

# A Corpus-Based Approach to Michelangelo's Epistolary Language

Gianluca Valenti

Université de Liège, Unité de recherches "Transitions"

## 1 Introduction

### 1.1 Sociolinguistic Background

Because of documentary and literary reasons, the language of Florence is probably the most studied among the Italian dialects. Its historical development, though, has not been as linear as one might think.<sup>1</sup> Indeed, as is the case for many other dialects, it begins to be widely written, also for literary works, during the Late Middle Ages, but—as is *not* the case for many other dialects—it suddenly becomes extremely popular, thanks to famous poets that used it and spread it all around the Peninsula in the 13<sup>th</sup> century.<sup>2</sup>

It goes without saying that the 14<sup>th</sup> century permanently sanctions the supremacy of the Florentine language. The 'Three Crowns'—Dante, Petrarch, and Boccaccio—ennoble it by writing masterpieces the caliber of the *Comedia*, the *Rerum Vulgarium Fragmenta*, and the *Decameron*.

Thus in Florence, in the subsequent centuries, the spoken language is constantly evolving over time, as is expected to be. However, people from outside Florence, who increasingly need a common code to communicate, take the language of the 14<sup>th</sup> century as a reference point. Thus, 14F acquires a greater value than other Italian dialects both because of the literary importance of the texts of Dante, Petrarch, and Boccaccio, and because—unlike other contemporary dialects—it is a written model, which could be studied and learned.

From the end of the *Quattrocento* and throughout the *Cinquecento*, some humanists begin to recommend taking as a linguistic model the 14F (often called 'volgar lingua'), disregarding any further development occurred at the spoken level. The first two grammar books are the *Regole grammaticali della volgar lingua* by Giovan Francesco Fortunio (1516) and the *Prose della volgar lingua* by Pietro Bembo (1525)—the latter being by far the most important and most widespread of its kind.<sup>3</sup>

---

<sup>1</sup> In the whole paper, when discussing the language of Florence, I will use the following abbreviations: 14F = fourteenth-century (also called 'golden') Florentine language; 16F = sixteenth-century (also called 'silver'—see Castellani (1970): 17) Florentine language.

<sup>2</sup> Among them, one could mention at least Dante da Maiano, Monte Andrea, Chiaro Davanzati, Bonagiunta Orbicciani and Guittone d'Arezzo.

<sup>3</sup> Obviously, I am simplifying a situation that is much more complex than that. Cf. at least Ghinassi (1961), Quondam (1983), Trovato (1991) and, lastly, Valenti (2018) on the problem of the linguistic norm

Surprisingly enough, in the sixteenth century, people from outside Florence are much more prone to learn *14F* than Florentine people themselves. Indeed, the foreigners study *14F* as a completely new language, without any concern for the fact that it is a 'dead' language, which dates back two hundred years. On the contrary, the Florentine people hardly accept to use a language different from the one they speak and write in daily life. Unfortunately for them, the more time passes, the more *14F* is perceived as the language of high society: people from Florence are increasingly required to use it, too, because this is how they are expected to communicate in cultured and educated milieus.

Indeed, in the whole *Cinquecento*, we notice in Florentine texts a tension between the will to keep the contemporary language (= the *16F*), and the need to use the *14F* to communicate with people from outside Florence. In this context, it is therefore of great interest to analyze the historical evolution of the Florentine language throughout the entire 16<sup>th</sup> century.

After performing a correspondence analysis and a correspondence regression on Michelangelo's entire epistolary corpus (about 500 letters), I verified an evolution over time in his use of the language, and I provided a historical explanation to the outcomes of the statistical tests.

## 1.2 Michelangelo's Language: An Open Question

In this paper, I focus on Michelangelo's epistolary language. On the one side, I have chosen to analyze only letters (leaving aside the many poems written by the sculptor) because the current linguistic studies increasingly show the need to focus mostly on practical texts instead of literary works.<sup>4</sup> Indeed, practical texts are the best choice for linguistic analyses, because they do not aim to be artistic, and frequently belong to authors without any specific literary education (Serianni (2007): 13).

In recent years, the relevance of private correspondence has been quickly perceived by scholars (cf. e.g., Magro (2014): 106). Letters provide a wealth of precious information for linguists, both because they often carry a date and because their language is frequently close to that of ordinary speech, thus offering access to useful data.<sup>5</sup> As is shown, for example, in Culpeper and Kytö (2010): 17, letters are included in the group of speech-related genres and listed among the speech-like typologies. Thus, despite in epistolary writing the interaction takes place asymmetrically over time, communication is similar to that of the oral speech.<sup>6</sup>

Specifically, this research is targeted towards Michelangelo because he can be considered one of those "intermediate individuals, neither erudite not uneducated people,"<sup>7</sup> who are nowadays drawing the attention of scholars. As is well known, since the appearance of the notion of semi-educated writers (Bruni (1978), Bruni (1984)), scholars increasingly discussed the topic, and today there is a strong tendency to consider the writers' level of education as a continuous, rather than as a set of discrete variables

---

in Italy in the 16<sup>th</sup> century. On Fortunio, cf. the updated bibliographic references in Moreno and Valenti (2017). On Bembo, cf. at least Patota (2017), together with the critical editions made by Dionisotti (1966), Vela (2001) and Tivosanis (2002).

<sup>4</sup> Cf e.g. Antonelli et al. (2014) for a recent overview of the Italian context.

<sup>5</sup> Recently, D'Achille and Stefinlongo (2016): 249 (translation mine) affirmed that the letters "witness the *langue* throughout the ages."

<sup>6</sup> In the last few years, many scientific essays analyze letters from a linguistic point of view: cf. e.g., Fitzmaurice (2002), Nevalainen and Tanskanen (2007) and Auer et al. (2015).

<sup>7</sup> See Testa (2014): 7 (translation mine). About the semi-educated writers (It. 'semicolti'), see at least D'Achille (1994) and Fresu (2014).

(Fresu (2004), Librandi (2004), Bianconi (2013)).

