sarah Hemming), un seul est à la fois blogueur et journaliste (Aleks Sierz). Les dix auteurs les plus représentés de ce corpus se partagent donc de manière égale : cinq journalistes d'un côté, et cinq blogueurs de l'autre.

Ce rapport égalitaire caractérise également la répartition des journaux et des blogs. Les trois premiers médias qui comptent le plus grand nombre de critiques sont trois blogs. Ces trois sites (Exeunt.com, AyoungerTheatre.com et EverythingTheatre.com) sont des plateformes collectives, ce qui explique leur grand nombre de publications.

Enfin, concernant la répartition des critiques dans le temps, nous pouvons observer deux tendances. La première concerne le nombre de critiques au début et à la fin du cadre temporel que nous avons choisi pour ce corpus. Tandis qu'au début des années 2010, il y a quasiment presque six fois plus de critiques dans le sous-corpus I que dans le sous-corpus II (2 432 critiques journalistiques contre 426 critiques numériques), les deux communautés présentent un nombre de critiques presque similaire à la fin des années 2020. Cet écart important au début de la chronologie se justifie par le fait que l'émergence de la blogosphère n'en est qu'à ses débuts.

Cinq blogs seulement sont répertoriés en 2010 dans notre base de données. C'est seulement à partir de 2014 qu'un nombre conséquent de plateformes en ligne commence à apparaître (18 au total dans notre corpus), ce qui correspond à un pic du nombre de critiques. Les années 2020 étant marquées par la fermeture des théâtres à cause de la Covid 19, le nombre de critiques dans les deux sous-corpus chute drastiquement (815 critiques pour le sous-corpus I, et 1 030 critiques pour le sous-corpus II).

Ces analyses reposent sur le tableau de bord créé à partir du corpus complet. Sa version interactive est disponible sur <https://dramacritiques.com/fr/categories/le-corpus/>. Elle permet notamment d'obtenir davantage d'informations sur un journal ou un blog particulier en cliquant directement sur son nom. Sa version image est présentée en figure 20 à la page suivante.

Critiques journalistiques et numériques (2010 - 2020) : Tour d'horizon



Fig. 20 : Tableau de bord des caractéristiques principales de l'ensemble du corpus.

II.B.2.2 La critique journalistique en quelques chiffres

Notre premier sous-corpus est constitué de 23 journaux. Tous ces organes de presse sont connus du grand public et forment ce que nous pourrions appeler le « canon » de la critique théâtrale anglaise. Une moitié d'entre eux apparaît au XIX^e siècle (11 journaux), et l'autre au XX^e siècle (10 journaux). Seulement deux voient le jour au XVIII^e siècle (*The Times* et *The Observer*).

Ce premier sous-corpus réunit 454 journalistes différents, dont 63,6% d'hommes et 36,4% de femmes. Des dix journalistes les plus prolifiques, nous pouvons remarquer que sept d'entre eux sont des hommes, contre trois femmes seulement. Les cinq premiers journaux qui comportent le plus grand nombre de critiques sont les suivants : *Time Out*, *The Times*, *The Guardian*, *The Evening Standard* et *The Daily Telegraph*.

Le premier tableau en figure 21 répertorie ces 23 journaux en fonction de leur nombre de critiques. Leur date de création, leur éditeur ainsi que leur nombre de journalistes respectifs sont aussi indiqués. Le deuxième tableau présenté en figure 22 est une représentation visuelle des chiffres que nous avons introduits plus haut. Leur version interactive est disponible sur le site de Drama Critiques, à l'adresse suivante : <https://dramacritiques.com/fr/categories/le-corpus/>. Les utilisateurs peuvent directement sélectionner un journal en particulier et ses caractéristiques correspondantes s'afficheront (répartition du genre des critiques, les dix journalistes / journaux les plus influents, répartition des critiques dans le temps, etc.).

Journaux	Création	Editeur	Journalistes	Critiques
TIME OUT	1968	Tony Elliott	92	2 333
THE TIMES	1785	John Walter	28	2 284
THE GUARDIAN	1821	John Edward Taylor	46	2 176
THE EVENING STANDARD	1859	Stanley Lees Giffard	22	1 874
THE DAILY TELEGRAPH	1855	Arthur B. Sleigh	43	1 509
THE FINANCIAL TIMES	1888	James Sheridan	28	1 344
THE SUNDAY TIMES	1821	Henry White	36	1 216
THE INDEPENDENT	1986	Andreas Whitam Smith	42	1 193
THE DAILY MAIL	1896	Alfred Harmsworth	19	1 026
THE OBSERVER	1791	W.S. Bourne	35	935
METRO (LONDON)	1999	Ted Young	27	793
THE STAGE	1880	Charles Lionel Carson	62	792
THE SUNDAY EXPRESS	1918	Arthur Pearson	12	711
THE MAIL ON SUNDAY	1982	Lord Rothermere	13	707
THE SPECTATOR	1828	Robert Stephen Rintoul	6	668
THE DAILY EXPRESS	1900	Arthur Pearson	19	588
THE JEWISH CHRONICLE	1841	Stephen Pollard	3	409
THE INDEPENDENT ON SUNDAY	1990	Stephen Glover	7	319
THE HERALD TRIBUNE	1887	James Gordon Bennett	13	317
THE SUNDAY TELEGRAPH	1918	Arthur Pearson	3	261
THE TRIBUNE	1937	George Strauss	5	220
THE TELEGRAPH	1855	Colonel Arthur B Sleigh	2	27
VARIETY	1905	Sime Silverman	2	25

Fig. 21 : Caractéristiques principales du corpus journalistique.

Corpus I : La Critique journalistique (2010 - 2020)

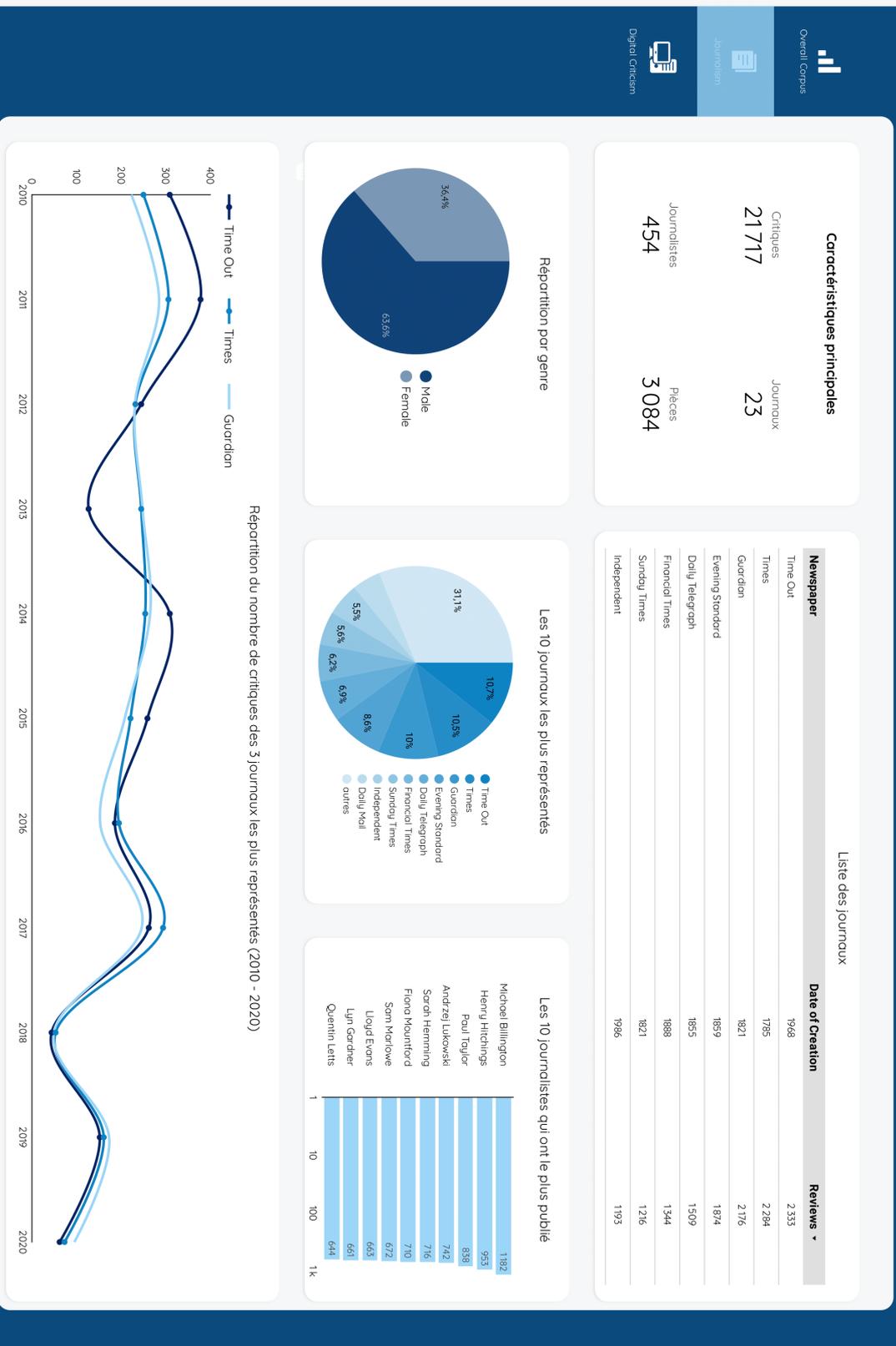


Fig. 22 : Tableau de bord des caractéristiques principales du corpus I (journalistique).

II.B.2.3 La critique numérique en quelques chiffres

Notre second sous-corpus est constitué de 28 sites Internet au total, plateformes collectives et blogs personnels inclus¹²⁰. Tous les blogs de notre base de données ont été créés dans les années 2000. Bien que ces deux bases de données aient un nombre de critiques similaires (21 717 dans le corpus I et 21 326 dans le corpus II), nous pouvons remarquer qu'il y a davantage de contributeurs dans le corpus numérique (655 blogueurs différents contre 454 journalistes), et plus de deux fois plus de pièces différentes analysées (8 351 pièces commentées par les blogueurs contre 3 084 pour les journalistes). La répartition du genre des critiques est, à l'inverse du premier corpus, à l'avantage du sexe féminin. Ce sont 53,7% des critiques qui sont des femmes, contre 43,3% d'hommes, le pourcentage restant n'ayant pas pu être calculé dû à l'origine ou à la mixité des prénoms. Ce rapport égalitaire entre le genre des blogueurs se retrouve également dans le classement des dix blogueurs les plus influents, puisque cinq d'entre eux sont des femmes et les cinq autres des hommes. Concernant les blogs et les plateformes qui comptabilisent le plus grand nombre de critiques, les trois premiers sont des plateformes collectives. Le reste concerne des blogs personnels, excepté WestEndWilma.com, en neuvième position, qui réunit plusieurs blogueurs.

Enfin, la répartition des critiques dans le temps des trois plateformes les plus influentes de ce sous-corpus présente une caractéristique notable pour les années 2010 à 2014. Pour ExeuntMagazine.com et Everything-Theatre.co.uk, les données que nous avons collectées commencent un an après la date d'apparition de ces deux plateformes. Il manque donc une année de critiques pour chacune d'entre elles. Concernant AyoungerTheatre.com, nous pouvons également observer une absence de critiques de 2011 à 2013. Dans les deux cas, cela s'explique par le fait qu'il était difficile de récupérer automatiquement ces publications car la structure HTML de ces sites changeait considérablement. Au lieu d'effectuer le travail manuellement, nous avons jugé qu'il était plus judicieux de les laisser de côté. Les pages suivantes répertorient chacun des sites Internet que nous avons intégrés à ce sous-corpus témoin. A nouveau, ces sites comportent ces trois caractéristiques communes :

1. Ils ne possèdent pas de version papier
2. Leur contenu est entièrement gratuit
3. Les auteurs de ces critiques ne sont pas rémunérés pour leurs publications

¹²⁰ Le tableau de bord (Fig. 23) sur lequel reposent les analyses que nous allons présenter se situe à la page 118.

Tous proviennent soit du classement des meilleurs sites de théâtre anglais établi par Vuelio en 2020, soit de la plateforme MyTheatreMates.com. Leur éditeur, leur date de création, leur nombre de critiques et de blogueurs sont indiqués dans chacune de leur description. Nous les avons classés en fonction de deux critères : selon la catégorie à laquelle ils appartiennent (plateforme collective ou blog personnel) et selon leur nombre de critiques. Pour chacun de ces 28 sites, nous avons pris soin de rédiger une brève description de leur éditeur, de leur ligne éditoriale et de leur contenu. Toutes ces informations sont extraites soit de leur blog personnel, soit de divers réseaux sociaux (Facebook, LinkedIn, Twitter, Instagram et YouTube).

LES PLATEFORMES COLLECTIVES

A Younger Theatre (2009)

<https://www.ayoungertheatre.com/>

2 747 critiques répertoriées dans notre corpus

185 blogueurs différents

A Younger Theatre a été fondé en 2009 par le producteur de théâtre Jake Orr. Cette plateforme collective est désormais administrée par Samuel Sims, écrivain et journaliste, qui se définit comme étant particulièrement sensible aux problématiques LGBTQ+ et à la pop culture. AYT a la particularité d'imposer une limite d'âge à ses blogueurs. Seules les personnes âgées entre 16 et 35 ans peuvent soit aider à la gestion de la plateforme, soit publier du contenu théâtral. L'identité numérique d'AYT se définit en opposition aux discours des médias traditionnels :

[AYT is] a platform for those who are often unheard. We champion the emerging generation. We celebrate excellence and pride ourselves on professionalism, while also leaving room for risk, failure and learning¹²¹.

En imposant une limite d'âge, AYT favorise ainsi un jeune public d'auteurs qui est encore novice dans l'art de la critique. Sa page de recrutement met également explicitement à l'honneur sa politique d'inclusion :

Are you a bit bored of seeing the same old 'traditional' reviews and want to shake things up a bit? (...) We are also keen to ensure that our writing team is representative of the society that we live in and so, we particularly welcome applications from people of colour, people with disabilities and the LGBTQ+ community¹²².

¹²¹<https://www.ayoungertheatre.com/about/> [En ligne], consulté le 19 juillet 2022.

¹²²<https://www.ayoungertheatre.com/jobs/> [En ligne], consulté le 19 juillet 2022.

Exeunt blog (2010)

<http://exeuntmagazine.com/>

2 820 critiques répertoriées dans notre corpus

146 blogueurs différents

Co-fondé par Natasha Tripney (éditrice pour le journal *The Stage* et journaliste freelance pour *The Guardian*) et Daniel B. Yates (ancien étudiant à la London School of Economics et à l'Université de Goldsmiths) en 2009, Exeunt entend également contrer le classicisme de la critique théâtrale journalistique :

Exeunt believes in making beautifully written, experimental, fierce and longform writing about theatre available for free. We give critics the freedom to expand beyond the tight deadlines and brutal word counts of more established media outlets¹²³.

Everything Theatre (2011)

<https://everything-theatre.co.uk/>

2 396 critiques répertoriées dans notre corpus

128 blogueurs différents

Everything Theatre fut créé en 2011 par deux étudiants londoniens qui commencèrent à publier en ligne les critiques des pièces pour lesquelles ils avaient obtenu des réductions. A partir de 2014, le blog migra vers Wordpress afin d'accueillir d'autres blogueurs qui désiraient eux aussi publier leurs critiques. Ne pas avoir de connaissances préalables sur le théâtre est qualifié comme « une force » dans leur page de présentation :

What if I know nothing technical about theatre? It doesn't matter. Our aim is to publish honest and *unpretentious writing* that *everybody can understand*. We're not a site aimed only at the theatre-buffs, we're a site for everybody. We don't mind if you don't know any *fancy theatrical terms*, as long as you can explain why you enjoyed (or didn't) what you've watched. Plus we'll help you along the way with advice on what makes a good review and what you should be looking for. You might start with limited knowledge but you'll be amazed how much you pick up when you start to look out for it¹²⁴.

¹²³<http://exeuntmagazine.com/about-exeunt/> [En ligne], consulté le 19 juillet 2022.

¹²⁴<https://everything-theatre.co.uk/join-our-team> [En ligne], consulté le 19 juillet 2022.

West End Wilma (2012)

<https://www.westendwilma.com/>

759 critiques répertoriées dans notre corpus

76 blogueurs différents

Premier du classement Vuelio 2020, West End Wilma se présente comme une plateforme familiale créée par « Ted et sa mère. » Leur description met en avant l'honnêteté de leurs publications. Selon eux, West End Wilma ne « joue pas aux mêmes règles que les autres sites Internet » puisqu'ils refusent de donner une note élevée à un spectacle s'il ne le mérite pas. Soutenu par un public particulièrement enthousiaste et fidèle, West End Wilma a même créé depuis 2014 sa propre cérémonie annuelle afin de décerner le « Wilma Awards » aux personnalités les plus méritantes du théâtre londonien. Depuis 2018, les « cérémonies » sont disponibles sur leur chaîne Youtube¹²⁵.

View from the Cheap Seat (2015)

<https://viewfromthecheapseat.com/>

509 critiques répertoriées dans notre corpus

25 blogueurs différents

Le titre View From the Cheap Seat est inspiré du type de places que Shanine Salmon obtenait lorsqu'elle a commencé à se rendre au théâtre en 2011. Ouverte à toutes les formes théâtrales, cette plateforme était originellement un blog personnel tenu par Salmon afin de garder une trace des spectacles auxquels elle assistait. Un nombre de contributeurs l'ont rejointe depuis.

London Theatre Reviews (2018)

<https://www.londontheatreviews.co.uk/>

459 critiques répertoriées dans notre corpus

34 blogueurs différents

¹²⁵<https://www.youtube.com/playlist?list=PLsR2bqLO2Y4hL1WaiI23m8Q89QuFMhw-b>[En ligne], consulté le 19 juillet 2022.

London Theatre Reviews a été créé par Luca Molinari en 2018. Aucune description de la plateforme n'est disponible. Son sous-titre cependant, « Simply and honestly », rappelle les valeurs mises en avant par les autres blogs. London Theatre Reviews a reçu deux prix : le Theatre & Technology Award en 2019, pour la meilleure publication en ligne, et le Online Influence Award en 2020.

Musical Theatre (2013)

<https://musicaltheatrereview.com/>

294 critiques répertoriées dans notre corpus

33 blogueurs différents

Lisa Martland a obtenu une licence en études cinématographiques à l'Université de Kent avant de passer une année à l'Ecole de journalisme de Cardiff. Elle est aujourd'hui rédactrice en chef à Synthesis Media, une maison d'édition spécialisée en sciences dans le nord de Londres. Martland crée la plateforme Musical Theatre en 2013 qui fut classée quatrième parmi les sites les plus influents de l'année 2020 par Vuelio. Elle couvre toutes les comédies musicales du London Fringe et du West End, mais également celles jouées à Broadway et à Edinburgh. Les sections « Interviews » permettent d'avoir de précieuses informations sur les acteurs, mais aussi sur les paroliers, les librettistes et les musiciens. Les productions menées par les étudiants sont également incluses au même titre que celles réalisées par les professionnels.

Rewrite this Story (2015)

<https://www.rewritestory.com/>

320 critiques répertoriées dans notre corpus

25 blogueurs différents

Rewrite This Story a été créé par Olivia Mitchell en 2015. Comme View From the Cheap Seat, cette plateforme était initialement administrée par Mitchell uniquement, et conçue comme un espace dans lequel elle pouvait partager ses « pensées » librement (« Olivia

begun site in late 2015 as a place to share her thoughts on theatre just for fun¹²⁶. »). Depuis, une vingtaine de blogueurs venant du Royaume-Uni mais aussi d’Australie et des Etats-Unis l’ont rejointe. RTS a gagné deux prix, le premier en 2017 pour le meilleur blog de critiques théâtrales, et le second en 2018 pour le meilleur blog d’art et de culture.

Pocket Size Theatre (2012)

<http://www.pocketsizetheatre.com/>

234 critiques répertoriées dans notre corpus

23 blogueurs différents

Pocket Size Theatre a été créé en 2012 par « Jay. » Cette plateforme a été classée parmi les dix blogs de théâtre les plus influents en 2022 par Feedspot¹²⁷ et premier par Score Big dans la même catégorie¹²⁸. Le sous-titre de leur plateforme, « The Independent Theatre Website » est explicite quant à leur positionnement dans la communauté théâtrale londonienne. La description de leur contenu numérique est également parlante : « All reviews are *simply opinions* of those who attend the show¹²⁹. » La qualification de leurs publications comme « simples opinions » rappelle la distinction établie par les critiques professionnels entre « critique légitime » et « opinion amatrice » écrites par les blogueurs.

¹²⁶[Who Am I? | Rewrite This Story](#)[En ligne], consulté le 19 juillet 2022.

¹²⁷https://blog.feedspot.com/theatre_blogs/[En ligne], consulté le 19 juillet 2022.

¹²⁸<https://www.scorebig.com/home/top-50-theater-blogs>[En ligne], consulté le 19 juillet 2022.

¹²⁹[About | Pocket Size Theatre](#) [En ligne], consulté le 19 juillet 2022.

LES BLOGS PERSONNELS

There Ought to be Clowns, Don't Bother, They are Here (2003)

<https://oughttobeclowns.com/>

1784 critiques répertoriées dans notre corpus

There Ought To Be Clowns a été classé parmi les 10 meilleurs blogs de théâtre au Royaume-Uni par LastMinute, Feedspot, Vuelio et Superbreak. Sa page « à propos » est lacunaire. Ian Foster, son éditeur, préfère rester discret quant à son identité. Il qualifie cependant son blog de « journal intime portant sur le théâtre » (*theatrical diary*)¹³⁰, une dénomination utilisée par Michael Billington pour faire la distinction entre une « critique légitime » rédigée par un journaliste, et celle écrite par un blogueur¹³¹.

Partially Obstructed View (2012)

<http://partially-obstructed-view.blogspot.com/>

1 581 critiques répertoriées dans notre corpus

Partially Obstructed View a été créé par « Nick » en 2012. Très discret quant à son identité, Nick est le seul blogueur qui a explicitement refusé que ses données soient mises en accès libre. C'est pourtant, des deux corpus confondus, celui qui a publié le plus de critiques. Il se définit comme quelqu'un ayant un diplôme en littérature théâtrale anglaise, même s'il ne travaille pas dans ce milieu. Probablement conscient de la discorde entre le journalisme et la critique numérique, Nick admet que si ses connaissances littéraires ne sont pas omniscientes, il possède néanmoins un bagage culturel qui lui donne la légitimité d'écrire, bien que « tout le monde ne soit pas d'accord avec lui sur ce point¹³². »

¹³⁰<https://oughttobeclowns.com/about-me/> [En ligne], consulté le 19 juillet 2022.

¹³¹ Michael BILLINGTON, « Who needs reviews?, *The Guardian* [En ligne], publié le 17 septembre 2007, URL :<https://www.theguardian.com/stage/theatreblog/2007/sep/17/whoneedsreviews>, consulté le 19 juillet 2022.

¹³²<http://partially-obstructed-view.blogspot.com/p/about-me.html> [En ligne], consulté le 19 juillet 2022.

The Plays The Thing UK (2015)

<https://theplaysthethinguk.com/>

1 052 critiques répertoriées dans notre corpus

The Plays The Thing UK a été créé par Laura Kressly, première auteure la plus influente de ce second sous-corpus, en 2015. Kressly vient de la périphérie de New York. Après s'être spécialisée en *Shakespeare Performance* à l'Université d'Exeter, elle emménagea à Londres afin de poursuivre une carrière d'actrice. Elle est aujourd'hui en train de terminer un doctorat en études théâtrales à la Central School of Speech and Drama, à Londres. Son blog a pour vocation de mettre en lumière les théâtres périphériques qui sont, selon ses propres mots, « peu commentés par les critiques journalistiques¹³³. »

Aleks Sierz Blog (2006)

<https://www.sierz.co.uk/>

828 critiques répertoriées dans notre corpus

Aleks Sierz se démarque des autres éditeurs de ce corpus témoin puisque ses travaux sont reconnus dans le monde académique. Auteur d'un certain nombre d'essais sur le théâtre contemporain britannique¹³⁴, il a également publié de nombreuses critiques dans *The Tribune*, *The Art Desk* et *The Independent* et a enseigné l'histoire théâtrale à l'Université de Boston ainsi que dans plusieurs universités allemandes. Depuis 2006, il poste régulièrement des critiques personnelles sur son blog et est membre de MyTheatreMates.com, plateforme sur laquelle sont republiées ses publications.

Theatre Cat (2013)

<https://theatre.cat.com/>

773 critiques répertoriées dans notre corpus

¹³³<https://theplaysthethinguk.com/about-me/> [En ligne], consulté le 19 juillet 2022.

¹³⁴*In-Yer-Face Theatre: British Drama Today* (2001), *The Theatre of Martin Crimp*(2006), *John Osborne's Look Back in Anger* (2008), *Rewriting the Nation: British Theatre Today* (2011) et *Modern British Playwriting: The 1990s*(2012).

Libby Purves est une autre blogueuse dont les publications se démarquent des autres critiques de ce corpus. De 2010 à 2013, Purves était critique en cheffe de la catégorie « Théâtre » dans le journal *The Times*. La vague de messages qu'elle reçut en septembre 2013 après l'annonce de son renvoi l'encouragea à créer ce blog personnel afin de continuer d'écrire. A la différence des autres blogueurs, Purves affirme que ses critiques respecteront les conventions journalistiques : « reasonably brief, tightly edited, informative, sometimes perhaps amusing¹³⁵. » L'utilisation d'illustrations humoristiques de chats pour remplacer le traditionnel système d'étoiles suggère néanmoins sa prise de recul face à la critique classique.

Life in the Cheap Seats - Webcow Girl (2003)

<https://webcowgirl.wordpress.com/>

610 critiques répertoriées dans notre corpus

Webcow Girl a été créé en 2003 par une blogueuse américaine expatriée à Londres qui préfère garder l'anonymat. Le sous-titre de son blog, « From the Cheap Seats », fait écho au site de Shanine Salmon. Comme pour beaucoup de blogueurs, l'absence de rémunération de leur activité de critique peut s'avérer problématique. C'est pourquoi de nombreuses astuces sur les tickets bon marché sont données au fil de ses publications, afin de permettre au plus grand nombre d'assister à des spectacles¹³⁶. L'honnêteté de ses critiques est également mise en exergue dans sa présentation :

Just be warned that I'm as likely to post negative things as positive about your production, and if you give me comps, you understand this is not a "you scratch my back, I scratch yours" proposition¹³⁷.

Breaking the Fourth Wall (2017)

<https://breaking-the-fourth-wall.com/>

460 critiques répertoriées dans notre corpus

¹³⁵<https://theatre.cat/about/>[En ligne], consulté le 19 juillet 2022.

¹³⁶<https://webcowgirl.wordpress.com/about/>[En ligne], consulté le 19 juillet 2022.

« I also put in information about good deals on shows I think are worth catching. I feel the cost of theater is what keeps a lot of people from going and I want everyone and anyone to feel like they can afford to attend, even if doesn't mean watching from the stalls. »

¹³⁷*Ibid.*

Breaking the Fourth Wall a été créé par Michael Davis qui est un ancien acteur et producteur de théâtre. Davis est particulièrement intéressé par tous les genres de théâtre qui échappent au canon littéraire et qui reçoivent peu d'attention de la part des journaux¹³⁸. Avant de créer son blog personnel, il a publié dans The Plays The Thing UK, édité par Laura Kressly et présenté dans la section « plateformes collectives », et a également écrit pour le journal Female Arts.

Stage Review (2014)

<https://www.stagereview.co.uk/>

444 critiques répertoriées dans notre corpus

Stage Review a été créé par Anne Cox en 2014. Cox se définit elle-même comme journaliste et blogueuse à la fois. Son site a été élu deux fois par Vuelio parmi les blogs les plus influents de la toile britannique. Son profil hybride se distingue de certains des autres blogueurs faisant partie de ce corpus. Bien que son site soit indépendant, ses publications sont néanmoins réservées à un théâtre qu'elle identifie comme « professionnel¹³⁹. »

Monkey Matters (2011)

<http://www.monkeymatterstheatre.com/>

442 critiques répertoriées dans notre corpus

Monkey Matters a été créé par Shyama Perera en 2011. Elle est auteure de romans, animatrice à la radio et a écrit pour plusieurs journaux locaux et pour la BBC de Londres. Perera a également fait partie du jury des Olivier Awards, l'équivalent britannique des Molières français. Elle ne publie plus depuis janvier 2018 dans son blog, mais ses critiques sont toujours disponibles.

West End Whingers (2006)

<https://westendwhingers.wordpress.com/>

429 critiques répertoriées dans notre corpus

¹³⁸<https://mytheatremates.com/author/michael-davis/> [En ligne], consulté le 19 juillet 2022.

¹³⁹<https://www.stagereview.co.uk/about/> [En ligne], consulté le 02 août 2022.

West End Whingers est un des premiers blogs à avoir réécrit ouvertement les codes de la critique littéraire. Comme ses deux auteurs, Phil Argent et Andrew Rogers, l'affichent explicitement sur chacune de leur page, leur contenu est « rempli d'opinions triviales. » (*Full of trivial opinions*). Leurs critiques à l'humour grinçant rappellent le journalisme gonzo.

Rev Stan (2007)

<https://theatre.revstan.com/>

412 critiques répertoriées dans notre corpus

Editrice et rédactrice indépendante, Stacey Meadwell a créé son blog personnel en 2007. L'honnêteté de son travail est, comme pour beaucoup d'autres blogueurs, d'une grande importance dans la rédaction de ses critiques : « I'm a fan but not a fan-girl, so I will highlight flaws, regardless of how much I love a particular actor/director/writer.¹⁴⁰ » Meadwell est une des rares blogueuses de ce corpus à être aussi active sur Facebook. Elle publie de courtes vidéos qui durent entre une et trois minutes dans lesquelles elle analyse diverses productions théâtrales. Son blog a été élu parmi les plus influents de la toile britannique en 2020 par Vuelio.

Mind The Blog (2014)

<https://pleasemindtheblog.wordpress.com/>

409 critiques répertoriées dans notre corpus

Debbie Gilpin a créé Mind the Blog en 2014, lorsqu'elle emménagea à Londres après avoir obtenu un diplôme en génétique humaine à l'Université de Newcastle. Passionnée par les pièces de Shakespeare, elle s'intéresse également aux nouvelles formes d'écriture. Son blog fut nommé en 2019 par les UK Blog Awards dans deux catégories : « Arts & Culture » et « Social Influencer. »

The Blog of Theatre Things (2015)

¹⁴⁰<https://www.vuelio.com/uk/social-media-index/top10-theatre-blogs/>[En ligne], consulté le 02 août 2022.

<https://theblogoftheatrethings.com/>

377 critiques répertoriées dans notre corpus

Avant de créer son blog personnel en 2015, Liz Dyer écrivait déjà pour diverses plateformes théâtrales en ligne, comme London Theatre Direct¹⁴¹, London Theatre 1¹⁴² ou Carn's Theatre Passion¹⁴³. A la différence d'autres blogueurs, elle admet donner des critiques particulièrement bienveillantes envers les productions théâtrales qui en sont à leurs débuts (« As a reviewer, while honesty is obviously essential, I tend to always try and aim for constructive positivity, particularly when it's a fringe production that may still be in development¹⁴⁴ »). Dyer précise qu'elle ne possède pas de bagage académique en études théâtrales. Elle préfère ainsi se définir comme une « simple spectatrice » qui adore le théâtre, plutôt que comme « critique. »

2nd From Bottom (2016)

<https://2ndfrombottom.wordpress.com/>

340 critiques répertoriées dans notre corpus

Illustrée par une assiette dans laquelle est disposé un homard avec une étiquette plantée indiquant « John », la première page du blog 2nd From Bottom se démarque de la mise en page traditionnelle des éditions numériques des journaux classiques. Son éditeur, John Chapman, exerce de multiples activités en même temps. Après avoir été enseignant assistant en littérature anglaise, Chapman est désormais conseiller pédagogique, écrivain et correcteur indépendant. Il a participé à plus de 150 productions théâtrales en tant qu'acteur ou que metteur en scène.

Cultural Capital (2013)

<https://maryamphilpottblog.wordpress.com/>

296 critiques répertoriées dans notre corpus

¹⁴¹<https://www.londontheatredirect.com/>[En ligne], consulté le 02 août 2022.

¹⁴²<https://www.londontheatre1.com/>[En ligne], consulté le 02 août 2022.

¹⁴³<https://carnstheatrepassion.com/>[En ligne], consulté le 02 août 2022.

¹⁴⁴<https://theblogoftheatrethings.com/about/>[En ligne], consulté le 02 août 2022.

Cultural Capital a été créé par Maryam Philpott. Docteure en histoire de l'Université de Londres, elle commence son blog la même année que la parution de son livre *Air and Sea Power in World War I: Combat and Experience in the Royal Flying Corps and the Royal Navy* (2013). Ses publications portent autant sur le théâtre que sur les expositions et les films à Londres. Avant d'écrire indépendamment sur son blog, Philpott a été critique professionnelle pour The Reviews Hub, une plateforme théâtrale pour laquelle elle a écrit plus de 800 critiques.

Lou Review (2012)

<https://loureviews.blog/>

181 critiques répertoriées dans notre corpus

Lou Review a été créé par Louise Penn, qui a effectué ses études en gestion de bibliothèque à l'Université de Manchester. Après avoir travaillé comme libraire pendant 25 ans, Penn est aujourd'hui rédactrice et éditrice indépendante. Son blog couvre un vaste champ d'activités culturelles à Londres : le théâtre en premier lieu, mais aussi le cinéma, la télévision, les livres sur le théâtre et la musique.

Kate in Brockley (2011)

<http://kateinbrockley.blogspot.com/>

135 critiques répertoriées dans notre corpus

Kate in Brockley est le blog dont l'identité de l'éditrice est la plus discrète de ce corpus. Il est tenu par « Kate », qui se définit comme journaliste et mère de deux enfants vivant dans le sud de Londres. Le sous-titre de son blog, « Not quite a Yummy Mummy », insiste sur son rôle de mère. Au contraire des autres blogs, le sien ne comporte aucun lien vers un réseau social. Depuis 2011, elle a posté plus de 500 publications (critiques théâtrales, interviews, activités pour les enfants, etc.) avant de s'arrêter en mars 2021. C'est la seule blogueuse que nous ne sommes pas parvenus à contacter.

Susan Elkin (2016)

<http://susanelkin.co.uk/>

131 critiques répertoriées dans notre corpus

Susan Elkin a un profil différent des autres blogueurs de ce corpus. Avant de créer son propre blog, elle a été critique professionnelle pour *The Times*, *The Daily Telegraph*, *The Sunday Times*, *The Daily Mail* et *The Stage*. Elle a passé également 36 ans de sa vie comme institutrice et a écrit plus d'une cinquantaine de livres sur le théâtre, la littérature anglaise en général et aussi un roman. Depuis 2016 elle publie indépendamment sur Susan Elkin Blog, et s'est syndiquée auprès de My Theatre Mates.

Scatter of Opinion (2015)

<https://scatterofopinion.com/>

104 critiques répertoriées dans notre corpus

Scatter of Opinion a été créé par Rosie Snell en 2015. Après avoir obtenu un diplôme en théâtre à l'Université de Middlesex, Snell a travaillé en tant qu'assistante adjointe chez Evrythng, une compagnie de conseil et de service en informatique. Elle travaille aujourd'hui dans le marketing et a arrêté de publier depuis 2018 dans son blog.

Corpus II : La Critique numérique (2010 - 2020)

Caractéristiques principales

Critiques
21 326

Bloggs
28

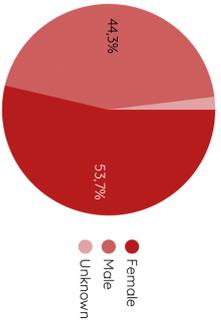
Biogqueurs
655

Pieces
8 351

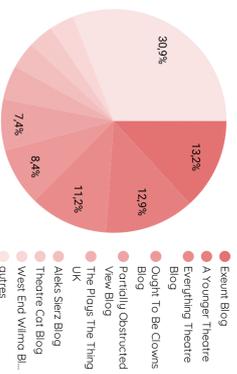
Liste des blogs

Blog	Date of Creation	Reviews
A Younger Theatre	2009	2 747
Exeunt Blog	2010	2 820
Everything Theatre Blog	2011	2 396
West End Wilms Blog	2012	759
London Theatre Review Blog	2018	459
Musical Theatre Review Blog	2013	294
View From The Cheap Seat	2011	509
Rewrite This Story Blog	2015	320

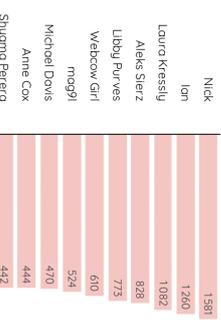
Répartition par genre



Les 10 blogs les plus représentés



Les 10 biogqueurs qui ont le plus publie



Répartition du nombre de critiques des 3 blogs les plus représentés (2010 - 2020)

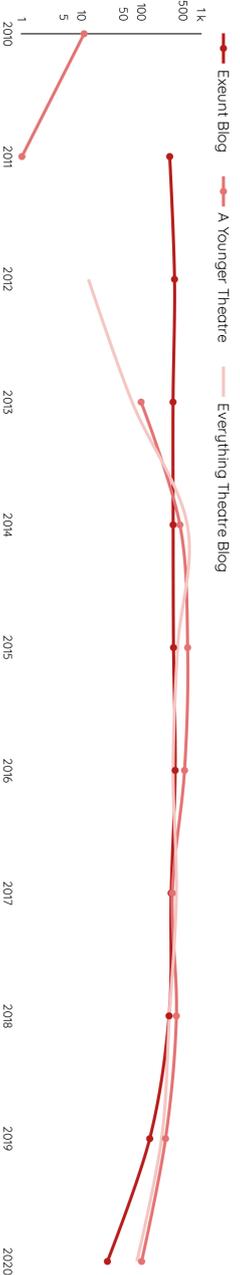


Fig. 23 : Tableau de bord des caractéristiques principales du corpus II (numérique).

Remarques conclusives sur le chapitre II

Après avoir passé plus d'une année à récupérer, structurer et nettoyer ces données, notre corpus était enfin terminé. C'est aujourd'hui la première base de données à réunir autant de critiques journalistiques et numériques contemporaines directement exploitables par la machine. Au-delà des aspects techniques qui ont représenté une tâche indéniablement chronophage, c'est surtout la sélection des sites qui a pris beaucoup de temps. Aucune étude n'ayant encore recensé le nombre de blogs disponibles sur Internet, il nous a fallu effectuer de longues recherches sur ces nouveaux acteurs numériques. Au total, ce sont donc plus de 40 000 critiques étalées sur dix années que nous avons répertoriées et qui vont désormais nous permettre d'analyser la réception du théâtre contemporain londonien à travers le regard de ces deux communautés critiques. Les deux sous-corpus sont relativement équilibrés. Nous avons 23 journaux d'un côté avec 21 717 critiques, et 28 blogs de l'autre avec 21 326 critiques. Concernant le nombre d'auteurs de critiques, nous pouvons observer qu'il y a légèrement plus de blogueurs que de journalistes (454 journalistes et 655 blogueurs).

Les chiffres que nous venons de présenter semblent déjà indiquer de fortes différences entre ces deux formes de réception. Tandis que les journaux ont tous été créés par un homme avant les années 2000, les blogs sont tous apparus après cette date et ont été majoritairement mis en place par une femme (60% des blogs ont été créés par une femme). Le nombre de pièces examinées varie considérablement d'un corpus à l'autre. Nous pouvons compter presque trois fois plus de pièces différentes vues par les blogueurs (3 084 pièces ont été commentées par les journalistes contre 8 351 par les blogueurs), ce qui laisse déjà sous-entendre qu'ils prônent probablement une culture plus diversifiée. L'appartenance de genre des critiques souligne également un rapport inégalitaire entre les deux sous-corpus. Alors que les blogueuses sont légèrement plus nombreuses que les blogueurs puisqu'elles représentent 53,7% de l'ensemble des critiques numériques, les femmes ne représentent que 36,4% du corpus journalistique.

Le chapitre suivant porte précisément sur l'analyse de ces différences. Il s'agira, grâce à l'utilisation de méthodes computationnelles, d'explorer en quoi ces deux communautés divergent. Sont-elles réellement si différentes ? Analysent-elles le théâtre contemporain londonien de manière radicalement opposée ? Et si oui, dans quelle mesure cela est-il vrai et comment le modéliser ? Avant de passer à l'exploration du corpus afin d'apporter des

réponses à ces questions, nous avons schématisé les caractéristiques principales de ces deux communautés ci-dessous :

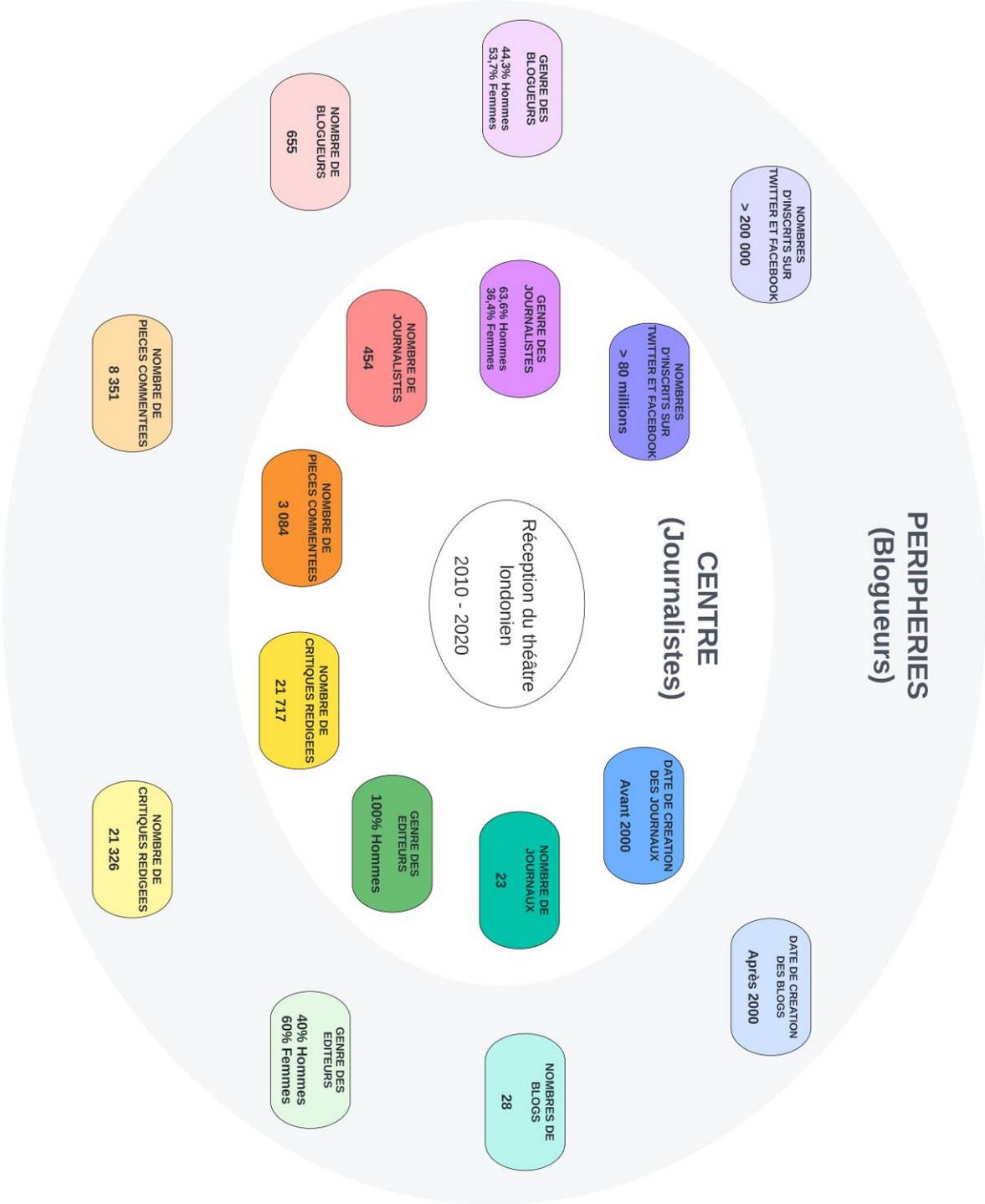


Fig. 24 : Schéma des caractéristiques principales du canon de la critique théâtrale londonienne et de ses périphéries.

Chapitre III

Exploration du corpus

Les deux chapitres précédents ont permis de présenter les fondements théoriques et techniques de notre recherche. Notre projet s'inscrit dans deux champs encore en construction : les humanités numériques théâtrales d'une part, et les études de la réception d'autre part. A l'aide d'outils informatiques, notre projet consiste à examiner la manière dont deux groupes de spectateurs analysent la scène contemporaine londonienne des années 2010 à 2020. Notre problématique est ainsi fondée sur la divergence des horizons d'attente entre ces deux communautés. Les journalistes et les blogueurs représentant deux populations dont la culture et les attentes esthétiques diffèrent, nous pouvons supposer que leur perception du théâtre en sera affectée. Comme nous l'avons montré dans le premier chapitre, la littérature disponible à leur égard confirme cette tendance. Tandis que les critiques numériques se définissent en opposition aux médias traditionnels, une partie des journalistes nie la légitimité des blogueurs. Ces deux publics participent-ils ainsi à construire deux discours différents sur le théâtre londonien ?

Pour répondre à cette problématique, nous allons explorer trois questions littéraires qui correspondent à trois domaines techniques différents. La première concerne la manière dont les journalistes et les blogueurs écrivent. Posée simplement, cette question consiste à déterminer si des différences stylistiques peuvent être observées d'un corpus à l'autre. Nous allons pour cela utiliser deux méthodes : la stylistique computationnelle et la technique dite de « text zoning. » La stylistique computationnelle se situe au carrefour de l'informatique, de la linguistique et des statistiques. Elle permet, entre autres, de modéliser la langue grâce à des données chiffrées. Le zoning est une technique qui permet de reconnaître automatiquement des « zones », c'est-à-dire des portions de texte récurrentes et homogènes sur le plan thématique ou rhétorique parmi des textes d'un domaine donné. Nous verrons comment l'association de ces deux méthodes nous permettra d'analyser le style des journalistes et des blogueurs, et quelles conclusions littéraires nous pourrons en tirer.

La deuxième question porte sur la répartition géographique de ces deux communautés dans Londres. La capacité d'accueil d'une salle de spectacle, ses subventions, sa localisation mais aussi sa popularité conditionnent la programmation des spectacles. Le Royal Court Theatre ne produit pas les mêmes dramaturges que les petits théâtres des banlieues de Londres par exemple. Le choix des lieux dans lesquels se rendent ces deux communautés nous renseigne donc sur le type de culture théâtrale qu'elles mettent en valeur. En utilisant un Système d'Information Géographique (SIG), nous pourrons produire une cartographie numérique des endroits visités par ces deux publics dans Londres. Cette expérience nous

permettra d'explorer ainsi la dichotomie « centre versus périphéries » d'un point de vue spatial.

Enfin, pour mieux comprendre les préférences esthétiques des journalistes et des blogueurs, nous analyserons dans un troisième temps ce qui constitue selon eux les « succès » du théâtre contemporain londonien. Dit autrement, la question directrice de cette partie sera la suivante : « Quelles pièces ont-ils préférées ? » Nous utiliserons un modèle algorithmique en analyse d'opinion, un domaine informatique qui permet de catégoriser automatiquement les opinions exprimées dans un texte donné, pour explorer cette thématique. Cette technique nous permettra de calculer les scores de positivité et de négativité pour chacune des pièces de notre corpus. Nous obtiendrons ainsi un classement des pièces acclamées par la critique journalistique et numérique. Cette expérience nous permettra de mieux cerner les préférences esthétiques de ces deux communautés, et d'en tirer des conclusions quant à leurs horizons d'attente respectifs.

Le chapitre suivant sera structuré en fonction de ces trois expériences. Pour chacune d'entre elles, nous commencerons par présenter le contexte littéraire qui nous a conduit à examiner notre problématique sous ces trois angles techniques. Après avoir détaillé un état de l'art de ces domaines, nous expliquerons l'expérience que nous réaliserons. Nous finirons chacune de ces parties par aborder les conclusions littéraires que nous pourrions tirer, ainsi que les limites de notre approche.

III.A Stylistique computationnelle et « zoning » : affect versus raison

III.A.1 Les journalistes et les blogueurs écrivent-ils de la même manière ?

III.A.1.1 La critique théâtrale et la notion de style

Le premier domaine que nous allons explorer a trait au style et à la forme des critiques. Bien qu'il ne soit jamais défini, le style dans la critique littéraire apparaît comme un point essentiel pour nombreux de ses auteurs. Selon Kenneth Tynnan, célèbre critique dans les années 1960 à Londres et à New York:

What counts [about critics] is not their opinion, but the art with which it is expressed [...] The subtlest and best-informed of men will still be a bad critic if his style is bad. It is irrelevant whether his opinion is 'right' or 'wrong': I learn more from GBS [George Bernard Shaw] when he is wrong than from Clement Scott when he is right¹⁴⁵.

