Indexing and Linking Text in a Large Body of Family Writings
Beatrice Dal Bo, Francesca Frontini, Giancarlo Luxardo, Agnès Steuckardt

To cite this version:
Beatrice Dal Bo, Francesca Frontini, Giancarlo Luxardo, Agnès Steuckardt. Indexing and Linking Text in a Large Body of Family Writings. Digital Humanities 2019, Jul 2019, Utrecht, Netherlands. 2019. hal-02187300

HAL Id: hal-02187300
https://hal.archives-ouvertes.fr/hal-02187300
Submitted on 17 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Indexing and Linking Text in a Large Body of Family Writings
Beatrice Dal Bo; Francesca Frontini; Giancarlo Luxardo; Agnès Steuckardt
Praxiling UMR 5267 - Montpellier, France

THE PROJECT
The Corpus 14 project aims at building a digital edition of correspondences of French soldiers in WW1. Writers are soldiers and their families, with little formal education. The main objectives are:
- the study of the Great War and its legacy
- the evolution of the linguistic uses in wartime

THE CORPUS
- The corpus is comprised of almost 1800 correspondences written by 37 writers in 11 areas for a total of almost 500,000 tokens.
- The transcripts are encoded in conformance with the TEI guidelines, which allowed to keep the alignment between the text and the facsimile.
- Sources are browsable and downloadable either from the ORTOLANG repository or from a TXM portal (with CQP queries).

THE VISUALISATION
1) Identify places and persons evoked in the correspondences between soldiers and their families
2) Trace the level of mobility of soldiers and family members

THE VISUALISATION INTERFACE

THE VISUALISATION
1) Identify places and persons evoked in the correspondences between soldiers and their families
2) Trace the level of mobility of soldiers and family members

RESEARCH OUTCOMES

RELATED STUDIES
- Most notable projects (Mapping the Republic of Letters; Early Modern Letters Online; Clavius on the Web) work on the correspondences of intellectuals. An exception is The Migrant Letter Digitised (O’Leary and Moreton, 2017).
- A TEI Correspondence Element was introduced in the 2.8.0 version of TEI P5.
- Project Testament des poilus - different genre, same writers - was an inspiration for the indexes.

ANNOTATION
TEIheader - <correspDesc> for the identification of sender(s), addressee(s), relationship, date and place of sending

Ongoing development:
- Annotation of places, people, organisations.
- Internal indexes for references + link to external resources (DBpedia, Geonames)

THE VISUALISATION

The visualisation interface was developed in collaboration with Pietro Barbieri, Chiara Capone and Luca Ciccone, MSc students in computer science, supervised by Marina Ribaudo, associate professor at DIBRIS, Università degli Studi di Genova.

Corpus 14 is a project of Laboratoire Praxiling UMR 5267 CNRS & Université Paul-Valéry Montpellier 3 Project: https://www.univ-montp3.fr/corpus14/ Laboratory: http://www.praxiling.fr