

state-of-the-art

Originally, all Italian dialects were divided into three main groups: Northern, Tuscan, and Southern. Although initially, the Roman language (*Romanesco*) formed part of the last group, before the second half of the 16th century, it came to resemble the Tuscan varieties. Its mutation was so deep that scholars tend to refer to it as 'disintegration' (Migliorini 1932), instead of 'evolution' (as is the usual case for most of the languages). The process is called **Tuscanization**, or **de-southernization**.

Nowadays, **two competing theories try to explain this process**. Mancini (1987) stated that the Tuscanization was a slow event, already in place in the half of the 15th century, while Trifone (1990) replied that the de-southernization of the spoken language of Rome was mainly caused by the demographic de-southernization of the city after the Sack of 1527.

I argue that—to solve the issue—we should run our analyses on bigger datasets, and pay more attention to the physical supports of the sources that transmit texts written in *Romanesco*. If we take into account those two variables, we could enhance our understanding of this process.

Visualizing *Romanesco*; or, Old Data, New Insights

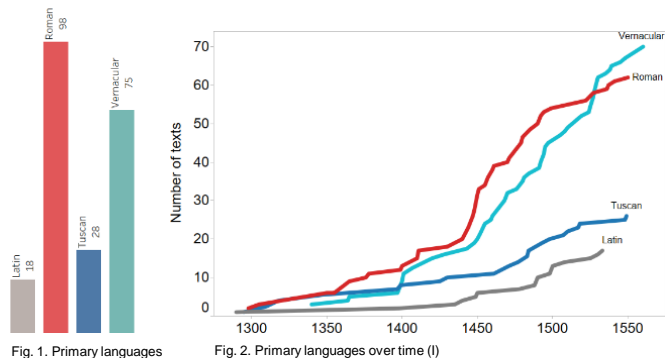


Fig. 1. Primary languages

Fig. 2. Primary languages over time (I)

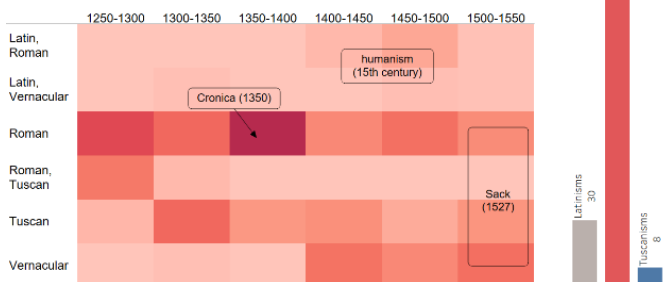


Fig. 3. Primary languages over time (II)

Fig. 4. Secondary languages

The visualizations of the **languages over time** (fig. 1–4 above) provide some insights that, in a way, strengthen both hypotheses of Mancini and Trifone.

- The number of texts written in *Romanesco* and Tuscan starts to diverge (in favor of the former) in the 2nd half of the 14th century, maybe partially due to the success of the *Cronica* (1357–1358)
- Tuscan and Vernacular languages are attested all over the time, from the Origins to the 16th century. We do not see a dramatic increase after 1527
- In the first half of the 16th century, texts written in *Romanesco* slightly decrease, while texts written in Tuscan and Vernacular languages increase

conclusions

Within the traditional approach, scholars tried to explain the Tuscanization by analyzing the linguistic features of a small selection of texts. The results of this approach did not lead to a sharp understanding of this particular linguistic process.