Pour John W. English, « it is axiomatic to say that the best critics are also the best writers¹⁴⁶. » Tandis que les journalistes reprochent aux blogueurs leur manque de rigueur, ces derniers, à l'inverse, revendiquent leur prise de liberté dans la rédaction de leurs textes. Que cela soit la critique de la pièce *Death and The King's Horseman* (1975) rédigée sous forme d'e-mail frauduleux¹⁴⁷ ou l'utilisation d'illustrations de souris¹⁴⁸ pour remplacer le traditionnel « star rating » des médias journalistiques, certains des blogueurs mettent un point d'honneur à réécrire les codes de la critique classique. La plateforme Exeunt.com propose elle aussi différentes formes de critiques qui présentent un aspect ludique. Ce désir de jouer avec les normes fait explicitement écho à la description de leur identité numérique :

We give critics the freedom to expand beyond the tight deadlines and brutal word counts of more established media outlets. To give theatre and performance the consideration and thought it deserves. To be playful, and to find new ways of bringing the experience of performed arts to readers.¹⁴⁹

¹⁴⁵ Kenneth TYNNAN, *Theatre Writing*, Londres, Nick Hern Books, 2007, p. 119.

¹⁴⁶ John W. ENGLISH, *Criticizing the Critics*, Londres, Hastings House, 1979.

¹⁴⁷ Phil et Andrew, *op. cit.*

¹⁴⁸ <https://theatre.cat.com/> [En ligne], consulté le 26 décembre 2021.

¹⁴⁹ <http://exeuntmagazine.com/about-exeunt/> [En ligne], consulté le 02 août 2022.

Un de leurs critiques, David Ralf, a par exemple rédigé un texte sous forme de sonnet afin de donner son analyse de *Lemons Lemons Lemons Lemons Lemons* (2015) écrite par Sam Steiner et jouée au Camden People's Theatre en 2015 :

Walked from Camden People's Theatre to Warren Street. Called love,
agreed meet. Hung up, marvelled at the number of words.
*Lemons*⁵ teaches you a game. One hundred and ten left.
Want to say 'Warwick'. Want to compare *Some People Talk*

About Violence. Want to say "acting was good" and hat-tip
but ninety- no, eighty-five left. Have to strip it down
Like the overnight press, elucidate quick, brutish, short. No qualifiers.
Mouths and throats can make or fake almost any sound.

Otherwise? We can strike the floor in Morse or fart.
The complexity of sound is why BSL (are acronyms cheating?)
Is so different to English, down to its core, but

they don't (twenty-seven?!) sign. Sam Steiner avoids rationalising his concept
And though politics is foregrounded with anti-'Hush law' marches, (ten!)
It's 'flotsam' that'd go unremarked. Exquisite words endangered. Us diminished.¹⁵⁰

Alice Saville, co-éditrice de la plateforme Exeunt.com, a pris soin de traduire une de ses critiques sous forme d'histoire illustrée¹⁵¹. Deux autres auteurs ont également fait preuve de créativité : tandis que Tim Bano a rédigé une courte pièce de théâtre pour faire part de ses observations sur la représentation de *The End of Longing* de Lindsay Posner¹⁵², Tom Wicker a imaginé un dialogue entre deux personnages ivres pour aborder la pièce *The Boy Who Kicked Pigs* qui fut jouée au théâtre Jackson's Lane en 2013¹⁵³. Ces deux exemples illustrent la créativité avec laquelle les blogueurs peuvent rédiger leurs analyses.

Du côté de la critique journalistique en revanche, le nombre limité de caractères autorisés implique un format rédactionnel plus strict. Cette liberté prise par les blogueurs dans la rédaction de leur contenu est perçue par certains journalistes comme un manque de professionnalisme qui donnerait lieu à des critiques jugées de moins bonne qualité. Pour

¹⁵⁰ David RALPH, « Lemons, Lemons, Lemons, Lemons, Lemons, A Sonnet », *The Exeunt Magazine* [En ligne], publié le 2 décembre 2015, URL : <http://exeuntmagazine.com/reviews/lemons-lemons-lemons-lemons-lemons/>, consulté le 02 août 2022.

¹⁵¹ Alice SAVILLE, « Huff, An Illustrated Review », *The Exeunt Magazine* [En ligne], publié le 23 août 2014, URL : <http://exeuntmagazine.com/reviews/huff/>, consulté le 02 août 2022.

¹⁵² Tim BANO, « Review : The End of Longing at Playhouse Theatre », *The Exeunt Magazine* [En ligne], publié le 15 février 2016, URL : <http://exeuntmagazine.com/reviews/review-the-end-of-longing-at-playhouse-theatre/>, consulté le 02 août 2022.

¹⁵³ Tom WICKER, « The Boy who Kicked Pigs », *The Exeunt Magazine* [En ligne], publié le 7 mars 2013, URL : <http://exeuntmagazine.com/reviews/the-boy-who-kicks-pigs/>, consulté le 02 août 2022.

Michael Billington, critique de longue date dans *The Guardian*, un blog ressemble davantage à une « lettre informelle » qu'à une véritable critique. « A review, if it's to have any impact, has to have a definable structure¹⁵⁴. » Danielle Tarento, directrice du théâtre Chocolate Menier Factory à Londres, va même jusqu'à affirmer que ces blogueurs ne sont pas de « réels écrivains » : « They do not have the intellectual background or historical background or time to know what they are writing about¹⁵⁵. »

Existe-t-il alors de réelles différences stylistiques entre ces deux communautés comme le sous-entendent les articles à leur propos ? Les journalistes et les blogueurs écrivent-ils de la même manière ? Utilisent-ils les mêmes arguments et les structurent-ils de manière similaire ? Ces interrogations nous amènent à poser deux questions : « S'il y a un écart de style entre les journalistes et les blogueurs, quels outils numériques mobiliser pour le mesurer ? » et surtout, « Comment interpréter ces différences d'écriture d'un corpus à l'autre ? »

Nous allons mener deux types d'expériences afin d'esquisser des réponses à ces questions. La première repose sur le zoning. Le zoning est une méthode informatique qui permet de déterminer automatiquement les « zones » d'un texte donné. Cette expérience nous permettra de comprendre quels types d'arguments les auteurs de ces deux corpus utilisent au sein de leurs critiques. La seconde expérience reprend des procédés utilisés en stylistique computationnelle. A l'intersection de l'informatique, de la linguistique et des statistiques, la linguistique computationnelle permet, entre autres, de modéliser des phénomènes de la langue naturelle grâce à des approches statistiques. Elle a notamment été largement utilisée dans les études sur la paternité d'une œuvre.

La notion de style étant poreuse, nous avons choisi celle élaborée par D.I Holmes, qui définit le style comme un ensemble de variables calculables qui contribuent à créer « l'empreinte digitale » (*fingerprint*) d'un auteur¹⁵⁶. La complémentarité de ces deux expériences nous permettra de mieux comprendre les particularités stylistiques des auteurs de ces deux communautés.

¹⁵⁴Billington, *op. cit.*

¹⁵⁵ Matthew HEMLEY, « Danielle Tarento: Unpaid bloggers often lack 'intellectual background' to write theatre reviews », *The Stage*[En ligne], publié le 10 mai 2016, URL: <https://www.thestage.co.uk/news/danielle-tarento-unpaid-bloggers-often-lack-intellectual-background-to-write-theatre-review>consulté le 13 décembre 2021.

¹⁵⁶ D. I. HOLMES, « The analysis of literary style – A review », *Journal of the Royal Statistical Society*, vol. 148, no. 4, p. 328.

III.A.1.2 Etat de l'art de la technique du « zoning »

La notion de text zoning a été premièrement introduite par Simone Teufel dans sa thèse de doctorat rédigée à l'université d'Edinburgh en 2000¹⁵⁷. Le travail de Teufel avait pour objectif de déterminer automatiquement la structure des articles de recherche scientifique. Dans le cadre de son travail, le zoning était défini comme « a rhetorical status of a sentence with respect to the communicative function of the whole paper. » L'utilisation du zoning permettait ainsi à Teufel de faire une distinction entre les « informations secondaires » (*background information*) et les « énoncés de l'objectif particulier du papier de recherche en question. » (*statements of the particular aim of the current paper*)¹⁵⁸.

Bien que les méthodes de text zoning aient principalement été appliquées aux textes scientifiques, elles se sont également révélées utiles dans d'autres domaines, comme dans l'analyse d'e-mails¹⁵⁹ ou d'annonces publicitaires¹⁶⁰. Le nombre de zones prises en considération varie d'un auteur et d'un domaine à l'autre, mais ce sont généralement dix zones qui sont examinées. L'annotation des zones est effectuée par un groupe d'experts dans le domaine étudié. La labellisation des zones est réalisée phrase par phrase, plus rarement en fonction des paragraphes. Enfin, un bon accord inter-annotateurs est nécessaire afin de valider la pertinence des étiquettes.

Une fois un corpus terminé, il est possible d'entraîner un classificateur afin de reconnaître automatiquement les zones identifiées. Les éléments pris en compte dans la reconnaissance des zones peuvent être de deux natures différentes. La première concerne des caractéristiques de bas niveau, comme les monogrammes, les bigrammes ou certains termes spécifiques¹⁶¹. La seconde fait référence à des éléments de haut niveau, comme les relations syntaxiques¹⁶² ou la prise en compte de la zone précédente. Ces informations contextuelles permettent de faciliter le repérage automatique des différentes catégories dans la mesure où la structure d'un article scientifique est relativement similaire d'un papier à l'autre. Concernant

¹⁵⁷ Simone TEUFEL, *Argumentative Zoning: Information Extraction from Scientific Texts*, Thèse de doctorat, Université d'Edimbourg, 1999.

¹⁵⁸ Simone TEUFEL et Marc MOENS, « Summarizing Scientific Articles: Experiments with Relevance and Rhetorical Status », *Computational Linguistics*, vol. 28, no. 4, 2002, p. 409 – 406.

¹⁵⁹ Andrew LAMPERT, Robert DALE et Cécile PARIS, « Segmenting Email Message Text into Zones », *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, 2009, p. 919-928.

¹⁶⁰ Ann-Sophie GNEHM et Simon CLEMATIDE, « Text Zoning and Classification for Job Advertisements in German, French and English », *Proceedings in the Fourth Workshop on Natural Language Processing and Computational Social Science*, 2020, p. 83 – 93.

¹⁶¹ TEUFEL, *op. cit.*

¹⁶² Yufan GUHO, Anna KORHONEN et Thierry POIBEAU, « A Weakly-Supervised Approach to Argumentative Zoning of Scientific Documents », *Empirical Methods in Natural Language Processing*, 2011.

l'apprentissage, plusieurs techniques de Machine Learning ont été investiguées, en passant par l'application de la classification naïve bayésienne¹⁶³, des LSTM¹⁶⁴, des CRF et des SVM¹⁶⁵.

III.A.1.3 Définir automatiquement les zones littéraires d'une critique théâtrale

La finalité de cette expérience est en partie similaire à celle exposée dans les articles que nous venons de présenter. Cependant, cette tâche présente d'autres obstacles dus à la nature de notre corpus. La structure de la critique théâtrale n'est pas aussi régulière que celle des articles scientifiques, surtout dans le cas de la critique numérique. Les auteurs peuvent commencer la rédaction de leur texte par une analyse de l'intrigue par exemple, avant d'écrire une introduction aux différentes thématiques de la pièce. Les blogueurs, mais également les journalistes dans une certaine mesure, ne suivent donc pas toujours le même schéma. Les zones à annoter sont également plus poreuses. A l'inverse d'un article de recherche, dans lequel l'introduction et la description de l'expérience seraient deux catégories aisément identifiables par exemple, plusieurs phrases d'une critique théâtrale peuvent faire partie de la même catégorie.

Bien que l'adaptation de cette méthode présente certaines difficultés dans le cadre de notre étude, elle nous permet néanmoins d'explorer un des aspects stylistiques de notre base de données en évaluant dans quelle mesure les auteurs de ces deux communautés utilisent des caractéristiques stylistiques similaires. Notre travail se divisera en trois étapes que nous allons détailler au cours de cette sous-partie. La première porte sur l'annotation d'un échantillon de données aléatoires. Pour entraîner un programme à repérer automatiquement les différentes zones que nous allons identifier, il faut lui donner un « modèle », c'est-à-dire un certain nombre de critiques qui auront été annotées manuellement. La deuxième étape consiste à préparer et à sélectionner les différents éléments qui permettront à l'algorithme de reconnaître ces catégories. La troisième étape repose sur l'implémentation de plusieurs modèles algorithmiques pour déterminer lequel d'entre eux nous permettra d'obtenir les résultats les plus performants. Enfin, nous comparerons les résultats obtenus, et présenterons les conclusions littéraires que nous pouvons tirer de cette expérience.

¹⁶³ TEUFEL, *op. cit.*

¹⁶⁴ GNEHM, *op. cit.*

¹⁶⁵ GUO, *op. cit.*

Étape 1 : Annotation d'un échantillon de critiques

En nous appuyant sur l'essai *How to Write About Theatre* (2015)¹⁶⁶, rédigé par le critique Mark Fisher, et sur nos propres analyses des textes, nous avons identifié huit catégories employées par l'ensemble des auteurs de ce corpus pour rédiger leurs critiques :

Numéro	Catégorie
1	Introduction
2	Description de l'intrigue
3	Performance des acteurs
4	Détails visuels, sonores et olfactifs
5	Structure de la pièce
6	Public
7	Analyses
8	Conclusion

Fig. 25 : Tableau des huit catégories représentées au sein d'une critique théâtrale.

- La première catégorie, l'« **introduction** », concerne les accroches rédigées par les auteurs ainsi que les quelques phrases qui permettent de situer la pièce dans le paysage théâtral contemporain britannique. Cette section contient généralement des informations sur le dramaturge, sur les pièces qu'il a écrites, sur les thèmes prédominants dans la pièce ainsi que sur le théâtre dans lequel elle a été jouée.
- La « **description de l'intrigue** » est une catégorie narrative. Elle est aisément identifiable : en utilisant souvent le présent simple, l'auteur y détaille le cadre spatio-temporel de l'histoire, les relations entre les personnages ainsi que le nœud de l'intrigue.
- La « **performance des acteurs** » est une autre section facilement repérable dans la mesure où elle consiste en quelques phrases dans lesquelles de nombreux noms propres apparaissent. Le critique y décrit la qualité du jeu des acteurs, c'est-à-dire les émotions

¹⁶⁶ FISHER, *op. cit.*

qu'ils ont réussi à transmettre au public et la manière dont ils ont interprété le script de la pièce.

- Les « **détails visuels, sonores et olfactifs** » portent davantage sur l'analyse des décors scéniques. Dans de nombreuses critiques, les auteurs prenaient soin d'énumérer ce qui avait retenu leur attention sur le plateau. Les décors, les costumes ou les lumières sont passés au crible dans cette section, de sorte que les lecteurs puissent avoir une idée de l'atmosphère qui imprégnait le plateau lors de la représentation.
- La section « **structure de la pièce** » fait référence aux remarques littéraires. Les auteurs y examinent la représentation théâtrale d'un point de vue théorique.
- Le « **public** » inclut toutes les mentions des spectateurs : leurs réactions, leur disposition dans la salle, les discussions que les auteurs ont pu avoir avec eux à la suite du spectacle ou s'ils étaient nombreux ou non lors de la représentation.
- La section « **analyses** » est plus subtile. Généralement très courte (ou une deux phrases), elle contient des observations sur les thématiques et sur le « message » de la pièce.
- Enfin, la « **conclusion** » est tout le temps, ou presque, située à la fin de la critique. Composée d'une ou de deux phrases, elle clôturait l'avis général du critique sur la représentation.

Afin d'entraîner notre modèle à reconnaître automatiquement la structure des critiques, nous sélectionnons 1 000 critiques de manière aléatoire que nous annotons manuellement en fonction de ces huit catégories. Ces critiques sont extraites du premier corpus uniquement. Nous considérons que les critiques journalistiques représentent la « norme », c'est pourquoi nous nous focalisons sur ce corpus. L'annotation est réalisée par une seule personne spécialiste des études théâtrales. Il faut entre 8 et 15 minutes par critique pour placer chacune des étiquettes, ce qui représente 170 heures de travail au total.

Pour faciliter la labellisation, nous travaillons directement sur un fichier Excel. Nous attribuons une couleur à chacune des catégories, et modifions ensuite la couleur des phrases des textes en fonction de la zone à laquelle elles appartiennent. Le code des couleurs est

ensuite récupéré sur Python, ce qui permet au script de différencier les huit catégories. Le tableau suivant (Fig. 26) répertorie l'association des différentes couleurs à leur zone :

Numéro	Catégorie	Couleur
1	Introduction	Violet
2	Description de l'intrigue	Rouge
3	Performance des acteurs	Orange
4	Détails visuels, sonores et olfactifs	Vert
5	Structure de la pièce	Marron
6	Public	Gris
7	Analyses	Bleu
8	Conclusion	Jaune

Fig. 26 : Tableau des huit catégories associées à leur couleur.

Pour illustrer notre travail de labellisation, voici un exemple extrait des critiques rédigées pour la pièce de Sam Shepard, *A Lie of the Mind* (1985), qui a été représentée en mai 2017 au Southwark Playhouse Theatre de Londres. Cette critique a été écrite le 11 mai 2017 par Fergus Morgan, critique dans le journal *The Stage* :

Critique originale

Sam Shepard plays gnaw away at you. They tease you with cryptic clues, disintegrating storylines and restless, febrile characters. His 1985 play *A Lie of the Mind* features the same symbolism heavy blend of redneck grit and warped American dreams as *Fool For Love* and *Buried Child*. This time, its themes - dysfunctional relationships, inescapable destinies, mortal love - land in frost-bitten rural Montana. Two families are joined by an abusive marriage. Jake has beaten Beth to a pulp and retreated to his childhood bedroom, plagued by guilt and grief. Brain-damaged and bewildered, Beth has been swallowed up by her clan too, cooped up with her bitter parents and her gun-toting brother in their snowy ranch. Shepard follows these converging narratives, tracing every character's inner geography in forceful, elliptical brushstrokes. James Hillier's dusky, smoke-filled staging bears striking resemblance to John Tiffany's recent production of *The Glass Menagerie*, with the action isolated on small, square platforms against a vast blackness, and a neon moon hovering balefully behind the stage. But despite a set of detailed performances - particularly from John Stahl as Beth's flinty father and Laura Rogers' as Jake's mousy sister Sally - and a contemplative live score from James Marples, it never evokes the requisite haunted atmosphere nor mines the murky depths of Shepard's dialogue. It's just too crowded, too cluttered, too clunky, and the play loses much of its unsettling power as a result.

Critique annotée

(Introduction) Sam Shepard plays gnaw away at you. They tease you with cryptic clues, disintegrating storylines and restless, febrile characters. His 1985 play *A Lie of the Mind* features the same symbolism heavy blend of redneck grit and warped American dreams as *Fool For Love* and *Buried Child*. This time, its themes - dysfunctional relationships, inescapable destinies, mortal love - land in frost-bitten rural Montana. **(Intrigue)** Two families are joined by an abusive marriage. Jake has beaten Beth to a pulp and retreated to his childhood bedroom, plagued by guilt and grief. Brain-damaged and bewildered, Beth has been swallowed up by her clan too, cooped up with her bitter parents and her gun-toting brother in their snowy ranch. **(Analyses)** Shepard follows these converging narratives, tracing every character's inner geography in forceful, elliptical brushstrokes. **(Détails visuels, sonores et olfactifs)** James Hillier's dusky, smoke-filled staging bears striking resemblance to John Tiffany's recent production of *The Glass Menagerie*, with the action isolated on small, square platforms against a vast blackness, and a neon moon hovering balefully behind the stage. **(Performance des acteurs)** But despite a set of detailed performances - particularly from John Stahl as Beth's flinty father and Laura Rogers' as Jake's mousy sister Sally - and a contemplative live score from James Marples, it never evokes the requisite haunted atmosphere nor mines the murky depths of Shepard's dialogue. **(Conclusion)** It's just too crowded, too cluttered, too clunky, and the play loses much of its unsettling power as a result.

Etape 2 : Préparation des données

Afin d'obtenir la meilleure prédiction possible, nous créons plus de 80 variables au total en nous inspirant des travaux précédents effectués par Teufel et Guo et Al. Le tableau ci-dessous résume les quatre catégories de variables que nous créons :

Catégorie de la variable	Exemples
Statistiques textuelles	- Longueur moyenne des mots ou des phrases, le nombre de mots, fréquence des pronoms personnels, adjectifs / adverbes / verbes les plus fréquents...
Temps des verbes	- Nombre de verbes au passé, au présent et au futur
Parties du discours	- Analyse de la syntaxe des phrases
Reconnaissance des entités nommées	- Lieux, organisations, personnes...

Fig. 27 : Catégorisation des types de variables créées pour la prédiction du modèle.

Si la création de ces différents types de variables permet d'affiner la précision de l'algorithme, seulement un certain nombre d'entre elles sont pertinentes pour les modèles que nous testons. Afin de sélectionner les prédicteurs les plus performants, nous utilisons dans un premier temps une matrice de corrélation ainsi qu'une Analyse des Composantes Principales (ACP). Dans un second temps, nous appliquons plusieurs modèles afin de déterminer lequel d'entre eux nous permet d'obtenir le score de précision le plus élevé.

Etape 3 : Sélection de quatre modèles pour la classification des phrases

Nous sélectionnons ensuite quatre modèles traditionnellement utilisés pour cette tâche :

- **La classification naïve bayésienne** est un algorithme qui se fonde sur le Théorème de Bayes, qui propose l'hypothèse selon laquelle une indépendance entre les prédicteurs existerait. Cela se traduit dans l'algorithme par la supposition que chaque classe possède une caractéristique particulière et que cette dernière est indépendante des autres. Comme on fait ici l'hypothèse d'indépendance des prédicteurs, le Naïve Bayes fonctionne très bien avec peu de données.
- **Random Forest**¹⁶⁷ : La forêt d'arbres de décisions repose sur des arbres décisionnels, c'est-à-dire sur une série de questions que l'on appelle « tests » en langage informatique et dont la réponse binaire « oui / non » permettra de prendre une décision finale. La forêt aléatoire est un type d'algorithme de classification (ou de régression), qui se fonde sur l'utilisation d'un grand nombre d'arbres de décisions. Ces arbres ont un fonctionnement séquentiel, c'est-à-dire que l'on sépare les données en deux groupes à chaque nœud de l'arbre. Cette séparation se fait à l'aide d'un seuil sur un prédicteur. Le prédicteur choisi est celui qui permet de maximiser l'homogénéité des sous-groupes créés. Dans la mesure où cette méthode peut prendre en compte de multiples paramètres, nous pouvons supposer qu'elle sera efficace pour résoudre notre problème lié à la reconnaissance des différentes catégories au sein d'une critique.

¹⁶⁷ Tin Kam HO, « RandomDecision Forests », *Proceedings of 3rd International Conference on Document Analysis and Recognition*, vol. 1, 1995, p. 278–282.

- **La méthode K-NN¹⁶⁸** : Pour réaliser une prédiction, nous regardons les classes des K plus proches voisins. La classe prédite correspond à la classe majoritaire parmi ces K classes. Les K plus proches voisins sont trouvés en calculant la distance entre l'individu et l'ensemble des individus d'apprentissage. La dimension de l'espace dans laquelle cette distance est calculée est égale au nombre de prédicteurs utilisés pour caractériser un individu.
- **Réseaux de neurones¹⁶⁹** : Le réseaux de neurones est un modèle de deep learning, ou d'apprentissage profond, qui peut prendre des formes très différentes. Pour la classification, on utilise des modèles simples avec seulement quelques couches. Une couche correspond à un ensemble de neurones. Lorsqu'une information (ici des nombres) passe d'un neurone à un autre, on la multiplie par un certain poids. Ce poids permet au neurone suivant de réagir aux différentes informations qu'il reçoit. C'est cette combinaison d'informations à laquelle on donne plus ou moins de poids qui permet d'obtenir en sortie une information sur la classe la plus probable.

Etape 4 : Comparaison des performances

Nous appliquons ces quatre modèles à notre corpus. Les résultats obtenus sont présentés dans le tableau ci-dessous :

Modèle	Performance
Classification naïve bayésienne	69%
Forêt d'arbres décisionnels	80%
Méthode KNN	72%
Réseaux de neurones	61%

Fig. 28 : Résultats des performances des quatre modèles algorithmiques utilisés.

¹⁶⁸ Naomi ALTMAN, «An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression», *The American Statistician*, vol.46, 1992 p.175–185.

¹⁶⁹ Alessandro SPERDUTI, Antonina STARITA, «Supervised Neural Networks for the Classification of Structures», *IEEE Transactions on Neural Networks*, vol. 8, no. 1997, p. 714-735.

C'est finalement le modèle de la forêt d'arbres décisionnels qui nous permet d'obtenir le score de précision le plus haut. Concernant la performance des trois autres modèles, nous pouvons observer que ce sont les réseaux de neurones qui obtiennent le score de précision le plus bas. Ce résultat s'explique par le « faible » nombre de données annotées pour ce type de modèle. La classification naïve bayésienne ne parvient pas à rendre compte de la complexité de la tâche et donne, par conséquent, des résultats insuffisants. Enfin, la méthode K-NN présente elle aussi des limites. Dans la mesure où les textes sont trop courts, leur classification reste superficielle.

III.A.2 Deux façons d’appréhender l’expérience théâtrale

III.A.2.1 Résultats

Bien que la performance des modèles pourrait être améliorée, ces premiers résultats nous permettent néanmoins d’obtenir une première vue d’ensemble différentes catégories utilisées par les journalistes et les blogueurs. Le tableau et les graphes suivants présentent les résultats obtenus :

Catégorie	Journaux	Blogs
Introduction	15.9% du corpus journalistique total sont dédiés à l’introduction	17% du corpus numérique total sont dédiés à l’introduction
Description de l’intrigue	32.5% du corpus journalistique total sont dédiés à l’intrigue	28% du corpus numérique total sont dédiés à l’intrigue
Performance des acteurs	15.2% du corpus journalistique total sont dédiés à la performance des acteurs	18.4% du corpus numérique total sont dédiés à la performance des acteurs
Détails visuels, sonores et olfactifs	4.4% du corpus journalistique total sont dédiés aux détails visuels, sonore et olfactifs	7.3% du corpus numérique total sont dédiés aux détails visuels, sonore et olfactifs
Structure de la pièce	8.6% du corpus journalistique total sont dédiés à la structure de la pièce	7.2% du corpus numérique total sont dédiés à la structure de la pièce
Public	0.9% du corpus journalistique total est dédié au public	2.1% du corpus numérique total sont dédiés au public
Analyses	13.3% du corpus journalistique total sont dédiés aux analyses	11.7% du corpus numérique total sont dédiés aux analyses
Conclusion	9% du corpus journalistique total sont dédiés à la conclusion	8.3% du corpus numérique total sont dédiés à la conclusion

Fig. 29 : Pourcentages des huit catégories au sein de chaque corpus.

Représentation des huit catégories littéraires utilisées par l'ensemble des journalistes

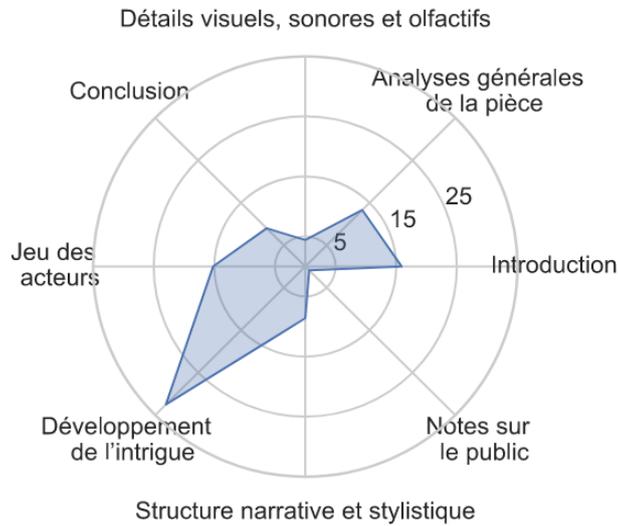


Fig. 30 : Représentation des pourcentages des huit catégories au sein du corpus I.

Représentation des huit catégories littéraires utilisées par l'ensemble des blogueurs

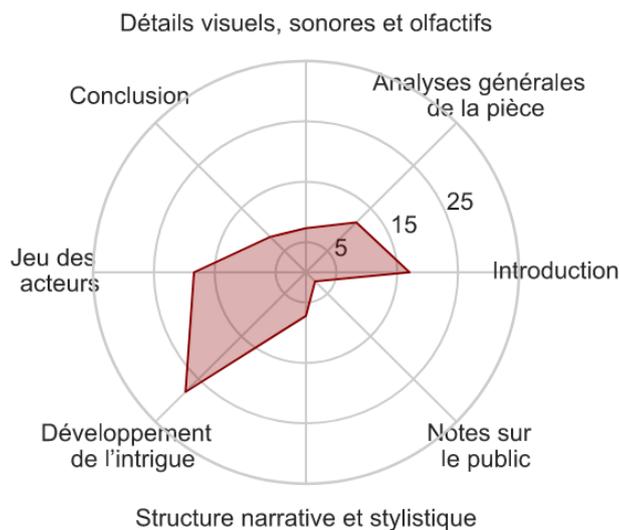


Fig. 31 : Représentation des pourcentages des huit catégories au sein du corpus II.

Représentation des huit catégories littéraires utilisées par l'ensemble des journalistes et des blogueurs

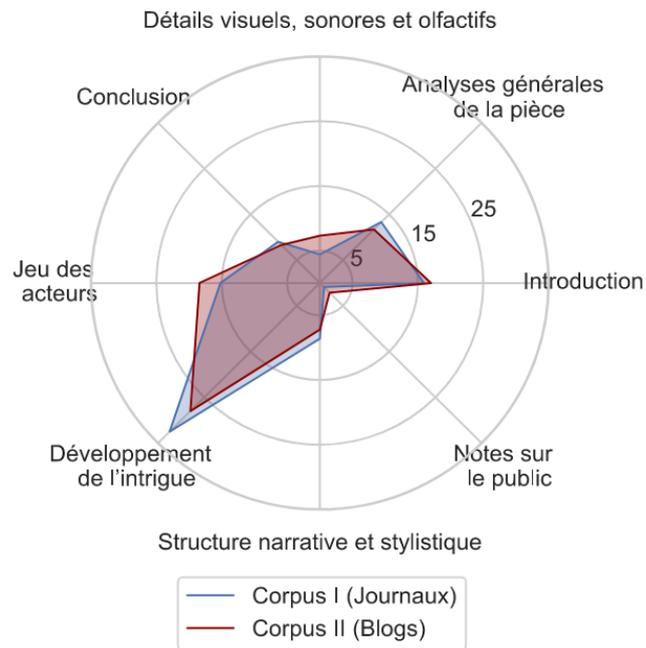


Fig. 32 : Représentation des pourcentages des huit catégories au sein du corpus I et II.

III.A.2.2 Interprétation

Si le débat dans la sphère artistique souligne principalement les différences stylistiques entre ces deux communautés, nous pouvons néanmoins observer que des points communs subsistent entre elles. Les huit catégories que nous avons définies sont présentes dans les deux corpus, ce qui prouve que les journalistes et les blogueurs utilisent des arguments similaires. Malgré ces similitudes, des divergences notables sont observables. Lorsque nous regardons de plus près les pourcentages de chacune de ces catégories au sein des deux corpus, nous pouvons nous rendre compte que les blogueurs et les journalistes semblent se focaliser sur des aspects différents des pièces auxquelles ils assistent.

Les « Détails visuels, sonores et olfactifs », la « Performance des acteurs » ainsi que les « Remarques sur le public » sont des catégories qui mettent au premier plan la subjectivité du critique et dont les pourcentages sont plus élevés chez les blogueurs que chez les journalistes. Nous avons reproduit une partie des résultats dans le tableau ci-dessous qui présente ces trois catégories :

Catégorie	Journaux	Blogs
Détails visuels, sonores et olfactifs	4.4%	7.3%
Performance des acteurs	15.2%	18.4%
Remarques sur le public	0.9%	2.1%

Fig. 33 : Tableau des trois catégories dans le corpus II (numérique) dont les pourcentages sont plus élevés que dans le corpus I (journalistique).

Pour le corpus I au contraire, c'est-à-dire pour la critique journalistique, les valeurs élevées en pourcentages concernent des catégories liées à des arguments de nature factuelle. Les « Analyses », l'« Intrigue » et les remarques sur la « Structure de la pièce » s'appuient davantage sur des analyses rationnelles. Le tableau suivant (Fig. 34) répertorie ces trois catégories de résultats :

Catégorie	Journaux	Blogs
Analyses	13.3%	11.7%
Description de l'intrigue	32.5%	28%
Structure de la pièce	8.6%	7.2%

Fig. 34 : Tableau des trois catégories dans le corpus I (journalistique) dont les pourcentages sont plus élevés que dans le corpus II (numérique).

Les analyses en stylistique computationnelle que nous avons effectuées semblent également confirmer la tendance de ces résultats. Deux expériences en particulier ont retenu notre attention. La première porte sur le calcul des noms communs les plus utilisés au sein des deux corpus, tandis que la seconde concerne l'utilisation des pronoms personnels. Les résultats que nous avons obtenus sont présentés ci-dessous en figures 35, 36 et 37 :

Corpus I (Journaux) - Top 15 des noms communs les plus fréquents

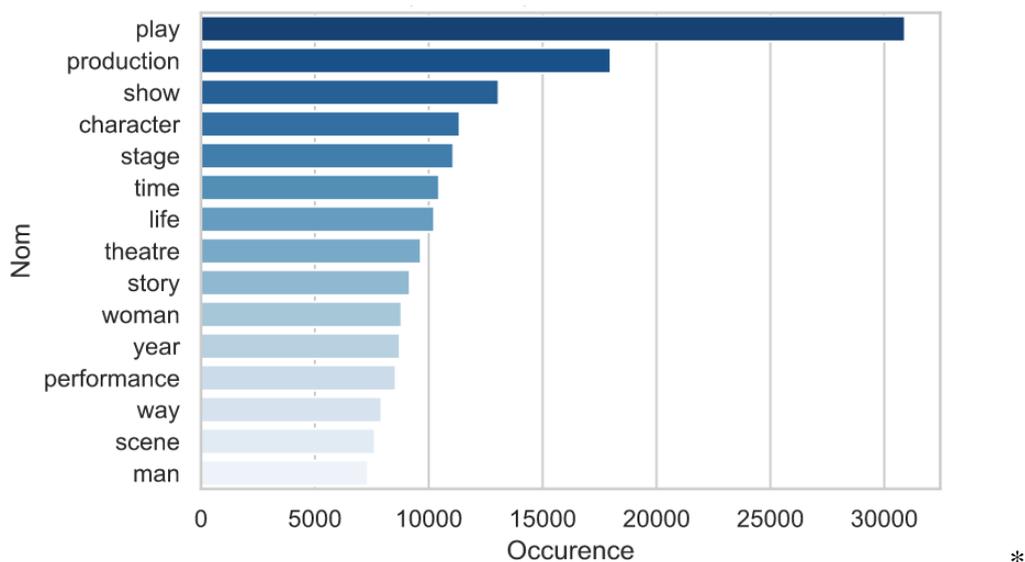


Fig. 35 : Classement des 15 noms communs les plus employés par les journalistes.

Corpus II (Blogs) - Top 15 des noms communs les plus fréquents

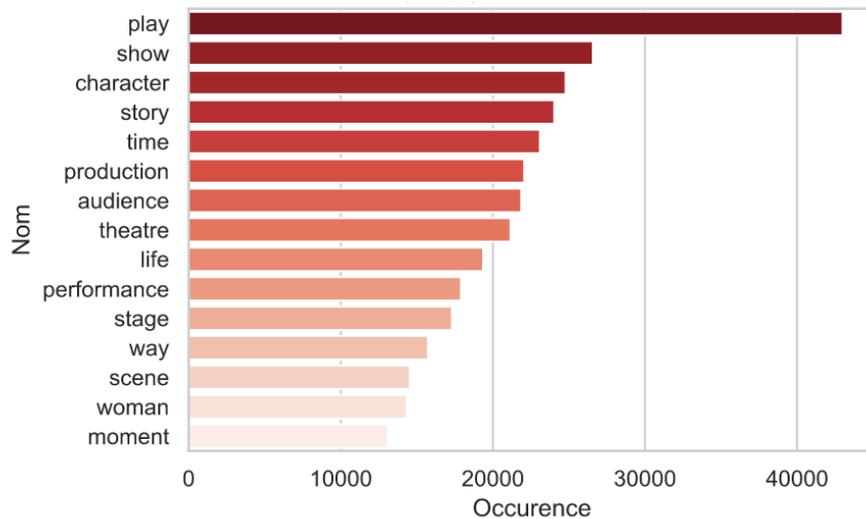


Fig. 36 : Classement des 15 noms communs les plus employés par les blogueurs.

A première vue, il semble que l'utilisation des noms communs est relativement similaire d'un corpus à l'autre puisque les mêmes termes sont employés par les critiques professionnels que par les critiques amateurs. Deux exceptions significatives peuvent être notées cependant. Tandis que les journalistes emploient les noms « Year » et « Man » qui sont absents du top 15 des noms les plus récurrents parmi les blogueurs, ce sont les termes « Moment » et « **Audience** » qui n'apparaissent pas dans le classement journalistique. La présence du terme « Audience » laisse sous-entendre que les blogueurs prêtent une attention particulière au public. Pourrait-on y voir ici deux manières d'appréhender l'expérience du théâtre ? L'une qui serait plutôt rationnelle, tournée vers l'analyse du spectacle, et l'autre qui serait plus émotionnelle, davantage tournée vers les réactions de l'humain, ou du public ?

La répartition de l'utilisation des pronoms personnels entre ces deux corpus semble confirmer la validité de cette hypothèse (Fig. 37). Alors que la première personne du singulier représente 10,9 % de l'utilisation totale des pronoms personnels dans le corpus I, sa présence est presque doublée dans le corpus II (18,5 %).

Pourcentages des pronoms personnels dans le corpus I et II

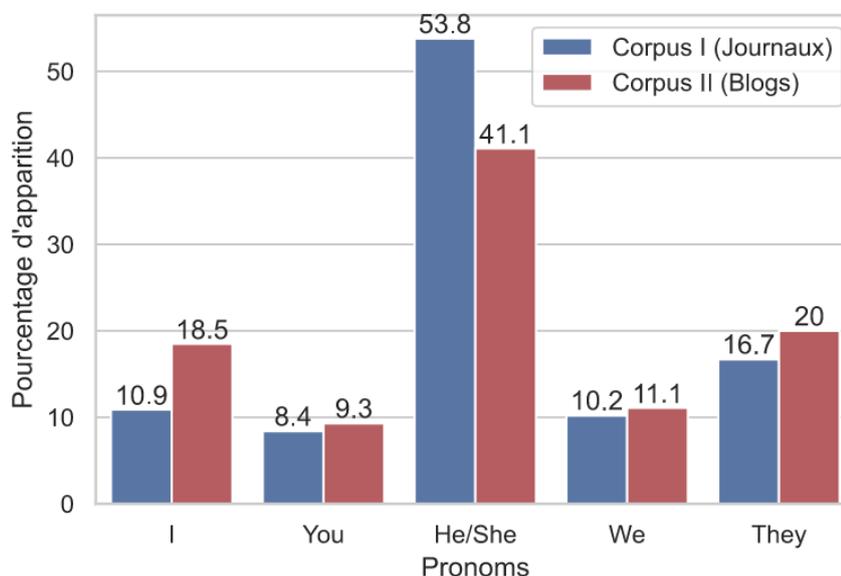


Fig. 37 : Représentation des différents pronoms personnels utilisés dans les deux corpus.

La prédominance de ce « Je » témoignerait-elle d'une écriture plus subjective qui laisserait davantage d'espace à la singularité et à la sensibilité de chaque opinion des critiques ? Des quelques réponses que nous avons obtenues de la part des blogueurs, tous mettaient en avant l'importance de faire entendre leur propre voix et de proposer des analyses avant tout personnelles. Lorsque nous avons demandé à Stacey Midwell, éditrice de RevStanBlog.com, pourquoi elle écrivait, sa réponse fut explicite : « writing the blog was different ; it was about writing for me and in my own voice¹⁷⁰. » John Chapman, à la tête de 2ndFromBottom.wordpress.com, nous a donné une réponse similaire : « For my first blog I wanted to create a permanent record of a unique opportunity as an archive for myself¹⁷¹. » Laura Kressly, éditrice de ThePlaysTheThingsUK.com, affirme également qu'à la différence des journalistes, les « blogueurs écrivent pour eux-mêmes (« bloggers write for themselves »)¹⁷². » Dans les prérequis pour postuler en tant que critique pour la plateforme AyoungerTheatre.com, les éditeurs demandent une analyse *personnelle* d'une représentation

¹⁷⁰ L'entretien est disponible en annexes p. 253 – 54.

¹⁷¹ L'entretien est disponible en annexes p. 255 – 56.

¹⁷² L'entretien est disponible en annexes p. 257 – 58.

théâtrale : « Remember we are looking for opinionated arguments, lots of passion and a clear, unique voice, as well as coherent use of the English language¹⁷³. »

L'absence de nombres de caractères imposés contribue probablement au déploiement de ce style subjectif. Comme le remarque justement Dan Rebellato, l'espace accordé dans les journaux aux critiques théâtrales s'amenuise progressivement depuis les années 1950, et contraint les critiques à réduire leurs arguments :

This has been cut back and cut back and now a critic is lucky to get 600 words on an evening of theatre, quite often 350. Once you've set the scene and complemented a couple of actors and the set, what do you have for analysis and discussion? 150 words? Critics rarely have the chance to think aloud, to speculate, to take an idea for a walk. They don't have the space.¹⁷⁴

Au lieu de condenser leurs idées comme un journaliste le ferait par manque d'espace et de temps, les blogueurs peuvent explorer la complexité de leurs réflexions. Cet espace-temps illimité les amène à développer leur créativité, en écrivant des sonnets ou des dialogues par exemple, mais aussi à enrichir leurs analyses en faisant référence à d'autres textes critiques. Beaucoup de blogueurs insèrent des liens hypertextes qui renvoient à d'autres blogs ou parfois même à des billets publiés dans des journaux en ligne.

A notre question de départ, « Existe-t-il de réelles différences stylistiques entre ces deux communautés, et comment les interpréter ? », nous pouvons répondre que oui, dans une certaine mesure, des variations dans leur manière d'écrire sont observables. Leur objet d'intérêt varie d'un corpus à l'autre. Tandis que les critiques professionnels semblent proposer un discours davantage neutre et objectif, les critiques amateurs se concentrent sur leur ressenti.

III.A.2.3 Limites de l'approche

Tant d'un point de vue technique que littéraire, nous pourrions améliorer nos résultats. Sur un plan informatique d'abord, bien que la performance des modèles implémentés soit satisfaisante, nous pourrions probablement obtenir des scores de précision plus élevés en annotant un plus grand nombre de données et en labellisant également des critiques du corpus II. L'accord inter annotateurs est aussi une composante importante dans ce type de travail que

¹⁷³ Le formulaire de candidature mis à disposition par AyoungerTheatre.com est disponible en annexes p. 259 – 261.

¹⁷⁴ Dan REBELLATO, « Critical Thinking », *Dan Rebellato*[En ligne], URL : Critical Thinking — Dan Rebellato, consulté le 09 mars 2022.

nous n'avons pas pu mettre en application par manque de temps. Pour mener à bien cette étude, il faudrait ainsi demander à un ou à plusieurs spécialistes en études théâtrales d'annoter ces textes.

Sur un plan littéraire, nous pourrions mener d'autres expériences à partir de ce que nous avons déjà effectué pour affiner notre analyse de la stylistique dans ces deux corpus. Nous pourrions, par exemple, étudier l'ordre d'apparition des catégories au sein d'une critique. Cela nous permettrait de comprendre quels aspects d'une pièce de théâtre les critiques mettent en avant. Nous pourrions également reproduire ces analyses à une plus petite échelle, c'est-à-dire en fonction des journaux et des blogs, afin de repérer s'il y a des différences stylistiques d'un média à l'autre. De la même manière, nous pourrions répéter ce processus à une échelle individuelle. Nous pourrions ainsi déterminer si les critiques au sein d'un journal ou d'un blog écrivent différemment, ou si des variations sont observables entre les hommes et les femmes.

Il ne s'agit pas d'affirmer, suite à cette première expérience, que tous les journalistes rédigent de manière objective, et que tous les blogueurs se concentrent sur leur ressenti. Certains journalistes se focalisent sûrement davantage sur les détails de la scène, tandis que certains blogueurs proposent probablement des analyses plus distantes des spectacles auxquels ils assistent. Ce que ces résultats montrent est plutôt une tendance générale. Les particularités de chaque individu sont effacées de sorte que nous puissions obtenir une vue d'ensemble des propriétés littéraires propre à chaque corpus. C'est ce qui fait la force et la faiblesse des humanités numériques, ou de nombreuses approches statistiques. L'informatique permet d'examiner un phénomène dans sa globalité, mais il efface en même temps les singularités.

Transition

Cette première sous-partie avait pour finalité d'examiner la manière dont les journalistes et les blogueurs écrivent. Nous avons utilisé le zoning, une méthode informatique qui permet de déterminer automatiquement les « zones » ou les séquences thématiques d'un texte pour étudier la structure des critiques. Cette technique, encore jamais utilisée pour l'étude des textes théâtraux, nous a permis de constater que les auteurs de ces deux corpus focalisent leur attention sur des éléments différents.

Tandis que les journalistes privilégient des catégories liées à des éléments rationnels (description de l'intrigue, analyses de l'auteur et remarques sur la structure narrative de la pièce), les blogueurs s'attardent sur des catégories qui sont davantage liées à l'affect (détails visuels, sonores et olfactifs, performance des acteurs et remarques sur le public). Les expériences en stylistique computationnelle que nous avons effectuées confirment cette tendance. Les noms communs les plus utilisés par les auteurs de ces deux corpus révèlent des priorités différentes dans leurs analyses. La présence du terme « Audience » parmi les blogueurs laisse supposer qu'ils adoptent une approche sensible du théâtre, ou du moins davantage tournée vers l'humain. De la même manière, le pronom personnel « Je » est presque deux fois plus employé dans la blogosphère que dans les médias classiques.

Nous avons déduit de ces expériences que les blogueurs adoptent un style plus subjectif qui laisse probablement davantage de place à leur ressenti. Libérés des contraintes liées au temps et au nombre de mots imposés, les blogueurs peuvent explorer et développer leurs idées. Cette liberté permise grâce à l'espace numérique explique sûrement en partie cette tendance stylistique. L'autre élément à prendre en considération est leur désir de faire entendre leur propre voix. En réitérant ce « Je », symbole d'une subjectivité affirmée, les blogueurs entendent donner une interprétation singulière des représentations théâtrales auxquelles ils assistent. Pourrait-on y voir ici les signes d'une pluralité de discours qui seraient moins présente dans la communauté journalistique ?

La partie suivante s'attachera à examiner ces deux publics sous un autre angle. Nous avons voulu savoir si des différences spatiales pouvaient être cartographiées d'un corpus à l'autre. La position d'un théâtre dans une ville est indicative, dans une certaine mesure, du type de population qui le fréquente, des subventions qu'il obtient et des compagnies qui y sont produites. Un théâtre situé au centre de Londres ne bénéficiera pas des mêmes avantages qu'un théâtre localisé dans les périphéries de la capitale. En déterminant les types de lieux

fréquentés par les journalistes et par les blogueurs, nous avons ainsi cherché à savoir quel type de culture ils mettaient en valeur. Formulé autrement, « Peut-on observer un clivage entre le canon de la critique théâtrale londonienne et ses périphéries d'un point de vue géographique ? Et si oui, que nous indique cet écart quant aux pratiques culturelles de ces deux communautés ? »

Cette expérience sera structurée de la même manière que celle que nous venons de présenter à propos du style des critiques. Nous commencerons par introduire le contexte littéraire qui nous a menés à adopter cette approche. Nous poursuivrons notre travail en présentant un état de l'art des différents travaux qui reposent sur la cartographie numérique en études théâtrales. Nous détaillerons ensuite les outils techniques que nous avons utilisés pour créer nos cartes. Enfin, nous présenterons les résultats graphiques, les conclusions que nous pouvons tirer et les limites de notre expérience.

III.B Cartographie numérique : Centre versus périphéries spatiales

III.B.1 Existe-t-il une rupture géographique entre le centre de la critique théâtrale et ses périphéries ?

III.B.1.1 Analyser le théâtre d'un point de vue spatial

Le deuxième domaine que nous allons explorer porte sur la répartition spatiale des journalistes et des blogueurs à Londres. La capacité d'accueil d'une salle de spectacle, ses subventions, sa localisation mais aussi sa popularité conditionnent la programmation des spectacles. La pièce *Hamlet* de Shakespeare n'aura pas le même rayonnement commercial si elle est jouée dans une bibliothèque municipale du nord de Londres qu'au National Theatre.

