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

Gianluca Valenti
Université de Liège
`gianluca.valenti@uliege.be`

Abstract

English. The evolution of the Roman language over time is a problematic topic, which has been analysed by many scholars and different points of view. Nonetheless, we do not know enough about the process of its Tuscanisation. With the help of some data visualisations, I show that from the sole analysis of the language of all known sources, it is not possible to clarify this issue. Conversely, new attention given to the physical supports that transmit the Roman documents can provide new, interesting insights.

Italiano. L'evoluzione dell'antico romanesco è un argomento analizzato da numerosi studiosi ma, ad oggi, ancora non compreso in ogni suo aspetto. In particolare, non si è giunti a un accordo su tempistiche e modalità con cui avvenne il processo della sua toscanizzazione. Uno sguardo d'insieme ai testi con una o più caratteristiche di romanesco antico chiarisce come la questione non possa essere risolta se l'attuale corpus è indagato esclusivamente nei suoi aspetti linguistici; per ottenere nuove informazioni bisognerà dunque volgersi allo studio dei supporti materiali che trasmettono tali documenti.

1 The sociolinguistic background

In the past decades, dialectology studies have exponentially increased. All linguists nowadays acknowledge the importance of the dialects in the social and cultural history of humanity, and put them on an equal footing with standard languages. In particular, in the context of Italian sociolinguistics, the study of the dialect of Rome is considered as an extremely relevant topic.

Since the beginnings, Italian dialects are divided into three main groups: Northern, Tuscan, and Southern varieties. The medieval and renaissance periods marked an irreversible revolution in the Roman social background, and consequently, in the Roman language—the *Romanesco*.

Although *Romanesco* formed part of the Southern dialects, before the second half of the 16th century at the latest, it came to resemble the Tuscan varieties—a process called 'Tuscanisation' (Ernst, 1970). This change, whose dynamics remain largely unclear, represents a unique episode in the history of the Italian language. The uniqueness of the 'Rome case' has been stressed on many occasions, and several explanations have been proposed for it, but no agreement has been reached yet.[1] Indeed, the mutation of *Romanesco* was so deep (and without similar precedents) that scholars tend to refer to it as 'disintegration, decay' (Migliorini, 1932), instead of 'evolution' (as is the usual case for most of the languages).

More than fifty years after Migliorini's work, Mancini (1987:41) argued that, until then, scholars 'placed too much emphasis on some demographically macroscopic events,' such as the Sack of Rome of 1527. Instead, according to him, the mutation of *Romanesco* was a slow event, already in place, at least in its germinal stages, in the *Trecento* and the *Quattrocento*. Trifone (1990:92) quickly reacted with a detailed linguistic analysis of new documents, and concluded that the demographic de-southernisation of Rome (caused by the sack) and the ensuing repopulation of the city post-1527 were the main reasons for the de-southernisation of the spoken language of its lowest social classes. The debate has continued for several years (cf. also Trifone 1992 and Mancini 1993) without a consensus.

From one side, it is unquestionable that, as De Caprio (1988:453) states, 'the sack of Rome of 1527 is a traumatic caesura in Roman cultural history,' but, from the other side, its role in the context of the Tuscanisation is admittedly unclear. Even at the present time, instead of positioning themselves on one side or

---

[1] Cf. e.g., Vignuzzi (1988, 1995), De Mauro (1989), Palermo (1991), and, recently, Coluccia (2011), who argues that the exceptionality of the phenomenon lies in the untimely Tuscanisation at a spoken level.

the other, scholars tend to report both opinions, at most trying to harmonise them into a single whole.[2] Currently, and maybe because no one has established a definitive answer to the issue, researchers seem to prefer to focus on the currently spoken *Romanesco*, a language that has still many points of contact with its renaissance variety.[3] Interest in the old *Romanesco* has not waned, though, as is clear by the recent organisation of the roundtable 'Rome in history, in linguistics and in literature' (Rome, 23th July 2016) and the international conference 'Il romanesco tra ieri e oggi', which I organised in Liège the 9th September 2019.

2   Old data, new insights

The old epistemological framework—where single scholars tried to explain the evolution of the Roman language by studying analytically one or few texts—has proved to be unfit for the task of understanding the dynamics of the Tuscanisation of *Romanesco*. At the present time, it is necessary to look at a broader picture, and consider the Roman texts as if they were a single whole. Indeed, up to now, scholars essentially based their findings on qualitative research, but did not exploit the potentialities of databases and digital tools. Most of the current papers focused on *Romanesco* analyse the language of a specific text (or a bunch of texts), while a general overview that takes into account the huge amount of data pertaining to the Roman sources collected by the scientific community during the past hundred years is still missing.