Finally, it is also of interest to study Michelangelo's language because it fits in a complex and much debated epistemological framework, that of the language of arts and artists. Starting with Folena (1951) and Folena (1957), scientific studies on this topic have multiplied, and have involved prominent scholars, from Barocchi (1984) to Nencioni (1995).<sup>8</sup>

Previous scholars have argued that Michelangelo's epistolary language constitutes a representative example of *16F*, and that it does not make use of most of the features that characterize the language of the Three Crowns.<sup>9</sup> Persuasive as this may seem, I suggest that this assumption can be challenged.<sup>10</sup>

The language of Michelangelo's letters testifies to an interesting tension between the contemporary linguistic usage typical of a sixteenth-century man of Florence, and the Old Florentine literary language prescribed by the grammarians. I present in Section 2.2 the results of an investigation conducted so as to determine the extent to which Michelangelo used *14F* and *16F*. Indeed, from my findings, it would seem that the artist was more aware of the Old Florentine linguistic system than was initially assumed, as it appears that, starting from 1530, he was not loath to borrow from it.

## 2 Methodological Framework

### 2.1 Limits and Constraints

Before I go any further, I cannot pass over in silence that the boundaries between *14F* and *16F* are less clear than suggested above. As is well known, some of the so-called fourteenth-century linguistic phenomena had already occurred by the end of the *Duecento* and the beginning of the *Trecento*, but the majority of them only appeared in its second half, and became more stable during the following century.<sup>11</sup> Moreover, it is not even clear when exactly those phenomena started to fade. It is probable that some phenomena were still in use in the first half of the sixteenth century, while others had spontaneously evolved, and others still suddenly found themselves in competition with the fourteenth-century linguistic system, which—at some point—replaced them. Accordingly, the labels *14F* and *16F* do not reflect a clear chronological distinction, and each phenomenon should be discussed and evaluated on a case by case basis.

Another issue is that I focus only on the diachronic variable, while I do not take into account the diaphasic variation. Obviously, for a more comprehensive approach, I should distinguish between letters sent e.g. to subordinates or relatives, and letters sent to the pope or to noblemen. The contents of the message hardly are the same, and the overall tone and style can vary significantly from letter to letter. However, because of the high number of documents taken into account, considering uniquely the diachronic variable can lead to interesting results, too.

Two other limits are somehow inherent to such research. First, I analyze only one writer, while—for a comprehensive study of the variation of the Florentine language in the 16<sup>th</sup> century—many epistolary corpora, written by different authors, should be compared. And second, the open debate about the possibility of determining (and

---

<sup>8</sup> It goes beyond the objectives of this paper to summarize all references on this topic. See, lastly, Aresti (2019) for recent bibliographical updates.

<sup>9</sup> See, e.g., Ciulich (1973), Nencioni (1965) and, more recently, D'Onghia (2014), D'Onghia (2015), Felici (2015) and Marazzini (2015).

<sup>10</sup> I already argued against this view in Valenti (2017b).

<sup>11</sup> See Palermo (1992) and Manni (1979), two of the most accurate works on the topic.

to what extent) new information about a spoken language from the analysis of its graphic representation, dates back to the creation of the word *scripta* itself (Remacle (1948)). However, as Arcangeli (2011): 10 (translation mine) notes: “if we are willing to formulate some hypothesis on the state of a language between the Middle Ages and the Renaissance, ... we are necessarily forced to base our conjectures on written texts.” Indeed, historical linguistics “is not a second-best solution by inevitable necessity, but just the best solution in those areas of study for which oral records are not available, especially when studying long-term developments of language variation and change” (Hernández-Campoy and Schilling (2012): 64).

Then, strictly speaking, I am analyzing only the *scripta*—not the language—of Michelangelo: but analyzing the *scripta* is the only way to get information about his language.

## 2.2 Results

To collect the corpus, I copy-pasted the texts from Memofonte (2008) to thirteen .txt files, split into time intervals, from 1495 to 1564.<sup>12</sup> At the end of this first step, each file contained all the letters written by Michelangelo over a range of five years.<sup>13</sup>

Then, I deleted all the special characters “.,:;!?’.” and I split the texts, one word per line. At the end of this step, I put every document (= every time interval) into vectors.<sup>14</sup> Subsequently, based on the current bibliography,<sup>15</sup> I selected the major features that differentiate *14F* and *16F*. For every feature, I identified the golden forms (i.e., *14F* forms) and the silver forms (i.e., *16F* forms). I display here all of them:

- **CruoG** = A tonic Latin ō, preceded by consonant plus *r*, becomes *uo* in *14F* (ex. Lat. PRŌBA > *pruova*)
- **CruoS** = A tonic Latin ō, preceded by consonant plus *r*, becomes *o* in *16F* (ex. Lat. PRŌBA > *prova*)
- **CrieG** = A tonic Latin ě, preceded by consonant plus *r*, becomes *ie* in *14F* (ex. Lat. BRĚVIS > *brieve*)
- **CrieS** = A tonic Latin ě, preceded by consonant plus *r*, becomes *e* in *16F* (ex. Lat. BRĚVIS > *breve*)
- **schivG** = Before vowel, /skj/ remains /skj/ (written <sch>) in *14F* (ex. Lat. SCLAVUS > *schivo* after a passage CL > *chi* in Medieval Latin)
- **schivS** = Before vowel, /skj/ becomes /stj/ (written <sti>) in *16F* (ex. Lat. SCLAVUS > *stivo*)
- **lliG** = A plural noun or adjective, ending in -LLI in Latin, ends in *-lli* in *14F* (ex. Lat. CABALLUS, pl. CABALLI > *cavallo*, pl. *cavalli*)
- **lliS** = A plural noun or adjective, ending in -LLI in Latin, ends in *-gli* in *16F* (ex. Lat. CABALLUS, pl. CABALLI > *cavallo*, pl. *cavagli*)

<sup>12</sup> Last viewed: 01.04.2020. This digital edition accurately reproduces the text of the most important critical edition of Michelangelo’s letters (Barocchi and Ristori (1983)).

<sup>13</sup> Notice that we do not possess any letter dated between 1500 and 1505.

<sup>14</sup> For all the steps that follow, I used the software *R* (Team (2020)), and the packages “ca”, “mclm”, “factoextra” and “FactoMineR.”

<sup>15</sup> See Migliorini (1955), Castellani (1970), Castellani (2000), Buck and Pfister (1971), Manni (1979), Palermo (1992), Salani (1992), Renzi and Salvi (2010).

