

Une édition au format TEI de la première traduction française de *La Cité de Dieu*

Section 16

Bertrand Gaiffe et Béatrice Stumpf

Atilf UMR 7118 (CNRS & Université de Lorraine)

Introduction

La traduction de *La Cité de Dieu* de Raoul de Presles que nous éditons à l'ATILF grâce au financement européen du programme de recherche ERC intitulé « Histoire du lexique politique français » provient de l'un des 58 manuscrits qui nous sont parvenus de ce texte composé entre 1371 et 1375 : le manuscrit français BNF, fr. 22912 et 22913. Pour établir au mieux l'édition de ce texte imposant (894 folios), les éditeurs ont opté pour un encodage TEI qui présente l'immense avantage sur n'importe quel autre traitement de texte de mettre à leur disposition un certain nombre d'outils informatiques qui les assistent tout au long des différentes phases éditoriales. Néanmoins, le choix de publier sous forme papier a été impulsé par l'idée de léguer à la communauté scientifique un document pérenne. Car comme le souligne très justement M. Burghart qui, dans sa communication sur les humanités numériques intitulée « Digital Editions as the Myth of Sisyphus » (Burghart 2011), s'inspire du mythe de Sisyphe qu'elle transpose allégoriquement à la problématique de la pérennité des éditions disponibles sur le WEB, beaucoup d'éditions ainsi disponibles ne sont pas maintenues sur le long terme¹.

Il faut préciser que dans notre cas, la version imprimable de l'édition est obtenue automatiquement à partir de l'encodage en TEI².

L'encodage en tant qu'abstraction

Les éditions critiques traditionnelles, nous entendons par là les éditions papier, se présentent formellement comme un texte : le texte édité, sur lequel portent des notes. L'édition visée, dans sa forme papier, ne déroge pas à cette règle : on y trouve trois étages de notes relatives respectivement à l'établissement du texte, aux variantes lexicales et aux commentaires destinés à l'identification des sources citées. On y trouve également une introduction et trois index exhaustifs³.

L'encodage TEI de l'édition renverse cette perspective : là où l'édition papier voit au premier chef le texte édité et les notes sur ce texte, la version TEI voit au premier chef la transcription d'un manuscrit et des annotations qui concernent cette transcription. En particulier, les choix des éditeurs qui se présentent explicitement comme des annotations (des modifications) de la transcription. À titre d'exemple, une leçon du manuscrit de base jugée erronée par le transcripateur est annotée comme fautive et est remplacée par la leçon corrigée accompagnée du ou des sigles des manuscrits de contrôle qui justifient cette correction.

Pour dire les choses de façon simple, les opérations d'édition (résolutions d'abréviations, déglutinations, ajouts et corrections) sont entièrement explicites dans l'encodage.

De même, les index se présentent comme des annotations de segments de texte par des clefs plus abstraites que les entrées d'index données dans l'édition papier. Ainsi, dans l'encodage on peut avoir la clef d'identification abstraite « Cicéron », alors que dans l'index, mis à part les renvois, on trouvera, avec leurs références, les segments de texte correspondants à cette identification non pas à l'entrée « Cicéron » mais à l'entrée « Tulle » qui se trouve être la forme du texte la plus fréquente.

L'encodage pour obtenir une meilleure édition

Jusqu'à présent nous avons montré en quoi l'encodage en TEI est plus abstrait que la forme papier finalement visée. Nous souhaitons maintenant montrer en quoi cet encodage améliore la qualité de l'édition produite. Une partie des améliorations tient au caractère électronique de l'édition : qui dit fichier électronique dit possibilité d'utiliser des outils que la forme papier interdirait. Une seconde partie des améliorations tient spécifiquement à l'édition en TEI.

1 Une revue des projets TEI (cf. Burghart, Rehbein 2012) montre que 97 % des projets encodés en TEI ne visent que la publication sur le WEB.

