

Les *topic models* au service de l'histoire d'un genre vidéoludique : vers une représentation non périodique de l'évolution du contenu textuel des jeux de rôle sur ordinateur entre 1992 et 2017

Fanny Barnabé

Laboratoire Méthodes numériques pour les sciences humaines et la société (MNSHS), Epitech. E-mail : fanny.barnabe@epitech.eu

Nicolas Bourgeois

Laboratoire Méthodes numériques pour les sciences humaines et la société (MNSHS), Epitech. E-mail : nicolas.bourgeois@epitech.eu

Résumé : Cet article propose de montrer comment des outils mathématiques encore peu employés par les historiens (*topic models*, classification hiérarchique, carte auto-organisatrice) peuvent être combinés et exploités pour l'étude d'un corpus historique daté, mais hétérogène, afin d'en caractériser les évolutions temporelles. Le fil conducteur de l'étude sera d'examiner l'évolution du vocabulaire employé par un ensemble de 21 jeux vidéo de rôle occidentaux à forte audience, publiés entre 1992 et 2017, pour un total de 17,5 millions de mots. Nous nous efforcerons, au travers de cette analyse, d'apporter un nouvel éclairage à l'histoire canonique du genre et de proposer un modèle alternatif à la périodisation afin de mettre au jour les tensions et influences avec lesquelles chaque jeu particulier doit négocier.

Mots-clés : topic models, lexicométrie, graphes, sciences du jeu, jeu vidéo de rôle sur ordinateur, CRPG, XX^e-XXI^e siècles

CHRONOKeywords: XX^e siècle, XXI^e siècle

Using Topic Models to Study the History of a Video Game Genre: Towards a Non-Periodic Representation of Changes in the Textual Content of Computer Role-Playing Games between 1992 and 2017

Abstract: This paper shows how researchers can combine and exploit mathematical tools still rarely used by historians (topic models, hierarchical classification, self-organizing maps) to study a dated but heterogeneous historical corpus and to characterize its evolution over time. The analysis focuses on changes in the vocabulary used by a set of 21 popular western role-playing video games published between 1992 and 2017 and comprising a total of 17.5 million words. Through this example, we aim to shed new light on the canonical history of the genre and to propose an alternative model to periodization for uncovering the trade-offs and influences underlying the development of each particular game must negotiate with.

Keywords: topic models, lexicometry, graphs, game studies, computer role-playing game (CRPG), twentieth-twenty-first centuries

Cet article entend participer au processus d'historicisation de la culture vidéoludique en analysant l'évolution du jeu vidéo de rôle (*computer role-playing game*, CRPG) occidentale – genre caractérisé par l'importance de son contenu textuel – par l'intermédiaire de son vocabulaire. Précisément,

l'étude poursuivra deux objectifs qui s'articuleront tout au long du texte : l'un d'ordre méthodologique, l'autre théorique.

D'un point de vue méthodologique, nous tâcherons de démontrer l'efficacité de trois outils mathématiques encore peu employés par les historiens comme par les chercheurs en *game studies* (*topic models*¹, classification hiérarchique et carte auto-organisatrice) pour mesurer et représenter les évolutions temporelles d'un corpus textuel caractérisé par son ampleur, sa discontinuité et son incomplétude (voir ci-dessous). En effet, si les outils de traitement automatique du langage sont déjà mobilisés en sciences du jeu, ils sont pour le moment largement réservés à l'analyse des discours métatextuels qui accompagnent les œuvres (critiques journalistiques², commentaires de joueurs³ ou de développeurs⁴, notices encyclopédiques⁵) plutôt qu'à celle des contenus des jeux. De plus, l'histoire du médium est souvent racontée depuis la perspective et les cadres de pensée fournis de manière descendante par l'industrie, servant ses propres intérêts⁶, et pourrait donc bénéficier d'une approche permettant de faire jaillir des résultats de façon inductive, à partir de l'examen des productions culturelles. Développer des méthodes pour caractériser l'évolution des discours portés par les jeux permettra donc, d'une part, d'ouvrir des pistes inédites pour les études historiques de ce champ et, d'autre part, de fournir des outils réutilisables dans le contexte d'autres recherches historiques se basant sur des discours (histoire littéraire, histoire politique, etc.).

D'un point de vue théorique, ensuite, cette recherche tâchera de démontrer la nécessité de développer des modèles alternatifs à la périodisation pour représenter efficacement l'évolution d'un genre vidéoludique. En examinant les mutations du vocabulaire du CRPG de façon diachronique par le biais d'une approche par *topic models* (détaillée ci-dessous), nous montrerons en effet la complexité des mouvements de rupture, reprise et transformation qui traversent l'histoire du genre et la difficulté de représenter ces phénomènes de façon linéaire. Le genre est un outil conceptuel régulièrement mobilisé pour penser l'histoire du médium vidéoludique⁷, or, si Jauss a insisté sur la nécessité de concevoir celui-ci via « des concepts non téléologiques⁸ », comme un « espace d'expérimentation non linéaire⁹ », les modèles décrivant la solidification des conventions génériques de façon diachronique ont parfois le défaut de présenter cette évolution comme un processus unidirectionnel. Ainsi, le modèle en trois étapes développé par Fowler¹⁰ et adapté au jeu vidéo par Arsenault propose de lire l'histoire des genres comme une progression (cyclique) démarrant avec une « œuvre innovatrice » qui pose de nouveaux codes sans obtenir de reconnaissance, suivie d'une « œuvre paradigmatique »

¹ Modélisation de sujet.

² J. P. ZAGAL, N. TOMURO & A. SHEPITSEN, 2011.

³ J. P. ZAGAL & N. TOMURO, 2013.

⁴ L. D. GRACE, 2014.

⁵ Pour un état de l'art détaillé sur ces travaux, voir J. O. RYAN *et al.*, 2015.

⁶ J. O. RYAN *et al.*, 2015.

⁷ Voir, pour exemple, les travaux du Laboratoire universitaire de documentation et d'observation vidéoludiques (LUDOV) sur le *survival horror* et le *first-person shooter* : D. ARSENAULT, 2011 ; C. THERRIEN, 2015 ; B. PERRON, 2018.

⁸ H. R. JAUSS, 1986, p. 58.

⁹ D. ARSENAULT, 2011, p. 163.

¹⁰ A. FOWLER, 1982.

qui fait reconnaître le genre comme tel en articulant les innovations avec des éléments plus conventionnels, puis d'une « œuvre ultime », qui ferme les horizons du genre et oblige ses successeurs à innover à nouveau¹¹. Cette construction théorique offre une grille de lecture particulièrement utile pour la formalisation de grandes tendances historiques. Néanmoins, nous verrons que les jeux d'influence rendus observables grâce à l'approche par *topic models* ne permettent pas de tracer des ruptures ou des héritages aussi clairs entre les titres du corpus.

L'examen de la manière dont le vocabulaire naît, meurt ou se maintient dans les classes de topics va ainsi nous permettre de donner une représentation plus dynamique de l'histoire récente du CRPG et d'affiner notre compréhension des éléments (thématiques, mécaniques, méta-discursifs) qui ont précisément participé au « consensus culturel¹² » consistant à le reconnaître comme un genre vidéoludique à différents moments. Loin de s'organiser en groupes compacts qui formeraient des « sous-genres » cohérents (par exemple : jeux modernes contre rétros, jeux d'action contre *tactical RPG*, etc.) ou qui permettraient de proposer une périodisation nette, les œuvres négocient chacune de façon originale avec toutes les tensions et contradictions internes à l'histoire du genre. La mise au jour de ces tensions représente donc un moyen de compléter les modèles historiques existants par des visualisations dynamiques du temps, incarnées par les différents graphes présentés au fil de l'article.

Pour servir ces deux objectifs, nous avons étudié un corpus de 21 jeux vidéo de rôle en langue anglaise et deux extensions, issus de séries ayant eu un fort impact sur la zone Europe-États-Unis, et dont la publication s'étale de 1992 à 2017. Le corpus se compose précisément de l'intégralité des contenus textuels (diégétiques ou non) des jeux suivants : *Baldurs Gate I et II* (ci-après abrégés *BG1* et *BG2*), *Darkest Dungeon*, *Divinity Original Sin 2* (abrégé *Divinity 2*), *Fallout 1 à 4* ainsi que l'extension *Fallout 4: Creation Club*, *Fallout New Vegas*, *Pillars of Eternity*, *The Elder Scrolls III à V* (abrégés *ES3*, *ES4* et *ES5*) ainsi que l'extension *Skyrim: Creation Club*, *Planescape: Torment*, *Torment: Tides of Numenera* (abrégés *Planescape* et *Tides of Numenera*), *Ultima 7 à 9*, *Wasteland 2* et *The Witcher 2* et 3¹³. Il s'agit d'un corpus original et d'une taille considérable (17,5 millions de mots), pour lequel le recours aux méthodes quantitatives s'impose. Il est cependant loin d'être exhaustif, si l'on songe aux centaines de titres publiés, dont le seul recensement est une gageure¹⁴. C'est là une propriété et une difficulté propres à l'objet de cet article, qui le différencie de nombreux corpus historiques : en raison du nombre de sources et des problèmes d'accessibilité technique, il est particulièrement difficile, à l'heure actuelle, de rassembler un corpus de textes de CRPG qui pourrait prétendre à la complétude, ce qui rend l'échantillonnage nécessaire.

Dans le cadre de l'expérimentation proposée dans cet article, l'échantillonnage du corpus a été guidé par plusieurs critères, répondant aux exigences classiques de l'analyse de discours : homogénéité, diachronicité et contrastivité¹⁵. Ainsi, les textes sélectionnés proviennent d'un même genre

¹¹ D. ARSENAULT, 2011, p. 167.

¹² D. ARSENAULT, 2011, p. 90.

¹³ Ce corpus a pu être constitué grâce au travail inestimable de Damien Hansen et Pierre-Yves Houlmont, chercheurs en traductologie, qui présenteront leur méthodologie de recueil des textes vidéoludiques dans l'article à paraître « A Snapshot into the Possibility of Video Game Machine Translation ».

¹⁴ L'enquête menée par le site *Unleash the Gamer* (URL : <https://unleashthegamer.com/best-rpg-games/#gref>) recense plus de 500 CRPG publiés au cours des quarante dernières années. L'ouvrage collectif dirigé par Felipe Pepe propose des fiches descriptives de plus de 400 jeux parus entre 1975 et 2015 (F. PEPE, 2019).

¹⁵ D. MAYAFFRE, 2002.

vidéoludique, mais traversent son histoire depuis « l'âge d'or¹⁶ » (*Ultima 7 et 8, BG1*) jusqu'aux productions les plus récentes, et regroupe les principales séries qui ont marqué cette histoire¹⁷. Comme pour tout autre secteur culturel, cette notion d'œuvre à fort impact est difficile à définir : nous nous sommes néanmoins appuyés sur la recension historico-critique effectuée par Matt Barton et Shane Stacks¹⁸ ainsi que sur les différents classements proposés par la presse spécialisée et les sites communautaires (les jeux étudiés comptabilisent ainsi des scores *Metacritic* particulièrement importants¹⁹, allant de 81 à 95). Le corpus intègre principalement des séries, mais aussi quelques œuvres isolées, ce qui permet d'opérer une comparaison entre les évolutions lexicales qui sont internes à une série, un studio (par exemple, Bethesda, qui édite *The Elder Scrolls* et *Fallout*) ou un genre thématique (le médiéval fantastique par opposition à la science-fiction), et celles qui caractérisent le genre dans son ensemble. Ce positionnement entraîne que certaines sagas ou certains studios (Bethesda et Interplay Productions, principalement) exercent un poids important sur les résultats (ce qui doit être gardé à l'esprit au moment de leur interprétation), mais permet des observations qui n'auraient pas été possibles si un seul opus avait été sélectionné pour chaque série. Il se prête ainsi particulièrement à la mise au jour de la difficulté de périodiser l'évolution d'un genre vidéoludique.

D'autre part, ce corpus comporte une originalité qui est à la fois un obstacle et un intérêt dans le cadre de cette étude, à savoir sa discontinuité. Loin d'être publiés en suivant une périodicité régulière (comme ce serait le cas des numéros d'un journal, par exemple), les jeux étudiés sont espacés par des intervalles de temps variables. L'avantage de prendre pour objet un corpus ainsi « troué » est que cette propriété le rapproche de la documentation historique, souvent marquée par des manques avec lesquels les historiens doivent composer. Les méthodes développées pour traiter un tel ensemble de textes pourront donc être réemployées sur d'autres corpus présentant des carences similaires. Enfin, notons que le modèle d'analyse que nous proposons ici est précisément pensé pour pouvoir intégrer des augmentations futures : les traitements lexicométriques opérés ci-dessous pourront ainsi être utilisés sur un corpus croissant de textes et, dans ce but, nous fournissons le code des outils employés sur l'espace Nakala associé à ce numéro²⁰.

Dans la suite de l'article, le propos s'organisera en deux temps : le premier exposera les principes méthodologiques qui ont guidé l'analyse et qui ont permis d'aboutir dans une série de graphes représentant les mutations du lexique du CRPG ; le second consistera en un commentaire analytique de trois axes d'évolution (stylistique, méta-discursif et social) dont le repérage a été permis par l'approche par *topic models*. Signalons toutefois que cinq axes avaient été dégagés au départ et que les deux analyses manquantes (sur le gameplay et les composantes des univers), supprimées faute d'espace, restent consultables sur Nakala²¹.

¹⁶ À savoir les années 1990 selon D. SCHULES, J. PETERSON & M. PICARD, 2018, p. 117, et la fin des années 1980-début des années 1990 selon M. BARTON & S. STACKS, 2019.