- **artG** = The singular masculine definite article is *il* in 14F
- **artS** = The singular masculine definite article is *el* in 16F
- **prIVamoG** = The indicative present IV p. ends in *-(i)amo* in 14F (ex. *noi scriviamo*)
- **prIVamoS** = The indicative present IV p. ends in *-(i)ano* or *-emo* in 16F (ex. *noi scriviano, noi scrivemo*)
- **prVianoG** = The indicative present VI p. ends in *-ano* in 14F (ex. *loro parlano*)
- **prVianoS** = The indicative present VI p. ends in *-ono* or *-eno* in 16F (ex. *loro parlono, loro parlono*)
- **impfVivanoG** = The indicative imperfect VI p. ends in *-vano* in 14F (ex. *loro scrivevano, loro mangiavano*)
- **impfVivanoS** = The indicative imperfect VI p. ends in *-vono* or *-veno* in 16F (ex. *loro scriveveno, loro mangiavano*)
- **futrG** = The endings of the indicative future have one intervocalic *r* in some verbs in 14F (ex. *scriverò, scriverai, scriverà...*)
- **futrS** = The endings of the indicative future have two intervocalic *r* in some verbs in 16F (ex. *scriverrò, scriverai, scriverà...*)
- **congG** = The subjunctive present ends in *-a* (I, II and III p.) and *-ano* (VI p.) in 14F (ex. *che tu abbia, che lui voglia, che loro debbano...*)
- **congS** = The subjunctive present ends in *-i* (I, II and III p.) and *-ino* (VI p.) in 16F (ex. *che tu abbi, che lui vogli, che loro debbino...*)
- **condG** = The endings of the conditional present have one intervocalic *r* in some verbs in 14F (ex. *scriverei, scriveresti, scriverebbe...*)
- **condS** = The endings of the conditional present have two intervocalic *r* in some verbs in 16F (ex. *scriverei, scriveresti, scriverebbe...*)
- **trG** = Lat. DE ÎNTRO, DE RËTRO become *dentro, dietro* in 14F
- **trS** = Lat. DE ÎNTRO, DE RËTRO become *drento, dreto* in 16F
- **senzaG** = Lat. ABSËNTIĀ becomes *senza* in 14F
- **senzaS** = Lat. ABSËNTIĀ becomes *sanza* in 16F
- **ultimG** = Lat. ULTIMUS, -A, -UM becomes *ultim-* in 14F
- **ultimS** = Lat. ULTIMUS, -A, -UM becomes *utim-* in 16F

Clearly, here I am only talking of formal variation, significantly different from conceptual variation. The latter refers to the authors' choice between the use of a word (for instance, 'oak') and, e.g., the use of an hyperonymous ('tree'), while the former only concerns the linguistic variation of the same term—such as, in sixteenth-century Florence, the choice between the forms 'senza' and 'sanza', both meaning 'without'. Within this approach, "the downside is that formal variation is only one aspect of a

much broader reality, but it is an aspect we claim is worth isolating” (Speelman et al. (2003): 319).

After listing all those features, I ran the corresponding queries all over the thirteen vectors, so to find the total number of occurrences of each feature for each time interval, and then I manually put the outcomes in a single .csv file. I underline that for running the queries, I had to choose between two options. Sometimes, I could search for the exact match. That was the easiest way. So, for example, to find the occurrences of **senzaG** and **senzaS** I could use those scripts: **senzaG** <- ‘`\\bsenza\\b`’ and **senzaS** <- ‘`\\bsanza\\b`’, and thereafter, I calculated the total number of occurrences with the function ‘length’. So, for counting the occurrences of the silver form ‘sanza’ in the letters written between 1495 and 1499, I used the script: **length(conc\_re(senzaS, a, as\_text = TRUE)\$match)**. After that, I recorded the numerical outcome in a separated file.

Sometimes, however, I could not search for the exact match. In this case, for every occurrence, I had to manually check whether the outcome was correct. I did it with the function ‘View’, like this: **View(conc\_re(CrieG, a, as\_text = TRUE))**. Below, I explain with a few examples why, under certain circumstances, it was impossible to search for the exact match in a completely automated way.

It can happen that a similar visual outcome represents different grammatical rules: for instance, a software cannot make the distinction between the form ‘scriviano’ (that belongs to the **prIVamoS** group) and the form ‘pregano’ (that belongs to the **prVIanoG** group), because—graphically—the two words have the same ending *-ano* (stressed in the first case, unstressed in the second case). The solution that I found, was to write the same script for the two features—‘`(ano\\b)`’—, and then disambiguate them on a case by case basis, depending on the context.

Sometimes—as in the case of the **CrieG**, **CrieS**, **CruoG**, **CruoS** groups—the rule applies only to words derived from Latin short vowels (for instance, from Lat. *PRĒCARI* we get *priego* in 14*F* and *prego* in 16*F*). If the Latin word has a long vowel, the outcome was a single vowel (and not a diphthong) in both 14*F* and 16*F* (for instance, from Lat. *CRĒDĒRE* we always obtain *credo*). Therefore, when the Latin word has a long vowel, the Florentine word is not included in the **CrieG** group. Of course, there was no way that I could automate a procedure to include a word such as *prego* in the **CrieG** group, while leaving aside a word such as *credo*, because the only difference relies on the Latin etymology. The best I could do was to write two scripts such as: **CrieG** <- ‘`(?mix)[bcdfgpt] (rie) [^\\b]`’ and **CrieS** <- ‘`(?mix)[bcdfgpt] (re) [^0-9] [^0-9]? [^0-9]? \\b`’, and after, check one by one all the results, discarding the words whose outcome *e* did not derive from *ĕ*, *ĝ*.

Often, I needed to know also the *meaning*—and not only the *etymology*—of the word that I was analyzing. The structure of the words ‘stiavo’ and ‘stiano’, for example, is identical, but the former is part of the **stiVS** group, while the latter is not. Similarly, a word like ‘begli’ is part of the **IliS** group, while ‘degli’ is not. The four scripts **schIVG** <- ‘`(?mix)(schi) [aeou] [^\\b]`’, **schIVS** <- ‘`(?mix)(sti) [aeou] [^\\b]`’, **IliG** <- ‘`(?mix) [aeiou] (Ili) \\b`’ and **IliS** <- ‘`(?mix) [aeiou] (gli) \\b`’ account for more results, if compared to the correct ones, and again, a manual check was needed.

When I was finally done counting the correct number of occurrences of the forms that I was searching for, I put the numeric outcomes (= the number of occurrences of every feature) in a separated file, just as I did with the **length()** function above.

Next, I did chi-squared test on the outcomes, divided by time intervals. The null hypothesis was that the golden and silver forms were randomly distributed over time.