Le choix des lieux est donc un indicateur qui nous permet de mieux comprendre ces deux communautés. Afin de cartographier les déplacements de ces critiques dans leur ensemble, les outils numériques sont nécessaires. Avec plus de 40 000 textes littéraires, et donc plus de 40 000 mentions géographiques, il nous est impossible de donner des réponses précises sans analyse statistique. Surtout, les outils de visualisation s'avèrent particulièrement utiles dans ce cas précis. La création d'une carte géographique permet de repérer des tendances qui seraient difficilement décelables à l'aide de simples résultats chiffrés.

Les questions centrales de cette deuxième partie sont les suivantes : « Les journalistes et les blogueurs se rendent-ils dans les mêmes lieux de spectacle ? Les critiques journalistiques tendent-ils à favoriser les théâtres au centre de la ville, ce qui laisserait supposer qu'ils mettent à l'honneur les grandes productions théâtrales ? A l'inverse, les critiques numériques visitent-ils davantage les lieux marginaux de la capitale, ce qui pourrait nous conduire à penser qu'ils mettent plus en avant de petites productions ? »

A ce jour, aucune étude ne propose une analyse spatialisée de cette envergure des lieux de spectacle de Londres au XXI^e siècle. Cette absence de travaux dans ce domaine est probablement due au manque de données disponibles pour réaliser une telle tâche. Ce n'est que très récemment que les institutions anglaises ont commencé à s'intéresser à leurs lieux de spectacle. En 2014, le National Theatre de Londres et la Society of London Theatre (SOLT)

ont fait la demande du premier rapport de l'état de l'art de l'industrie du théâtre en Angleterre. Cette étude, composée de 48 pages, est la première à se fonder sur des données chiffrées afin de cartographier les lieux de spectacle de Londres et de ses visiteurs¹⁷⁵. Alistair Smith, éditeur au journal *The Stage* et chargé de ce rapport, reconnaît cependant que seuls les théâtres dits « professionnels » ont été pris en considération. Autrement dit, une partie des lieux qui abritent des performances moins conventionnelles reste donc absente de ce rapport. Ce travail réalisé sur les données cartographiques de notre corpus permet ainsi non seulement de mieux comprendre comment se positionnent ces deux communautés d'un point de vue géographique, mais aussi d'avoir une vision générale des différents lieux de représentations théâtrales dans Londres aujourd'hui.

La question directrice de ce sous-chapitre sera la suivante : « La dichotomie « centre / périphéries » de la critique théâtrale londonienne est-elle observable d'un point de vue géographique ? Et si oui, comment interpréter ces résultats ? » Comme dans le sous-chapitre précédent, nous commencerons par présenter l'état de l'art des Systèmes d'Information Géographique (SIG). Nous verrons en quoi consistent ces technologies récentes et comment elles ont pu être utilisées par d'autres chercheurs dans les études théâtrales. L'état de l'art effectué, nous aborderons le travail que nous réaliserons. Dans un dernier temps, nous exposerons les résultats obtenus et les conclusions que nous pouvons en tirer.

III.B.1.2 Que sont les Systèmes d'Information Géographique (SIG) et comment ont-ils été utilisés dans les études théâtrales ?

Un système d'information est un « ensemble de ressources (...) permettant la collecte, le stockage, la structuration, la modélisation (...) des informations (textes, images, sons, vidéo...) »¹⁷⁶. Un système d'information géographique est donc un système qui permet de traiter l'information spatiale, c'est-à-dire l'information liée à un territoire donné. Les SIG ont deux caractéristiques principales :

¹⁷⁵ <https://www.londontheatre1.com/londontheatrereportv7.pdf> [En ligne], consulté le 12 juin 2022.

¹⁷⁶ Sylvie SERVIGNE, « SYSTÈMES INFORMATIQUES - Conception, architecture et urbanisation des systèmes d'information », *Encyclopædia Universalis*, URL : <https://www.universalis.fr/encyclopedie/systemes-informatiques-conception-architecture-et-urbanisation-des-systemes-d-information/>, consulté le 12 juin 2022.

1. La capacité de gérer et de traiter les relations spatiales entre objets ou phénomènes dans l'espace terrestre, ce qui implique des fonctions d'*analyse spatiale* (non courantes dans les traitements d'information classiques) et de *synthèse* pour l'aide à la décision ;
2. La représentation visuelle de cet espace sous la forme d'une *carte* ou d'un *plan*, ce qui implique des fonctions de conception et de production cartographiques, constituant en elles-mêmes un langage différent du langage ordinaire (textuel ou numérique)¹⁷⁷.

Ce domaine prend son essor dans les années 1970 et est appliqué à un nombre grandissant de disciplines, comme l'histoire, l'archéologie ou l'histoire de l'art. Avec les humanités numériques, les SIG ont considérablement évolué et démontrent « l'intérêt et l'utilité de l'analyse spatiale et de son intégration dans la « boîte à outils » des praticiens du numérique en SHS¹⁷⁸. » De nombreuses librairies, comme le module MapSF¹⁷⁹ avec R ou Folium¹⁸⁰ avec Python, permettent à ceux qui maîtrisent la programmation informatique de créer des cartes aisément. Plus récemment, des logiciels libres ont également été conçus pour ceux qui ne disposent pas de compétences techniques et qui voudraient créer leur propre carte. Khartis¹⁸¹, Leaflet¹⁸², CartoLibre¹⁸³, Framacarte¹⁸⁴, parmi d'autres, sont autant d'applications en ligne qui facilitent le travail de visualisation des données spatiales.

Dans les études théâtrales, ces technologies ont également trouvé leur place afin de mettre en lumière des dynamiques spatiales qui seraient imperceptibles à l'œil nu. Nous choisissons trois exemples afin d'illustrer la manière dont les SIG contribuent à une autre compréhension des cultures théâtrales.

Exemple 1 : Ibsen Stage

La base de données IbsenStage, que nous avons déjà mentionnée dans le premier chapitre¹⁸⁵, met à disposition de ses utilisateurs plusieurs cartes numériques fondées sur leur corpus. En fonction des filtres choisis (« Titres des pièces », « Contributeurs », « Pays » et

¹⁷⁷Jean DENEGRE, François SALGE, « Introduction aux systèmes d'information géographique », *Les systèmes d'information géographique*, Paris, Presses Universitaires de France, 2004, p. 5 – 11.

¹⁷⁸Carmen BRANDO, Francesca FRONTINI, Dominic MOREAU, Mathieu ROCHE et Éric MASSON, « Introduction. Humanités numériques et analyses spatiales : enjeux et perspectives », *Humanités numériques* [En ligne], no. 3, 2021, publié 01 mai 2021, URL : <http://journals.openedition.org/revuehn/2038>, consulté le 16 mai 2022.

¹⁷⁹<https://cran.r-project.org/web/packages/mapsf/index.html>[En ligne], consulté le 16 mai 2022.

¹⁸⁰<https://python-visualization.github.io/folium/modules.html>[En ligne], consulté le 16 mai 2022.

¹⁸¹<https://www.sciencespo.fr/cartographie/khartis/>[En ligne], consulté le 16 mai 2022.

¹⁸²<https://leafletjs.com/>[En ligne], consulté le 16 mai 2022.

¹⁸³<https://carto.framasoft.org/>[En ligne], consulté le 16 mai 2022.

¹⁸⁴<https://framacarte.org/fr/>[En ligne], consulté le 16 mai 2022.

¹⁸⁵ Voir pages p. 33 – 35.

« Langues »), il est possible d'examiner la répartition et l'évolution des caractéristiques des pièces d'Ibsen dans le monde à partir de la fin du XIXe siècle jusqu'à nos jours.

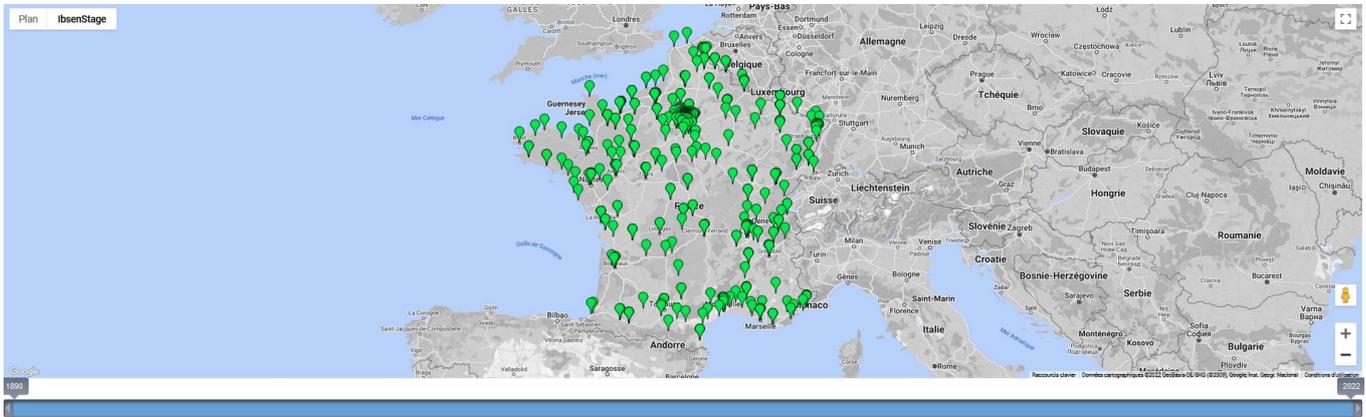


Fig. 38 : Représentation des pièces d'Ibsen jouées en français dans le monde entre 1890 et 2022.

Exemple 2 : TheatreScapes, Mapping Global Theatre Histories

L'outil de recherche « TheatreScapes, Mapping Global Theatre Histories¹⁸⁶ », mis au point à l'Université de Munich à partir de 2016, est un autre exemple qui met en son centre les SIG pour explorer l'histoire du théâtre à plus grande échelle. A la croisée de l'ethnologie, de la sociologie et de l'histoire, ce projet s'inscrit dans une perspective transnationale afin d'étudier le théâtre et l'opéra du point de vue des circulations dont ils résultent. En se focalisant sur la période de la première mondialisation (1850 – 1920), TheatreScapes a pour finalité de montrer comment les échanges et les flux culturels de cette période ont façonné le théâtre. Bien qu'elle soit encore en construction, leur base de données réunit déjà plus de 1000 théâtres dans le monde ainsi que 85 personnes différentes (acteurs, impresari, managers...).

¹⁸⁶ <https://www.theatrescapes.theaterwissenschaft.uni-muenchen.de/index.html> [En ligne], consulté le 19 mai 2022.

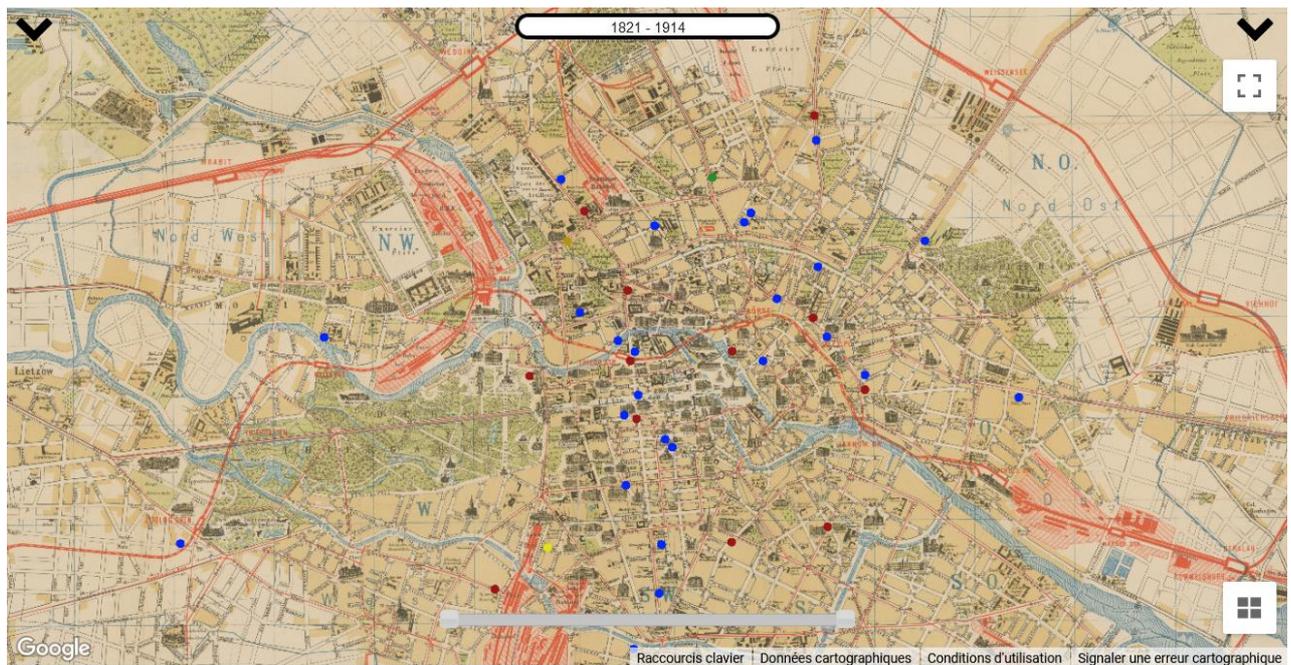


Fig. 39 : Représentation des théâtres à Berlin de 1821 et 1914.

Exemple 3 :Circuit: Mapping Theatre Performances in Victoria

Enfin, la plateforme « Circuit: Mapping Theatre Performances in Victoria » est générée à partir de la base de données AusStage¹⁸⁷ afin de représenter tous les types de spectacles (résidences d’artistes, mises en scène créées par des étudiants, ateliers...) qui ont eut lieu en Australie de 1963 jusqu’à nos jours. Grâce à la mise en place d’une barre de recherche, les utilisateurs peuvent sélectionner les caractéristiques qui les intéressent afin d’obtenir les visualisations de leur choix.

¹⁸⁷<https://blogs.unimelb.edu.au/creative-convergence/creative-convergence-the-circuit/>[En ligne], consulté le 16 mai 2022.

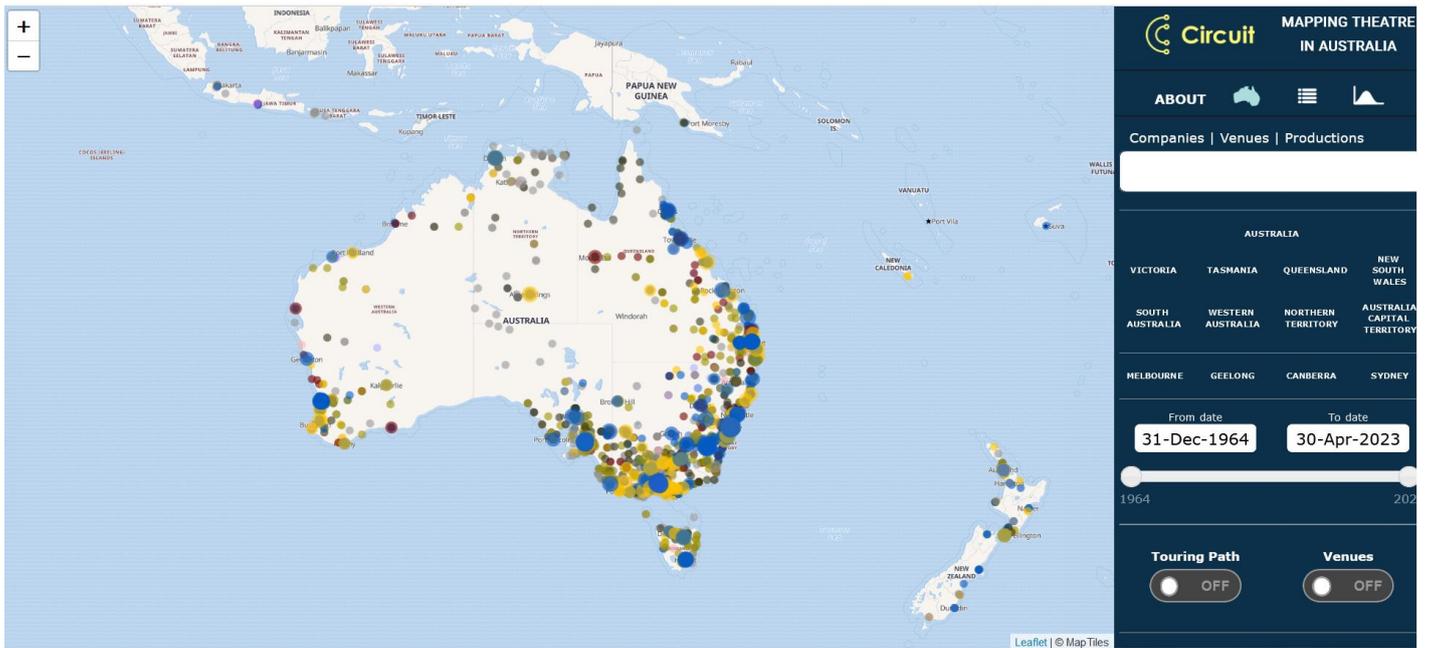


Fig. 40 : Représentation des théâtres en Australie de 1963 à nos jours.

III.B.1.3 Création d'une carte géographique des lieux fréquentés par la critique journalistique et numérique

Etape 1 : Nettoyage et uniformisation des données spatiales

Comme pour les exemples précédents, nous avons également généré une carte à partir de notre corpus afin d'examiner la critique théâtrale d'un point de vue spatial. Notre travail se décompose en trois étapes. La première repose sur un travail de nettoyage et d'uniformisation des noms de théâtre. Cette première partie consiste à re-examiner les mentions géographiques de l'ensemble de nos fichiers Excel. Parmi les critiques que nous avons extraites automatiquement d'Internet, certaines d'entre elles portent sur des théâtres à l'extérieur de Londres. Cela est dû à la structure des blogs. Un grand nombre de blogueurs publient leurs textes les uns à la suite des autres, sans faire de distinction entre ceux qui ont été rédigés à Londres et les autres. A l'inverse, les critiques du premier corpus ont préalablement été divisées entre celles qui ont été écrites pour les spectacles de la capitale et celles pour les théâtres régionaux. Il nous faut donc d'abord éliminer toutes les critiques qui n'analysent pas les spectacles de Londres.

Dans un deuxième temps, nous uniformisons tous les noms de théâtres. Tandis que certains ont changé de dénomination, comme le Coronet Theatre qui était anciennement appelé le Print Room Theatre, d'autres comportent plusieurs mentions pour le même lieu. Cela est le cas pour le National Theatre, qui est parfois appelé Olivier Theatre, Lyttelton Theatre ou Dorfman Theatre. Les trois noms font référence à la même localisation géographique, mais à des salles différentes au sein du même théâtre. Outre ces cas particuliers, nous devons également attribuer la même orthographe à tous les noms de théâtre. Pour un système informatique, ces trois orthographes ne correspondent pas au même lieu par exemple : « National Theatre, « The National » ou « national theatre. » Ce travail d'uniformisation nous permet ainsi de passer de plus de 2000 localisations différentes à plus de 400 distinctes.

Etape 2 : Enrichissement des métadonnées

La deuxième étape consiste à enrichir notre base de données afin d'obtenir des résultats plus détaillés pour nos analyses. Nous ajoutons trois types de métadonnées que nous avons déjà présentés brièvement dans le chapitre précédent¹⁸⁸ :

- **La capacité d'accueil des lieux** (lorsque cette information est disponible), ce qui nous permet de savoir si les deux communautés favorisent des endroits avec un grand nombre de places ou non.
- **Le lien vers le site Internet du théâtre**, ce qui nous permet non seulement d'avoir un aperçu de l'architecture des lieux, mais aussi une meilleure compréhension de leur politique culturelle grâce aux informations données sur leur site.
- **La catégorie de lieu**. Au-delà des théâtres, les critiques se rendent également dans d'autres endroits qui accueillent des spectacles. Les institutions académiques, les restaurants, les hôtels, parmi d'autres, sont autant de lieux mentionnés dans notre corpus. Afin de mieux saisir dans quels types de lieu les journalistes et les blogueurs se rendent, nous les classons en 24 catégories différentes. Le tableau suivant (Fig. 41) présente cette classification¹⁸⁹ :

¹⁸⁸ Voir pages 92 – 93.

¹⁸⁹ La liste de toutes les catégories de lieux est disponible en annexes p. 240 – 250.

Numéro	Catégorie de lieu	Exemples
1	Academic Place	King's College London, Goldsmith University, London College of Music...
2	Arts Association	Rag Factory, Swan Wharf, Applegarth Arts...
3	Arts Centre	Space Arts Centre, Waterman's Arts Centre, Bernie Grant Arts Centre...
4	Building	Tower of London, Bussey Building, Balfour Tower...
5	Charity Association	Pushkin House, Trinity House, Sydenham Centre...
6	Cinema	Empire Cinema Haymarket, Rio Cinema, Riverside Studios...
7	Concert Place	Wilton Music Hall, Wigmore Hall, Fairfield Halls...
8	Cultural Centre	South Bank Centre, Barbican, British Library...
9	Gallery	Saatchi Gallery, Hoxton Gallery, Guildhall Art Gallery...
10	Hotel	Charing Cross Hotel, May Fair Hotel, Green Rooms Hotel...
11	Library	London Library, Putney Library, Kensington Central Library...
12	Miscellaneous	Stratford Circus, London Zoo, Lords Disco...
13	Museum	British Museum, Victoria and Albert Museum, National Portrait Gallery...
14	Natural Space	Russell Square Gardens, Finsbury Park, Marylebone Gardens...
15	Online	Online
16	Pub	Glory Pub, Vibe Bar, New Moon Pub...
17	Pub Theatre	Bread and Roses Theatre, Landor Theatre, Drayton Arms Theatre...
18	Public Space	Brunel Tunnel, Borough Market, Covent Garden...
19	Reception venue	Gray's Inn Hall, Free Word Centre, Cecil Sharp House...
20	Religious Place	Saint Paul's Church, Caroline Gardens Chapel, Saint James Church...
21	Restaurant Area	Zedel Brasserie, Cock Tavern, Canvas Café...
22	Studio	Abbey Road Studios, Dance Attic Studios, Cable Street Studios...
23	Theatre	National Theatre, Royal Court Theatre, Orange Tree Theatre...
24	Town Hall	Shoreditch Town Hall, Limehouse Town Hall, Islington Town Hall...

Fig. 41 : Tableau des 24 catégories de lieux.

Etape 3 : Visualisation et analyse des données

La dernière étape repose sur la visualisation et l'analyse des données que nous avons collectées et organisées. Afin de générer une carte, nous devons d'abord récupérer automatiquement la latitude et la longitude de chacun des lieux que nous avons identifiés. Pour cela, nous ré-utilisons une méthode que nous avons déjà employée pour récupérer automatiquement le contenu des blogs : le web scraping. Une fois ces coordonnées géographiques répertoriées, il suffit ensuite de les replacer sur une carte contemporaine de Londres. Enfin, afin d'analyser plus finement ces données, nous utilisons Google Studio pour produire plusieurs graphes qui sont présentés en figure 45. Nous décidons de mettre en avant quatre caractéristiques :

- Le nombre de lieux visités par les journalistes et les blogueurs
- Le pourcentage des différentes catégories de lieux fréquentés par ces deux communautés (en excluant les théâtres)
- Le top 10 des théâtres les plus populaires selon ces deux groupes
- Une liste de tous les lieux, avec leur nombre de places, la catégorie à laquelle ils appartiennent et le nombre de fois qu'ils ont été fréquentés par les critiques du premier et du second corpus

Les résultats graphiques ainsi que leurs analyses se trouvent dans la partie suivante. Avant de les présenter, nous avons résumé le protocole de création des cartes que nous venons de détailler sous forme de schéma (Fig. 42) :

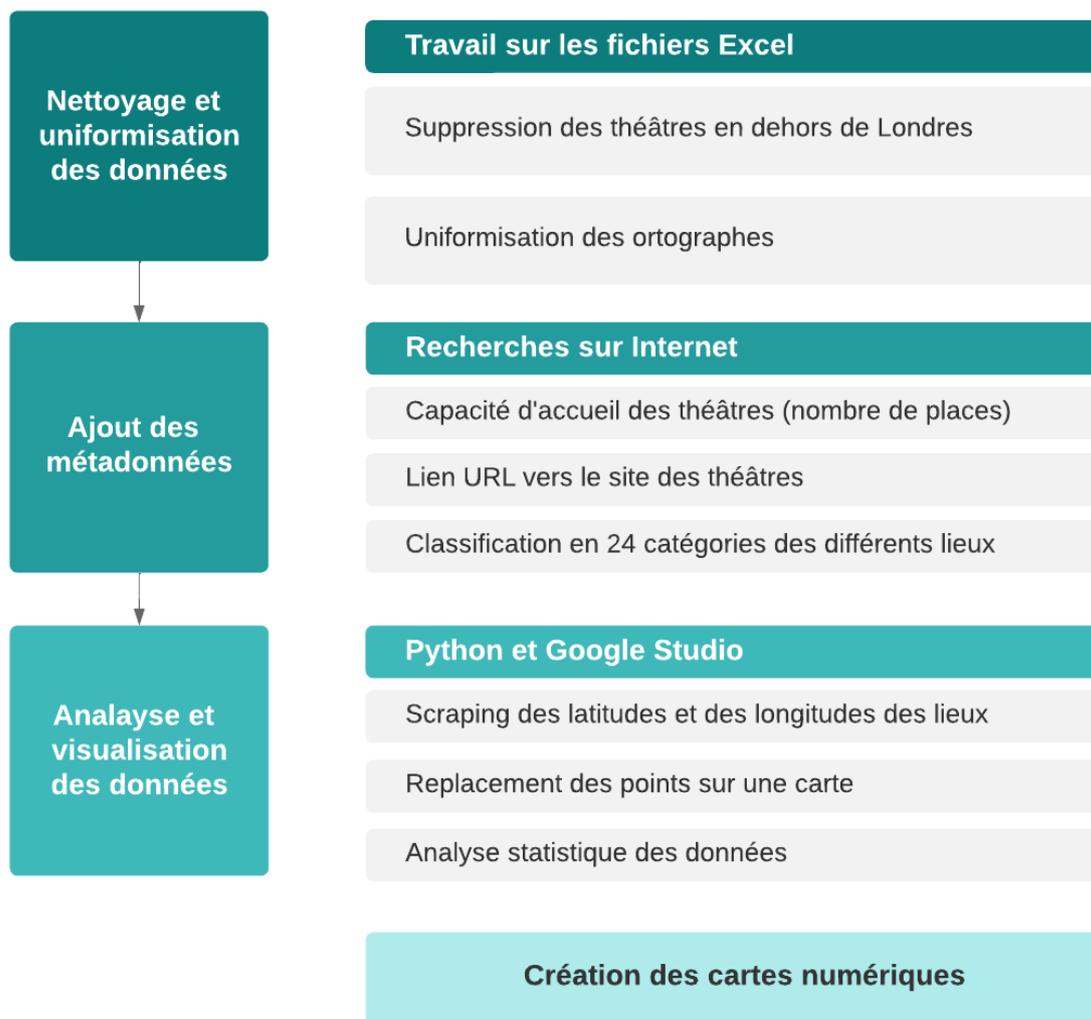


Fig. 42 : Schéma des trois étapes successives pour la réalisation de la cartographie numérique.

III.B.2 Deux communautés aux intérêts divergents

III.B.2.1 Résultats

Lieux visités par les journalistes de Londres (2010 - 2020)

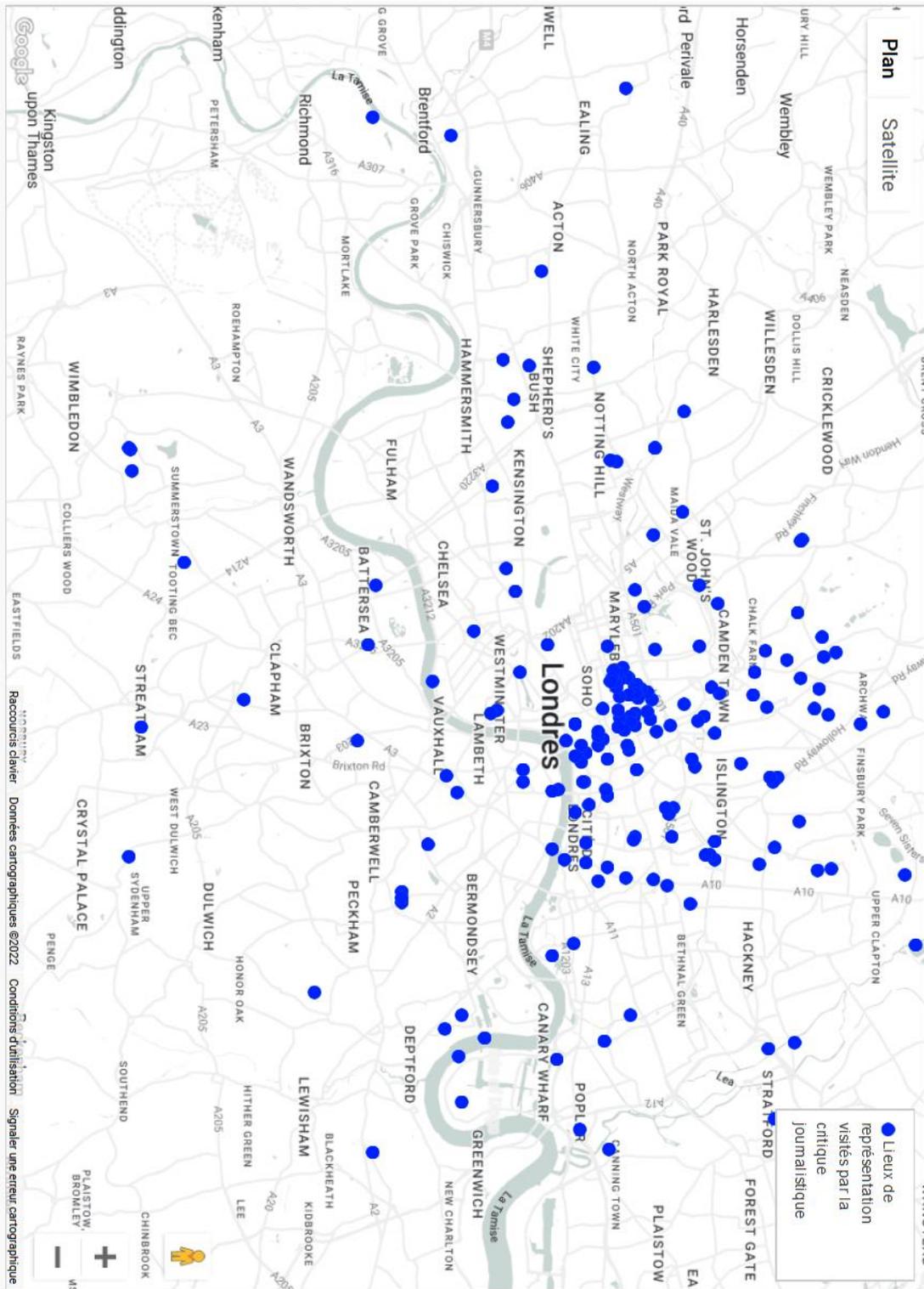


Fig. 43 : Capture d'écran de la cartographie numérique réalisée pour le corpus I.

Lieux visités par les blogueurs de Londres (2010 - 2020)

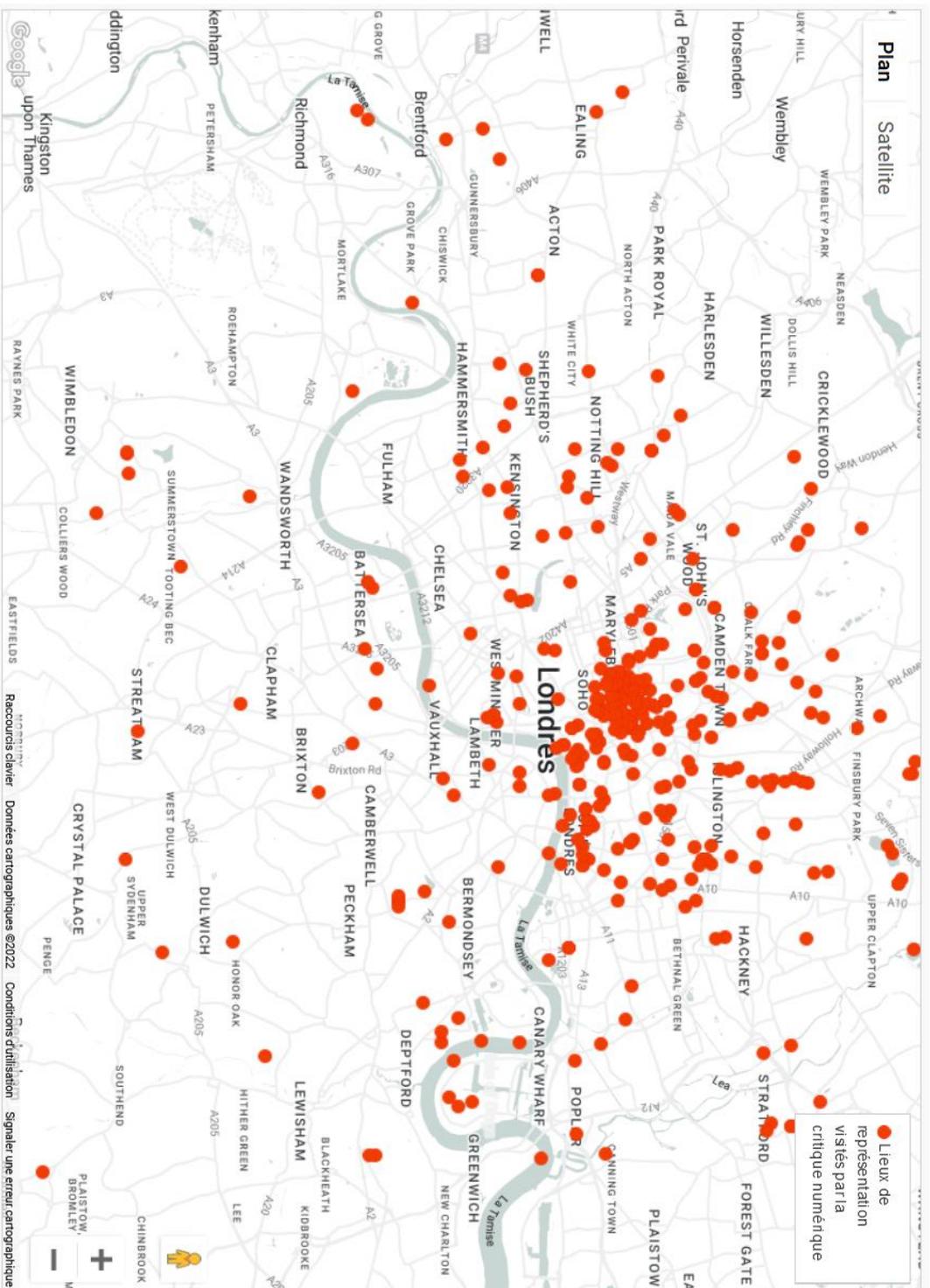
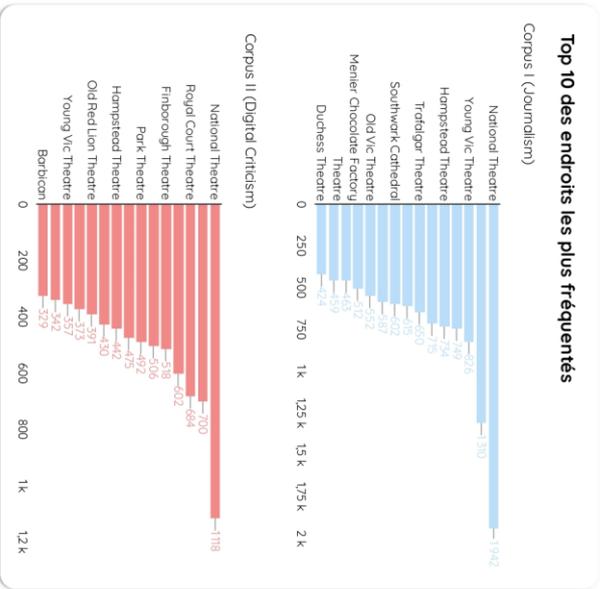
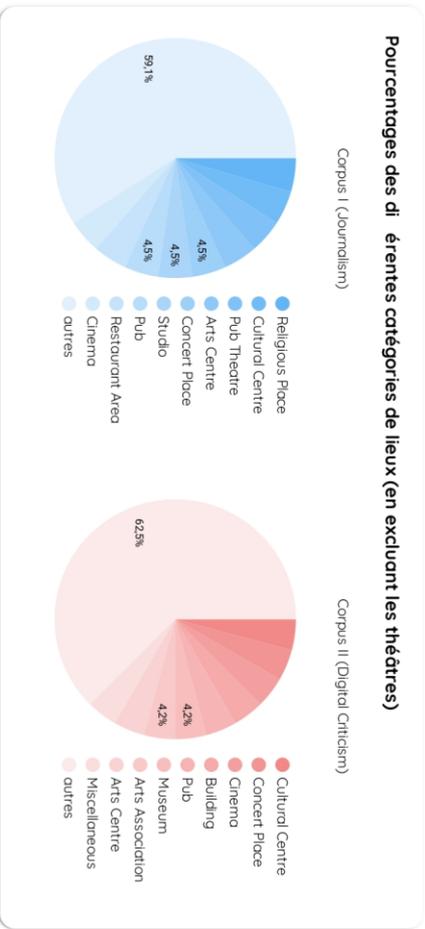
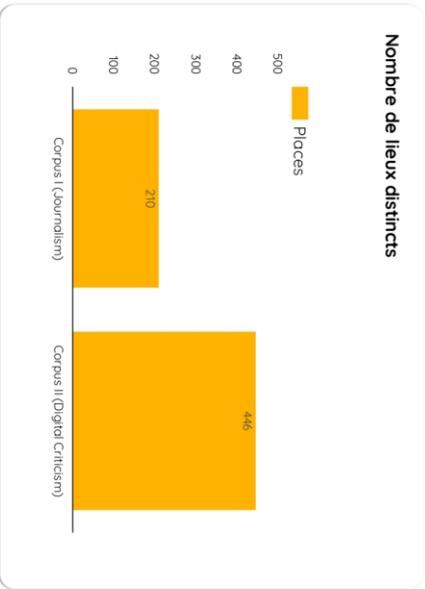


Fig. 44 : Capture d'écran de la cartographie numérique réalisée pour le corpus II.

Analyse des lieux visités par la critique journalistique et numérique londonienne (2010 - 2020)



Name of the Place	Category	Seating Capacity	Corpus	Visited
National Theatre	Theatre	1676	Corpus I (Journalism)	1942
Royal Court Theatre	Theatre	465	Corpus I (Journalism)	1310
National Theatre	Theatre	1676	Corpus II (Digital Criticism)	1188
Young Vic Theatre	Theatre	420	Corpus I (Journalism)	826
Shakespeare's Globe Theatre	Theatre	1570	Corpus I (Journalism)	749
Hampstead Theatre	Theatre	405	Corpus I (Journalism)	734
Barbican	Cultural Centre	null	Corpus I (Journalism)	715
Southwark Playhouse Theatre	Theatre	null	Corpus II (Digital Criticism)	700
Royal Court Theatre	Theatre	465	Corpus II (Digital Criticism)	684
Trafalgar Theatre	Theatre	630	Corpus I (Journalism)	650
Donmar Theatre	Theatre	251	Corpus I (Journalism)	615
Soho Theatre	Theatre	140	Corpus II (Digital Criticism)	602
Southwark Cathedral	Religious Place	null	Corpus I (Journalism)	602
Almeida Theatre	Theatre	null	Corpus I (Journalism)	587
Old Vic Theatre	Theatre	1067	Corpus I (Journalism)	552
Finborough Theatre	Theatre	50	Corpus II (Digital Criticism)	518
Arcola Theatre	Theatre	null	Corpus I (Journalism)	512
Vaults Theatre	Theatre	176	Corpus II (Digital Criticism)	506

Fig. 45 : Tableau de bord des caractéristiques principales des lieux fréquentés par les journalistes et les blogueurs.

Intégration des deux cartographies numériques sur DramaCritiques.com avec un formulaire de recherche à six filtres

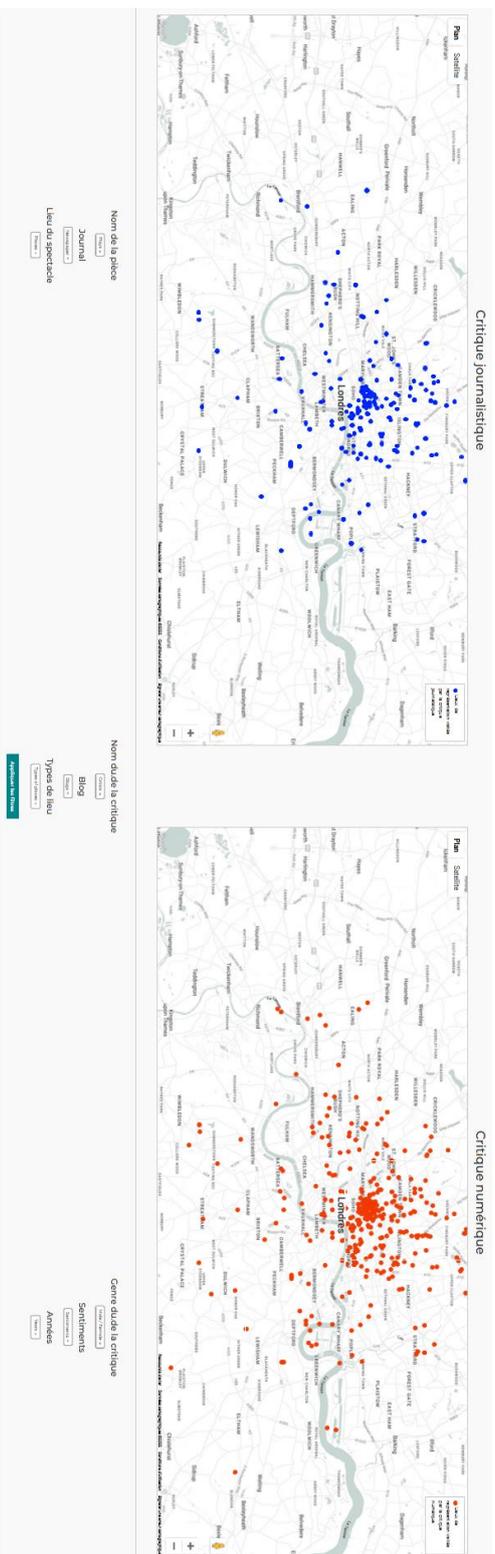


Fig. 46 : Présentation des deux cartographies numériques sur DramaCritiques.com

Intégration de la liste des pièces jouées dans chacun des lieux avec leur nombre de critiques

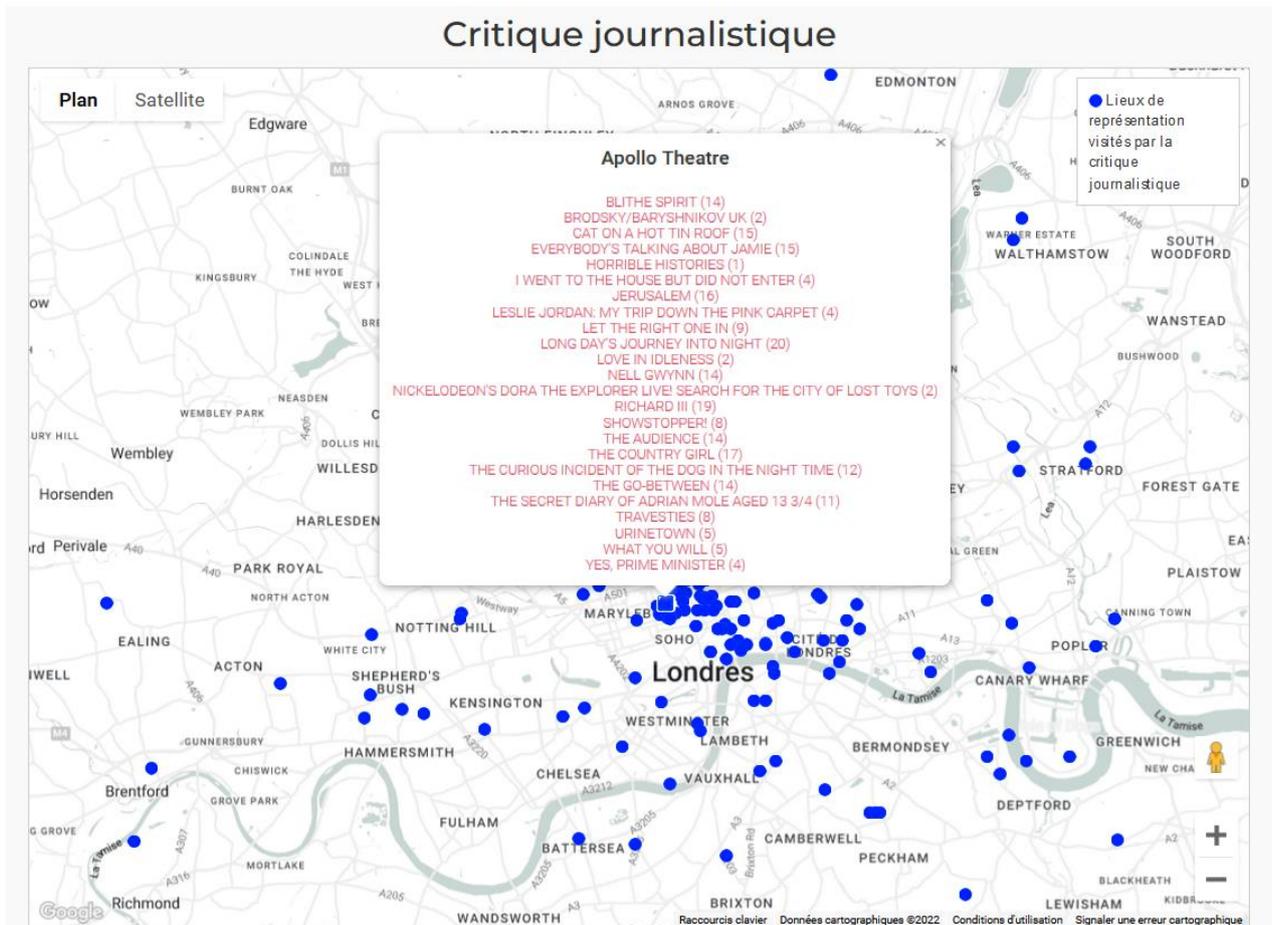


Fig. 47 : Présentation des fonctionnalités ajoutées aux deux cartes numériques sur DramaCritiques.com

Mise en place d'une page avec le site Internet du théâtre et accès à la visualisation de chacune des critiques

APOLLO THEATRE
<https://www.apollotheatrelondon.co.uk/>

An independent guide to

APOLLO THEATRE

31 Shaftesbury Avenue, London W1D 7ES

HOMECRUISEBLIPPI THE MUSICALTHE UPSTART CROWDERREN BROWN - SHOWMANTICKETS🔍

SEATINGMAP AND NEARBYABOUT US

APOLLO THEATRE

Cruise is a poignant and moving examination of London's LGBTQ community in the late 1980s

THE URGENT, MOVING, INSPIRATIONAL NEW PLAY
ARIA ENTERTAINMENT AND LAWRETT JACKSON PRESENTS



RECENT POSTS

- [Blippi The Musical](#)
- [Derren Brown - SHOWMAN](#)
- [The Upstart Crow](#)
- [Cruise](#)
- [Adam Kay: This is Going to Hurt](#)

Pièces

- [BLITHE SPIRIT \(14\)](#)
- [BRODSKY/BARYSHNIKOV UK \(2\)](#)
- [CAT ON A HOT TIN ROOF \(15\)](#)
- [EVERYBODY'S TALKING ABOUT JAMIE \(15\)](#)
- [HORRIBLE HISTORIES \(1\)](#)
- [I WENT TO THE HOUSE BUT DID NOT ENTER \(4\)](#)
- [JERUSALEM \(16\)](#)
- [LESLIE JORDAN: MY TRIP DOWN THE PINK CARPET \(4\)](#)
- [LET THE RIGHT ONE IN \(9\)](#)
- [LONG DAY'S JOURNEY INTO NIGHT \(20\)](#)
- [LOVE IN IDLENESS \(2\)](#)
- [NELL GWYNN \(14\)](#)
- [NICKELODEON'S DORA THE EXPLORER LIVE! SEARCH FOR THE CITY OF LOST TOYS \(2\)](#)
- [RICHARD III \(19\)](#)
- [SHOWSTOPPER! \(8\)](#)

Fig. 48 : Ajout de différentes fonctionnalités aux cartographies numériques sur DramaCritiques.com

III.B.2.2 Interprétation

Ces résultats permettent de visualiser que la rupture entre le canon de la critique théâtrale londonienne et ses périphéries est très marquante d'un point de vue géographique. Comme le montrent les deux premières cartes géographiques (figures 43 et 44), au contraire de la critique journalistique, les points qui représentent la critique numérique ne se concentrent pas uniquement au centre de Londres, mais aussi dans ses périphéries. Le nombre de théâtres diverge également d'un corpus à l'autre. Les statistiques du tableau de bord que nous avons produit en figure 45 le prouvent : tandis que l'on peut compter 210 lieux différents dans le corpus I, il y en a 446 dans le second corpus, soit deux fois plus. Cette diversité de lieux laisse sous-entendre que les blogueurs se rendent dans des endroits plus marginaux qui, nous pouvons le supposer, proposent d'autres types de spectacles que ceux disponibles dans les grands théâtres du centre de Londres.

