

## MONDES ARABES ET COMPARATISME : PRATIQUES PEDAGOGIQUES

10 juin 2015, 9h30-17h Université Paris Ouest Nanterre La Défense, salle LR5



1. Le réseau « Littérature Générale et Comparée-Mondes Arabes » et les enjeux de la diffusion numérique.
2. La littérature arabe, la littérature « classique » et les « textes fondateurs ».
3. La littérature arabe et le canon littéraire contemporain – catégories, classements, brouillages.

Carole Boidin (Paris Ouest), Zoé Carle (Paris 3), Eve de Dampierre (Bordeaux 3), Elise Duclos (Paris Ouest), Ines Horchani (Paris 3), Tristan Mauffrey (Paris Diderot), Emilie Picherot (Lille 3), Claire Placial (Metz-Lorraine), Pablo Roza Candas (INALCO)

Organisation : Carole Boidin et Emilie Picherot - Contact : [cboidin@u-paris10.fr](mailto:cboidin@u-paris10.fr)

# Une approche méthodologique des humanités numériques appliquées à l'étude du corpus *aljamiado*.

PABLO ROZA CANDÁS  
Seminario de Estudios Árabo-Románicos  
Universidad de Oviedo

On peut définir, en peu de mots, l'*aljamía* comme la *variété islamique* de l'espagnol qui, du XVI<sup>e</sup> au XVII<sup>e</sup> siècles, utilise les caractères arabes à l'écrit. Cette langue arabisée est le produit de la minorité musulmane hispanophone qui avait déjà perdu sa connaissance de l'arabe et qui, afin de maintenir son identité vivante, a versé son patrimoine culturel islamique dans une langue romane. De cette façon, l'espagnol est la seule langue romane qui fait partie du *Sprachbund* islamique, selon la terminologie de Troubetzkoy (1930; 17-18), c'est-à-dire: «l'alliance linguistique déterminée par l'influence d'une langue sacrée», dans ce cas l'arabe, et formé par des langues de diverses origines, telles que le persan, l'ourdou, le berbère, le bosniaque, le grec, l'albanais ou le turc. La variété aljamiado représente ainsi le dernier effort des musulmans espagnols pour créer une identité islamo-hispanique dans un environnement non-arabe, parallèle à celles qui se sont produites en Orient, dans les territoires islamo-turques ou islamo-persanes.

De cette manière, la langue de la littérature aljamiado représente le moyen de communication d'une minorité plutôt religieuse qu'ethnique et, à cet égard, son caractère islamique déterminera aussi ses particularités. Par conséquent, comme il est arrivé dans d'autres langues européennes, africaines et asiatiques, l'utilisation de la graphie arabe donne à la langue le prestige nécessaire pour la transformer en un moyen digne de transmettre le savoir islamique.

Il s'agit ainsi d'une langue hybride: islamisée mais nettement hispanique; romane mais arabisée dans son lexique et sa syntaxe; dialectale et parfois archaïsante par rapport à la norme cultivée de la Castille du Siècle d'Or. Ce patrimoine textuel des Morisques est arrivé jusqu'à nos jours dans plus de 300 codex, beaucoup d'eux encore inédits, conservés, entre autres, dans les bibliothèques de Madrid, Paris, Londres, Rome, Florence, Alger ou La Valette. Fréquemment, sous la forme de mélanges, les Morisques réunissaient dans ces manuscrits des sujets et types littéraires divers comme la littérature narrative, les traités juridiques, la littérature eschatologique, la prose didactique, les recueils de croyances populaires, la poésie religieuse et profane, etc.

\*

Pendant les dernières décennies, les études d'aljamiado ont vu une croissance significative en ce qui concerne le nombre de monographies, de dissertations et d'articles consacrés à l'édition de cette production manuscrite. Parmi ces études, deux projets développés à l'Université d'Oviedo et lancés par Álvaro Galmés de Fuentes ont été décisifs pour l'avancée scientifique dans le domaine: la *Colección de literatura española aljamiado-morisca* (CLEAM) [<http://www.arabicaetromanica.com/cleam>] aujourd'hui une plate-forme bien consolidée pour l'édition de nouvelles compilations et d'études, et en ce qui concerne la lexicographie aljamiado, le *Glosario de voces aljamiado-moriscas* (GVAM), édité par la Bibliotheca