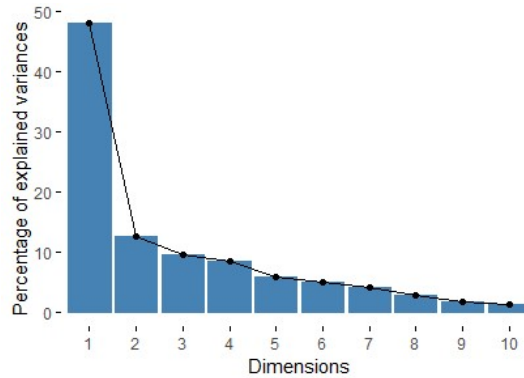


Figure 1: Correspondence analysis scree plot.

I obtained a p-value  $< 2.2^{-16}$ , far below the commonly accepted threshold (0.05).<sup>16</sup>

This means that the outcomes were statistically significant, and consequently, that the use of golden and silver forms varies in a non-random way over the years. But the test does not say in what way non-random choices affected Michelangelo’s use of golden and silver forms: that is why I needed to run also the correspondence analysis, “a statistical technique that provides a graphical representation of cross tabulations ... Cross tabulations arise whenever it is possible to place events into two or more different sets of categories.”<sup>17</sup>

With the function: `features_ca <- ca(features)` I run the correspondence analysis, and I printed the scree plot (Figure 1).

Given these results, I considered only dimension 1 (time-related), which accounts for 50% of the total.<sup>18</sup> Furthermore, since I was interested in differences among time intervals, I focused on rows (`map="rowprincipal"`, cf. Nenadic and Greenacre (2007)). At that point, I could finally print the plot of the correspondence analysis (Figure 2).

The data and the subsequent plot show a clear cut-off date around 1530: indeed, values consistently diverge before and after this date. On the left side of Figure 2, together with all the time intervals from 1495 to 1530, are grouped most of the silver forms (= 16F, recognizable by a capital ‘S’ at the end of their names). On the contrary, on the right side of the plot, together with the time intervals from 1530 to 1564, are grouped most of the golden forms (= 14F, recognizable by a capital ‘G’ at the end of the name). Moreover, since the first dimension, corresponding to the horizontal axis, is time-sensitive, I could deduce from the plot a strong separation between silver forms, most of them at the very left side of the plot, and golden forms, most of them—with the only exception of **ultimS**—at the right side.

However, correspondence analysis could be also sensitive to variation different than

<sup>16</sup> The establishment of a threshold has always been a topic of debate; many scholars steadily warned against all possible misinterpretations. In general, if the p-value is much greater or much lower than 0.05, then it is possible to reject (or not)  $H_0$ , but if its value is close to 0.05, then all we can do is *not completely reject*  $H_0$ . Fisher himself cautioned against a strict exploitation of the threshold as a benchmark for rejecting or accepting  $H_0$ : in his words, “if p is between 0.1 and 0.9 there is certainly no reason to suspect the hypothesis tested. If it is below 0.02 it is strongly indicated that the hypothesis fails to account for the whole of the facts. We shall not often be astray if we draw a conventional line at 0.05” (Fisher’s quote is taken from Biau et al. (2010): 886b). Furthermore, on “four misconceptions about what a p-value tells us,” cf. Nicenboim and Vasishth (2016a) and Nicenboim and Vasishth (2016b).

<sup>17</sup> See Yelland (2010): 1. On correspondence analysis, Greenacre’s works are still a great point of departure: see Greenacre (1984), Greenacre (1993) and Greenacre and Blasius (1994).

<sup>18</sup> An extended discussion on the contribution of the axes is in Bendixen (1995).

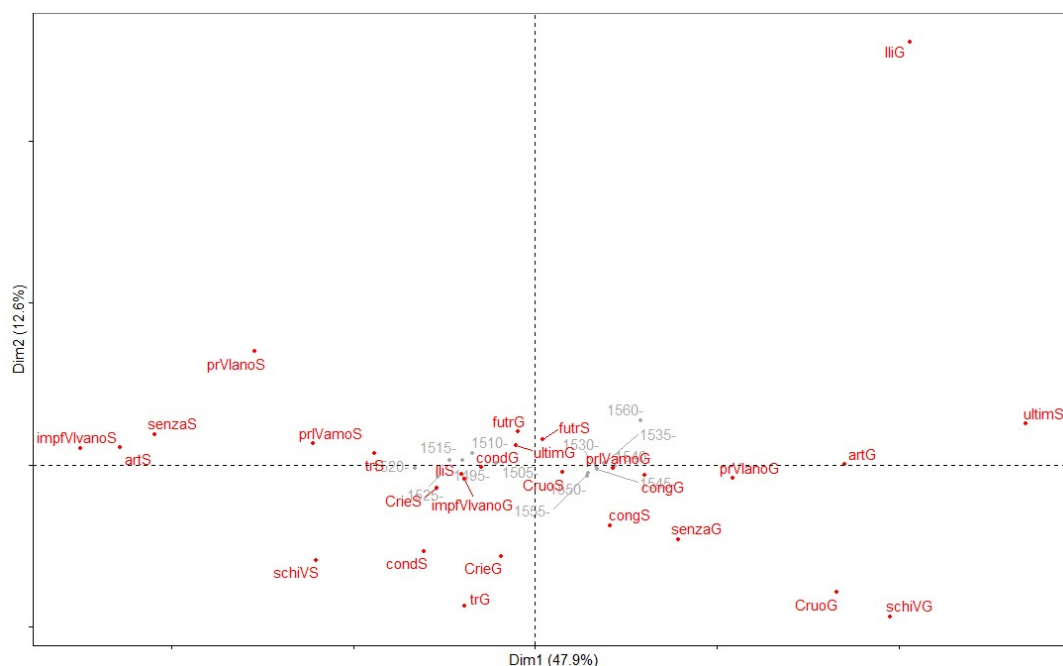


Figure 2: Correspondence analysis.

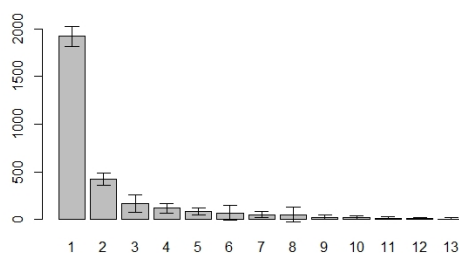


Figure 3: Correspondence regression scree plot.

time, or gold/silver variation, because the features under scrutiny are not completely independent, but come in pairs of alternative variants (*14F* vs *16F* forms). One could argue that in cases like this, the frequencies of the features are not only determined by the writer’s preference for the one or the other variant, but also by the overall frequency of the lexical items at hand, as is fully explained in Speelman et al. (2003). To address this issue, I applied correspondence regression, using the *R* package “corregp” (cf. Plevoets (2015)).<sup>19</sup>

I then reshaped the data, so to obtain Table 1. At this point, I performed a correspondence regression of the response variable “feature” in function of the main effect of “time” + the main effect of “measure” + the interaction between “time” and “measure” (cf. Plevoets (2018): 2–3). The plot in Figure 3 shows that dimension 1 (time-related) and dimension 2 (related to the golden and silvery alternation) account for most of the variation (to be precise, 79% of it).

