

What is Elena Ferrante?

A comparative analysis of a secretive bestselling Italian writer

Arjuna Tuzzi

Dipartimento di Filosofia, Sociologia, Pedagogia e Psicologia Applicata (FISPPA), University of Padova, Italy

Michele A. Cortelazzo

Dipartimento di Studi linguistici e letterari (DISLL), University of Padova, Italy

Correspondence:

Arjuna Tuzzi, Dipartimento FISPPA, via M. Cesarotti, 10/12, 35123 Padova, Italy.

E-mail:

arjuna.tuzzi@unipd.it

Abstract

This article looks at the case of Elena Ferrante, the (presumed) pseudonym of an internationally successful Italian novelist, and has two objectives: first, to observe how her novels are positioned in the panorama of modern Italian literature (represented by an *ad hoc* reference corpus—composed of 150 novels by forty different authors) and, second, to attempt to understand whether, amongst the authors in the corpus, there are any that can be considered candidates for involvement in the writing of the novels signed Ferrante. Consistent with these two objectives, the analyses also use two methods: correspondence analysis for the content mapping of the novels and Labbé's intertextual distances to establish a measure of similarity between the novels. In the results, we do not see the expected similarities with writers from the Naples area as Elena Ferrante distinguishes herself with original literary products that, both in terms of theme and style, show her strong individuality. Amongst the authors included, Domenico Starnone, who has been previously identified by other investigations as the possible hand behind this pen name, is the author who has written novels most similar to those of Ferrante and which, over time, has become progressively more similar.

1 Introduction

Elena Ferrante is a rather particular Italian publishing and journalistic phenomenon. She first became famous for her stories outside of Italy, mainly in the USA, while at home she became known—at least in the beginning—not as much for book sales or for recognition from academic literary critics as she did for her collection of successes overseas and significantly also for the curiosity that resulted from the question of her anonymity. She took part in the

Strega Prize (Premio Strega) in 1992 with *L'amore molesto* [Troubling Love] (Ferrante, 1992), which also became a film directed by Mario Martone. Some years later, in 2007, the Italian publishing house UTET (Turin) chose to include the book in the 100 most relevant works in the collection *Primo tesoro della lingua letteraria italiana del Novecento* [First Companion to Italian Literature in the 20th Century], compiled by Tullio De Mauro; more recently, in 2015, she was a finalist for the Strega Prize with the last volume of her *L'amica geniale* [My

Brilliant Friend] tetralogy, *Storia della bambina perduta* [The Story of the Lost Child] (Ferrante, 2014). Today, *Ferrante Fever*¹ is a successful brand for librarians and for the publishers E/O (who publish Elena Ferrante's books) and the author has much international exposure, despite still being underrepresented in academic studies.

As far as critical studies of Ferrante's works are concerned, it can be surprisingly observed that there is relatively little scientific literature that discusses her books. Of the articles we have been able to gather, they were mostly published in the USA (Mullenneaux, 2007; Milkova, 2013; Alsop, 2014; Deutsch, 2015; Bakopoulos, 2016) or in other European countries (France: Bovo-Romoeuf, 2006; UK: Segnini, 2017; Spain: Ferrer, 2016). There are some studies that evaluate the works of Ferrante from a gender studies perspective (Chemotti, 2009; Dow, 2016; Lee, 2016; Ceccoli, 2017) or that focus their analysis on representations of social life in Naples (Benedetti, 2012; Caldwell, 2012; Falotico, 2015; Ricciotti, 2016).

As far as authorship attribution is concerned, investigators have endeavoured to uncover the identity of the author from different angles: Galella (2005) and Gatto (2016) both adopted a thematic perspective that found affinities between the novels of Elena Ferrante and those of Domenico Starnone, and came to the conclusion of identifying Starnone as the author hiding behind the pseudonym, as did some studies based on quantitative linguistic approaches conducted by the Italian physicist Vittorio Loreto, who was cited by Galella (2006), by the Swiss OrphAnalytics company, cited by Rastelli (2016), and by an Italian research team (Cortelazzo *et al.*, 2016; Cortelazzo and Tuzzi, 2017). While research from a biographical perspective by Santagata (2016) concluded that Elena Ferrante is the Neapolitan historian Marcella Marmo. Lastly, by using one of the oldest methods of investigation—follow the money—Gatti (2016) claimed to have identified Anita Raja, a translator for the publishing house E/O and wife of Domenico Starnone, as the person who receives the author's royalties for the works signed Elena Ferrante.

In this study the problem of authorship attribution (AA) is only a part of the work that has been carried out, as we wished, first and foremost, to position the novels of Elena Ferrante in the panorama that is

Italian literature and understand the similarities and differences that emerge when they are compared to a collection of works and authors which should share unmistakable (thematic and lexical) common traits. Following on from ideas in the critical literature on Elena Ferrante, we have taken into consideration, on the one hand, entering the writing of Elena Ferrante in the list of gender literature as a (presumed) female writer, and on the other hand its belonging to regional literature as a (presumed) Neapolitan writer. We expect that the diastatic variation (gender) and the diatopic one (geographical origin) might play a role in terms of topics as well as linguistic features. Moreover, it is worth remembering that geographical origin plays an important role in Italian written texts because, unlike other more standardized languages, some traits of a dialect or of regional variations are recognizable, in the syntax and lexicon, even when the text is written in Italian.

2 The Corpus

Paraphrasing the Italian writer Alessandro Baricco,² we can say that, when partaking in distant reading, you are not completely done for as long as you have on your side a good corpus, and a method to examine it. The decision to study Elena Ferrante, not just as a case of AA but as a relevant literary phenomenon, determined many of the choices made in corpus collection. Even though all of the Italian novelists 'suspected' of being the hand hiding behind the veil of the pseudonym have been included in the corpus, it also takes into consideration a much wider range of authors. In contemporary Italian literature there is no immediately available reference corpus that the scientific community recognizes as exemplary and meaningful; therefore, we are obliged to bring together an *ad hoc* corpus, adopting transparent and systematic criteria. In an attempt to identify the similarities and differences between her works and those of other authors, we have put together a corpus of novels from the past thirty years, that fall into four categories: novels written by authors from the same geographical area as that declared by Elena Ferrante (the region of Campania is a valid representation of Naples and its surroundings both culturally

and linguistically); novels by authors who have been named—by scholars, writers, journalists, and essayists—as the true identity of Elena Ferrante; novels by authors who have gained public success (as can be based on the number of copies sold or the awarding of literary prizes); novels by authors whose literary value has been attested to by respected critics. Despite Anita Raja being identified as a possible author of the novels of Elena Ferrante, neither her translations nor the works of further authors that cannot be attributed to genres comparable to novels have been taken into consideration. For the same reasons, we have also not included one work by Elena Ferrante herself, *La frantumaglia* (latest, extended, edition: Ferrante, 2016), which is a collection of letters by the writer and her co-correspondents, interviews, and other materials. Nevertheless, from a qualitative viewpoint, this volume allowed us to better understand her thoughts, her relationship with writing and the context into which her works are placed. As far as the target is concerned: all of the novels were written for adult readers. For this reason, the children's tale *La spiaggia di notte* (Ferrante, 2007) [The Beach at Night] has also been excluded.

All of the authors are Italian novelists and their novels in the corpus were originally written in Italian and the majority were published between 1987 and 2016. For two authors, Dacia Maraini and Marta Morazzoni, we must resort to older novels to bring together a relevant collection of works; for Michele Prisco we need to go back to the sixties to find material that we can work with. The choice to include the latter was determined by the fact that Prisco was initially called into question as a possible secret identity for Elena Ferrante.

Elena Ferrante has published seven novels (Ferrante, 1992, 2002, 2006, 2011, 2012, 2013, 2014). The latest four make up *L'amica geniale* [My brilliant friend] tetralogy, which is a story in four volumes that tells the tale of two Neapolitan women, Elena (Lenucia or Lenù) Greco and Raffaella (Lila) Cerullo, who are united from infancy by the deepest of connections but divided by differing life choices.

Taking into account the Elena Ferrante's novels we have put together a corpus of 150 works (see Supplementary Material) written by forty different authors (Table 1) which gathers almost 10 million

word tokens ($N=9,837,851$). There are contributions from thirteen women for a total of fifty novels (in addition to Ferrante, twelve authors, and forty-three novels) and eleven authors from the Campania region who contribute forty-six novels (Ferrante, plus ten authors with thirty-nine novels). The novels' dimensions in terms of word tokens, word types, and, as a rough measure of lexical richness, an estimation of the mean number of word types per 1,000 word tokens (based on repeated measures of the number of word types in samples of 1,000 text chunks of 1,000 word tokens in length) show that the novels approximately resemble each other. The works of only two authors stand out: Eraldo Affinati, with two novels that are the richest in terms of the number of different words used; and Paolo Nori, who provides the three most redundant novels (in that they contain an abundance of colloquial language).

3 Content Mapping

It was deemed useful to have an initial distant reading (Moretti, 2013) of the novels, that would make it possible to see at a first glance the entire corpus and the reciprocal collocation of all the authors and novels included in our study. To position the literary products of Elena Ferrante in the context generated by the corpus as a whole, content mapping was conducted by means of a classical (lexical) Correspondence Analysis (CA).