Arabo-Romanica et Islamica [[www.arabicaetromanica.com/biblioteca-arabo-romanica-1](http://www.arabicaetromanica.com/biblioteca-arabo-romanica-1)] de la même université asturienne

On doit préciser néanmoins que la production aljamiado continue toujours à être un domaine marginal dans les études hispaniques, cela étant dû en partie aux solutions ecclésiastiques, parfois très complexes, appliquées à ce corpus littéraire. L'un des aspects principaux de l'exploitation du corpus aljamiado (au-delà de l'intérêt littéraire) est, sans aucun doute, la connaissance approfondie de la langue quotidienne et de la variation dialectale des XVI<sup>e</sup> et XVII<sup>e</sup> siècles, puisque ces codex «marginiaux» reflètent souvent des aspects linguistiques qui apparaissent rarement dans les textes «chrétiens» du Siècle d'Or.

À cet égard, le projet que je présente ici, essaie de rendre plus accessible cette littérature hispano-musulmane en utilisant les nouvelles ressources des humanités numériques. Nous avons besoin d'outils précis qui nous permettent d'effectuer des éditions fiables ainsi que d'accéder rapidement aux textes. Des supports solides et une classification correcte de ces matériaux nous permettront de répondre à des questions jusqu'à présent étudiées d'une manière partielle comme celles relatives à la dimension variationnelle de la modalité aljamiado ou la filiation textuelle des différents manuscrits du corpus.

\*

Jusqu'à présent, la seule expérience de recherche dans le domaine des humanités numériques appliquées au corpus aljamiado se réduit à l'étude que en collaboration avec Raquel Suárez nous développons dans le cadre du projet «Portal Andrés de Poza» de l'Université de Deusto (Pays Basque) [<http://andresdepoza.com>]. Cette équipe de recherche, menée par Carmen Isasi, s'est consacrée à l'édition et à la publication numérique de textes multi-version, y compris des traductions dans certaines langues européennes. Dans ce projet à Deusto, nous avons exploré les possibilités d'une édition numérique des textes aljamiado sur la base d'une étude, encore en élaboration, des différents récits de la tradition narrative morisque. Comme résultat de cette première expérience, nous avons pu établir les fondements pour une édition numérique des textes romans en caractères arabes en utilisant une méthodologie adaptée et simple mais, à notre avis, très efficace pour ce type de textes.

À grands traits, le but de notre travail est d'étudier la tradition textuelle crypto-islamique en langue romane (castillan et aragonais) en réalisant l'édition numérique XML (*Extensible Mark-up Language*) de divers récits aljamiado, entre autres, la légende de *La doncella Carayona*. Sur la base de cette narration, nous nous proposons d'étudier les problèmes théoriques et pratiques liés à l'édition numérique, tels que la segmentation et l'alignement de versions ou le balisage pour lequel nous suivrons les standards généralement reconnus du TEI (*Text Encoding Initiative*) [<http://www.tei-c.org>], ainsi que les critères fixés par le réseau international CHARTA (*Corpus hispánico y americano en la red: Textos antiguos*) [<http://www.charta.es>] un projet global dans lequel sont rassemblées plusieurs équipes de recherche dans le domaine des humanités numériques.

\*

Pour notre projet on a choisi, donc, un récit appartenant à la littérature narrative aljamiado, *La doncella Carayona* (la histoire de «La jeune fille aux mains coupées»), une narration très populaire chez les Morisques, et dont on connaît également des versions (plus ou moins proches) dans d'autres traditions orientales, africaines et asiatiques. En ce qui concerne la tradition aljamiado, les versions connues sont au nombre de six, bien que dans notre projet nous n'en ayons pris en considération que cinq puisque l'une d'elles est une version de type abrégé. On a ainsi les versions des mss. Junta 57 et Junta 3 du CSIC de Madrid, ms. 5313 de la Biblioteca Nacional de España, ms. aljamiado de la Biblioteca Pública de Lleida et ms. 1944

de la Bibliothèque Nationale d'Algérie.

