**Françoise Rubellin, professeur à l'Université de Nantes, CETHEFI/L'AMO**

# Spectacles des Lumières et éclairage numérique :

# de nouveaux outils pour l'histoire culturelle

Contexte :

Au siècle des Lumières, le théâtre ne se résume pas aux célèbres pièces de Marivaux et de Beaumarchais. Comment se fait-il que l'histoire littéraire et plus largement l'histoire culturelle aient occulté la multitude de pièces le plus souvent accompagnées de musique et de danses, un répertoire extrêmement riche et protéiforme ? Pièces par écriteaux, par pantomime, en monologue, en jargon, pièces bilingues, parodies, parades etc ? Ces spectacles variés se développent dans les théâtres de la Foire (Foire Saint-Germain et Foire Saint-Laurent à Paris) pendant tout le XVIIIème siècle. Il s'agit de théâtres non-officiels, qui ne bénéficient pas d'un privilège contrairement à la Comédie-Française et à l'Opéra. De ce fait ils sont la cible d'innombrables procès. De même, la Comédie-Italienne de Paris, dépourvue de privilège, est souvent attaquée par lees deux grands théâtres parisiens. Ces théâtres non-officiels ou non-privilégiés de l'âge d'or du théâtre ont longtemps pâti de la méconnaissance, de la sélection, de la difficulté d'accès aux sources manuscrites, et de la non prise en compte d'une masse d'archives administratives. Nous voudrions montrer comment l'ère du numérique a commencé à changer la face de cette époque de spectacles : de grands précurseurs, Barry Russell (fondateur du site foires.net), David Trott et Jeff Ravel fondèrent en 2002 le site CESAR (Calendrier Electronique des Spectacles de l'Ancien régime) qui fut vite fréquenté par les chercheurs du monde entier. A notre tour, nous avons développé deux projets, l'un terminé, l'autre en cours, dont nous montrerons en ligne, en image et en musique les capacités.

## I – Un nouvel outil numérique depuis 2012 : THEAVILLE

Un programme ANR obtenu et développé en 2008-2011, nous a permis de construire une base de données innovante, theaville.org, qui regroupe théâtre et musique. Le but du programme POIESIS *(Parodies d'Opéra : Intertextualité, Etablissement des Sources Interdisciplinaires des Spectacles sous l'Ancien Régime)* était de mettre au jour un vaste pan du patrimoine culturel considéré au XVIII$^{e}$ siècle comme « non officiel », les parodies dramatiques d'opéra. Elles souffraient depuis leur naissance d'un discrédit lié à leur statut d'œuvres mineures (en marge des grandes créations de l'Académie Royale de Musique), à la méconnaissance d'une grande partie de ces productions (seuls 40% des textes étaient édités et souvent sans leur musique) et à la difficulté d'appréhender des pièces mettant en jeu plusieurs degrés de compréhension. Pourtant, l'étude du phénomène de la parodie d'opéra permet non seulement de mieux comprendre les chefs-d'œuvre de l'art lyrique, mais surtout de mieux connaître les conditions de création sous l'Ancien Régime. Concrètement, le programme POIESIS avait pour objectif d'étudier et de publier en ligne l'ensemble du corpus des parodies d'opéra avec leur musique, les vaudevilles – ces airs issus de la musique savante ou populaire – et de créer une base littéraire et musicale, THEAVILLE (Théâtre et Vaudeville).

Au terme de 4 ans de travail collaboratif et interdisciplinaire, nous avons ouvert la base de données *THEAVILLE* (http://theaville.org) qui contient, outre le texte de toutes les parodies connues, plus de deux mille partitions (fichier son et PDF). C'est une masse de données pour les chercheurs en littérature, art du spectacle et musicologie, mais l'accès gratuit à la base THEAVILLE a aussi pour but de permettre aux artistes de récupérer chez eux les textes et leur musique de pièces rares, comiques, originales et souvent subversives. Cet outil numérique a débouché sur quelques recréations scéniques (spectacles en parlé-

chanté, par écriteaux, ou pour marionnettes) et nous espérons que les résultats encourageront la recréation de spectacles dans la veine de ceux qui ont été réalisés pendant le projet.

Pendant le symposium, nous montrerons en direct quelques cas d'utilisation de l'outil THEAVILLE : par exemple comment on peut entendre un air dont le titre figure dans une pièce, comment on peut trouver un air qu'on cherche sur le clavier électronique, ou bien découvrir un air populaire grivois caché derrière un titre religieux.