Peut-être plus parlant encore est le type d'endroits dans lesquels se rendent les critiques de ces deux communautés. Le classement de ces 446 lieux en 24 catégories distinctes révèle des différences significatives concernant le type d'endroit que les journalistes et les blogueurs fréquentent. Lorsque nous regardons plus en détail les pourcentages de ces catégories sur le tableau de bord en figure 45, nous pouvons remarquer que quatre types d'endroits distinguent ces deux communautés. Concernant les journalistes, ce sont les « Religious places », les « Pub theatres », les « Studios » et les « Restaurant aeras » qui les démarquent des blogueurs puisque ces derniers ne fréquentent quasiment pas ces lieux. Pour les blogueurs, les catégories qui les différencient des journalistes sont les suivantes : les « Buildings », les « Museums », les « Arts associations » et tous les lieux que nous n'avons pas pu classés et que nous avons répertoriés sous le terme de « Miscellaneous », ou « Divers. »

Les lieux peu visités par les journalistes mais préférés par les blogueurs sont caractérisés par leur ouverture, tant physique que métaphorique. La catégorie « Building » par exemple inclut des représentations théâtrales qui ont eu lieu dehors (Tower of London, Great Eastern Quay, Millbank Tower...). Concernant les associations artistiques (« Arts associations »), ce sont la pluralité des activités et l'inclusion de tous les types de population qui sont mises à l'honneur dans ces lieux. Danse, chant, couture et expositions, pour ne citer que quelques exemples, sont autant d'animations proposées dans ces lieux destinés, pour la plupart, aux personnes avec un revenu financier limité. Afin de rendre compte de la diversité

prônée par les associations d'art, nous avons répertorié les phrases de présentation disponibles sur leur site. Quatorze lieux au total constituent cette catégorie :

Catégorie « Arts Associations »	Présentation sur leur site
Applecarts Arts	We at Applecarts Arts believe that stories are potent and have the power to challenge, change and strengthen communities. (...) We work with a wide range of artistic disciplines – creating original music, film, visual art and drama. ¹⁹⁰
Bell House	<i>Bell House</i> is an educational centre that offers support outside the mainstream school curriculum, lifelong learning, exhibitions, talks and musical events. ¹⁹¹
Chats Palace	Chats Palace is an arts centre with a civic role , empowering local people. (...) Our key focus areas are with those who are marginalised, excluded or without a voice (...) . ¹⁹²
Greenwich Dance	Greenwich Dance finds inspiring spaces for innovative dance creation, performance and community participation: creating quality opportunities to make, watch and take part in dance regardless of age, background or ability . ¹⁹³
Highbury Opera Theatre	Highbury Opera Theatre is run by a team of artists and arts-lovers with an association with Islington, most of whom live or work around Highbury. They come from all walks of life and backgrounds , sharing a love of bringing performing arts to their local community . ¹⁹⁴
Jaguar Shoes Collective	JAGUARSHOES COLLECTIVE constitutes a group of creatively motivated businesses and individuals working in art, film, fashion, music, publishing and design . We provide many platforms for creative talent; from exhibitions and events, to collaborative products and retail opportunities . ¹⁹⁵
Lauderdale House	We offer different things to a wide range of different people , always seeking to reflect the original charitable gift of the house and park as a centre for arts and heritage for all and a ' garden for the gardenless '. ¹⁹⁶
National Youth Music Theatre	National Youth Music Theatre represents the very best in work with young people through musical theatre enabling thousands of youngsters across the UK (...) ¹⁹⁷

¹⁹⁰https://applecartsarts.com/about_us.php[En ligne], consulté le 22 juin 2022.

¹⁹¹<https://www.bellhouse.co.uk/>[En ligne], consulté le 22 juin 2022.

¹⁹²<https://chatspalace.com/past-present/>[En ligne], consulté le 22 juin 2022.

¹⁹³<https://greenwichdance.org.uk/about-us/>[En ligne], consulté le 22 juin 2022.

¹⁹⁴<https://www.highburyoperatheatre.com/who-we-are-1/>[En ligne], consulté le 22 juin 2022.

¹⁹⁵<http://www.jaguarshoes.com/>[En ligne], consulté le 22 juin 2022.

¹⁹⁶<https://www.lauderdalehouse.org.uk/about-us>[En ligne], consulté le 22 juin 2022.

¹⁹⁷<https://nymt.org.uk/>[En ligne], consulté le 22 juin 2022.

National Youth Theatre	We seek out talented creative young people from a diverse range of backgrounds through an active and inclusive audition programme and community engagement (...) We empower our young people with creative leadership, inclusive practice and intersectionality training and actively work to address young people who are underrepresented in our community and industry, proactively breaking down barriers that prevent young people from taking part in our work. ¹⁹⁸
Platform Southwark	Platform is a contemporary art gallery, theatre, and studio space in Waterloo. (...) We connect emerging practitioners in the arts and encourage interdisciplinary collaboration. All our public events are free for the local community of Southwark. ¹⁹⁹
Rag Factory	The Rag Factory was set up in early 2006 to provide affordable space for people working in the creative industries, whether performance-based, such as theatre, film, dance, music and fashion or visual arts such as painting, sculpture, photography, installation and graphic design. ²⁰⁰
Redbridge Drama Centre	Redbridge Drama Centre (RDC) is located in South Woodford and is a centre for creating, performing and watching theatre for all ages. ²⁰¹
Swan Wharf	Swan Wharf is a unique centre for creativity and arts, collaboration and culture. Combining artist studios, woodworking and metalworking workshops, tech offices, gallery & exhibition spaces, event space, collaborative areas and a stunning canalside bar. ²⁰²
Ugly Duck	Ugly Duck is a London based arts organisation that supports under-represented voices and emerging artists. Established in August 2012, their programme enables makers, community groups, professionals and the public to come together around unique cultural experiences and curated events. ²⁰³

Fig. 49 : Description des différents lieux inclus dans la catégorie « Arts Association. »

« Those who are marginalised, excluded or without a voice, » « individuals working in art, film, fashion, music, publishing and design », « a garden for the gardenless », « inclusive practice and intersectionality training »... Toutes ces dénominations illustrent la volonté des

¹⁹⁸<https://www.nyt.org.uk/about-us>[En ligne], consulté le 22 juin 2022.

¹⁹⁹<http://www.platformsouthwark.co.uk/about/>[En ligne], consulté le 22 juin 2022.

²⁰⁰<https://www.towerhamletsarts.org.uk/?cat=30&cid=44376&guide=venues>[En ligne], consulté le 22 juin 2022.

²⁰¹<https://www.redbridge.gov.uk/leisure-sport-and-the-arts/redbridge-drama-centre/>[En ligne], consulté le 22 juin 2022.

²⁰²<https://www.hackneywicked.co.uk/venues/swan-wharf/>[En ligne], consulté le 22 juin 2022.

²⁰³<https://uglyduck.org.uk/about-us/>[En ligne], consulté le 22 juin 2022.

acteurs de ces lieux d'inclure les populations marginales aux sphères artistiques. Ce constat renforce l'hypothèse selon laquelle les blogueurs favorisent la culture périphérique de la capitale.

Une autre catégorie de lieux très fréquentés par les blogueurs et peu visités par les journalistes concerne tous les endroits que nous n'avons pas pu classer et que nous avons ainsi regroupés sous le terme de « Miscellaneous », ou « Divers ». Parmi ces 25 lieux qui constituent cette catégorie, figurent deux parkings (Smithfield car park et Doon car park), le zoo de Londres, la discothèque Lords, le cirque de Stratford, la librairie Calder Bookshop and Theatre, la grande enseigne Selfridges et le Rowans Bowling. Tous ces endroits originaux dans lesquels les critiques numériques s'aventurent pour y découvrir des performances confirment leur curiosité et leur désir de prendre en compte la culture théâtrale londonienne dans son ensemble. Tant d'un point de vue géographique qu'artistique, ces lieux éloignés des grandes institutions théâtrales proposent d'autres types de spectacle avec, nous pouvons le supposer, des moyens humains, techniques et financiers bien moins importants que ceux dont disposent les grands théâtres du centre de Londres.

Bien que les journalistes fréquentent également d'autres types de lieux que les seules institutions théâtrales, ils tendent néanmoins à fréquenter des endroits soit plus touristiques, soit plus normés. Les « Religious places » par exemple font référence aux lieux de cultes de Londres, comme Westminster Abbey, Southwark Cathedral ou Saint Paul's Church, pour n'en citer que quelques-uns. Même si certaines de ces performances ont eu lieu en extérieur, ce qui dénote un autre type de théâtre, il n'en reste pas moins que ces lieux touristiques attirent davantage de spectateurs et s'éloignent des performances marginalisées sur les parkings que les blogueurs analysent. De la même manière, nous pouvons supposer que les spectacles qui ont eu lieu dans les « pub theatres » ou dans les « restaurant aeras » s'adressent à une certaine catégorie de la population : celle qui peut se permettre de manger régulièrement à l'extérieur. L'analyse géographique du canon de la critique londonienne et de ses périphéries permet ainsi de déduire qu'il y a une véritable scission entre ces deux communautés d'un point de vue spatial.

III.B.2.3 Limites de l'approche

Bien que ces résultats nous permettent d'aborder notre problématique sous un angle nouveau et d'approfondir notre réflexion sur la relation qui caractérise ces deux communautés, de nombreuses améliorations restent envisageables, tant d'un point de vue esthétique qu'analytique. Premièrement, la différence du nombre de théâtres d'un corpus à l'autre pourrait être rendue plus frappante grâce à l'amélioration du visuel de nos cartographies. N'étant pas accoutumés aux Systèmes d'Information Géographique, nous avons eu recours à des méthodes simples dans ce domaine. Cependant, nous avons pensé à adapter la taille et la couleur des points sur la carte en fonction du nombre de critiques rédigées dans les théâtres. Cela aurait permis par exemple de comprendre instantanément dans quels lieux les critiques se rendaient le plus. Nous pourrions également indiquer par un système de couleur la taille des théâtres en fonction de leur nombre de places. Cela compléterait nos analyses.

D'un point de vue analytique, d'autres expériences pourraient être menées à partir de notre base de données afin d'obtenir des résultats plus précis. Nous pourrions effectuer des calculs en fonction des trente-deux quartiers, ou *boroughs*, de Londres. Cela nous permettrait de faire une distinction entre les théâtres appartenant à l'*Inner London*, qui est composé de douze quartiers, et ceux faisant partie de l'*Outer London*, qui est constitué de vingt quartiers. Le nombre de théâtres, le nombre moyen de places assises par théâtre et le nombre de visites des critiques dans ces lieux en fonction de chacun des quartiers nous apporteraient une vision plus précise de la répartition géographique des journalistes et des blogueurs dans la capitale. Ces analyses nécessiteraient cependant une amélioration de notre cartographie numérique afin d'extraire les données propres à chacun de ces quartiers.

Examiner plus en détail l'histoire et la politique des lieux que nous avons répertoriés nous aiderait également à apporter des résultats plus nuancés. C'est pourquoi nous avons intégré les liens des sites Internet des théâtres à nos analyses. Ces sites Internet permettent d'obtenir de précieuses informations liées à la politique culturelle de ces endroits. Comme le montrent les figures 46, 47 et 48, en cliquant sur les théâtres sur la carte, une autre page s'affiche dans le navigateur web. Cette nouvelle page permet non seulement aux utilisateurs de lire les critiques qui ont été rédigées sur les spectacles qui ont été joués dans ces lieux, mais aussi d'avoir accès au site Internet du théâtre.

Effectuer des recherches sur la politique et l'histoire de ces lieux nous permettrait de répondre à des questions comme : « Quand ont été créés ces lieux ? », « Sont-ils dirigés par un homme ou une femme ? », « Ont-ils tendance à produire des dramaturges femmes ou hommes ? », ou bien « Combien de spectacles y sont joués en moyenne par année ? » Répondre à ces questions demanderait un travail supplémentaire pour ajouter les métadonnées nécessaires (date de création des théâtres, genre du directeur du théâtre, etc.), mais contribuerait à cartographier plus finement le rapport qu'entretiennent ces deux communautés avec les espaces théâtraux.

Enfin, compléter cette « lecture distante » des cartes par une « lecture proche » des textes qui ont été rédigés dans ces lieux de représentation nous permettrait de comprendre ce qui attire les critiques à fréquenter certains endroits plutôt que d'autres. « Pourquoi les blogueurs se rendent-ils davantage dans les associations d'art et très peu auprès ou dans les églises par exemple ? Quels sont les éléments théâtraux qui retiennent leur attention dans ces lieux ? Préfèrent-ils les thématiques des pièces jouées dans ces lieux, les décors ou le jeu des acteurs ? » Procéder à des analyses « classiques » des textes des critiques pourrait ainsi esquisser des réponses à ces questions. Nous pourrions également réinvestir les techniques de zoning que nous avons déjà utilisées. La catégorie « Détails visuels, olfactifs et sonores » se concentre sur la description des lieux et prend en compte des éléments comme la scène, les costumes, les couleurs et l'atmosphère créée. En extrayant cette catégorie, nous pourrions essayer de comprendre d'un point de vue informatique ce qui séduit les critiques dans ces lieux.

Transition

Cette deuxième sous-partie avait pour finalité d'examiner la répartition géographique des journalistes et des blogueurs dans Londres, des années 2010 à 2020. La question directrice qui a structuré ce travail était la suivante : « Peut-on observer un clivage entre le canon de la critique théâtrale londonienne et ses périphéries d'un point de vue géographique ? » Cette thématique était pertinente pour mieux comprendre quels types de théâtre les spectateurs de ces deux communautés fréquentent. Nous avons utilisé les Systèmes d'Informations Géographiques et différents calculs statistiques pour répondre à notre question. Premièrement, en créant deux cartes de Londres, l'une qui représente les lieux dans lesquels vont les journalistes, l'autre qui se focalise sur les blogueurs, nous avons pu obtenir une visualisation des lieux visités par les critiques. Deuxièmement, en mettant au point un tableau de bord, nous avons obtenu une vision détaillée des types de lieux qu'ils fréquentent.

Ces résultats graphiques et statistiques nous ont permis d'observer deux éléments principaux. Le premier concerne le nombre de lieux fréquentés par les critiques. En se rendant dans deux fois plus d'endroits différents qui ne se situent pas seulement dans le centre de Londres, les blogueurs s'intéressent à une culture bien plus diversifiée que celle examinée par les journalistes. Le second élément que nous avons relevé confirme et renforce ce premier résultat. En classant les 446 endroits mentionnés dans notre corpus en 24 catégories différentes, nous avons pu observer que les blogueurs se rendent dans des lieux qui tendent à favoriser une culture dite « populaire. » Nous avons ainsi pu déduire de ces expériences que la fracture « centre / périphéries » est particulièrement présente d'un point de vue géographique entre les journalistes et les blogueurs.

La partie suivante s'attachera à examiner ces deux publics sous un autre angle. Nous avons voulu savoir si les préférences esthétiques de ces communautés variaient d'un corpus à l'autre. Dit autrement, « Les journalistes et les blogueurs mettent-ils en avant le même type de pièce ? » Pour répondre à cette question, nous avons fondé notre étude sur l'analyse d'opinion. L'analyse d'opinion consiste à déterminer automatiquement l'opinion d'un auteur dans un texte donné. Cette méthode informatique, populaire dans les humanités numériques, nous permettra d'évaluer sur l'ensemble de notre corpus les pièces qui ont été acclamées par la critique et celles qui, au contraire, ont reçu des avis peu favorables.

A nouveau, cette expérience sera structurée de la même manière que celle que nous venons de présenter à propos de la répartition spatiale de ces deux groupes. Nous commencerons par introduire le contexte littéraire qui nous a menés à adopter cette approche, avant d'établir un état de l'art des différents travaux qui reposent sur l'analyse d'opinion. Nous détaillerons ensuite les différents modèles algorithmiques que nous avons implémentés. Enfin, nous présenterons les résultats graphiques, les conclusions que nous pouvons tirer et les limites de notre expérience.

III.C Analyse d'opinion : étude des succès sur la scène londonienne

III.C.1 Les préférences esthétiques de ces deux groupes de spectateurs sont-elles similaires ?

III.C.1.1 Quel type de pièces les journalistes et les blogueurs préfèrent-ils ?

Le succès d'une représentation théâtrale représente d'importants enjeux pour les compagnies. Une pièce acclamée par le public et surtout par la critique littéraire n'aura pas le même rayonnement artistique que celle qui aura été jugée comme médiocre. La renommée du dramaturge, du metteur en scène, des comédiens, mais aussi de tous ceux qui participent à la mise en place d'une pièce s'en trouve par conséquent affectée. La polarité de ces deux notions n'est cependant pas si simple. Le discours qui se construit autour de la réussite ou de l'échec d'une pièce dépend du regard posé sur cette dernière, ou pour reprendre un des concepts-clés des théories de la réception, de « l'horizon d'attente » du spectateur. Un public d'adolescents n'apprécierait probablement pas de la même manière la tragédie shakespearienne *Hamlet* qu'un professeur de littérature classique par exemple.

La nature de notre corpus laisse supposer que la notion de succès variera d'une communauté littéraire à l'autre. Tandis que les journalistes incarnent la culture dite « classique », les blogueurs prennent le parti de représenter les marges du théâtre anglais. En 2012, la mise en scène par Sebastian Nübling de la pièce de Simon Stephens *Three Kingdoms* (2012) a démontré à quel point les préférences esthétiques de ces deux groupes divergent. Très peu d'étoiles furent accordées par la critique journalistique à cette pièce. Quentin Letts, critique dans le *Daily Mail*, décrivit la représentation comme « magnificently bad, laughably awful, a real honking turkey. » Pour Michael Coveney, critique dans le journal *What's on Stage*, toute personne qui appréciait la pièce était qualifiée de « debauched beyond redemption²⁰⁴. » A l'inverse, les blogueurs reçurent la pièce avec curiosité, admiration et enthousiasme.

²⁰⁴https://www.whatsonstage.com/west-end-theatre/news/review-round-up-three-kingdoms-divides-the-critics_4378.html[En ligne], consulté le 07 juillet 2022.

Nous avons ainsi voulu savoir si cet exemple était un cas particulier ou s'il était représentatif d'une dynamique plus générale entre les deux groupes. Dit autrement, « Est-il possible d'observer une opposition entre les pièces mises en avant par les journalistes et les blogueurs ? Leurs préférences esthétiques sont-elles également divergentes, ou au contraire, peut-on observer des points de convergence ? »

Afin de répondre à ces questions, nous allons avoir recours à l'analyse d'opinion pour explorer les textes de notre corpus. L'analyse d'opinion permet de définir automatiquement les opinions, les sentiments ou les attitudes d'un auteur dans un texte. Ces méthodes sont souvent utilisées par de grandes firmes pour comprendre si un produit a été aimé ou non par ses acheteurs. Bien que l'extraction d'informations subjectives soit plus délicate que l'extraction d'informations objectives, elle reste néanmoins un moyen pertinent pour aborder notre corpus. Elle nous permettra d'obtenir une vue d'ensemble des préférences esthétiques de ces deux groupes.

Nous avons divisé cette expérience en trois parties. La première sera consacrée à l'état de l'art. Nous expliquerons en quoi consiste l'analyse d'opinion et comment elle a été utilisée dans de précédents travaux en sciences humaines et sociales. Dans la deuxième partie, nous détaillerons les modèles informatiques que nous avons choisis et implémentés. Si l'analyse d'opinion semble particulièrement bien adaptée pour étudier notre corpus, plusieurs obstacles se sont néanmoins présentés à nous. Après avoir exposé les solutions que nous avons trouvées, nous concluons par les résultats obtenus et les analyses littéraires que nous pouvons en tirer.

III.C.1.2 Etat de l'art de l'analyse d'opinion

L'analyse d'opinion consiste à classer la polarité d'un texte donné, c'est-à-dire à déterminer si l'opinion exprimée est positive ou négative. Cette tâche est aussi parfois appelée analyse des sentiments. Dans ce cas-là, plusieurs éléments sont pris en compte, comme le dégoût, la joie, la peur, la surprise... Pour ce travail, nous nous sommes intéressés qu'à une classification binaire. Cette méthode nous semblait être la plus pertinente pour les critiques de théâtre. Certains systèmes proposent une classification à trois niveaux : positif, négatif et neutre, pour les textes qui ne véhiculent pas d'opinion particulière. Dans notre étude, nous avons choisi de ne retenir que deux catégories (positif versus négatif) car les critiques de

théâtre doivent généralement prendre position (explicitement ou implicitement, à travers le vocabulaire utilisé par exemple).

L'analyse d'opinion est populaire depuis plusieurs décennies²⁰⁵, et divers modèles ont été proposés, suite aux développements de la recherche en apprentissage automatique appliqué au Traitement du Langage Naturel (en particulier les classificateurs). Récemment, de nouveaux modèles de langage capables de représenter avec précision la sémantique des mots en contexte ont été mis au point. L'analyse d'opinion bénéficie ainsi de modèles capables de capturer le sens de toute une phrase (ou de tout un texte) grâce à une représentation fine de l'interaction des différents mots impliqués (adjectifs, adverbes, négations, etc.). Les approches populaires ont été les LSTM²⁰⁶ et surtout les modèles de langage dont le plus populaire est BERT²⁰⁷.

III.C.1.3 Calcul des pourcentages de positivité et de négativité des critiques

Présentation du modèle XLNet

Pour notre corpus, nous avons choisi d'utiliser XLNet²⁰⁸ qui est un modèle autorégressif récent. XLNet a été critiqué²⁰⁹ pour sa grande complexité de calcul et d'espace, mais le modèle a obtenu de bons résultats parmi les ensembles de tests standards, en particulier pour la tâche d'analyse de sentiments. C'est la principale raison pour laquelle nous avons choisi ce modèle. Dans le domaine du Traitement Automatique des Langues, l'extraction d'opinion est aujourd'hui l'une des tâches les plus populaires. La mise à disposition de grands ensembles de données annotées facilite le développement de ces différents systèmes et la comparaison de ces performances avec d'autres approches. Pour

²⁰⁵ Bo PANG et Lillian LEE, « Opinion Mining and Sentiment Analysis », *Foundations and Trends in Information Retrieval*, vol. 2, no. 1 – 2, 2008, p. 1 - 135.

²⁰⁶ Devendra Singh SACHAN, Manzil ZAHEER et Ruslan SALAKHUTDINOV, « Revisiting LSTM Networks for Semi-Supervised Text Classification via Mixed-Objective Function », *Proceedings of the AAAI Conference on Artificial Intelligence*, no. 33, 2019, p. 6940 – 6948.

²⁰⁷ Qizhe XIE, et al. « Unsupervised data augmentation for consistency training », *Proceedings of the 34th Conference on Neural Information Processing Systems*, no. 525, 2020, p. 6256 – 6268.

²⁰⁸ Zhilin YANG, et al. « XLNet: Generalized Autoregressive Pretraining for Language Understanding », Cite arxiv:1906.08237.

²⁰⁹ Michael HAHN, Dan JURAFSKY et Richard FUTRELL, « Sensitivity as a Complexity Measure for Sequence Classification Tasks », *Transactions of the Association for Computational Linguistics*, no. 9, 2021, p. 891 – 908.

l'anglais, les ensembles de données tels que IMDb²¹⁰ (critiques de films) ou Yelp (critiques de produits) ont été les plus utilisés. Quelques jeux de données similaires existent pour d'autres langues (comme le corpus Allocine pour le français, un corpus de critiques de films comparable à IMDb), mais à une échelle beaucoup plus réduite. Les critiques de théâtre semblent relativement proches de la base de données IMDb, nous supposons donc que l'algorithme XLNet peut être appliqué avec succès aux critiques de théâtre.

Comme nous l'avons déjà mentionné dans l'état de l'art, XLNet a été entraîné sur la base de données IMDb, un corpus constitué de critiques de cinéma. La proximité entre les critiques de cinéma et les critiques de théâtre a motivé le choix de cet algorithme. Si ces critiques n'ont pas le même objet d'étude, toutes deux utilisent néanmoins des arguments similaires, comme la description des performances des acteurs, l'introduction de l'intrigue ou l'analyse des décors. Au lieu d'annoter manuellement les critiques de notre corpus nous-mêmes, ce qui aurait représenté une tâche chronophage, nous avons décidé de ré-utiliser ce modèle pré-entraîné.

Réutilisation du « text zoning » pour filtrer le contenu des critiques

Afin d'obtenir des résultats plus précis, nous faisons le choix de supprimer certaines parties des textes des critiques avant d'implémenter XLNet. Nous nous rendons compte que le genre d'une pièce peut affecter les performances de l'algorithme. La critique d'une tragédie comporte nécessairement plus de vocabulaire négatif que la critique d'une comédie, ce qui induit en erreur les analyses algorithmiques. Pour contourner ce problème, nous investissons notre travail sur le zoning²¹¹ à nouveau afin de retirer les parties des textes qui peuvent fausser les traitements informatiques.

Comme nous l'avons déjà expliqué dans la première sous-partie de ce chapitre, une critique peut être divisée en huit arguments ou huit « zones », auxquelles nous avons attribué huit couleurs :

Numéro	Catégorie	Couleur
1	Introduction	Violet

²¹⁰Andrew MAAS, et al., « Learning Word Vectors for Sentiment Analysis », *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 2011, p. 142–150.

²¹¹ Voir pages 129 – 137.

2	Description de l'intrigue	Rouge
3	Performance des acteurs	Jaune
4	Détails visuels, sonores et olfactifs	Vert
5	Structure de la pièce	Marron
6	Public	Gris
7	Analyses	Bleu
8	Conclusion	Jaune

Fig. 50 : Tableau des différentes catégories et de leur couleur associée.

Nous considérons que les catégories « Introduction » et « Description de l'intrigue » présentent des éléments factuels qui ne sont pas liés à l'opinion du critique. L'introduction permet de situer la pièce dans le paysage théâtral contemporain. L'auteur y donne généralement des informations relatives au lieu de la représentation, au dramaturge, à ses précédents succès et aux autres pièces qu'il a écrites. La description de l'intrigue se concentre presque exclusivement sur une présentation du cadre spatio-temporel de l'histoire, des personnages et des thématiques principales de la pièce. Bien que ces éléments soient nécessaires à la rédaction d'une critique, ils ne comportent pas, ou très peu, de traces de la subjectivité de l'auteur. Surtout, la description de l'intrigue risquait de fausser les résultats.

Le principe de cette étape de filtrage consiste ainsi à supprimer ces phrases descriptives pour ne garder que les phrases liées à l'expression de la subjectivité de l'auteur. Pour reprendre le même exemple que nous avons donné lors de la description de notre expérience sur le zoning, voilà à quoi ressemble la critique après que les catégories « Introduction » et « Description de l'intrigue » soient retirées :

Critique originale

Sam Shepard plays gnaw away at you. They tease you with cryptic clues, disintegrating storylines and restless, febrile characters. His 1985 play *A Lie of the Mind* features the same symbolism heavy blend of redneck grit and warped American dreams as *Fool For Love* and *Buried Child*. This time, its themes - dysfunctional relationships, inescapable destinies, mortal love - land in frost-bitten rural Montana. Two families are joined by an abusive marriage. Jake has beaten Beth to a pulp and retreated to his childhood bedroom, plagued by guilt and grief. Brain-damaged and bewildered, Beth has been swallowed up by her clan too, cooped up with her bitter parents and her gun-toting brother in their snowy ranch. Shepard follows these converging narratives, tracing every character's inner geography in forceful, elliptical brushstrokes. James Hillier's dusky, smoke-filled staging bears striking resemblance to John

Tiffany's recent production of *The Glass Menagerie*, with the action isolated on small, square platforms against a vast blackness, and a neon moon hovering balefully behind the stage. But despite a set of detailed performances - particularly from John Stahl as Beth's flinty father and Laura Rogers' as Jake's mousy sister Sally - and a contemplative live score from James Marples, it never evokes the requisite haunted atmosphere nor mines the murky depths of Shepard's dialogue. It's just too crowded, too cluttered, too clunky, and the play loses much of its unsettling power as a result.

Critique annotée

Sam Shepard plays gnaw away at you. They tease you with cryptic clues, disintegrating storylines and restless, febrile characters. His 1985 play *A Lie of the Mind* features the same symbolism heavy blend of redneck grit and warped American dreams as *Fool For Love* and *Buried Child*. This time, its themes - dysfunctional relationships, inescapable destinies, mortal love - land in frost-bitten rural Montana. Two families are joined by an abusive marriage. Jake has beaten Beth to a pulp and retreated to his childhood bedroom, plagued by guilt and grief. Brain-damaged and bewildered, Beth has been swallowed up by her clan too, cooped up with her bitter parents and her gun-toting brother in their snowy ranch. Shepard follows these converging narratives, tracing every character's inner geography in forceful, elliptical brushstrokes. James Hillier's dusky, smoke-filled staging bears striking resemblance to John Tiffany's recent production of *The Glass Menagerie*, with the action isolated on small, square platforms against a vast blackness, and a neon moon hovering balefully behind the stage. But despite a set of detailed performances - particularly from John Stahl as Beth's flinty father and Laura Rogers' as Jake's mousy sister Sally - and a contemplative live score from James Marples, it never evokes the requisite haunted atmosphere nor mines the murky depths of Shepard's dialogue. It's just too crowded, too cluttered, too clunky, and the play loses much of its unsettling power as a result.

Critique filtrée

~~Sam Shepard plays gnaw away at you. They tease you with cryptic clues, disintegrating storylines and restless, febrile characters. His 1985 play *A Lie of the Mind* features the same symbolism heavy blend of redneck grit and warped American dreams as *Fool For Love* and *Buried Child*. This time, its themes - dysfunctional relationships, inescapable destinies, mortal love - land in frost-bitten rural Montana. Two families are joined by an abusive marriage. Jake has beaten Beth to a pulp and retreated to his childhood bedroom, plagued by guilt and grief. Brain-damaged and bewildered, Beth has been swallowed up by her clan too, cooped up with her bitter parents and her gun-toting brother in their snowy ranch.~~ Shepard follows these converging narratives, tracing every character's inner geography in forceful, elliptical brushstrokes. James Hillier's dusky, smoke-filled staging bears striking resemblance to John Tiffany's recent production of *The Glass Menagerie*, with the action isolated on small, square platforms against a vast blackness, and a neon moon hovering balefully behind the stage. But despite a set of detailed performances - particularly from John Stahl as Beth's flinty father and Laura Rogers' as Jake's mousy sister Sally - and a contemplative live score from James Marples, it never evokes the requisite haunted atmosphere nor mines the murky depths of Shepard's dialogue. It's just too crowded, too cluttered, too clunky, and the play loses much of its unsettling power as a result.

Pour réaliser cette tâche de filtrage, nous nous appuyons sur les précédents travaux effectués sur le zoning. Nous améliorons cependant notre modèle d'arbres décisionnels de sorte que chacune des phrases soit prise en compte. L'entraînement du classificateur se fait donc phrase par phrase. Pour représenter chaque phrase au sein de chaque texte, nous sélectionnons les meilleurs prédicteurs en termes d'importance. Pour rappel, l'algorithme que nous utilisons correspond aux forêts de classification. Les arbres de cette forêt sont construits en utilisant l'entropie. L'entropie est un paramètre des arbres de décision qui permet de sélectionner les meilleurs prédicteurs lors de la construction de notre arbre.

Nous n'avons utilisé que cinq prédicteurs pour déterminer la classe d'une phrase (Nombre de caractères dans la phrase, position de la phrase dans la critique, phrases consécutives de même type, nombre de mots dans la phrase, somme du score positif et négatif attribués à la critique). Ce sont ces cinq paramètres qui permettent de minimiser l'entropie globale de notre algorithme. Connaître les raisons de cette sélection demanderait un travail d'approfondissement qui demanderait une expertise technique élevée.

Afin d'augmenter la précision du modèle, nous décidons de réduire le problème de classification à huit zones à un problème binaire : 0 si la phrase correspond à la zone « introduction » ou « intrigue », 1 si la phrase correspond à une autre zone. Avec ces paramètres, nous obtenons une précision de 86 % sur les phrases tests. Les sections « Introduction » et « Intrigue » étant finalement supprimées, nous pouvons désormais appliquer le modèle à l'ensemble de notre corpus filtré.

Application du modèle XLNet

La première étape consiste à fournir à XLNet un ensemble de données appropriées afin qu'il apprenne à déterminer automatiquement si un texte peut être considéré comme positif ou négatif. Nous utilisons 24 000 critiques de films issues de la base de données IMDb en entrée de notre modèle. Ces critiques ont été préalablement labellisées par des milliers d'annotateurs qui ont attribué un score de positivité ou de négativité à ces textes. Après trois *epoch*, c'est-à-dire après que l'ensemble des données ait été présentées trois fois au réseau de neurones, le score de précision obtenu était de 94%. Ces étapes terminées, nous appliquons le modèle à notre corpus.

III.C.2 Des choix esthétiques similaires

III.C.2.1 Résultats

Représentation des pièces vues le plus grand nombre de fois à Londres (2010 – 2020)

CORPUS I - Critique journalistique

Numéro	Pièce (toute représentation confondue)	Dramaturge	Nombre de fois mentionnée
1	<i>A Midsummer Night's Dream</i>	William Shakespeare	137
2	<i>Richard III</i>	William Shakespeare	104
3	<i>Hamlet</i>	William Shakespeare	98
4	<i>Uncle Vanya</i>	Anton Tchekhov	92
5	<i>MacBeth</i>	William Shakespeare	90
6	<i>The Tempest</i>	William Shakespeare	76
7	<i>Three Sisters</i>	Anton Tchekhov	59
8	<i>As You Like It</i>	William Shakespeare	59
9	<i>Much Ado About Nothing</i>	William Shakespeare	58
10	<i>The Importance of Being Earnest</i>	Oscar Wilde	54
11	<i>Twelfth Night</i>	William Shakespeare	52
12	<i>King Lear</i>	William Shakespeare	52
13	<i>Julius Caesar</i>	William Shakespeare	52
14	<i>The Comedy of Errors</i>	William Shakespeare	52
15	<i>The Duchess of Malfi</i>	John Webster	50

CORPUS II - Critique numérique

Numéro	Pièce (toute représentation confondue)	Dramaturge	Nombre de fois mentionnée
1	<i>MacBeth</i>	William Shakespeare	110
2	<i>A Midsummer Night's Dream</i>	William Shakespeare	83
3	<i>Hamlet</i>	William Shakespeare	76
4	<i>A Christmas Carol</i>	Charles Dickens	67
5	<i>Twelfth Night</i>	William Shakespeare	65
6	<i>Romeo and Juliet</i>	William Shakespeare	61
7	<i>Othello</i>	William Shakespeare	54
8	<i>Richard III</i>	William Shakespeare	51
9	<i>The Tempest</i>	William Shakespeare	48
10	<i>King Lear</i>	William Shakespeare	46
11	<i>Much Ado About Nothing</i>	William Shakespeare	45
12	<i>Henri V</i>	William Shakespeare	33
13	<i>As You Like It</i>	William Shakespeare	33
14	<i>Julius Caesar</i>	William Shakespeare	32
15	<i>The Taming of the Shrew</i>	William Shakespeare	30

Fig. 51 : Classement du nombre de pièces les plus vues par les journalistes et les blogueurs.

Classement des pièces ayant reçu les meilleurs scores de positivité

CORPUS I - Critique journalistique

Numéro	Pièce	Dramaturge	Théâtre	Capacité d'accueil
1	<i>Springstorm</i>	William Tennessee	National Theatre	1 676
2	<i>Now This Is Not The End</i>	Rose Lewenstein	Arcola Theatre	197
3	<i>Martine</i>	Jean-Jacques Bernard	National Theatre	1 676
4	<i>Trestle</i>	Stewart Pringle	Southwark Playhouse	200
5	<i>Mass Observation</i>	Lu Kemp	Almeida Theatre	325
6	<i>The Kindness of Strangers</i>	Russell Woodhead et Jack Lowe	Southwark Playhouse	200
7	<i>I Wanna Be Yours</i>	Zia Ahmed	Bush Theatre	180
8	<i>So Many Reasons</i>	Racheal Ofori	Camden People's Theatre	60
9	<i>Girls</i>	Theresa Ikoko	Soho Theatre	140
10	<i>Betrayal</i>	Harold Pinter	Harold Pinter Theatre	796

CORPUS II - Critique numérique

Numéro	Pièce	Dramaturge	Théâtre	Capacité d'accueil
1	<i>Rosenbaum's Rescue</i>	A. Bodin Saphir	Park Theatre	200
2	<i>Srill</i>	Paul Hodson	Brixton House	200
3	<i>A Dark Night in Dalston</i>	Tim Spark	Park Theatre	200
4	<i>Wet Bread</i>	Tom Glover	King's Head Theatre	110
5	<i>Eventide</i>	Barney Norris	Arcola Theatre	197
6	<i>Minotaur</i>	Michael Fentiman	Polka Theatre	300
7	<i>Steel Magnolias</i>	Robert Harling	Hope Theatre	50
8	<i>Southern Belles</i>	Tennessee Williams	King's Head Theatre	110
9	<i>The High Table</i>	Lynette Linton	Bush Theatre	180
10	<i>The Canary and The Crow</i>	Daniel Ward	Arcola Theatre	197

Fig. 52 : Représentation des pièces ayant reçu le meilleur score de positivité.

III.C.2.2 Interprétation

La notion de succès peut être analysée de deux manières. Soit nous pouvons considérer les pièces qui ont été vues le plus grand nombre de fois, soit nous pouvons prendre en compte leur score de positivité après l'analyse algorithmique. Lorsque nous examinons le tableau présenté en figure 48, soit le top 10 des pièces qui ont été vues le plus grand nombre de fois, nous pouvons observer que dans les deux cas, ce sont les classiques de la littérature anglaise qui apparaissent. William Shakespeare, Anton Tchekhov, Oscar Wilde et Charles Dickens sont les dramaturges dont les représentations ont attiré le plus l'attention des journalistes et des blogueurs.

Lorsque nous passons ensuite aux analyses des pièces qui ont reçu les commentaires les plus positifs, nous pouvons voir dans un premier temps des différences claires entre les deux corpus. Chacune des deux communautés affiche un palmarès distinct. Nous pouvons également noter que pour les journalistes, une des pièces qui a obtenu le plus de succès a reçu un prix en 2017, le Papatango New Writing Prize, pour la pièce *Trestle* de Stewart Pringle, alors qu'aucune des pièces acclamées par la critique numérique n'a bénéficié d'un tel prestige.

Concernant les lieux dans lesquels ont été représentées ces pièces, les journalistes ont tendance à préférer des pièces jouées dans de grands théâtres, au contraire des blogueurs. *Springstorm*, *Martine*, *Mass Observation* et *Betrayal* ont été jouées au National Theatre, au Almeida Theatre et au Harold Pinter Theatre, qui sont des lieux qui offrent entre 300 et 1600 places. Du côté de la critique numérique, il n'y a que l'adaptation de *Minotaur* de Michael Fentiman qui a été jouée au Polka theatre qui contient 300 places. Toutes les autres pièces ont été jouées dans des plus petits théâtres de moins de 300 places, ce qui laisse penser que c'étaient des pièces plus discrètes, qui ne bénéficiaient pas de la même aura que celles jouées dans les grands théâtres.

Malgré ces différences, nous pouvons surtout remarquer des similitudes entre les préférences esthétiques de ces deux communautés. Dans les deux cas, lorsque l'on prête attention aux textes de ces critiques, nous pouvons voir que ce sont les mêmes éléments théâtraux qu'ils apprécient, et surtout les mêmes thématiques qui reviennent, même si ce n'est pas dans les mêmes pièces et pas sous la même forme. Les thèmes liés aux minorités, qu'elles soient sexuelles, ethniques ou religieuses, à la politique ou à la vie intérieure des personnages sont très présents dans les 10 pièces favorisées par ces deux communautés.

Contrairement à ce que nous aurions pu penser au départ, le top 10 des pièces ayant obtenu les commentaires les plus positifs ne démontre pas de scission entre le canon de la critique théâtrale londonienne et ses périphéries. Au contraire, il y aurait plutôt des ponts entre ces deux communautés.

III.C.2.3 Limites

Les limites de notre approche sont autant techniques que littéraires. D'un point de vue informatique, il nous est difficile d'évaluer la pertinence des résultats obtenus avec le modèle XLNet. Dans la mesure où le réseau de neurones a été entraîné sur la base de données IMDb, c'est-à-dire sur des données cinématographiques, il est possible que les scores obtenus manquent d'exactitude. Bien que les critiques de notre corpus et celles d'IMDb présentent des caractéristiques communes indéniables (analyse des décors, du jeu des acteurs, de l'intrigue, etc.), nous sommes néanmoins contraints d'émettre des réserves quant aux performances de notre modèle. Une des solutions consisterait à annoter nous-mêmes une partie de notre corpus afin d'entraîner un algorithme à reconnaître les scores de positivité et de négativité des critiques théâtrales.

D'un point de vue littéraire, nous pourrions approfondir ces résultats en effectuant d'autres expériences. Nous pourrions par exemple calculer les scores de positivité et de négativité pour les pièces communes aux deux corpus. Cela nous permettrait d'observer si pour une même pièce donnée, les analyses divergent d'un corpus à l'autre. Nous pourrions également calculer les scores de négativité les plus élevés afin de déterminer quelles pièces les journalistes et les blogueurs ont appréciées le moins. Nous pourrions également réinvestir le zoning pour comprendre quels éléments théâtraux ont le plus séduit les critiques. Nous pourrions étendre notre sélection aux 50 pièces acclamées par les journalistes et les blogueurs, et essayer de définir si des éléments récurrents peuvent être observés. Cette approche nous permettrait de répondre à des questions comme : « Ont-ils préféré les décors ou la performance des acteurs ? » ou « Peut-on repérer des individus communs (acteurs, metteurs en scène, éclairagistes, etc.) à toutes ces représentations ? » Nous pourrions reproduire la même démarche, mais pour les pièces considérées comme des échecs.

Remarques conclusives sur le chapitre III

Ce chapitre avait pour dessein de présenter les expériences techniques que nous avons réalisées afin d'explorer notre corpus et de pouvoir esquisser des réponses à notre problématique. Afin d'examiner les points de convergence et de divergence entre ces deux groupes de spectateurs, nous avons mené trois études qui abordaient trois questions distinctes. La première était liée à leur manière d'écrire, la deuxième à leur répartition géographique dans Londres, et la troisième à leurs préférences esthétiques. Quand cela était possible, notre point de départ était fondé sur des exemples précis donnés par les critiques. La finalité de notre travail consistait ensuite à évaluer si ces exemples étaient représentatifs d'une dynamique générale ou s'ils constituaient des cas isolés.

Bien que dans les trois cas nous avons identifié des limites littéraires et techniques à nos approches, les résultats obtenus nous permettent néanmoins de conclure qu'il existe de réelles différences entre les journalistes et les blogueurs. Les horizons d'attente de ces deux publics sont reflétés dans leurs analyses de la scène. Tandis que les blogueurs mettent à l'honneur la diversité et la marginalité, tant dans leur manière d'écrire en affirmant leur subjectivité que dans les lieux qu'ils fréquentent, les journalistes tendent à respecter la « norme. » En proposant un discours davantage tourné vers l'analyse rationnelle d'un spectacle et en se focalisant principalement sur les théâtres du centre de Londres, les journalistes incarnent, dans une certaine mesure, le « centre », ou le « canon » de la culture théâtrale londonienne.

Ces conclusions sont cependant à nuancer. D'une part, l'étude que nous avons menée sur les préférences esthétiques de ces deux communautés prouve le contraire et, d'autre part, ces analyses résultent de traitements informatiques. Les algorithmes nous permettent d'obtenir une vision d'ensemble de la critique théâtrale contemporaine. Cette lecture distancée efface ainsi toutes les particularités. Tandis que certains journalistes se rendent dans les associations culturelles pour assister à de petites productions, certains blogueurs fréquentent régulièrement les grands théâtres du centre de Londres. Il y a donc bien, dans une certaine mesure, un fossé entre le canon de la critique théâtrale et ses périphéries. Cependant, cela ne signifie pas que l'ensemble des auteurs de ces deux corpus peuvent être catégorisés de manière binaire.

Chapitre IV
Science ouverte :
Partage et mise en valeur des données

Les trois chapitres précédents ont permis de montrer les étapes successives de notre projet. Le premier chapitre démontrait en quoi les deux disciplines dans lesquelles s'inscrit notre étude ont été peu explorées jusqu'ici. Tant dans les humanités numériques théâtrales que dans les études de la réception, un nombre restreint de travaux ont été menés. Surtout, personne ne s'était encore intéressé à la réception du théâtre contemporain londonien à l'aide d'outils numériques. De par son originalité, notre recherche contribue ainsi à la construction de ces deux champs.

Dans le deuxième chapitre, nous nous sommes concentrés sur la constitution et la présentation de notre corpus. Pour répondre à notre problématique, c'est-à-dire pour analyser la manière dont était perçu le théâtre par le centre de la critique théâtrale et par ses périphéries, il nous a fallu réunir et organiser un grand nombre de données représentatives de ces deux communautés. Ce travail, qui s'est étalé sur plus d'une année, nous a permis de créer la première base de données à réunir autant de critiques journalistiques et numériques directement exploitables par l'ordinateur.

Dans le troisième chapitre, nous avons exploré trois questions qui correspondaient à trois domaines techniques et qui ont permis d'esquisser des réponses à notre problématique. En explorant la manière dont écrivent les critiques, les lieux dans lesquels ils se rendent ainsi que leurs préférences esthétiques, nous avons pu mieux comprendre leur perception singulière du théâtre. Nous avons conclu que bien qu'il existe de réelles différences entre ces deux communautés, des points de convergence subsistent néanmoins entre elles deux.

Ce quatrième chapitre aura pour finalité de présenter un aspect différent de notre travail, mais complémentaire de notre démarche. Ce projet de thèse a donné naissance à une plateforme en ligne intitulée DramaCritiques.com. Son objectif est double : d'une part, donner davantage de visibilité aux marges de la culture théâtrale londonienne et, d'autre part, participer à la vulgarisation des humanités numériques. Travailler à l'intersection de deux disciplines opposées présente des obstacles. En détaillant et en illustrant nos analyses techniques et en mettant en libre accès nos données ainsi que nos scripts informatiques, nous espérons contribuer à la science ouverte.

Présenter un site Internet pour clôturer un projet de thèse est peu commun. Cependant, dans la mesure où notre travail s'inscrit dans une perspective numérique et que cette plateforme contribue à partager nos recherches avec le plus grand nombre, nous avons décidé de l'intégrer pleinement à notre manuscrit. De plus, la création de DramaCritiques.com a requis beaucoup de temps et de réflexion. La mise en place d'un site Internet nécessite l'apprentissage et l'application de compétences techniques diverses. Ce projet a duré plus de

huit mois au total et présente toutes les expériences décrites dans cette thèse. Au-delà des connaissances nouvelles sur le théâtre contemporain londonien qu'il propose, cette plateforme peut également être utilisée comme outil pédagogique. DramaCritiques.com a été conçu pour permettre aux étudiants de découvrir les humanités numériques et le théâtre contemporain britannique. C'est pour toutes ces raisons que nous venons d'évoquer que nous avons fait le choix d'intégrer sa présentation à la rédaction de notre manuscrit.

Ce dernier chapitre sera divisé en deux parties, elles-mêmes composées de deux sous-parties. La première, intitulée « Créer un pont entre les humanités et le numérique », sera consacrée à la présentation de notre site Internet. Nous présenterons dans un premier temps l'identité numérique de Drama Critiques, ainsi que sa page d'accueil. La première page d'un site Internet a une fonction importante. Elle doit permettre aux visiteurs de comprendre rapidement les thématiques abordées et l'inviter à découvrir ses articles. Nous verrons quelles stratégies nous avons mises en place afin de créer une page d'accueil attractive et, surtout, qui met en évidence le mariage des sciences et des lettres. Dans un second temps, nous présenterons le contenu de notre plateforme, et plus particulièrement l'application « Derrière la machine ». Cette dernière a été précisément pensée pour illustrer comment l'informatique et la littérature peuvent s'allier.

Dans la seconde sous-partie de ce dernier chapitre, intitulée « Un potentiel d'exploitation pour les communautés littéraires et scientifiques », nous montrerons en quoi ce travail et cette plateforme ouvrent de nouvelles perspectives de recherche pour quiconque serait intéressé soit par le théâtre contemporain britannique, soit par l'informatique. Premièrement, nous nous focaliserons sur l'ouverture de nos données. Ce sujet épineux dans les humanités numériques est pourtant capital dans la mesure où il garantit la reproductibilité des expériences et le partage des données à l'ensemble de la communauté académique. Nous présenterons les « FAIR Principles », une série de critères qui ont pour finalité de faciliter la recherche, l'accessibilité, l'interopérabilité et la réutilisation des données. Nous verrons plus précisément en quoi notre projet répond à ces principes directeurs.