<sup>19</sup> On the other side, Geeraerts et al. (1994), together with Geeraerts et al. (1999), provide a different approach to this issue.



Table 1: Golden and silver forms in Michelangelo's letters

Feature	Measure	Time												
		1495	1505	1510	1515	1520	1525	1530	1535	1540	1545	1550	1555	1560
art	G	1	34	13	11	3	2	31	7	79	143	46	66	38
	S	5	35	38	80	54	36	5	1	5	5	2	12	2
cond	G	3	32	21	35	19	25	16	2	12	37	19	29	12
	S	0	6	0	2	1	3	2	0	0	2	3	1	0
cong	G	0	10	13	12	4	4	7	1	14	35	7	14	4
	S	3	41	14	31	15	13	10	0	27	88	45	23	4
Crie	G	0	3	0	8	3	4	1	0	2	11	6	2	0
	S	3	82	53	82	56	36	8	3	51	82	45	43	6
Cruo	G	0	2	1	1	0	0	1	0	5	14	8	10	0
	S	2	15	9	19	8	7	5	1	17	29	12	9	5
futr	G	12	162	103	143	47	40	30	5	77	196	50	82	49
	S	0	8	9	7	0	0	4	0	2	8	4	6	1
impfVivano	G	0	5	0	1	1	0	0	0	2	1	0	1	0
	S	0	1	1	3	3	0	0	0	0	0	1	0	0
lli	G	0	0	3	2	0	0	1	0	3	7	3	2	14
	S	0	9	9	11	8	6	1	0	8	17	7	0	2
prIVamo	G	2	0	0	3	1	2	0	0	3	3	4	3	2
	S	1	12	5	8	5	3	0	0	1	6	2	2	1
prVIano	G	0	3	1	3	2	2	4	0	4	11	9	5	4
	S	0	1	2	2	0	0	0	0	0	1	0	0	0
schiV	G	0	0	0	0	0	0	0	0	0	1	0	1	0
	S	1	0	0	0	0	1	0	0	0	1	0	0	0
senza	G	0	2	1	4	7	7	3	1	8	16	8	23	4
	S	0	8	4	2	3	2	0	0	0	0	0	1	0
tr	G	0	1	1	1	3	11	2	0	3	11	2	1	0
	S	0	0	0	2	1	0	0	0	2	0	0	0	0
ultim	G	0	12	18	22	11	3	5	1	3	26	19	13	7
	S	0	0	0	0	0	0	0	0	9	0	0	1	1

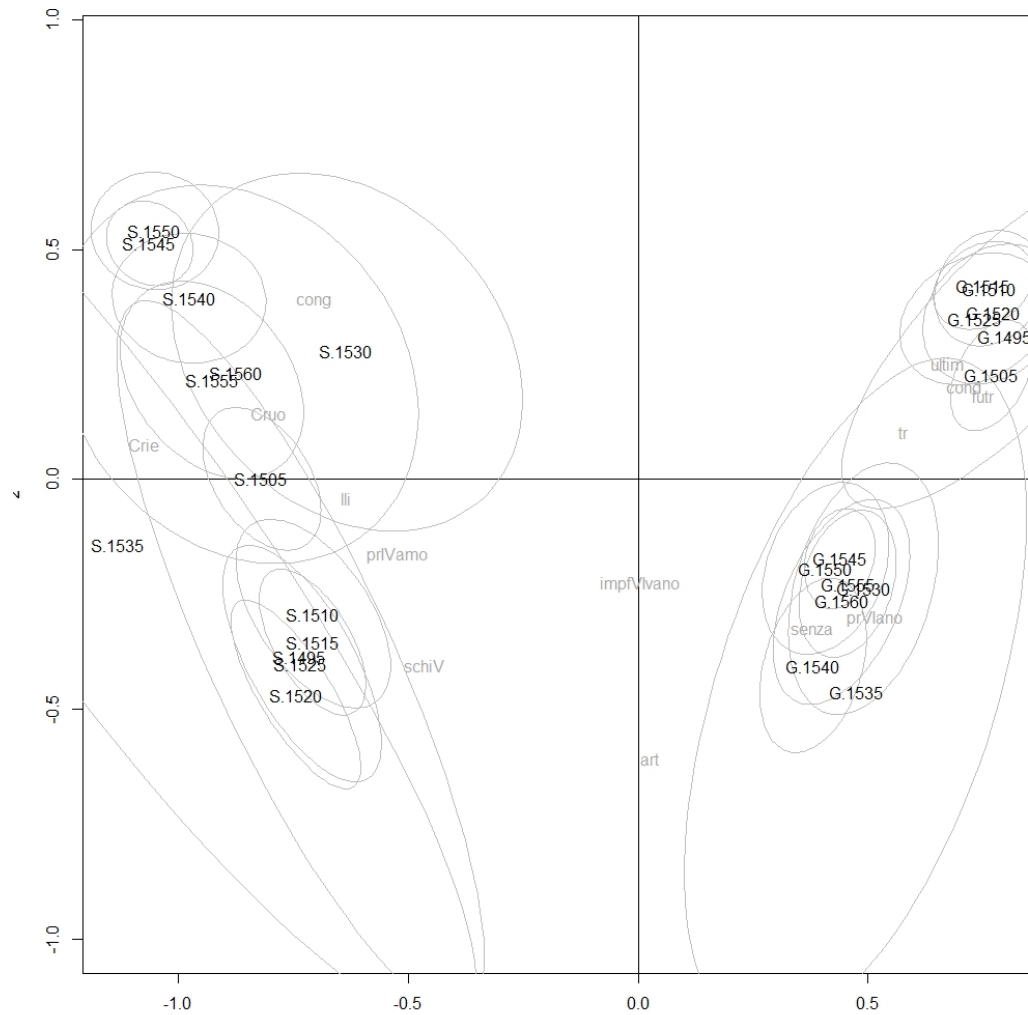


Figure 4: Correspondence regression.