¹⁷ On déplorera toutefois deux grands absents dans notre corpus, *Mass Effect* et *Diablo*, aux textes desquels nous n'avons pu avoir accès.

¹⁸ M. BARTON & S. STACKS, 2019.

¹⁹ *Metacritic* est un site qui agrège les notes attribuées aux œuvres culturelles dans les médias. Notons toutefois que la série *Ultima* et les *Creation Clubs* n'y bénéficient pas encore d'une évaluation.

²⁰ URL : <https://nakala.fr/u/collections/10.34847/nkl.8ae1v35l>.

²¹ URL : <https://nakala.fr/10.34847/nkl.cf5f3a32>.

1. Principes méthodologiques

Méthodes et outils

Notre travail se base sur la combinaison de cinq outils algorithmiques. Le cœur de l'analyse est constitué par l'utilisation de *topic models* afin d'extraire du corpus des ensembles lexicaux cohérents. Cette extraction est itérée plusieurs fois afin d'en limiter le caractère aléatoire, ce qui génère un nombre important de topics. Par la suite, une classification automatique est appliquée, de façon à ramener la quantité de topics ainsi générés à un plus petit nombre de classes, afin d'en permettre une interprétation qualitative. En utilisant une carte de Kohonen, nous produisons alors une première représentation, statique, de ces classes. Par la suite nous construisons un graphe biparti entre classes et jeux en nous basant sur la prévalence des topics composant les premières dans le vocabulaire des seconds. Cette seconde représentation nous permet alors d'introduire la dimension temporelle en positionnant les sommets (c'est-à-dire les points du graphe) selon la date de production des jeux et la temporalité de leur diégèse. Permettant de mesurer de façon simultanée les cycles de vie des classes tout en nous prémunissant contre l'hétérogénéité des années de production, ce graphe constitue non seulement une visualisation originale, mais également un support privilégié pour l'analyse lexicale qui constitue la dernière partie de l'article.

La simple série des fréquences d'utilisation d'un terme dans un corpus daté est certainement l'exemple le plus populaire d'insertion d'une dimension temporelle dans une étude lexicométrique ; c'est également l'un des moins fiables tant l'équivalence entre un terme et un concept est rarement vérifiée²². Une façon de se prémunir contre cette variabilité du vocabulaire est d'étudier plutôt des ensembles de termes dont l'utilisation est généralement conjointe, par exemple en utilisant des *topic models*²³. L'étude de la prévalence des topics dans le temps permet la constitution de séries temporelles plus pertinentes, car portant sur des ensembles de mots statistiquement reliés, qui, par définition, résistent mieux aux substitutions de termes, aux erreurs de transcription ou encore aux variations partielles de lexique entre auteurs pour exprimer des concepts similaires²⁴. Sans entrer dans le détail de la modélisation mathématique²⁵, nous présentons ici le principe général de l'outil employé (Encadré 1)

Encadré 1. Modélisation de sujets (*topic model*)

L'hypothèse du modèle est que le vocabulaire à disposition des auteurs d'un corpus peut être distribué en un grand nombre de topics (ensemble de mots ayant une forte homogénéité thématique) et que, lors de la rédaction d'un texte spécifique, son auteur ne puisera pas indifféremment dans l'ensemble du lexique mais seulement dans un nombre limité de ces topics. Ainsi, un rapport de session parlementaire mobilisera probablement le vocabulaire du fonctionnement des institutions ainsi que celui de plusieurs idéologies politiques, mais pas celui de la navigation, du tricot ou de l'arithmétique. Les textes étudiés seraient ainsi chacun le résultat d'un processus génératif qui, puisant alternativement dans quelques topics pertinents, élaborerait terme après terme le

²² F. CHATEURAYNAUD & J. DEBAZ, 2010.

²³ Pour une présentation de l'outil destinée aux chercheurs et chercheuses en sciences humaines et sociales, D. BLEI, 2012.

²⁴ X. WANG & A. MCCALLUM, 2006.

²⁵ Pour celle-ci, se référer à D. BLEI, A. NG & M. JORDAN, 2003.

document final. Le fait que des termes aient une distribution statistique cohérente s'interprète dans ce cadre comme la traduction de leur appartenance commune à un même topic.

Le rôle du modèle est de ne pas définir les topics *a priori*, mais de tenter de les découvrir en inversant ce processus génératif. Puisqu'il est impossible de connaître le mécanisme réel de conception, nous cherchons à le modéliser par un processus aléatoire dont les paramètres nous permettraient d'obtenir une distribution statistique des termes aussi proche que possible de celle observée. Ainsi, si nous examinons les discours électoraux, nous pouvons rejeter comme absolument improbable des milliards de segmentations (par exemple la séparation entre un topic regroupant tous les verbes et un topic contenant tous les adjectifs n'a aucune chance d'aboutir à la génération du corpus), alors que certaines, par exemple un topic sur la nation, un sur l'écologie et un sur la lutte des classes, sont statistiquement beaucoup plus en phase avec le résultat observé. À l'issue d'une itération, également appelé *run*, nous nous retrouvons donc avec un ensemble de topics, c'est-à-dire d'un ensemble non étiqueté de mots que l'algorithme estime être étroitement associés.

Lorsque nous effectuons une analyse par *topic models*, nous faisons face à deux difficultés. La première est qu'il est difficile de savoir *a priori* quel nombre de topics est pertinent. Un nombre de topics trop faible risque de regrouper artificiellement des termes appartenant à des champs différents et de ne pas percevoir certains topics ayant des profils de répartition singuliers entre les différents jeux. Un nombre trop élevé va à l'inverse forcer la segmentation d'ensembles cohérents. La seconde difficulté vient du caractère aléatoire de l'algorithme : il est tout à fait possible qu'une itération donnée mette l'accent sur une segmentation mais en laisse de côté une autre qui aurait pu être tout aussi intéressante²⁶. Il existe des indicateurs mathématiques simples comme la perplexité, mais dont l'efficacité empirique n'est pas garantie²⁷.

Pour y remédier, nous effectuons plusieurs *runs* successives (7) de l'algorithme, avec un nombre élevé de topics (100), et regroupons ensuite les topics en classes, que ce soit selon une base horizontale (plusieurs topics d'une même *run* dont on pense la segmentation arbitraire) ou longitudinale (plusieurs topics provenant de *runs* successives)²⁸. Pour créer ces classes, il convient de choisir, d'une part, une mesure de similarité et, d'autre part, un algorithme de classification. Ici, la fonction de similarité choisie entre deux topics est le nombre de mots qu'ils ont en commun parmi les 50 plus représentatifs, et l'algorithme employé est la classification hiérarchique ascendante, qui, à chaque étape, agrège les deux classes les plus proches jusqu'à l'atteinte d'un certain seuil, fixé à une distance intra-classe maximale inférieure à 8²⁹. Autrement dit, notre algorithme s'interrompt dès qu'il ne peut plus fusionner de classes sans qu'il existe dans la classe résultante au moins deux topics ayant moins de 8 mots en commun parmi les 50 mots les plus fréquents. Nous avons décidé de conserver les classes robustes, c'est-à-dire présentes sur au moins 4 runs sur 7, classes que nous avons (manuellement

²⁶ M. V. MANTYLA, M. CLAES & U. FAROOQ, 2018.

²⁷ J. CHANG *et al.*, 2009.

²⁸ Notons que, dans le cas où on aurait confiance dans la robustesse du résultat, il serait possible d'effectuer cette démarche de classification à partir d'une seule *run*. Pour cette étude, il nous a semblé *a posteriori* que les classes obtenues à partir de plusieurs *runs* étaient plus cohérentes. Les deux séries de graphes peuvent être comparées (URL : <https://nakala.fr/10.34847/nkl.b1ecnu5>).

²⁹ N. BOURGEOIS *et al.*, 2015.

cette fois) regroupées en super-classes en fonction des champs lexicaux qu'elles convoquent, de façon à faciliter la lisibilité.

Une fois ces classes obtenues, la question de leur représentation demeure. En raison de la grande dimension, l'utilisation d'une méthode linéaire comme l'analyse factorielle s'avère insuffisante, car les premiers axes ne capturent qu'une faible partie de l'information : en l'occurrence les deux premiers axes de l'analyse en composantes principales (ACP) ne représentent en cumulé que 32,6 % de la variance³⁰. Nous utilisons plutôt une carte de Kohonen (Encadré 2), afin d'éviter de générer des proximités apparentes qui masquent des distances importantes selon les axes ultérieurs.

Encadré 2. Carte de Kohonen (*self-organising map*)

Une carte de Kohonen (ou *self organising map*, SOM) est une méthode de représentation sur une grille, généralement bi-dimensionnelle, de données prises dans un espace de large dimension³¹. Elle fonctionne par adaptations successives de la grille aux données, chaque adaptation fonctionnant en deux temps : d'abord l'identification du neurone (la case de la grille) le plus proche d'une donnée arbitraire, puis le déplacement de ce neurone et des neurones voisins en direction de ces données. Son principal intérêt est de préserver la proximité immédiate, c'est-à-dire que des classes qui sont placées dans la même case ou dans des cases contiguës sont effectivement proches dans l'espace initial eu égard à la mesure de similarité choisie, et cela même en très grande dimension. Elle nous permet ici de mieux appréhender la proximité entre les classes constituées et de confronter les étiquetages réalisés à l'étape précédente.

Il serait tentant d'exploiter les classes de topics comme des indicateurs dont nous pourrions étudier la prévalence selon les années (*topic intensity*) avec plus de pertinence que de simples ensembles de termes constitués *ad hoc*. Cette approche peut être très fructueuse lorsqu'on dispose d'un grand nombre de textes distribués de façon homogène sur l'ensemble de la période considérée, par exemple des articles de journaux³². Hélas, avec un corpus formé de quelques très grands textes publiés selon un rythme discontinu, elle se révèle peu lisible.

Pour pallier cette difficulté, nous avons recours à une autre forme de représentation. Un graphe biparti est composé de deux ensembles de sommets (ici, respectivement les 21 jeux et les 64 classes) et d'un ensemble d'arêtes, matérialisant chacune une relation entre un élément du premier ensemble et un élément du second (ici, la contribution importante d'un jeu au vocabulaire d'une classe). Si la construction d'un tel graphe est simple, sa présentation l'est moins et est généralement déléguée à un algorithme qui vise à optimiser certains critères, comme la proximité des sous-graphes fortement connectés ou le croisement des arêtes, positionnement qui a le défaut d'être achronique. Ici, nous faisons un choix différent, qui est de forcer la position des sommets en fonction de critères temporels, afin de mesurer l'évolution du vocabulaire dans le temps. En choisissant de fixer horizontalement les jeux sur un axe temporel associé à leur publication et en positionnant les classes de

³⁰ On trouvera sur l'espace Nakala une visualisation et un commentaire de cette ACP.

³¹ Pour une définition mathématique des SOM nous renvoyons à T. KOHONEN, 1982 ; M. COTTRELL *et al.*, 2018.

³² X. WANG & A. MCCALLUM, 2006.

topics au barycentre pondéré par la prévalence de la classe dans le jeu, nous sommes amenés à placer chaque classe au niveau de l'année moyenne de son utilisation.

Lecture des ruptures et continuités

Le graphique de la Figure 1 permet de suivre le déroulé de la classification hiérarchique ascendante. À chaque nouvelle étape, deux classes sont fusionnées et les indicateurs mis à jour. Après 486 étapes (chiffre lu en abscisse), nous nous retrouvons avec une classification en 213 classes, dont la plus grande contient 41 topics, de sorte que chaque paire de topics à l'intérieur d'une classe a au moins $50 - 42 = 8$ mots en commun. À ce moment, nous avons 64 classes robustes, c'est-à-dire présentes sur une majorité de *runs*, ce qui nous situe sur le plateau optimal. En effet, en-deçà de 475 étapes nous avons moins de 60 classes robustes, car il reste des topics similaires présents sur plusieurs *runs* qui n'ont pas encore été assimilés ; à l'inverse, au-delà de 620 étapes le nombre de classes s'effondre mécaniquement et avec lui le nombre de classes robustes en particulier.

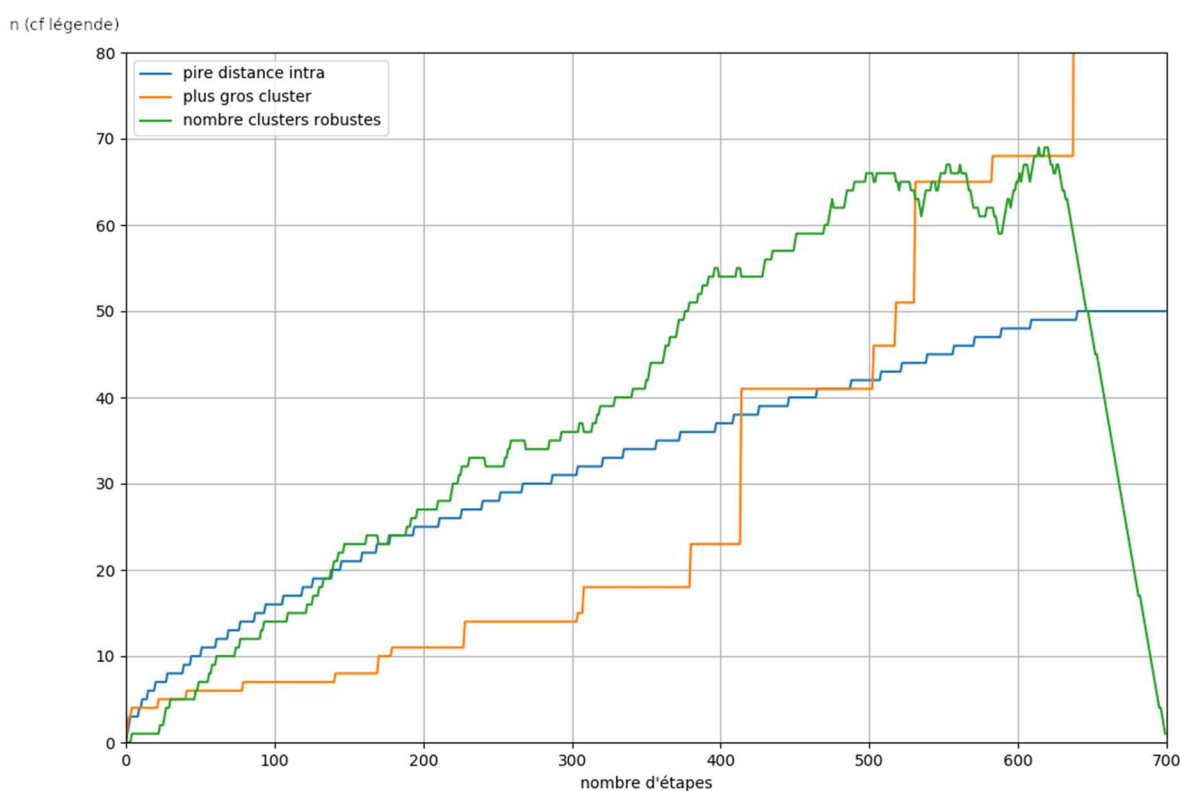


Figure 1. Indicateurs caractérisant la classification hiérarchique

Note. Le seuil d'arrêt est atteint après 486 étapes.