I recently made the first step in order to fill this gap, by building and putting online a database that allows users to make queries into the whole corpus of texts written in *Romanesco* from the Origins to 1550. It is available online, free of charges, at the address `http://www.romanesco.uliege.be/`. Regularly updated, the database makes available metadata concerning not only the texts, but also the physical supports (printed books, manuscripts, and places, such as churches or catacombs) that transmit them.

Working on digital data leads scholars to new findings, and allows them to answer old research questions, which could not be solved with the traditional approach. In this paper, with the help of some visualisations, I show that—due to the scarcity of the sources—we do not have sufficient data to get new insights about the language of Rome and its evolution through time if we only look at the languages of the texts. Therefore, to have a sharp understanding of the Tuscanisation process, we need to reanalyse the linguistic features of the epigraphs, whose language has been often defined, maybe too quickly, as generically 'Vernacular'.

2.1   Corpus and tool

To conduct this study, I put in plain text files the metadata of 372 texts, written from 800 to 1550, with at least some features of *Romanesco* in them. I took the data from D'Achille and Giovanardi (1984). With regard to the languages, notice that a text can contain some features that do not belong to its original linguistic system. In consequence, for each text, I identified its primary language and all the potential secondary languages (i.e., the languages that occur to a lesser extent). The condition for a text to be included in this corpus is to have *Romanesco* as its primary or, at least, its secondary language. The corpus is thus composed of texts written in *Romanesco*—which may contain some pieces of Tuscan or Latin—, but also of texts that contain only a small amount of features of *Romanesco*, while their primary language is Tuscan, Latin, or Vernacular.

Each text is transmitted by one physical support: a) 'places' transmit epigraphic texts; b) 'manuscripts' transmit handwritten texts; and c) 'printed books' transmit printed texts.

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

2.2   Results

Figures 1–4 show the total number of occurrences of primary and secondary languages, and their evolution over time. The figures provide some interesting insights that, in a way, strengthen both hypotheses of Mario Mancini and Pietro Trifone.

---

[2] Cf. e.g., Vignuzzi (1994), Giovanardi (1998:61), D'Achille and Petrocchi (2004:122–123).
[3] Cf. e.g. the projects *VRC. Vocabolario del Romanesco Contemporaneo* and *ERC. Etimologie del romanesco contemporaneo.*
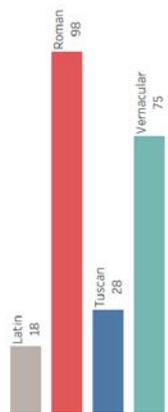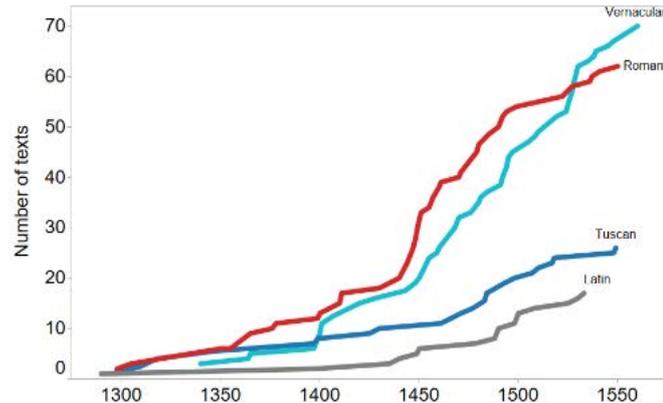
Figure 1. Primary languages

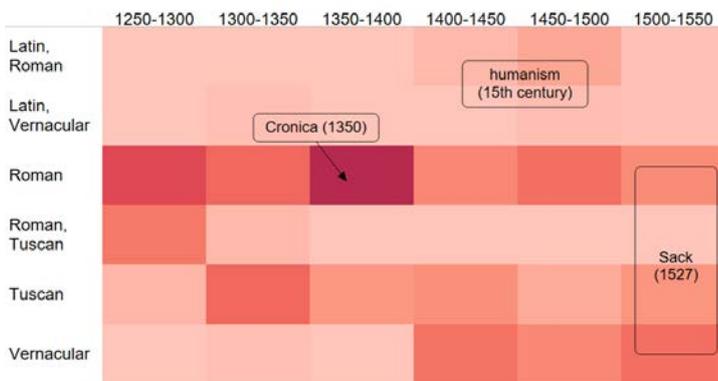*Figure 2. Primary languages over time (I)*

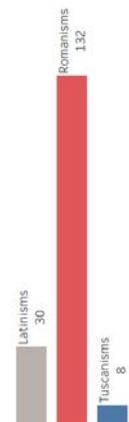*Figure 3. Primary languages over time (II)*

Figure 4. Secondary languages

We notice that the total amount of occurrences of texts written in *Romanesco* and Tuscan starts to diverge (in favour of the former) in the 2nd half of the 14th century, maybe due to the success of Anonimo Romano's *Cronica* (1357–1358), the most important Roman text of its century.