Based on singular value decomposition (eigenvalues and eigenvectors) as well as Principal Component Analysis, CA (Greenacre, 1984, 2007; Lebart *et al.*, 1984, 1998; Murtagh, 2005, 2010, 2017) aims to transform the frequencies of words of a two-way contingency table into coordinates on the axes of a multidimensional space. When a contingency table 'words \times texts' reports the number of occurrences of each word type in each novel, CA positions all novels and all words in a low-dimensional space, e.g. a factorial plan, by mapping the chi-squared distance into a Euclidean distance. CA is an explorative data analysis that is widely utilized in text mining to both deepen global understanding of the main features of a large corpus and to obtain

Table 1 Authors' profile and number of novels included in the corpus

Name	Surname	Gender	Year of birth	Place of birth	Region	Number of novels
Eraldo	Affinati	m	1956	Roma	Lazio	2
Niccolò	Ammaniti	m	1966	Roma	Lazio	4
Andrea	Bajani	m	1975	Roma	Lazio	3
Marco	Balzano	m	1978	Milano	Lombardia	2
Alessandro	Baricco	m	1958	Torino	Piemonte	4
Stefano	Benni	m	1947	Bologna	Emilia Romagna	3
Enrico	Brizzi	m	1974	Bologna	Emilia Romagna	3
Gianrico	Carofiglio	m	1961	Bari	Puglia	9
Mauro	Covacich	m	1965	Trieste	Friuli Venezia-Giulia	2
Erri	De Luca	m	1950	Napoli	Campania	4
Diego	De Silva	m	1964	Napoli	Campania	5
Giorgio	Faletti	m	1950	Asti	Piemonte	5
Elena	Ferrante					7
Marcello	Fois	m	1960	Nuoro	Sardegna	3
Paolo	Giordano	m	1982	Torino	Piemonte	3
Nicola	Lagioia	m	1973	Bari	Puglia	3
Dacia	Maraini	f	1936	Fiesole	Toscana	5
Margareth	Mazzantini	f	1961	Dublino - Tivoli	Lazio	4
Melania G.	Mazzucco	f	1966	Roma	Lazio	5
Rossella	Milone	f	1979	Pompei	Campania	2
Giuseppe	Montesano	m	1959	Napoli	Campania	2
Marta	Morazzoni	f	1950	Milano	Lombardia	2
Michela	Murgia	f	1972	Cabras	Sardegna	5
Edoardo	Nesi	m	1964	Firenze	Toscana	3
Paolo	Nori	m	1963	Parma	Emilia Romagna	3
Valeria	Parrella	f	1974	Torre del Greco	Campania	2
Francesco	Piccolo	m	1964	Caserta	Campania	7
Tommaso	Pincio	m	1963	Roma	Lazio	3
Michele	Prisco	m	1920	Torre Annunziata	Campania	2
Christian	Raimo	m	1975	Roma	Lazio	2
Fabrizia	Ramondino	f	1936	Napoli	Campania	2
Ermanno	Rea	m	1927	Napoli	Campania	3
Tiziano	Scarpa	m	1963	Venezia	Veneto	4
Clara	Sereni	f	1946	Roma	Lazio	6
Domenico	Starnone	m	1943	Saviano	Campania	10
Susanna	Tamaro	f	1957	Trieste	Friuli Venezia-Giulia	5
Chiara	Valerio	f	1978	Scauri Minturno	Lazio	3
Giorgio	Vasta	m	1970	Palermo	Sicilia	2
Sandro	Veronesi	m	1959	Firenze	Toscana	4
Simona	Vinci	f	1970	Milano	Lombardia	2

an effective visualization of available textual information. CA may uncover hidden relationships among texts, and may expose correspondences and similarities among texts, among words as well as between texts and words. CA may also lead to the discovery of new knowledge.

For the corpus as a whole, this study looks at the contingency table 'words \times novels' that represents 150 novels and 24,846 word types (forms) with a frequency higher than or equal to a threshold of

twenty (with a corpus coverage rate of 94.8%, and a vocabulary coverage rate of 15.6%) and the contingency table 'words \times authors' that represents forty subcorpora that pool all of the works of a single author together. Figure 1 represents the position of authors (A) and novels (B) on the first Cartesian plane as generated by factors first and second. For the sake of simplicity and readability, labels for only 75% of the authors have been made visible on the graphic (thirty squares out of forty)

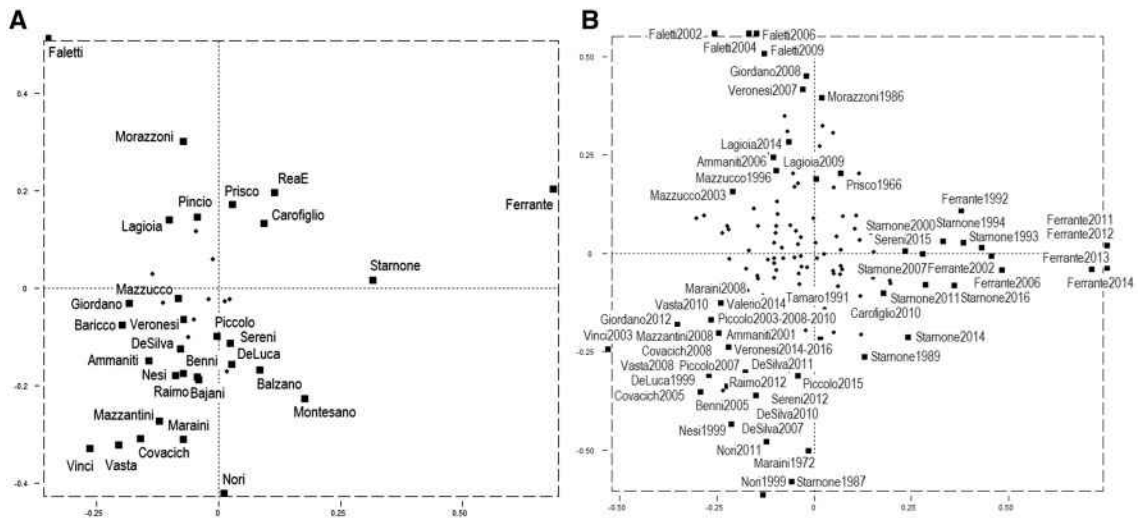


Fig. 1. First plan of CA: contingency tables for 24,846 word types with a frequency ≥ 20 per forty authors (A) and 150 novels (B). Projection of 75% of authors and 40% of novels with the highest contributions on axes

and for only 40% of the novels (sixty squares out of 150). These thirty authors and sixty novels, respectively, were chosen based on their having provided the highest contributions in determining the result. The remaining ten authors and ninety novels, which are mainly grouped in the area surrounding the origin of the axes, are represented with tiny dots that are unlabelled.

The subsequent figures depict the same analyses being carried out on a subcorpus that was selected from the larger corpus of novels.

Figure 2 shows the results that are obtained if one takes into consideration only the subcorpus made up of the thirteen female writers (A) and their corresponding fifty novels (B). The authors included in these analyses are, in addition to Elena Ferrante, Dacia Maraini, Margareth Mazzantini, Melania G. Mazzucco, Rossella Milone, Marta Morazzoni, Michela Murgia, Valeria Parrella, Fabrizia Ramondino, Clara Sereni, Susanna Tamaro, Chiara Valerio, and Simona Vinci. The analyses include a total of 22,754 word types (forms) with a frequency higher than or equal to a threshold of eight (corpus coverage rate 94.9%, vocabulary coverage rate 23.2%). In Fig. 2A all of the labels for the authors are visible (even if it can be stated that the contributions of Milone, Parrella, Sereni, and Tamaro are

amongst the lowest), while, again for the sake of readability, only 80% of the labels for the novels (forty squares out of fifty, Fig. 2B) are visible, which as before provided the highest contributions in determining the result. The remaining ten novels, which are mainly grouped in the area surrounding the origin of the axes, are represented with tiny dots that are unlabelled.

Figure 3 shows the results that are obtained if one takes into consideration only the subcorpus composed of the eleven writers from the Campania region (A) and the corresponding forty-six novels (B). The authors included in these analyses are, in addition to Elena Ferrante, Erri De Luca, Diego De Silva, Rossella Milone, Giuseppe Montesano, Valeria Parrella, Francesco Piccolo, Michele Prisco, Fabrizia Ramondino, Ermanno Rea, and Domenico Starnone. The analyses include a total of 19,800 word types (forms) with a frequency higher than or equal to a threshold of eight (corpus coverage rate 94.9%, vocabulary coverage rate 23.0%). In Fig. 3A all of the labels for these authors from Campania are visible (and we can add that the contributions of Milone and Parrella are amongst the lowest), as are the labels (Fig. 3B) for all of their novels.

Of the most interesting personalities that are set apart by CA due to some original traits, mentions

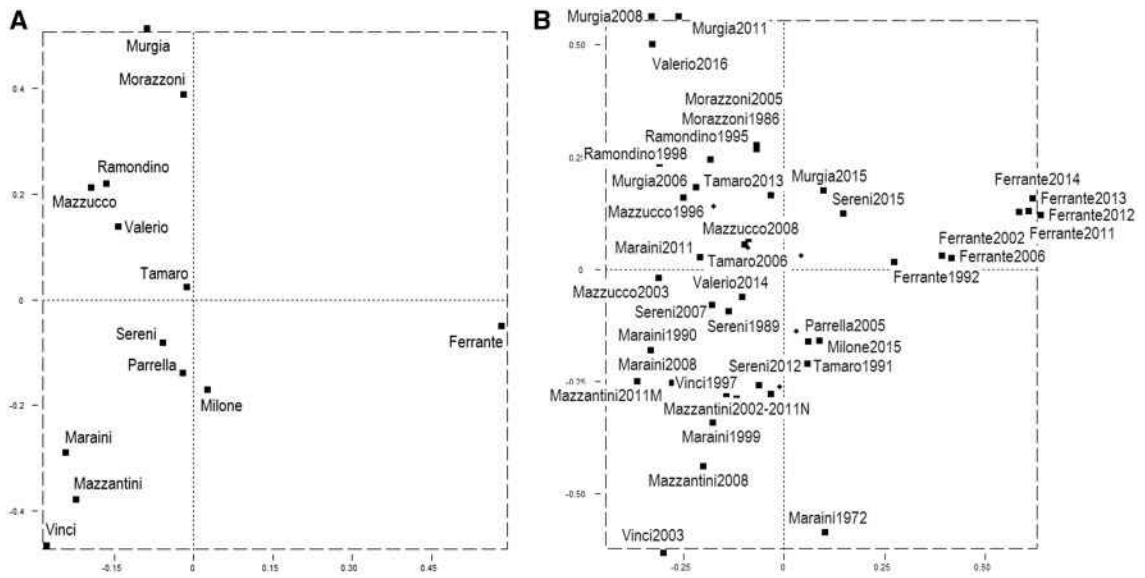


Fig. 2. First plan of CA: contingency tables for 22,754 word types with a frequency ≥ 8 per thirteen women (A) and fifty novels (B). Projection of all authors and 80% of novels with the highest contributions on axes