La construction de la base THEAVILLE marque une avancée dans la recherche avec l'analyse et la constitution d'un corpus inédit et permet de mieux appréhender la diversité du théâtre comique et l'évolution des genres lyriques français au XVIIIe siècle. Elle invite à envisager l'extension du corpus à d'autres genres de théâtre musical, notamment l'opéra-comique ou les parodies d'opéra qui se développent au XIXe siècle.

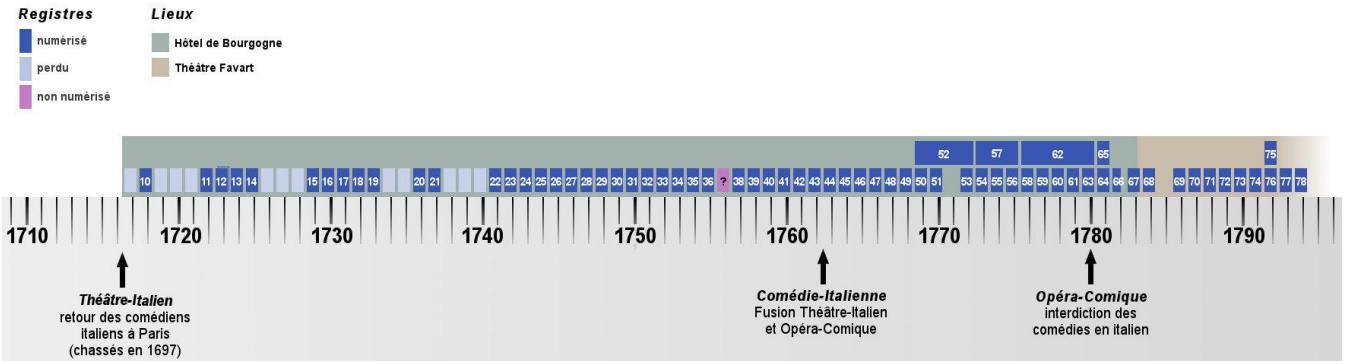
**II – Un outil en construction : REGITAL (les registres du Théâtre-Italien)**

Le centre CETHEFI de l'université de Nantes (Centre d'études des théâtres de la Foire et de la Comédie Italienne) a acquis la numérisation de tous les registres de la Comédie-Italienne de Paris pour le XVIIIe siècle, de 1717à 1793 : il s'agit de plus de 70 cahiers d'archives qui totalisent 27544 pages et qui consignent jour après jour toutes les pièces jouées, le nombre de spectateurs, les recettes, les dépenses etc.

On comprend devant l'ampleur de ces archives, que personne n'ait jamais réussi à les éditer ni à les exploiter (un chercheur américain en 1960, Cl. Brenner, a relevé la liste des titres et des entrées, ce qui ne représente que l'utilisation de 5% des informations des registres). Or, conscient des immenses possibilités offertes par l'ère numérique , le CETHEFI s'est proposé de tenter de relever un défi technologique en faisant collaborer les spécialistes du théâtre avec des chercheurs réputés en STIC pour travailler à la création d'un outil de reconnaissance de l'écriture manuscrite, afin de faciliter la saisie de milliers de pages pour la construction d'une vaste base de données interactive. Il s'agira à terme de pouvoir mettre en relation tous les éléments et événements qui concernent un spectacle de la Comédie-Italienne, depuis le coût de la production (registres), les accessoires utilisés, les acteurs employés, la composition sociale du public, les airs (retrouver les partitions et les instruments utilisés par l'orchestre), les danses (retrouver les types de danse), le texte de la pièce (déchiffrer et étudier les manuscrits, analyser les marques des censeurs).

De plus, alors qu'il existe des chronologies sous forme de calendrier en ligne pour les deux théâtres privilégiés, (pour l'Opéra : http://chronopera.free.fr/ dir. M. Noiray et S. Serre et pour la Comédie-Française, http://hyperstudio.mit.edu/projects/comedie-francaise-registers-project/ J. Ravel, Ch. Biet, projet RCF ANR Corpus), il n'en va pas de même pour les deux autres lieux de spectacle de Paris.

Les registres de la Comédie-Italienne sont lacunaires (il manque quatorze années), comme le montre le schéma suivant que nous avons réalisé pour rendre compte de l'état de la situation :



Sur le plan technologique, comment construire cet outil ? La collaboration avec l'équipe IVD de l'IRCCyN en cours permet de préciser comment le travail va se mener.

