.

Editing a Medieval Hispanic Poetry Corpus with TextGrid.

María Gimena del Rio Riande.

Cita:
María Gimena del Rio Riande (2021). *Editing a Medieval Hispanic Poetry Corpus with TextGrid*.

Text+ User Story

Editing a Medieval Hispanic Poetry Corpus with TextGrid

Gimena del Rio Riande (CONICET), Buenos Aires, Argentina

DFG subject area: 105 Literary Studies

Text+ data domain: Editions

Motivation

In 2015, I was working with the digital edition of a selection of the debate and dialogue poems transcribed in a medieval 16th century songbook, the *Cancionero de Baena*, as part of my research work at the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) in Argentina. These poems were part of the corpus of the free-access online database ReMetCa, a Digital Repertoire on Medieval Castilian Poetry and Diálogo Medieval, a project I was developing at CONICET Argentina (repository in GitHub: https://github.com/Gimena/poesiaymetrica). This project aimed at establishing the basic characteristics of verse structure in dialogue poetry and give account of possible imitations or contrafactures, common verse, and metrical and rhyming patterns, a very important feature in medieval versification studies.

One year before, in 2014, Dr. Frank Fischer participated in the first Digital Humanities Conference organized in Argentina and showed some digital editions developed with TextGrid. I became interested in knowing more about this software and applied for a CENDARI research fellowship to work at the Göttingen Centre in Digital Humanities (GCDH), where different teams related to digital library studies and digital scholarly edition were working on different projects with this XML-TEI software.

By those times, Diálogo Medieval was the only DH project dealing with Medieval poetry and digital editing in Argentina, and I had very little funding from a national project, that didn't allow me to buy licenses for text editors such as Oxygen (*Primera fase de un Proyecto de Investigación en Humanidades Digitales: Estudio, edición y etiquetado de la poesía castellana medieval dialogada (siglos XII-XV) para la Base de Datos* DIÁLOGO MEDIEVAL. Miembro del Grupo responsable. (PICT FONCYT 2013, Temas Abiertos — Equipos de Reciente Formación) (9/2014 al 8/2017).). My Argentinian research team was also very new to Digital Humanties, and I was interested in learning more about developing digital editions. I had attended a Dixit Spring Institute that year in Graz and felt I needed to work together with a larger group of scholars interested in digital editions (I was awarded with a DIXIT Bursary to attend the Spring Institute "Advanced XML/TEI technologies for Digital Scholarly Editions", Graz, Austria, April 13–17, 2015).

Between July 20th and September 16th, 2015, and thanks to a CENDARI fellowship, I spent a 10-week fellowship at the Göttingen Centre for Digital Humanities (GCDH). I worked on the TEI encoding (mainly verse module) of almost fifty debate and dialogue poems in the Cancionero the Baena, and analyze and compare their formal characteristics. The project meant to argue that data-rich analysis of poetic meters offers humanistic insights into medieval Iberian poetry.

The research questions underlying the project focussed on:

1. the use of digital methods to create and make available knowledge that could not be achieved otherwise, and that in the case of this project were related to the coincidences revealed by the digital edition and comparison of a large amount of verse structures and metrical and rhyming patterns, in order to identify coincidences between poets and groups of poems,

2. the development of new methodologies that enabled original research about formal metrical practice on the basis of large digital corpora, e.g., the use of XPath/XQuery to process XML-TEI-encoded editions applied to texts edited and studied through a traditional and rigorous philological approach,

3. the use of digital tools that make it possible to address new research needs.

The GCDH was in my opinion the best choice to conduct this research, given its dedicated involvement into the developments of the TEI-XML standards, the know-how assembled around projects developed with TextGrid, and a digitally-oriented library that also holds a lot of Cancionero source material. I was able to edit the whole corpus and to put it online with the help of Mathias Göbel.

With the Göttingen team we organized a TextGrid workshop that year at UNED (Madrid, Spain, La edición digital académica: de lo analógico a lo digital, TextGrid at LINHD, UNED, 2015). In this workshop, students found useful and intuitive the way that TextGrid let them work with images from manuscripts while working on their digital editions. I completely agree with that opinion.

We also helped to prepare a preliminary translation of TextGridLab interface to Spanish which has not been implemented yet (https://github.com/lehkost/textgrid-es).

Sadly, I was not able to publish the edition in TextGridRep. For a time, I could not have access to my digital edition in TextGrid and now it's only accessible through the files kept in the DARIAH database and my GitHub repository. I lost contact with the GCDH team, that by those times was running through a different direction.

Objectives

To enhance the use of TextGridRep and TextGridLab by researcher from other disciplines and the international community, following aspects should be considered:

1. A wider acceptance of the X-Technologies inside TextGridLab. One of the difficulties of working with TextGrid was that it didn't work with some TEI standards, such as the use of a Relax-ng schema. I had already developed my schema, so I was force continuing working with the text editor Oxygen and only worked on the transformations in TextGrid. TextGridLab should be able to accept the different formats of schema that are used in the community, at least those that can be generated with TEI tools such as Roma.

2. To ensure the archiving and long-term accessibility of editions through repositories such as TextGridRep.

3. TextGrid should have a more multilingual sight than it currently has. Currently, TextGridLab offers the possibility of selecting either English or German for the interface. The documentation of Rep is available both in German and English, but many sections of the documentation of Lab is only in English. These should be offered in other languages.

Solution

1. Integration of other X-Technologies which are yet not part of TextGrid (such as schemata formats).

2. Support in the editorial process, and specially in the publication in TextGridRep. This could have the form of a help desk, tutorials (written or in video), workshops or documentation with a clear perspective on the user.

3. For that, documentation, teaching materials and the interfaces should be translated into several languages.