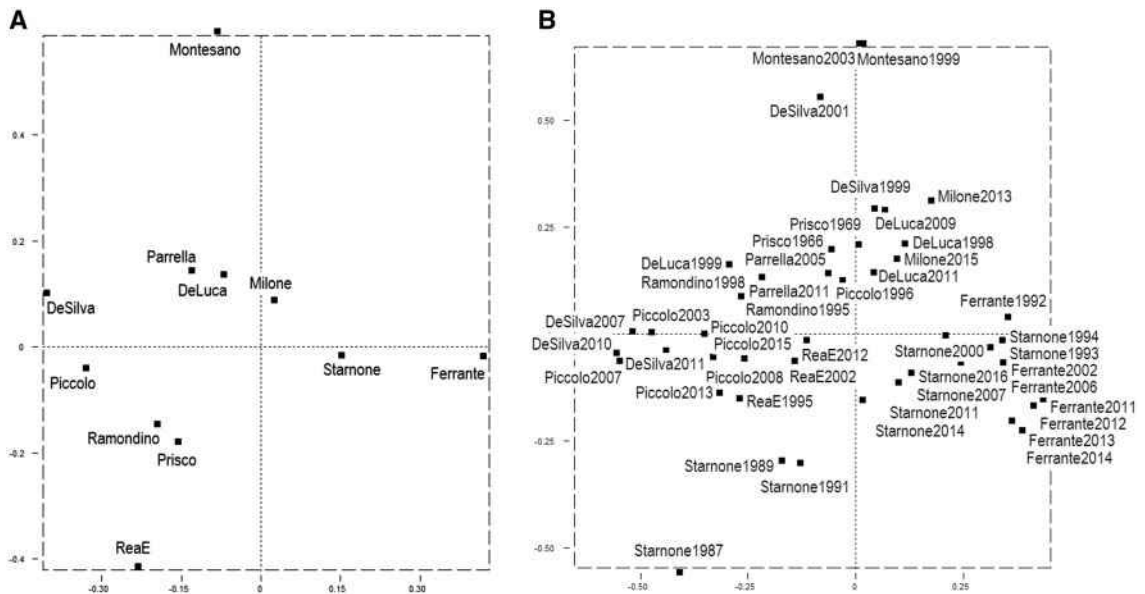


Fig. 3. First plan of CA: contingency table for 19,800 word types with frequency ≥ 8 per eleven authors from the Campania Region (A) and forty-six novels (B). Projection of all authors and all novels

should certainly be given to Paolo Nori and Giorgio Faletti in the general analysis of the 150 novels (Fig. 1), Simona Vinci and Michela Murgia amongst the women (Fig. 2), and Giuseppe Montesano, Domenico Starnone, and Ermanno Rea from the writers of Campania (Fig. 3). In all of the analyses performed, Elena Ferrante always comes out as the most unique and peculiar novelist. Furthermore, the author and her novels represent the units of the highest relevance in the determination of the first (and most important in terms of explained inertia) axis. The tetralogy *My Brilliant Friend* is always clearly visible as a compact and autonomous group, an indication that the topics and style of all four volumes are consistent as a whole when compared with the rest.

Going into more detail, the position of Elena Ferrante in the general analysis of the corpus (Fig. 1) suggests, first and foremost, a marked individuality in theme and lexicon; however, in the right hand half plane, which is dominated by Ferrante, we also see grouped together other Neapolitan or southerners writers: Prisco, Rea, Carofiglio, Starnone, De Luca, Montesano, and the Milanese writer, of southern family, Balzano. The author who appears closest, and therefore most similar in terms of lexical profile, to Elena Ferrante is Domenico Starnone. Of the non-southerners one notable presence is Clara Sereni and in particular the novel *Via Ripetta 155*.

The picture that is formed by literary works of female authors (Fig. 2) is very clear and requires little commentary: the position of Elena Ferrante, in total isolation to the right hand side of the diagram (first axis), shows a clear and absolute autonomy in relation to all the other women, with only the slightest of similarities to Rossella Milone. The configuration for the individual novels shows some moderate affinities only with specific works by Michela Murgia (*Chirù*) and Clara Sereni (*Via Ripetta 155*).

A comparison of Elena Ferrante with male authors is not presented here, as it does not add any knowledge to the general analyses of the corpus as a whole.

Regarding only the works by authors from the Campania region (Fig. 3), Elena Ferrante's novels once again appear very closely grouped together;

only her first novel (Ferrante, 1992) seems partially different in that it is located in a quadrant (upper right) that is occupied by the novels of De Luca, Milone, and Da Silva, yet it still collocates in a point of the horizontal axis (first axis) that is very close to where the other novels written by Ferrante can be found. In contrast, the novels by Domenico Starnone (Fig. 3B) are grouped together in two distinct groups: in one location are those from his first period (*Ex cattedra*, *Il salto con le aste*, *Fuori registro*), which are situated solitarily and autonomously in the third quadrant (bottom left), while the positions for his novels from 1993 on fall near those written by Elena Ferrante (bottom right quadrant). Thus, when looked at in the light of literature from Campania, the writings of Domenico Starnone appear, therefore, to be clearly divided into two blocks: The first block, that of his origins, is characterized by a stand-out uniqueness and traits of individuality that clearly set it apart; while the second block, that of the past twenty-five years, is characterized by being analogous to Ferrante's works. The turning point is found between 1991 and 1993, which is also the same period in which we see the appearance—in a style that is more uncertain and less clearly defined—of the first novel by Ferrante (1992).

4 Similarity Measures and Text Clustering

There are many methods that can be used for the purposes of AA in a computational frame (Rudman, 1998; Stamatatos *et al.*, 2001; Binongo, 2003; Grieve, 2007; Zheng *et al.*, 2006; Juola, 2008, 2012; Koppel *et al.*, 2008; Stamatatos, 2009; Savoy, 2010, 2013; Juola and Stamatatos, 2013; Rybicki *et al.*, 2016), e.g. stylometry and similarity measures (Holmes, 1998; Labbé and Labbé, 2001; Burrows, 2002; Eder 2013, 2015, 2016; Ratinaud and Marchand, 2016), compression algorithms, entropic methods, and long-range correlations (Benedetto *et al.*, 2002, 2013; Basile *et al.*, 2008; Altmann *et al.*, 2012; Oliveira *et al.*, 2013), and machine learning and profiling (Jockers and Witten, 2010; Mikros, 2013a, 2013b; Mikros and Perifanos, 2015), to cite

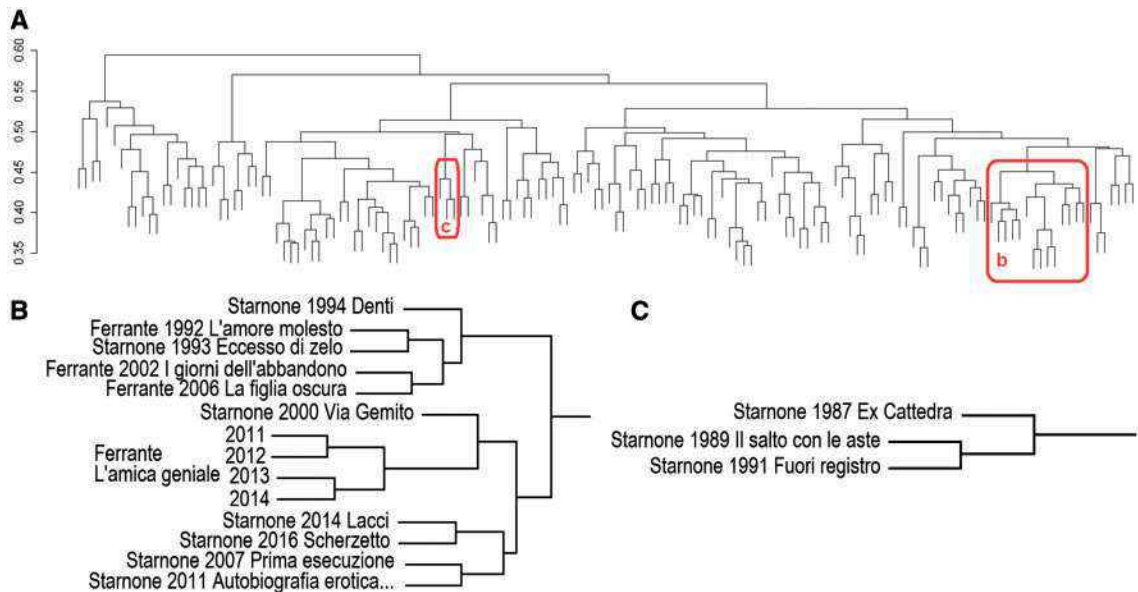


Fig. 4. Agglomerative hierarchical cluster algorithm with complete linkage (A) and close-ups (B, C)

only a few. Even though scientific literature is rich in proposals, no one method can be considered preferable to the alternatives, as the results depend on the types of text that are studied and the objectives of the analyses. AA can still be considered a field of research that is in its development phase because no standard parameters and protocols are available to compare and contrast results achieved according to different procedures (Juola, 2015).