I then plotted the outcomes of the correspondence regression, as in Figure 4. The plot confirms the results previously obtained with the correspondence analysis. In the horizontal axis—that accounts for 65% of the total variation—we notice a strong separation between silver (S.) and golden (G.) forms over time. But also the vertical axis shows a connection between two groups of forms:

- 1) golden forms, 1495–1530 and silver forms, 1530–1560 (top quadrants);
- 2) silver forms, 1495–1530 and golden forms, 1530–1560 (bottom quadrants).<sup>20</sup>

The results suggest, one more time, that the key period when Michelangelo started to modify his use of the language is around 1530; moreover, this outcome seems to be independent of the type of lexical items taken into account and their overall frequency.

<sup>20</sup> The only exception is S.1535 in the bottom left quadrant.

### 3 Historical Interpretation of the Statistical Tests

During his life, Michelangelo repeatedly denied his interest in the contemporary debate on language and grammar, and all previous scholars that studied his texts from a linguistic point of view have insisted that his written style is an excellent example of *16F*. I hypothesize that—despite his repeated claims to be grammatically ignorant—Michelangelo was aware of the existence of manuals prescribing the Old Florentine and may have used some *14F* features more or less deliberately. In this paper, I explored on a quantitative basis Michelangelo's use of the language in his letters.

Among many parameters, I choose to consider *diachrony*. First, I split Michelangelo's letters into documents representing time intervals. Then, I selected the most relevant features that differentiate *14F* from *16F* and I counted their occurrences in the corpus, aiming to see whether there is a difference in his use of the language, and—if so—whether this difference can be related with historical reasons.

The analysis has shown that most of the silver forms are used before 1530, and most of the golden forms are used after 1530, and that their use varies in a non-random way over time. This result is of extreme interest, because the most important Italian grammar book of that time—the *Prose della volgar lingua*, written by the renowned humanist Pietro Bembo—was published in 1525. In that book, Bembo prescribes the use of *14F* (i.e., the golden forms) as a common language for all Italian people.

So far, there was no evidence that Michelangelo ever read the *Prose*. Interestingly enough, my data shows that he consciously started using the *14F* forms a few years after the publication of that book. I argue that Michelangelo was more informed of the contemporary grammatical dispute than we previously thought. It seems reasonable to affirm that in the years following the publication of the *Prose*, Michelangelo read a copy of it, and began modifying his written language, following the *14F* rules prescribed by Bembo. He did so, because in those days he was no more a simple artisan, as he was at the beginning of his life. On the contrary, in 1525–1530 he already became a public figure, and wanted to emancipate from his humble origins: but, to do so, he needed to polish his language from the most marked *16F* phonetic and morphological features, at that time perceived as 'popular'.

The linguistic evidence showing Michelangelo's use of the *14F* forms prescribed by Bembo not only is a reasonable hypothesis, but also perfectly matches with the historical documentation. Indeed, in the 1520s and 1530s, Michelangelo was in Florence, frequenting the *Orti Oricellari* together with his friends Donato Giannotti, Battista della Palla and Antonio Brucioli.<sup>21</sup> They were all devotees to Bembo's ideals, prone to use the Old Florentine language, and they could easily have introduced Michelangelo to that linguistic system. In particular, in those years, Antonio Brucioli was translating Christian texts into vernacular language: he published the New Testament in 1530, the Psalms in 1531 and the Bible in 1532. We also know that in 1529, when Michelangelo was living in Venice, Michelangelo and Brucioli regularly met.<sup>22</sup> Not only then "it is likely that on that occasion, Michelangelo ... has been faced for the first time with the topics of the Protestant Reformation."<sup>23</sup> Moreover, I would like to emphasize that Michelangelo's Venetian stay and the reading of Brucioli's translations may have had some consequences in terms of his linguistic beliefs, too.<sup>24</sup>

---

<sup>21</sup> On Michelangelo's relationships with other intellectuals, see at least Corsaro (2008).

<sup>22</sup> See Campi (1994): 156.

<sup>23</sup> See Forcellino (2014): 240 (translation mine).

<sup>24</sup> I discussed in further detail the relationship between Brucioli and Michelangelo in Valenti (2017a): 193–195.

Furthermore, in the following years, Michelangelo and Bembo had both stayed at the papal court in Rome. Unfortunately, in the absence of documents witnessing Bembo and Michelangelo friendship, we cannot say much about it, but we are supported by Vasari's words. In the *Life of Michelangelo*, he states:

The illustrious Cardinal Polo was his close friend, and Michelangelo loved his virtue and goodness. Other friends were Cardinal Farnese and Santa Croce, who later became Pope Marcellus II; Cardinal Ridolfi and Cardinal Maffeo *and Sir Bembo*, Carpi and many other cardinals and bishops and prelates that we do not mention.<sup>25</sup>

Despite there is no evidence of it, the two of them are likely to have discussed grammar and language, and it is possible that Michelangelo showed some kind of interest in the recently printed *Prose della volgar lingua*, the grammar book that was revolutionizing the entire linguistic debate in the whole Peninsula.

Indeed, the preference given by the sculptor in those very years for the use of linguistic features characteristic of the 14F system may be a reflection of his learned dissertations and his increasing social status, and consequently, might reveal his wish to align his language with the 14F grammatical rules prescribed by Bembo in 1525. Therefore, the historical documents that witness his frequentations—starting from 1520—with the key players of the sixteenth-century grammatical dispute, confirm and support the results of the correspondence analysis and the correspondence regression.

Likewise, a few years later (1542) Michelangelo asked his friends Donato Giannotti and Luigi del Riccio to amend the language of his poems:

Sir Luigi, you who have the spirit of poetry, I beg you to shorten and improve one of these madrigals, which at the moment is imperfect, because I must give it to a friend of ours.<sup>26</sup>

As showed in Valenti (2019), he was probably asking for a review of the linguistic features that did not match with Bembo's grammatical rules. This is the last piece of evidence that call into question the old assumption that Michelangelo was unaware of the grammar books prescribing the use of fourteenth-century Florentine language. In fact, his linguistic choices did not always reflect the contemporary use and, sometimes, he was more inclined to employ the archaic forms than we would expect.

Michelangelo Buonarroti once defined himself "grammatically mistaken."<sup>27</sup> Maybe he did not know all the rules of 14F listed in the *Prose*, as other people of his time did, but this analysis shows that he was far from being completely unaware of them.

