

Littérature d'anticipation et Humanités Numériques

16-17 mars 2023, Lyon

Colloque organisé par Claire Barel-Moisan (CNRS, UMR IHRIM) et Delphine Gleizes (Université Grenoble Alpes, UMR Litt&Arts)

Ce colloque est conçu dans le cadre d'une collaboration entre l'ANR Numapresse (pôle IHRIM), le projet « PhraséoAnticipation » porté par Litt&Arts et le LIDILEM (Université Grenoble Alpes), et l'ANR Anticipation.

Par la multiplicité de ses désignations, la diversité des supports éditoriaux et des publics visés, la variété de ses orientations esthétiques, la littérature d'anticipation constitue un exemple particulièrement riche pour analyser les mécanismes de construction d'un genre littéraire. Comme pour tout autre genre littéraire, les critères de reconnaissance du genre (voire les termes choisis pour le nommer) sont éminemment variables en fonction des instances concernées : auteurs, éditeurs, critiques, lectorat. On peut notamment distinguer d'une part des critères externes d'identification du genre, comme les caractéristiques des supports de publication (éditeurs spécialisés, collections spécifiques), la réception critique (mise en avant de termes stratégiquement choisis pour désigner le genre, sélection de grands auteurs comme figures tutélaires du genre, pratiques de légitimation par la création de prix littéraires). D'autre part, on peut repérer les traits communs du genre à travers des critères internes aux œuvres : des thématiques spécifiques (voyages réalisés à l'aide de machines extraordinaires, dangers des sciences, projection dans le futur), des structures narratives caractéristiques, des personnages types (ingénieur, journaliste, scientifique, explorateur), ou des traits stylistiques (usage des néologismes, place de la description). Quels que soient les critères choisis, on constate à la fois un effet de reconnaissance de caractéristiques communes des œuvres d'anticipation, bien relevées par la critique et les lecteurs, et en même temps un mouvement d'étoilement, les auteurs explorant des directions nettement divergentes.

Le colloque « Littérature d'anticipation et Humanités Numériques » se propose de confronter trois méthodologies s'appuyant sur les Humanités Numériques pour mettre en évidence les caractéristiques de la littérature d'anticipation que chacune de ces approches fait apparaître. Il s'appuiera donc sur les trois outils suivants :

- **la base de données « Anticipation »** avec les visualisations produites à partir des résultats des requêtes formulées sur son moteur de recherche avancée
- **la « générothèque Numapresse »** : apprentissage automatique de classification des genres élaboré par Pierre-Carl Langlais dans le cadre de l'ANR Numapresse
- **le « Lexicoscope »**, outil d'exploration linguistique, lexicologique et morphosyntaxique de corpus, produit par le LIDILEM à l'Université de Grenoble.

L'enjeu sera non seulement d'aboutir à une meilleure connaissance de l'anticipation et à une analyse fine de certains de ses traits génériques, mais aussi d'utiliser l'anticipation comme un « corpus test » pour adopter une approche réflexive sur les apports et les biais des Humanités Numériques.

La première journée du colloque débutera ainsi avec une session axée sur les méthodologies. Chaque outil numérique sera présenté à travers ses modalités de

fonctionnement, les corpus sur lesquels il s'appuie, en fournissant des exemples des résultats qu'il permet d'obtenir, en pointant les éventuelles difficultés de sa mise en œuvre ainsi que les biais inévitables (ou éventuellement corrigibles) liés à cette méthodologie. Le projet est donc d'ouvrir le colloque avec un regard réflexif sur la mobilisation des différentes approches d'Humanités numériques mises en œuvre, en débattant de leurs atouts et de leurs limites, avant d'entrer dans le détail de leurs applications, dans les communications qui suivront.

Les communications pourront mobiliser **l'un de ces trois outils d'Humanités Numériques** pour mettre en évidence et explorer une caractéristique spécifique de la littérature d'anticipation. Elles pourront également **travailler directement sur des récits d'anticipation** pour analyser une des caractéristiques du genre – et ce, sans s'appuyer sur des approches digitales. L'objectif du colloque sera précisément de **confronter les images de l'anticipation et les différentes délimitations du genre que chacune de ces approches produit**.

Le **corpus envisagé dans le cadre du colloque** est celui de la « littérature d'anticipation » au sens où cette désignation générique a été mobilisée dans le cadre du programme ANR « Anticipation ». Nous avons étudié un corpus de littérature française et francophone défini selon les critères suivants : il s'agit de récits fictionnels publiés entre 1860 et 1940, intégrant une conjecture rationnelle (c'est-à-dire un décrochage par rapport à l'univers référentiel du lecteur) qui se déploie dans trois domaines principaux – la projection dans le futur, la mise en œuvre d'une science imaginaire, ou la description de sociétés imaginaires. Il suffit de remplir l'un de ces trois critères pour entrer dans le corpus que nous avons nommé « Anticipation ». Un grand nombre des récits d'anticipation se déroulent par conséquent dans une temporalité actuelle et non dans le futur. Sur la période s'étendant de 1860 à 1940, nous avons recensé 2400 œuvres correspondant à cette définition de la littérature d'anticipation. Lors du colloque, certaines communications pourront également confronter le corpus de la littérature d'anticipation à des corpus de contraste, comme celui de la littérature de science-fiction.