Il s'agira **dans un premier temps** d'utiliser les travaux existants sur l'analyse de documents pour catégoriser automatiquement les différents types de pages. Les images ci-dessous montrent les grandes catégories de types de pages découvertes par un échantillonnage aléatoire du corpus des registres de la Comédie-Italienne : des pages d'introduction (environ 1 % du corpus), des comptes quotidiens (environ 30 à 90 % des pages, suivants les années), des comptes mensuels (5 à 30 % suivant les années), des pages blanches (environ 4 % des pages) ; auxquelles s'ajoutent certaines pages volantes non reliées comprenant des notes manuscrites. L'analyse plus approfondie permettra sûrement la découverte de nouvelles catégories ou sous-catégories vu l'évolution des habitudes pendant la période. Cette première analyse permettra une indexation simple des documents pour faciliter la navigation au sein de la collection.

Introduction

Comptes quotidiens

Comptes mensuels

Pages vierges

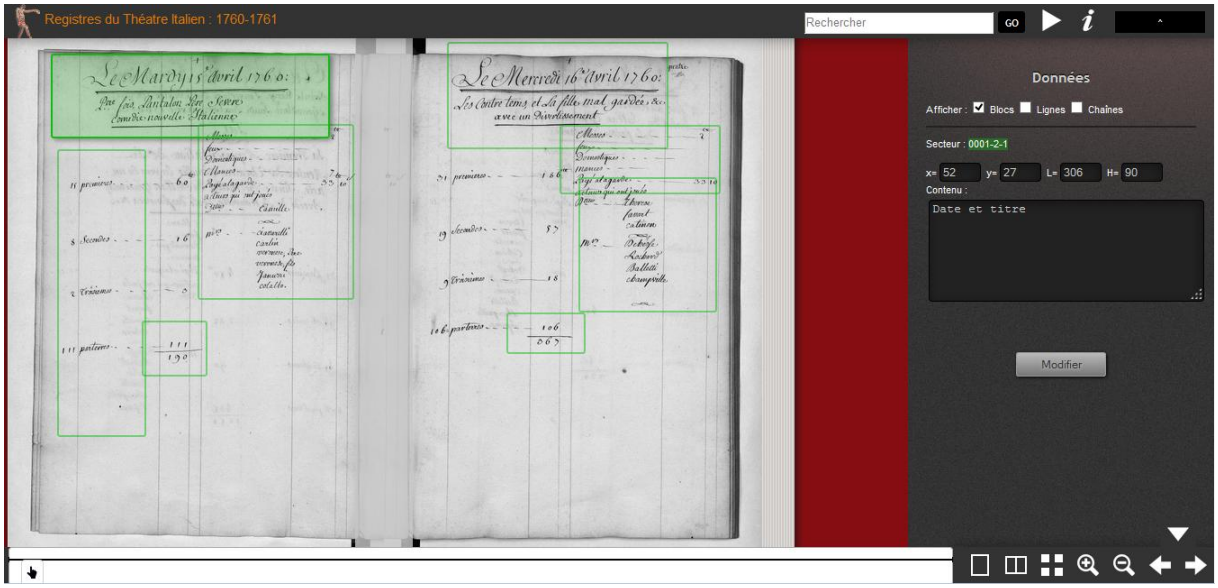
**Dans un second temps** on se focalisera sur le contenu des documents, ce qui représentera la partie principale du projet. En effet les techniques de reconnaissance de l'écriture actuelle trouvent leur limite dans le cas de l'écriture manuscrite ancienne, très irrégulière et dans un contexte très bruité.

Malgré ces difficultés, on recherchera les traces de stabilité locale (sur quelques années), dépendant beaucoup du scripteur. En effet, celui qui tient les comptes et écrit les notes est le même pendant quelques années, gardant son style d'écriture, son vocabulaire ou sa mise en page.

**Une plate-forme de consultation et d'annotation** de la collection. Le contenu de cette plate-forme sera initialisé avec les données manuelles existantes déjà collectées et mis à disposition des

utilisateurs (chercheurs, ingénieurs d'étude, bénévoles ...). Ensuite cette plate-forme servira d'un côté à alimenter en données réelles les mécanismes de recherche et d'exploration, ainsi que de fouille de données mis en œuvre pour le projet, et d'un autre côté elle sera complétée par les résultats de la catégorisation et reconnaissance automatique des pages de la collection.

La figure suivante montre un prototype (non fonctionnel pour l'instant, basé sur l'outil BookReader) de ce que pourrait être cette plate-forme, avec une partie gauche pour visualiser le document et localiser les différents champs dans chaque page ; la partie droite permettrait de consulter la transcription automatique ou l'annotation manuelle. En haut se trouve un champ de recherche et en bas une barre de navigation permettant de poser des signets (manuels ou résultant d'une recherche).