Enfin, nous terminerons en présentant toutes les pistes de réflexion que nous avons soit commencé à examiner, soit perçu au cours de nos analyses mais que nous n'avons pas encore explorées. Tant d'un point de vue littéraire que technique, ce projet de thèse offre de nombreuses possibilités d'analyses.

IV.A Créer un pont entre les humanités et le numérique

IV.A.1 La plateforme DramaCritiques.com

Le projet de construire une plateforme en ligne a émergé au bout de la deuxième année de travail. Après avoir établi notre corpus et présenté quelques-uns de nos travaux lors de conférences, nous nous sommes rendu compte que les humanités numériques laissaient de nombreuses questions en suspens. Le mariage de deux disciplines qui sont rarement associées l'une à l'autre interroge. Tant pour les informaticiens que pour les littéraires, cette combinaison peu conventionnelle ne fait pas toujours immédiatement sens.

La première finalité de cette plateforme est donc d'ouvrir nos travaux au plus grand nombre afin de clarifier les concepts qui sous-tendent notre recherche et ceux des humanités numériques de manière plus générale. Tandis que le fonctionnement d'un algorithme pour une personne avec un bagage littéraire est sûrement obscur, l'utilisation de la science des données pour aborder le théâtre représente également une curieuse démarche pour un informaticien. DramaCritiques.com a ainsi pour vocation de participer à la vulgarisation des humanités numériques en expliquant clairement et simplement comment allier informatique et littérature.

IV.A.1.1 Identité numérique de la plateforme

Transmettre des informations par le biais d'un site Internet demande une organisation rigoureuse des idées. L'utilisation des couleurs, des typographies, des tailles d'écriture et de la disposition des textes sur une page sont autant d'éléments que le chercheur doit prendre en compte avant de mettre en place un nouvel outil numérique. L'informatique impose ses propres contraintes qu'il convient d'adopter et d'adapter en fonction de la nature du contenu qui sera mis en ligne. C'est pourquoi passer d'une feuille de papier classique à une page Internet structurée et interactive requiert une véritable organisation des idées.

Nous choisissons d'appeler notre plateforme DramaCritiques car nous voulons une appellation qui puisse être comprise par un public francophone et anglophone à la fois. Pour notre logo, afin de rendre compte des aspects littéraires et techniques de notre recherche, nous ajoutons une racine carrée dessinée à l'encre de Chine pour encadrer le nom de notre site « DramaCritiques » :

√ Drama Critiques

Fig. 53 : Logo de la plateforme DramaCritiques.com.

De nombreux sites qui abordent la science des données utilisent la couleur bleue. Nous avons repris cette couleur en l'associant au corail afin de diminuer ses consonances trop techniques.

Concernant l'aspect esthétique de notre plateforme, nous demandons à Anaïs Rault, une artiste parisienne, de créer des dessins qui représentent au mieux le mariage de la science et des lettres. Nous organisons plusieurs sessions de travail afin de réfléchir aux formats et aux matériaux pour produire nos illustrations. Nous voulons mettre en avant l'aspect numérique de notre projet, c'est pourquoi nous optons pour des vignettes carrées qui rappellent le format des images publiées dans certains réseaux sociaux utilisés par les blogueurs.

Nous faisons également le choix d'utiliser l'encre de Chine afin de faire ressortir les contrastes de nos dessins. Enfin, nous décidons de choisir plusieurs symboles clairs et représentatifs du théâtre et de la science des données, comme un ordinateur, les masques de la Commedia Dell'Arte ou une scène de théâtre, afin de jeter la lumière sur les deux disciplines opposées de notre projet. Chaque dessin est d'abord réalisé sur une feuille de papier, avant d'être retravaillé sur Photoshop. Voici quelques productions :



Fig. 54 : Représentation de huit dessins de la plateforme DramaCritiques.com.

IV.A.1.2 Construction de la page d'accueil

La finalité de cette plateforme étant d'expliquer les enjeux des humanités numériques, nous essayons de structurer au mieux les différentes catégories de notre projet. La première étape consiste à créer une page d'accueil qui reflète la double dimension de notre travail. Une page d'accueil est un élément important dans la création d'un site Internet. Elle doit non seulement résumer les fondements d'un projet, mais aussi inviter les utilisateurs à vouloir se rendre sur les autres pages du site.

Pour la création de notre site Web, nous décidons d'utiliser WordPress. WordPress est un Système de Gestion de Contenu (SGC) libre, gratuit et open-source qui apparaît en 2003. WordPress fonctionne à partir de « blocs de textes » que le créateur peut insérer à sa convenance et moduler en fonction du contenu qu'il veut intégrer (images, formulaires, vidéos, etc.). Après huit mois de travail, voici la page d'accueil de DramaCritiques.com que nous avons mise en place à partir d'avril 2021 :



LITTÉRATURE ET ALGORITHMES

Drama Critiques explore 10 ans du **théâtre contemporain anglais** grâce à la **Data Science**. En se focalisant sur deux communautés littéraires : les journalistes d'un côté, et les blogueurs.euses de l'autre, nous examinons à travers leurs critiques théâtrales le discours poétique et politique qu'ils.elles construisent. Fondé sur un corpus de **plus de 40 000 critiques théâtrales** publiées entre 2010 et 2020 sur la scène londonienne, Drama Critiques explore ainsi les points de convergence et de divergence entre ces deux communautés grâce aux **algorithmes**.

LE PROJET

A partir des années 2010, une vague de blogs rédigés par des auteurs venant d'**horizons multiples** émerge sur Internet. **Etudiant.e.s**, **professionnel.le.s du théâtre** mais aussi **simples amateurs.trices** s'emparent de la toile pour publier leurs critiques. Ces nouvelles voix marginales de l'**espace numérique** décentralisent le canon de la critique journalistique et redessinent progressivement les contours de l'**histoire théâtrale anglaise**.

LA CRITIQUE LITTÉRAIRE



LA BASE DE DONNÉES

Avec plus de **40 000 critiques** théâtrales anglaise, le corpus est composé des **23 journaux** les plus influents de la presse anglaise (*The Guardian*, *The Times*, *The Independent*, etc.) et des **28 blogs** les plus populaires de la toile (*West End Wilma*, *Everything Theatre*, *There Ought To Be Clowns*, etc.). C'est grâce à cette base de données que nous pouvons examiner ces deux communautés avec des **outils numériques**.

DÉCOUVRIR LE CORPUS

 **40 000**
CRITIQUES LITTÉRAIRES

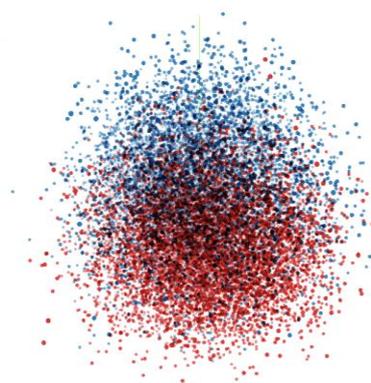
 **10**
ANNÉES DE THÉÂTRE (2010-2020)

 **600**
BLOGUEURS.EUSES

 **400**
JOURNALISTES

 **9 000**
PIÈCES DIFFÉRENTES

 **400**
LIEUX DE REPRÉSENTATION À LONDRES



LES ANALYSES LITTÉRAIRES / ALGORITHMIQUES



Machine Learning et NLP

Comment déterminer le style littéraire de ces critiques ? Peut-on examiner des différences ?



Prédiction des sentiments

Quelles sont les préférences esthétiques des journalistes et des blogueurs.euses ?



Cartographie Numérique

Comment se répartissent géographiquement ces deux communautés ?



PARCOURIR LES 40 000 CRITIQUES

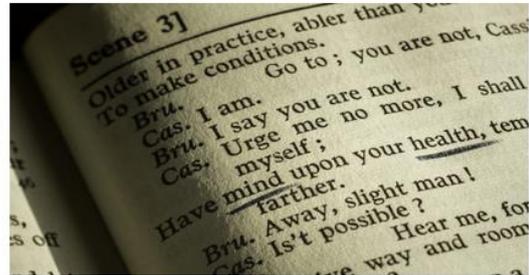
Un **moteur de recherche** a été créé afin d'explorer le corpus. Découvrez des pièces que vous ne connaissez pas encore, cherchez des œuvres **par thématique** ou bien lisez simplement toutes les critiques rédigées sur Hamlet !

[LE MOTEUR DE RECHERCHE](#)

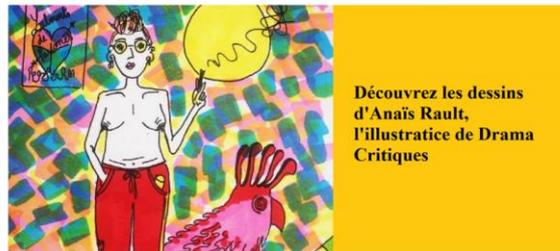
L'ORDINATEUR SAIT-IL LIRE ?

L'ordinateur sait-il lire ? Les algorithmes remplacent-ils désormais l'humain ? D'ailleurs, **qu'est-ce qu'un algorithme** ? Derrière la machine est une application créée pour répondre à ces questions.

[DERRIÈRE LA MACHINE](#)



EN SAVOIR PLUS



Découvrez les dessins d'Anaïs Rault, l'illustratrice de Drama Critiques

Fig. 55 : Capture d'écran de la page d'accueil de DramaCritiques.com.

Cette version est celle qui est en ligne aujourd'hui, et que nous avons diffusée auprès de différents publics. La page d'accueil répond aux critères principaux d'une page principale d'un site Internet. Trois premières sections accompagnées d'un bref paragraphe explicatif ainsi que d'un lien hypertexte permettent de présenter :

- L'ambivalence de notre projet qui repose sur les sciences et les lettres (« Littérature et algorithmes »)
- L'objet de notre étude qui porte sur la réception du théâtre contemporain londonien des années 2010 à 2020 (« Le Projet »)
- Le matériau sur lequel nous avons fondé nos analyses (« La Base de données »)

Nous ajoutons également un compteur interactif à la suite de ces trois sections, afin de donner un aperçu des chiffres-clés de notre projet aux utilisateurs (nombre de critiques, cadre temporel de notre étude, nombre de blogueurs et de journalistes, nombre de pièces différentes examinées par l'ensemble des critiques du corpus et nombre de lieux visités par les critiques). La sphère qui apparaît à droite de ce compteur est tirée de notre expérience sur le topic modeling. Nous ne l'avons pas présentée dans ce travail car elle manquait de conclusions pertinentes. Le visuel obtenu étant néanmoins esthétique et faisant écho à nos deux corpus (le bleu représente le corpus journalistique et le rouge le corpus numérique), nous l'avons intégré comme élément graphique.

Nous ajoutons aussi un bandeau qui présente les trois expériences principales que nous avons décrites au cours du précédent chapitre : la stylistique computationnelle, l'analyse d'opinion et la cartographie numérique. Au-dessous de chacune de ces mentions techniques, la question littéraire est posée, de sorte que l'association de la littérature et de la science des données soit claire. Les deux sections qui suivent mettent en valeur deux outils centraux à notre plateforme. Tandis que la section « Parcourir les 40 000 critiques » redirige les utilisateurs vers le moteur de recherche que nous avons implémenté, la section « L'Ordinateur sait-il lire ? » les invite à découvrir notre application « Derrière la machine » qui explique comment les algorithmes rendent un texte intelligible pour un ordinateur.

Enfin, nous terminons notre page d'accueil par l'insertion d'un carrousel d'images qui renvoient à d'autres sites liés à notre recherche (articles parus dans le magazine de l'Ecole normale supérieure, compte Instagram de notre illustratrice et lien vers la base de données *Theatre Record*). Cette page d'accueil réunit ainsi toutes les informations essentielles pour comprendre les enjeux de notre projet.

IV.A.2 Un Contenu clair, pédagogique et interactif

IV.A.2.1 Présentation du contenu du site web

Afin de présenter le plus clairement possible notre recherche, nous créons un menu déroulant dans le haut de notre site qui s'affiche sur toutes les pages et qui permet aux utilisateurs de découvrir toutes les étapes de notre projet. La capture d'écran ci-dessous met en évidence ce procédé :



Fig. 56 : Capture d'écran du menu déroulant « Catégories. »

Nous répertorions notre recherche en six catégories qui reprennent l'ordre des chapitres que nous avons rédigés. Ces six catégories (« Le Projet », « Le Corpus », « Stylistique computationnelle », « Analyse de sentiments », « Cartographie numérique » et « Derrière la machine ») correspondent à six pages distinctes qui suivent la progression de la construction de notre projet.

Catégorie 1 : Page « Le Projet »

La première catégorie, « Le Projet », résume les fondements théoriques littéraires de ces travaux. Nous rédigeons six sous-parties sur cette page afin d'explicitier notre problématique :

1. Que signifie « Critiquer ? »
2. Quel est le rôle du critique ?
3. Quand commence la critique théâtrale anglaise ?
4. Reformulation de la critique à l'ère numérique
5. Deux communautés différentes : la critique journalistique versus la critique numérique
6. Drama Critiques comme observatoire de ces changements culturels

Nous condensons toutes les informations importantes sur le contexte littéraire de la critique théâtrale au XXI^e siècle et intégrons une frise chronologique afin que les utilisateurs puissent saisir les enjeux de notre travail.

Catégorie 2 : « Le Corpus »

Cette deuxième catégorie a pour finalité d'introduire les deux communautés critiques que nous avons étudiées. Les tableaux de bord qui figurent aux pages 99, 102 et 118 de ce manuscrit ont été intégrés à cette catégorie. Son interactivité permet aux utilisateurs de découvrir le corpus dans sa globalité, ou les corpus I et II séparément à travers ses principales caractéristiques chiffrées (nombres de critiques, nombres de pièces commentées, répartition des journaux / blogs au cours du temps, etc.).

Le logo du site *Theatre Record* est également mis en évidence et permet aux utilisateurs, en cliquant dessus, d'être automatiquement redirigés vers leur site. Enfin, nous prenons soin de présenter les 28 blogs de notre second corpus. Pour chacun d'entre eux, nous ajoutons une image de leur page d'accueil ainsi qu'un lien vers leurs réseaux sociaux lorsqu'ils en disposent d'un (Twitter, Instagram et Facebook). En cliquant sur leur image de présentation, les utilisateurs sont directement redirigés vers leur blog, ce qui leur permet de découvrir l'univers propre à chacun des blogueurs. La capture d'écran ci-dessous permet de visualiser une partie de ce travail :

Les 28 blogs du corpus II

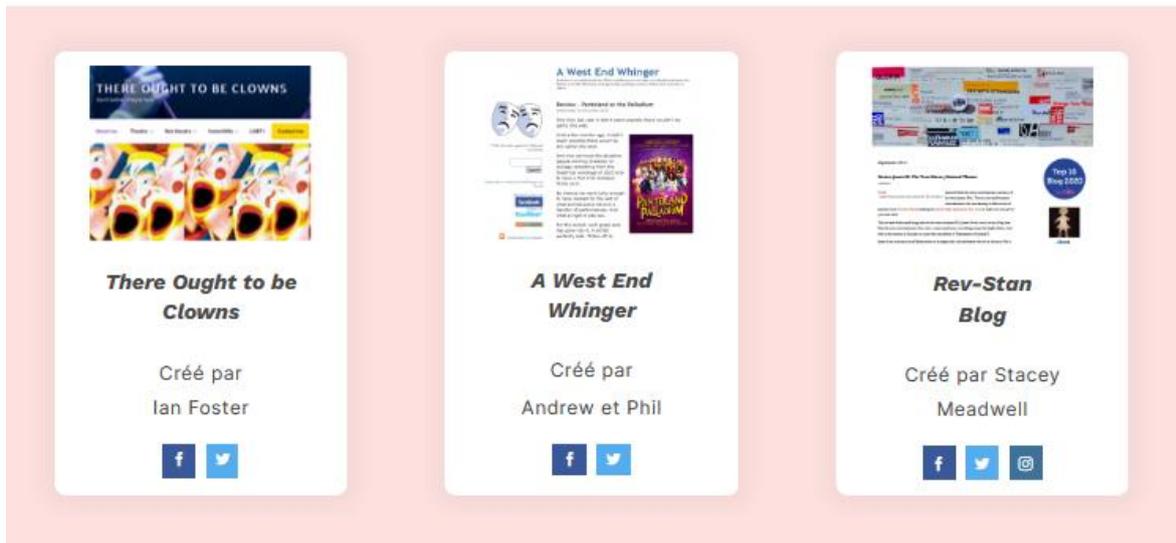


Fig. 57 : Capture d'écran de la présentation des blogs dans DramaCritiques.com

Enfin, nous intégrons la représentation schématique des caractéristiques principales de nos deux corpus. Cela permet à nos visiteurs d'obtenir un résumé clair des différences entre le canon de la critique théâtrale londonienne et ses périphéries :

Et pour résumer :

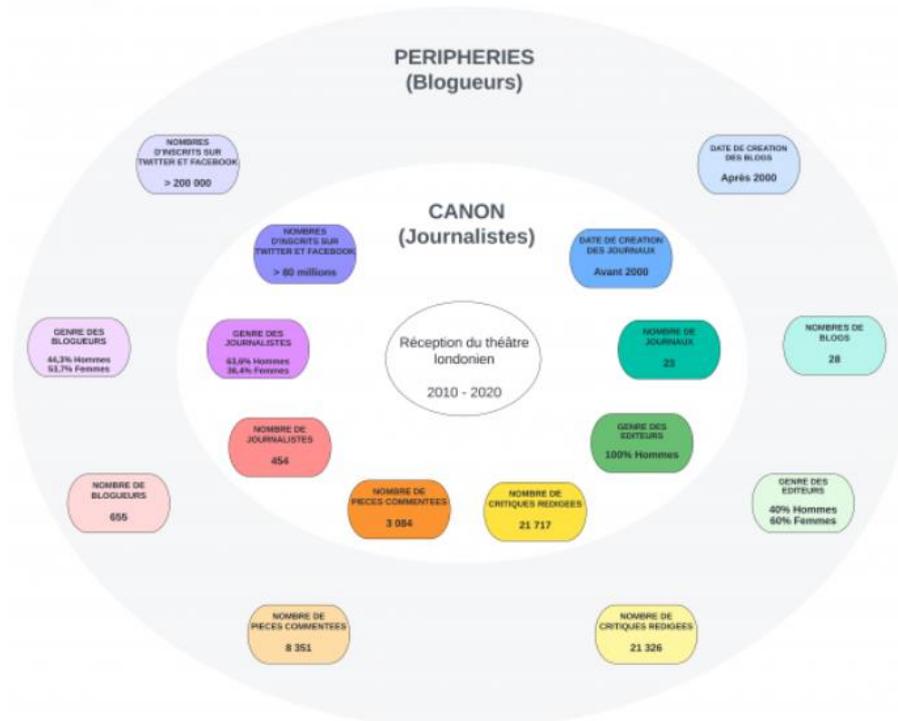


Fig. 58 : Représentation schématique des deux corpus sur la page « Le Corpus ».

Catégories 3, 4 et 5 : Présentation des expériences

La présentation de la problématique et du corpus terminée, nous pouvons aborder les expériences que nous avons menées. Les trois pages suivantes, intitulées « Stylistique computationnelle », « Analyse de sentiments » et « Cartographie numérique » ont toutes une structure similaire qui a pour finalité d'aider le lecteur à saisir rapidement la double dimension de ces travaux. Après avoir posé la question littéraire, nous rédigeons un premier paragraphe pour expliciter la pertinence de notre question. Nous abordons ensuite le choix des outils numériques, avant de présenter les résultats et d'en proposer une analyse conclusive.

Ces billets publiés sur notre plateforme reprennent les arguments que nous avons développés dans ce manuscrit, même s'ils sont plus courts et s'ils reposent sur un vocabulaire moins technique. Cette vulgarisation de notre recherche a pour dessein d'ouvrir nos travaux au plus grand nombre. Pour illustrer notre démarche, nous prenons ici l'exemple de la page « Stylistique computationnelle ». Nous avons mis en valeur la structure de notre article par un encadré bleu :

√Drama Critiques Accueil Catégories ▾ Focus sur les femmes ▾ Base de données ▾ A Propos 

Stylistique computationnelle

Traitement Automatique des Langues et style littéraire

Comment écrivent ces auteurs ?

Intérêt littéraire de cette question

Nous avons commencé par nous intéresser au style littéraire des auteurs de ces deux corpus. Avant d'analyser leurs idées, nous voulions comprendre comment ils structuraient leurs écrits. Bien que la notion de « style » soit difficile à définir, elle est pourtant autant au coeur de l'activité de critique que des débats actuels entre la critique papier et la critique numérique.

Tandis que les auteurs des journaux reprochent aux blogueurs.euses leur manque de rigueur, ces derniers, à l'inverse, revendiquent leur droit d'écrire différemment. Dans sa page de présentation, la plateforme Exeunt résume très bien les enjeux stylistiques que sous-tendent ces nouvelles formes de critiques numériques :

Exeunt believes in making beautifully written, experimental, fierce and longform writing about theatre available for free.

Pour Michael Billington en revanche, critique de longue date dans *The Guardian*, un blog ressemble davantage à une « lettre informelle » qu'à une véritable critique. Danielle Tarento, directrice du théâtre Chocolate Menier Factory à Londres, va même jusqu'à affirmer que ces blogueurs.euses ne sont pas de « réels écrivains » :

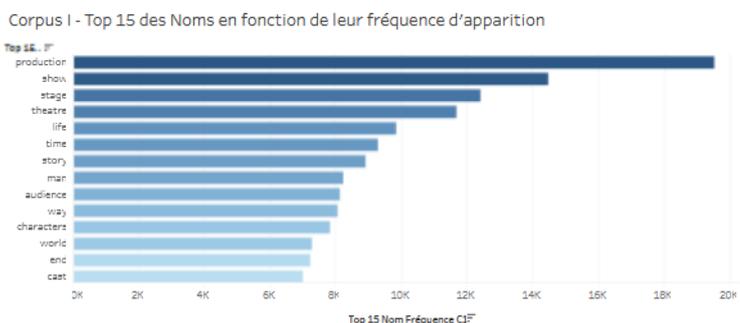
They do not have the intellectual background or historical background or time to know what they are writing about.

Quels moyens techniques utiliser pour y répondre ?

Nous nous sommes appuyés sur des travaux en linguistique computationnelle pour répondre à cette question. Cette discipline se situe à l'intersection de l'informatique, de la linguistique et de la statistique. Elle permet, entre autres, de modéliser des phénomènes de la langue naturelle grâce à des approches logiques. Nous avons pris pour point de départ les travaux de D.I Holmes qui définit le style comme un ensemble de variables mesurables qui participent à constituer « l'empreinte digitale » (*fingerprint*) d'un.e auteur.e.

Cette première piste de recherche consistait à examiner une série de caractéristiques stylistiques simples dans les deux corpus et de les comparer (nombre de mots / phrases par critique, les noms communs / verbes / adjectifs / les plus récurrents, la répartition des différents types de phrases, l'utilisation de la ponctuation, etc.).

Les résultats



Quelles conclusions ?

1. Les 5 premiers noms les plus employés dans ces deux corpus semblent indiquer que ces deux communautés portent leur attention sur deux objets différents. Lorsque l'on regarde les deux premiers termes les plus récurrents, ils sont sensiblement similaires (« Production » et « Show » pour le corpus I – « Theatre » et « Show » pour le corpus II). Le troisième terme le plus employé par la critique journalistique révèle davantage d'informations quant à leur centre d'intérêt. Le mot « Stage » laisse supposer que ces critiques se focalisent davantage sur la scène., ou du moins, que leur regard se porte sur le spectacle, sur ce qui se passe devant eux. Lorsque l'on compare avec les mots les plus fréquents dans le corpus II, en troisième position vient un terme assez similaire aux deux précédents (« Theatre » – « Show » – « Production »). Le quatrième terme est en revanche plus intéressant. « Audience » sous-entend que le regard des critiques numériques porte en priorité sur ce qui se passe à côté deux, sur les spectateurs, et non ce qui se passe en face d'eux, c'est-à-dire sur la scène. Pourrait-on y voir ici deux manières d'appréhender l'expérience du théâtre ? L'une qui serait rationnelle, tournée vers l'analyse du spectacle, et l'autre qui serait plus émotionnelle, davantage tournée vers les réactions de l'humain, ou du public ?

2. La répartition de l'utilisation des pronoms personnels entre ces deux corpus semble confirmer la validité de cette piste hypothèse. C'est le deuxième graphe qui représente les disparités les plus importantes en terme de pourcentages entre ces deux bases de données. Alors que la première personne du singulier représente 10 % de l'utilisation totale des pronoms personnels dans le corpus I, sa présence est doublée dans le corpus II (20 %). Cela signifie que les blogueurs.ueuses emploient deux fois plus le pronom « I » ou « je » dans leurs critiques. Serait-ce la marque d'une subjectivité plus affirmée de la part de la critique numérique ?

Fig. 59 : Capture d'écran de la page « Stylistique computationnelle ».

En exposant de manière concise les processus techniques et les éléments littéraires qui sous-tendent notre recherche, nous espérons que ces articles sur notre plateforme permettront à ceux qui les liront de comprendre en quoi l'alliance de ces deux champs peut donner naissance à une autre manière de concevoir l'objet littéraire.

IV.A.2.2 L'application « Derrière la machine »

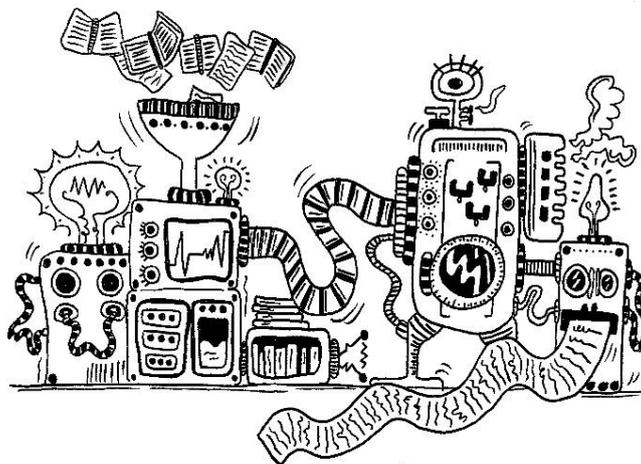


Fig. 60 : Illustration de l'application « Derrière la machine ».

La dernière page de notre menu « Catégories » est intitulée « Derrière la machine. » Cet onglet est celui qui met le plus explicitement en avant le mariage des sciences et des lettres. Les conférences auxquelles nous avons pu montrer notre travail nous ont permis de prendre conscience du fossé creusé entre deux communautés académiques qui s'ignorent. L'application « Derrière la machine » est ainsi conçue pour lever les interrogations sur l'utilisation de l'informatique pour l'étude des lettres. Analyser des critiques littéraires avec un ordinateur ne va pas de soi. Des questions comme « L'ordinateur remplace-t-il l'humain ? », « Mettre la littérature sous algorithmes ne revient-il pas à considérer les lettres comme une série de 0 et 1 ? » ou « Comment une machine peut-elle lire un texte ? » sont pertinentes et légitimes.

Pour tenter d'y répondre, « Derrière la machine » explique et illustre toutes les étapes techniques nécessaires pour rendre un texte intelligible pour l'ordinateur. La finalité de cette application est de montrer que la machine n'est pas une entité autonome. Un ordinateur ne peut pas lire un texte tout seul. Il ne peut pas le comprendre seul non plus. C'est l'humain, et en l'occurrence le chercheur qui, en appliquant une série de programmes informatiques, rend un texte analysable pour l'ordinateur.

Le principe de cette application est simple : l'utilisateur commence par choisir une critique journalistique de son choix qui s'affiche ensuite sur le tableau de bord situé à gauche de l'écran :

Etape 1 : Sélection de la critique

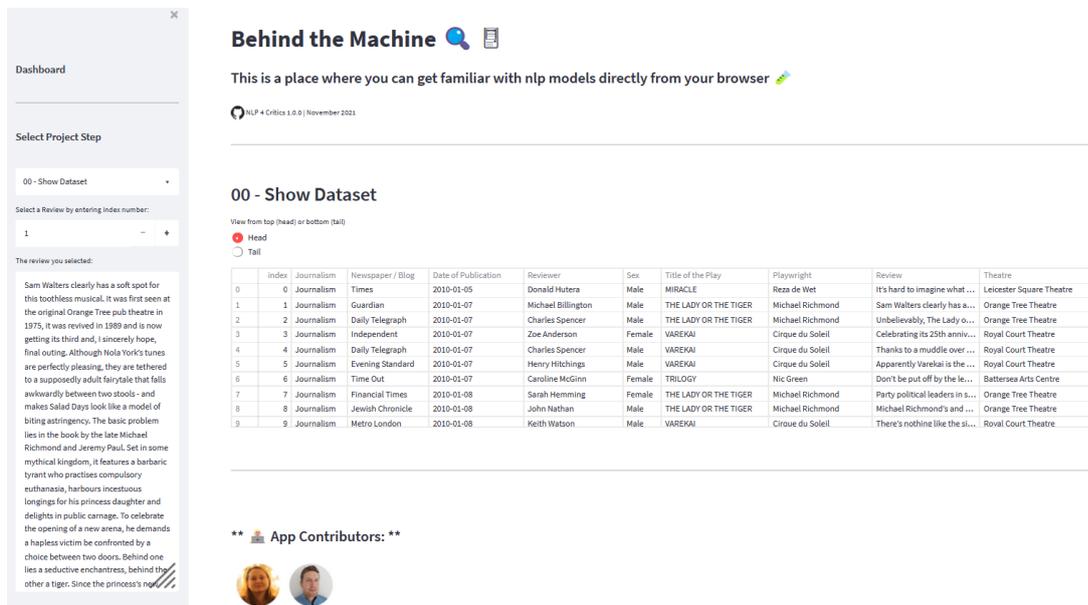


Fig. 61 : Capture d'écran de la première étape de l'application « Derrière la machine ».

Etape 2 : Tokenisation de la critique

En cliquant ensuite sur le premier onglet dans le tableau de bord, situé sous la mention « Select Project Step », l'utilisateur peut sélectionner la première étape nécessaire pour rendre un texte intelligible pour la machine. Comme nous l'avons expliqué dans le deuxième chapitre de ce manuscrit, la tokenisation consiste à diviser un texte en « tokens », c'est-à-dire en une « séquence de caractères qui sont regroupés en une unité sémantique utile pour un traitement²¹². » L'application que nous avons créée donne non seulement le rendu visuel de cette étape technique, ce qui permet aux utilisateurs de comprendre en quoi consiste concrètement la tokenisation, mais aussi les codes informatiques pour effectuer une telle démarche. Ces codes sont réutilisables sur Python et peuvent être appliqués sur n'importe quel texte.

²¹²<https://nlp.stanford.edu/IR-book/html/htmledition/tokenization-1.html> [En ligne], consulté le 19 mai 2022.

It's | hard | to | imagine | what | impact | a | change | of | venue | will | have, | on | this | uneven | but | tantalising | show | by | the | young | British | company | Ruby | in | the | Dust. | Written | in | 1992 | by | the | much-lauded | South African | playwright | Reza | de | Wet, | Miracle | is | a | Chekhovian | allegorical | drama | about | a | clapped-out | little | theatrical | troupe | circa | the | mid-1930s | that | pitches | up | in | a | small | town | on | a | cold | winter's | night. | The | five | actors | are | tired, | hungry | and | all | too | ready | to | take | their | frustrations | out | on | each | other. | Hospitality' | arrives | in | the | form | of | a | middle-aged, | black-clad | matron | (Lynne | Miller), | who | happens | to | be | the | wife | of | the | company's | errant | and | virulently | suspicious | male | ingenu. | As | the | forlorn | band | of | thespians | eventually | realise, | this | single | spumed | woman | poses | a | greater | threat | than | they | had | any | right | to | expect. | It's | an | intriguing | premise | that | borders | on | melodrama | before | tipping | over | into | an | almost | metaphysical | dimension. | Although | it's | set | during | the | Depression, | de | Wet's | script | has | a | teasing, | timeless | quality | that | lingers. | It | has | certainly | stayed | with | me. | Linnie | Reedman's | production | had | a | five-day | run | in | the | crypt | of | St | Andrew | Church, | Holborn, | in | Central | London, | just | before | the | holidays. | Although | her | staging | could | have | made | more | of | the | setting, | it | still | provided | a | suitably | atmospheric | backdrop | for | de | Wet's | mildly | Gothic | tale | of | uneasy | bargains | and | unexpected | redemption. | From | today | Miracle | will | be | squeezed | into | the | tiny | basement | of | the | Leicester | Square | Theatre. | With | luck, | Reedman | should | be | able | to | turn | the | space's | extreme | intimacy | to | her | advantage. | She | might | also | do | more | to | bring | out | the | sardonic | humour | underlying | a | text | that | has | been | aptly | dubbed | a | tragicomedy | Her | cast | will | certainly | benefit | from | a | longer | run. | In | the | crypt | they | were | still | discovering | the | play's | internal | rhythms. | Although | this | can | be | exciting | to | witness, | there | it | yielded | a | faltering | collective | performance. | Reedman | is | fortunate | to | have | experienced | actors | such | as | Susannah | York | and | Tim | Woodward | in | her | corner. | The | latter | plays | an | actor-manager | of | the | old | school | who | uses | a | shameless | blend | of | orotund | phrases | and | flaring | temperament | to | manipulate | those | around | him. | York, | luminous | as | always, | is | his | faded | leading | lady. | Together | these | two | reveal | a | good | deal | about | the | vanity, | self-delusion | and | cruelty | coiled | inside | the | lives | of | those | devoted | to | the | theatre. | And | yet, | perhaps | because | de | Wet | herself | is | an | actress, | they | and | her | other | characters | also | demonstrate | more | tender | and | noble | attributes | of | forgiveness, | loyalty | and | a | show-must-go-on | fortitude.

Fig. 62 : Capture d'écran de la tokenisation du texte choisi.

Etape 3 : Lemmatisation de la critique

La troisième étape, la lemmatisation de la critique, consiste à attribuer la forme non fléchi des mots à chacun des termes de la critique. Pour ce travail, un tableau s'affiche au-dessous du texte choisi. Chacun des termes y est répertorié en fonction de son ordre d'apparition et son lemme est également indiqué.

	Text	Position	Unique Code	Lemma
0	It	PRON	10239237003504588839	it
1	's	ADV	614914527630368944	's
2	hard	ADJ	5791351258821946298	hard
3	to	PART	3791531372978436496	to
4	imagine	VERB	11158042576763907092	imagine
5	what	PRON	5865838185239622912	what
6	impact	NOUN	10122832419266696032	impact
7	a	DET	11901859001352538922	a
8	change	NOUN	3979101002134440957	change
9	of	ADP	886050111519832510	of
10	venue	NOUN	2277755173160143686	venue
11	will	AUX	18307573501153647118	will
12	have	VERB	14692702688101715474	have
13	,	PUNCT	2593208677638477497	,
14	on	ADP	5640369432778651323	on

Fig. 63 : Capture d'écran de la lemmatisation du texte choisi.

Etape 4 : Reconnaissance des entités nommées de la critique

Une entité nommée est une expression linguistique qui permet d'affecter une catégorie, comme les personnes, les lieux ou les organisations, aux noms. Ce traitement informatique pourra aider le chercheur à effectuer des recherches sur des éléments précis. Si nous voulions obtenir un panorama de toutes les personnes mentionnées dans les critiques de théâtre par exemple, il nous suffirait de rédiger un programme informatique qui récupérerait toutes les entités nommées « Personne. » Comme pour les exemples précédents, la critique sélectionnée au départ s'affiche au milieu de l'écran avec les entités nommées mises en évidence par des étiquettes de couleur différentes :

It's hard to imagine what impact a change of venue will have, on this uneven but tantalising show by the young **British NORP** company Ruby in the Dust. Written in **1992 DATE** by the much-lauded **South African NORP** playwright **Reza de Wet PERSON**, **Miracle ORG** is a **Chekhovian NORP** allegorical d
 their frustrations out on each other. Hospitality' arrives in the form of a middle-aged, black-clad matron (**Lynne Miller PERSON**), who happens to be the wife of the company's errant and virulently suspicious male ingenu. As the forlorn band of thespians eventually realise, this single spumed woman poses a greater thn
 teasing, timeless quality that lingers. It has certainly stayed with me. **Linnie Reedman's PERSON** production had a **five-day DATE** run in the crypt of **St Andrew Church ORG**, **Holborn GPE**, in **Central London LOC**, just before **the holidays DATE**. Although her staging could have made more of the sett
 the **Leicester Square Theatre FAC**. With luck, **Reedman PERSON** should be able to turn the space's extreme intimacy to her advantage. She might also do more to bring out the sardonic humour underlying a text that has been aptly dubbed a tragicomedy Her cast will certainly benefit from a longer run. In the crypt th
Tim Woodward PERSON in her corner. The latter plays an actor-manager of the old school who uses a shameless blend of orotund phrases and flaring temperament to manipulate those around him. York, luminous as always, is his faded leading lady. Together these **two CARDINAL** reveal a good deal about the vanity',
 forgiveness, loyalty and a show-must-go-on fortitude.

Fig. 64 : Capture d'écran de la reconnaissance des entités nommées du texte choisi.

Etape 5 : Analyse en dépendance de la critique

Enfin, la dernière étape consiste à annoter automatiquement les relations syntaxiques entre les mots dans une phrase. Cela permet notamment de comprendre le fonctionnement interne d'une phrase et d'avoir un aperçu de la manière dont sont organisés les discours d'un point de vue syntaxique. A nouveau, la représentation visuelle de ce traitement informatique facilite la compréhension de cette étape. Chaque phrase de la critique choisie apparaît à l'écran. De nombreuses flèches permettent d'indiquer les relations syntaxiques entre les mots.

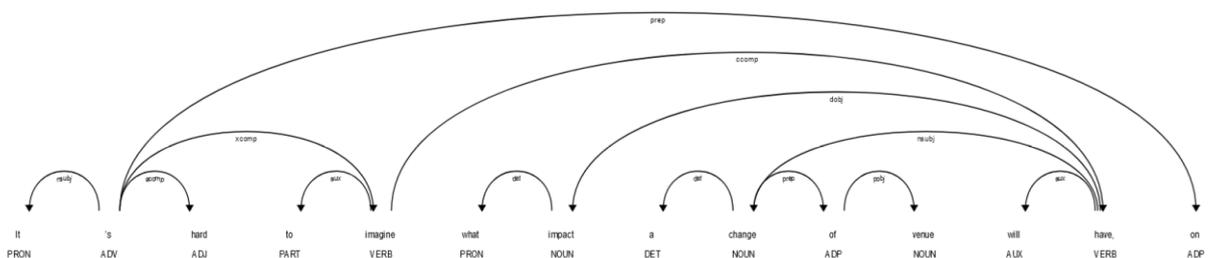


Fig. 65 : Capture d'écran de l'analyse en dépendance du texte choisi.

Transition

Cette première partie avait pour finalité de présenter l'identité numérique de notre plateforme ainsi que son contenu. Nous avons vu quelles stratégies numériques nous avons mises en place afin de jeter la lumière sur la double dimension littéraire et technique de notre projet. Concernant l'identité numérique de DramaCritiques.com, nous avons utilisé des

illustrations simples mais symboliques du théâtre et de l'informatique. Nous avons également pris soin de créer une page d'accueil qui résume tous les enjeux de notre recherche.

Concernant le contenu de notre site Internet, nous avons repris les trois expériences décrites dans ce manuscrit en proposant une structure similaire pour chacune d'entre elles. En mettant en avant des titres précis comme « Intérêt littéraire de cette question », « Quels moyens techniques utiliser pour y répondre ? », « Résultats » et « Interprétation », nous espérons avoir partagé nos connaissances de manière claire et concise.

Enfin, l'application « Derrière la machine » aborde frontalement les thématiques que nous avons présentées tout au long de cette partie. Cette application propose aux utilisateurs d'interagir avec les données. En choisissant une critique, ils peuvent découvrir toutes les étapes nécessaires pour rendre un texte intelligible pour l'ordinateur. Au-delà des brèves explications qui sont données, une illustration ainsi que le code informatique sont fournis à chaque fois. Cela permet non seulement de mieux comprendre les processus techniques, mais aussi de les réutiliser pour ses propres données.

L'importance de la reproduction des données sera centrale à la partie suivante. Nous avons divisé cette dernière sous-partie en deux sous-parties. Dans la première, nous nous focaliserons sur les Principes FAIR. Cette série de recommandations établies en 2016 par un groupe de chercheurs venant de multiples universités a pour objectif d'assurer la traçabilité, l'accessibilité, l'interopérabilité et la réutilisation d'un jeu de données. Nous montrerons comment notre corpus répond à tous ces critères. Enfin, dans la seconde sous-partie, nous aborderons toutes les pistes de recherche que nous avons esquissées mais que nous n'avons pas eu le temps d'explorer. Nous commencerons par étudier la question des femmes dans le théâtre contemporain londonien avant de présenter les multiples pistes auxquelles nous avons pensé.

IV.B Un potentiel d'exploitation pour les communautés littéraires et scientifiques

IV.B.1 Ouverture et accès aux données : les Principes FAIR

IV.B.1.1 Principes 1 et 2 : « Findable » et « Accessible »

En 2016, un groupe de chercheurs de différentes universités a mis en avant la nécessité de penser les enjeux du partage des données. Dans le contexte de l'utilisation croissante du big data, il fallait réfléchir aux conditions qui permettraient aux chercheurs d'ouvrir leurs données de sorte qu'elles soient exploitables par l'ensemble de la communauté scientifique. Ce groupe de chercheurs a mis ainsi au point une liste de principes à suivre, intitulés les « FAIR Principles », qui avaient pour objectif d'améliorer l'accessibilité aux données. Voici le résumé de l'article fondateur qu'ils ont publié :

There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measurable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals.²¹³

L'acronyme FAIR regroupe quatre types de principes : Findable, Accessible, Interoperable et Reusable. Nous détaillerons ici ces quatre principes, et montrerons en quoi notre recherche y répond. A ce jour, plus de 36 000 données de notre corpus sont en libre accès sur notre site et sur la plateforme Zenodo²¹⁴. Ces données téléchargeables rassemblent l'ensemble des critiques journalistiques et une partie des critiques numériques. Après avoir obtenu les droits de publication par *Theatre Record*, nous avons envoyé une demande d'autorisation individuelle à chaque blogueur. Au total, ce sont 22 blogueurs sur 28 qui nous ont accordé les droits de reproduction²¹⁵. L'ensemble de notre jeu de données est en revanche disponible sur demande.

Principe 1 : Findable, ou identifiable

²¹³ Marc D. WILINSON, et al., « The FAIR Guiding Principles for scientific data management and stewardship », *Sci Data*3 [En ligne], publié le 15 mars 2016, URL : <https://www.nature.com/articles/sdata201618>, consulté le 21 septembre 2022.

²¹⁴ <https://zenodo.org/record/6309656#.Y0MTyEzP2so> [En ligne], consulté le 21 septembre 2022.

²¹⁵ Voir annexes p. 262 – 269.

La première caractéristique importante d'un jeu de données est sa traçabilité. L'humain et l'ordinateur devraient pouvoir retrouver un corpus facilement, que cela soit de manière automatique ou par des recherches effectuées par un individu sur Internet.

Attribution d'un PID (Persistent Identifier)

Pour cela, il est nécessaire d'attribuer dans un premier temps une licence à la base de données en question sur un entrepôt. Un entrepôt est un endroit où les données sont hébergées et archivées. Nous avons déposé notre corpus sur Zenodo²¹⁶, une plateforme de dépôt d'archives ouvertes en libre accès. Notre base de données possède désormais une licence (Creative Commons Attribution 4.0 International) ainsi qu'un DOI. Un DOI (Digital Object Identifier) est un code d'identification unique et pérenne qui est attribué aux ressources numériques. Il permet, entre autres, de retrouver l'objet même si son adresse URL est changée et de faciliter son référencement dans les papiers de recherche. Notre corpus est aujourd'hui disponible à l'adresse suivante : <https://zenodo.org/record/6799656#.YzYBgUzP25c>

Description des métadonnées

La description des métadonnées est un autre élément qui permettra à un corpus d'être facilement identifiable par un ensemble de chercheurs, notamment en permettant d'automatiser des tâches de tri et de hiérarchisation dans la quête d'informations. Elles sont également utiles pour comprendre le contexte d'un corpus (types d'auteurs, chronologie, nature des données recueillies, etc.). Notre corpus est constitué de 17 types de métadonnées :

1. Le type de corpus (journalistique / numérique)
2. Le nom du journal / du blog
3. La date de création du journal / du blog
4. La date de publication de la critique
5. Le nom et le prénom du critique

²¹⁶<https://zenodo.org/>, [En ligne] consulté le 21 septembre 2022.

6. Le genre du critique
7. Le titre de la pièce
8. Le nom et le prénom du dramaturge (lorsque ces informations étaient disponibles)
9. La note attribuée par le critique (lorsque ces informations étaient disponibles)
10. Le lieu où a été jouée la pièce
11. Les coordonnées géographiques du théâtre
12. Le lien URL vers le site du théâtre
13. Le type de lieu dans lequel a été jouée la pièce (24 au total)
14. Les capacités d'accueil du théâtre (lorsque ces informations étaient disponibles)
15. Le coefficient de chaque thématique au sein d'une critique²¹⁷
16. Le coefficient de « positivité », de « négativité » et de « neutralité » au sein de chaque texte²¹⁸
17. Le coefficient des sentiments suivants : « Colère », « Peur », « Joie », « Tristesse » et « Surprise »²¹⁹

Toutes ces métadonnées ont été longuement nettoyées, ce qui garantit une certaine fiabilité de leur contenu. Ces informations permettent non seulement de trouver notre corpus plus facilement sur Internet grâce à une recherche automatique, mais elles ouvrent également un ensemble de possibles pour l'analyse du théâtre contemporain britannique.

Principe 2 : Accessible, ou accessibilité

Une fois que l'utilisateur a trouvé les ressources qu'il cherchait, il doit pouvoir y avoir accès facilement, c'est-à-dire pouvoir les télécharger rapidement sur son ordinateur personnel.

²¹⁷Afin d'obtenir une idée générale des sujets abordés dans les critiques, nous avons défini une liste de 20 thématiques (la nature, la science, les femmes, etc.). Nous avons ensuite calculé le nombre de synonymes en lien avec ces thématiques dans chacun des textes de notre corpus. Cette expérience a été réalisée afin de compléter une étude menée sur le *topic modeling*. Les résultats obtenus manquant de pertinence, nous avons décidé de ne pas les présenter dans ce manuscrit. Nous avons néanmoins choisi de laisser le coefficient de chacune de ces thématiques au sein de notre corpus afin qu'ils puissent servir à d'autres chercheurs. La liste de ces 20 thématiques est disponible en annexes, p. 270.

²¹⁸Ces coefficients ont été calculés grâce à l'algorithme XLNet. Tous les renseignements sur ce modèle ont été donnés dans le chapitre III, p. 175 – 176.

²¹⁹ Ces coefficients ont été calculés grâce à la librairie « text2emotion » avec Python. Nous avons également mené plusieurs expériences qui ne sont pas présentées dans ce manuscrit à partir de cette librairie. Nous avons néanmoins décidé de laisser ces coefficients à disposition des utilisateurs afin qu'ils puissent les utiliser pour réaliser d'autres études.

Utilisation d'un protocole standard

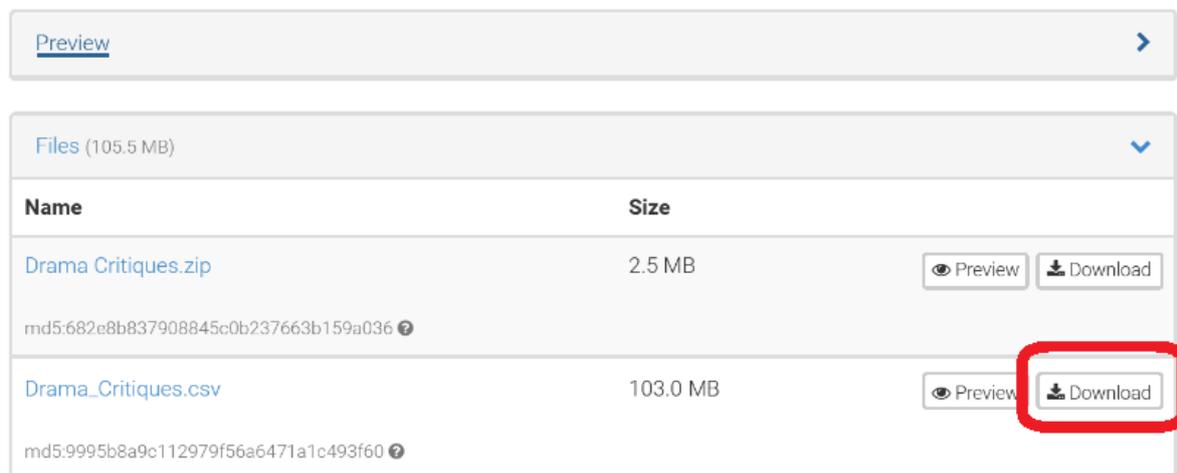
Pour rendre les données disponibles, les auteurs des Principes FAIR préconisent l'utilisation d'un protocole standard de communication. Ce dernier fait référence aux « procédures que suivent les machines pour communiquer entre elles²²⁰. » Dit autrement, pour chaque tâche effectuée, un protocole de communication est requis. Tandis qu'une page web sera ouverte avec HTTP (HyperText Transfer Protocol) par exemple, un partage de fichier sera effectué avec FTP (File Transfer Protocol). Pour que les données soient accessibles, il faut utiliser un protocole standard. C'est ce que nous faisons en déposant notre corpus sur Zenodo et sur notre site DramaCritiques.com, qui utilisent tous les deux un protocole libre et ouvert (le HTTP). Sur Zenodo, le téléchargement de notre corpus est facilement repérable. Il suffit de se rendre sur le bas de la page qui correspond à DramaCritiques et cliquer sur le bouton « Download. » La capture d'écran ci-dessous illustre ce procédé :

²²⁰https://doranum.fr/enjeux-benefices/principes-fair_10_13143_z7s6-ed26/[En ligne] consulté le 21 septembre 2022.

