
Encoding sociolinguistic variables in a corpus of Medieval Sardinian texts

Freitag
10.03.2017
13:30 – 14:00
B31, 0.12

Nicoletta Puddu
University of Cagliari

n.puddu@unica.it

Medieval texts are characterized by a high degree of graphical and linguistic variation and by limited and partial documentation. On the one hand, these features pose clear problems in the creation of a corpus which, by definition, should be balanced, representative and normalized. On the other hand, the peculiar variation which emerges from Medieval texts can be fruitfully investigated in a socio-historical perspective. The lack of standardization can indeed provide clues of historical, regional and social variation, and the widespread presence of code mixing can reveal repertoires of text producers, shedding light on their level of literacy and on the distribution of the varieties in the community (Mazzon 2015). In this perspective, Sornicola (2015) clearly demonstrates the importance of providing an adequate sociolinguistic characterization of the document compilers in order to correctly interpret variation in Medieval texts.

In this contribution we will deal with the problem of annotating the *condaghes*, Sardinian medieval texts that record transactions of churches and monasteries and often contain transcriptions of legal disputes. We will show that a combined XML annotation of extra-linguistic and philological information according to TEIP5 can help the correct meaning to be assigned to the variation phenomena present in the texts. In particular, we shall focus on the interference of Italian in the *condaghes*, which can be assigned to different levels of text production, and, as a consequence, to different levels of annotation.

References: • Mazzon, G. (2015): Historical sociolinguistics and history of English. In Molinelli, P.& I. Putzu (eds.), *Modelli epistemologici, metodologie della ricerca e qualità del dato*. Milano: FrancoAngeli, 50-68. • Sornicola, R. (2015): Curiales, Notarii, Presbyteri nella Campania Alto-Medievale. In Consani, C. (ed.), *Contatto interlinguistico fra presente e passato*. Milano: Edizioni Universitarie Lettere, Economia e Diritto, 237-282.