**A terme : une redécouverte de l'économie des spectacles**

On pourra faire des bilans pour globalement cerner : l'importance de la musique et de la danse (salaires des danseurs, des musiciens, des copistes de partition, présence des instruments) ; l'apparition des droits d'auteur et leur courbe ; le coût des non-artistes du « monde invisible de l'art » (domestiques, gardes, perruquiers, habilleuses, porteurs de chandelles, souffleurs, menuisiers, tapissiers, imprimeurs d'affiches etc.). À partir des registres de la Comédie-Italienne, des éléments important d'information seront portés à la connaissance et entrés dans la base de données qui permettra requêtes et statistiques :
– la **répartition du public**, selon les places en première loges, secondes, troisièmes, ou au parterre. Il sera possible de savoir quelles pièces attirent le plus le public le moins élevé socialement. Les registres mentionnent parfois des « personnalités », des grands de la Cour qui viennent et font des dons ;
– **les dépenses régulières :** les bougies, les gants, les tissus, les « arguments » imprimés (on fait imprimer le résumé des pièces en italien, car le public ne comprenait plus la langue) ;
– **les salaires** des **auteurs** (source de renseignements extrêmement précieuse, même si elle n'est pas portée régulièrement), des **acteurs** (mieux comprendre le système de parts), des **domestiques**, des

« **gardes** » (le système de police des théâtres sera à étudier), du **souffleur** (contrairement à l'idée reçue, les comédiens italiens ont besoin d'un souffleur ; on voit dans les registres le prix de la fosse qu'on lui a construite).

Notre projet est innovant dans son sujet (réévaluation de la **face non académique** de la culture dramatique), ses méthodes (approche qualitative et quantitative ; ressources de **milliers de pages** exhaustivement exploitées grâce à la construction d'un outil numérique) et ambitieux dans ses objectifs (sur le plan technologique comme sur le plan scientifique). L'étude de ces espaces de création offre un magnifique exemple de cas où s'entrelacent **tradition populaire et tradition savante**. C'est un **défi matériel et herméneutique** pour montrer un autre visage culturel du siècle des Lumières que celui académiquement consacré. Porteur d'enjeux aussi bien **esthétiques que sociaux**, il permet de poser des questions qui dépassent le fait littéraire, auquel on limite trop souvent l'étude du théâtre.Un tel projet place au centre l'usage du numérique au service de la patrimonialisation, et grâce à une collaboration forte entre les SHS et les STIC, toutes les données brutes déchiffrées et retranscrites pourront encourager et faciliter de futures études historiques, linguistiques et littéraires tirant profit de ces objets d'archive. De surcroît, l'exposition et la valorisation des données s'inscrivent dans la perspective du Web des données (Linked Open Data ou LOD) pour favoriser l'interconnexion avec d'autres sources de données et améliorer la visibilité et donc « l'usabilité » des données d'archives

Les utilisateurs des résultats pourront être aussi bien les chercheurs (théâtre, danse, musique, histoire, linguistes) que les amateurs de théâtre mais aussi tous les praticiens de spectacle vivant : acteurs, danseurs, musiciens, metteurs en scène, scénographes, machinistes… L'outil numérique qui sera développé pour la reconnaissance de l'écriture manuscrite des documents anciens constituera une avancée certaine dans le domaine. Il devrait pouvoir être appliqué ou adapté par d'autres chercheurs travaillant sur des champs voisins (archives militaires, archives d'état-civil, fonds patrimoniaux de bibliothèques…).

**Une réévaluation pédagogique**

Le projet offrira un nouveau matériel d'enseignement pour dévoiler le visage du patrimoine culturel franco-italien : moins d'idées fausses, plus de pratique. Le théâtre de l'Ancien Régime a très longtemps été appréhendé comme un ensemble de textes, au détriment de sa dimension spectaculaire. Les théâtres de la Foire et la Comédie-Italienne ne sont pas de la littérature mineure, mais le lieu d'un art majeur dans le spectaculaire où se coordonnent musique, danse et pantomime. On insistera sur la dimension interactive – et donc ludique – de la base de données, qui comportera aussi les fichiers sons (format MIDI) des airs utilisés dans les théâtres forains et italiens (lien avec Theaville.org ; on peut entendre et chanter soi-même les airs en lisant les textes) et des extraits vidéos. Nous espérons que le projet contribuera vraiment à une réévaluation de ce patrimoine qui entrera enfin dans les manuels scolaires.

**Conclusion**

Nous voudrions montrer que ce type d'entreprise, impensable avant l'ère numérique, permettra de jeter a un regard résolument neuf sur des formes émergentes de création et sur les mutations de l'économie du spectacle ; il comporte des enjeux aussi bien esthétiques que sociaux. C'est un défi matériel et herméneutique pour montrer un autre visage culturel du patrimoine dramatique des Lumières que celui qui lui est académiquement consacré.