Drama Critiques gathers 10 years (2010 – 2020) of London contemporary theatre reviews. By focusing on two literary communities (journalists on the one hand and bloggers on the other hand), this corpus enables one to examine the poetical and political discourse journalistic and digital reviewers build. While our initial corpus is composed of more than 43 000 theatre reviews, the version made available here is constituted of 36 766 reviews. This is explained by the fact that we still have not received the authorisation of some of the bloggers to publish their data in open access. To have more information about our project, the different analyses of the corpus can be found here: <https://dramacritiques.com/en/home/>

The corpus based on journalism was created thanks to *Theatre Record*, a paper magazine originally created by the English critic Ian Herbert. *Theatre Record* reprints in full all the national drama critics' reviews of the latest productions in and out of London. Published every two weeks in England since January 1981, it is in January 2019 that its archives were digitized thanks to Julian Oddy (<https://www.theatrerecord.com/>). We have selected 23 newspapers in total which correspond to 21 717 theatre reviews. All of them were initially available in a PDF format. After having converted all the files in a textual format, a massive work of automatic and manual corrections was done on each of the files. This task represents more than 1050 hours of work.

The corpus based on blog platforms is constituted of the most popular 28 blog platforms on the Internet. They can be divided into two sub-categories: collective blog platforms on the one hand, and individual blog platforms on the other. Either these digital platforms are run by a publisher who invites other reviewers to post on his website, or the publisher publishes all of his reviews himself. In both cases, these authors are not paid for their activity, the content of their blog does not have a printed version and is completely free. In this version there are 21 blogs, or 15 049 reviews. All of them were automatically extracted thanks to web scraping techniques before being corrected (250 hours of work).



The screenshot shows a Zenodo file listing interface. At the top, there is a 'Preview' button with a right-pointing arrow. Below it, a section titled 'Files (105.5 MB)' contains a table with two rows. The first row is for 'Drama Critiques.zip' (2.5 MB) with 'Preview' and 'Download' buttons. The second row is for 'Drama_Critiques.csv' (103.0 MB) with 'Preview' and 'Download' buttons. The 'Download' button for the CSV file is highlighted with a red rectangle. Each row also includes an MD5 hash and a lock icon.

Name	Size	Preview	Download
Drama Critiques.zip md5:682e8b837908845c0b237663b159a036	2.5 MB	Preview	Download
Drama_Critiques.csv md5:9995b8a9c112979f56a6471a1c493f60	103.0 MB	Preview	Download

Fig. 68 : Capture d'écran de la page Drama Critiques sur Zenodo.

Nous créons également une page dédiée au téléchargement du corpus sur notre plateforme. Dans l'onglet « Base de données », puis dans « Téléchargements », il est possible d'accéder au corpus ainsi qu'aux scripts qui nous ont permis de réaliser nos analyses. Ces fichiers sont téléchargeables facilement. Il suffit de cliquer sur les boutons « Télécharger » pour les enregistrer sur un ordinateur personnel.

Téléchargements

Télécharger le corpus et les scripts

Afin de contribuer à l'ouverture et au partage des données, Drama Critiques met à disposition non seulement une partie de sa base de données (plus de 36 000 critiques en accès libre, l'intégralité du corpus est disponible sur demande), mais aussi tous les scripts qui ont permis la réalisation des analyses techniques. Ce corpus est également téléchargeable sur Zenodo (Licence : Creative Commons Attribution 4.0 International). Vous pouvez ainsi implémenter

à nouveau tous les algorithmes sur l'ensemble du corpus pour mieux comprendre les résultats, ou les adapter à vos propres données. Cela nous permet également d'assurer la transparence de nos résultats. La finalité de cette page est de permettre à toutes celles et à tous ceux qui aimeraient travailler aussi bien en littérature qu'en informatique de bénéficier d'un large corpus entièrement propre (plus de 1 400 heures de nettoyage) et novateur.

Pour télécharger le corpus sur le journalisme :

DRAMA_CRITIQUES_CORPUS_I

Pour télécharger le corpus sur les blogs :

DRAMA_CRITIQUES_CORPUS_II

Pour télécharger tous les scripts :

DRAMA_CRITIQUES_SCRIPTS

Pour télécharger l'application « Derrière la machine »

DERRIERE_LA_MACHINE

Fig. 69 : Capture d'écran de la section « Téléchargements » sur DramaCritiques.com.

Préservation de l'accès aux métadonnées

Conserver des jeux de données en ligne présente un risque. Un jeu de données peut disparaître pour un ensemble de raisons. C'est pourquoi les principes FAIR prônent de garder une trace des métadonnées. Dans le cas où le corpus deviendrait inaccessible, les chercheurs pourront ainsi soit contacter les auteurs à l'origine de la création du corpus, soit retrouver des articles publiés à son sujet. Pour pallier ce risque, nous prenons soin de rédiger une documentation concernant notre base de données et nos métadonnées. Nous laissons à disposition des utilisateurs six fichiers en version PDF qui sont disponibles sur notre répertoire Zenodo :

18. Un fichier intitulé « Dashboards.pdf » qui reprend les trois tableaux de bord que nous avons présentés dans le deuxième chapitre. Ces tableaux de bord permettent d'avoir une vue générale du corpus dans son entièreté ainsi que des deux corpus séparément (répartition des critiques au cours du temps, genre des auteurs des critiques, les dix journaux / blogs les plus représentés, etc.).
19. Un fichier intitulé « List of all the Newspapers and Blogs Represented in Drama Critiques (Open Access Dataset).pdf » qui contient une liste de tous les journaux et les blogs présents dans notre corpus disponible pour le téléchargement. Ce fichier contient la mention de 23 journaux et de 21 blogs.
20. Un fichier intitulé « List of all the Newspapers and Blogs Represented in Drama Critiques.pdf » qui détaille l'ensemble des journaux et des blogs qui font partie de notre corpus, soit 23 journaux et 28 blogs.
21. Un fichier intitulé « List of the 20 Themes Represented in Drama Critiques.pdf » qui présente les 20 thèmes que nous avons calculés à partir des synonymes de ces termes (« Body », « Love », « Religion », etc.).
22. Un fichier intitulé « List of the 24 Spatial Categories Represented in Drama Critiques.pdf » qui détaille les 24 catégories que nous avons identifiées pour répertorier les 446 lieux différents de notre corpus (« Theatre », « Natural Space », « Museum », etc.).
23. Un fichier intitulé « Main Characteristics of Drama Critiques' Corpus.pdf » qui contient deux tableaux avec les mêmes données. Le premier porte sur le corpus dans sa globalité, tandis que le second concerne le corpus disponible pour le téléchargement. Nous y avons résumé les éléments principaux : la nature des sources de nos corpus (PDF / HTML), le temps de création de ces deux bases de données, le nombre de journaux et de blogs, le nombre d'auteurs de critiques, le nombre de critiques rédigées et le nombre de mots.

Le détail de ces fichiers permettra ainsi, dans le cas où le corpus venait à disparaître ou à devenir inaccessible, de conserver une trace pérenne de notre corpus.

Création d'un moteur de recherche

Nous implémentons également un moteur de recherche à notre plateforme DramaCritiques.com afin de rendre nos données facilement accessibles. Grâce à un système de huit filtres distincts, les utilisateurs peuvent découvrir l'ensemble de notre corpus en ligne. La capture d'écran ci-dessous permet de visualiser l'interface de notre formulaire de recherche :



Explorez 10 ans du théâtre anglais à travers la critique littéraire !

Ce moteur de recherche vous permet de parcourir les quelques 43 000 critiques théâtrales de notre base de données. Issues des plus grands journaux de la presse anglaise et des blogs les plus populaires de la toile, ces critiques vous permettront de découvrir la scène contemporaine londonienne à travers le regard de ces deux communautés littéraires

Recherchez parmi 43043 résultats

Corpus

Journalistique
 Numérique
 Les deux

Genre du.de la critique

Homme
 Femme
 Les deux

Nom du.de la critique

Entrez le nom du.de la critique...

+

Date de publication de la critique

Année >

Titre de la pièce

Entrez le titre de la pièce...

+

Lieu

Entrez le nom du lieu...

+

Thèmes de la pièce

Sélectionnez les thèmes >

Journal / Blog

Sélectionnez les journaux / blogs >

RECHERCHER

Fig. 70 : Capture d'écran du moteur de recherche sur DramaCritiques.com.

Si nous cochons l'option « Femme » dans la catégorie « Genre du.de la critique », voici la page qui s'affiche dans un nouvel onglet :


[Accueil](#)
[Catégories](#)
[Focus sur les femmes](#)
[Base de données](#)
[A Propos](#)

19350 résultats correspondant à votre recherche pour :

Corpus : Both Genre du.de la critique : Female

CORPUS	JOURNAL / BLOG	DATE	NOM DU.DE LA CRITIQUE	TITRE	CRITIQUE	LIEU
Journalistique	Independent	07/01/2010	Zoe Anderson	VAREKAI	Celebrating its 25th anniversary this year, Cirque du Soleil has become the bigg...	Royal Court Theatre
Journalistique	Time Out	07/01/2010	Caroline McGinn	TRILOGY	Don't be put off by the length or the enigmatic title. This three-part, three-...	Battersea Arts Centre
Journalistique	Financial Times	08/01/2010	Sarah Hemming	THE LADY OR THE TIGER	Party political leaders in search of eye-catching policies to spice up their man...	Orange Tree Theatre
Journalistique	Observer	10/01/2010	Susannah Clapp	THE LADY OR THE TIGER	This is the season in which theatres revisit their histories. In the crumbling g...	Orange Tree Theatre
Journalistique	Independent on Sunday	10/01/2010	Jenny Gilbert	VAREKAI	Extravagant That's the word that best sums up Cirque du Soleil. That will cert...	Royal Court Theatre

Fig. 71 : Capture d'écran du résultat de la recherche en appliquant le filtre « Femme. »

Le nombre de requêtes qui correspondent au filtre « Femme » est indiqué en haut de la page. La liste de toutes les critiques rédigées par une femme est ensuite affichée. Il suffit de cliquer sur l'une d'entre elles pour accéder à son texte et à ses diverses informations annexes (journal ou blog dans lequel la critique a été publiée, date de parution, lieu dans lequel la représentation s'est tenue, etc.). Ce moteur de recherche est créé pour rendre nos données aisément accessibles. Un fichier csv avec plus de 40 000 lignes s'ouvre plus ou moins facilement selon les ordinateurs, ce qui peut compliquer l'accès aux textes. De plus, la manipulation d'Excel requiert quelques connaissances techniques que tout le monde ne maîtrise pas. C'est pourquoi la création de ce formulaire nous semble nécessaire pour contribuer au deuxième critère des Principes FAIR (Accessibility).

IV.B.1.2 Principes 3 et 4 : « Interoperable » et « Reusable »

Principe 3 : « Interoperable » ou Interopérabilité

Le troisième principe FAIR, nommé « Interoperable », consiste à faire en sorte que les données puissent être réutilisées quel que soit l'environnement informatique choisi.

Utilisation d'un vocabulaire contrôlé

Pour assurer l'interopérabilité, les auteurs des Principes FAIR recommandent d'utiliser un « vocabulaire contrôlé ». Un vocabulaire contrôlé est une « liste de termes prédéfinis servant à organiser des informations afin d'en faciliter la recherche et l'accès. Cette liste suit une structure bien définie afin de hiérarchiser le contenu²²¹ ». Ce vocabulaire permet aux différents systèmes informatiques de partager un langage commun et d'assurer l'interopérabilité entre eux. Le référencement de notre base de données sur un dépôt libre d'accès (Zenodo), l'attribution d'un DOI, la description des métadonnées ainsi que la mise à disposition de notre corpus en format csv sont autant d'éléments qui assurent l'échange de données d'un système à l'autre.

Métadonnées liées

Le Web des données constitue un réseau global d'informations scientifiques. En liant les métadonnées d'un corpus à d'autres données ou métadonnées grâce aux outils d'un dépôt comme Zenodo, le chercheur permet ainsi de créer une communauté de jeux de données liés entre eux. Les mots-clés que nous ajoutons à la version déposée dans Zenodo permettent la génération de tels liens. Ainsi, un chercheur qui tapera « Digital Humanities », « Theatre Reviews », « English Criticism », « Journalists », « Bloggers » ou « Reception Theories » trouvera soit notre corpus, soit d'autres corpus sur ces thématiques.

²²¹https://doranum.fr/enjeux-benefices/principes-fair_10_13143_z7s6-ed26/ [En ligne], consulté le 15 septembre 2022.

Principe 4 : Reusable, ou réutilisation

Description du contexte de la création du corpus

Le quatrième principe, « Reusable » ou « Réutilisation » assure la bonne reproductibilité des informations. Pour faciliter la prise en main d'un corpus et son analyse, un nombre d'informations relatives à sa création sont nécessaires. C'est pourquoi nous précisons dans la description disponible sur Zenodo dans quel cadre notre jeu de données a été produit, quand s'est déroulée sa réalisation, quels logiciels nous avons utilisés et quelles personnes ont contribué à sa mise en place. Nous incluons également les schémas récapitulatifs présentés au chapitre II de ce manuscrit²²² afin que les lecteurs puissent obtenir une vue d'ensemble des différentes étapes techniques qui nous ont menées à la création finale de notre jeu de données.

L'article que nous avons publié sur notre corpus²²³ apporte également de nombreux éléments informatifs quant au contexte littéraire de cette étude et récapitule toutes les démarches que nous avons entreprises, qu'elles soient littéraires ou techniques. Nous y résumons aussi les résultats obtenus ainsi que les pistes qu'ouvre notre travail.

Enfin, la mise à disposition des scripts qui ont permis de mener ces expériences sont mis à disposition des utilisateurs sur trois plateformes différentes : sur DramaCritiques.com, sur notre dépôt GitHub dont le lien figure sur notre site Internet ainsi que sur Zenodo. Tous ces éléments garantissent ainsi la bonne réutilisation de notre jeu de données.

Licence

Le dernier élément mis en avant par les auteurs des Principes FAIR pour assurer la réutilisation d'un jeu de données concerne l'obtention d'une licence. Une licence permet de délimiter et de définir les droits des utilisateurs. Plusieurs licences sont disponibles : la Creative Commons, l'Open Database ou la d'Etalab. Notre corpus est sous la Creative Commons, adaptée pour l'Open Data.

²²² Voir p. 81 et 88.

²²³ Mylène MAIGNANT, Damien PELLE, Gaëtan BRISON et Thierry POIBEAU, « DramaCritiques' Database » [En ligne], *Journal of Open Humanities Data*, vol. 8, no. 21, 2022, disponible sur : <https://openhumanitiesdata.metajnl.com/article/10.5334/johd.81/> [Dernier accès le 01/10/22].

Transition

Cette partie avait pour objectif de montrer quelles stratégies nous avons adoptées pour participer à la science ouverte afin de partager notre corpus. Nous avons détaillé les quatre critères prônés par les auteurs des Principes FAIR et avons montré comment nous y avons répondu. La traçabilité (Findable), l'accessibilité (Accessible), l'interopérabilité (Interoperable) et la réutilisation (Reusable) des données représentent quatre piliers dans tout travail qui comporte un aspect technique.

Même si cette thématique semble éloignée de notre objet d'étude sur le théâtre, elle doit pourtant être prise en compte afin d'assurer la reproductibilité et la transparence de nos résultats. En déposant notre base de données sous licence sur le dépôt Zenodo et sur notre plateforme, en documentant notre recherche et nos métadonnées, nous espérons que ces travaux pourront être réinvestis par d'autres chercheurs afin de faire émerger de nouvelles études sur le théâtre contemporain britannique.

La partie suivante abordera les pistes de recherche qu'offre l'ouverture de nos données. Nous avons choisi un exemple précis en lien avec la thématique des périphéries : celui de la présence des femmes dans le théâtre. Dans un premier temps, nous montrerons comment nous avons réutilisé notre base de données pour créer un nouveau moteur de recherche dédié aux dramaturges féminines exclusivement. Dans un second temps, nous détaillerons l'étude que nous avons menée sur la place des femmes dans l'industrie théâtrale londonienne des années 1995 à 2010.

Ces deux expériences sont conçues comme des points de départ pour d'autres recherches plus précises et plus abouties sur ce sujet. Le recours aux théories du genre serait par exemple particulièrement éclairant pour mettre en perspective certains de ces résultats. La question des femmes n'étant pas au centre de notre objet d'étude, et n'ayant pas eu assez de temps pour examiner ces thématiques plus en profondeur, nous exposerons ici brièvement nos découvertes. Ces deux exemples nous permettront cependant d'illustrer les multiples perspectives de recherche qu'ouvre notre travail.

IV.B.2 Un exemple de piste de recherche

IV.B.2.1 Un moteur de recherche dédié aux dramaturges féminines

Ce moteur de recherche est créé à partir des données disponibles dans notre corpus et celles mises à disposition par Doollee.com²²⁴. Doollee est une base de données rattachée à *Theatre Record*. Elle réunit sous forme de listes plus de 53 000 dramaturges du monde entier. Nous récupérons automatiquement ces données grâce au web scraping²²⁵ et nous les fusionnons avec notre propre corpus. Nous déterminons ensuite à l'aide de la librairie Gender.Detector²²⁶ le genre des dramaturges. La dernière étape consiste à mettre sous forme de formulaire l'ensemble de ces données sur notre site Internet.

The screenshot shows the website 'Drama Critiques' with a navigation bar containing 'Accueil', 'Catégories', 'Focus sur les femmes', 'Base de données', and 'A Propos'. The main heading is 'Le Théâtre écrit par les femmes'. Below it, a text block states: 'Cet index répertorie plus de 13 000 dramaturges femmes différentes venant des quatre coins du monde. Plus de 35 000 pièces sont ainsi référencées au total. En cliquant sur chacune des lettres de l'alphabet, vous accéderez à la liste des noms des dramaturges. La ou les pièces qu'elles ont écrite(s) s'afficheront ensuite sur votre droite.' Below this is an alphabetical navigation bar with letters A through Z. The letter 'A' is highlighted. Two columns of results are shown. The left column lists names: AARON Caroline, AARON Joyce, AASENG Mikhaala, ABATISIS Elisa, ABBA Marta, ABBINANTI Jill, ABBOT June, ABBOTT Laura, ABBOTT Sally, ABDEL-RAZEK Mariam, ABEL Leah, **ABELA Donna**, ABERLIN Brooke, ABERNETHY Margaret, ABLESER Deanna, and ABOUD Morgan. The right column lists works under the heading 'ABELA Donna': Arabian Nights (Drama), Aurora's Lament (18 min Poetic Drama, radio play), Caylee' Ukulele (Surreal comedy with songs 60 min), Circus Caravan (60 mins. Children's play. One act), Daphne Massacre, The (70 mins. absurd comedy), Edam Or Gouda (2 acts Play/Drama), Fathom (30 min radio play), Greatest Show On Earth, The (Musical comedy adventure), Hessa Gets Her Body Back (TYA), Highest Mountain Fastest River (60 mins. Play for young people.), Jump for Jordan (Play/Drama), Life Is A Funny Business (Gay/Lesbian, musical), Princess And the Pea, The (15 mins. Children's play), Stories From the 428 (short play), Summer Reign, A (genre unknown), and Tales From the Arabian Nights (60 mins. Children's play).

Fig. 72 : Capture d'écran du moteur de recherche sur les femmes sur DramaCritiques.com.

²²⁴ <http://www.doollee.com/index.php> [En ligne], consulté le 24 juin 2022.

²²⁵ Les explications de cette méthode sont données dans le chapitre II, p. 84 – 88.

²²⁶ <https://pypi.org/project/gender-detector/> [En ligne], consulté le 24 juin 2022.

Il suffit désormais aux utilisateurs de cliquer sur la lettre de leur choix pour que la liste des dramaturges apparaisse. En sélectionnant ensuite le nom de l'auteur, un rectangle s'affiche à droite de l'écran dans lequel sont inscrites toutes les mentions de ses pièces. Au total, ce sont plus de 13 000 dramaturges féminines venant du monde entier et plus de 35 000 pièces différentes qui sont répertoriées dans ce moteur de recherche. Au-delà de mettre en valeur le travail de milliers de femmes, ce moteur de recherche peut également servir d'outil aux chercheurs qui désireraient travailler sur les études du genre.

IV.B.2.2 Etude de cas : la place des femmes dans l'industrie théâtrale londonienne (1995 – 2010)

La seconde expérience réalisée à partir de notre corpus porte sur la place des femmes dans l'industrie théâtrale londonienne des années 1995 à 2010. Au début de chaque magazine publié par *Theatre Record*, un index qui répertorie tous les individus qui ont participé de près ou de loin aux pièces mentionnées dans les critiques est proposé. Une longue liste de noms et de prénoms accompagnés de la fonction que ces personnes ont occupée (metteurs en scène, acteurs, dramaturges, éclairagistes, etc.) est ainsi mise à disposition des lecteurs. A partir de ces index, nous allons déterminer dans quels corps de métier les femmes sont les plus présentes. Cette expérience nous permettra d'obtenir une vue générale sur une période donnée de la place accordée aux femmes dans l'industrie théâtrale londonienne.

Afin de répondre à cette question, la première étape consiste à récupérer, à structurer et à nettoyer ces index. Nous réalisons cette première étape grâce à l'utilisation d'expressions régulières avec Python. Après avoir harmonisé ces listes, nous les transformons en *dataframes*, ou en matrices de données, c'est-à-dire en tableaux. Nous faisons ensuite appel à la librairie « Gender.Detector » qui contient des milliers de prénoms qui ont été labellisés « féminin » ou « masculin ».

Il s'agit ensuite de rédiger un programme qui compare chaque prénom de notre fichier avec celui de la base de données « Gender.Detector ». La dernière étape consiste à effectuer des pourcentages et à les représenter sous forme de graphes. Huit professions ont été représentées : les éclairagistes, les compositeurs de musique, les dramaturges, les metteurs en scène, les traducteurs, les acteurs, les chorégraphes et les couturiers.

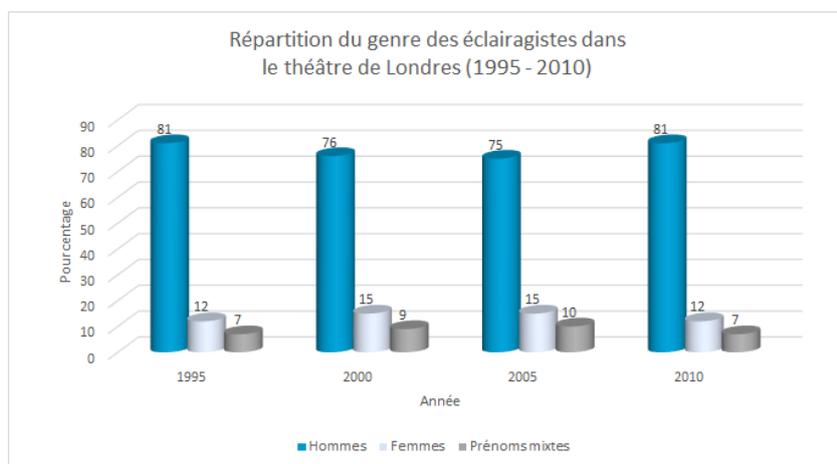


Fig. 73 : Représentation du pourcentage des hommes et des femmes éclairagistes.

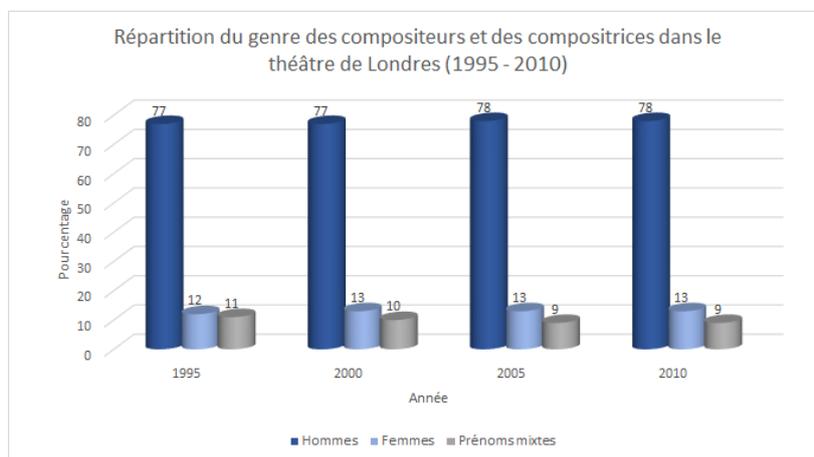


Fig. 74 : Représentation du pourcentage des hommes et des femmes compositeurs.

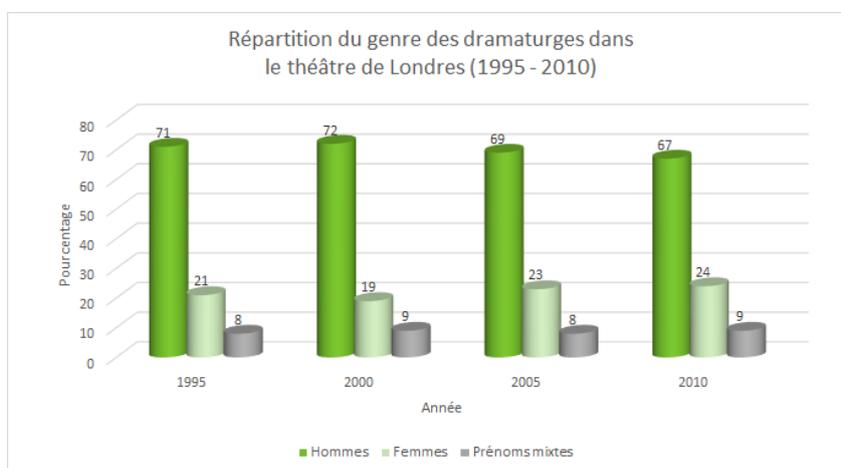


Fig. 75 : Représentation du pourcentage des hommes et des femmes dramaturges.

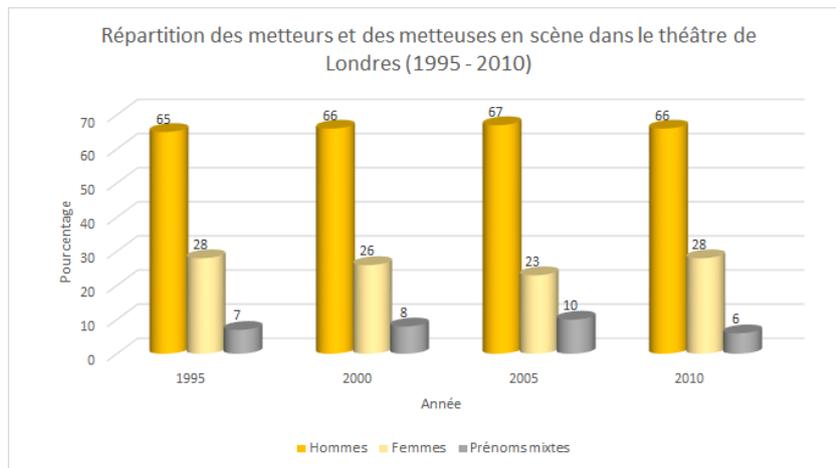


Fig. 76 : Représentation du pourcentage des hommes et des femmes metteurs en scène.

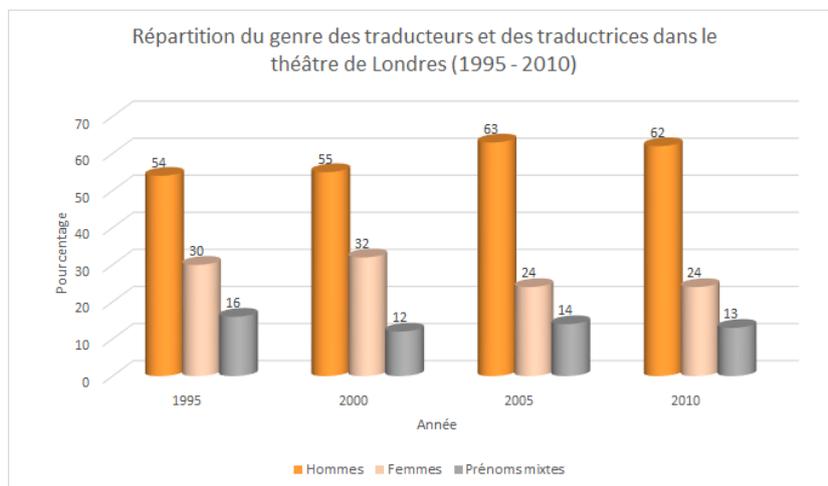


Fig. 77 : Représentation du pourcentage des hommes et des femmes traducteurs.

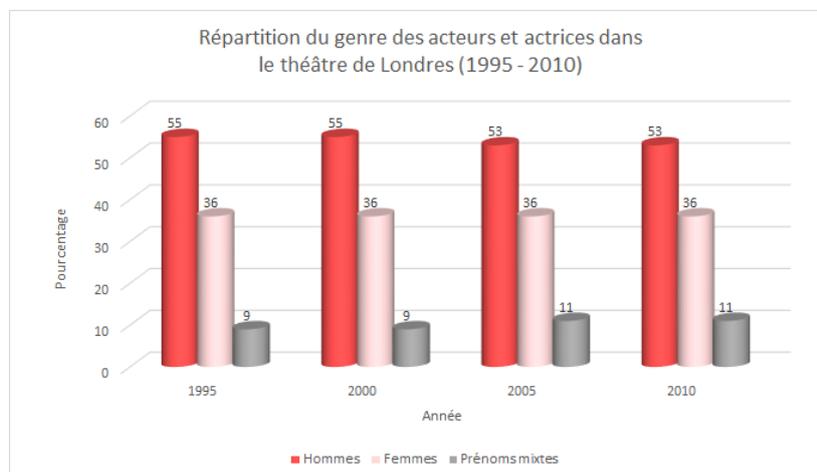


Fig. 78 : Représentation du pourcentage des hommes et des femmes acteurs.

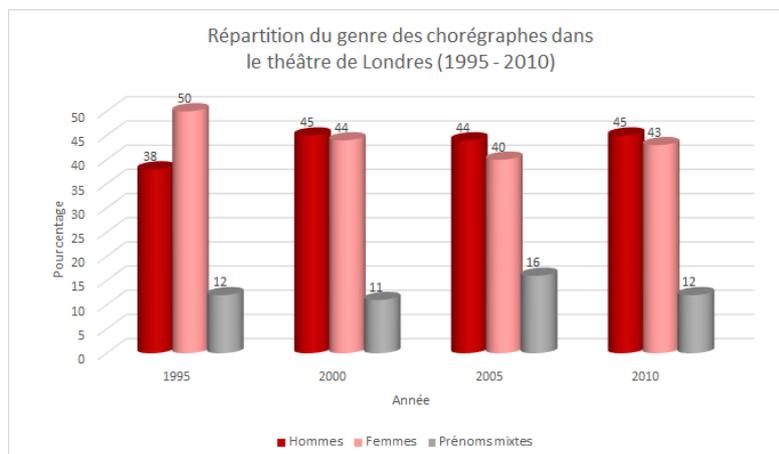


Fig. 79 : Représentation du pourcentage des hommes et des femmes chorégraphes.

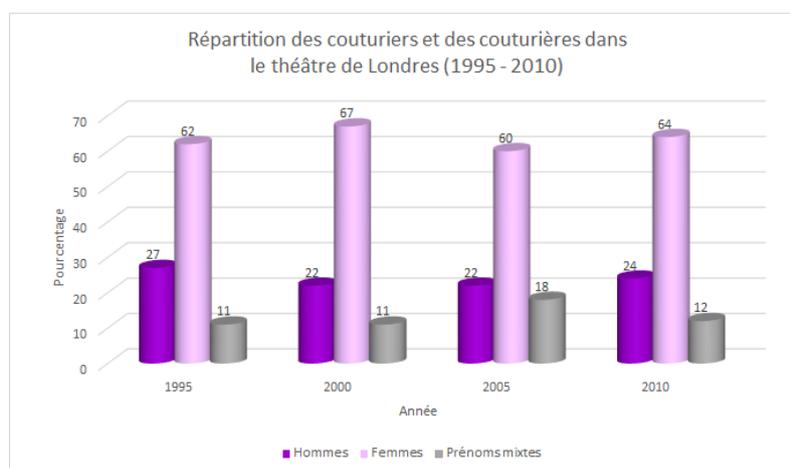


Fig. 80 : Représentation du pourcentage des hommes et des femmes couturiers.

Ces résultats jettent la lumière sur la sous-représentation des femmes dans la sphère théâtrale londonienne entre 1995 et 2010. En ce qui concerne le pourcentage des femmes dramaturges, nous pouvons observer une légère progression en 15 ans (+ 4 % entre 1995 et 2010). Les metteuses en scène sont légèrement plus présentes que les dramaturges, mais leur pourcentage reste nettement inférieur par rapport à leur homologue masculin. En 1995, elles représentent 28% de la profession de metteur.euse en scène contre 65% pour les hommes. Vingt ans plus tard, les chiffres montrent un rapport quasiment similaire : 28% des femmes sont metteuses en scène, contre 66% pour les hommes.

Les disparités tendent à s'effacer lorsque l'on considère la proportion d'acteurs et d'actrices. Bien que les femmes soient toujours en minorité, elles représentent en moyenne

36% des effectifs contre 53% pour les hommes. C'est dans le domaine de la danse que le ratio hommes / femmes trouve presque une forme d'égalité. Bien que le nombre de femmes décroisse depuis 1995 et celui des hommes augmente, les chiffres démontrent cependant que cette profession se caractérise par une certaine forme de parité entre les deux sexes.

C'est finalement la profession de costumière qui permet d'inverser la relation de pouvoir entre les hommes et les femmes. On retrouve des pourcentages quasiment identiques à ceux représentatifs de la profession de metteur en scène, excepté que dans ce cas, le rapport de force est inversé. C'est ainsi dans un domaine traditionnellement associé au travail féminin et à la sphère domestique que les femmes trouvent encore leur place au XXI^e siècle.

A nouveau, ces deux études de cas ont pour finalité d'illustrer les nombreuses projections qu'offre l'ouverture de nos données. Nous nous sommes concentrés sur la création d'un moteur de recherche dédié aux femmes ainsi qu'à l'analyse de leur présence dans les différentes professions théâtrales, mais nous aurions pu également nous intéresser aux distinctions stylistiques entre les critiques masculines et féminines par exemple. En réimplémentant l'intégralité de nos scripts sur notre corpus mais en ajoutant les filtres « Homme » et « Femme » qui sont disponibles dans notre fichier, il serait donc possible d'explorer la thématique « centre / périphéries » du point de vue du genre. La mise en accès libre de notre base de données ainsi que de nos scripts permet à l'ensemble de la communauté littéraire et scientifique d'explorer le théâtre londonien à travers de nombreux angles de vue différents.

CONCLUSION

Rappel de la problématique et de la méthode d'analyse

A partir du constat d'un décalage entre les critiques journalistiques et les critiques numériques, nous nous sommes posé les questions de recherche suivantes : « Est-il possible d'observer une rupture entre les journalistes et les blogueurs, ou entre le centre de la critique théâtrale londonienne et ses périphéries ? Quels outils informatiques mobiliser pour modéliser cette problématique et comment interpréter les résultats d'un point de vue littéraire ? » Pour répondre à ces questions, nous avons adopté une approche computationnelle. Nous avons ainsi réuni 43 043 critiques publiées entre 2010 et 2020 qui portent uniquement sur les spectacles joués à Londres. Ce corpus, composé des 23 plus grands noms de la presse écrite papier et des 28 blogs les plus populaires d'Internet, est le premier à réunir autant de critiques théâtrales contemporaines directement exploitables par la machine.

Nous avons divisé nos recherches en trois catégories qui abordaient trois questions littéraires auxquelles correspondaient trois domaines techniques. Nous avons ainsi examiné le style de ces auteurs, leur répartition géographique dans Londres et enfin leurs préférences esthétiques. Ces trois thématiques nous ont amenés à avoir recours à la stylistique computationnelle, à l'analyse d'opinion et aux Systèmes d'Information Géographique. Nous aurions pu nous concentrer sur un domaine en particulier et l'approfondir. La rupture entre les journalistes et les blogueurs n'ayant jamais été étudiée sous cet angle auparavant, nous avons choisi de dresser un portrait global de ces deux communautés plutôt que de nous focaliser sur un aspect en particulier. Cette approche nous permettait ainsi d'obtenir une vue d'ensemble de leurs rapports.

Découverte de deux communautés aux intérêts divergents...

Les résultats que nous avons obtenus prouvent qu'il y a plus de points de divergence entre ces deux groupes de spectateurs que de similitudes entre eux. Leurs horizons d'attente respectifs se reflètent dans leurs discours. Les blogueurs tendent effectivement à représenter les marges de la culture théâtrale londonienne en incluant davantage les minorités. Le résultat le plus significatif à cet égard concerne leur répartition géographique dans la capitale. Une rupture nette entre le centre et les périphéries de la critique est observable grâce à la cartographie numérique. Tandis que 446 lieux sont fréquentés par les critiques numériques,

deux fois moins d'endroits (210) sont visités par les critiques professionnels. Surtout, les blogueurs s'intéressent à une variété de lieux périphériques, tant d'un point de vue géographique, social que culturel.

Concernant le style de leurs écrits, nos analyses algorithmiques ont permis de révéler que les auteurs de ces deux communautés rédigent leurs critiques de manière différente. Leur intérêt ne porte pas sur les mêmes éléments d'un spectacle. Tandis que les blogueurs tendent à se focaliser sur des arguments qui relèvent de l'affect, les journalistes posent un regard plus neutre sur la scène. En affirmant leur subjectivité par l'utilisation récurrente du pronom personnel « Je », il semble que les blogueurs adoptent une approche plus sensible du théâtre que les journalistes. L'espace numérique leur offre une liberté spatiale et temporelle qui leur permet d'explorer et de développer leurs réflexions. La rédaction de sonnets, d'histoires illustrées ou encore de dialogues fictifs témoignent de la créativité dont ils font preuve. Ces nouvelles formes d'écriture participent non seulement à remettre en question les codes stylistiques du canon de la critique littéraire, mais aussi à construire un discours pluriel sur le théâtre. La multiplicité de leurs points de vue permet de rendre compte de la diversité d'analyses possibles d'une pièce.

Il y a donc bien un centre de la critique théâtrale qui est incarné par les journalistes. Ce centre impose la norme et domine par son prestige la masse de son lectorat et la densité de ses instances de reconnaissance. A l'inverse, sa périphérie, représentée par les blogueurs, est définie par un éloignement spatial, social, linguistique et politique. En réécrivant les codes de la critique classique, en affirmant leur subjectivité et en se rendant dans les lieux périphériques de Londres, les blogueurs remettent ainsi en question la culture théâtrale dominante symbolisée par les journalistes. La présence des blogueurs permet ainsi d'assurer une représentativité plus juste et plus diverse de la culture théâtrale londonienne.

... qui partagent néanmoins des points communs.

Si ces deux études réalisées sur la répartition géographique et sur le style révèlent une scission entre ces groupes de spectateurs, il est néanmoins nécessaire de nuancer ce propos. Premièrement, l'examen de leurs préférences esthétiques témoigne de points communs entre eux. Contrairement à ce que nous aurions pu penser au départ, les journalistes et les blogueurs favorisent les œuvres du canon de la littérature anglaise. Les pièces de Shakespeare, Dickens, Wilde et Tchekhov ont obtenu le plus grand nombre de critiques dans nos deux corpus.

Concernant les textes des critiques qui ont obtenu les meilleures appréciations, leur classement permet d'observer que dans les deux cas, les auteurs mettent à l'honneur des pièces aux thématiques similaires. Deuxièmement, nous n'avons procédé qu'à un nombre restreint d'études. En approfondissant certaines de nos expériences, d'autres points communs pourraient être observés entre les journalistes et les blogueurs.

Limites de notre approche

De plus, l'informatique est un outil qui présente des limites. L'utilisation des algorithmes n'est jamais neutre. De la même manière, un corpus, aussi important soit-il, reflète les motivations personnelles de son créateur, même si ce dernier tente d'être le plus objectif possible. L'ordinateur est une machine puissante mais pas autonome. Comme le montre notre application « Derrière la machine », c'est l'humain qui rédige et applique les scripts informatiques sur un ensemble de textes. En fonction du choix des données et des modèles implémentés, les résultats peuvent donc varier sensiblement d'une méthode à l'autre.

Il est donc important de souligner les limites de nos conclusions. La présence de la science des données n'implique pas une vérité absolue. Certes, elle permet d'obtenir une vision globale de la critique théâtrale londonienne, mais pas totale. Pour comprendre plus finement le rapport entre les critiques professionnels et les critiques amateurs, il faudrait poursuivre ces recherches, tant d'un point de vue littéraire, historique et sociologique, que technique.

Approfondir les théories de Bourdieu sur la distinction entre la culture dominante et la culture populaire apporterait un éclairage pertinent sur nos analyses. Aborder la notion d'« hégémonie culturelle », telle qu'elle est définie par Antonio Gramsci, contribuerait également à mettre en perspective nos résultats. De la même manière, appliquer d'autres traitements informatiques sur notre corpus permettrait d'apporter des réponses plus fines à nos questions.

La construction de ce corpus et les expériences techniques ayant occupé une grande partie de notre travail, nous n'avons pas pu examiner toutes les pistes de recherche que nous avons identifiées. Notre travail a cependant permis de jeter la lumière sur un phénomène encore peu étudié jusqu'ici : le canon de la critique théâtrale londonienne et ses marges. En créant un corpus qui réunit la majeure partie de leurs productions écrites, nous avons pu apporter des réponses chiffrées à nos questions littéraires.

Pistes d'ouverture

Les pistes de recherche qu'offre notre travail sont nombreuses. C'est en partie pourquoi nous avons créé DramaCritiques.com. Cette plateforme permet non seulement de donner davantage de visibilité aux marges de la critique théâtrale, mais aussi de présenter, d'expliquer et de partager notre démarche. En respectant les Principes FAIR, nous nous sommes assurés que nos données puissent être facilement réutilisées par l'ensemble de la communauté scientifique et littéraire. Nous avons détaillé dans le dernier chapitre l'étude que nous avons commencée sur la place des femmes dans l'industrie théâtrale londonienne. Cette piste de recherche ne constitue qu'un exemple parmi tant d'autres.

Premièrement, nous pourrions approfondir notre étude en nous concentrant sur les différences et les similitudes au sein même de nos deux sous-corpus. Il s'agirait d'effectuer les mêmes expériences mais à une échelle plus réduite. Cela nous permettrait de répondre à des questions comme : « Quels blogueurs fréquentent le plus d'endroits différents ? » ou « Quels journalistes tendent à adopter le style le plus neutre ? »

Deuxièmement, nous pourrions également élargir notre étude en examinant les liens entre les critiques journalistiques et numériques à une échelle nationale et internationale. Dans le premier cas, *Theatre Record* met à disposition de ses utilisateurs les critiques parues dans les journaux régionaux anglais. Dans le second cas, de nombreux blogs ont également été créés à propos du théâtre américain et australien. En reprenant les scripts informatiques qui sont en accès libre sur différentes plateformes, il serait ainsi possible de créer d'autres corpus qui permettraient d'analyser cette problématique à des échelles différentes

Nous pourrions également envisager d'étudier la réception d'une seule œuvre, comme celles de Shakespeare dont les mentions sont nombreuses dans notre corpus. Nous avons répertorié 1974 critiques au total qui portent uniquement sur les adaptations de ses pièces. Nous pourrions déterminer quelles sont celles qui ont obtenu le plus de succès, dans quels théâtres elles ont été jouées et pourquoi elles ont été autant appréciées. Faire un état des lieux de l'histoire visuelle de la scène contemporaine londonienne pourrait constituer une autre piste de recherche intéressante. Grâce à la labellisation de notre corpus, nous avons désormais un algorithme capable de reconnaître automatiquement les différentes catégories d'une critique théâtrale. En ne conservant que celles qui sont liées aux « Détails visuels, olfactifs et sonores », nous pourrions obtenir une vue d'ensemble de l'évolution des décors scéniques au cours de ces dix dernières années dans le théâtre londonien.

Les pistes de recherche que nous venons de présenter brièvement ne sont que quelques exemples parmi d'autres. Les 40 043 critiques de notre corpus offrent une multitude de perspectives qu'il reste à découvrir. Les scripts informatiques laissés à disposition de l'ensemble de la communauté littéraire et scientifique faciliteront le travail de ceux qui voudraient poursuivre notre démarche. Au-delà de la présentation de nos expériences dans ce manuscrit, la plateforme DramaCritiques.com est également un autre moyen d'accéder facilement et rapidement à notre travail. Nous espérons que cette recherche ouvrira la voie à de nouvelles connaissances sur le théâtre contemporain britannique.

ANNEXES

Présentation des trois solutions mises en place pour la conversion des fichiers PDF en fichiers Texte

Solution 1

La première solution consiste à avoir recours l’outil Onlineocr²²⁷. OnlineOCR.net est un logiciel en ligne de reconnaissance optique de caractères (« *Optical Character Recognition* » en anglais). Gratuit si l'utilisateur ne convertit que 15 pages par heure, il devient en revanche payant si le volume de fichiers augmente. Nous testons d’abord Onlineocr sur quelques pages seulement. Si le traitement des fichiers PDF est rapide (quelques secondes environ pour 10 pages), le résultat est cependant non-satisfaisant. Le format en colonnes n’étant pas, ou partiellement, respecté, le fichier de sortie n’est pas exploitable pour de futures analyses.

Solution 2

Dans un deuxième temps, nous utilisons GROBID (GeneRation Of Bibliographic Data)²²⁸. GROBID est une librairie qui repose sur le machine learning. Il s'agit d'entraîner un modèle sur un fichier de notre choix, de préférence un article scientifique, afin que le logiciel mémorise la structure de ce dernier pour le convertir en fichier TEI ou XML. Cette solution offre un gain de temps considérable si l'opération est bien réalisée. L'encodage du texte à l'aide de balises permet un traitement précis des données et la production d'analyses détaillées. Le format de nos fichiers est cependant trop complexe pour GROBID, car trop éloigné de la forme des articles scientifiques.

²²⁷<https://www.onlineocr.net/fr/> [Dernier accès le 28/08/20].

²²⁸<https://grobid.readthedocs.io/en/latest/Introduction/> [Dernier accès le 10/09/20].

Solution 3

Nous essayons d'autres plateformes en ligne, comme Transkribus et Kraken, mais les résultats restent toujours insatisfaisants. Nous utilisons alors le package « pdftools²²⁹ » disponible sous le langage de programmation R. Cet outil permet d'extraire et de convertir les informations d'un fichier PDF en fichier text. La reconnaissance de caractères sur le fichier de sortie est relativement bien effectuée, et l'enchaînement des critiques sous forme de colonnes est également pris en compte. En revanche, nous sommes confrontés au même problème que celui que présentaient les fichiers PDF originaux. Il nous est impossible d'extraire les colonnes une à une. Nous sommes donc contraints de laisser cette troisième solution de côté.

²²⁹<https://cran.r-project.org/web/packages/pdftools/pdftools.pdf> [Dernier accès le 10/09/20].

Script informatique rédigé pour l'extraction des informations sur site Internet WestEndWilma.com

Import des bibliothèques nécessaires à la réalisation de cette tâche :

```
import requests
from bs4 import BeautifulSoup #bs4 pour BeautifulSoup4
import pickle
import string
```

Nous procédons en deux temps. Le premier consiste à récupérer les liens des pages où se trouve le texte. Pour réaliser cette tâche, il va falloir aller sur chacune des 240 pages. Dans un second temps, nous aspirerons le texte des pages.

Partie I

Nous commençons par l'analyse d'une page :

```
r = requests.get("https://www.westendwilma.com/category/reviews/plays/")
```

Nous vérifions le code source :

```
print(r.text[0:100])
```

Nous cherchons ensuite à rendre le texte intelligible pour la machine grâce à la bibliothèque BeautifulSoup. :

```
soup = BeautifulSoup (r.text, 'html.parser')
```

Nous pouvons remarquer que le block qui nous interesse est une liste (ul) de class block2. Dans cette liste se trouvent les informations dont nous avons besoin.