In many instances AA is approached through the identification of a suitable and appropriate measure of the similarity of (or distance between) two texts and as a particular case of text clustering (Berry, 2004; Argamon *et al.*, 2007; Tuzzi, 2010; Ratinaud and Marchand, 2012, 2016; Savoy, 2015). On the basis of certain previous experiences (Tuzzi, 2010; Cortelazzo *et al.*, 2013, 2016), we chose Labbé's intertextual distance (Labbé and Labbé, 2001; Labbé, 2007) and, given that the iterative version proved more effective when texts are long and their size varies considerably, we also chose to work with repeated measures of distance from several samples of equal-sized chunks. In this iterative procedure a sample is extracted for each replication. For this study each sample included 150 text chunks

(one for each novel) of equal dimensions ($n=10,000$ word tokens) and Labbé's distance has been calculated for every pair of text chunks. Given a pair of text chunks A and B of the same size n , the Labbé distance d is calculated as the difference (1) between the frequencies f_i in A and in B for all of the word types of the union set of word types $V_{A \cup B}$:

$$d(A, B) = \frac{\sum_{i \in V_{A \cup B}} |f_{i,A} - f_{i,B}|}{2n} \quad (1)$$

The operation is repeated a number of times ($k=500$ replications) and the distance between two novels is the mean distance calculated on text chunks drawn from that pair of novels. Given the properties of the distance, at the end of the procedure you have a squared matrix that includes 150×150 cells and 11,175 positive non-zero non-redundant values given by the formula $150(150 - 1)/2$. This matrix can be used for the automatic classification of the texts. Figure 4 shows these distances through a dendrogram obtained by an agglomerative hierarchical cluster algorithm with complete linkage (Everitt, 1980). The dendrogram, clearly in keeping with the results of CA, identifies a cluster

Table 2 The first twenty novels closest to those of Elena Ferrante

L'amore molesto 1992	I giorni dell'abbandono 2002	La figlia oscura 2006	L'amica geniale 2011	Storia del nuovo cognome 2012	Storia di chi fugge e di chi resta 2013	Storia della bambina perduta 2014
Starnone 1993	Ferrante 2006	Ferrante 2002	Ferrante 2012	Ferrante 2011	Ferrante 2014	Ferrante 2013
Ferrante 2006	Starnone 1993	Starnone 1993	Ferrante 2014	Ferrante 2014	Ferrante 2012	Ferrante 2012
Ferrante 2002	Ferrante 1992	Ferrante 2013	Ferrante 2013	Ferrante 2013	Ferrante 2011	Ferrante 2011
Starnone 1994	Starnone 2016	Ferrante 2014	Ferrante 2006	Ferrante 2006	Ferrante 2006	Ferrante 2006
Ferrante 2011	Starnone 1994	Ferrante 2012	Ferrante 1992	Starnone 1993	Starnone 2014	Starnone 2014
Ferrante 2012	Ferrante 2013	Ferrante 1992	Starnone 1993	Starnone 2014	Starnone 1993	Starnone 1993
Starnone 2000	Starnone 2007	Ferrante 2011	Starnone 2000	Starnone 2011	Starnone 2016	Starnone 2016
Ferrante 2014	Ferrante 2012	Starnone 1994	Starnone 2014	Starnone 2016	Starnone 2011	Ferrante 2002
Ferrante 2013	Ferrante 2014	Milone 2015	Carofiglio 2004	Starnone 2000	Ferrante 2002	Starnone 2011
Milone 2015	Murgia 2015	Starnone 2011	Milone 2015	Carofiglio 2011	Murgia 2015	Ferrante 1992
Carofiglio 2004	Ferrante 2011	Starnone 2007	Starnone 2011	Ferrante 1992	Carofiglio 2011	Tamaro 1994
De Luca 1998	Milone 2015	Tamaro 1994	Balzano 2014	Ferrante 2002	Starnone 2000	Murgia 2015
Murgia 2015	Tamaro 1994	Starnone 2014	De Luca 1998	Milone 2015	Carofiglio 2010	Starnone 2000
Starnone 2007	Starnone 2011	Starnone 2016	Ferrante 2002	De Silva 1999	Starnone 2007	Piccolo 2008
Starnone 2016	De Luca 1998	De Luca 1998	Sereni 2015	Murgia 2015	Milone 2015	Carofiglio 2011
Starnone 2011	Starnone 2014	Murgia 2015	Carofiglio 2011	Balzano 2014	Ferrante 1992	Carofiglio 2004
Starnone 2014	Starnone 2000	Carofiglio 2004	Starnone 2016	Carofiglio 2004	Tamaro 1994	Carofiglio 2010
Milone 2013	Carofiglio 2004	Balzano 2014	Murgia 2015	De Luca 1998	Balzano 2014	Carofiglio 2003
Tamaro 1994	Balzano 2014	Tamaro 1991	Piccolo 1996	Piccolo 2008	Carofiglio 2004	Milone 2015
Carofiglio 2003	De Luca 2011	De Luca 2011	Carofiglio 2003	Carofiglio 2010	Carofiglio 2006	Carofiglio 2006

Note. Novels in increasing order of intertextual distance based on entire vocabularies.

that includes all of the novels written by Ferrante (B) and most of the novels by Starnone, with the exception of the first three (C) which can be found in a small cluster of their own.

Given that this tree (dendrogram), with its 150 leaves (novels) and so many branches (clusters) is difficult to read, it is worthwhile that we attempt to read the distance matrix from an alternative perspective, i.e. as a simple ranking system. The rows of the matrix (or its columns) represent the distance of a novel from all the others and thus, when given any novel, all other texts can be sorted from the closest (minimum distance) to the furthest (maximum distance). Table 2 shows the first 20 novels in order of proximity with references to the seven novels by Ferrante; Table 3 shows references to the ten novels by Starnone.

If we look at the novels closest to those claimed by Elena Ferrante (Table 2), we systematically find in the top positions other novels written by the same author and novels penned by Domenico Starnone. In particular, we see that the four volumes of the *My Brilliant Friend* tetralogy resemble each other, as one would expect from a single story in four episodes,

but in the very next positions—demonstrating the highest levels of similarity—are Starnone's novels (*Eccesso di zelo*, *Lacci*, and *Scherzetto*). The novel *Eccesso di zelo* is actually the closest to the first novel by Ferrante (1992) and lies in the second position for the following two (Ferrante, 2002, 2006).

If we look at the novels that are closest to those of Domenico Starnone (Table 3), we see that the work *Ex cattedra* shows no similarities to Ferrante's novels. In the rankings for the novel *Il salto con le aste* there appear all four volumes of the *My Brilliant Friend* tetralogy in seventh (Ferrante, 2013), ninth (Ferrante, 2012), thirteenth (Ferrante, 2011), and fourteenth (Ferrante, 2014) position. In the rankings for the novel *Fuori registro* the tetralogy appears again in eighth (Ferrante, 2011), twelfth (Ferrante, 2012), thirteenth (Ferrante, 2013), and twentieth (Ferrante, 2014) position, and we also find the inclusion in fourteenth place of a further Ferrante (2006) novel, *La figlia oscura* [The Lost Daughter]. In each of these rankings the top places are occupied by other novels by Domenico Starnone and often those of Francesco Piccolo, another Neapolitan author. However, from *Eccesso di zelo* on, at least