However, the Tuscan language is attested all over the centuries, from the 14th to the 16th century. I do not register any dramatic increase right after 1527 (cf. figure 2). This observation seems to endorse Mancini's view of the Tuscanisation as a slow process, already in place in the *Quattrocento*.

On the other hand, if we look at figure 3 we notice, in the first half of the 16th century, a slight decrease of texts written in *Romanesco*, and a parallel growth of texts written in the Tuscan language. Admittedly, this outcome may be related—as Trifone states—to the de-southernisation of the Roman population after the Sack, and the subsequent increase in the number of Tuscan people moving to the town.

These visualisations improve significantly our perception of the evolution of *Romanesco* through time. Nonetheless, they do not provide any irrefutable evidence that would end the debate on the timing and modalities of its Tuscanisation. Therefore, we need to look at the problem from another perspective.

Indeed, an aspect has escaped the scrutiny of most of the past scholars: observing the physical supports that transmit Roman texts over the years, we notice some interesting insights.
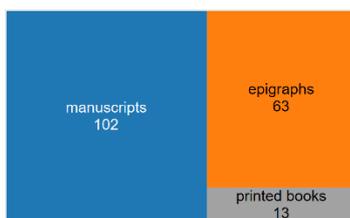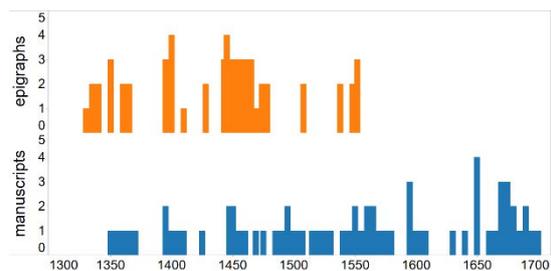

Figure 5. Physical supports

Figure 6. Physical supports over time

The prevalent supports of literary texts are, up until and including the 15th century, manuscripts, and afterwards, printed books. In the present corpus, though, the total number of epigraphs is significantly high (cf. figure 5). Furthermore, if we consider only the sources ranging from 800 to 1550, texts transmitted by epigraphs are even more than texts transmitted by manuscripts (cf. figure 6).[4] This is due to the fact that most of the texts of this corpus are practical documents (such as receipts, letters, and private notes), and have no literary value. Indeed, texts that are transmitted by perishable material—such as pieces of paper—get easily lost, while texts that are carved on the column of a church are more likely to be preserved.

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 high number of epigraphs included in this corpus may be related to the high number of occurrences of texts written in a language that has been defined, generically, as 'Vernacular'. Therefore—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 low number of handwritten documents of *Romanesco* and, in contrast, the high number of epigraphic texts, make the latter a critical source to understand the linguistic mutations in the medieval and renaissance Rome. Moreover, we should not forget that, as a starting point for the research, we have at disposal a solid documentary basis, the volume on Vernacular texts found in churches, edited by Sabatini et al. (1987). There, the authors provide detailed linguistic analyses, which could serve as a model for further studies focused on the language of the newly discovered epigraphic texts of the past thirty years.

Once we have collected a significant amount of new data, it will be possible to look again at the linguistic features of the Roman sources—including but not limited to manuscripts and printed books—thus refining our theories and reaching new conclusions about the Tuscanisation process of *Romanesco*.

3 Conclusions and new perspectives

Within the traditional approach, scholars tried to explain timing and causes of the Tuscanisation of *Romanesco* by analysing the linguistic features of a small selection of texts. The results of this approach—albeit essential in many respects—did not lead to a sharp understanding of this particular linguistic process. I have shown that the reason for this failure is not entirely due to the little number of texts analysed. Indeed, even though we consider all texts at our disposal, we are not able to recognise a clear pattern in favour of one or the other theory. In order to resolve this issue, we need more data, i.e., we need to look back at those texts that—until now—have been catalogued as written in 'Vernacular' language.

The high number of texts that we did not assign to any specific linguistic system is probably related to the high number of epigraphs that transmit them. Nowadays, scholars should approach the epigraphic texts with renewed attention, looking for pieces of evidence of their linguistic features. While awaiting additional archival findings, this is the only way to increase the number of texts whose language is known, which is our only chance to make new assumptions that could explain the Tuscanisation of *Romanesco*.