Les Tableaux 1 et 2 ci-dessous listent respectivement les 11 super-classes définies manuellement et les 64 classes initiales de topics à partir desquelles elles sont générées, avec une tentative d'étiquetage basée sur leur vocabulaire. Nous avons autorisé deux classes à figurer à l'intersection de deux super-classes plutôt que de les découper *a posteriori*, car nous avons constaté qu'elles étaient fusionnées à chaque fois assez tôt. La proximité entre les classes et entre les super-classes est figurée par la carte de Kohonen (Figure 2).

Super-classes		
Champ lexical mobilisé	Classes retenues	Exemples d'extraits où la super-classe est fortement mobilisée (en gras, les occurrences des topics qui la composent)

Méta-discours (interface utilisateur, instructions)	104, 119, 124, 165	Drag heroes into slots at the bottom of the screen to form your party and default party order. (<i>Darkest Dungeon</i>)
Vocabulaire de dialogues	3, 9, 48 , 62, 66, 138, 166	Man , I could learn to like you. You're either an incredible ass-hole , or you got guts . Either way is fine with Ton, as long as you don't jam me up. You want to join up? (<i>Fallout 1</i>)
Vocabulaire générique	12, 90	Some kind of gas in the lab. We broke in, it got out. Hit the vents, went everywhere . Everyone just snapped. Locke and Harrald drew on me. Shot them dead. I think . (<i>Fallout 4</i>)
Corps, Sens, Perception	7, 11, 17, 23, 75, 131	Get a good look at the body parts ? I saw ... a head , bobbing - eyes bulging , the tongue blue and popped out. (<i>The Witcher 3</i>)
Combat (capacités, dommages, soins, sorts)	5, 67, 76, 105, 120, 130, 134, 191, 194	Upon entering combat , the wielder will immediately go berserk , killing everything within reach until either calming down or falling unconscious . A very powerful sword , but one must decide whether or not it is worth the risk. (<i>BG 1</i>)
Armes et équipements	28, 41, 42, 55, 126	Most Energy Weapons fall into one of two categories: laser , which is fast, accurate , and low- damage , and plasma , which is slow-moving and very high- damage . (<i>Fallout New Vegas</i>)
Composantes de l'univers (lieux, matières, plantes, bestiaire)	0, 6, 21, 47, 59 , 74, 83, 88, 144	I know the alleys and streets of Sentinel intimately from decades worth of ambling. I know which bridges creak, which buildings cast long irregular shadows, the intervals at which the native birds begin the ululations of their evening songs. (<i>ES 4</i>)
Quête (mission, récompense, chasse, verbes liés au devoir ou à l'investigation)	14, 40, 46, 48	As Geralt searched the burned village for clues , he could not help but notice a pervading scent of sulfur. Finding no other traces of the beast, however, Geralt noted his suspicion and set off for the dwarven catacombs to pursue his other lead. (<i>The Witcher 2</i>)
Société et collectivités (famille, religion, politique, guerre, guildes, peuples)	1, 2, 44, 45, 59 , 61, 65, 72, 79, 146, 178	My most august and wise friends, members of the Elder Council , I am but a provincial queen , and I can only assume to bring to issue what you yourselves must have already pondered . (<i>ES 3</i>)
Cosmogonie, Histoire (destinée, vie, mort, passé)	8, 15, 18, 92, 182	During this conflict the forces of the Northern Realms perpetrated unpardonable atrocities , for which reprisals came first with the Battle of the Marnadal Valley and then with the Cintrian Incident . (<i>The Witcher 3</i>)
Talents, mini-jeux (crochetage, piratage)	22, 39, 96, 98	Collecting seed samples... Successfully collected 5 packets of genetically modified corn seeds. Depositing packets to collection container ... Success . (<i>Fallout 2</i>)

Tableau 1. Regroupement des 64 classes en 11 super-classes faisant référence à des thèmes communs

Indice	Étiquetage des classes robustes
0	Agriculture, plantes, bâtiments
1	Noblesse, Cour, contrat
2	Religion, politique (militaire)
3	Vocabulaire de dialogue (informel) en contexte post-apocalyptique
5	Soin, maladie, poison
6	Cadavres, squelette
7	Vue, description d'objets

8	Histoire (temporalité, possibilité, titres et statuts, batailles, livres) (<i>Witcher</i>)
9	Vocabulaire de dialogue (dialectal) : interjections, directions, actions
11	Parties du corps
12	Vocabulaire générique (verbes) : observations, impressions, opinions
14	Quête, chasse au monstre (<i>Witcher</i>)
15	Lore, récit historique <i>fantasy</i>
17	Apparitions, phénomènes étranges
18	Destinée (vie, mort, temps, pouvoir)
21	Minage, métaux
22	Données, système informatique, ordinateur
23	Perception, esprit, pensées
28	Classes et équipements de personnage
39	Crochetage, pièges
40	Quête, mission, récompense
41	Équipement, armes
42	Armes magiques (armes et éléments naturels)
44	Famille, foyer
45	Famille, religion
46	Devoir, nécessité, service
47	Craft, artisanat, matériaux
48	Vocabulaire de dialogue (informel) et vocabulaire de quête
55	Équipement médiéval
59	Bâtiments officiels, politique et religion
61	Assassinat, vengeance, politique
62	Vocabulaire de dialogue (neutre) : verbes génériques, formules de politesse, récompenses
65	Guilde, politique médiévale
66	Salutations, interpellations
67	Système de combat, statistiques, capacités
72	Guerre contemporaine (<i>Fallout</i>)
74	Bâtiments
75	Descriptions de réactions humaines (corps, expressions, émotions)
76	État de santé (mort, vie)
79	Groupes, meute, patrouille (<i>Fallout</i>)
83	Stockage, rangement, conteneur (post-apocalyptique)
88	Commerce citadin
90	Verbes de base, vocabulaire générique
92	Destin, bien et mal
96	Piratage informatique, ordinateurs, infrastructures
98	Bâtiments, sécurité (entrer, ouvrir, portes, crochetage)
104	Interface utilisateur, instructions
105	Mage, électricité, invocation
119	Tutoriel, interface utilisateur, instructions
120	Statistiques du personnage, compétences
124	Inventaire et progression
126	Armes (science-fiction)
130	Combat, logistique, corps (science-fiction)
131	Corps et esprit (mémoire, conscience, pensées, perception)
134	Magie en combat
138	Vocabulaire de dialogue : questions, salutations, communication
144	Lieux, déplacements, voyage, carte
146	Religion, politique, peuples (<i>Divinity</i>)
165	Interface utilisateur

- 166 Dialogues formels, archaïques
- 178 Bataille médiévale, quête (*Witcher*)
- 182 Monde, vie, destinée, métaphysique
- 191 Portée, vision, précision (attaques à distance), équipement moderne
- 194 Sorts en combat, magie, stratégie

Tableau 2. Liste des 64 classes robustes de topics et interprétation de leur vocabulaire par les auteur-e-s

Note. Les indices s'entendent par rapport aux 213 classes initialement générées. Les noms des classes ont été attribués par les auteurs.



Figure 2. Carte de Kohonen illustrant la proximité entre les classes

Note. Cette proximité s'entend par rapport à la mesure de similarité entre topics définie plus haut. Les classes en gras sont celles qui sont robustes. Les couleurs correspondent aux super-classes majoritaires sur la case (en cas d'égalité, les deux sont figurées).

Nous constatons que la carte isole en haut à gauche un noyau très connecté de classes robustes, tandis que le gros des classes volages est renvoyé en bas à droite. De surcroît, les super-classes qui ont été générées selon des critères purement sémantiques s'avèrent positionnées de façon très cohérente par l'algorithme. On voit en particulier que certains espaces lexicaux (le vocabulaire des dialogues, en jaune, celui des quêtes, en saumon, celui sur la famille et la société, en orange, celui sur l'histoire et la cosmogonie, en beige, et les verbes et termes génériques, en olive) constituent un bloc homogène central dans le coin supérieur gauche, sur lequel viennent s'accrocher deux blocs secondaires : en haut à droite celui centré sur le corps et les perceptions (vert), en bas à gauche celui qui regroupe les talents, les mini-jeux, les compétences (bleu foncé), la description de l'univers (rose clair), et le vocabulaire extra-diégétique (rose foncé). Enfin, la super-classe du combat (bleu clair) et celle de l'équipement (violet) sont non seulement repoussées à la périphérie, mais également éclatées par l'algorithme : en haut à droite, du côté du corps, sont renvoyés tous les topics d'orientation majoritairement contemporaine ou futuriste ; en bas à gauche, tous ceux d'inspiration médiévale-fantastique.

Répartition des topics dans le temps (réel et diégétique)

Nous pouvons désormais composer notre graphe (Figure 3), avec un premier ensemble de sommets représentant les classes, identifiées par leur numéro (en rouge), et un second représentant les différents jeux de notre corpus (en bleu). Les jeux ont été placés à l'intersection de leur année de commercialisation (en abscisse) et d'un indice arbitraire³³ d'archaïsme/futurisme relatif aux éléments de civilisation mentionnés dans leur vocabulaire diégétique (en ordonnée), tels qu'identifiés à partir des topics sur l'équipement³⁴. Les différentes classes ont quant à elles été positionnées au barycentre pondéré des jeux dont elles fournissent une contribution significative au vocabulaire – c'est-à-dire que la distance entre une classe et un jeu est d'autant plus faible que le vocabulaire de ce jeu est fortement représenté dans la classe. Enfin, une arête entre une classe et un jeu est ajoutée si le jeu contribue à lui seul à au moins 10 % des formes attribuées à l'ensemble des topics de la classe.

³³ Étant donné l'absence d'une unité de mesure objectivée qui permettrait de « dater » les périodes fictionnelles représentées.

³⁴ Cet indice utilisé pour le second axe a été constitué selon une formule arbitraire définie par les auteurs afin d'aérer la présentation qui, sinon, ne reposerait que sur un axe (la chronologie de publication). Cette formule consiste à compter les mentions d'éléments de vocabulaire référant à des technologies matérielles médiévales (*sword, bow...*), modernes (*paper, pistol...*), contemporaines (*plane, newspaper...*) ou futuristes (*laser, spaceship...*) et à comparer leur importance respective.