## Acknowledgements

I am deeply grateful to Pietro Mercuri (Sapienza University, Rome) for his help with the statistical tests. The codes that I used for my analyses are available here: <https://github.com/gianlucavalenti/A-Corpus-Based-Approach-to-Michelangelo-s-Epistolary-Language>.

---

<sup>25</sup> "Fu suo amicissimo lo illustrissimo cardinale Polo, innamorato Michelagnolo delle virtù e bontà di lui; il cardinale Farnese e Santa Croce, che fu poi papa Marcello; il cardinale Ridolfi e 'l cardinale Maffeo e *monsignor Bembo*, Carpi e molti altri cardinali e vescovi e prelati che non accade nominargli." See Vasari (2011) (last viewed: 01.04.2020. The italic is mine).

<sup>26</sup> "Messer Luigi, voi ch'avete spirito di poesia, vi prego che m'abbreviate e rachonciate uno di questi madrigali, quale vi pare il manco tristo, perché l'ò a dare a un nostro amico." See Barocchi and Ristori (1983), vol. IV: 144 (letter n° CMXCVIII).

<sup>27</sup> In his words: "scorrecto in gramatica." Cf. the letter sent to Luigi Del Riccio on February (?), 1544, and preserved in Florence, Archivio Buonarroti, XIII, n. 41 (See Barocchi and Ristori (1983), vol. IV: 177).

## References

- Antonelli, G., M. Motolese, and L. Tomasin  
2014. *Storia dell'italiano scritto. 3 vols. III. Italiano dell'uso*. Carocci.
- Arcangeli, M.  
2011. L'antico romanesco: gli studi recenti e una vecchia questione. *Bollettino dell'Atlante Lessicale degli antichi volgari italiani*, 4:9–41.
- Aresti, A.  
2019. *Lingua delle arti e lingua di artisti in Italia fra Medioevo e Rinascimento*. Franco Cesati.
- Auer, A., D. Schreier, and R. J. Watts  
2015. *Letter Writing and Language Change*. Cambridge University Press.
- Barocchi, P.  
1984. Storiografia artistica: lessico tecnico e lessico letterario. In *Studi vasariani*, Pp. 134–156. Einaudi.
- Barocchi, P. and R. Ristori  
1965-1983. *Il carteggio di Michelangelo. Edizione postuma di G. Poggi. 5 vols.* SPES.
- Bendixen, M.  
1995. Compositional perceptual mapping using chi-squared trees analysis and correspondence analysis. *Journal of Marketing Management*, 11:571–581.
- Bianconi, S.  
2013. *Italiano lingua popolare. La comunicazione scritta e parlata dei 'senza lettere' nella Svizzera italiana dal Cinquecento al Novecento*. Accademia della Crusca/Edizioni Casagrande.
- Biau, D. J. M., B. M. Jolles, and R. Porcher  
2010. P value and the theory of hypothesis testing: An explanation for new researchers. *Clinical Orthopaedics and Related Research*, 468:885–892.
- Bruni, F.  
1978. Traduzione, tradizione e diffusione della cultura: contributo alla lingua dei semicolti. In *Alfabetismo e cultura scritta. Atti del Seminario di Perugia (29-30 marzo 1977)*, Pp. 195–234.
- Bruni, F.  
1984. *L'italiano. Elementi di storia della lingua e della cultura*. UTET.
- Buck, A. and M. Pfister  
1971. *Studien zur Prosa des Florentiner Vulgärhumanismus im 15. Jahrhundert*. Fink.
- Campi, E.  
1994. *Michelangelo e Vittoria Colonna: un dialogo artistico-teologico ispirato da Bernardino Ochino e altri saggi di storia della Riforma*. Claudiana.
- Castellani, A.  
1967-1970. Italiano e fiorentino argenteo. *Studi linguistici italiani*, 7:3–19.

- Castellani, A.  
2000. *Grammatica storica della lingua italiana. I. Introduzione*. il Mulino.
- Ciulich, L. B.  
1973. Costanza ed evoluzione nella grafia di michelangelo. *Studi di grammatica italiana*, 3:5–138.
- Corsaro, A.  
2008. Michelangelo e i letterati. In *Officine del nuovo. Sodalizi fra letterati, artisti ed editori nella cultura italiana fra Riforma e Controriforma. Atti del Simposio Internazionale (Utrecht, 8-10 novembre 2007)*, Pp. 383–425.
- Culpeper, J. and M. Kytö  
2010. *Early Modern English Dialogues: Spoken Interaction as Writing*. Cambridge: University Press.
- Dionisotti, C.  
1966. *Pietro Bembo. Prose della volgar lingua, Asolani, Rime*. UTET.
- D'Achille, P.  
1994. L'italiano dei semicolti. In *Storia della lingua italiana. Vol. II*, Pp. 41–79. Einaudi.
- D'Achille, P. and A. Stefinlongo  
2016. Note linguistiche su un corpus di epistolari cinquecenteschi. In *Scrivere lettere nel Cinquecento. Corrispondenze in prosa e in versi*, Pp. 245–262. Edizioni di storia e letteratura.
- D'Onghia, L.  
2014. Michelangelo in prosa. sulla lingua del carteggio e dei ricordi. *Nuova Rivista di Letteratura Italiana*, 17:89–113.
- D'Onghia, L.  
2015. Fu vero stile? noterelle su michelangelo epistolografo. *L'Ellisse. Studi storici di letteratura italiana*, 10:135–146.
- Felici, A.  
2015. *Michelangelo a San Lorenzo (1515–1534). Il linguaggio architettonico del Cinquecento fiorentino*. Olschki.
- Fitzmaurice, S.  
2002. *The Familiar Letter in Early Modern English: A Pragmatic Approach*. Benjamins.
- Folena, G.  
1951. Chiaroscuro leonardesco. *Lingua nostra*, 12:57–63.
- Folena, G.  
1957. Noterelle lessicali albertiane. *Lingua Nostra*, 18:6–10.
- Forcellino, A.  
2014. *Michelangelo. Una vita inquieta*. Laterza.
- Fresu, R.  
2004. Alla ricerca delle varietà 'intermedie' della scrittura femminile tra xv e xvi secolo: lettere private di lucrezia borgia e di vannozza cattanei. *Contributi di Filologia dell'Italia mediana*, 18:14–82.