Présentation des trois outils numériques exploités lors du colloque

- La **base de données « Anticipation »** a été constituée grâce à la lecture et l'analyse détaillée d'un peu moins de 500 œuvres (romans et nouvelles). Une série d'informations a été relevée au fil de la lecture de chaque œuvre : données spatio-temporelles de la fiction, références scientifiques, éléments de science imaginaire, références intertextuelles, typologie des personnages, traits de la société imaginaire décrite, etc. D'autres informations sont issues de recherches sur les **supports éditoriaux** de chaque œuvre (rééditions, illustrations, types d'éditeurs et de collections, adaptations, traductions). Enfin, un troisième domaine exploré dans cette base de données porte sur la **réception critique** de chaque œuvre depuis sa parution jusqu'à nos jours, ainsi que sur les **catégories génériques** qui ont été employées par la critique, par les éditeurs ou par l'auteur lui-même pour désigner son œuvre.

Un moteur de recherche avancée permet d'interroger la base de données en fonction de n'importe lequel de ces critères, et les réponses pour ces requêtes peuvent ensuite être **visualisées** grâce à des **graphiques générés en direct en ligne** : histogrammes, courbes, diagrammes circulaires, réseaux d'auteurs, nuages de mots, carte interactive des lieux de la fiction, frise chronologique des éditions, etc.

La base de données propose dans de nombreux cas un lien vers le texte en ligne, mais contrairement aux deux autres outils numériques mobilisés dans le colloque, elle ne propose **pas d'outil de recherche plein texte dans les œuvres**.

- Le « **Lexicoscope** » est un outil de **fouille et de visualisations lexicologiques et morphosyntaxiques** développé par le LIDILEM, Laboratoire de Linguistique et de Didactique des langues étrangères et maternelles de l'Université de Grenoble. À partir de corpus de textes annotés en dépendances syntaxiques, le Lexicoscope permet, entre autres, de mettre en évidence des unités polylexicales formant des motifs phraséologiques récurrents qui apparaissent comme spécifiques au corpus envisagé. Cet outil a déjà été exploité sur un corpus d'œuvres de science-fiction contemporaines (104 romans analysés dans le projet PhraseoRom). Un **corpus test de littérature d'anticipation** est en cours de constitution pour engager l'exploration de motifs lexicosyntaxiques récurrents révélateurs de phénomènes phraséologiques et stylistiques.

- La « **Générothèque Numapresse** » est une bibliothèque de modèles qui permettent de prédire le genre d'un texte (ou d'autres formes de classification), en s'appuyant sur des **corpus d'entraînement** qui ont servi à la construction du modèle. Il s'agit de **classification automatique par apprentissage supervisé**. À l'issue de l'entraînement du modèle, des mots principaux (« **main tokens** ») apparaissent comme récurrents et spécifiques : ces groupes de 7 « main tokens » permettent de définir un genre donné. Par exemple, le modèle des genres journalistiques (1840-1860) a mis en évidence 23 types de textes journalistiques : débats politiques, débats judiciaires, roman-feuilleton, critique musicale, etc. En proposant alors un texte choisi aléatoirement, l'outil fournira une **prédiction de son statut générique** (avec un pourcentage d'attribution), comme 70% roman-feuilleton, 20% débat judiciaire, 10% fait-divers. Cet outil a notamment été entraîné par Pierre-Carl Langlais et Matthieu Letourneux pour comparer les « main tokens » d'un corpus de romans d'aventures et ceux d'un corpus de romans d'anticipation.

Chaque communication présentée lors du colloque pourra donc choisir de mobiliser l'un de ces outils ou d'étudier directement un corpus de textes d'anticipation pour **analyser une caractéristique du genre** ou une **thématique donnée**. L'objectif est d'observer les **recoupements et les divergences de caractérisation de la littérature d'anticipation** que ces études de cas feront apparaître. En diachronie, on pourra s'intéresser aux différentes **périodisations** du genre en fonction de l'objet de la communication. En synchronie, on pourra repérer des caractéristiques dominantes, analyser en détail des **sous-corpus d'anticipation en fonction des supports, des thématiques** (par exemple les romans de guerre future, les romans de monde perdu ou les récits de fin du monde).

Les propositions de communication (environ 2000 signes) sont attendues pour le 15 décembre 2022 aux adresses suivantes : claire.barel-moisan@ens-lyon.fr et delphine.gleizes@univ-grenoble-alpes.fr