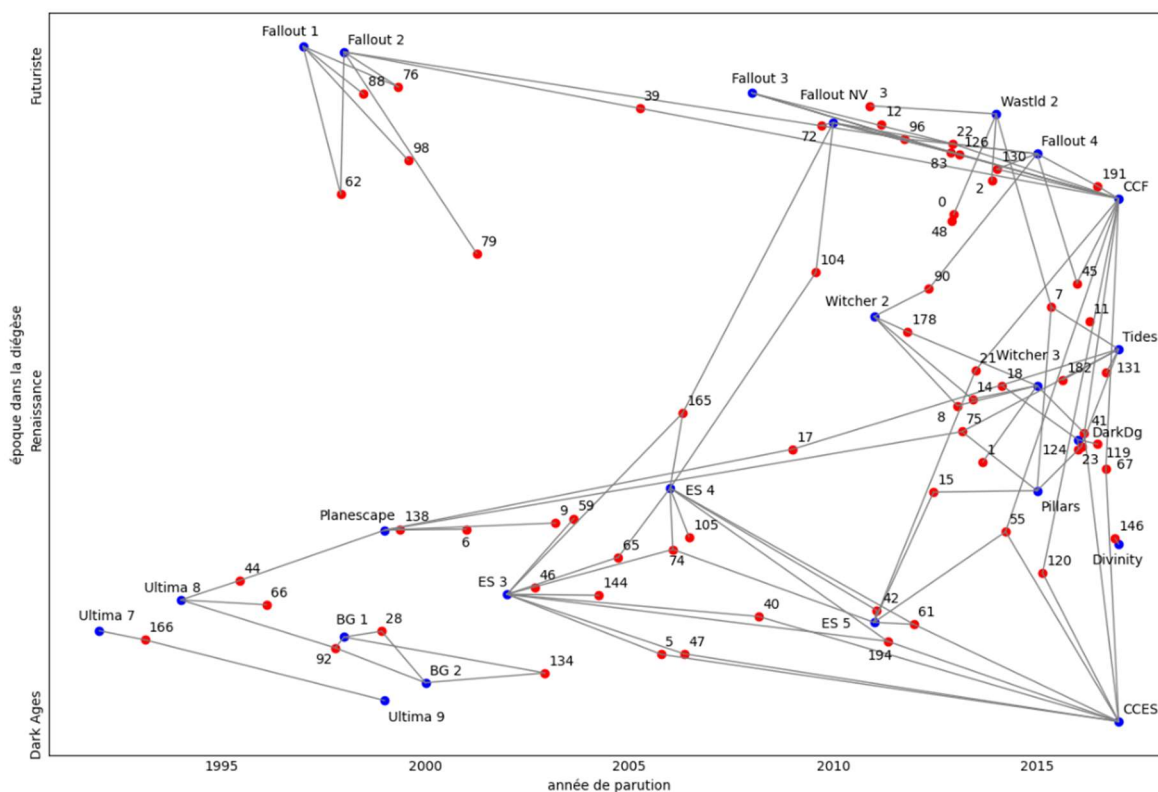


Figure 3. Positionnement des jeux (en bleu) et des classes de topics (en rouge) selon le temps réel (en abscisse) et le temps diégétique (en ordonnée)

On voit assez distinctement qu’il existe, au début de notre période (avant 2002), deux lignes de vocabulaire bien distinctes, qui marquent la double inspiration thématique structurant le genre du RPG depuis ses débuts³⁵ : en haut, les jeux futuristes-apocalyptiques (*Fallout 1* et *2*) ; en bas, les jeux médiévaux-fantastiques (*Ultima 7, 8* et *9*, *BG1*, *BG2* ainsi que *Planescape*, dont le positionnement plus élevé sur l’axe vertical s’explique par le fait que ce jeu représente un multivers où plusieurs plans d’existence sont enchâssés³⁶). Non seulement ces deux catégories de jeux n’ont rien en commun, mais, à l’intérieur de chaque ligne, la cohésion est assez forte. On trouve ainsi davantage de topics communs (classes 28, 44, 66 et 92) aux univers de *fantasy*, pourtant très différents, que de topics communs à *Fallout 1* et *Fallout 3*, qui est plus tardif. Notons que cette rupture sur la ligne supérieure – apparaissant au sein d’une même série et d’un même univers – laisse entrevoir l’influence des équipes de production sur l’évolution du vocabulaire, puisque *Fallout 3* correspond au moment où la franchise (jusque là éditée par Interplay Productions) a été rachetée par Bethesda.

Une autre rupture est introduite, à partir de 2003, par la série des *Elder Scrolls*, des jeux situés dans une temporalité médiévale-fantastique, mais avec un *gameplay* sensiblement différent de ceux de la génération précédente. La série, outre sa forte cohérence en termes de vocabulaire, emprunte à la fois aux topics des jeux médiévaux plus anciens (6, 9, 134, 138) et à ceux des post-apocalyptiques *Fallout 3* et *4* (104, 165).

Une troisième rupture intervient en 2011, date à partir de laquelle s’estompent les frontières de genres thématiques et iconographiques. Cette rupture se marque dans le graphe, d’une part, par le

³⁵ C. S. DETERDING & J. P. ZAGAL, 2018, p. 5.

³⁶ L’univers de *Planescape* est plus amplement décrit dans l’ouvrage de M. BARTON & S. STACKS, 2019, p. 299.

fait que les classes de topics ne sont plus strictement réparties en deux lignes, mais sont davantage partagées par les jeux (voir, pour exemple, les classes 7, 21, 67, 55, etc.), et, d'autre part, par la répartition plus dispersée des jeux sur l'axe vertical. Les titres du corpus se caractérisent ainsi, à partir de ce moment, par une plus grande hétérogénéité des périodes imaginaires auxquelles ils font allusion. On y trouve en effet des titres qui, tout en conservant certaines représentations héritées de la *high fantasy*, intègrent des composantes évoquant la Renaissance (comme *The Witcher 2 et 3* et *Pillars of Eternity*³⁷), voire les XVII^e et XVIII^e siècles (*Darkest Dungeon*, jeu d'inspiration lovecraftienne, mentionne aussi bien des croisés ou des médecins de peste que des mousquetaires et des pistolets). *Tides of Numenera*, pour sa part, fusionne les substrats de *fantasy* et de science-fiction, puisque le jeu prend place dans un futur lointain, où les reliquats technologiques de civilisations disparues sont considérés comme magiques par les humains, qui vivent dans des conditions quasi médiévales³⁸. Les classes de topics 0, 7, 11, 45, 48 et 90 font d'ailleurs le lien entre ces jeux et les derniers *Fallout* ainsi qu'avec *Wastelands 2*, un jeu post-apocalyptique plus récent.

Ce moment de diversification apparaît dans les périodes de l'histoire du CRPG que Barton et Stacks identifient comme « l'âge moderne tardif » (2006 jusqu'à aujourd'hui) et la « Renaissance kickstartée³⁹ », caractérisée par le recours massif aux financements participatifs (dans le corpus, c'est le cas de *Darkest Dungeon*, *Divinity 2*, *Pillars of Eternity*, *Tides of Numenera* et *Wasteland 2*), qui ont permis aux studios de relancer d'anciennes séries ou des formules jugées datées (telles que le combat au tour par tour, considéré comme dépassé par les gros éditeurs⁴⁰). Le fait que ces jeux se distinguent par une plus grande variété thématique est à mettre en lien avec la normalisation de ce nouveau mode de production qui, en limitant les investissements requis par les studios, offre un terreau plus favorable à l'expérimentation.

2. Analyse des résultats suivant trois axes d'interprétation

L'analyse par *topic models* rend possible l'identification de cinq axes d'évolution structurants : thématique (regroupant les super-classes référant aux composantes du monde fictionnel représenté), stylistique (laissant apparaître les changements de registres langagiers employés par les jeux), méta-discursif (mesurant l'utilisation des termes techniques et extradiégétiques), mécanique (composé par les classes référant au *gameplay*) et social (reprenant le vocabulaire utilisé par les jeux pour décrire l'organisation des sociétés qu'ils simulent). Ces cinq axes constituent un premier résultat de la méthode employée : l'étiquetage des classes de topics stables puis de super-classes (rassemblées en fonction d'une proximité sémantique) permet en effet de mettre au jour l'existence de certains champs lexicaux particulièrement cohérents et prégnants, dont on peut supposer qu'ils jouent un rôle important dans l'évolution du CRPG en tant que genre. Les axes d'interprétation annoncés ici constituent en réalité la formalisation des dynamiques évolutives qui apparaissent dans le graphe,

³⁷ « Il a un côté Renaissance [...]. Les aventuriers portent des arquebuses et des pistolets, les caravelles sillonnent les mers en transportant explorateurs, marchands et colons vers de nouvelles frontières, et les sociétés luttent pour faire face à des découvertes transformatrices » ; F. PEPE, 2019, p. 489. Notre traduction.

³⁸ M. BARTON & S. STACKS, 2019, p. 541-542.

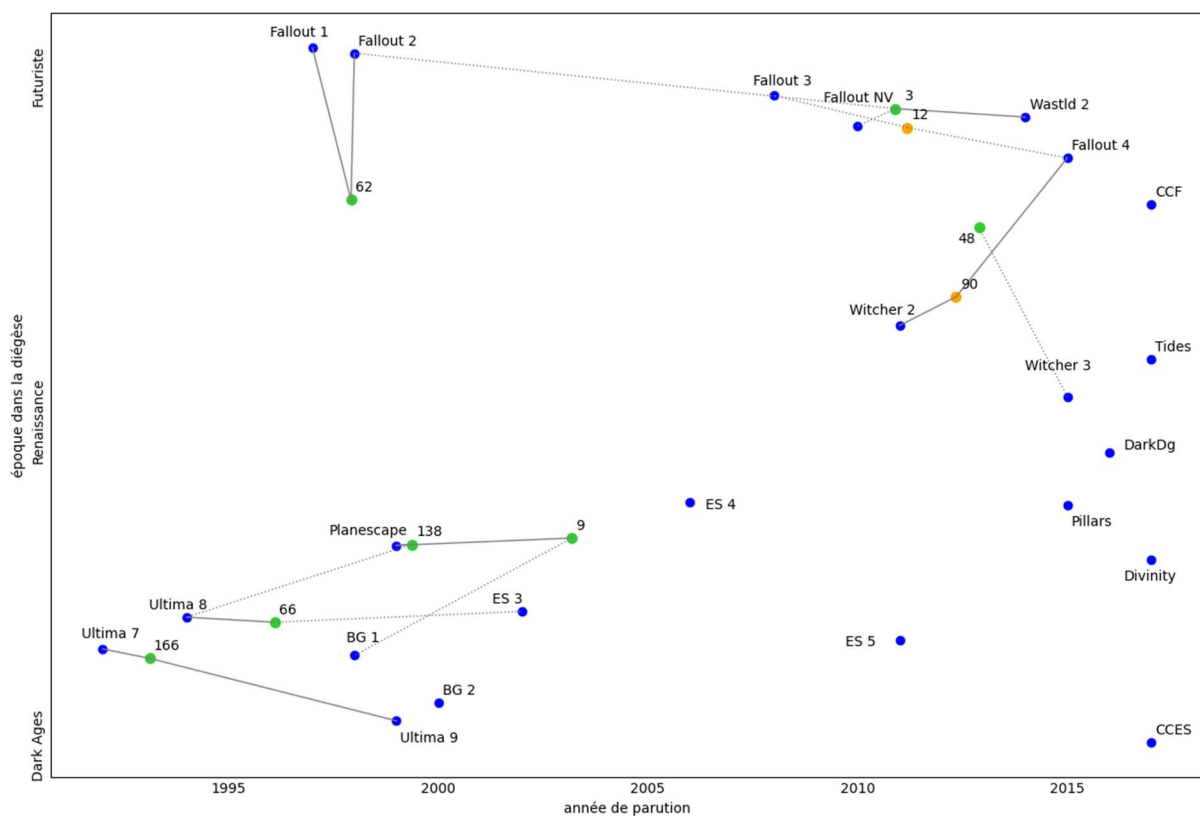
³⁹ M. BARTON & S. STACKS, 2019, p. 499 et 527. L'appellation de la période dérive de *Kickstarter*, nom d'une plateforme de financement participatif parmi les plus célèbres.

⁴⁰ Voir M. BARTON & S. STACKS, 2019, p. 527.

pour l'ensemble des super-classes. Cependant, pour des raisons d'espace, seuls trois de ces axes d'analyse seront explorés ci-dessous : les axes stylistique, méta-discursif et social.

Variations stylistiques des registres de langue employés

L'une des propriétés des CRPG les plus régulièrement commentées est sans doute leur dimension littéraire⁴¹, héritée de deux des trois grands ancêtres qui ont inspiré le genre, le jeu de rôle sur table et les littératures de l'imaginaire⁴². Le *gameplay* du jeu de rôle est en effet souvent fondé sur un format « *dialog-based*⁴³ ». Or cette prépondérance de la conversation est également l'un des résultats rendus apparents par l'analyse des topics, dont plusieurs sont formés autour d'un lexique renvoyant au discours oral. Parmi les classes solides qui ont été formées, plusieurs sont définies par une propriété d'ordre non pas thématique, mais formel : elles réunissent des marqueurs discursifs (ou pragmatiques⁴⁴) connotant divers registres de langage. L'observation de la répartition des super-classes « Vocabulaire de dialogues » (3, 9, 48, 62, 66, 138, 166) et « Verbes de base, vocabulaire générique » (12, 90) permet de montrer que ces marqueurs ont un poids certain dans la caractérisation des jeux du corpus.



⁴¹ Voir, entre autres, D. JARA & E. TORNER, 2018, p. 265 et J. ARJORANTA, 2011.

⁴² Tolkien et Lovecraft, tout particulièrement, voir P. HARRIGAN & N. WARDROP-FRUIIN, 2007, p. 3 et C. S. DETERDING & J. P. ZAGAL, 2018, p. 4-5.

⁴³ J. PETERSON, 2018, p. 55.

⁴⁴ D. SCHIFFRIN, 1987.

Figure 4. Évolution des super-classes « Vocabulaire de dialogues » (en vert) et « Verbes de base, vocabulaire générique » (en orange)

Note. Dans ce graphe et tous les suivants, en plus des arêtes pleines qui sont celles héritées de la Figure 3, on a ajouté des arêtes pointillées pour les contributions un peu moins importantes (au moins 7,5 % des formes).

Un premier élément rendu visible par le graphe est que les séries de jeux médiévaux-fantastiques pré-2005 (*Ultima*, *BG* et *Planescape*, avec une légère connexion les reliant aux débuts des *ES*) semblent partager un même vocabulaire discursif. Plusieurs classes de topics relient ces titres, et on peut noter que la présence de classes formelles est particulièrement déterminante pour cette période, par opposition à l'époque post-2005, où la stylisation des dialogues est moins structurante.

Au sein de ce groupe, la série *Ultima* se distingue par son recours à un lexique particulièrement archaïsant, incarné par la classe 166 (*dost, thou, thine, wouldst*). Cette propriété stylistique – que Barton et Stacks identifient comme une trace énonciative de Richard Garriott, le créateur de la série⁴⁵ – ne se limite d'ailleurs pas aux jeux, mais marquait aussi leurs manuels⁴⁶.