Nous commençons par vérifier qu'un seul block correspond à ce que nous désirons extraire :

```
len(soup.find_all('ul', attrs = {'class': "block2"})) # ok
```

Nous conservons désormais le block :

```
block = soup.find('ul', attrs = {'class': "block2"})
```

```
print(block)
```

```
len(block)
```

Nous enregistrons désormais les liens vers les autres pages :

```
liste = block.find_all({'p'})
```

```
len(liste) # nombre de liens
```

Puis nous initialisons la liste où seront stockés les liens :

```
liens = []
for i in liste :
    x = BeautifulSoup (str(i), 'html.parser')
    liens.append(x.find('a').get('href'))
```

Nous créons désormais une fonction pour automatiser cette tâche :

```
def recup_liens (path) :

    liens = []
    r = requests.get(path)

    soup = BeautifulSoup (r.text, 'html.parser')

    block = soup.find('ul', attrs = {'class' : "block2"})

    liste = block.find_all({'p'})

    for i in liste :
        x = BeautifulSoup (str(i), 'html.parser')
        liens.append(x.find('a').get('href'))

    return(liens)
```

Nous vérifions que le code fonctionne bien :

```
liens_test = recup_liens ('https://www.westendwilma.com/category/reviews/plays/page/2/')
print(liens_test)
```

Désormais, nous récupérons les liens de toutes les pages :

```
liens_grande_liste = []
```

Nous commençons par ajouter les liens trouvés sur la première page :

```
liens_grande_liste += liens
```

Puis, nous recommençons l'opération :

```
for i in range(2,242):
    liens_grande_liste +=
    recup_liens('https://www.westendwilma.com/category/reviews/plays/page/' + str(i) + '/')
    print(i)
```

Enfin, nous enregistrons les liens :

```
pickle.dump(liens_grande_liste, open('liens_scraping_mylene', 'wb'))
```

Partie II

Nous effectuons la même démarche que dans la première partie. Nous analysons une première page prise au hasard :

```
r = requests.get(liens_grande_liste[200])
```

```
soup = BeautifulSoup (r.text, 'html.parser')
```

```
# Nous récupérons la date de publication de la critique :
```

```
date = soup.find('meta', {'property' : "article:published_time"}).get('content')
```

```
print(date)
```

```
date = date[0:10]
```

```
# Nous récupérons le titre de la critique :
```

```
title = soup.find('meta', {'property' : "og:title"}).get('content')
```

```
title = title.split("(")[0].split(" ")[1:]
```

```
title.remove("(")
```

```
title += [date]
```

```
title = '_' .join(title)
```

```
print(title)
```

```
block = soup.find('div', {'class' : "main-content "})
```

```
block.find_all('p')
```

```
raw_text = block.text.split('_____')[0]
```

```
# Nous ne conservons que le texte qui nous intéresse :
```

```
text = raw_text[raw_text.find('R'):]
```

Nous faisons deux choses à la fois ici : nous créons dans un premier temps un fichier texte, puis nous écrivons une fonction afin de pouvoir écrire dedans :

```
open(title+'.txt','w', encoding="utf-8").write(text)
```

Il ne reste plus qu'à automatiser l'ensemble du script :

```
def get_and_write_text (path):
```

```
    r = requests.get(path)
```

```
    soup = BeautifulSoup (r.text, 'html.parser')
```

```
    date = soup.find('meta', {'property' : "article:published_time"}).get('content')
```

```
    date = date[0:10]
```

```
    file_name = soup.find('meta', {'property' : "og:title"}).get('content')
```

```
    file_name = file_name.split("(")[0].split(" ")
```

```
    if "" in file_name:
```

```
        file_name.remove("")
```

```
for i in range(len(file_name)) :  
    file_name[i] = ".join([str(e) for e in file_name[i] if e.isalnum()])  
  
file_name += [date]  
  
file_name = '_'.join(file_name)  
  
block = soup.find('div', {'class': "main-content"})  
  
block.find_all('p')  
  
raw_text = block.text.split('_____')[0]  
  
text = file_name + '\n' + raw_text[raw_text.find('R'):]  
  
open('fichier_bis/' + file_name + '.txt','w', encoding="utf-8").write(text)
```

Classification des 446 lieux mentionnés dans notre corpus en 24 catégories

Academic places

Aylward Academy
Drama Studio London
Dulwich College
English National Ballet School
Guildhall School of Music and Drama
King's College London
London Academy of Music and Dramatic Art
London College of Music
London School of Economics
London School of Musical Theatre
Mountview Academy of Theatre Arts
Musical Theatre Academy
Queen Mary University of London
Royal Academy of Arts
Royal Academy of Dramatic Arts
Royal Central School of Speech and Drama
Trinity Laban Conservatoire of Music and Dance

Arts Association

Applecart Arts
Bell House
Chats Palace
Greenwich Dance Borough Hall
Highbury Opera Theatre
Jaguar Shoes Collective
Lauderdale House
National Youth Music Theatre
National Youth Theatre
Platform Southwark
Rag Factory
Redbridge Drama Centre
Safehouse 1
Swan Wharf
Ugly Duck
Vault Festival

Arts Centre

Bernie Grant Arts Centre
Bishopsgate Institute (LMTO)
Rich Mix Theatre
Rivington Place
Space Arts Centre
Tramshed Theatre
Trinity Buoy Wharf
Waterman's Arts Centre
Woolwich Theatre

Building

Balfron Tower
Bussey Building
Great Eastern Quay
London County Hall
London House of Commons
Millbank Tower
Paynes and Borthwick Warf
Rose Lipman Building
Royal Artillery Barracks
Tower of London

Cinema

Crystal Theatre
Empire Cinema
Haymarket
Rio Cinema
Riverside Studios

Concert Place

Acklam Village Market
Cadogan Hall
CLF Art Cafe
Fairfield Halls
O2 Arena
Royal Albert Hall Theatre
Rudolf Steiner Theatre
Wigmore Hall

Wilton Music Hall

Cultural Centre

Barbican
British Film Institute
British Library
Draper Hall Theatre
House of Illustration
Institute of Contemporary Arts
John MacIntosh Arts Centre
Lincoln Center
London Irish Cultural Centre
London Latvian Centre
London Welsh Centre
Middle Temple Hall
Pauline Quirke Academy of Performing Arts
Queen Elizabeth Hall
Royal Festival Hall
Somerset House
South Bank Centre
Steiner House Theatre
Tate Modern
Two Temple Place
Union Chapel

Gallery

Archivist Gallery
Crypt Gallery
Guildhall Art Gallery
Hayward Gallery
Hoxton Gallery
Saatchi Gallery

Hotel

Ace Hotel
Bermondsey Square Hotel
Corinthia Hotel
Grange Holborn Hotel
Hotel Café Royal
Kingsway Hall Hotel

Langham Hotel
May Fair Hotel
Rockwell House
Town Hall Hotel

Library

John Harvard Library
Kensington Central Library
Leytonstone Library
Library Foyles
London Library
Putney Library
Westminster Reference Library

Miscellaneous

Calder Bookshop and Theatre
Clerkenwell House of Detention
COG ARTSpace
Doon Car Park
Kensington Palace
Kenwood House
London Hippodrome
London Zoo
Lords Disco
Museum of Comedy
Only Connect
Rowans Bowling
Selfridges
Smithfield Car Park
Smithfield Market
Stratford Circus
Swedenborg House

Museum

British Museum
Brunel Museum
Cinema Museum
Grant Museum of Zoology
Jewish Museum London
London Film Museum

London Museum of Water and Steam
London Natural History Museum
Madame Tussauds Museum
Museum of Childhood
Museum of London
Museum of London Docklands
Museum of the Order of Saint John
National Portrait Gallery
Old Operating Theatre
Victoria and Albert Museum
Wellcome Collection

Natural Space

Abney Park Cemetery
Arundel and Ladbrooke Gardens
Caledonian Park
Camley Street Natural Park
Coram's Fields
Finsbury Park
Kensington Gardens
Kew Gardens
Marylebone Gardens
Meanwhile Gardens
Normand Park
Park Hill Park
Rookery Gardens
Russell Square Gardens
Saint George's Gardens
Stephens House and Gardens
The Jetty
Tower Hamlets Cemetery Park

Online

Pub Theatre

Bread and Roses Theatre
Drayton Arms Theatre
Londor Theatre
Lion and Unicorn Theatre
Old Red Lion Theatre

Pedley Street Station
Pheasantry Restaurant
Rosemary Branch Theatre
Ye Olde Rose and Crown

Public Space

Brunel Tunnel
Covent Garden
Leicester Square Spiegeltent
Lewis Cubitt Square
Saint Katharine Docks
South Kentish Town Tube
Station
Waterloo Tunnels

Reception Venue

2 Northdown
Cecil Sharp House
East Winter Garden
Free Word Centre
Gray's Inn Hall
Printworks
SLAM King's Cross

Religious Place

Asylum Chapel
Caroline Gardens Chapel
French Protestant Church
Saint Giles-in-the-Fields
Saint James's Church
Saint John's Smith Square
Saint Paul's Church
Southwark Cathedral
St Leonard's Church
St Mary Magdalene Church
St Matthews Church
St Paul's Actors Church
St Saviour's Church

Swiss Church
Temple Church
Westminster Abbey

Restaurant Aeras

Bargehouse
Cahoots Bar
Camden Head Pub
Canal Café Theatre
Canvas Cafe
Cock Tavern
CoffeeWorks Project
Drink, Shop & Do Cafe
Front Room Café
Harringtons Pie & Mash
London Old Fire Station
Maison Bertaux
Matchstick Piehouse
Matthews Yard Restaurant
Tea House Theatre
Zedel Brasserie
Ziferblat

Studio

Abbey Road Studios
Cable Street Studio
Dance Attic Studios
Map Studio Cafe
Map Studio Cafe

Theatre

Abbey Theatre
Above The Arts Theatre
Above the Stag Theatre
Adelphi Theatre
Albany Theatre
Aldwych Theatre
Alexandra Palace Theatre
Almeida Theatre
Ambassadors Theatre

Angel Comedy Club
Apollo Theatre
Arcola Theatre
Arts Depot Theatre
Arts Theatre
Bakehouse Theatre
Barn Theatre
Barons Court Theatre
Battersea Arts Centre
Big House Theatre
Blackheath Halls Theatre
Bloomsbury Theatre
Blue Elephant Theatre
Boulevard Theatre
Bridewell Theatre
Bridge Theatre
Brixton House
Broadway Theatre
Bunker Theatre
Bush Theatre
Cambridge Theatre
Camden Comedy Club
Camden People's Theatre
Canada Water Theatre
Centre 17
Cervantes Theatre
Chapel Playhouse Theatre
Charing Cross Theatre
Chickenshed Theatre
Chiswick Playhouse
Churchill Theatre
Cockpit Theatre
Colab Theatre
Coliseum Theatre
Coronet Theatre
County Hall
Courtyard Theatre
Criterion Theatre
Curve Theatre
Dominion Theatre
Donmar Theatre
Drury Lane Theatre
Duchess Theatre
Duke of York's Theatre
Etcetera Theatre
Exchange Theatre
Finborough Theatre
Forest Community Theatre

Fortune Theatre
Found 111 Theatre
Garden Theatre
Garrick Theatre
Gate Theatre
Gatehouse Theatre
Gielgud Theatre
Golden Goose Theatre
Greenwich Theatre
Hackney Theatre
Half Moon Theatre
Hampstead Theatre
Harold Pinter Theatre
Hope Theatre
Hoxton Hall
Intimate Theatre
Jack Studio Theatre
Jacksons Lane Theatre
Jermyn Street Theatre
Katzpace Theatre
Kiln Theatre
King's Cross Theatre
King's Head Theatre
Laban Theatre
Leicester Square Theatre
Linbury Theatre
Little Angel Theatre
London Improv Theatre
London Palladium
London Theatre Workshop
London Wonderground?
Lost Theatre
Lyceum Theatre
Lyric Hammersmith Theatre
Marble Arch Theatre
Menier Chocolate Factory Theatre
Millfield Arts Centre
N16 Theatre
National Theatre
Network Theatre
New Diorama Theatre
New End Theatre
New London Theatre
New Wimbledon Theatre
Noël Coward Theatre
Novello Theatre
Nursery Theatre
Old Vic Theatre

Omnibus Theatre
Orange Tree Theatre
Orchard Theatre
OSO Arts Centre
Other Palace Theatre
Palace Theatre
Park Theatre
Peacock Theatre
Peckham Theatre
Pentameters Theatre
Phoenix Theatre
Piccadilly Theatre
Place Theatre
Playground Theatre
Playhouse Theatre
Pleasance Theatre
Polka Theatre
Prince Edward Theatre
Prince of Wales Theatre
Proud Cabaret Camben
Puppet Barge Theatre
Putney Arts Theatre
Questors Theatre
Regent's Park Open Air Theatre
Richmond Theatre
Rose Theatre
Roundabout Theatre
Roundhouse Theatre
Royal Court Theatre
Sadler's Wells Theatre
Savoy Theatre
Scoop Theatre
Seven Dials Playhouse
Shaftesbury Theatre
Shakespeare's Globe Theatre
Shaw Theatre
Shed Theatre
So & So Art Club
Soho Theatre
Sondheim Theatre
Southwark Playhouse Theatre
Space 18
St James Theatre
St Martin's Theatre
Stockwell Playhouse
Streatham Hill Theatre
Tabard Theatre
Tabernacle Theatre

Tara Arts Theatre
Temple Studios
The Last Refuge
Theatre 503
Theatre Delicatessen
Theatre Royal Haymarket
Theatre Royal Stratford East
Theatro Technis
Tooting Arts Club
Tower Theatre
Toynbee Studios
Trafalgar Theatre
Tristan Bates Theatre
Troubadour White City Theatre
Turbine Theatre
Unicorn Theatre
Union Theatre
University of the Arts London
Vaudeville Theatre
Vaults Theatre
Victoria Palace Theatre
Waterloo East Theatre
Watermill Theatre
White Bear Theatre
Wimbledon Studio Theatre
Woodbridge Theatre
Wyndham's Theatre
Yard Theatre
Young Vic Theatre
Yvonne Arnaud Theatre

Town Hall

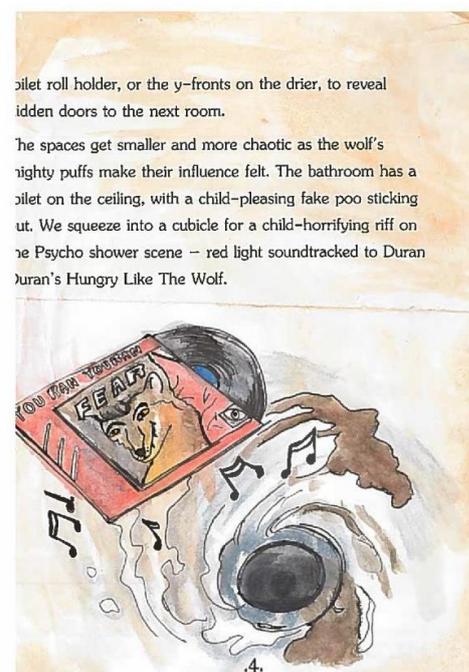
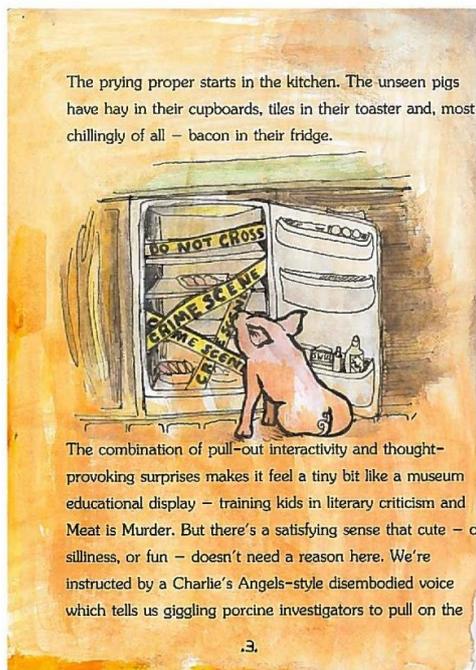
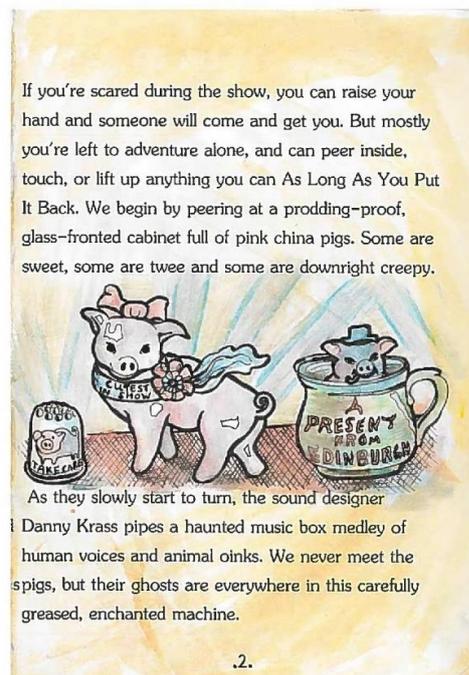
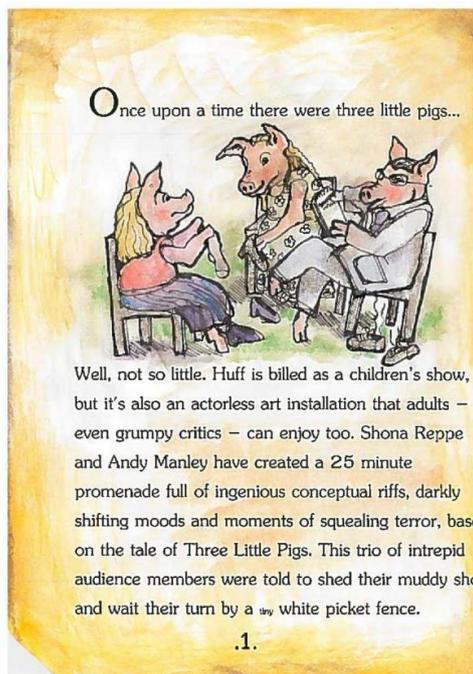
Bethnal Green Town Hall
Hornsey Town Hall
Islington Town Hall
Limehouse Town Hall
Shoreditch Town Hall
Stoke Newington Town Hall

Illustration de la pièce *Huff* (2014)

Critique de la pièce *Huff* (2004)

La représentation a eu lieu au Traverse Theatre de Londres, du 1^{er} au 24 août 2014.

Les illustrations ont été rédigées par la blogueuse Alice SAVILLE et ont été publiées publiées sur Exeunt.com à l'adresse suivante : <https://www.shonareppe.co.uk/huff.html> [Dernier accès le 02/10/22].

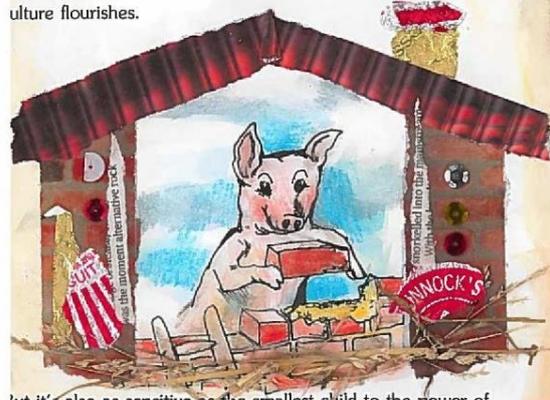


After the horror, the show's most gorgeous moment is a still life of ravishing beauty. Where Dutch masters painted half-eaten meals or shrivelled, caterpillar eaten roses to demonstrate the futility of earthly vanity and the inexorability of time, Huff manages to suggest both these themes, and the immediate presence of a great big wolf with lungs the size of bouncy castles. A table spread with waxy fruit and Tunnock's tea cakes is tilted by an invisible gale, cutlery and wrappers hang in the air in chaotic disorder that, in Sergey Jakovsky's constantly ingenious lighting design, spells out "Help!" on the wall in hectic shadow.



.5.

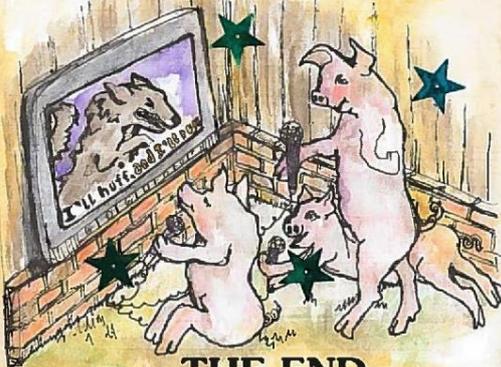
Huff isn't a story told in soft watercolour washes. Instead, it lays with a whole paintbox of visual references. The dog-ared original fairytale has been collaged over and rebuilt in straw, sticks and bricks, and decorated with shiny pop culture flourishes.



but it's also as sensitive as the smallest child to the power of cleverly pointed spotlight and of subtle shifts in perspective that make a tiny plastic toy become a Big Bad Wolf, prowling across the wall in bleak shadow.

.6.

Huff's sense of reality and order being overturned recalls older, Grimmier fairy tales. But its emphasis on magical shifts of mood and atmosphere is foil-wrapped fresh – a triumph sealed by three discordant pigs, squealing along.



THE END

.7.

Entretien avec Stacey Meadwell, éditrice du blog RevStanBlog.com.

**Questions réalisées par Mylène Maignant, les réponses ont été
obtenues le 19 mai 2022 par e-mail.**

I. About you (you do not necessarily have to answer these questions if you do not want to)

- **Where and when were you born?**
- **Where and what did you study?** I studied English and Psychology at Hope University in Liverpool

II. About your blog

- **When did you start your blog?** 2010
- **Where does the title of your blog come from?**
To cut a long story short, Rev Stan was a nickname that I started using on my personal Twitter account to keep it separate from my work Twitter account. When I set up the blog, I was already Tweeting about theatre as Rev Stan, so it felt like an obvious continuation of that persona.
- **Why did you start writing / Where does your passion for theatre come from?**
I started going to the theatre as a student. I had friends who were studying drama, and I was studying English, and you could get cheap student tickets.
When I moved to London for work, I wasn't earning very much, so didn't think I could afford to go to the theatre. So I got out of the habit. It wasn't until years later that I remembered how much I enjoyed going to the theatre and started dipping my toe in again. It quickly grew into an obsession.
I started writing for a number of reasons. I wanted to keep a record of all the plays I was seeing, plus I enjoyed writing. My day job was a business journalist but writing the blog was different; it was about writing for me and in my own voice.
- **As an independent blogger, what were the challenges you had to face?**
Time is the biggest challenge. Blogging doesn't pay anything, so I have to fit in theatre trips and writing around work. Sometimes I have to sacrifice the blogging to prioritise work because that is what pays the bills ...and buys the theatre tickets.
The other big challenge is how theatres and professional critics view bloggers. Theatre-land has been very slow to embrace digital and social media. And in some quarters, there is still a snobbishness.

II. About English theatre criticism in general

- **Would you say that there is a difference between journalists and bloggers? If so, how would you define it?**

Journalists are paid, bloggers generally aren't. Journalists are writing for publications which are aimed at a particular readership, and that, I'm sure, feeds into how they review. For example, someone writing for the Daily Mail readership vs writing for the The Guardian. Bloggers, I feel, have more freedom.

Then there are the tickets. Bloggers get some complimentary press tickets but also buy theatre tickets. Review tickets are generally good seats, but if you are spending your own money, it might be the cheaper seats. I think that gives bloggers a more realistic experience of theatre-going.

I've seen plays where the set design severely limited how much I could see, or the direction/blocking meant the actors were hidden from view for large periods of time. It affects how you experience a play.

A theatre critic would never be put in seats like that, but theatres are selling those tickets to the public. It's something I reflect in my reviews.

- **Do you feel your reviews are as legitimate as the ones written by paid journalists?**

Yes. Why wouldn't my impression/thoughts about a play be any less valid than someone who is being paid?

- **How would you characterise the London blogosphere? Do you know any other bloggers?**

I've got to know a few bloggers through social media and bumping into each other at press nights. In fact, I've made some great friends. Generally, the London blogosphere is a mostly friendly and supportive community.

- **Do you think independent bloggers have a role to play within the theatre community?**

Definitely, for two key reasons. It allows more diversity of voices reviewing theatre. And, given that people have far more choice of where and how they consume information, it broadens the reach of theatre criticism.

- **Do you think it is more challenging to write as a woman?**

If you mean how am I viewed as a female writer as opposed to a male then I'm not conscious of any discrimination. But then I can only judge based on my own experience – I don't have anything to compare it with.

Entretien avec John Chapman, éditeur du blog 2ndFromBottom.com.

**Questions réalisées par Mylène Maignant, les réponses ont été
obtenues le 21 juillet 2022 par e-mail.**

I. About you (you do not necessarily have to answer these questions if you do not want to)

- **Where and when were you born?** London 1954
- **Where and what did you study?** English Literature and Drama

II. About your blog

Where does the title of your blog come from?

2nd From Bottom started in 2016. It followed my first blog Notes From The Bottom. This covered 18 months of my involvement with the project *A Midsummer Night's Dream: A Play For the Nation* working with The Royal Shakespeare Company and in which I played...Bottom. This reached a natural conclusion in the summer of 2016. After that there was only one title to choose for the follow up. 2nd From Bottom really started to take off when the pandemic began. I decided to specialise in reviewing online theatre in the form of recorded material, live streams or shows constructed through media platforms such as Zoom. Between April 1st 2020 and December 31st 2021 reviews were continuous on 640 unbroken consecutive days. During this time I reviewed 1,000 online shows. Since January this year the site has continued in less frantic form with a mix of live and online reviews. As a "sideline" I maintain a curated list on online theatre opportunities.

- **Why did you start writing / Where does your passion for theatre come from?**
For my first blog I wanted to create a permanent record of a unique opportunity as an archive for myself, the group with which I was working and anyone else interested. I enjoyed writing so much that I decided to start the second blog, but it wasn't really until the pandemic that I found a position for it.
- **As an independent blogger, what were the challenges you had to face?**
Building an audience was hard and maintaining it is time consuming. Choosing and learning to use blogging software (WordPress) was an interesting challenge and a chance to learn new skills. It was (and still is) hard to get the blog onto PR, venue and company contact lists particularly West End commercial theatre

II. About English theatre criticism in general

- **Would you say that there is a difference between journalists and bloggers? If so, how would you define it?**

Yes there is – journalists are generally paid while bloggers are not. Of course many of the latter have found ways of monetizing their sites although personally I have never gone down that route

- **Do you feel your reviews are as legitimate as the ones written by paid journalists?**

I do, but there are definitely areas of the industry which do not. Generally they would still rather take star ratings and pull quotes from the “big names” in criticism such as those employed by national news media or bigger theatre review sites.

- **How would you characterise the London blogosphere? Do you know any other bloggers?**

There’s a definite community out there promoted by sites such as [My Theatre Mates](#). I interact on social media with several other reviewers and bloggers, read their reviews (only after writing my own!) and follow their sites.

- **Do you think independent bloggers have a role to play within the theatre community?**

Yes, our reviews are as legitimate as any others and from what I have observed we are often kinder and more considerate in our commenting trying to find an angle on positive criticism

- **Do you think it is more challenging to write as a woman?**

If there was one aspect I could ban it would be star ratings. I think they are a lazy route to critical reaction and mean that readers often ignore the carefully crafted writing which goes into a review. My refusal to comply means that PR advertising often ignores what I have to say - they seem to want to pepper a poster with 4 or 5 star review tags for a quick fix solution. I hope this doesn’t sound arrogant, but I regard it as their loss

Entretien avec Laura Kressly, éditrice du blog ThePlaysTheThingsUK.com

**Questions réalisées par Mylène Mignant, les réponses ont été
obtenues le 16 juin 2022 par e-mail.**

I. About you (you do not necessarily have to answer these questions if you do not want to)

- **Where and when were you born?** I was born in 1982, in the US.
- **Where and what did you study?** I did my BA in Theatre Performance at Marymount Manhattan College, my MFA in Shakespeare Performance at the University of Exeter, and started my PhD at the Royal Central School of Speech and Drama in 2019. I am due to submit my thesis in 2023.

II. About your blog

Where does the title of your blog come from?

I started my website in 2015. It pays homage to my background and specialism in Shakespeare and signifies my approach to criticism - the most important thing is writing about the work.

- **Why did you start writing / Where does your passion for theatre come from?**
I started reviewing in 2013 for a couple of other sites. At the time I was working full time in secondary education, and felt that I was losing touch with developments in the industry. I thought reviewing would be a good way to put my training and background in theatre to good use, keep up with the field, and see shows for free. I started performing at a young age, first as a musician, but then I joined my elementary school's drama club in 5th grade. That was a lightbulb moment - I decided I wanted to be an actor.
- **As an independent blogger, what were the challenges you had to face?**
Though this has slowly been changing since I started reviewing, there used to be a clear hierarchical distinction between critics and bloggers/unpaid critics. This is still present to some extent, particularly with larger venues and PR firms some of whom only give press comps to newspapers and larger publications. However, this has been breaking down as paid critics lose their jobs and start writing for free, and as bloggers/unpaid critics get paid writing work. There are some dinosaur critics and artists who see unpaid critics/bloggers as less skilled or informed, which is particularly frustrating because there are many of us who often write for free but work in theatre and/or have studied/trained in theatre.

II. About English theatre criticism in general

- **Would you say that there is a difference between journalists and bloggers? If so, how would you define it?**

In the context of theatre, yes - journalists tend to be paid whereas bloggers don't. However, not all theatre journalists are critics, and not all critics are journalists. Journalists and paid critics tend to undergo more rigorous editorial processes and have tighter deadlines. Bloggers write for themselves or for unpaid platforms and tend to have more freedom.

- **Do you feel your reviews are as legitimate as the ones written by paid journalists?**
Yes.

- **How would you characterise the London blogosphere? Do you know any other bloggers?**

Changeable. Because so many critics work for free, they come and go with little consequence. I review more or less depending on my paid workload and demands from my PhD and I suspect this dynamic is similar across other blogs and small sites. People review less or stop completely when they're busy or burnt out. Because I've been reviewing for so long I know a lot of other unpaid critics, and people who used to review for free and no longer do. Press nights can be social occasions where we catch up with each other and meet new unpaid critics.

- **Do you think independent bloggers have a role to play within the theatre community?**

Absolutely, especially as paid arts coverage is increasingly subject to budget cuts.

- **Do you think it is more challenging to write as a woman?**

I don't think I find writing harder than a man does, though due to pervasive misogyny in theatre women critics there are those who consider women's reviews less authoritative. I'm sure these same people would have the same feelings about reviews from other marginalised people, too.

Formulaire de candidature pour être critique sur la plateforme AyoungerTheatre.com



Thank you for your interest in being a blogger for A Younger Theatre. Please fill in the below application form and return it to the Managing Director: sam@ayoungertheatre.com to be considered for the role.

Please note: If you are not interested in this role but want to write a one-off piece for A Younger Theatre, get in touch with the editor directly.

Roles

Role one is to write a blog between 600-800 words to be hosted on A Younger Theatre. Bloggers will work closely with the editor to develop blog ideas and to arrange regular deadlines and they are usually expected to produce a minimum of two blogs per month.

Blogs are usually focused around a set theme. Our most current blog is based on being a theatre-maker on a budget. Previous blogs have been based on topics such as university life, being a first-time producer and the Edinburgh Festival.

Role two is for someone interested in writing one-off opinion pieces and essays between 800-1100 words.

Both roles will involve working on topics relating to theatre and the arts, with an emphasis on the current state of theatre, on latest news, and on the changing arts world. They are usually expected to produce a minimum of two blogs per month.

Both roles are suited to someone opinionated, who can bring a sense of their own personality to topics of wider interest. Being able to write coherently is essential. The aim of the AYT Opinion section is to provide critical comment on theatre; to engage with young people and to be the start of a dialogue with AYT readers.

You can also request to become part of the AYT reviewing team.

Application

Please note: The application is open to aged 35 and under and the position is unpaid. This is due to the nature of the project and to AYT's lack of funding. All details must be given in full. Please return the completed application to sam@ayoungertheatre.com with the title **BLOGGER Application** in the subject line.

Name:
Age:
Date of Birth:
Contact Email:
Contact Number:
Location:
Twitter:

Why do you want to write for A Younger Theatre? (max 250 words)

Tell us about the last show you went to see. What was it? Why did you go? (max 100 words)

Role One: Blogger(if you are applying for Role Two, please write N/A in this section)

What theme would you ideally like your blog to explore and why? (max 250 words)

Please write a proposed blog on a theatre topic or a recent arts/theatre news related item. Remember we are looking for opinionated arguments as well as coherent use of the English language. (600-800 words)

How often would you intend to write your blog? Weekly, bi-monthly or on an ad hoc basis?

Do you have a current blog? If yes, please provide a link for reference.

Blogs tend to be published on set days with the copy being provided approximately a week before for editing purposes. Is there a particular day of the week you would like to see your blog published? Please note: we do publish at the weekend and this can be arranged should it suit the blogger.

Role Two: Opinion/ Essay writer/ (if you are applying for Role One, please write N/A in this section)

Looking at the arts industry, what drives you mad and what would you most like to change? (max 100 words)

What attracts you to this role? (max 100 words)

Why do you see AYT as the platform with which to express your opinion? (max 250 words)

Please write an opinion piece or essay on a theatre topic or a recent arts/theatre news related item. Remember we are looking for opinionated arguments, lots of passion and a clear, unique voice, as well as coherent use of the English language. This may be a piece previously published (800-1100 words)

Please include any other comments you might have below, including whether you will need any extra support.

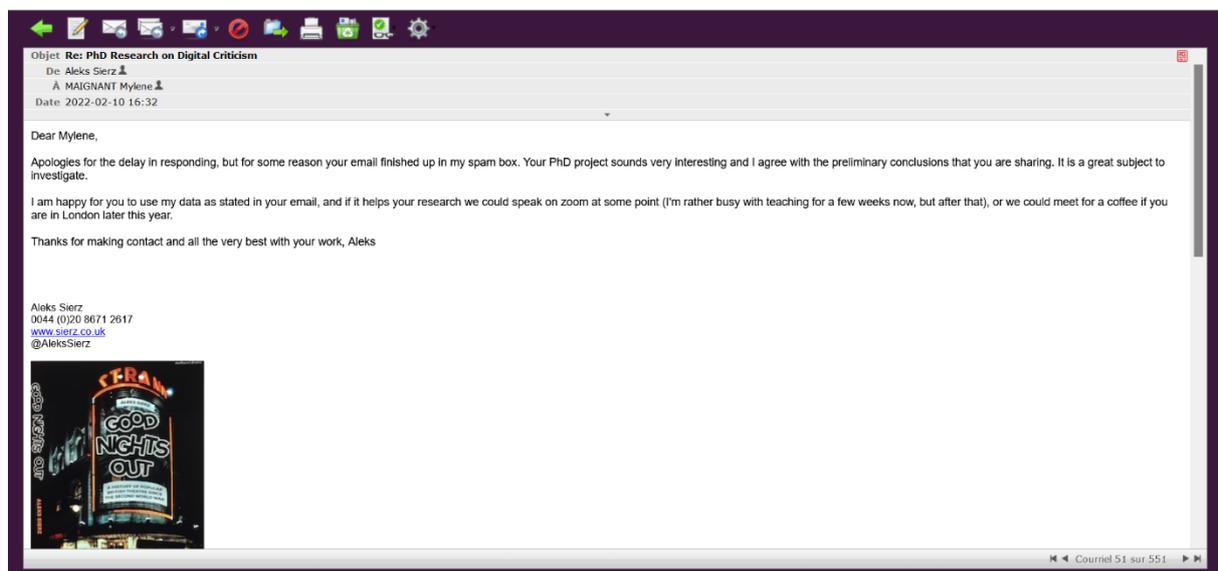
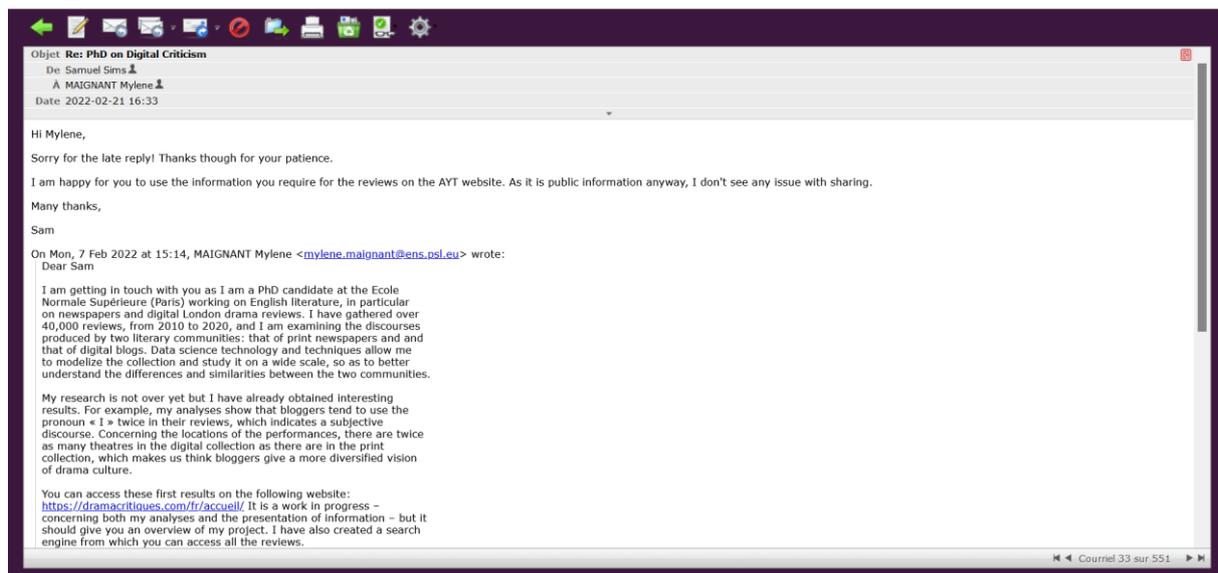
Thanks for filling in the application and for your interest in A Younger Theatre. We aim to get back to you as soon as possible, usually within a week.

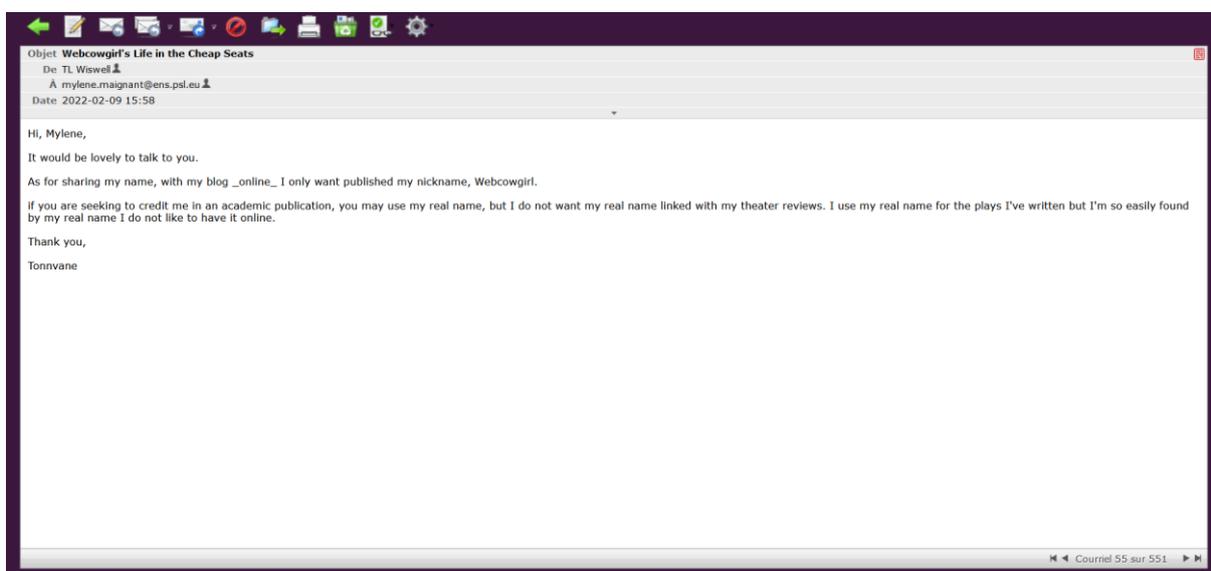
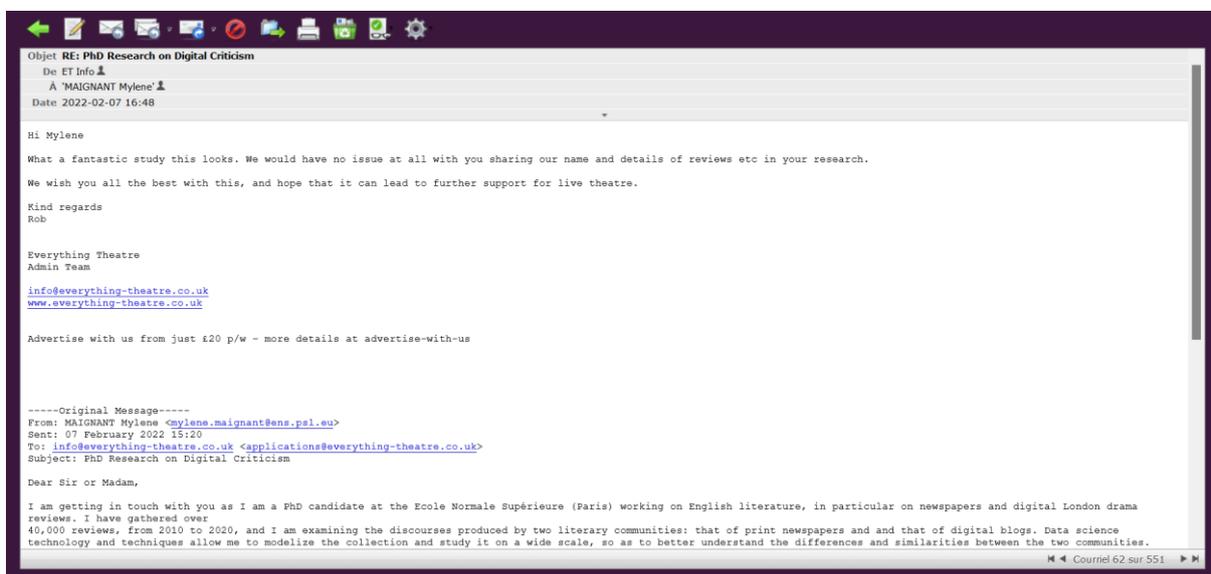
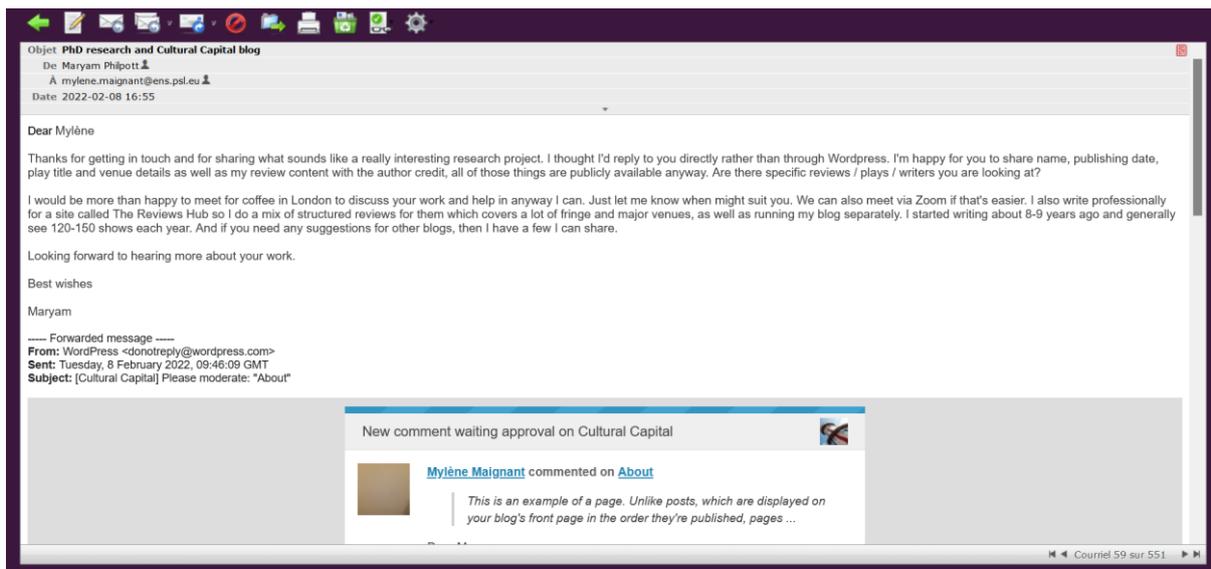
Samuel Sims

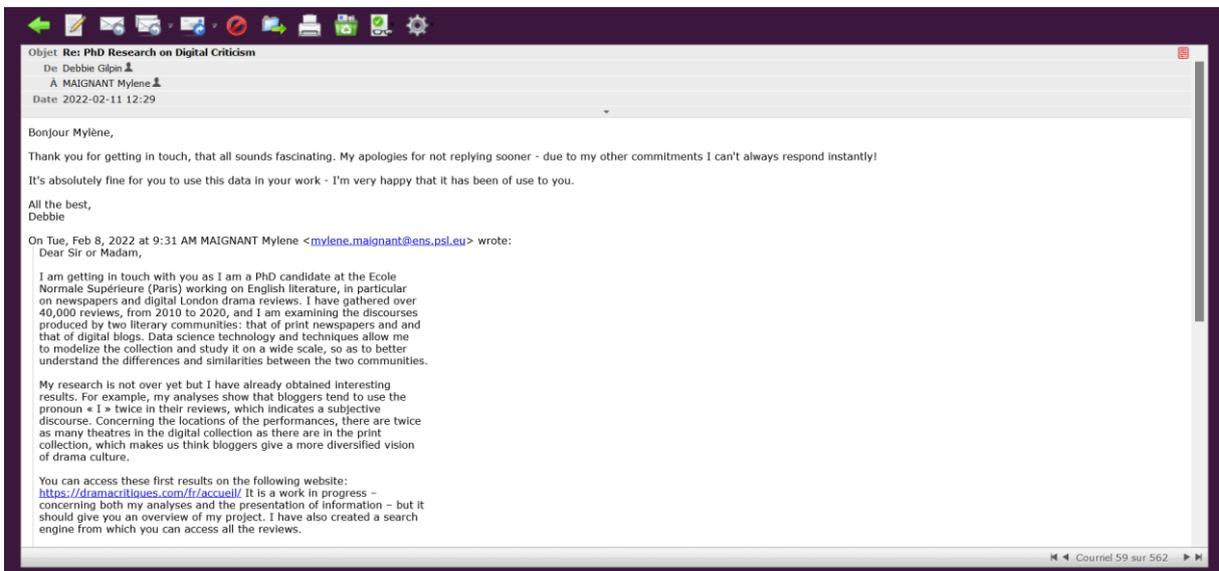
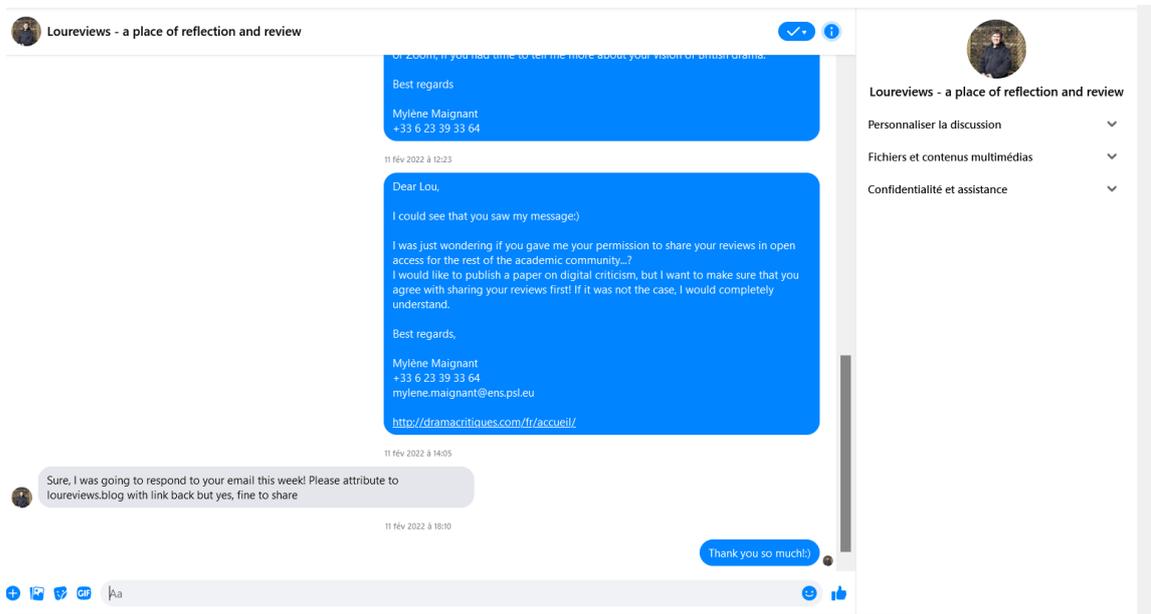
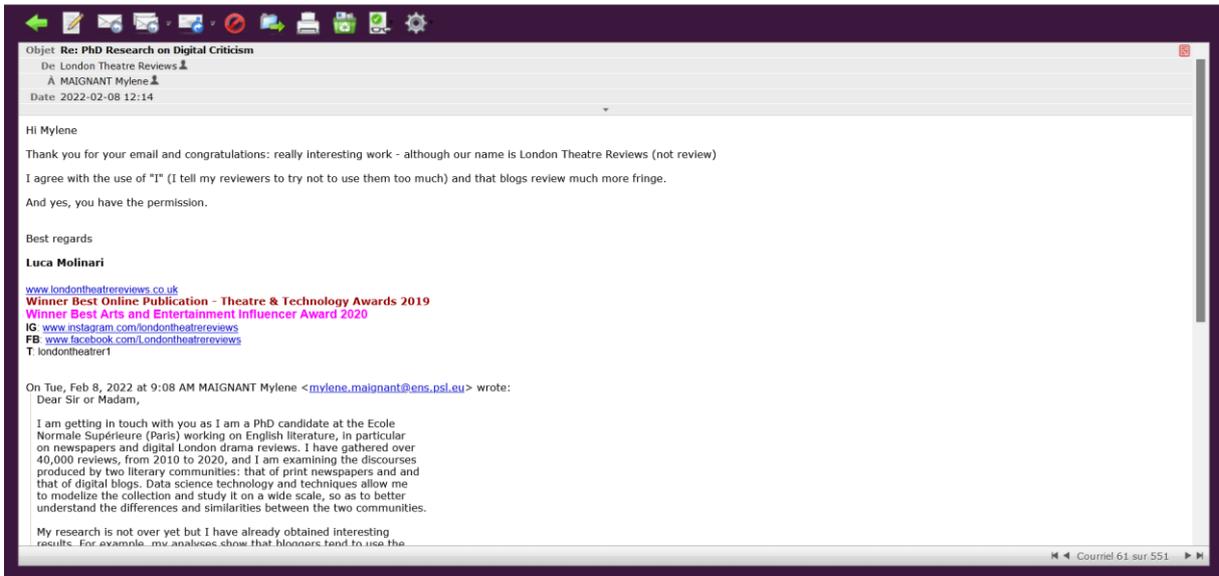
A Younger Theatre