2 Nous utilisons le logiciel de mise en page LaTeX comme format intermédiaire. Beaucoup de projets (surtout étrangers) n'utilisent pas la TEI, mais encodent directement en LaTeX. Un exemple français de taille significative est donné dans Rochebouet 2009.

3 Il y aura aussi un glossaire dans l'édition. Nous ne mentionnons pas cet aspect dans le résumé.

Outils liés au caractère électronique de l'édition

Parmi les outils utilisés pour améliorer l'édition, le lemmatiseur LGeRM développé par G. Souvay (Souvay, Pierrel 2009) tient une place de choix. Cet outil permet en effet de chercher automatiquement chaque mot du texte transcrit dans le dictionnaire du moyen français ; si ce mot est inconnu du lemmatiseur il appartient à l'éditeur de décider si on a affaire à une erreur de transcription, auquel cas la transcription doit être corrigée, ou si on a affaire à une incomplétude du DMF qui gagnerait à être enrichi. Inutile de préciser qu'un tel travail, à l'échelle d'un texte de cette taille serait inenvisageable sans automatisation, inutile aussi d'insister sur l'amélioration de la qualité du texte qui découle directement de cette vérification.

Outils liés à la forme spécifiquement TEI de l'édition

L'encodage en TEI participe au perfectionnement de la qualité de l'édition sous deux aspects. Le premier tient à la forme de l'édition papier finalement produite et le second tient à la qualité de l'édition proprement dite.

En ce qui concerne la qualité de l'édition papier, un bon exemple est probablement celui des index que nous avons déjà mentionnés. Dans l'édition produite, les entrées principales des index sont les formes les plus fréquentes du texte pour une clef donnée. Inutile de dire qu'il serait hors de question de prétendre produire à la main de tels index.

En ce qui concerne la qualité de l'édition, le schéma TEI impose des contraintes de forme aux éditeurs : par exemple, toute opération de correction impose une justification sur la base des manuscrits de contrôle contre le manuscrit transcrit. Cette contrainte force l'éditeur à préciser quels sont les manuscrits qui appuient la correction. Dans notre présentation, nous examinerons de façon plus détaillée et sur d'autres exemples en quoi l'encodage TEI, par les contraintes qu'il impose participe à l'amélioration de la qualité de l'édition.

Conclusion

Il peut paraître paradoxal, lorsqu'on vise *in fine* l'édition papier d'une édition critique de s'abstraire de la forme visée pour adopter une représentation abstraite. Outre l'avantage évident de permettre plusieurs supports d'édition (papier et web), nous montrerons dans notre communication que le choix d'une édition TEI, est pleinement justifié aux regards des contraintes qui guident les éditeurs et des outils auxquels ce format donne accès. Nous montrerons également que ce format abstrait permet d'engendrer automatiquement la forme papier traditionnellement commune à toutes les éditions de textes.

Bibliographie

Burghard (2011), « Digital Editions as the Myth of Sisyphus », TEI Members Meeting and Conference, Würzburg (Germany) 2011.

Burghart, Rehbein (2012), « The Present and Future of the TEI Community for Manuscript Encoding », *in* Journal of the Text Encoding Initiative. Issue 2, Feb 2012. URL : <http://jtei.revues.org/372> ; DOI : 10.4000/jtei.372

Gaiffe, Stumpf (2011), « A large scale critical edition : first translation of St Augustine's City of God by Raoul de Presle » TEI Members Meeting and Conference, Würzburg (Germany) 2011.

Rochebouet (2009), « D'une pel toute entiere sans nulle cousture. La cinquième mise en prose du Roman de Troie, édition critique et commentaire », thèse de doctorat de l'université Paris 4, 2009.

Souvay, Pierrel (2009), « LGeRM, Lemmatisation des mots en moyen français », *in* TAL (Traitement automatique des langues) vol. 50. 2009. URL : <http://halshs.archives-ouvertes.fr/halshs-00396452>