Table 3 The first twenty novels closest to those of Domenico Starnone

Ex Cattedra 1987	Il salto con le aste 1989	Fuori registro 1991	Eccesso di zelo 1993	Denti 1994
Starnone 1991	Starnone 1991	Starnone 1989	Ferrante 1992	Starnone 1993
Starnone 1989	Starnone 2014	Starnone 2014	Ferrante 2002	Ferrante 1992
Faletti 2011	Piccolo 2008	Piccolo 2008	Ferrante 2006	Ferrante 2006
Veronesi 2006	Piccolo 2015	Balzano 2014	Starnone 1994	Ferrante 2002
Sereni 2012	Veronesi 2006	Carofiglio 2011	Ferrante 2013	Ferrante 2011
Piccolo 2010	Carofiglio 2011	Starnone 2011	Ferrante 2014	Ferrante 2012
Piccolo 2015	Ferrante 2013	Starnone 1993	Ferrante 2012	Milone 2015
Piccolo 2007	Starnone 2011	Ferrante 2011	Starnone 2016	Ferrante 2013
Parrella 2011	Ferrante 2012	Faletti 2011	Ferrante 2011	Starnone 2007
Maraini 1999	Balzano 2014	Starnone 2007	Milone 2015	Starnone 2016
Balzano 2014	Veronesi 2014	Sereni 2012	Carofiglio 2004	Ferrante 2014
Covacich 2008	Sereni 2012	Ferrante 2012	Starnone 2007	Starnone 2011
Veronesi 2014	Ferrante 2011	Ferrante 2013	De Luca 1998	De Luca 1998
Sereni 2007	Ferrante 2014	Ferrante 2006	Starnone 2011	Starnone 2000
Giordano 2012	Starnone 1987	Piccolo 2015	Carofiglio 2003	Carofiglio 2004
Benni 2005	Starnone 2000	Tamaro 1994	De Silva 1999	Parrella 2005
Piccolo 2008	Tamaro 1994	Starnone 1987	Murgia 2015	Balzano 2014
Maraini 2008	De Silva 2010	Carofiglio 2013	Starnone 2014	De Silva 1999
Piccolo 2003	Faletti 2011	Carofiglio 2010	Carofiglio 2011	Murgia 2015
Piccolo 2013	Carofiglio 2010	Ferrante 2014	Starnone 2000	Giordano 2008
Via Gemito 2000	Prima esecuzione 2007	Autobiografia erotica di Aristide Gambia 2011	Lacci 2014	Scherzetto 2016
Ferrante 2011	Ferrante 2006	Ferrante 2006	Ferrante 2014	Ferrante 2002
Ferrante 2012	Starnone 2011	Starnone 2007	Ferrante 2013	Starnone 1993
Ferrante 1992	Ferrante 2002	Ferrante 2013	Piccolo 2008	Ferrante 2013
Ferrante 2013	Starnone 1993	Ferrante 2012	Starnone 2016	Starnone 2014
Ferrante 2014	Ferrante 2013	Ferrante 2014	Ferrante 2012	Ferrante 2014
Starnone 2007	Tamaro 1994	Starnone 2014	Tamaro 1994	Ferrante 2012
Starnone 1993	Carofiglio 2011	Ferrante 2011	Ferrante 2006	Ferrante 2006
Starnone 2011	Murgia 2015	Carofiglio 2011	Ferrante 2011	Ferrante 2011
Balzano 2010	Starnone 2016	Starnone 1993	Starnone 2011	Starnone 2007
Ferrante 2006	Ferrante 2012	Tamaro 1994	Piccolo 2015	Ferrante 1992
Ferrante 2002	Ferrante 1992	Piccolo 2008	Starnone 1989	Starnone 2011
Balzano 2014	Starnone 2000	Murgia 2015	Carofiglio 2011	Milone 2015
Murgia 2015	Ferrante 2014	Carofiglio 2013	Carofiglio 2010	Balzano 2014
Starnone 2016	Tamaro 2006	Ferrante 1992	Starnone 1993	Starnone 1994
Carofiglio 2011	Carofiglio 2003	Starnone 2000	Starnone 1991	Murgia 2015
Starnone 2014	Carofiglio 2004	Ferrante 2002	Tamaro 1991	Carofiglio 2002
Carofiglio 2004	Starnone 2014	Starnone 2016	Balzano 2014	Carofiglio 2011
Milone 2015	Starnone 1994	De Silva 1999	Scarpa 2010	Carofiglio 2004
De Silva 1999	Carofiglio 2013	Milone 2015	Ferrante 2002	Carofiglio 2006
Tamaro 1994	Ferrante 2011	Balzano 2014	Murgia 2015	De Silva 1999

Note. Novels in increasing order of intertextual distance based on entire vocabularies.

one piece of work by Elena Ferrante takes first position in the rankings, almost suggesting that she has more in common with Domenico Starnone than he does with himself.

A limitation of this analysis based on the entire lexicon is that of including in the distance an

enormous quantity of information that is relative to the contents of the novels, mostly conveyed through verbs, adjectives, and nouns (covering most of the vocabulary). As a consequence, there could be objections that it is not so much the style of the authors but rather the story, the plot, and the setting that give

Table 4 The first twenty novels closest to those of Elena Ferrante

L'amore molesto 1992	I giorni dell'abbandono 2002	La figlia oscura 2006	L'amica geniale 2011	Storia del nuovo cognome 2012	Storia di chi fugge e di chi resta 2013	Storia della bambina perduta 2014
Starnone 1993	Starnone 1994	Starnone 1994	Ferrante 2012	Ferrante 2011	Ferrante 2014	Ferrante 2013
Starnone 1994	Starnone 1993	Starnone 1993	Ferrante 2014	Ferrante 2014	Ferrante 2012	Ferrante 2012
Starnone 1991	Ferrante 2006	Ferrante 2002	Starnone 2014	Ferrante 2013	Ferrante 2011	Ferrante 2011
Starnone 2000	Starnone 2016	Ferrante 1992	Ferrante 2013	Starnone 2014	Starnone 2014	Starnone 2014
Ferrante 2011	Starnone 2007	Milone 2015	Starnone 1989	Starnone 1989	Starnone 1993	Starnone 1993
Ferrante 2006	Ferrante 1992	Starnone 2011	Starnone 1991	Ferrante 1992	Starnone 1989	Starnone 1989
Starnone 1989	Murgia 2015	Ferrante 2013	Ferrante 1992	Starnone 1991	Starnone 2016	Ferrante 1992
Ferrante 2013	Ferrante 2013	Starnone 1991	Starnone 1993	Starnone 1993	Ferrante 1992	Starnone 2016
Murgia 2015	Starnone 1991	Ferrante 2014	Starnone 1994	Starnone 2016	Starnone 1991	Starnone 2011
Starnone 2016	Covacich 2008	Ferrante 2011	Starnone 2000	Starnone 2011	Starnone 2011	Ferrante 2006
Starnone 2011	Starnone 2000	Starnone 2007	Balzano 2014	Starnone 1994	Ferrante 2006	Starnone 1994
Ferrante 2002	Starnone 2011	Starnone 2014	Starnone 2011	Balzano 2014	Murgia 2015	Starnone 1991
Ferrante 2012	De Luca 1998	Starnone 2016	Starnone 2016	Starnone 2000	Starnone 2000	Starnone 2000
Ferrante 2014	Milone 2015	Ferrante 2012	Sereni 2012	Milone 2015	Starnone 1994	Piccolo 2008
Starnone 2014	Ferrante 2012	De Luca 1998	Veronesi 2006	Parrella 2015	Balzano 2014	Sereni 2012
De Luca 1998	Covacich 2005	Sereni 2012	Ferrante 2006	Veronesi 2006	Milone 2015	Balzano 2014
Faletti 2011	Ferrante 2014	Balzano 2014	Milone 2015	Ferrante 2006	Ferrante 2002	Veronesi 2006
Starnone 2007	Faletti 2011	Murgia 2015	Parrella 2015	Sereni 2012	Raimo 2012	Milone 2015
Balzano 2014	Ferrante 2011	Tamaro 1994	Piccolo 2008	Piccolo 2008	Veronesi 2014	Murgia 2015
Carofiglio 2003	Balzano 2014	Covacich 2005	Veronesi 2014	Murgia 2015	Sereni 2012	Ferrante 2002

Note. Novels in increasing order of intertextual distance based on grammatical words.

rise to similarities between their novels. In the case of Starnone, the fact he centred some of his works around families based in Naples in the 1950s and following decades, which is also true for Ferrante's novels, may make it inevitable—and potentially of little importance—there are many lexical affinities.

To create a distance that disregards the content of the novels, it is possible to perform the calculations considering only the grammatical words. In practice, a calculation based on replications of the intertextual distance was carried out with an initial selection from the vocabulary where only occurrences of articles, prepositions, conjunctions, and pronouns were considered (430 items³). Literature has shown us that grammatical vocabulary is a crucial element for recognizing the hand of an author and a much clearer indication than that of content words (Binongo, 2003; Argamon and Levitan, 2005; Zhao and Zobel, 2005; Argamon *et al.*, 2007; Stamatas, 2009; Tuzzi, 2010). It could be hypothesized that this is true because grammatical words tend to be used unconsciously and, rather than conveying the most superficial and contingent parts of a person's lexicon, carry in them the deepest traits of

syntax. In other words, an author's hand is exposed more by the grammar than the content.

Table 4 shows the first twenty novels, ordered from the closest to the furthest, with reference to Ferrante's novels and to the intertextual distance calculated with the iterative method on the occurrences of grammatical words ($k = 500$ replications, $n = 5,000$ word tokens). Table 5 contains the orders with reference to Starnone's novels.

Once they have been cleansed of the influence of the content, the lines of research that have been traced by an analysis of the entire vocabulary become clearer. The Ferrante-Starnone link becomes more evident; for example, the inclusion of novels by any other authors in the top positions of the rankings is reduced for both writers (Tables 4 and 5). For Starnone's novels, the similarity with Ferrante in terms of writing style starts to become obvious sooner, as the novels by Ferrante occupy the first places beginning with *Il salto con le aste* and *Fuori registro*, and their presence in these positions is confirmed and strengthened by the measures of similarity based only on grammatical words.