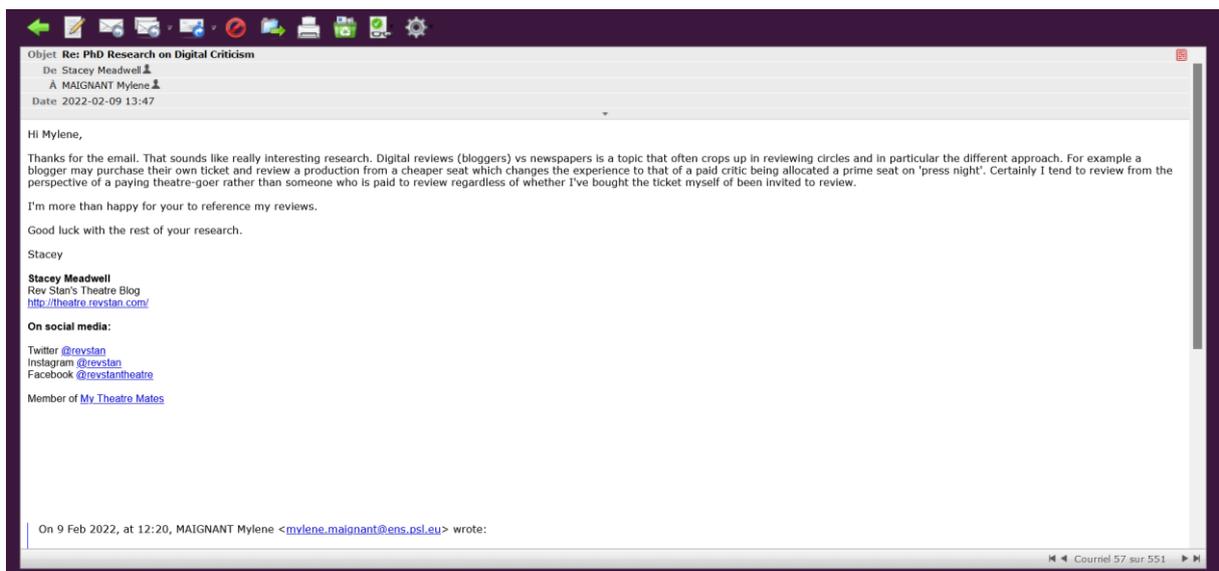
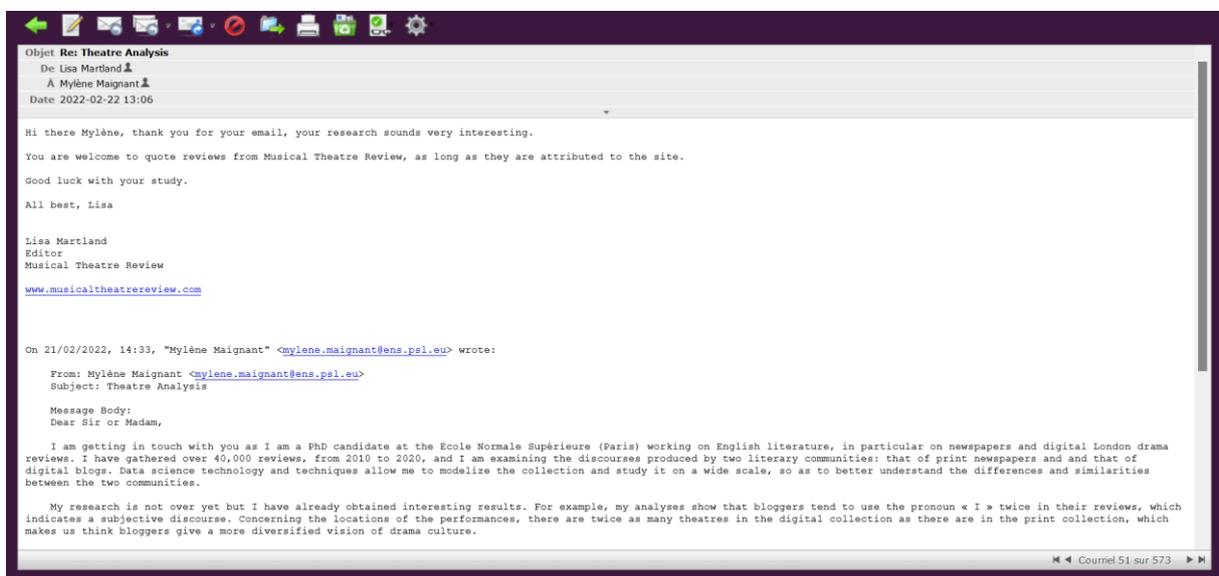
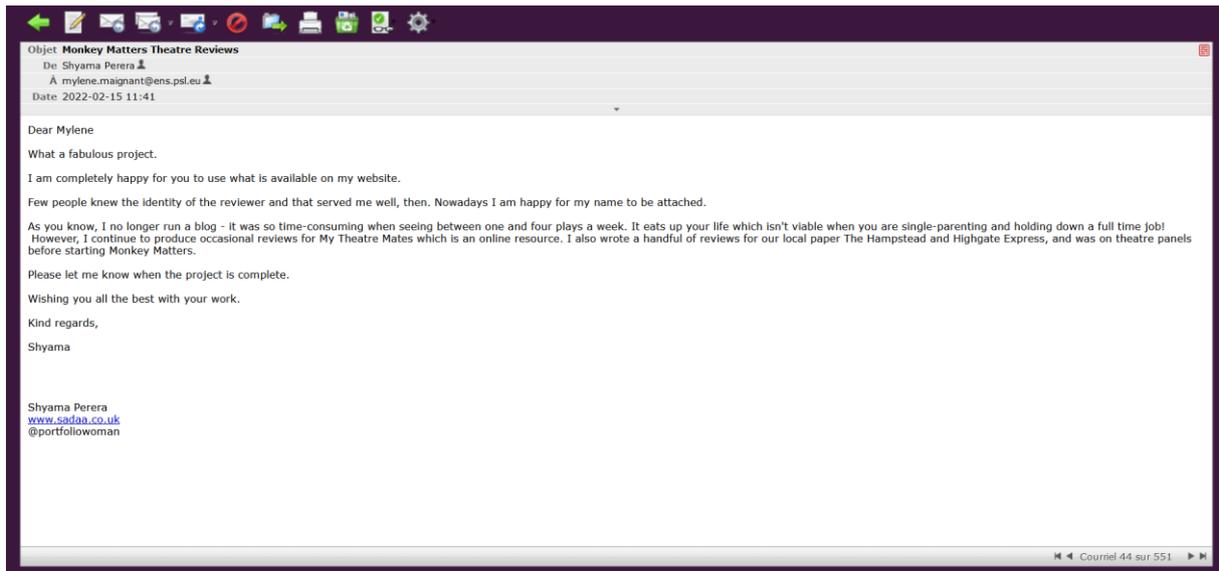
Managing Director

Autorisations attribuées par les blogueurs pour publier leurs critiques en accès libre











Mylène Maignant



Vous avez reçu une réponse à votre InMail

Nous avons crédité votre compte d'un InMail



Rosie Snell

Marketing Officer at Masterclass Trust

Hello Mylène,
 Thank you for getting in touch. **Scatter of Opinion** hasn't been running for the last few years, but if it is useful to your paper I'm more than happy for you to use it.
 Thanks
 Rosie
 -

Répondre



Mylène



Stage Review

the digital reviews. I would like your permission to share your data (name, publishing date, title of the play, the location where it was performed, as well as your review). It would be open access data (not for-profit) and used for academic purposes only.

I am aware data sharing is a sensitive topic nowadays and I am available should you have any concern. I would be very happy to meet you, either in London or via Skype or Zoom, if you had time to tell me more about your vision of British drama.

Best regards

Mylène Maignant
 +33 6 23 39 33 64
 mylene.maignant@ens.psl.eu

15 Fév 2022 à 16:23

Yes, happy to use my data. As a qualified journalist and now blogger I can tell you that the difference between print media & digital is that the former is produced by highly skilled, fully trained reporters and the latter by well intentioned, almost universally untrained amateurs.

Print reviews are written by experts who have either worked in theatre in some way or perhaps studied drama/English/journalism at university. It is their job to cover opening nights and they are salaried professionals.

But, certainly in London and to a lesser extent, in regional theatre, there has been a proliferation of bloggers over the last decade. These are usually the retired or theatre fans who may have day jobs as dentists, bankers, baristas, accountants, plumbers etc and who, let's be honest, blag free theatre tickets in exchange for seats in the stalls and reviews.

Fringe theatre is impoverished. Producers & companies work almost for nothing, usually at a loss, and don't have marketing budgets so they are grateful for any publicity their shows can get.



Stage Review

Personnaliser la discussion

Fichiers et contenus multimédias

Confidentialité et assistance

Facebook, Twitter, LinkedIn, WhatsApp, Messenger



Home Profile 9+ + Mylène

Stage Review

Professional reviewers don't have the time or inclination to attend minor productions especially as they may be staged in a 50-seat theatre above a London pub. Their milieu are the 1,500 seat West End theatres where their reviews are very influential and considerably more theatre-goers attend shows.

But bloggers are more than happy for a free night at any theatre, no matter how small, to feed their passion for the theatre, in exchange for often well-meaning but poorly written reviews.

These bloggers have no formal training in journalism or writing in general but anything is better than nothing.

Very, very rarely you get someone like me who was a print journalist for 37 years and who, on being made redundant, decided to launch her own website/ blog www.stagereview.co.uk (plus Twitter, FB, YouTube & Insta) to continue providing well written, professional coverage of theatre, both fringe, regional & mainstream.

The downside of this is that I'm not paid a salary (other than in free theatre tickets & glasses of gratis cheap interval wine) and travelling around the UK five nights a week, covering fringe, regional and West End productions is both very expensive and exhausting.

After a break due to the covid pandemic, I have cut back on my coverage as I was writing/ reviewing 14-15hrs a day for zero income.

Hope this is of help. There is a wealth of talent on stage in the hundreds of small fringe, Off-West End and regional theatres around Britain but it is impossible for print media to cover them all, it's not cost effective, which is why they concentrate on the big West End productions. The bloggers are far more adaptable, using their free, leisure time to support the industry at grassroots level.

Please feel free to use anything from my website and credit me by name 'Anne Cox, Stage Review' or 'Anne Cox www.stagereview.co.uk'

Personnaliser la discussion
Fichiers et contenus multimédias
Confidentialité et assistance

Messages

Dr Hannah Greenstreet @hg...
Mylène - happy for you...

Ben Kulvichit @B_Kulvichit Feb 21
You sent a link

Alice Saville @RaddingtonB Feb 14
You sent a link

Rewrite This Story @Rewrite... Feb 10
You sent a link

Mind the Blog @Mind_the_Bl... Feb 10
You sent a link

Susan Elkin @SusanElkinJou... Feb 10
Thank you so much Susa...

Aleks Sierz @AleksSierz Feb 9
Hello Aleks!

Susan Elkin
@SusanElkinJourn

I am aware data sharing is a sensitive topic nowadays and I am available should you have any concern. I would be very happy to meet you, either in London or via Skype or Zoom, if you had time to tell me more about your vision of British drama.

Best regards

Mylène Maignant
+33 6 23 39 33 64

Feb 10, 2022, 10:22 AM ✓

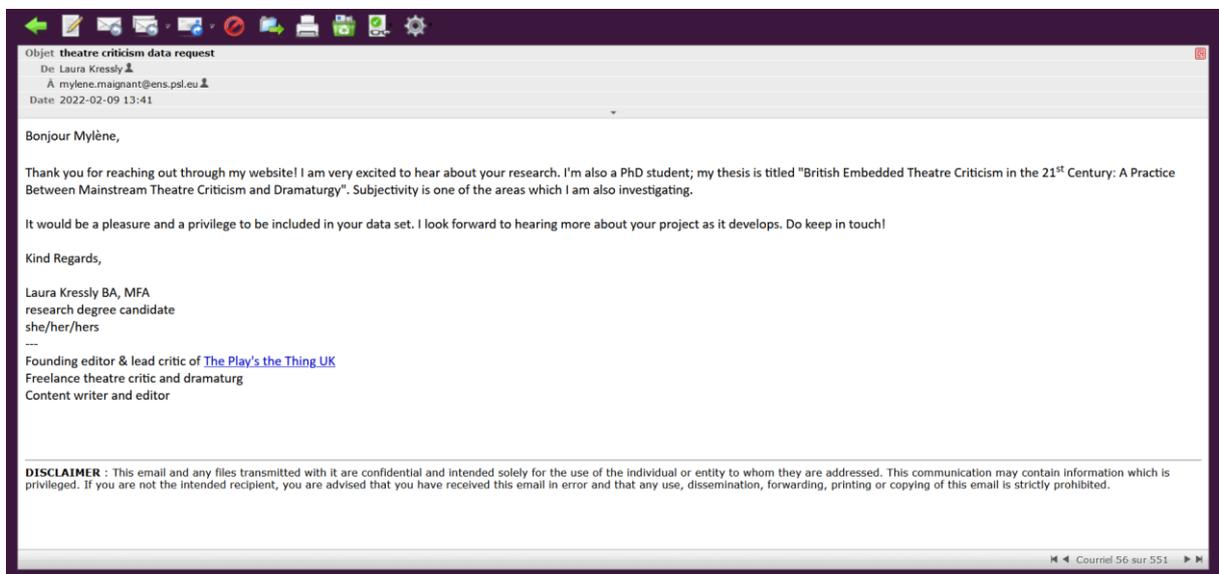
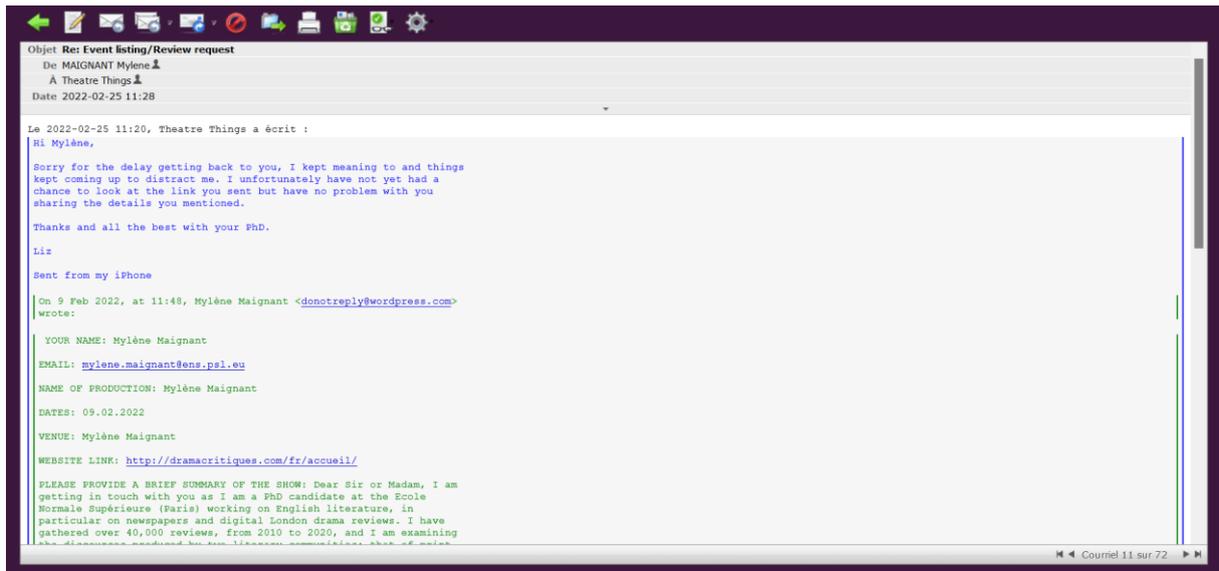
Hello Mylene. Interesting project - well done. Yes you have my permission to use my written material. Best, Susan

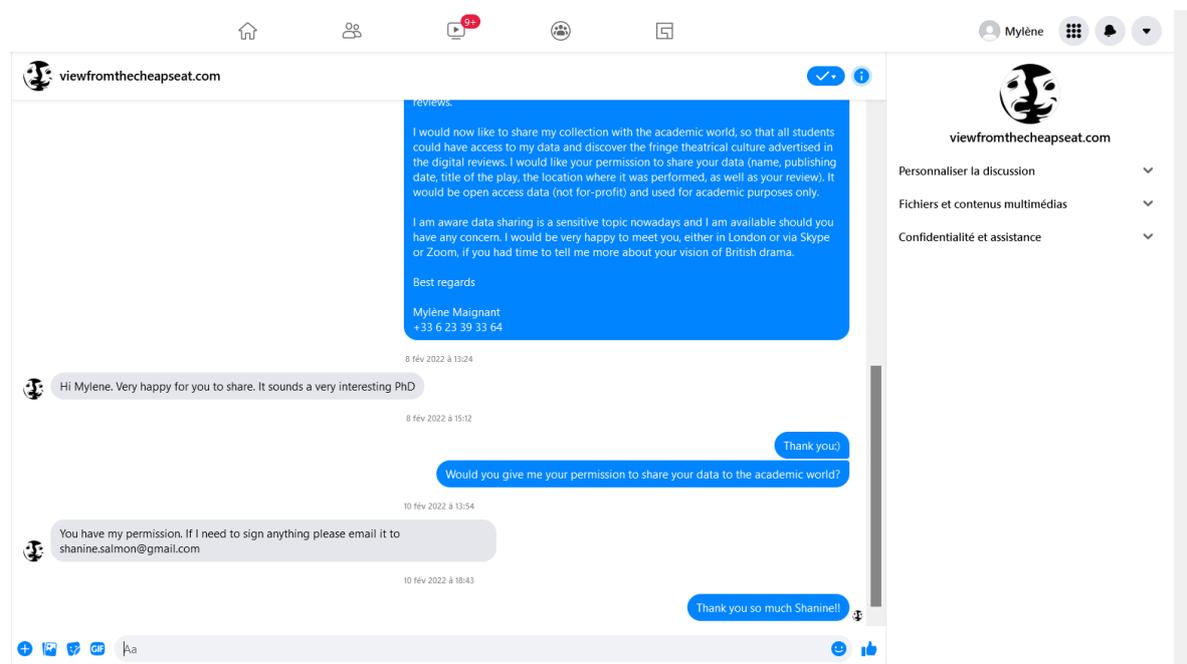
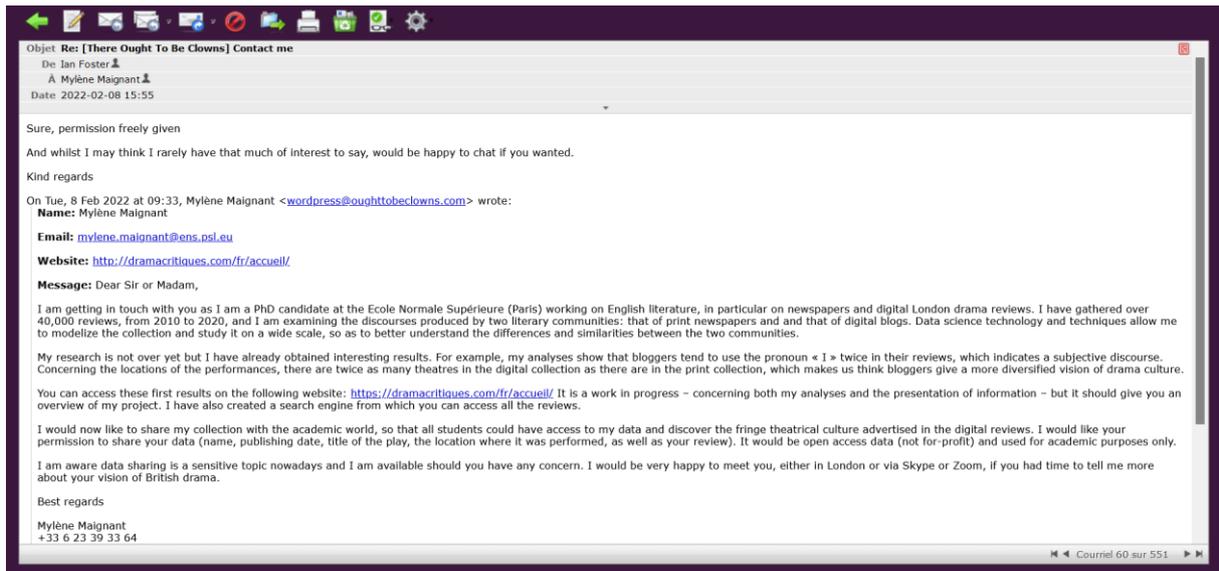
Feb 10, 2022, 10:28 AM

Thank you so much Susan!!

Feb 10, 2022, 10:33 AM ✓

Start a new message





Liste des vingt thématiques disponibles dans les métadonnées de notre corpus

Numéro	Thème
1	Body
2	Childhood
3	Cultural Difference / Race
4	Death
5	Disability
6	Education
7	Environment
8	Family
9	Friendship
10	Identity
11	LGBT / Queer / Sexuality / Gender
12	Love
13	Memory
14	Politics
15	Relationship
16	Religion
17	Science
18	Supernatural
19	Violence
20	Women / Feminism

BIBLIOGRAPHIE

Sources littéraires

Essais et articles universitaires

BAIN David, *Actors and Audience: A Study of Asides and Related Conventions in Greek Drama*, Oxford, Oxford University Press, 1987.

BARDIOT Clarisse, « Arts de la scène et Culture Analytics », *Revue d'Historiographie du Théâtre [En ligne]*, no. 4, 2017. URL : <https://sht.asso.fr/etudes-theatrales-et-humanites-numeriques-une-introduction/>, consulté le 24 janvier 2020.

BAY-CHENG Sarah, 2017. « Digital Historiography and Performance », *Theatre Journal*, vol. 68, no. 4, 2017, p. 507 –27.

BENNETT Susan, *The Role of the Theatre Audience*, Thèse de doctorat, Université de MacMaster, 1988, p. vi.

BENNETT Susan, *Theatre Audiences*, Londres et New York, Routledge, 1990.

BLANCHARD Nelly et MANNAIG Thomas, *Des littératures périphériques*, Rennes, Presses universitaires de Rennes, 2014, p. 11.

BOURDIEU Pierre, *La distinction : critique sociale du jugement*, Paris, Editions de Minuit, 1979.

BROOKE Peter, *The Empty Space, A Book About the Theatre: Deadly, Holy, Rough, Immediate*, New York, Scribner, 1968, p. 9.

BURNARD Lou, « Introduction », *Qu'est-ce que la Text Encoding Initiative ? [En ligne]*, OpenEdition Press, coll. « Encyclopédie numérique », 2015, URL : <http://books.openedition.org/oep/1297>, consulté le 08 mai 2022.

CAPLAN Debra, « Notes from the Frontier: Digital Scholarship and the Future of Theatre Studies », *Theatre Journal*, vol. 67, no. 2, 2015, p. 347 – 59.

CARROLL Noel, *On Criticism*, Londres, Routledge, 2008.

CHEVREL Yves, « Les études de réception », *Précis de littérature comparée*, Paris, Presses Universitaires de France, 1989, p. 177.

CRAIG Hugh et KINNEY Arthur, *Shakespeare, Computers, and the Mystery of Authorship*, Cambridge, Cambridge University Press, 2009.

DACOS Marin, « Manifeste des Digital Humanities », *ThatCamp Paris 2010 [En ligne]*, publié le 26 mars 2011, URL : <http://tcp.hypotheses.org/318>, consulté le 12 janvier 2022.

DASU Tamrapani et JOHNSON Theodore, *Exploratory Data Mining and Data Cleaning*, New Jersey, Wiley, 2003, p. ix.

DAVIS Jim et EMELJANOW Victor, *Reflecting the Audience: London Theatregoing, 1840 – 1880*, Hertfordshire, University of Hertfordshire Press, 2001.

DE CERTEAU Michel, *L'Invention du quotidien. T. 1. Arts de faire*, Paris, Gallimard, 1980, p. 241.

DONOHUE Joseph, « The London Stage 1800-1900: A Data Base for a Calendar of Performances on the Nineteenth-Century London Stage », *Computers and the Humanities*, Vol. 9, no. 4, 1975.

DUPRE Xavier, « Python pour un data scientist / Economiste, Web scraping », *Python dans tous ses états* [En ligne], URL : http://www.xavierdupre.fr/app/ensae_teaching_cs/helpsphinx/notebooks/TD2A_Eco_Web_Scraping.html, consulté le 09 novembre 2021.

FIEVET Cyril, Emily TURRETTINI, *Blog Story*, Paris, Eyrolles, 2004.

FISHER Mark, *How to Write about Theatre*, Londres, Bloomsbury, 2015, p. 12 – 16.

FRESHWATER Helen, *Theatre and Audience*, Basingstoke, Palgrave Macmillan, 2009.

GARDE-TAMINE Joëlle et HUBERT Marie-Claude, *Dictionnaire de critique littéraire*, Paris, Edition Armand Colin, 2002, p. 174.

GITELMAN Lisa, *Raw Data is an Oxymoron*, Cambridge, The MIT Press, 2013.

GRAZIA TURRI Maria, *A Psychoanalytic Perspective on Theories of Spectator-Character and Actor Character Identification in the Theatre*, Thèse de doctorat, Université d'Exeter, 2015.

GROTOWSKI Jerzy, *Towards a Poor Theatre*, Londres, Methuen, 1968, p. 32.

HANSSSEN Jens-Morten, *Ibsen on the German Stage 1876 – 1918, A Quantitative Study*, Narr Francke Attempto Verlag, Tübingen, 2018.

HOLLEDGE Julie, BOLLEN Jonathan, HELLAND Frode et TOMPKINS Joanne Elizabeth, *A Global Doll's House: Ibsen and Distant Visions*, Londres, Palgrave Macmillan, 2016.

HOLMES D. I., « The analysis of literary style – A review », *Journal of the Royal Statistical Society*, vol. 148, no. 4, p. 328.

HOPE Jonathan et WITMORE Michael, « Quantification and the language of later Shakespeare », *Actes des congrès de la Société française Shakespeare [En ligne]*, vol. 31, 2014. URL : <https://journals.openedition.org/shakespeare/2830>, consulté le 11 mai 2021.

JOCKERS Matthew, *Macroanalysis: Digital Methods & Literary History*, Champaign, University of Illinois Press, 2013.

Kenneth TYNAN, *Theatre Writing*, Londres, Nick Hern Books, 2007, p. 119.

KNOWLES Rick, *Reading the Material Theatre*, Cambridge, Cambridge University Press, 2004, p. 17.

MARPEAU Elsa, François-René MARTIN, « RÉCEPTION, *art et littérature* », *Encyclopædia Universalis* [En ligne], URL : <https://www.universalis.fr/encyclopedie/reception-art-et-litterature/>, consulté le 12 septembre 2021.

MCDONALD Ronan, *The Death of the Critic*, Londres, Bloomsbury, 2007.

MILLER Derek, « Average Broadway », *Theatre Journal*, vol. 68, no. 4, 2016.

MITCHELL Ryan, *Web Scraping with Python: Collecting Data from the Modern Web*. Sebastopol, O'Reilly Media, 2015.

MORETTI Franco, *Graphs, Maps, Trees, Abstract Models for Literary History*, Londres, Verso, 2005, p.1

MOUNIER Pierre, *Les humanités numériques, une histoire critique*, Paris, Éditions de la Maison des sciences de l'homme, 2018, p. 24.

OLSEN Dan, « A Quantitative Research Method in Reception Theory in International Theatre: Case Studies: 'Transit', Brasil 'Motherland', Jamaica 'Sandhom', 'FIF' and 'Visit to Turkey', Denmark, 'Vida de Los Perros' and 'El Meón', Bolivia, '5PSA – o Filho', Brazil. », *Caribbean Quarterly* [En ligne], no. 1/2, vol. 53, 2007, p. 181 – 93, URL : <http://www.jstor.org/stable/40654983>, consulté le 13 octobre 2021.

P. MACDONALD Jackson, « Authorship and the Evidence of Stylometrics », *Shakespeare beyond Doubt: Evidence, Argument, Controversy*, Cambridge, Cambridge University Press, 2013, p. 100 – 110.

PAVIS Patrice, « The Conditions of Reception in the Theatre: Psychological and Psychoanalytical Approaches », *New Approaches to Theatre Studies and Performance Analysis: Papers Presented at the Colston Symposium, Bristol, 21 mars 1997*, Berlin, New York, Max Niemeyer Verlag, p. 13 – 30.

POLLOCK Griselda, « Des canons et des guerres culturelles », *Cahiers du genre* [En ligne], vol. 2, no. 47, 2007, URL : <https://www.cairn.info/revue-cahiers-du-genre-2007-2-page-45.htm>, consulté le 29 août 2022.

RADOSAVLJEVIC Duška, *Theatre Criticism: Changing Landscapes*, Londres, Bloomsbury, 2016.

REBELLATO Dan, « Critical Thinking », *Dan Rebellato* [En ligne], URL : [Critical Thinking — Dan Rebellato](https://www.danrebellato.com/critical-thinking/), consulté le 09 mars 2022.

RETTBERG Walker, « Final Version of weblog definition », *Jill/txt* [En ligne], publié le 28 juin 2003, URL : [final version of weblog definition – jill/txt \(jilltxt.net\)](http://jilltxt.net), consulté le 09 novembre 2021.

ROBERT JAUSS Hans, *Pour une esthétique de la réception*, Paris, Gallimard, 1978, p. 259.

ROBERTS-SMITH Jennifer, « Visualizing Theatrical Text: From Watching the Script to the Simulated Environment for Theatre (SET) », *Digital Humanities Quarterly* [En ligne], vol. 7, no. 3, 2013, URL : <http://www.digitalhumanities.org/dhq/vol/7/3/000166/000166.html>, consulté le 02 mai 2021.

ROUSSILLON Marine et SCHUWEY Christophe, « Questionnaire – Écrire l’histoire des spectacles avec des bases de données », *Revue d’Historiographie du Théâtre* [En ligne], no. 5, 2017, URL : <https://sht.asso.fr/les-confidants-du-theatre-classique-essai-denquete-quantitative/>, consulté le 25 janvier 2021.

SAPIRO Gisèle, « LITTÉRATURE - Sociologie de la littérature », *Encyclopædia Universalis* [En ligne], disponible sur : <https://www.universalis.fr/encyclopedie/litterature-sociologie-de-la-litterature/>, consulté le 15 septembre 2021.

SCHREIBMAN Susan, SIEMENS Ray et UNSWORTH John, *A Companion to Digital Humanities*, Oxford, Blackwell, 2004.

SPENCE Paul, « Teatro clásico y humanidades digitales: el cruce entre método, proceso y nuevas tecnologías », *Teatro de Palabras*, no. 7, 2013, p. 9 – 38.

UBERSFELD Anne, « The Pleasure of the Spectator », Trans. Pierre Bouillaguet et Charles Jose, *Modern Drama*, vol. 25, no. 1, 1982, p. 138.

VARELA ESCOBAR Miguel, *Theater as Data*, Michigan, University of Michigan Press, 2021, p. 2.

VAUGHAN Meghan, *Theatre Blogging: The Emergence of a Critical Culture*, Londres, Bloomsbury, 2020.

VICKERS Brian, « Shakespeare and Authorship Studies in the Twenty-First Century », *Shakespeare Quarterly*, vol. 62, no. 1, 2011, p. 106 – 42.

W. ENGLISH John, *Criticizing the Critics*, Londres, Hastings House, 1979.

W. MEISTER Charles, *Dramatic Criticism: A History*, Londres, McFarland and Company, 1985.

WARDLE Irving, *Theatre Criticism*, Londres, Routledge, 1992.

WICKHAM Hadley, « Tidy Data », *Journal of Statistical Software*, vol. 59, no. 10, 2014, p. 4.

Articles dans les journaux et dans les blogs

Andrew et Phil, « Death and the King's Horseman, National Theatre », *The West End Whingers* [En ligne], publié le 8 avril 2009, URL : <https://westendwhingers.wordpress.com/2009/04/08/review-death-and-the-kings-horseman-national-theatre/>, [En ligne], consulté le 22 novembre 2021.

Andrew et Phil, « Resurrection Blues : Well it Made us Cross », *The West End Whingers* [En ligne], publié le 8 mars 2006, URL : <https://westendwhingers.wordpress.com/2006/03/08/resurrection-blues-well-it-made-us-cross/>, consulté le 22 novembre 2021.

BANO Tim, « Review : The End of Longing at Playhouse Theatre », *The Exeunt Magazine* [En ligne], publié le 15 février 2016, URL : <http://exeuntmagazine.com/reviews/review-the-end-of-longing-at-playhouse-theatre/>, consulté le 02 août 2022.

BILLINGTON Michael, « Who needs reviews? », *The Guardian* [En ligne], publié le 17 septembre 2007, URL : <https://www.theguardian.com/stage/theatreblog/2007/sep/17/whoneedsreviews>, consulté le 19 juillet 2022.

GARDNER Lyn, « Are Critics and Bloggers on the Same Side? », *The Guardian* [En ligne], publié le 22 août 2013, URL : <https://www.theguardian.com/stage/theatreblog/2013/aug/22/theatre-edinburghfestival>, consulté le 04 mai 2022.

HEMLEY Matthew, « Danielle Tarento: Unpaid bloggers often lack 'intellectual background' to write theatre reviews », *The Stage* [En ligne], publié le 10 mai 2016, URL : <https://www.thestage.co.uk/news/danielle-tarento-unpaid-bloggers-often-lack-intellectual-background-to-write-theatre-review> consulté le 13 décembre 2021.

HOYLE Ben, « Dead White Men in the Critic's Chair Scorning Work of Women Director », *The Guardian* [En ligne], publié le 14 mai 200, URL : <https://www.thetimes.co.uk/article/dead-white-men-in-the-critics-chair-scorning-work-of-women-directors-6jn7025295h>, consulté le 04 mai 2022.

Karen FRICKE, « Blogging and Theatre Criticism: The Debate Continues », *The Guardian* [En ligne], publié le 28 novembre 2008, URL : <https://www.theguardian.com/stage/theatreblog/2008/nov/28/blogging-critics>, consulté le 04 mai 2022.

RALPH David, « Lemons, Lemons, Lemons, Lemons, Lemons, A Sonnet », *The Exeunt Magazine* [En ligne], publié le 2 décembre 2015, URL : <http://exeuntmagazine.com/reviews/lemons-lemons-lemons-lemons/>, consulté le 02 août 2022.

RAYNER Jay, « Is it Curtains for Critics? », *The Guardian* [En ligne], URL : <https://www.theguardian.com/artanddesign/2008/jul/13/art.comedy>, consulté le 04 mai 2022.

SAVILLE Alice, « Huff, An Illustrated Review », *The Exeunt Magazine* [En ligne], publié le 23 août 2014, URL : <http://exeuntmagazine.com/reviews/huff/>, consulté le 02 août 2022.

SMURTHWAITE Nick, « How Theatre Record's Survival is a Beacon of Hope while Arts Critics are Axed », *The Stage* [En ligne], publié le 23 juin 2017, URL : <https://www.thestage.co.uk/obituaries--archive/archive/how-theatre-records-survival-is-a-beacon-of-hope-while-arts-critics-are-axed>, consulté le 12 janvier 2022.

WICKER Tom, « The Boy who Kicked Pigs », *The Exeunt Magazine* [En ligne], publié le 7 mars 2013, URL : <http://exeuntmagazine.com/reviews/the-boy-who-kicks-pigs/>, consulté le 02 août 2022.

Plateformes en ligne

[About | Pocket Size Theatre](#) [En ligne], consulté le 19 juillet 2022.
<http://beescope.blogspot.com/> [En ligne], consulté le 22 novembre 2021.
<http://exeuntmagazine.com/about-exeunt/> [En ligne], consulté le 19 juillet 2022.
<http://partially-obstructed-view.blogspot.com/p/about-me.html> [En ligne], consulté le 19 juillet 2022.
<http://thearcadesproject.blogspot.com/>, [En ligne], consulté le 22 novembre 2021.
<http://www.doollee.com/index.php> [En ligne], consulté le 24 juin 2022.
<http://www.jaguarshoes.com/> [En ligne], consulté le 22 juin 2022.
<http://www.notexactlybillington.com/> [En ligne], consulté le 02 mai 2021.
<http://www.platformsouthwark.co.uk/about/> [En ligne], consulté le 22 juin 2022.
https://applecartarts.com/about_us.php [En ligne], consulté le 22 juin 2022.
https://blog.feedspot.com/theatre_blogs/ [En ligne], consulté le 19 juillet 2022.
<https://carnstheatrepassion.com/> [En ligne], consulté le 02 août 2022.
<https://chatspalace.com/past-present/> [En ligne], consulté le 22 juin 2022.
<https://criticsofcolour.tumblr.com/> [En ligne], consulté le 11 mai 2021.
<https://greenwichdance.org.uk/about-us/> [En ligne], consulté le 22 juin 2022.
<https://ibsenstage.hf.uio.no/> [En ligne], consulté le 23 mai 2021.
<https://memorekall.com/fr/> consulté le 02 mai 2021.
<https://mytheatremates.com/about-us/authors/> [En ligne], consulté le 24 novembre 2021.
<https://mytheatremates.com/author/michael-davis/> [En ligne], consulté le 19 juillet 2022.
<https://nymt.org.uk/> [En ligne], consulté le 22 juin 2022.
<https://oughttobeclowns.com/about-me/> [En ligne], consulté le 19 juillet 2022.
<https://theatrecat.com/about/> [En ligne], consulté le 19 juillet 2022.
<https://theblogoftheatrethings.com/about/> [En ligne], consulté le 02 août 2022.
<https://theplaysthethinguk.com/about-me/> [En ligne], consulté le 19 juillet 2022.
<https://uglyduck.org.uk/about-us/> [En ligne], consulté le 22 juin 2022.
<https://viewfromthecheapseat.com/> [En ligne], consulté le 23 mai 2021.
<https://visualizingbroadway.com/index.html> [En ligne], consulté le 23 mai 2021.
<https://webcowgirl.wordpress.com/about/> [En ligne], consulté le 19 juillet 2022.
<https://westendwhingers.wordpress.com/> [En ligne], consulté le 22 novembre 2021.
<https://westendwhingers.wordpress.com/zero-stars-hall-of-fame/> [En ligne], consulté le 22 novembre 2021.
<https://www.ausstage.edu.au/pages/browse/> [En ligne], consulté le 23 mai 2021.
<https://www.ayoungentheatre.com/> [En ligne], consulté le 04 juillet 2021.
<https://www.ayoungentheatre.com/about/> [En ligne], consulté le 19 juillet 2022.
<https://www.ayoungentheatre.com/jobs/> [En ligne], consulté le 19 juillet 2022.
<https://www.beckettarchive.org/> [En ligne], consulté le 11 mai 2021.
<https://www.bellhouse.co.uk/> [En ligne], consulté le 22 juin 2022.
<https://www.britishtheatreinfo.com/> [En ligne], consulté le 22 novembre 2021.
<https://www.hackneywicked.co.uk/venues/swan-wharf/> [En ligne], consulté le 22 juin 2022.
<https://www.highburyoperatheatre.com/who-we-are-1/> [En ligne], consulté le 22 juin 2022.
<https://www.ibdb.com/> [En ligne], consulté le 23 mai 2021.
<https://www.lauderdalehouse.org.uk/about-us> [En ligne], consulté le 22 juin 2022.
<https://www.londontheatre1.com/> [En ligne], consulté le 02 août 2022.
<https://www.londontheatre1.com/londontheatrereportv7.pdf> [En ligne], consulté le 12 juin 2022.
<https://www.londontheatredirect.com/> [En ligne], consulté le 02 août 2022.

<https://www.nyt.org.uk/about-us> [En ligne], consulté le 22 juin 2022.

<https://www.playbill.com/vault> [En ligne], consulté le 23 mai 2021.

<https://www.redbridge.gov.uk/leisure-sport-and-the-arts/redbridge-drama-centre/> [En ligne], consulté le 22 juin 2022.

<https://www.scorebig.com/home/top-50-theater-blogs> [En ligne], consulté le 19 juillet 2022.

<https://www.stagereview.co.uk/about/> [En ligne], consulté le 02 août 2022.

<https://www.theatrerecord.com/> [En ligne], consulté le 11 mai 2021.

<https://www.theatrescapes.theaterwissenschaft.uni-muenchen.de/index.html> [En ligne], consulté le 19 mai 2022.

<https://www.towerhamletsarts.org.uk/?cat=30&cid=44376&guide=venues> [En ligne], consulté le 22 juin 2022.

<https://www.vuelio.com/uk/social-media-index/top10-theatre-blogs/> [En ligne], consulté le 02 août 2022.

<https://www.whatsonstage.com/> [En ligne], consulté le 22 novembre 2021.

https://www.whatsonstage.com/west-end-theatre/news/review-round-up-three-kingdoms-divides-the-critics_4378.html [En ligne], consulté le 07 juillet 2022.

<https://www.youtube.com/playlist?list=PLsR2bqLO2Y4hL1WaI23m8Q89QuFMhw-b> [En ligne], consulté le 19 juillet 2022.

[Who Am I? | Rewrite This Story](#) [En ligne], consulté le 19 juillet 2022.

Sources informatiques

Articles universitaires

ALTMAN Naomi, « An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression », *The American Statistician*, vol. 46, 1992 p.175 – 185.

BRANDO Carmen, FRONTINI Francesca, MOREAU Dominic, ROCHE Mathieu et MASSON Éric, « Introduction. Humanités numériques et analyses spatiales : enjeux et perspectives », *Humanités numériques [En ligne]*, no. 3, 2021, publié 01 mai 2021, URL : <http://journals.openedition.org/revuehn/2038>, consulté le 16 mai 2022.

D. WILINSON Marc, et al., « The FAIR Guiding Principles for scientific data management and stewardship », *Sci Data* 3 [En ligne], publié le 15 mars 2016, URL : <https://www.nature.com/articles/sdata201618>, consulté le 21 septembre 2022.

DENEGRE Jean et SALGE François, « Introduction aux systèmes d'information géographique », *Les systèmes d'information géographique*, Paris, Presses Universitaires de France, 2004, p. 5 – 11.

GNEHM Ann-Sophie et CLEMATIDE Simon, « Text Zoning and Classification for Job Advertisements in German, French and English », *Proceedings in the Fourth Workshop on Natural Language Processing and Computational Social Science*, 2020, p. 83 – 93.

GUHO Yufan, KORHONEN Anna et POIBEAU Thierry, « A Weakly-Supervised Approach to Argumentative Zoning of Scientific Documents », *Empirical Methods in Natural Language Processing*, 2011.

HAHN Michael, JURAFSKY Dan et FUTRELL Richard, « Sensitivity as a Complexity Measure for Sequence Classification Tasks », *Transactions of the Association for Computational Linguistics*, no. 9, 2021, p. 891 – 908.

HO Tin Kam, « Random Decision Forests », *Proceedings of 3rd International Conference on Document Analysis and Recognition*, vol. 1, 1995, p. 278 – 282.

LAMPERT Andrew, DALE Robert et PARIS Cécile, « Segmenting Email Message Text into Zones », *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, 2009, p. 919 – 928.

MAAS Andrew, et al., « Learning Word Vectors for Sentiment Analysis », *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 2011, p. 142 – 150.

MAIGNANT Mylène, PELLE Damien, BRISON Gaëtan et POIBEAU Thierry, « DramaCritiques' Database » [En ligne], *Journal of Open Humanities Data*, vol. 8, no. 21,

2022, disponible sur : <https://openhumanitiesdata.metajnl.com/article/10.5334/johd.81/> [Dernier accès le 01/10/22].

PANG Bo et LEE Lillian, « Opinion Mining and Sentiment Analysis », *Foundations and Trends in Information Retrieval*, vol. 2, no. 1 – 2, 2008, p. 1 - 135.

SACHAN Devendra Singh, ZAHEER Manzil et SALAKHUTDINOV Ruslan, « Revisiting LSTM Networks for Semi-Supervised Text Classification via Mixed-Objective Function », *Proceedings of the AAAI Conference on Artificial Intelligence*, no. 33, 2019, p. 6940 – 6948.

SERVIGNE Sylvie, « SYSTÈMES INFORMATIQUES - Conception, architecture et urbanisation des systèmes d'information », *Encyclopædia Universalis*, URL : <https://www.universalis.fr/encyclopedie/systemes-informatiques-conception-architecture-et-urbanisation-des-systemes-d-information/>, consulté le 12 juin 2022.

Simone TEUFEL et Marc MOENS, « Summarizing Scientific Articles: Experiments with Relevance and Rhetorical Status », *Computational Linguistics*, vol. 28, no. 4, 2002, p. 409 – 406.

SPERDUTI Alessandro et STARITA Antonina, « Supervised Neural Networks for the Classification of Structures », *IEEE Transactions on Neural Networks*, vol. 8, no. 1997, p. 714 - 735.

TEUFEL Simone, *Argumentative Zoning: Information Extraction from Scientific Texts*, Thèse de doctorat, Université d'Edimbourg, 1999.

XIE Qizhe, et al. « Unsupervised data augmentation for consistency training », *Proceedings of the 34th Conference on Neural Information Processing Systems*, no. 525, 2020, p. 6256 – 6268.

YANG Zhilin, et al. « XLNet: Generalized Autoregressive Pretraining for Language Understanding », Cite arxiv:1906.08237.

Plateformes en ligne

<https://chrome.google.com/webstore/detail/web-scraper-free-web-scra/jnhgnonknehpejjnehehllkliplmbmhn> [En ligne], consulté le 09 novembre 2021.

<https://docs.python.org/3/library/re.html> [En ligne], consulté le 09 novembre 2021.

https://doranum.fr/enjeux-benefices/principes-fair_10_13143_z7s6-ed26/ [En ligne] consulté le 21 septembre 2022.

<https://dramacritiques.com/fr/base-de-donnees/> [En ligne], consulté le 01 juillet 2022.

<https://dramacritiques.com/fr/telechargements/> [En ligne], consulté le 28 août 2022.

<https://nlp.stanford.edu/IR-book/html/htmledition/tokenization-1.html> [En ligne], consulté le 19 mai 2022.

<https://pdf.abbyy.com/fr/> [En ligne], consulté le 09 novembre 2021.

<https://pypi.org/project/gender-detector/> [En ligne], consulté le 24 juin 2022.

<https://scantailor.org/> [En ligne], consulté le 09 novembre 2021.

<https://spacy.io/> [En ligne], consulté le 09 novembre 2021.

<https://zenodo.org/> [En ligne] consulté le 21 septembre 2022.

<https://zenodo.org/record/6799656#.Y0hSDkzP2so> [En ligne], consulté le 28 août 2022.

<https://zenodo.org/record/6799656#.Y0mXFEzP2so> [En ligne], consulté le 01 juillet 2022.

<https://dramacritiques.com/fr/telechargements/> [En ligne], consulté le 01 juillet 2022.