Notre projet comprend, donc, les phases de travail suivantes:

1) **La translittération des textes.**

Dû en partie au manque d'études précédentes sur l'édition numérique appliquée aux textes espagnols en caractères arabes, on a réalisé, comme première étape essentielle, une révision précise du système de translittération aljamiado (c'est-à-dire, la conversion des caractères arabes en caractères latins) afin d'adapter les textes à l'édition numérique proposée. Sur ce point, nous avons choisi le système CLEAM (Colección de Literatura Española Aljamiado-Morisca) de l'Université d'Oviedo, dans lequel on combine parfaitement la lisibilité avec une fidélité maximale aux caractéristiques graphématiques des manuscrits. De cette façon, nous obtenons une translittération semi-paléographique qui nous permettra de restituer à tout moment le texte en caractères arabes.

2) **La segmentation et l'alignement des textes.**

On a choisi dans cette phase du projet un logiciel qui nous permet d'aligner d'une manière semi-automatique les textes. Puisque le codex CSIC Junta 57 [<http://simurg.bibliotecas.csic.es/viewer/image/CSIC001227720/1>] contient la version la plus ancienne et la plus complète, on l'emploie comme texte base pour la segmentation. À ce point, la structure du texte (souvent de type dialogique) elle-même nous fournit les directrices pour cette segmentation. À partir de l'alignement avec les versions des autres manuscrits, nous pourrions identifier correctement les variations entre les différentes traditions.

3) **L'édition et le balisage.**

L'édition et le balisage, la troisième phase du projet, sont deux processus complémentaires et souvent simultanés; en même temps que l'on établit le texte, on ajoute une série de balises relatives, par exemple à des ratures du copiste, des détériorations du manuscrit, etc. Dans une seconde phase, on ajoute des nouvelles balises concernant la classification du lexique, les toponymes et anthroponymes, les expressions arabes, etc.

Puisque notre intérêt est fondamentalement linguistique, le balisage qu'on a initialement considéré s'est focalisé, d'une part, sur la notation d'éléments de la langue, comme les traits dialectaux notables; d'autre part, nous classons les aspects codicologiques, comme les déchirements du papier, ou quelques autres dommages physiques des manuscrits qui peuvent conditionner l'édition d'une façon ou d'une autre.

Dans une deuxième phase du processus de balisage, nous nous concentrerons sur le développement *ad hoc* de nouvelles balises à partir du langage TEI afin de noter d'autres particularités linguistiques des manuscrits aljamiado qui présentent une problématique associée, comme les phénomènes d'hybridation linguistique arabo-romane, des formes adventices et la possible subsistance de traits dialectaux arabes ou la variation graphématique du système aljamiado.

4) **La visualisation et la diffusion des résultats:**

Les résultats de notre projet seront mis à disposition dans la plate-forme du «Portal Andrés de Poza» de l'Université de Deusto [<http://andresdepoza.com>]. De cette manière, nous offrirons une édition numérique multiple, alignée de façon dynamique; c'est-à-dire, les différentes versions du texte sont alignées en fonction des options choisies par l'utilisateur. De la même

façon, l'utilisateur aura l'option de visualiser le texte amendé ou la version originale sans les interventions proposées par l'éditeur. Par ailleurs, toutes les informations concernant le document (la localisation, la description, la date, l'auteur, etc.) et l'édition (les responsables, les caractéristiques techniques, les critères, etc.) seront contenues dans l'en-tête à laquelle l'utilisateur pourra accéder en cliquant sur l'option *cabecera* d'un menu général que l'on trouvera à côté des différentes versions.

\*

Ce projet ne propose qu'une première étape vers la création à moyen terme d'un *Corpus numérique aljamiado* plus ample et dans lequel on rassemblera les divers types littéraires de la tradition morisque. En définitive, le but de notre travail est de mettre à disposition toute une série de matériaux qui seront d'une grande utilité pour des futures études, comme l'analyse de la variation linguistique, notamment en ce qui concerne les aspects diachroniques et diatopiques de la production textuelle aljamiado; la filiation et l'établissement des familles textuelles ou les études comparatives de la tradition aljamiado avec d'autres traditions.

Le projet est, bien entendu, ouvert à d'autres possibilités de développement et d'étude, comme, par exemple, la création d'indices conceptuels, à partir du balisage TEI, qui répondent à la structure du domaine sémantique et qui permettent l'accès aux textes à partir de concepts.

L'édition numérique alignée de textes ajoute sans doute une nouvelle valeur à l'édition d'un témoignage littéraire concret. De cette façon, on croit que le développement de cette méthodologie sera d'utilité pour l'édition d'autres traditions textuelles qui, comme l'aljamiado, se caractérisent par la diversité de leurs versions.