

# **De nouvelles perspectives pour l'étude scriptologique des manuscrits médiévaux : l'encodage en XML-TEI**

Camilla TALFANI

Université Toulouse Jean-Jaurès/ UPVM

La TEI (Text Encoding Initiative) est un puissant outil de plus en plus adopté dans le domaine des sciences humaines pour l'édition des textes en format numérique. La raison de ce choix consiste dans les différentes possibilités d'interrogation efficace d'un texte édité en TEI, comme par exemple l'extraction semi-automatique des informations pour la réalisation de bases de données.

En plus de la recherche lexicale et sémantique, le texte encodé en XML-TEI permet à l'heure actuelle d'effectuer l'analyse linguistique. Cette possibilité donnée par la TEI est assez récente et à découvrir.

En particulier, pour ce qui concerne l'étude de la *scripta* des manuscrits médiévaux de genre littéraire, le balisage des graphèmes simples et complexes permet de résoudre certains problèmes relatifs à la stratification linguistique, ainsi que, dans une phase ultérieure, la création de statistiques sur la composition linguistique des textes et sur les habitudes d'écriture du copiste. Plus précisément, par rapport à la stratigraphie de type synchronique il est possible, grâce aux balises, de rétablir la relation entre les phonèmes de la langue du document et ses différentes représentations graphématiques. D'autre part, au regard de la stratification en diachronie, les balises nous permettent de préciser les interprétations géolinguistiques des formes et, par conséquent, de les fournir à l'utilisateur ou bien de les repérer par traitement automatique et de les employer dans une étude complète et approfondie des différentes strates de la *scripta* du manuscrit.