Un certain archaïsme et une recherche poétique se manifestent également dans la classe 9 (connectant *Planescape* et *BG1*), bien que celle-ci se distingue plus spécifiquement par le caractère dialectal des termes qui la composent (*aye, fer, ya*). Les classes 138 et 66, qui jouent un rôle de connecteurs entre séries, mobilisent pour leur part à la fois le champ lexical de la communication (*questions, speak, nods, hear*), des salutations formelles ou tournures de politesses (*farewell, greetings, thanks, please*) et des marqueurs discursifs servant à l'organisation de la conversation⁴⁷ (*well, yes*). Si le caractère archaïsant du vocabulaire n'est pas uniformément réparti dans cette catégorie de jeux, ces derniers partagent un registre de langage plutôt soutenu.

Ce trait distingue fortement les titres médiévaux-fantastiques pré-2005 des titres post-apocalyptiques. Les premiers *Fallout* (1 et 2), d'une part, sont caractérisés par un registre plus neutre ou courant : la classe 62 mêle en effet des verbes génériques (*tell, need, talk, ask*) à des formules basiques de politesse (*sorry, thanks, goodbye*) et à des marqueurs conversationnels moins châtiés (*uh, oh, well, sure*), se rapprochant des expressions de l'époque contemporaine. D'autre part, *Fallout 2* tend déjà vers un trait qui va être significatif pour la suite de la saga (et qui est partagé par *Wasteland 2*), à savoir l'utilisation d'interjections prosaïques, voire vulgaires (*hey, like, well, fuck, shit*).

L'analyse révèle donc que la parenté entre les jeux de la ligne supérieure n'est pas uniquement thématique, mais aussi linguistique et stylistique : le sous-genre post-apocalyptique ne tient pas uniquement dans les composantes des univers représentés, mais aussi dans la tonalité via laquelle s'expriment leurs personnages. Cette intrication de la forme et du fond est d'ailleurs incarnée par l'ambivalence de la classe 3, qui comprend autant de marqueurs d'oralité que de termes renvoyant à des composantes de la diégèse (*desert, robots, mayor, ranger*).

Une connexion est néanmoins assurée entre ces fictions post-apocalyptiques et la série *The Witcher* par la classe 90 (un vocabulaire générique) et la classe 48, qui multiplie les signes d'expressivité informels et grossiers (*oh, sure, fine, huh, damn, fuck*). Ainsi, à travers le cas de *The Witcher*, on voit que

⁴⁵ M. BARTON & S. STACKS, 2019.

⁴⁶ M. BARTON & S. STACKS, 2019, p. 89.

⁴⁷ D. SCHIFFRIN, 1987, p. 102.

les mutations apportées au genre par les jeux de l'« âge moderne tardif⁴⁸ » ne consistent pas uniquement en une diversification des composantes diégétiques représentées et en un traitement plus libéral des thèmes de la *high fantasy* : elle passe aussi par l'intégration, dans les dialogues, de registres de langage inférieurs, qu'on peut interpréter comme un marqueur de modernité.

Apparition des tutoriels et du méta-discours

Les CRPG ne se définissent toutefois pas uniquement par leurs univers inspirés des littératures de l'imaginaire et par leur style linguistique : le genre est aussi construit par des mécaniques de *game-play*, par le recours à des règles explicites qui déterminent les interactions que le joueur peut avoir avec le monde fictionnel. Ces règles se manifestent à des degrés divers dans le vocabulaire des jeux.

Quatre classes de topics (104, 119, 124 et 165) ont en commun l'inclusion de termes méta-discursifs. Ceux-ci servent aux jeux à expliquer leur propre fonctionnement à l'utilisateur et renvoient tantôt à l'interface physique (*screen, mouse, button*) ou graphique (*menu, inventory, options*), à des instructions tutorielles livrées au joueur (*hold, select, click*) ou à des éléments de la fiction ludique reconnus en tant que tels (*character, quest, game, player*). L'association de ces différents champs lexicaux permet de distinguer les quatre classes méta-discursives des classes 22 et 96 (associées aux jeux *Fallout* modernes), qui comprennent également du vocabulaire lié à l'informatique, mais où celui-ci n'est associé qu'à des composantes existant au sein de la diégèse (l'univers de *Fallout* contient bel et bien des ordinateurs). L'articulation de termes tels que *player, click, button, character* et *menu*, au contraire, laisse entrevoir une adresse directe à l'utilisateur qui suppose la production d'une métalepse, c'est-à-dire d'une rupture des frontières de la fiction⁴⁹, par l'intermédiaire de laquelle « des niveaux narratifs normalement étanches se retrouvent alors reliés, allant jusqu'à l'interpénétration du monde raconté et du monde depuis lequel on raconte⁵⁰ ». Par ces transgressions, le jeu vidéo reconnaît temporairement sa propre artificialité. Or, si la métalepse est le mode d'expression privilégié du tutoriel dans le jeu vidéo de manière générale⁵¹, au point d'être considérée comme une véritable convention vidéoludique⁵², le graphe ci-dessous montre que le recours à ces figures ne caractérise qu'un nombre très limité de jeux du corpus.

⁴⁸ M. BARTON & S. STACKS, 2019, p. 499.

⁴⁹ G. GENETTE, 1982, p. 527.

⁵⁰ S. ALLAIN, 2018.

⁵¹ Les phrases telles que « appuie sur B pour courir » sont métaleptiques en ce qu'elles renvoient simultanément au monde fictionnel et au monde empirique. Voir B. NELIDE-MOUNIAPIN, 2005, p. 245.

⁵² Voir M.-L. RYAN, 2004 et A. ENSSLIN, 2015.

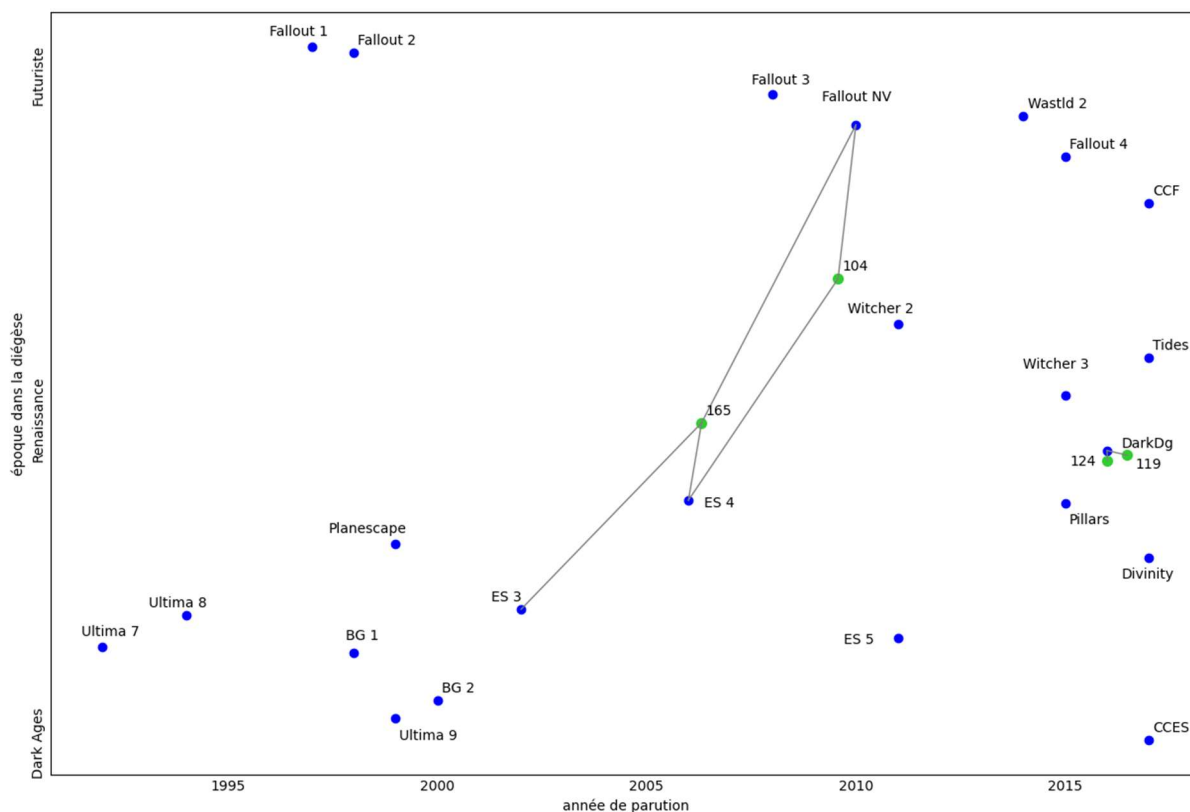


Figure 5. Classes mobilisant du vocabulaire méta-discursif (tutoriels, interfaces, etc.)

On y voit que le lexique méta-discursif n'apparaît de façon déterminante qu'à partir de *Morrowind*, en 2002. On peut en déduire que les jeux de la période précédente ne recourent pas autant à des termes extradiégétiques pour expliciter leur fonctionnement (ils « diégétisent⁵³ » leurs explications) ou qu'ils donnent moins d'instructions au joueur. Rappelons que les CPRG des premiers temps étaient principalement destinés à un public de joueurs déjà familiers des règles du jeu de rôle sur table et que les jeux vidéo de cette époque reposaient largement sur le recours à des ressources paratextuelles pour guider leurs utilisateurs, notamment des manuels d'instruction⁵⁴.

Ainsi, le fait que ce n'est qu'à la période moderne (post-2000) qu'apparaissent des classes méta-discursives (104 et 165) montre que les jeux du corpus s'inscrivent tardivement dans la « montée du paradigme d'assistance dans le design du jeu vidéo » décrite par Therrien et Julien⁵⁵. De plus, ces classes méta-discursives forment un vocabulaire commun aux *ES* et aux jeux *Fallout* tardifs (après le rachat de la franchise par Bethesda), ce qui constitue un nouveau marqueur de l'influence des studios de développement et d'édition sur la détermination du vocabulaire : bien que les univers des *ES* et *Fallout* ne pourraient être plus différents, ces jeux partagent des similarités structurelles (le même type de tutoriels) et des conventions communes (ils utilisent un lexique identique pour décrire leur mode de fonctionnement). Ces conventions portent d'ailleurs autant sur les outils informationnels de

⁵³ Stratégie narrative décrite plus amplement dans F. BARNABE, 2018.

⁵⁴ C. THERRIEN & M. JULIEN, 2015, p. 11.

⁵⁵ C. THERRIEN & M. JULIEN, 2015, p. 12.

l'interface (*script, journal, timer, target*) que sur la manière par laquelle ces jeux s'adressent au joueur pour lui donner des directives (*disable, reset, enable, start*).

Le fait que le méta-discours soit introduit par les jeux Bethesda invite à formuler deux remarques. Premièrement, l'apparition de ces classes indique la prise d'importance, au sein des textes, des mentions de l'expérience de jeu. Là où d'autres titres (même postérieurs : *The Witcher, Divinity, Tides of Numenera*) discutent prioritairement sur leur univers, les jeux Bethesda parlent explicitement de l'activité ludique en tant que telle et mentionnent directement le « joueur ». Deuxièmement, la série *ES* est connue pour avoir fait radicalement basculer le genre du CRPG vers un *gameplay* plus orienté « action », là où les titres des premiers temps privilégiaient des systèmes de combat « au tour par tour » ou « en temps réel avec pause »⁵⁶. En d'autres termes, ces jeux sont basés sur un rapport plus direct du joueur à l'univers, puisque ses gestes y sont immédiatement retranscrits. On peut s'étonner, en conséquence, que ces titres supposément plus « immersifs » soient précisément ceux qui recourent de façon privilégiée à un méta-discours, mettant à distance les contenus ainsi décrits.

Ce résultat permet en réalité de déconstruire les présupposés régulièrement énoncés concernant « l'intuitivité » de certains genres vidéoludiques : on voit ici que, dans les jeux Bethesda, bien que le joueur puisse agir sans passer par certains intermédiaires (des icônes, des menus, des statistiques...), cette action n'en est pas moins préalablement encadrée par des tutoriels, des instructions ou des explications. Toutefois, on peut également noter que le méta-discours est particulièrement présent au moment du basculement du genre vers l'action, puis que sa présence diminue à nouveau dans les jeux plus récents des séries *Fallout* et *ES* : il semble ainsi qu'une fois les conventions de l'*action role-playing game* (ARPG) établies, la nécessité d'un discours d'escorte se fait moins sentir. On perçoit ici, une fois encore, l'intérêt de l'étude diachronique du vocabulaire pour l'historicisation du jeu vidéo et l'histoire culturelle : l'analyse des mutations lexicales permet d'éclairer autrement les conditions d'exploitation ou de réception des jeux examinés.

Systèmes de réputation et de moralité : représentations de l'organisation sociale et morale de la diégèse

L'un des résultats de l'analyse par *topics models* peut, de prime abord, paraître surprenant : il s'agit de l'importance du vocabulaire lié à l'organisation sociale du monde représenté. La super-classe Société et collectivités – qui couvre des thèmes tels que la famille, les guildes et groupes ou les structures politico-religieuses – est effectivement celle qui rassemble le plus grand nombre de classes. Une telle présence indique que le vocabulaire lié à ces thématiques apparaît dans les jeux de façon assez cohérente et qu'il caractérise fortement certains titres.

Ce résultat invite à remettre en avant la capacité du jeu de rôle – principalement étudié comme *pratique sociale*⁵⁷ – à être le vecteur d'un *commentaire social*⁵⁸. Comme le rappellent Christoph Deterding et José Zagal, les rôles qui donnent leur nom au genre « sont un élément fondamental des

⁵⁶ Voir D. SCHULES, J. PETERSON & M. PICARD, 2018, p. 111.

⁵⁷ Pour un exemple, voir l'article de R. BARTLE, 1997. Pour un état de l'art plus détaillé, voir J. P. WILLIAMS *et al.* 2018.