Table 5 The first twenty novels closest to those of Domenico Starnone

Ex Cattedra 1987	Il salto con le aste 1989	Fuori registro 1991	Eccesso di zelo 1993	Denti 1994
Starnone 1989	Starnone 1991	Starnone 1993	Starnone 1994	Starnone 1993
Starnone 1991	Ferrante 2011	Starnone 1989	Ferrante 1992	Ferrante 2006
Brizzi 2008	Ferrante 2012	Ferrante 1992	Starnone 1991	Ferrante 1992
Balzano 2010	Ferrante 2013	Starnone 1994	Ferrante 2006	Starnone 1991
Maraini 2011	Ferrante 1992	Ferrante 2011	Starnone 2016	Ferrante 2002
Brizzi 1994	Starnone 1993	Starnone 2000	Ferrante 2002	Starnone 2016
Brizzi 2015	Ferrante 2014	Starnone 2011	Starnone 2011	Starnone 2011
Piccolo 2013	Veronesi 2006	Ferrante 2012	Ferrante 2013	Ferrante 2011
Benni 1987	Starnone 2000	Ferrante 2013	Ferrante 2011	Starnone 2007
Starnone 2000	Starnone 2014	Starnone 2014	Milone 2015	Balzano 2014
Parrella 2005	Starnone 1987	Brizzi 2015	Ferrante 2014	Milone 2015
Carofiglio 2002	Veronesi 2014	Starnone 2016	Starnone 2007	De Luca 1998
Covacich 2005	Starnone 2011	Starnone 2007	Starnone 1989	Ferrante 2012
Mazzucco 2005	Balzano 2014	Balzano 2014	Balzano 2014	Ferrante 2013
Benni 2005	Parrella 2005	Faletti 2011	Ferrante 2012	Starnone 2000
Balzano 2014	Carofiglio 2011	De Luca 1998	De Luca 1998	Ferrante 2014
Covacich 2008	Balzano 2010	Murgia 2015	Starnone 2014	Starnone 1989
Starnone 2007	Starnone 2016	Ferrante 2006	Sereni 2012	Sereni 2012
Veronesi 2007	Starnone 1994	Maraini 2011	Murgia 2015	Faletti 2011
Ferrante 2011	Sereni 2012	Sereni 2015	Starnone 2000	De Luca 2011
Via Gemitto 2000	Prima esecuzione 2007	Autobiografia erotica di Aristide Gambia 2011	Lacci 2014	Scherzetto 2016
Ferrante 1992	Starnone 2011	Starnone 1993	Ferrante 2014	Starnone 1993
Starnone 1991	Ferrante 2002	Starnone 2007	Ferrante 2011	Ferrante 2002
Starnone 1989	Starnone 1993	Starnone 1991	Ferrante 2013	Ferrante 2013
Ferrante 2011	Murgia 2015	Starnone 1994	Ferrante 2012	Starnone 1994
Brizzi 2015	Starnone 1994	Ferrante 1992	Starnone 2016	Ferrante 1992
Balzano 2010	Starnone 1991	Ferrante 2011	Starnone 1991	Balzano 2014
Ferrante 2013	Ferrante 1992	Ferrante 2013	Ferrante 1992	Starnone 2014
Starnone 2011	Tamaro 2006	Ferrante 2012	Starnone 1989	Starnone 1991
Starnone 1993	Scarpa 2016	Ferrante 2006	Starnone 1993	Ferrante 2014
Ferrante 2012	Ferrante 2006	Starnone 1989	Ferrante 2006	Ferrante 2012
Starnone 2016	Starnone 2016	Starnone 2016	Sereni 2012	Ferrante 2011
Starnone 1994	Starnone 2000	Ferrante 2014	Balzano 2014	Starnone 2011
Starnone 2007	Covacich 2008	Starnone 2000	Starnone 2011	Faletti 2011
Murgia 2015	Carofiglio 2014	Murgia 2015	Raimo 2012	Starnone 2000
Ferrante 2014	Faletti 2011	Starnone 2014	Starnone 1994	Sereni 2012
Veronesi 1995	Carofiglio 2011	Milone 2015	Carofiglio 2010	Murgia 2015
Ferrante 2002	Carofiglio 2014	Carofiglio 2011	Piccolo 2008	Starnone 2007
Maraini 2011	Giordano 2012	De Luca 1998	Murgia 2015	Ferrante 2006
Giordano 2014	Carofiglio 2002	Ferrante 2002	Starnone 2000	Carofiglio 2002
Veronesi 2007	Faletti 2002	Covacich 2005	Veronesi 2006	Milone 2015

Note. Novels in increasing order of intertextual distance based on grammatical words.

5 An Example of Qualitative Analysis

Quantitative methods are called upon to detect the lexical profile of a given text and are based on the

idea that style is determined by both the words and the syntactic structures that a writer decides to use—either consciously or unconsciously—when drafting his or her text and that these individual traits are, to some extent, measurable.

Table 6 Three examples of words

lemma	occ. corpus	occ. Ferrante	occ. Starnone	occ. Others
risatella	30	20	10	0
sfottente	73	43	28	2
malodore	17	12	5	0
maleodore	13	0	0	13

Note. Occurrences in the corpus as a whole and in the collections of novels written by Ferrante, Starnone, and other authors.

After all of these statistical analyses, we cannot do without an acknowledgement of a more exquisitely qualitative analysis of the texts that have been studied through concordances. There are no surprises: there are numerous examples of words that are present only in Ferrante and Starnone, or mostly in Ferrante and Starnone, with isolated occurrences from other authors or alternate variations that are used seldom by other authors. We can show (Table 6) three emblematic examples of words that are not common in Italian but occur in Elena Ferrante and Domenico Starnone's novels: *risatella* [little laugh, noun], present, both in the singular and the plural, only in Ferrante and Starnone; *sfottente* [teasing, adjective], extensively seen in Ferrante (forty-three occurrences) and Starnone (twenty-eight occurrences) and found in the remainder of the entire corpus only twice more (in two novels by two different authors); *malodore* [stink, noun], this variation is present solely in Ferrante and Starnone, while in the rest of the corpus only one author, Francesco Piccolo, uses a similar word with a variation in the spelling, *maleodore*.

Even qualitative analyses, therefore, lead us to reach a similar conclusion.

6 Discussion

In the analyses carried out with CA methods, the first axis is monopolized by the author Elena Ferrante and her novels, which, to some extent, demonstrate a certain individual trait in the overall panorama of modern Italian literature. Amongst the women, the uniqueness of the novels by Ferrante clearly stands out, as they systematically appear separated from the others (it is almost like she is not a woman). And when compared to other

authors from Campania the only author that can be considered similar is Domenico Starnone (thus regional traits do not appear to be particularly relevant). The case of Starnone-Ferrante is at its most intriguing when looking closely at Fig. 3. What we see are two moments in the history of the literary production of these authors—clearer for Starnone, much vaguer for Ferrante—which bring the writers progressively closer together as they begin to reveal increasingly overlapping and indistinguishable traits. The Starnone-Ferrante similarity emerges as the priority, over comparisons of gender or of Neapolitan origins, and can also be placed on a common time scale in the early years of the nineties.

Additionally, Labbé's intertextual distance also recognizes the novels of Domenico Starnone as those closest to the novels of Elena Ferrante and, most importantly, in the classifications calculated by the hierarchical cluster analysis with complete linkage, the works of the two authors tends to cross-over. Working with the entire vocabulary, which puts together similarities in content with those in writing style, we discover that the works that are most similar to those of Elena Ferrante are, once again, those by Domenico Starnone. More interesting still, is the fact that the closest novels to Domenico Starnone's are, after his own early writing experiences, those by Elena Ferrante. Furthermore, in some cases, we see that Elena Ferrante has written novels that are more similar to those by Starnone, than other novels written by Starnone himself. Working with only the grammatical words, thereby eliminating any possible similarities due to factors linked to historical, social, and cultural contexts and general content in the novels, the similarity between Starnone and Ferrante is reinforced and the classification method mix up Ferrante and Starnone as if dealing with the same hand.

7 Conclusions

The language and style of Elena Ferrante prove to be quite unique when compared to the general panorama of modern Italian literature that emerged from our corpus of novels. At least in theory, our corpus should reflect a context made up of authors that are comparable to the Ferrante phenomenon and Elena Ferrante stands out as having a linguistic profile that is clearly defined and undeniably peculiar. To what this individuality is due, is not clear. Is Elena Ferrante representative of a different genre? Is she the initiator of a new method of writing? Or is her work the fruit of a clear sit-down project, a meeting of minds—and hands—that consequentially resulted in writing that is unique to drafting and editing methods? We do not have an answer to these questions as they require further research—both quantitative and qualitative.

There is also the possibility that the author who hides behind the pen name Elena Ferrante has never been identified or taken into consideration, either in our research or in that of other scholars. It is also possible their only creative act has been the writing of the novels signed by Elena Ferrante; if he or she is a writer who has never written other novels under their real name, then we would be lacking the textual data necessary for comparison against the work produced by Elena Ferrante. In this case it would be practically impossible to identify the author.

From our analyses, which confirm some previous discussions, one piece of incontestable information nevertheless emerges: there is a strong affinity between the writings of Elena Ferrante and those of Domenico Starnone, to the point that, in graphical representations of the measures of similarity, the novels of these two authors are almost inextricably intertwined. This result has been confirmed in all of the analyses we have carried out, with different methods: Of the thirty-nine authors that have been taken into consideration, Starnone is the only author who demonstrates clear-cut and consistent similarities with Ferrante. This affinity cannot be simply explained on the grounds of belonging to a common social-cultural group. At most, it may only make sense in the light of a

strong link between Starnone and the author who signs their work as Elena Ferrante.

Domenico Starnone, both in one of his books⁴ and his interactions with the press, has strenuously denied all of the hypotheses that implicate him as the mysterious hand behind the pseudonym of Elena Ferrante. According to what Starnone (2011) wrote in his novel *Autobiografia erotica di Aristide Gambia* [Erotic Autobiography of Aristide Gambia] he and Ferrante are similar because they tell the stories of similar families, set in a context that is the fruit of their common experiences in Naples and of a precise period in history. But in truth, it is rather difficult to imagine that Starnone has not played any role in the planning and/or the drafting of Ferrante's work. It is difficult to precisely define his role: he could also be just one of the hands and heads that have made up the Ferrante phenomenon, but he has left his mark on it in some way. There is a good chance that Domenico Starnone knows 'who is', or rather, 'what is' Elena Ferrante.

Supplementary Data

Supplementary data are available at *LLC* online.

References

- Alsop, E.** (2014). Femmes Fatales: “La fascinazione di morte” in Elena Ferrante’s “L’amore molesto” and “I giorni dell’abbandono”. *Italica*, **91**(3): 466–85.
- Altmann, E. G., Cristadoro, G., and Degli Esposti M.** (2012). On the origin of long-range correlations in texts. *Proc Natl Acad Sci USA*, **109**(29): 11582–7.
- Argamon, S. and Levitan, S.** (2005). Measuring the usefulness of function words for authorship attribution. In *Proceeding of the 2005 ACH/ALLC Conference*. Victoria, BC, Canada, June 2005.
- Argamon, S., Whitelaw, C., Chase, P., Raj Hota S., Garg, N., and Levitan, S.** (2007). Stylistic text classification using functional lexical features. *Journal of the American Society for Information Science and Technology*, **58**(6): 802–22.
- Bakopoulos, N.** (2016). We are always us: the boundaries of Elena Ferrante. *Michigan Quarterly Review*, **55**(3): 396–419.