- Fresu, R.  
2014. Scritture dei semicolti. In *Storia dell'italiano scritto*. 3 vols. III. *Italiano dell'uso*, Pp. 195–223. Carocci.
- Geeraerts, D., S. Grondelaers, and P. Bakema  
1994. *The Structure of Lexical Variation. Meaning, Naming and Context*. De Gruyter.
- Geeraerts, D., S. Grondelaers, and D. Speelman  
1999. *Convergentie en divergentie in de Nederlandse woordenschat. Een onderzoek naar kleding- en voetbaltermen*. Meertensinstituut.
- Ghinassi, G.  
1961. Correzioni editoriali di un grammatico cinquecentesco. *Studi di filologia italiana*, 19:33–93.
- Greenacre, M. J.  
1984. *Theory and Applications of Correspondence Analysis*. London Academic Press.
- Greenacre, M. J.  
1993. *Correspondence Analysis in Practice*. London Academic Press.
- Greenacre, M. J. and J. Blasius  
1994. *Correspondence Analysis in the Social Sciences: Recent Developments and Applications*. London Academic Press.
- Hernández-Campoy, J. M. and N. Schilling  
2012. The application of the quantitative paradigm to historical sociolinguistics: Problems with the generalizability principle. In *The Handbook of Historical Sociolinguistics*, Pp. 63–79. Blackwell.
- Librandi, R.  
2004. Varietà intermedie di italiano in testi preunitari. *Studies in Language*, 8:77–103.
- Magro, F.  
2014. Lettere familiari. In *Storia dell'italiano scritto*, volume 3, Pp. 101–157. Carocci.
- Manni, P.  
1979. Ricerche sui tratti fonetici e morfologici del fiorentino quattrocentesco. *Studi di grammatica italiana*, 8:115–171.
- Marazzini, C.  
2015. La lingua di michelangelo. *Annali aretini*, 23:125–132.
- Memofonte  
2008. Memofonte. <http://www.memofonte.it/ricerche/michelangelo-buonarroti/>.
- Migliorini, B.  
1955. Note sulla grafia italiana nel rinascimento. *Studi di filologia italiana*, 13:258–296.
- Moreno, P. and G. Valenti  
2017. «Un pelago di scientia con amore». *Le Regole di Fortunio a cinquecento anni dalla stampa*. Salerno Editrice.

- Nenadic, O. and M. Greenacre  
2007. Correspondence analysis in r, with two- and three-dimensional graphics: The ca package. *Journal of Statistical Software*, 20:1–13.
- Nencioni, G.  
1965. La lingua di michelangelo. In *Michelangelo artista, pensatore, scrittore. Vol. II*, Pp. 569–576. Istituto Geografico De Agostini.
- Nencioni, G.  
1995. Sulla formazione di un lessico nazionale dell'architettura. *Bollettino d'informazioni del Centro di Ricerche Informatiche per i Beni Culturali*, 5:7–33.
- Nevalainen, T. and S.-K. Tanskanen  
2007. *Letter Writing*. Benjamins.
- Nicenboim, B. and S. Vasishth  
2016a. Statistical methods for linguistic research: Foundational ideas – part i. *Language and Linguistics Compass*, 10:349–369.
- Nicenboim, B. and S. Vasishth  
2016b. Statistical methods for linguistic research: Foundational ideas – part ii. *Language and Linguistics Compass*, 10:591–613.
- Palermo, M.  
1990-1992. Sull'evoluzione del fiorentino nel tre-quattrocento. *Nuovi Annali della Facoltà di Magistero dell'Università di Messina*, 8-10:131–156.
- Patota, G.  
2017. *La Quarta Corona. Pietro Bembo e la codificazione dell'italiano scritto*. il Mulino.
- Plevoets, K.  
2015. *corregp: Functions and Methods for Correspondence Regression*. Ghent University.
- Plevoets, K.  
2018. Correspondence regression: A tutorial. <https://cran.r-project.org/web/packages/corregp/vignettes/corregp.pdf>.
- Quondam, A.  
1983. La letteratura in tipografia. In *Letteratura italiana. Vol II: Produzione e consumo*, Pp. 555–686. Einaudi.
- Remacle, L.  
1948. *Le problème de l'ancien wallon*. Bibliothèque de la Faculté de Philosophie et Lettres de l'Université de Liège.
- Renzi, L. and G. Salvi  
2010. *Grammatica dell'italiano antico. 2 vols*. il Mulino.
- Salani, T. P.  
1992. La toscana. In *L'italiano nelle regioni. Lingua nazionale e identità regionali*, Pp. 402–461. UTET.
- Serianni, L.  
2007. La storia della lingua italiana, oggi. *Bollettino di italianistica n.s.*, 4:5–19.



- Speelman, D., S. Grondelaers, and D. Geeraerts  
2003. Profile-based linguistic uniformity as a generic method for comparing language. *Varieties Computers and the Humanities*, 37:317–337.
- Tavosanis, M.  
2002. *La prima stesura delle «Prose della volgar lingua»: fonti e correzioni*. ETS.
- Team, R. C.  
2020. R: A language and environment for statistical computing. r foundation for statistical computing. <https://www.R-project.org/>.
- Testa, E.  
2014. *L'italiano nascosto. Una storia linguistica e culturale*. Einaudi.
- Trovato, P.  
1991. *Con ogni diligenza corretto. La stampa e le revisioni editoriali dei testi letterari italiani (1470-1570)*. il Mulino.
- Valenti, G.  
2017a. Le lettere di michelangelo. auto-promozione e auto-percezione nel contesto del dibattito linguistico contemporaneo. *Studi di Memofonte*, 18:182–210.
- Valenti, G.  
2017b. Towards an analysis of michelangelo's epistolary language: Some remarks. *Italica*, 94:685–708.
- Valenti, G.  
2018. «de l'uso frequentato si fan norme»: L'italie au xvie siècle, entre normalité et normativité. In «*Modello, regola, ordine*». *Parcours normatifs dans l'Italie du Cinquecento*, Pp. 323–336. Presses Universitaires de Rennes.
- Valenti, G.  
2019. Le revisioni linguistiche del riccio e del giannotti alle lettere di michelangelo. In *Lingua delle arti e lingua di artisti in Italia fra Medioevo e Rinascimento*, Pp. 263–290. Franco Cesati.
- Vasari, G.  
2011. Vita di michelangelo (1568). <http://vasariscrittore.memofonte.it/home>.
- Vela, C.  
2001. *Pietro Bembo. Prose della volgar lingua. L'«editio princeps» del 1525 riscontrata con l'autografo Vaticano latino 3210*. CLUEB.
- Yelland, P. M.  
2010. An introduction to correspondence analysis. *The Mathematica Journal*, 12:1–23.