⁵⁸ La critique sociale et politique était d'ailleurs un objectif explicitement revendiqué par les designers des premiers *Fallout* (M. BARTON & S. STACKS, 2019, p. 386).

structures et des processus du pouvoir dans une société⁵⁹ » et il n'est donc pas étonnant que les jeux déploient un lexique particulier pour nommer, décrire et qualifier ces structures. Une idée similaire est au cœur de l'article d'Erik Champion, qui défend la capacité des CRPG solo à « transmettre l'impression de mondes partagés dotés d'une présence sociale et d'une agentivité sociale⁶⁰. » Au-delà de la simple représentation, le genre fait d'ailleurs du commentaire social une mécanique de *gameplay* à part entière à travers la mise en place de systèmes dits « de réputation » ou « de moralité », qui existent depuis les débuts de son histoire⁶¹. Ces deux modèles ont un fonctionnement globalement similaire (ils consistent à attribuer des points aux actions et choix du joueur), mais différent en fonction des variables qu'ils affectent : soit la popularité de l'avatar auprès de différentes factions, soit son alignement sur un axe moral.

Dans ce point, nous allons donc détailler la manière dont les jeux du corpus parlent de l'organisation sociale de leur univers et la manière dont ces discours s'articulent avec une certaine vision du monde. Pour ce faire, nous observerons en parallèle l'évolution des classes de topics appartenant aux super-classes Société et collectivités (1, 2, 44, 45, 59, 61, 65, 72, 79, 146, 178) et Cosmogonie, Histoire (8, 15, 18, 92, 182).

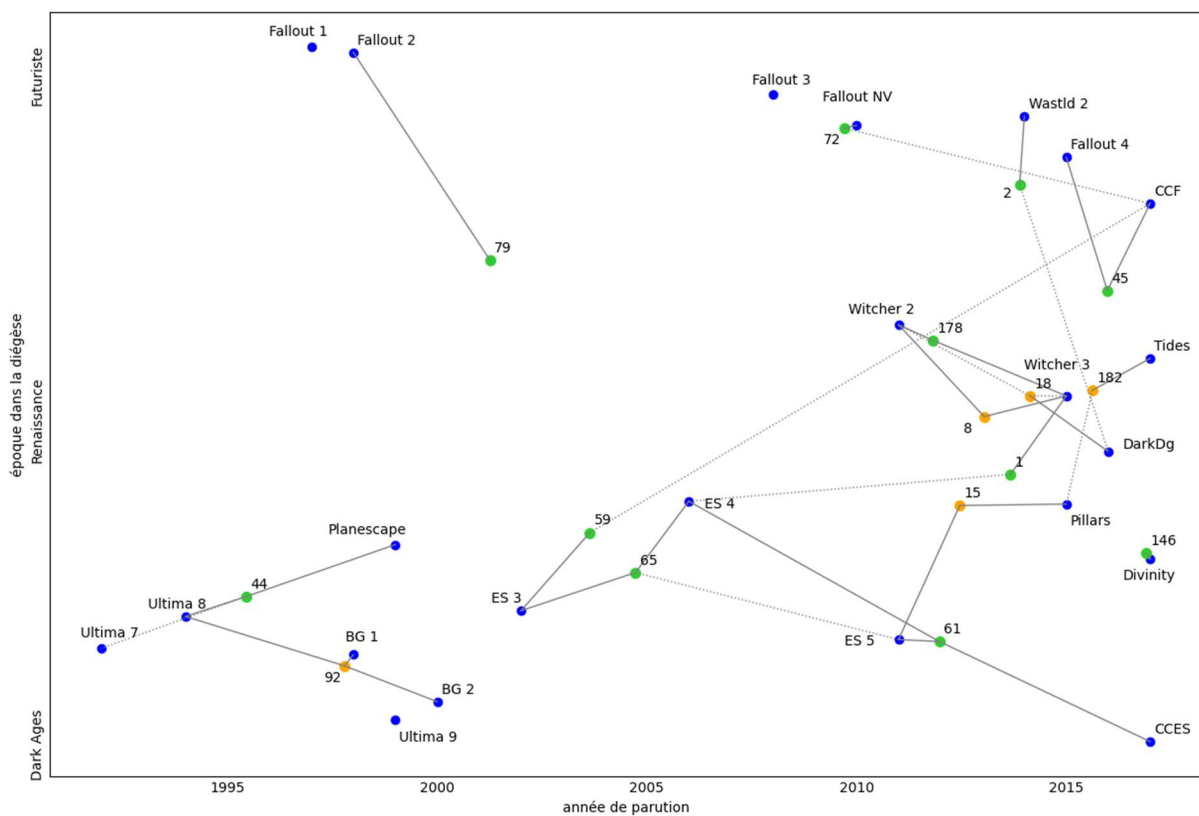


Figure 6. Variation des super-classes Société et collectivités (en vert) et Cosmogonie, Histoire (en jaune)

⁵⁹ C. S. DETERDING & J. P. ZAGAL, 2018, p. 3. Notre traduction.

⁶⁰ E. CHAMPION, 2009, p. 37. Notre traduction.

⁶¹ Voir M. BARTON & S. STACKS, 2019, p. 187.

La mise en parallèle de ces classes permet de remarquer, tout d'abord, l'isolement des deux premiers *Fallout*, qui se caractérisent par un vocabulaire renvoyant à des formes de collectivités dispersées, claniques (*group, hunters, villagers, pack*), qui ne semblent pas être traversées par un quelconque système d'organisation politique, ni associées à des systèmes de valeur. Cette juxtaposition de noms de groupes isolés n'étonne pas dans le cadre d'une fiction post-apocalyptique, qui repose justement sur la mise en scène d'un monde qui a perdu ses institutions et structures sociétales. Cette dimension tribale continue d'ailleurs de caractériser la série jusque *Fallout New Vegas*, qui se distingue lui aussi par la mention de factions multiples (*legion, khans, troops, boomers, followers*), bien que son vocabulaire intègre davantage de termes renvoyant à une organisation militaire (*army, battle, president*).

À l'opposé, les jeux médiévaux-fantastiques pré-2000 (*Ultima, Planescape* et *BG*) forment eux aussi un tout cohérent, marqué par un lexique mettant peu en avant les castes propres à ce type d'univers. Loin de laisser apparaître des thèmes politiques (qui ne sont pourtant pas absents de ces jeux), la classe 44 fait référence à une structuration plus familiale (*people, father, mother, house, name*), notamment associée à des principes positivement marqués (*love, truth, good*). La classe 92, qui connecte les séries *Ultima* et *BG*, aborde quant à elle le thème de la « destinée » en mélangeant des verbes d'action génériques liés au devoir ou à la quête (*shall, help, seek, order, etc.*) avec des figures de pouvoir (*lord, power*) et des termes renvoyant au temps (*long, end, never*), à des valeurs (*evil, great, good*) ou à l'existence (*life, die*).

Contrairement à ce que l'on peut voir dans le domaine post-apocalyptique à la même époque, où le système de réputation n'est pas lié à une structuration morale de l'univers, les premiers jeux médiévaux-fantastiques se caractérisent donc par une articulation entre les mentions d'une institution sociale (la famille) et l'expression d'un système de valeur (le bien et le mal).

L'ensemble suivant, formé par la série des *ES*, marque l'apparition conjointe de deux nouveaux thèmes : le politique et le religieux. Si la classe 65 rassemble surtout du vocabulaire lié aux guildes (*city, guild, members, join*), la 59 incarne particulièrement cette combinaison en renvoyant autant au pouvoir civil (*lord, law, council, tribunal*) qu'au religieux (*prophecies, priest, temple*). Or notons que cette association ne se limite pas à la série ni au domaine médiéval-fantastique, mais caractérise aussi très fortement les jeux post-apocalyptiques les plus modernes, à savoir *Fallout 4* (ainsi que son *Creation Club*) et *Wasteland 2*. On retrouve en effet, dans les classes 2 et 45, autant de termes théologiques liés au salut (*church, faith, salvation, sin ; sister, holy, confessor*) que de termes évoquant une organisation militaire (*bastion, king, militia ; captain, peace, division*).

Le fait que ce trait soit partagé par *Wasteland 2* exclut l'idée selon laquelle l'organisation politico-religieuse serait l'apanage des jeux du studio Bethesda et invite plutôt à lire dans cette résurgence du motif une véritable évolution générique (d'autant que la classe 2 est également assez présente dans *Darkest Dungeon*, un jeu par ailleurs assez isolé des autres). Dans la *fantasy* comme dans le post-apocalyptique, on passerait ainsi de jeux mettant en avant des structures sociales de proximité (les groupes du monde chaotique des premiers *Fallout* ou les cercles familiaux dans les jeux médiévaux-fantastiques pré-2000) à des titres représentant un monde social structuré par de multiples institutions. Les collectivités qui les peuplent peuvent être formées par un principe martial (*Fallout 4, Wasteland 2*), ou caractériser un groupe social (les guildes des *ES*), mais le vocabulaire qui s'y réfère est en tout cas parsemé de religiosité. La dimension spirituelle n'est pas non plus absente de *Tides of Numenera* ou *Divinity 2*, avec un accent mis sur la métaphysique dans le premier cas, et sur les divisions ethno-religieuses dans le second.

Les jeux *The Witcher* se singularisent en n'associant plus le pouvoir politique à des forces ou institutions spirituelles, mais aux lexiques de la mission, de la guerre et de la temporalité historique. Les classes 1 et 178 combinent des titres ou fonctions (*duchess, knights ; king, soldiers*) à des instructions (*contract, must, order ; help, kill, wanted*), laissant entendre que les instances politiques mentionnées apparaissent régulièrement dans ces jeux comme des donneurs et/ou des objets de quêtes. Ces jeux confèrent à l'avatar (et au joueur) une véritable fonction sociale : celui-ci n'est pas seulement intégré à une société fictionnelle au sein de laquelle il est invité à choisir un camp, mais exerce un pouvoir transformateur sur cette société via les missions qu'il accomplit. D'autre part, l'absence de factions ou de termes moraux laisse entendre que les systèmes de moralité et de réputation sont moins structurants au sein de cet univers. Ces lexiques sont remplacés par ceux du temps et de la guerre : à la classe 18, qui rassemble les termes absolus de la destinée (*life, death, world, time*), s'ajoute ainsi la classe 8, qui réfère au récit historique d'une guerre médiévale (*years, happened, emperor, queen, battle, died, war*). Cette dernière thématique est d'ailleurs aussi au cœur de la classe 15 (*war, battle, land, time*), qui connecte le dernier *ES* à *Pillars of Eternity*, maintenant ainsi une ligne d'influence thématique (l'organisation du monde par le conflit guerrier) au sein des jeux médiévaux-fantastiques modernes.

En synthèse, si la série *Fallout* se distingue en préservant une forte stabilité de son modèle à travers le temps (un amalgame chaotique de groupes locaux), le reste du corpus témoigne d'une double évolution. Premièrement, on passe autour de l'an 2000 d'un ensemble de micro-structures évoluant dans un univers pétri de valeurs morales (*Ultima, BG, Planescape*) à une scénographie complexe impliquant de nombreux acteurs politico-religieux aux narrations concurrentes (*ES, Divinity, Darkest Dungeon*). Dans un second temps (2011), l'universalisme réapparaît dans plusieurs titres à travers la mise en scène de populations accablées par l'Histoire et la guerre (*Skyrim, Pillars of Eternity, The Witcher*).

Comme le souligne Champion, « dans un sens, le jeu de rôle est [un travail] de conservateur, nous choisissons les aspects de la culture qui valent la peine d'être conservés et nous nous débarrassons du reste des informations⁶². » Ainsi, la résurgence du lexique théologique peut être interprétée comme une manifestation de « l'enchantement désenchanté⁶³ » dont Deterding et Zagal considèrent le RPG comme un représentant emblématique⁶⁴. Dans le même ordre d'idées, la présence croissante de la guerre et du vocabulaire eschatologique n'est sans doute pas sans lien avec les préoccupations collectives qui marquent notre propre société.

Conclusions et prolongements : influence des studios et analyse de réseaux

Nous avons choisi, dans ce travail, de recourir à des outils encore peu utilisés dans le domaine des sciences historiques, et ce parti pris méthodologique a été déterminant pour l'obtention des résultats que nous venons de présenter. Le recours aux *topic models* nous a permis de nous affranchir des

⁶² E. CHAMPION, 2009, p. 39. Notre traduction.

⁶³ M. SALER, 2012, p. 12.

⁶⁴ « Alors que la sécularisation, la rationalisation et la bureaucratisation ont débarrassé notre monde moderne d'expériences plus profondes de signification et d'émerveillement spirituel, magique ou sublime, les auteurs de fiction ont rétabli des mondes imaginaires remplis d'un tel enchantement, bien qu'avec une conscience ironique de leur statut de "comme si" » (C. S. DETERDING & J. P. ZAGAL, 2018, p. 5). Notre traduction.

variations de vocabulaire dans un corpus qui compte plus d'auteurs que de textes, et de mieux saisir des évolutions sémantiques significatives. Par la suite, le choix de représentations souples comme la carte auto-organisatrice ou le graphe biparti jeu-topic nous a évité d'être limités par une troncature brutale de dimensionnalité, en préférant au seul critère de maximisation de la variance la représentation d'un maximum de relations de proximité. Enfin, le positionnement contrôlé des sommets du graphe nous a amenés à construire une représentation focalisée sur la temporalité, mais s'affranchissant des discontinuités qui auraient rendu illisibles de simples séries temporelles définies sur moins de quinze dates.