- Baricco, A.** (1994). *Novecento. Un monologo*. Milano: Feltrinelli.
- Basile, C., Benedetto, D., Caglioti, E., and Degli Esposti, M.** (2008). An example of mathematical authorship attribution. *Journal of Mathematical Physics*, **49**(12): 125211.
- Benedetti, L.** (2012). Il linguaggio dell'amicizia e della città: L'amica geniale di Elena Ferrante tra continuità e cambiamento. *Quaderni d'italianistica*, **33**(2): 171–87.
- Benedetto, D., Caglioti, E., and Loreto, V.** (2002). Language trees and zipping. *Physical Review Letters*, **88**(4): 487021–4.
- Benedetto, D., Degli Esposti, M., and Maspero, G.** (2013). Authorship attribution and small scales analysis applied to a real philological problem in Greek patristics. In Obradović, I., Kelih, E., and Köhler, R. (eds), *Methods and Applications of Quantitative Linguistics. Selected papers of the 8th International Conference on Quantitative Linguistics (QUALICO), Belgrade, Serbia, Aprile, 2012*. Belgrade: Academic Mind, pp. 11–20.
- Berry, M. W.** (ed.) (2004). *Survey of Text Mining. Clustering, Classification, and Retrieval*. New York, NY: Springer-Verlag.
- Binongo, J. N. G.** (2003). Who wrote the 15th book of Oz? An application of multivariate analysis to authorship attribution. *Chance*, **16**(2): 9–17.
- Bovo-Romoeuf, M.** (2006). Sensualité et obscénité dans “L'amore molesto” et “I giorni dell'abbandono” d'Elena Ferrante. *Cahiers d'études italiennes*, **5**: 129–138.
- Burrows, J. F.** (2002). Delta: a measure of stylistic difference and a guide to likely authorship. *Literary and Linguistic Computing*, **17**(3): 267–87.
- Caldwell, L.** (2012). Imagining naples: the senses of the city. In Bridge G. and Watson S. (eds), *The New Blackwell Companion to the City*. Malden: Wiley-Blackwell, pp. 337–46.
- Ceccoli, V. C.** (2017). On being bad and good: my brilliant friend. *Studies in Gender and Sexuality*, **18**(2): 110–14.
- Chemotti S.** (2009). *L'inchiostro bianco. Madri e figlie nella narrativa italiana contemporanea*. Padova: Il Poligrafo.
- Cortelazzo, M. A., Nadalutti, P., Ondelli, S., and Tuzzi, A.** (2016). *Authorship Attribution and Text Clustering for Contemporary Italian Novels*. QUALICO 2016 book of abstracts. Germany: University of Trier, August 2016.
- Cortelazzo, M. A., Nadalutti, P., and Tuzzi, A.** (2013). Improving Labbé's intertextual distance: testing a revised version on a large corpus of Italian literature. *Journal of Quantitative Linguistics*, **20**(2): 125–52.
- Cortelazzo, M. A. and Tuzzi, A.** (2017). Sulle tracce di Elena Ferrante: questioni di metodo e primi risultati. In Palumbo, G. (ed.), *Testi, corpora, confronti interlinguistici: approcci qualitativi e quantitativi*. Trieste: EUT Edizioni Università di Trieste, pp. 11–24.
- Deutsch, A.** (2015). Fiction in review: Elena Ferrante. *Yale Review*, **103**(2): 158–65.
- Dow, G.** (2016). The ‘biographical impulse’ and pan-European women's writing. In Batchelor J. and Dow G. (eds), *Women's Writing, 1660-1830: Feminisms and Futures*. London: Palgrave Macmillan, pp. 193–213.
- Eder, M.** (2013). Mind your corpus: systematic errors in authorship attribution. *Literary and Linguistic Computing*, **28**(4): 603–14.
- Eder, M.** (2015). Does size matter? Authorship attribution, small samples, big problem. *Digital Scholarship in the Humanities*, **30**(2): 167–82.
- Eder, M.** (2016). Rolling stylometry. *Digital Scholarship in the Humanities*, **31**(3): 457–69.
- Everitt, B.** (1980). *Cluster analysis*. New York, NY: Halsted Press.
- Falotico, C.** (2015). Elena Ferrante: Il ciclo dell'Amica geniale tra autobiografia, storia e metaletteratura. *Forum Italicum*, **49**(1): 92–118.
- Ferrante, E.** (1992). *L'amore molesto*. Roma: E/O.
- Ferrante, E.** (2002). *I giorni dell'abbandono*. Roma: E/O.
- Ferrante, E.** (2006). *La figlia oscura*. Roma: E/O.
- Ferrante, E.** (2007). *La spiaggia di notte*. Roma: E/O.
- Ferrante, E.** (2011). *L'amica geniale. Infanzia, adolescenza*. Roma: E/O.
- Ferrante, E.** (2012). *Storia del nuovo cognome. L'amica geniale volume secondo*. Roma: E/O.
- Ferrante, E.** (2013). *Storia di chi fugge e di chi resta. L'amica geniale volume terzo*. Roma: E/O.
- Ferrante, E.** (2014). *Storia della bambina perduta. L'amica geniale volume quarto*. Roma: E/O.
- Ferrante, E.** (2016). *La Frantumaglia*. Roma: E/O.
- Ferrer, M. R.** (2016). The role of the mirror in Elena Ferrante's novel *I giorni dell'abbandono*. *Cuadernos de Filología Italiana* **23**: 221–36.
- Galella, L.** (2005). Ferrante-Starnone. Un amore molesto in via Gemito. *La Stampa*. Torino, 16 January 2005, p. 27.
- Galella, L.** (2006). Ferrante è Starnone. Parola di computer. *L'Unità*. Roma, 23 November 2006.

- Gatti, C.** (2016). Elena Ferrante, le «tracce» dell'autrice identificata. *Il Sole 24 Ore – Domenica*. Milano, 2 October 2016, pp. 1–2.
- Gatto, S.** (2016). *Una biografia, due autofiction. Ferrante-Starnone: cancellare le tracce*. Lo Specchio di carta. Osservatorio sul romanzo italiano contemporaneo, 22 October 2016. www.lospecchiodicarta.it (accessed 19 October 2017).
- Greenacre, M. J.** (1984). *Theory and Application of Correspondence Analysis*. London: Academic Press.
- Greenacre, M. J.** (2007). *Correspondence Analysis in Practice*. London: Chapman & Hall.
- Grieve, J.** (2007). Quantitative authorship attribution: an evaluation of techniques. *Literary and Linguistic Computing*, **22**(3): 251–70.
- Holmes, D. I.** (1998). The evolution of stylometry in humanities scholarship. *Literary and Linguistic Computing*, **13**(3): 111–17.
- Jockers M. L. and Witten D. M.** (2010). A comparative study of machine learning methods for authorship attribution. *Literary and Linguistic Computing*, **25**(2): 215–23.
- Juola, P.** (2008). *Authorship Attribution*, vol. 1(3). Foundations and Trends® in Information Retrieval, pp. 233–334.
- Juola, P.** (2012). Large-scale experiments in authorship attribution. *English Studies*, **93**: 275–83.
- Juola, P.** (2015). The Rowling Case: A Proposed Standard Analytic Protocol for Authorship Questions. *Digital Scholarship in the Humanities*, **30**(1): 100–13.
- Juola, P. and Stamatatos, E.** (2013). Overview of the authorship identification task. In *Proceedings of PAN/CLEF 2013*. Valencia, Spain.
- Koppel, M., Schler, J., and Argamon, S.** (2008). Computational methods in authorship attribution. *Journal of the American Society for Information Science and Technology*, **60**(1): 9–26.
- Labbé, C. and Labbé, D.** (2001). Inter-textual distance and authorship attribution. Corneille and Molière. *Journal of Quantitative Linguistics*, **8**(3): 213–31.
- Labbé, D.** (2007). Experiments on authorship attribution by intertextual distance in English. *Journal of Quantitative Linguistics*, **14**(1): 33–80.
- Lebart, L., Morineau, A., and Warwick, K. M.** (1984). *Multivariate Descriptive Statistical Analysis. Correspondence Analysis and Related Techniques for Large Matrices*. New York, NY: Wiley.
- Lebart, L., Salem, A., and Berry, L.** (1998). *Exploring Textual Data*. Dordrecht: Kluwer Ac. Pub.
- Lee, A.** (2016). Feminine identity and female friendships in the 'Neapolitan' novels of Elena Ferrante. *British Journal of Psychotherapy*, **32**(4): 491–501.
- Mikros, G. K.** (2013a). Authorship Attribution and Gender Identification in Greek Blogs. In Obradović, I., Kelih, E., and Köhler, R. (eds), *Methods and Applications of Quantitative Linguistics. Selected papers of the 8th International Conference on Quantitative Linguistics (QUALICO), Belgrade, Serbia*, Aprile, 2012. Belgrade: Academic Mind, pp. 21–32.
- Mikros, G. K.** (2013b). Systematic stylometric differences in men and women authors: a corpus-based study. In Köhler, R. and Altmann, G. (eds), *Issues in Quantitative Linguistics 3. Dedicated to Karl-Heinz Best on the occasion of his 70th birthday*. Lüdenscheid: RAM – Verlag, pp. 206–23.
- Mikros, G. K. and Perifanos, K.** (2015). Gender identification in modern Greek tweets. In Tuzzi, A., Benešová, M. and Macutek J. (eds), *Recent Contributions to Quantitative Linguistics*, vol. 70. Berlin: De Gruyter, pp. 75–88.
- Milkova, S.** (2013). Mothers, daughters, dolls: on disgust in Elena Ferrante's *La figlia oscura*. *Italian Culture*, **31**(2): 91–109.
- Moretti, F.** (2013). *Distant Reading*. London: Verso.
- Mullenneaux, L.** (2007). Burying mother's ghost: Elena Ferrante's "troubling love". *Forum Italicum*, **41**(1): 246–50.
- Murtagh, F.** (2005). *Correspondence Analysis and Data Coding with Java and R*. London: Chapman & Hall/CRC.
- Murtagh, F.** (2010). The correspondence analysis platform for uncovering deep structure in data and information, 6th Boole lecture. *The Computer Journal*, **53**(3): 304–15.
- Murtagh, F.** (2017). Big data scaling through metric mapping: exploiting the remarkable simplicity of very high dimensional spaces using correspondence analysis. In Palumbo, F., Montanari, A., and Vichi, M. (eds), *Data Science - Innovative Developments in Data Analysis and Clustering*. Cham: Springer International Publishing, pp. 295–306.
- Oliveira, W. Jr., Justino, E., and Oliveira, L. S.** (2013). Comparing compression models for authorship attribution. *Forensic Science International*, **228**: 100–4.
- Rastelli, A.** (2016). Elena Ferrante, lo studio statistico richiama in causa Starnone. *Corriere della Sera*, Milano, 12 October 2016, p. 37.