In order to resolve this issue, we need to look back at those texts that—until now—have been catalogued as written in 'Vernacular' language. **Looking for pieces of evidence of the linguistic features of epigraphs is the only way to increase the number of texts whose language is known**. This is our best chance to expand the dataset and thus provide new insights that will explain the *Tuscanisation* of *Romanesco*.

corpus & tools

Dataset I put in plain text files the metadata of all known texts with at least some features of *Romanesco* in them, written from 800 to 1550. I took the data from D'Achille & Giovanardi (1984)

Languages A text can contain some features that do not belong to its original linguistic system. Therefore, for each text of my corpus, I identified its **primary language** and its **secondary languages**. The languages of these texts are Roman, Tuscan, Latin and Vernacular. (Notice that scholars usually apply the label «Vernacular» every time that the linguistic system of a given text cannot be established surely)

Supports The texts are transmitted by different physical supports:

- **Places** (like churches, or catacombs) transmit epigraphic texts
- **Manuscripts** transmit handwritten texts
- **Printed books** transmit printed texts

Visualizations All the visualizations are made with the software **Tableau**

A deep look at the **physical supports** (fig. 5–6 below) shows that the number of epigraphs is significantly high. The overall number is definitely above average, if compared to other geographic areas in the same period. In particular, fig. 6 shows that epigraphs are the most frequent support for Roman texts up to the end of the 16th century. From that date, manuscripts prevail.

Then what?

By their very nature, epigraphs are dramatically short, and in consequence, linguistic features that are typical of a given area are less likely to be detected in epigraphs than in other textual typologies. **The large number of epigraphs in this corpus may be related to the large number of occurrences of texts written in a language that has been defined, generically, as 'Vernacular'**. Therefore—and maybe because of their **apparently low** linguistic value—the past surveys on the Tuscanisation of *Romanesco* did not take enough into account epigraphic texts. However, the limited amount of handwritten sources of *Romanesco* makes the epigraphic texts a critical source to understand the linguistic mutation of the Roman language. Thus, we should look at the epigraphs with renewed attention, searching for new pieces of evidence of their linguistic features. This will lead to new assumptions in the context of the issue of the Tuscanization.

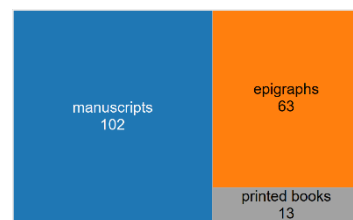


Fig. 5. Physical supports

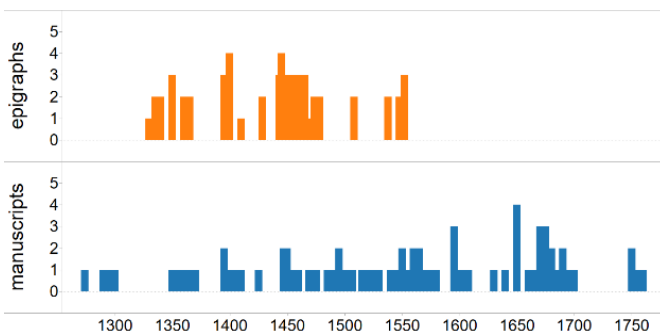


Fig. 6. Physical supports over time

references

- D'Achille, P. & Giovanardi, C. (1984). *La letteratura volgare e i dialetti di Roma e del Lazio. Bibliografia dei testi e degli studi. Vol. 1: Dalle origini al 1550*. Rome: Bonacci
- Mancini, M. (1987). «Aspetti sociolinguistici del romanesco nel Quattrocento». In *RR. Roma nel Rinascimento. Bibliografia e note* 3, 38–75
- Migliorini, B. (1932). «Dialetto e lingua nazionale a Roma». In *Capitolium* 10, 350–356
- Trifone, P. (1990) «La svolta del romanesco tra Quattro e Cinquecento». In *Studi in memoria di Ernesto Giamarco*, Pisa: Giardini, 425–452

gianluca valenti



- PhD in Romance Philology (Sapienza University)
- Advanced Master of Digital Humanities (KULeuven)
- Currently post-doc researcher (Université de Liège)
- Topics of interest
 - Historical linguistics
 - Philology of practical texts
 - Epistolary language
 - Italian grammars of the 16th century

gianluca.valenti@uliege.be