C'est ainsi que nous sommes parvenus à éviter aussi bien l'illusion d'une réinvention permanente que celle d'un simple découpage par sous-genres ou périodes. En d'autres termes, à travers l'examen des mots en contexte, l'analyse a permis de représenter de façon dynamique les mutations d'un genre dont l'évolution est « tout sauf linéaire⁶⁵ » et d'apporter un éclairage neuf aux approches historiques et tentatives de périodisation existantes. Nous avons dégagé cinq grands axes d'évolution du genre (thématique, stylistique, méta-discursif, mécanique et social), sur lesquels s'exercent conjointement plusieurs tensions avec lesquelles chaque jeu doit négocier. Bien que seul trois d'entre eux aient pu être approfondis dans les limites de cet article, nous proposons une visualisation récapitulative des évolutions observées dans le vocabulaire sur l'ensemble des cinq axes, auxquels s'ajoute une ligne (nommée « temps diégétique ») reprenant la répartition des temporalités imaginaires représentées (Figure 7). Cette représentation est nécessairement schématique, mais permet d'embrasser d'un seul regard les multiples paradigmes qui interviennent dans la construction du genre CRPG et de mesurer la difficulté de définir, au sein de ces dynamiques, des périodes qui formeraient des ensembles cohérents sur tous les niveaux à la fois – difficulté à laquelle l'approche par *topic models* fournit une alternative.

⁶⁵ M. BARTON & S. STACKS, 2019, p. 383. Notre traduction.

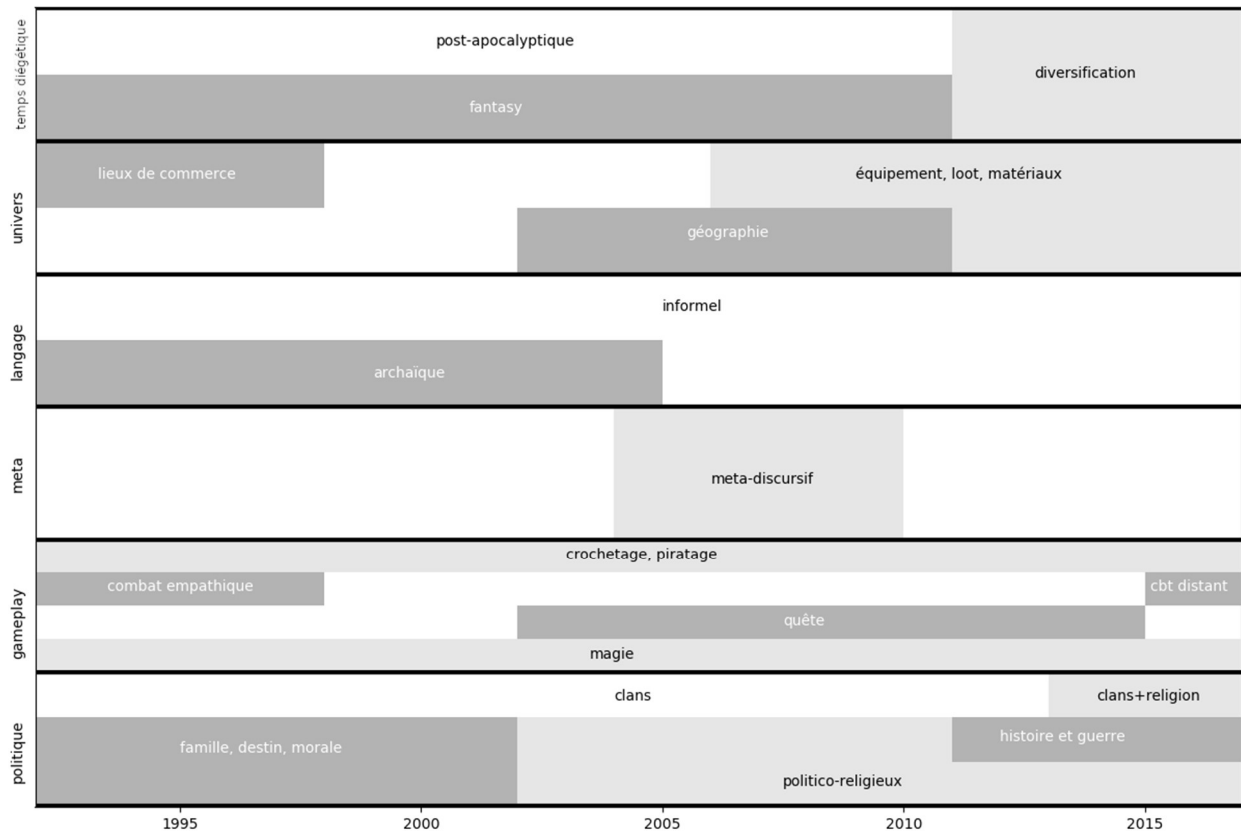


Figure 7. Frise récapitulative des évolutions observées sur les différents axes (en ordonnée) dans le temps (en abscisse)

D'un point de vue thématique, on a effectivement vu que l'opposition entre les univers de *fantasy* et de science-fiction, au départ très marquée, se fait moins ferme à partir des années 2010, qui voient apparaître la représentation d'une plus grande diversité de périodes imaginaires, ce qui incarne dans le lexique l'avènement d'une période de renouveau dans les logiques de production des CRPG, la « Renaissance kickstartée ». Au niveau stylistique, on a signalé le rôle déterminant des marqueurs discursifs dans le vocabulaire des jeux, qui montrent notamment une opposition entre les titres mobilisant un ton archaïque et formel, qui disparaissent après 2005, et ceux employant un registre familier ou (plus récemment) vulgaire. À travers l'analyse des termes extradiégétiques, on a pu mettre au jour l'apparition tardive (après 2002) du vocabulaire méta-discursif, qui trace un pont entre les jeux du studio Bethesda et marque l'entrée du CRPG dans le paradigme de l'assistance et du *design* coopératif⁶⁶ (qui prend son origine dans les années 1980). Enfin, en ce qui concerne les représentations sociales, nous avons observé après 2000 l'abandon d'une narration unifiée ancrée dans un système moral fort, au profit de la multiplication de structures politiques disposant d'agendas propres. Ce phénomène est cependant amoindri à partir de 2011 dans plusieurs jeux par le retour d'une part d'universel sous forme de trame eschatologique.

Afin de mieux comprendre le rôle du jeu vidéo en tant que discours social, il serait utile de mettre en relation ces premières observations avec une analyse approfondie des contextes socio-culturels dans lesquels ces jeux voient le jour et sont reçus. Les événements historiques ou politiques majeurs de

⁶⁶ C. THERRIEN & M. JULIEN, 2015, p. 12.

l'époque contemporaine ont-ils une répercussion sur la manière dont les jeux pensent le social ? Certaines structures politiques connues ou identifiables se retrouvent-elles dans les œuvres ludiques selon les époques ou les lieux de production ? Nous voyons que l'intérêt d'une étude diachronique du lexique du CRPG ne se borne pas à contribuer à une meilleure définition du genre : en repérant et en représentant les filiations, ruptures et réapparitions qui rythment les transformations du vocabulaire au fil du temps, ce travail participe aussi à inscrire ce genre dans le temps historique. Dans un contexte où le jeu vidéo est encore parfois considéré comme un objet « nouveau », prisonnier d'une rhétorique des « lendemains qui chantent »⁶⁷ et dont les évolutions ne sont pensées qu'en termes techniques ou commerciaux, cet article espère ainsi avoir apporté des éléments empiriques qui alimenteront l'appréhension de ce phénomène culturel dans le temps.

Enfin, il importe de souligner qu'un dernier axe a montré son influence tout au long de nos analyses : le fait que certains des jeux étudiés partagent un même studio de développement ou un même éditeur. En effet, l'évolution des corpus textuels observée ici s'inscrit elle-même dans une évolution des conditions matérielles de production : les studios de développement naissent, disparaissent ou se rachètent ; les licences changent de main et les employés passent d'un projet à un autre. Notre corpus, en particulier, est le lieu de jeux d'influence multiples et complexes : si la série *Fallout* a initialement été inspirée par *Wasteland*, elle n'a elle-même pas été sans impact sur la conception de *Wasteland 2* ; de même, *Planescape* a motivé la création de *Tides of Numenera*, mais celle-ci a été réalisée par un tout autre studio (inXile Entertainment, le même que celui de *Wasteland 2*, ce qui explique sans doute en partie le fait que deux jeux aussi différents partagent plusieurs classes) ; *Planescape* a quant à lui été produit par Black Isle Studios (développeurs de *Fallout 2*, avec qui il ne partage pourtant aucune classe de topics), mais avec le moteur de Bioware (développeurs de *BG1* et *2*) tout en étant édité par Interplay Productions (aussi éditeurs de *BG1* et *2* et de *Fallout 1* et *2*) ; *Pillars of Eternity* est un descendant direct de *BG*, bien que ses producteurs soient différents, et il est développé par Obsidian Entertainment, le studio de *Fallout: New Vegas* ; etc. Ces nombreux héritages ou transferts apparaissent de façon plus lisible dans le graphe ci-dessous (Figure 8)).

⁶⁷ M. TRICLOT, 2016.

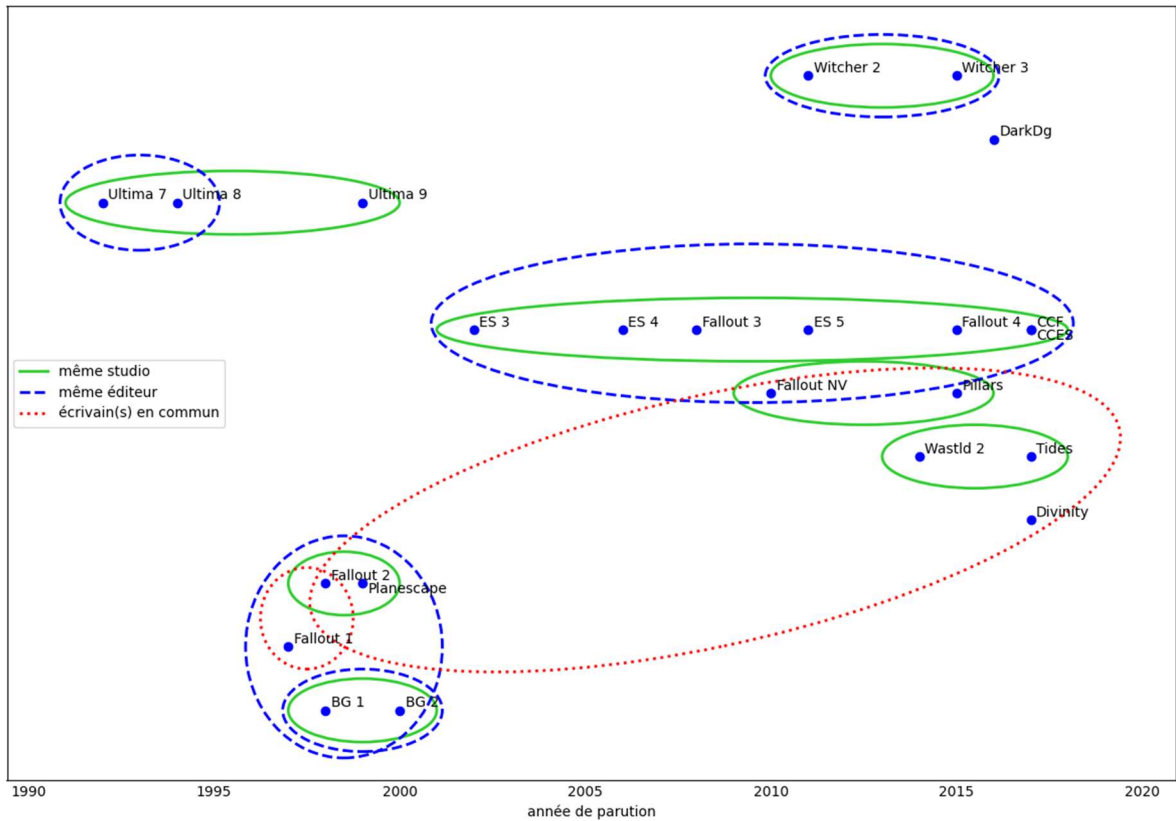


Figure 8. Représentation informelle sous forme d'hypergraphe d'une partie des connexions entre les différentes productions

L'analyse de ces jeux d'influence – auxquels s'ajoutent les effets de cohérence des séries et les répartitions de vocabulaire liées à la mode des *reboots* – permettrait de jeter un éclairage supplémentaire à nos résultats (on a, entre autres, déjà noté les nombreux partages entre les jeux Bethesda) et de soulever de nouvelles pistes de questionnement. Pourquoi, par exemple, *Divinity 2* (développé par Larian, un studio ancien) partage-t-il aussi peu de vocabulaire avec les autres titres, avec lesquels sa série a pourtant cohabité, alors qu'un jeu comme *Tides of Numenera* prolonge fidèlement le lexique d'œuvres anciennes ? Développer ces interrogations dépasserait néanmoins les limites du présent article, d'autant qu'une évaluation efficace de l'influence des équipes de développement sur la circulation du vocabulaire ne pourrait se faire qu'à condition de cartographier précisément la mobilité des acteurs au sein de ces studios. Cette première recherche ouvre donc la porte à une analyse de réseaux, qui permettra de mesurer l'impact du contexte de production sur l'évolution du genre.

Au-delà de ces perspectives, rappelons néanmoins que l'utilité des méthodes et outils détaillés au sein de cet article ne se veut pas limitée au domaine vidéoludique. En effet, s'ils servent efficacement l'entreprise en cours d'historicisation de ce médium (notamment en permettant de faire jaillir des résultats de façon inductive), leur application à un corpus de grande ampleur, hétérogène, discontinu et possédant des composantes sérielles avait également pour but de démontrer leur applicabilité à d'autres corpus historiques marqués par des propriétés similaires. C'est précisément pour faciliter cette appropriation des outils par les historiens que l'intégralité des codes mobilisés dans ce travail a été rendu disponible sur l'espace Nakala associé au présent numéro.