- Ratinaud, P. and Marchand, P.** (2012). Application de la méthode ALCESTE à de “gros” corpus et stabilité des “mondes lexicaux”: analyse du “CableGate” avec IRaMuTeQ. In Dister, A., Longrée, D., and Purnelle, G. (eds), *JADT 2012 Actes des 11es Journées internationales d’analyse statistique des données textuelles*. Liège – Bruxelles: LASLA – SESLA, pp. 835–44.
- Ratinaud, P. and Marchand, P.** (2016). Quelques méthodes pour l’étude des relations entre classifications lexicales de corpus hétérogènes: application aux débats à l’Assemblée Nationale et aux sites web de partis politiques. In Mayaffre, D., Poudat, C., Vanni, L., Magri, V., and Follette, P. (eds), *JADT 2016 Proceedings of 13th International Conference on Statistical Analysis of Textual Data, Nice 7-10 June 2016*, vol. 1. Nice: Presses de Fac Imprimeur France, pp. 193–202.
- Ricciotti, A.** (2016). Un confronto tra Elena Ferrante e Anna Maria Ortese: la città di Napoli, la fuga, l’identità, Zibaldone. *Estudios italianos de La Torre del Virrey*, 4(2): 111–22.
- Rudman, J.** (1998). The state of authorship attribution studies: some problems and solutions. *Computers and the Humanities*, 31: 351–65.
- Rybicki, J., Eder, M., and Hoover, D.** (2016). Computational stylistics and text analysis. In Crompton, C., Lane, R. L., and Siemens R. (eds), *Doing Digital Humanities*. London; New York, NY: Routledge, pp. 123–44.
- Santagata, M.** (2016). Elena Ferrante è *La lettura – Corriere della Sera*, Milano, 13 March 2016, p. 2 and p. 5.
- Savoy, J.** (2010). Lexical analysis of US political speeches. *Journal of Quantitative Linguistics*, 17(2): 123–41.
- Savoy, J.** (2013). The federalist papers revisited: a collaborative attribution scheme. *Proceedings of the Association for Information Science and Technology*, 50(1): 1–8.
- Savoy, J.** (2015). Text clustering: an application with the State of the Union Addresses. *Journal of the American Society for Information Science and Technology*, 66(8): 1645–54.
- Segnini, E.** (2017). Local flavour vs global readerships: the Elena Ferrante project and translatability. *Italianist*, 37(1): 100–18.
- Stamatatos, E.** (2009). A survey of modern authorship attribution methods. *Journal of the American Society for Information Science and Technology*, 60(3): 538–56.
- Stamatatos, E., Fakotakis, N., and Kokkinakis G.** (2001). Computer-based authorship attribution without lexical measures. *Computer and the Humanities*, 35: 193–214.
- Starnone, D.** (2011). *Autobiografia erotica di Aristide Gambia*. Torino: Einaudi.
- Tuzzi, A.** (2010). What to put in the bag? Comparing and contrasting procedures for text clustering. *Italian Journal of Applied Statistics/Statistica Applicata*, 22(1): 77–94.
- Zhao, Y. and Zobel, J.** (2005). Effective authorship attribution using function word. In *Proceedings of the 2nd AIRS Asian Information Retrieval Symposium*. Berlin: Springer, pp. 174–90.
- Zheng, R., Li, J., Chen, H., and Huang, Z.** (2006). A framework for authorship identification of online messages: writing-style features and classification techniques. *Journal of the American Society for Information Science and Technology*, 57(3): 378–93.

Notes

1. *Ferrante Fever* is also a recent documentary directed by Giacomo Durzi and distributed in cinemas in October 2017.
2. ‘Non sei fregato veramente finché hai da parte una buona storia, e qualcuno a cui raccontarla’ [You’re not completely done for as long as you have a good story in you, and someone to tell it to] (Baricco, 1994: 17, Trans. author’s own).
3. *a, ad, affinché, agl’, agli, ai, al, alcuna, alunché, alcune, alcuni, alcuno, all’, alla, alle, allo, allora, allorché, alquanta, alquanto, altr’, altra, altre, altrettanta, altrettante, altrettanti, altrettanto, altri, altro, altrui, ambedue, anch’, anche, ancorché, anzi, anziché, appena, appresso, attraverso, avanti, benché, bensì, c’, cadauno, casomai, ce, certe, certi, ch’, che, ché, checché, chi, chicchessia, chiunque, ci, ciascun, ciascuna, ciascuno, ciò, cioè, cionondimeno, circa, codesta, codeste, codesti, codesto, cogli, coi, col, colei, coll’, colla, colle, collo, coloro, colui, com’, come, comunque, con, conforme, contro, cos, così, cosicché, costei, costoro, costui, cotanto, cui, d’, da, dacché, dagli, dai, dal, dall’, dalla, dalle, dallo, de, degl’, degli, dei, del, dell’, della, delle, dello, dentro, di, dietro, difatti, dimodoché, dinanzi, dinnanzi, diverse, diversi, dopo, dov’, dove, dunque, durante, e, ebbene, eccetto, ed, egli, ei, ella, entrambe, entrambi, entro, eppure, ergo, essa, esse, essi, esso, extra, fin, finché, fino, fintantoché, fra, fuorché, fuori, giacché, giusta, gl’, gli, gliel’, gliela, gliele, glieli, glielo, glien’, gliene, granché, i, idem, il, in, infatti, innanzi, io, l’, la, laddove, le, lei, li, lo, lontano, loro, lui, lungo, m’, ma,*

malgrado, me, meco, medesima, medesime, medesimi, medesimo, mediante, meno, mentre, mezz', mi, mia, mie, miei, mio, molta, molte, molti, molto, n', ne, né, neanche', neanche, negli, nei, nel, nell', nella, nelle, nello, nemmeno, nemmeno, neppure, nessun, nessuna, nessuno, nient', niente, noi, noialtre, noialtri, nonché, nondimeno, nonostante, nostra, nostre, nostri, nostro, null', nulla, nulle, nulli, nullo, o, od, ogniqualvolta, ognuna, ognuno, oltre, oltreché, onde, oppure, ora, ordunque, ossia, ove, ovvero, parecchi, parecchia, parecchie, parecchio, pei, penultima, penultimo, per, peraltro, perché, perciò, permesso, però, pertanto, più, piuttosto, poc', poca, poche, pochi, poco, poiché, presso, prim', prima, prime, primi, primo, pro, propri, propria, proprie, proprio, prossima, prossime, prossimi, prossimo, pur, puranco, purché, pure, qual, qualcheduno, qualcos', qualcosa, qualcun', qualcuna, qualcuno, quale, quali, qualora, qualunque, qualvolta, quand', quando, quant', quanta, quante, quanti, quanto, quantunque, quasi, quegli, quei, quel, quell', quella, quelle, quelli, quello, quest', questa, queste, questi, questo, quindi, rispetto, s', salvo, se, sé, sebbene, seco, seconda, seconde, secondi, secondo, semmai, sempreché, sennonché, senonché, senz', senza, seppur, seppure, si, sia, sicché, siccome, sin, sino, solo, sopra, sott, sotto, stante, stessa, stesse, stessi, stesso, su, sua, sue, sugli, sui, sul, sull',

sulla, sulle, sullo, suo, suoi, t', tal, talaltra, tale, tali, talune, taluni, taluno, tant', tanta, tante, tanti, tantino, tanto, te, ti, tot, tra, tramite, tranne, troppa, troppe, troppi, troppo, tu, tua, tue, tuo, tuoi, tutt', tutta, tuttavia, tutte, tutti, tutto, ultima, ultime, ultimi, ultimo, un, un', una, une, uni, uno, v', vari, varie, ve, verso, vi, viceversa, voi, voialtre, voialtri, vostra, vostre, vostri, vostro.

4. Domenico Starnone wrote 'bel lavoro, professore; peccato che io resto maschio e Elena Ferrante femmina, io autore, lei autrice; sì, sono spiacente, non sa quanto mi consolerebbe in questo momento poterle dire è vero, mi travesto da donna, mi inciprio il naso allo specchio, sono lo scrittore che si fece donna, maschio e femmina come il primo Adamo; ma non è così; però, grazie, mi ha divertito' [good job, professor! Unfortunately I am still a man and Elena Ferrante a woman. I am a male writer, she is a female writer. I am sorry, Heaven knows I would really like to tell you you are right, that I wear women's clothes, sit in front of a mirror and powder my nose, that I am a male writer turned into a woman, who is both man and woman like Adam, the first human being; unfortunately this is not the case. However, thank you: you amused me] (Starnone, 2011: 431, Trans. author's own).