References

Alberto Asor Rosa. 2009. *Storia europea della letteratura italiana. I. Le origini e il Rinascimento*. Einaudi, Turin.

Rosario Coluccia. 2011. Scripta. In: R. Simone (ed.), *Enciclopedia dell'italiano*, volume 2, Istituto della Enciclopedia Italiana, Rome: 1287–1290.

Paolo D'Achille and Claudio Giovanardi. 1984. *La letteratura volgare e i dialetti di Roma e del Lazio. Bibliografia dei testi e degli studi. Vol. 1: Dalle origini al 1550*. Bonacci, Rome.

Paolo D'Achille and Stefano Petrocchi. 2004. *Limes* linguistico e *limes* artistico nella Roma del Rinascimento. In: V. Casale and P. D'Achille (eds.), *Storia della lingua e storia dell'arte in Italia. Dissimmetrie e intersezioni. Atti del III Convegno ASLI (Roma, 30–31 maggio 2002)*. Cesati, Florence: 99–137.

Vincenzo De Caprio. 1988. Roma. In: A. Asor Rosa (ed.), *Letteratura italiana. Storia e geografia*, volume 2/1, Einaudi, Turin: 327–472.

---

[4] I did not put in the plot the printed books, because only three of them are attested before the end of the 16th century. Similarly, I did not put in the plot the sources from the 9th to the 13th century, because they are extremely rare (only four epigraphs in total). Notice also that, even though I only consider texts ranging from 800 to 1550, the dating of the sources can be later.

Tullio De Mauro. 1989. Per una storia linguistica della città di Roma. In: Id. (ed.). *Il romanesco ieri e oggi. Atti del Convegno del centro romanesco Trilussa e del Dipartimento di Scienze del linguaggio dell'Università 'La Sapienza' (Roma, 12–13 ottobre 1984)*, Bonacci, Rome: XIII–XXXVII.

Gerhard Ernst. 1970. *Die Toskanisierung des romischen Dialekts im 15. und 16. Jahrhundert*. Niemeyer, Tübingen.

Claudio Giovanardi. 1998. *La teoria cortigiana e il dibattito linguistico nel primo Cinquecento*. Bulzoni, Rome.

Mario Mancini. 1987. Aspetti sociolinguistici del romanesco nel Quattrocento. *RR. Roma nel Rinascimento. Bibliografia e note* 3:38–75.

Mario Mancini. 1993. Nuove prospettive sulla storia del romanesco. In: Istituto Nazionale di Studi Romani (ed.), *«Effetto Roma». Romababilonia*, Bulzoni, Rome: 9–40.

Bruno Migliorini. 1932. Dialetto e lingua nazionale a Roma. Reprint in: Id. 1948. *Lingua e cultura*. Tumminelli, Rome: 109–123.

Massimo Palermo. 1991. Fenomeni di standardizzazione a Roma nel primo Cinquecento. *Contributi di Filologia dell'Italia mediana* 5:23–52.

Francesco Sabatini, Sergio Raffaelli and Paolo D'Achille. 1987. *Il volgare nelle chiese di Roma. Messaggi graffiti, dipinti e incisi dal IX al XVI secolo*. Bonacci, Rome.

Pietro Trifone. 1990. La svolta del romanesco tra Quattro e Cinquecento. Reprint in: Id. 2006. *Rinascimento dal basso. Il nuovo spazio del volgare tra Quattrocento e Cinquecento*. Bulzoni, Rome: 61–94.

Pietro Trifone. 1992. *L'italiano nelle regioni. Roma e il Lazio*. Utet, Turin.

Ugo Vignuzzi. 1988. Italienisch: Areallinguistik VII. Marche, Umbrien, Lazio / *Aree linguistiche VII. Marche, Umbria, Lazio*. In: G. Holtus, M. Metzeltin and C. Schmitt (eds.), *Lexikon der Romanistischen Linguistik*, volume 4, Niemeyer, Tübingen: 606–642.

Ugo Vignuzzi. 1994. Il volgare nell'Italia mediana. In L. Serianni and P. Trifone (eds.), *Storia della lingua italiana*, volume 2, Einaudi, Turin: 329–372.

Ugo Vignuzzi. 1995. Marche, Umbrien, Lazio / *Marche, Umbria, Lazio*. In: G. Holtus, M. Metzeltin and C. Schmitt (eds.), *Lexikon der Romanistischen Linguistik*, volume 2/2, Niemeyer, Tübingen: 151–169.