Bibliographie

Sources

- Baldurs Gate I*, BioWare, Interplay Productions & Black Isle Studios, 1998.
- Baldurs Gate II: Shadows of Amn*, BioWare, Interplay Productions & Black Isle Studios, 2000.
- Creation Club (Fallout 4)*, Bethesda Game Studios, Bethesda Softworks, 2017.
- Creation Club (Skyrim)*, Bethesda Game Studios, Bethesda Softworks, 2017.
- Darkest Dungeon*, Red Hook Studios, Red Hook Studios & Merge Games, 2016 (early access: 2015).
- Divinity Original Sin 2*, Larian Studios NV, Larian Studios NV, 2017.
- Fallout 1*, Interplay Productions, Interplay Productions, 1997.
- Fallout 2*, Black Isle Studios, Interplay Productions, 1998.
- Fallout 3*, Bethesda Game Studios, Bethesda Softworks, 2008.
- Fallout 4*, Bethesda Game Studios, Bethesda Softworks, 2015.
- Fallout: New Vegas*, Obsidian Entertainment, Bethesda Softworks, 2010.
- Pillars of Eternity*, Obsidian Entertainment, Paradox Interactive, 2015.
- Planescape: Torment*, Black Isle Studios, Interplay Productions, 1999.
- The Elder Scrolls III: Morrowind*, Bethesda Game Studios, Bethesda Softworks, 2002.
- The Elder Scrolls IV: Oblivion*, Bethesda Game Studios, Bethesda Softworks, 2006.
- The Elder Scrolls V: Skyrim*, Bethesda Game Studios, Bethesda Softworks & 2K Games, 2011.
- The Witcher 2: Assassins of Kings*, CD Projekt RED, CD Projekt & Atari, 2011.
- The Witcher 3: Wild Hunt*, CD Projekt RED, CD Projekt, 2015.
- Torment: Tides of Numenera*, inXile Entertainment, Techland Publishing, 2017.
- Ultima IX: Ascension*, Origin Systems, Electronic Arts, 1999.
- Ultima VII: The Black Gate*, Origin Systems, Origin Systems, 1992.
- Ultima VIII: Pagan*, Origin Systems, Origin Systems, 1994.
- Wasteland 2*, inXile Entertainment, Deep Silver, 2014.

Travaux

- ALLAIN, Sébastien, « Métalepses du récit vidéoludique et reviviscence du sentiment de transgression », *Sciences du jeu* (en ligne), n° 9, 2018.
URL : <https://journals.openedition.org/sdj/909>, consulté le 17/02/2021.
- ARJORANTA, Jonne, « Defining Role-Playing Games as Language-Games », *International Journal of Role-Playing*, n° 2, 2011, p. 3-17.
URL : <http://www.ijrp.subcultures.nl/wp-content/issue2/IJRPissue2-Article1.pdf>, consulté le 16/02/2021.

- ARSENAULT, Dominic, « Des typologies mécaniques à l'expérience esthétique. Fonctions et mutations du genre dans le jeu vidéo », thèse en histoire de l'art et études cinématographiques, Université de Montréal, 2011.
- BARNABE, Fanny, *Narration et jeu vidéo. Pour une exploration des univers fictionnels*, Liège, Presses universitaires de Liège, 2018.
URL : <http://books.openedition.org/pulg/2613>, consulté le 17/02/2021.
- BARTLE, Richard, « Hearts, Clubs, Diamonds, Spades: Players Who Suit MUDs », *The Journal of Virtual Environments*, vol. 1, n° 1, 1997.
URL : <https://www.hayseed.net/MOO/JOVE/bartle.html>, consulté le 16/02/2021.
- BARTON, Matt & STACKS, Shane, *Dungeons and Desktops: The History of Computer Role-Playing Games*, 2^e éd., Londres et New York, CRC Press, 2019.
- BLEI, David, « Topic Modeling and Digital Humanities », *Journal of Digital Humanities* (en ligne), vol. 2, n° 1, 2012.
URL : <http://journalofdigitalhumanities.org/2-1/topic-modeling-and-digital-humanities-by-david-m-blei/>
- BLEI, David, NG, Andrew & JORDAN, Michael, « Latent Dirichlet Allocation », *Journal of Machine Learning Research*, vol. 3, 2003, p. 993-1022.
- BOURGEOIS, Nicolas, COTTRELL, Marie, LAMASSÉ, Stéphane & OLTEANU, Madalina, « Search for Meaning Through the Study of Co-occurrences in Texts », communication au colloque *International Work-Conference on Artificial Neural Networks, Palma de Majorque, juin 2015*, document déposé en ligne, 2015.
URL : <https://hal.archives-ouvertes.fr/hal-01519217/document>.
- CHAMPION, Erik, « Roles and Worlds in the Hybrid RPG Game of Oblivion », *International Journal of Role-Playing*, n° 1, 2009, p. 37-52.
URL : http://www.iirp.subcultures.nl/wp-content/uploads/2009/01/champion_roles_worlds_in_oblivion.pdf, consulté le 16/02/2021.
- CHANG, Jonathan, BOYD-GRABER, Jordan, WANG, Chong, GERRISH, Sean & BLEI, David M., « Reading Tea Leaves: How Humans Interpret Topic Models », in *NIPS'09: Proceedings of the 22nd International Conference on Neural Information Processing Systems (Vancouver, December 2009)*, Red Hook, Curran Associates, 2009, p. 288-296.
URL : <https://proceedings.neurips.cc/paper/2009/file/f92586a25bb3145facd64ab20fd554ff-Paper.pdf>
- CHATEAURAYNAUD, Francis & DEBAZ, Josquin, « Prodiges et vertiges de la lexicométrie », billet de blog (*Socio-informatique et argumentation*), 23 décembre 2010.
URL : <https://socioargu.hypotheses.org/1963#more-1963>.
- COTTRELL, Marie, OLTEANU, Madalina, ROSSI, Fabrice & VILLA-VIALANEIX, Nathalie, « Self-Organizing Maps, Theory and Applications », *Revista de Investigacion Operacional*, vol. 39, n° 1, p. 1-22, 2018.
- DETERDING, Christoph Sebastian & ZAGAL, José P., « The Many Faces of Role-Playing Game Studies », in Christoph Sebastian Deterding & José P. Zagal (dir.), *Role-Playing Game Studies*, New York, Routledge, 2018, p. 1-16.
- DOZO, Björn-Olav, « Pour une histoire polyphonique du jeu video », in LIÈGE GAME LAB, *Culture vidéoludique !*, Liège, Presses universitaires de Liège, 2019.
- ENSSLIN, Astrid, « Video Games as Unnatural Narratives », in Mathias FUCHS (dir.), *Diversity of Play*, Lunebourg, Meson Press, 2015, p. 41-72.

- FOWLER, Alastair, *Kinds of Literature: An Introduction to the Theory of Genres and Modes*, Cambridge, Harvard University Press, 1982.
- GENETTE, Gérard, *Palimpsestes. La littérature au second degré*, Paris, Seuil, 1982.
- GRACE, Lindsay D., « A Linguistic Analysis of Mobile Games: Verbs and Nouns for Content Estimation », communication au colloque *FDG 2014: 9th International Conference on the Foundations of Digital Games, Fort Lauderdale, avril 2014* (en ligne), 2014.
URL : https://www.fdg2014.org/papers/fdg2014_wip_07.pdf, consulté le 08/03/2022.
- HARRIGAN, Pat & WARDRIP-FRUIIN, Noah (dir.), *Second Person: Role-Playing and Story in Games and Playable Media*, Cambridge, MIT Press, 2007.
- JARA, David & TORNER, Evan, « Literary Studies and Role-Playing Games », in Christoph Sebastian Deterding & José P. Zagal (dir.), *Role-Playing Game Studies*, New York, Routledge, 2018, p. 265-282.
- JAUSS, Hans Robert, « Littérature médiévale et théorie des genres », in Gérard GENETTE, Hans Robert JAUSS, Jean-Marie SCHAEFFER, Robert SCHOLE, Wolf Dieter STEMPEL, Karl VIËTOR, *Théorie des genres*, Paris, Seuil, 1986, p. 37-76.
- KOHONEN, Teuvo, « Self-Organized Formation of Topologically Correct Feature Maps », *Biological Cybernetics*, vol. 43, n° 1, 1982, p. 59-69.
- MANTYLA, Myka V., CLAES, Maelick & FAROOQ, Umar, « Measuring LDA Topic Stability from Clusters of Replicated Runs », in *ESEM'18: Proceedings of the 12th ACM/IEEE Symposium on Empirical Software Engineering and Measurement (Oulu, October 2018)*, New York, Association for Computing Machinery, p. 1-4.
URL : <https://doi.org/10.1145/3239235.3267435>
- MAYAFFRE, Damon, « Les corpus réflexifs : entre architextualité et hypertextualité », *Corpus* (en ligne), n° 1, 2002.
URL : <http://corpus.revues.org/11>, consulté le 04/08/2021.
- NELIDE-MOUNIAPIN, Bernadette, « Exemple d'énonciation dans un jeu vidéo », in Sébastien Genvo (dir.), *Le game design de jeux vidéo. Approches de l'expression vidéoludique*, Paris, L'Harmattan, 2005, p. 239-251.
- PEPE, Felipe (dir.), *The CRPG Book: A Guide to Computer Role-Playing Games, v2.0* (en ligne), 2019.
URL : <https://crpgbook.wordpress.com/>, consulté le 16/02/2021.
- PERRON, Bernard, *The World of Scary Video Games: A Study in Videoludic Horror*, Londres, Bloomsbury Academic, 2018.
- PETERSON, Jon, « Precursors », in Christoph Sebastian Deterding & José P. Zagal (dir.), *Role-Playing Game Studies*, New York, Routledge, 2018, p. 55-62.
- ROSSI, Fabrice, VIALANEIX, Nathalie & HAUTEFEUILLE, Florent, « Exploration of a Large Database of French Notarial Acts with Social Network Methods », *Digital Medievalist* (en ligne), vol. 9, 2014.
DOI : <http://doi.org/10.16995/dm.52>
- RYAN, James O., KALTMAN, Eric, MATEAS, Michael & WARDRIP-FRUIIN, Noah, « What We Talk about When We Talk about Games: Bottom-Up Game Studies Using Natural Language Processing », communication au colloque *FDG 2015: 10th International Conference on the Foundations of Digital Games, Pacific Grove, juin 2015* (en ligne), 2015.
URL : https://www.fdg2015.org/papers/fdg2015_paper_47.pdf
- RYAN, Marie-Laure, « Metaleptic Machines », *Semiotica*, n° 150, 2004, p. 439-469.

- SALER, Michael, *As If: Modern Enchantment and the Literary Prehistory of Virtual, Reality*, Oxford, Oxford University Press, 2012.
- SCHIFFRIN, Deborah, *Discourse Markers*, Cambridge, Cambridge University Press, 1987.
- SCHULES, Douglas, Peterson, Jon & Picard, Martin, « Single-Player Computer Role-Playing Games », in Christoph Sebastian DETERDING & José P. ZAGAL (dir.), *Role-Playing Game Studies*, New York, Routledge, 2018, p. 107-129.
- THERRIEN, Carl & JULIEN, Mikaël, « “Pour obtenir de l’aide, appuyez sur X”. La montée du paradigme d’assistance dans le design du jeu vidéo », *Sciences du jeu* (en ligne), n° 4, 2015.
URL : <https://journals.openedition.org/sdj/508>, consulté le 27/03/2021.
- THERRIEN, Carl, « Inspecting Video Game Historiography Through Critical Lens: Etymology of the First-Person Shooter Genre », *Game Studies* (en ligne), vol. 15, n° 2, 2015.
URL : <http://gamestudies.org/1502/articles/therrien>, consulté le 27/03/2021.
- TRICLOT, Mathieu, « Les lendemains qui chantent : une histoire de l’avenir des jeux vidéo », communication au colloque *La presse de jeu vidéo francophone*, 2016.
- WANG, Xuerui & MCCALLUM, Andrew, « Topics over Time: A Non-Markov Continuous-Time Model of Topical Trends », in *KDD’06: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (Philadelphia, August 2006)*, New York, Association for Computing Machinery, 2006, p. 424-433.
DOI : <https://doi.org/10.1145/1150402.1150450>
- WILLIAMS, J. Patrick, KIRSCHNER, David, MIZER, Nicholas & DETERDING, Sebastian, « Sociology and Role-Playing Games », in Christoph Sebastian DETERDING & José P. ZAGAL (dir.), *Role-Playing Game Studies*, New York, Routledge, 2018, p. 227-244.
- ZAGAL, José P. & ALTIZER, Roger, « Examining “RPG Elements”: Systems of Character Progression », communication au colloque *FDG 2014: 9th International Conference on the Foundations of Digital Games, Fort Lauderdale, avril 2014* (en ligne), 2014.
URL : https://www.fdg2014.org/papers/fdg2014_paper_38.pdf, consulté le 16/02/2021.
- ZAGAL, José P. & TOMURO, Noriko, « Cultural Differences in Game Appreciation: A Study of Player Game Reviews », communication au colloque *FDG 2013: 8th International Conference on the Foundations of Digital Games, Chania, mai 2013* (en ligne), 2013.
URL : http://www.fdg2013.org/program/papers/paper12_zagal_tomuro.pdf, consulté le 03/08/2021.
- ZAGAL, José P., TOMURO, Noriko & SHEPITSEN, Andriy, « Natural Language Processing in Game Studies Research: An Overview », *Simulation & Gaming*, vol. 43, n° 3, 2011, p. 356-373.